Moral Deviance Across Cultures

Investigations of the Human Moral Mind

A thesis submitted for the degree of

Doctor of Philosophy

in

Psychology

by

PAY LAURIN JESSEN

Bremen, 02.12.2024

Under the Supervision of

Prof. Dr. Johannes Huinink

Prof. Dr. Ulrich Kühnen

Prof. Dr. Vivian Vignoles

The dissertation was reviewed by

Prof. Dr. J. Huinink, Prof. Dr. U. Kühnen & Prof. Dr. V. Vignoles,

& successfully defended from P. L. Jessen on February 12, 2025

Bremen International Graduate School of Social Sciences

Universität Bremen / Constructor University

To My Loved Ones

Acknowledgement

At the festivities for his 90th birthday, my grandfather gave a speech in front of the assembled guests. Looking back on his great life experience, he came to a kind of conclusion at the end of his speech. My grandfather's words still resonate with me years later. At 34 years old, I have only received a fraction of the diverse experiences and insights of my grandfather, but I am gradually coming to understand more and more what he said that day. He ended his speech by saying, "*It is the others who make us who we are.*"

On the one hand, as we shall see, this sentence, enriched with experience, is deeply rooted in what I have brought to expression in the work at hand over the span of the last three years. For indeed, our selves and our morals are fundamentally socially constituted. On the other hand, and this gives me great joy, I am fortunate to experience the content of this sentence in my life in a very practical way. Over the course of my 34 years, I have met so many wonderful people who have enriched and shaped me deeply. My heartfelt thanks go to you!

My loved ones, my partner, my family, and my friends, are a great source of support, joy, and constant broadening of my horizon. There are probably few people who have experienced the often stressed, mostly restless everyday life of a doctoral student, who has far too little time for anything other than work, as much as my partner, Insa. I want to thank you from the bottom of my heart for your endless patience with me, for your love and care, for the great fun that our time together always means to me, and especially for the perseverance you have shown and the support you have given me. You have been a pillar of strength in so many situations in life, a faithful friend and the most loving partner. Without you, I would never have come this far. I am deeply grateful and so happy to have you by my side! The same goes for my family — I would like to thank you from the bottom of my heart for your existence, your support, your love and your devotion. I feel very fortunate to have two people by my side who have been with me my whole life: my sister Liv and my brother Finn. You are my deepest friends and soulmates. I would like to thank my dear mom, Gesche, and my dear dad, Gerd, for your deep love, patience, great support in every situation and the trust that you have always placed in me. My gratitude encompasses so much more, for which I have no words. I couldn't be happier to have you by my side! I would also like to thank my brother's dear wife, Marie, and my dear uncle Sønnie (Mange tak, min kære morbror) for their constant and heartfelt support, for which I am very grateful. My dear family, you are source of deep happiness and joy for me; I will always be there for you. My heartfelt thanks also go to my dear friends — we live in a relational mobile society and it is therefore not at all self-evident that you stand by me

as you do. Especially in the last 18 months, I have certainly been unable to see you often enough and spend time with you. My daily routine, which results from this work, simply did not allow for this. I am very grateful that you were always there for me, that you thought of me, and that in the few hours we spent together, you were always the friends I wished for. Dear Tanya, Carina, Gamze, Daria, Catalina, Janna, dear Torben, Roberto, Johannes, Merlin, Patrick, Michel, Gonzalo, Maurice, David, Sönke, and all those not mentioned in this far-too-short list, I would like to thank all of you and look forward to the time together that awaits us.

Dear Sabine Sommer, dear Regina Arant, dear Mandy Boehnke, dear Mandi Larsen, dear Stefanie Bittner, dear Monika Liorentaite-Rossdeutscher, dear Inga Köster, and dear Christian Peters: I would like to thank you for your fantastic organizational skills, without which I would have been lost at many points in the dissertation. But above all, I am happy and grateful for the open arms and open ears; the exchange with you was an enrichment every day. I would also like to thank Sonja Drobnič, Patrick Sachweh and Adalbert Wilhelm for their support and for what I was able to learn from them. My thanks go to the entire BIGSSS family for the exciting and insightful glimpses into your work and the enriching intellectual debates I was able to engage in with you.

There are also people who have accompanied and influenced me on my academic path and to whom I would like to express my sincere appreciation. Dear Michael Windzio and dear Alexander Gattig, your approachability and your sense of recognizing and fostering a sense of curiosity about the world were a great blessing that I wouldn't want to have missed. I am very grateful for our conversations and for what I was able to learn from you in the academic field and beyond. You were always a great support for me in many ways on the way to my dissertation.

Torsten Selck and Murat Ardag certainly shaped the foundation of my academic path. I will always be grateful for your early support and what I was able to learn from you!

I would also like to express my sincere gratitude to Stefanie Kley for her tremendous support and for creating the space for me to be able to complete this work.

My deep gratitude also goes to Mrs. Lafon for the enriching, instructive and amusing (!) conversations and the enormous support in difficult moments.

The following people were crucial to the creation of this thesis: Johannes Huinink, Ulrich Kühnen, Vivian Vignoles, Klaus Boehnke and Franziska Deutsch. Without your enthusiasm, your words of caution and your foresight, without your creativity, your honesty, the good and deep conversations, the very warm and personal interaction and your trust, this work would not have been possible. Dear Johannes, dear Uli, dear Vivian, dear Klaus, dear

Franzi, your support during the highs and lows that my dissertation demanded of me was extraordinary! I am grateful to you with all my heart. Summarizing in a few lines what a decisive contribution you have made to this work, but also to my personal development over the last three years, is an undertaking I cannot succeed at, because your contribution was enormous. The way you are able to steer the confusing game of thoughts into fruitful paths, how you are able to look so far and deep and share your view with others, has not only impressed me deeply, but also shaped me. I am extremely grateful to you that you did not let the interpersonal aspect fade into the background, but that the clarity that characterizes your actions was always accompanied by it. I would like to sincerely thank you for everything! As already mentioned, it is impossible to capture your contribution to this work in a few lines. Therefore, please take it as a sign of my gratitude and appreciation for your efforts that I have decided to write about "we" when positions, findings and interpretations are shared in the course of this thesis. As readers will see, this work contains many numbers, among other things, and many numbers are also many sources of potential error. The readers should therefore note that I am solely responsible for any errors in this work regarding numbers and other things. But when it comes to the contribution that this work can hopefully make, I would like to speak of a *we* and thus honor the wonderful people who have supported and enriched me so much.

I thank you all from the bottom of my heart, and indeed I understand more and more that it is the others who make us what we are.

Abstract

This dissertation deals theoretically and empirically with human morality. More specifically, morally deviant actions are the focus of the research that we will present in the course of this work. In addition to *morality*, the human *self* and *culture* are the further pillars on which this work is based. As an overarching goal, we pursue the research *question which moral system guides cooperation in different cultures*? In seven chapters, we will first deal theoretically, but then mainly empirically, with human morality, the self, and culture. *Chapter 1* discusses the three theoretical foci of this work, mainly against the background of evolutionary theories. At the end of this chapter, we will also derive several hypotheses to be tested. In *Chapter 2*, we take a look at comparative cultural research and prepare the ground for subsequent investigations of cross-cultural similarities and differences in morality. With regard to moral tendencies, our research focuses on comparisons between *Egypt*, *Germany*, *Japan* and the *United States of America*. *Chapters 3* through *6* address independent yet complementary cross-cultural investigations of the human moral mind. In these chapters, we aim to approach our overarching research question by means of a canon of different methods. Eventually, *Chapter 7* provides a summary discussion and conclusion. Cross-cultural research on the human moral mind is situated in the field of tension between the poles of the *culturally specific* and the *universally human*. The results of our investigations also fall within this field of tension. We will be able to provide strong empirical indications of universal moral domains on the one hand, and present results that demonstrate the massive influence of culture on the calibration of our moral mind on the other. Overall, this dissertation aims to make three contributions: *First*, we attempt to make a theoretical contribution by synthesizing two leading moral theories and proposing our own moral approach based on them. *Second*, we develop and test three different instruments that are meant to expand our toolbox for cross-cultural research on morality. *Third*, we attempt to make an empirical contribution by examining the moral systems of four heterogeneous cultural entities. We are guided by the hope that this dissertation will shed some light on the human disposition that drives us to self-regulate and that enables us to cooperate so profoundly and extensively with others: our morality.

Table of Contents

List of Tables

List of Figures

Chapter 1: Culture, the Self and Human Morality

1.1 Introduction

Have you ever felt a flash of anger because you were treated unfairly? Or do you remember impulsive negative feelings because dear friends did not stand at your side in public? In general, can you recall a gut feeling of sudden negative arousal, accompanied by a slight inclination to punish, because certain standards you consider to be good were violated? Most likely you can. Yet what about situations in which you have acted: Have you ever felt immediately ashamed when you realized you had failed to keep a promise, or was there a sense of guilt that rose up in you all of a sudden because you recognized you had been disrespectful to an authority? Again, in general, ask yourself if you can recall a situation in your everyday life in which you intuitively felt a negative emotion because you violated something that is considered a standard of good human behavior.

Probably every one of us has experienced and felt either of these events and the instant flash of respective emotions. From time to time, we are exposed to other people who deviate from what we consider to be good and acceptable. And on top, it would be hypocritical and inflated to think of oneself as having never done something that deviates from other people's realm of what is socially acceptable. However, more striking is the fact, that most of the time these standards of what is good and acceptable among us humans are not violated!

With this in mind, we would like to take a leap to other examples that we often take for granted in our social lives. So, have you ever wondered why we are able to possess and even accumulate particular objects in the first place? It's because people recognize to a certain extent that objects can be owned, that they can be private property. Further, have you also come across the phenomenon of so many people being drawn to superhero movies and wondered why that actually is? Probably part of the explanation is that we are attracted to those who courageously intervene to defend others who are in need, because we perceive heroic characters as good and righteous. Essentially, we have a gut feeling towards heroism and a tendency to reward heroic figures good deeds, if not materially than at least with positive social reputation. When we glance at the news, we sometimes read about union strikes for fair wages or see justice movements demonstrating in the streets and raising their voices for equity. Most of us are also familiar with the pleasant experience that families, as well as our groups of friends, generally tend to support us both emotionally and materially. We also often defer to certain hierarchies (think of how unsatisfying and dysfunctional your work life would be if you and others did not

respect certain authorities and work instructions). In addition, we often return a favor when owe it (consider how many times you and your friends helped each other to move), and sometimes we even share some of our innermost secrets with others because we trust them to keep our affairs confidential.

If some or even all of what we have just described seems familiar to you, then the experiences of *your* life suggest that people apparently care about others in regard to several socially cooperative domains. To be precise, the domains we referred to in the examples are: *fairness*, *trustworthiness*, *property*, *family*, *in-group*, *deference*, *heroism* and *reciprocity*. We will discuss these domains in more detail in the course of this text. But let us come back to the argument that we like to put forward; Instead of facing constant violations to these domains of the social, our daily life is rather traversed by manifold ways of *cooperation* with other people. Now, what is the driving force that enables cooperative interactions of the kind mentioned above among humans, and why do we feel ashamed or guilty when we deviate from acceptable standards in these domains?

An integral component to the answer of this question lies in our *morality*. By aiming for no more than a fuzzy working definition at the moment, we can state that morality refers to ideas, to standards of judgment of what is socially acceptable or unacceptable (Carlo et al., 2016). Essentially, morality is a force of self-regulation promoting interpersonal cooperation and social order (Baumeister, 2010; Tomasello & Vaish, 2013; Ellemers et al., 2019). Clearly, when we referred above to *our* morality, we implied the view that morality is a characteristic of all humanity. We follow the view of a *universal human morality* throughout our work. However, this is not yet the whole story of morality, as we suggest below and across the investigation we undertake here. It is precisely when we treat the semantic content of the word universal (i.e., belonging to all; see: Hügly & Lübcke, 2013, p. 906) from a social, human-centered perspective that we must also embrace its relation to the other side of the coin, i.e., to the *culture specific*. Human cultures are diverse and different. Even though the daily lives of people around the world call for morality to guide and regulate social interactions, this does not mean that the same demands are placed on social interactions across cultures. It is therefore possible that the calibration of our moral tendencies is not uniform across different cultures. We will explain this idea in detail in the course of this project. For now, however, let us briefly elaborate on morality and other phenomena with regard to universalism and the culturally specific.

We have social standards for what is considered acceptable, standards for what is good and right. Unlike social conventions (group norms), when we talk about morality, these standards are not seen as something that is tied to a particular group life (Skitka & Conway,

2019). Some scholars argue that the nature of morality is different from that of norms because human morality is seen as something that is shared across social groups, it is a universal standard for what is good and what is bad (Hoerster, 2022). This view implies that morality includes standards for what is considered to be a good or bad action for all people. The widely known categorical imperative by Immanuel Kant essentially reflects one strand of the universalist concept of morality. Basically, the categorical imperative states that an action can only be considered moral if the determination of will to act is guided by the notion that one´s action is principally capable of becoming a universal law (see: Kant, 1788/2011, p. 738). In simple terms, the reason underlying the action must take into account the well-being of all in order for the action to be considered moral. Kant grounds his deliberations on morality on abstract reasoning nested in complex structures of argumentation. Notably, in line with philosophical tradition, the categorical imperative is deductively derived from (pure) thought referring to a rationalist view (Moses & Knutson, 2012) on what is considered to be moral.

However, the study of human morality is no longer limited exclusively to philosophy and the primacy of a rationalist perspective. Within psychology and the social sciences (ethnology, anthropology, and sociology etc.) we can find an evolutionary oriented approach to morality that is different to a rationalist position. This approach is more empirically oriented and searches for the emotional, cognitive and behavioral underpinnings of human morality, and in an evolutionary sense for the primal antecedents of our moral mind ultimately. Though it may seem striking at first, however, evolutionary guided streams of psychology nowadays strongly indicate that particular domains of human social interaction are policed and guided by our *evolved moral mind* (Haidt & Joseph, 2007). Central to this view is the notion that the evolution of our moral mind solved recurring *adaptive challenges* faced by our species throughout its evolutionary history (Kurzban et al., 2015; Hare, 2017; Henrich & Muthukrishna, 2021). One line of particularly influential adaptive challenges is concerned with social interaction problems of human cooperation, as highlighted by the Morality as Cooperation Theory (Curry, 2016). We have already referred to domains of human cooperation, for example by touching on fairness, which becomes relevant as a moral domain with regard to the (fair) distribution of resources. Furthermore, this view implies two main propositions. First, the content of morality is *plural* and consists of several moral domains that have developed in response to various recurring challenges (Graham et al., 2013). Second, the human moral mind can be regarded as bearer of particular functions that serve to solve interaction problems of human *cooperation* (Diekmann & Lindenberg, 2015; Curry et al., 2019a). This is nevertheless not to claim a metaphysical functionalism and also not comparable to the base from which Kant

started. Rather, the functionalistic notion here is a methodological functionalism that guides as a heuristic the search for mechanisms and patterns to be found in correlations between our mind and events realized in the world (Brüntrup, 2004). Clearly, the evolutionary approach to morality involves a *universalist notion*, though one that is different to the categorical imperative. Within an evolutionary framework to morality a universalist stance is emphasized by the notion that the evolved human moral mind is a feature of our species in general.

After these introductory reflections on the universality of our evolved moral mind, we now move on to a brief look at what is meant by the *culture specific*. Cultures encompass historically evolved social structures, language(s), customs and traditions, rituals, taboos, norms and may also include particular clothing styles, architectural styles or food preferences (e.g.).[1] Cultural contexts can also imply particular natural conditions, as cultures may be regarded as spatially localized, at least to a certain extent. Overall, cultures represent a rich bouquet of meaning systems, worldviews and practices that differ around the world (Esser, 2010; Smith, 2014; Henrich, 2020; Brown et al., 2022; Rippl & Seipel, 2022).

Since the planet we inhabit comprises a variety of natural ecologies it is not surprisingly that dimensions of human culture(s) are diverse too. Let's briefly look at some examples of cultural dimensions to get a first impression. Cultural dimensions include (e.g.): proximity, i.e., variable understandings of physical closeness and distance; relational mobility, i.e., rather static or fluid structures of social relations; and a focus on nuclear family or extended family relationships, a distinction representative of the kinship intensity dimension (Barmeyer, 2010; Thomson et al., 2018; Schulz et al., 2019). Another cultural dimension that we would explicitly like to highlight here is *individualism-collectivism*. This dimension encompasses various attributes, but at its core it is characterized by different social focal points. Put simple, in *individualist* cultures the focus is on the individual and interests of the individual take precedence over those of the group. In contrast, in *collectivist* cultures the focus lies on the (in-)group and the interests of the group take precedence over those of individuals (Triandis et al., 1990; Triandis, 2001; Krys et al., 2022; Minkov & Kaasa, 2022; Żemojtel-Piotrowska & Piotrowski, 2023; Kitayama & Salvador, 2024). With regard to cultural dimensions, it can therefore be seen that although culture itself is universal for humans, cultures around the world differ in several aspects.

Our psyche and culture are interwoven. In the course of human history, culture has become indispensable for our species and has had a major influence on the development and

---

[1] This listing serves illustrative purposes and is not intended to be in any way exhaustive in terms of the characteristics of cultural entities.

spread of human life (Henrich, 2016; 2020). Environmental and sociocultural factors influenced each other in reciprocal ways throughout human history, resulting in the emergence of distinct cultural ecologies that are associated with different social needs and affordances (Henrich & McElreath, 2007; Chudek et al., 2016; Mesoudi, 2017; Whiten et al., 2017; Heyes & Moore, 2021). In fact, culture is so central to human living that it can be regarded as the human's biological niche for it comprises all the environmental factors that are necessary for human survival and reproduction of our species (Brown et al., 2022). Thus, as culture ensures survival it is no wonder that our species adapts to its cultural environment. Our adaptive minds and the respective cultural ecologies of humans are hence central to understanding what drives people (Cosmides & Tooby, 1997; Boyd et al., 2011; Kitayama & Salvador, 2024). To be precise, our mind is calibrated to navigate the sociocultural world, which is the primary source of our experiences and social interactions (Henrich, 2020). This view entails that the human psychological apparatus adapts to the sociocultural environment and ensures that feelings, thoughts and actions are sufficiently effective within the immediate sociocultural context. A fundamental and crucial survival related quality of our evolved mind is revealed. However, our evolved psychological apparatus also affects the (re-)production and preservation of culture. The latter relation is particularly evident in our evolved learning biases, which we take as an example here. Our learning psychology leads us to select, process and imitate more information from certain social models (i.e., individuals) than from others. This means that our minds are biased towards particular social cues. Same ethnicity and gender, prestige, skill and success are social cues that attract our learning attention and guide the selection of the social model we learn from (Henrich, 2016). Consequently, not all cultural information is passed on with the same probability, as our evolved learning biases act as gatekeepers in this regard. Thus, as illustrated in the example, the relation between human psychology and culture is one of reciprocal influence.

In the course of the present project, we will argue that the notion of adaptation and reciprocal influence between culture and psychology is visible in the way how people construe their own selfhood and in the cultural configuration of moral systems. To be able to build a link between selfhood and morality we will therefore turn shortly to cultural ways of *self-construal*. The influence of culture on our human psychology is essentially so pervasive that even the way we construe our own selves — the "I am ..." when we refer to how we see ourselves in relation to others — is subject to the powerful influence of cultural context (Markus & Kitayama, 1991; Vignoles et al., 2016; San Martin et al., 2018; Uskul et al., 2023). Researchers have identified cross-cultural differences in the way how people construe their own self. In particular, a

distinction is made between two modes of self-construal: the *independent self* and the *interdependent self*. On the one hand, the first mode is characterized by the emphasis of being an autonomous, independent individual. Values center around autonomy, emotions are rather ego-focused and information tends to be processed analytically by people with a predominant independent selfhood. The second mode, on the other hand, is characterized by the emphasis on being embedded in and in harmony with one's social group of relevant others. For people with predominant interdependent self-construal groups are the focal point of values, emotions are rather other- than ego-focused and information tends to be processed holistically (Cross et al., 2011). Although each mode of selfhood is found across cultures, research shows that cultures clearly differ in whether the independent self or the interdependent self predominates (Kitayama & Salvador, 2024). Our self is accordingly fundamentally social and shaped by culture. As an acting individual, however, our self is also the entity that (re-)produces the elements of the sociocultural world in the first place (Ridgeway, 2006; Markus & Kitayama, 2010; Morf & Koole, 2014). We are therefore both the product and (re-)producer of our sociocultural ecology (Greshoff, 2008; Esser, 2010; Berger & Luckmann, 2013). Ultimately, the human self reflects the sociocultural requirements of a context, reveals itself as an element of the (re-)production of that context, and expresses the psychological functioning of the individual in that context (Baumeister, 2022).

Here we pick up the red thread of our topic again: the constitution of morality is probably no exception in terms of cultural contingent configuration, and human morality is therefore likely to vary, at least in part, across cultures. Against the background of the adaptationist view, it follows that morality serves a purpose related to the cooperative requirements of societies and the social order of a particular sociocultural system (Ellemers et al., 2019; Henrich, 2020; Baumeister, 2022). In line with this is the notion that the adapted moral mind promotes human survival by fitting moral domain endorsement preferences to the conditions of the immediate cultural context. Although we follow the argument of moral universalism, we now have an indication to assume differences in human morality due to different socio-cultural environments in which individuals find themselves embedded. In this line a prominent theory in evolutionary moral psychology forwards the idea of *binding* and *individualizing* morality (Haidt, 2008). The *Moral Foundations Theory* (Haidt & Joseph, 2007) suggests a view comprising universal moral foundations that are yet culturally adapted. The binding approach to morality is thought to reflect a prevalence of group focused moral values and virtues, whereas the individualizing approach to morality is considered indicative of individual focused moral values and virtues. Moreover, as empirical evidence suggests, cultures differ in terms of *particularistic* and

impartial moral tendencies (Henrich et al., 2005; 2010b; Enke, 2019; Waytz et al., 2019). Furthermore, an emphasis on one or the other tendency is probably related to historical processes of socio-cultural development that have promoted either a main focus on group orientations and interactions or on orientations towards and interactions with individuals (Thomson et al., 2018; Schulz et al., 2019; Talhelm, 2022). These socially distinct orientations and modes of interaction also seem to be characteristic of the individualizing and binding approach to morality.

We would now like to make a brief synthesis of the themes broached. Recall that we asked you if you have experienced a flash of emotion and a tendency to punish when confronted with others who deviate from social standards. Now what is your take, do you consider your responsive emotional flash to be the same in intensity when someone treats you *unfair* or does not *defer* to your orders? Are fairness and deference of the same *relevance* to you? Let us integrate this thought to the notions of the (evolved) moral mind and the culture specific, and formulate some more general questions. Morality is characteristic to all humankind, yet human cultures vary considerably. Hence, we wonder: are moral systems in different cultural contexts the same in terms of the *relevance* they attach to different domains of cooperation? Are systemic requirements for cooperative action between people the same across cultures? Or asked less abstract: do different sociocultural entities (i.e., societies) today attribute the same relevance to fairness as to deference? In addition to morality, we have also introduced several cultural dimensions. We have touched the central characteristic of individualism-collectivism and elaborated briefly on cross-cultural differences in self-construal. The independent and the interdependent self-construal were highlighted as culturally different modes of selfhood. Moreover, the self was emphasized as culturally shaped and at the same time characterized as an agentic entity that brings forth the social and cultural in the first place. Against this background, further questions arose in our heads: is there a systematic correspondence between the modes of selfhood and the cultural configuration of morality? Moreover, we have also briefly extended our introductory remarks to binding and individualizing morality and to moral tendencies of particularism and impartiality. With regard to the central features of these concepts, we further ask ourselves whether there is a systematic connection between them?

Before we condense these questions into more precise theses, we will try to develop an initial and brief but consistent theorization to approach them. Central to our theoretical considerations is a link between the respective configuration of the moral mind and the cultural configuration of the self. Essentially, we argue that morality is an aspect of the human self (Baumeister, 2022) and to support life under given sociocultural conditions our self and

morality are configured contingent on the cultural context. The basis for our argument comprises the following: All societal living is to some degree dependent on inter-individual cooperation. Our evolved moral mind bears the capacity to regulate human cooperation. So, we regard morality to be a universal facet of the evolved human mind. Culture is the biological niche of humans supporting human survival and reproduction. Human cultures are sociohistorical formed diverse systems of knowledge and meaning. In order to meaningfully promote interindividual cooperative interaction in a given sociocultural context, we assume our moral mind to be at least partly adapted to the requirements of its cultural environment. Since cultural contexts differ, this reasoning does imply cross-cultural variation in the configuration of morality and thus moral tendencies. We argue furthermore that morality is a part of our self. One central aspect of cross-cultural differences related to selfhood is the predominance of an overall group or overall individual social orientation. This aspect is expressed, among other things, in collectivist or individualist tendencies of cultural entities. Culture and self "make each other up" (Markus & Kitayama, 2010, p. 421). So, respective cultural requirements are mirrored in the self. Against this background we expect that culturally induced differences in self-construal are systematically related to cross-cultural differences in moral systems. In other words: We assume that the respective relevance of different moral domains depends to a certain degree on their (historical yet also actual) functional utility, their compatibility in a given sociocultural system. This utility in turn reflects systemic demands on and of the individual, and is therefore mirrored in the construal of selfhood. More so, groups imply a social boundary between the inner and outer realm of belonging (Hogg et al., 2004). The interdependent self has a characteristic relationship (group) focus while the independent self-construal has a characteristic focus on the individual. Furthermore, we expect these foci to be reflected in differential preferences for either binding or individualizing morality, and in tendencies toward particularistic or impartial morality. Thus, a consistent stream of thought emerges from the heuristic considerations outlined above. This reasoning focuses on substantially different sociocultural focal points, i.e., the individual or the group, and ranges from *individualistic* and *collectivistic* cultures to *independent* and *interdependent* self-construal, *individualizing* and *binding* morality, and *impartiality* and *particularism* in moral tendencies. Our reasoning draws on cumulative cultural evolution, associated path dependencies and the mutual constitution of individual actors and sociocultural structures. Against this background we assume different cultural contexts and correspondence in variation of self-construal and moral configuration.

Now we culminate our thoughts in the formulation of central theses and research questions. We will state these considerations here shortly and elaborate them further in the

course of this text. Essentially, we forward *four underlying theses*: *first* we assume several domains of human cooperation as cross-culturally shared universals. *Second*, we expect cross-cultural variation in moral systems, i.e., we expect cultures to vary with regard to the relevance they attach to different moral domains. *Third*, we suppose cross-culturally a systematic correspondence between ways of selfhood and the configuration of moral systems. *Fourth*, we presume cross-cultural differences in moral particularism/impartiality tendencies and expect relations of these tendencies to binding and individualizing morality. Ultimately, we are concerned with the overarching research question **Which moral system guides cooperation in different cultures?**

In this **thesis** we will **outline** a quantitative cross-cultural project to investigate the human moral mind. Our project deals with *culture*, *morality* and *self-construal* primarily from a psychological perspective. Specifically, data for comparative purposes will be gathered in four countries. These are: Egypt, Germany, Japan and the United States. The cases (countries/nations) that we have selected for our study serve as proxies for cultural entities (Smith, 2014; Minkov et al., 2021), that are differentiated along cultural dimensions that should promote in an ideal-typical sense either an individual or group social orientation.

Overall, we pursue three major research goals, of which two are primarily theoretical in nature and one is focusing mainly a methodological research gap. *At first* this project is designed to investigate cross-culturally the psychometric properties of a new self-report instrument that is conceptualized to capture the deviance relevance of 8 moral domains. So, our goal is to develop a scale that is committed to moral pluralism and centers on moral deviance. By utilizing model based statistical approaches we aim to examine our new scale, the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS), in terms of factor structure, validity, reliability and measurement invariance across four cultural entities. This project will be the first study to gather and inspect data with MaC-DRS cross-culturally and strives to close the methodological gap of insufficient self-report measures of moral relevance (Graham et al., 2011; Curry et al., 2019a Iurino & Saucier, 2020; Atari et al., 2022a). Provided that our investigations are able to demonstrate the plural structure of MaC-DRS across cultures, our investigations will also contribute to our understanding of moral universalism. In addition, we supplement the scale with a new set of 9 moral dilemma scenarios and a newly developed Moral Deviance Factorial Survey. These instruments focus in one way or another on *moral deviance* and are intended to expand our methodological repertoire in the field of moral research.

Cross-cultural research always moves between the poles of the human universal and the culture-specific. Within this framework, we ask the question **Which moral system guides**

cooperation in different cultures? and were not able to find an adequate answer in the relevant literature. Motivated by this research question, the present project investigates *at second* possible differences in the relevance of different moral domains across four target cultures. More so, we see a striking theoretical similarity between the *independent self-construal* and the higher order construct of *individualizing morality*, and between the *interdependent self-construal* and the higher order construct of *binding morality* (Markus & Kitayama, 1991; 1998; 2010; Haidt, 2008; Graham et al., 2011; Mooijman et al., 2017). In this regard we reason that morality is an aspect of the self and ask the following sub-question: *is there a systematic correspondence between the cultural configuration of self-construal and the cultural configuration of morality?* This thesis aims therefore to test also several hypotheses about theoretically inferred associations between culture, variant ways of selfhood, and differences in the relevance of moral domains. In doing so, we will first focus on what is referred to as moral intuition. Moral intuitions are the quick, automatic, and uncontrollable reactions we have to moral content in the social world (Haidt, 2001). Later, we will also touch on areas of deliberate moral cognition. Here, we present respondents with scenarios in the form of moral dilemmas and offer only two options for resolving the dilemma, pitting the pursuit of a group-focused approach to morality against an individual-centered approach to morality (and vice versa).

Third, based on previous findings (Triandis, 2001; Thomson et al, 2018; Waytz et al, 2019; Schulz et al, 2019; Enke, 2019; Aldering & Böhm, 2020; Henrich, 2020; Talhelm, 2022; Kirkland et al, 2023), we hypothesize that moral tendencies may differ between (as well as within) cultures in terms of *impartiality* and *particularism*. Furthermore, we assume that these tendencies are reflected in what we regard as binding and individualizing morality (Haidt, 2008). To address this and other theoretical propositions, we conduct a study to collect data using the Moral Deviance Factorial Survey (MDFS) that we developed, which allows us to examine across cultures whether evaluations of moral breaches vary with respect to differing social relationships (i.e., in-group and family vs. stranger). In addition, we enrich this empirical investigation by examining whether moral deviance relevance and moral deviance judgment vary in their extent and potentially cover different aspects of human morality. The superordinate goal of the analysis of the MDFS is to gain further insights into our overarching research question.

In terms of **structure**, this **thesis** consists of seven chapters. ***Chapter 1*** deals with the theoretical background and discusses the topics of culture, self-construal and morality. **Culture** will be discussed briefly in terms of constitutional elements and with an emphasis on cultural evolution. The topic of different cultural dimensions is also touched. In the subsequent section

on **self-construal**, we will address cultural differences in selfhood. We focus on the independent and interdependent self-construal. Here we will highlight central aspects of the culturally contingent ways of selfhood and introduce to the functions of the self. Furthermore, we will describe in more detail how our self and our socio-cultural context constitute each other. In doing so, we will integrate the self-construal approach into the **Model of Sociological Explanation** (Greshoff, 2008; Esser, 2010). This integration ultimately serves to derive different situational logics, which in turn form the basis for our argument of correspondence between the cultural configuration of self-construal and morality. In the transition from self to morality, we emphasize that we consider the latter to be part of the self, which is particularly reflected in self-regulation, the central function of morality. When we turn to **morality**, we move from a general introduction to the **Moral Foundations Theory** (MFT) and the **Morality as Cooperation Theory** (MaC). In what follows, we present a partially expanding but mostly converging perspective on MFT and MaC as our own approach to morality. Guided by our theoretical considerations, we will then introduce the research instruments we have developed. In this context, a new self-report instrument for assessing the relevance of moral deviance, a factorial survey on moral deviance, and a set of moral dilemma scenarios are presented as means of empirically measuring morality. After this excursion, we resume the theoretical discourse and address several various cross-cultural variations in human morality. We conclude the theoretical discussion by emphasizing that we consider morality to be part of the self. Following the Model of Sociological Explanation, we will thereafter work out specific ideal-type situational logics and finally present a theory guided **research model**. From the preceding theoretical discussions, we derive various **hypotheses**. These hypotheses revolve around moral deviance. Ultimately, we seek to test our previously shared considerations through cross-cultural investigations of the human moral mind.

Chapter 2 forms the transition between the theoretical discussion and following empirical investigations. Three primary data collections are introduced in this chapter, with a particular spotlight on the third, cross-cultural data collection. The overarching focus of this chapter is on the methodology and the preparation of our cross-cultural empirical project. We start by outlining our data collections; the respective designs, instruments, and participants are briefly described. After these introductory remarks we will address the **case selection strategy** behind the decision to focus on comparisons between four cultural entities. Our cross-cultural investigations will take a comparative perspective, centering on **Egypt**, **Germany**, **Japan**, and the **United States of America**. In discussing the case selection, we will refer to empirical evidence that supports the assumption that the countries we selected to investigate are

characterized by specific situational logics. This section also contains a research model that is adapted to the four countries in our cross-cultural study. Next, we will introduce you to the instruments used for the cross-cultural data collection. This is where we present the **questionnaire** and the **research tools**. Hereafter we will lead over to primarily methodological discussions. First, **sample size** and **power** considerations are shared, before we will in more detail discuss **equivalence** and **bias** in cross-cultural research. This part is followed by a section on **measurement invariance** and **approaches** to be used **in the light of non-invariance**. In the latter part of this chapter, we will identify challenges of cross-cultural research and our strategies for addressing them prior to data collection, as well as our statistical strategies for safeguarding our research after data collection.

The overarching focus of *Chapter 3* is the development and testing of a scale to capture moral pluralism. We present and discuss the **Morality as Cooperation—Deviance Relevance Scale** (MaC-DRS) as a new instrument for cross-cultural moral research. After introductory remarks, this chapter begins with a concise retrospective of the core elements of our theoretical position, before moving on to self-report measures in the context of empirical moral research and the difference between moral judgment and moral relevance. We then disclose the development process of our scale and provide examples of items. In this chapter, we will draw on all three primary data collections for our empirical analyses. In **study 1** (N = 792, German student sample), we subjected MaC-DRS to an initial psychometric examination. The focus here is on exploratory and confirmatory factor analyses. In the next step, we compare MaC-DRS with leading scales from Moral Foundations Theory and Morality as Cooperation Theory in terms of psychometric properties (**study 2**: N = 2,326, German-wide non-student sample). Subsequently, we will refer for the first time empirically to the primary database of our further investigations. In **study 3** (N = 2,982, cross-cultural study comprising data from Egypt, Germany, Japan and the United States), we will test hypotheses regarding the universalism of the human moral mind as well as the psychometric properties of MaC-DRS across cultures. In this section, we will show that MaC-DRS is a useful supplement to the repertoire of cross-cultural moral research. In addition to our moral scale, we also examine a scale developed by Vignoles and colleagues (2016) to measure 8 ways of being independent and interdependent. Exploratory and confirmatory factor analyses, reliability tests, and tests for measurement invariance are used for this purpose. Overall, Chapter 3 provides the psychometric basis for the investigations in the following chapter.

Chapter 4 is the first of the following three substantial empirical investigations of the human moral mind. Overall, this chapter centers on our overarching research question —

Which moral system guides cooperation in different cultural entities? — which we will examine in this part of the thesis via **MaC-DRS** and a focus on moral intuitions. By way of introduction, we turn first to descriptive findings from our cross-cultural database to familiarize the reader with the data. We then empirically refer to culture-specific response styles. After addressing our hypotheses, we present covariate models for our analyses. Subsequently, we focus on covariate effects. Here we discuss main and interaction effects and present initial results. Thereafter, the main part of our MaC-DRS investigation begins and we estimate average marginal effects for 8 different moral domains to uncover **similarities and differences in the context of moral deviance relevance across cultures**. We will examine various cultural difference hypotheses by comparing deviance relevance margins in a pairwise fashion between the four cultural entities in our study. Furthermore, the massive **influence of culture on the calibration of the human moral mind** will be demonstrated. Thereafter we address an empirical fact that is of particular importance to us: **individualizing morality seems to be relevant in WEIRD cultures and beyond**.[2] Eventually, we focus on the dominant moral domains of the four countries we are studying. We start by interpreting the results of our analyses for Germany, the United States, Japan, and Egypt and then propose **four different moral systems** based on the MaC-DRS data. Chapter 4 concludes with a comprehensive synthesis of our analyses and results.

In *Chapter 5*, we focus on deliberate moral cognition in dilemma scenarios that contrast binding and individualizing morality. After some introductory remarks, we come to our basic model, which we will use as a starting point for further analyses. Following initial descriptive insights, we then delve deeper into the statistical analysis of 9 different dilemma scenarios. In addition to influential covariates, cultural comparisons are at the center of our analyses. The **binding vs. individualizing dilemma scenario** results corroborate our previous findings by indicating that the importance of individualizing morality is not limited to WEIRD cultures. Hence, Chapter 5 primarily complements our previous MaC-DRS investigations and expands our insight to include **deliberate moral preferences**. Furthermore, our findings, which highlight both cross-cultural similarities and differences, fit into the discourse on intuitive and deliberate moral cognition and demonstrate that the two processes do not necessarily have to lead to the same result. We conclude this chapter with a discussion of deliberate choices in moral dilemmas across cultures.

---

[2] WEIRD is an anacronym coined by Henrich and colleagues (2010a) and stands for *Western*, *Educated*, *Industrialized*, *Rich* and *Democratic* countries (Henrich, 2020). The anacronym applies to people and corresponding psychological calibrations.

The final empirical study of the human moral mind that we conducted is presented in *Chapter 6*, which revolves around the **Moral Deviance Factorial Survey** (MDFS) analysis. At the beginning of this chapter, we present, among other things, the formal structure of our research instrument and provide exemplary insights into the vignettes. We present our hypotheses and introduce ways of analyzing the Moral Deviance Factorial Survey. Thereafter, we will briefly refer to descriptive findings before commenting on our basic model and further research aspects. Next, we start the analyses. **Cultural similarities and differences** are examined across 7 moral domains. We thus continue to pursue a cross-cultural comparative approach. The focus of our analyses is particularly on whether it makes a difference in moral valuations whether a stranger, an in-group member or a family member is harmed by the act of moral deviance. Besides tabular insights into our results, we also provide easy-to-interpret graphical representations. In our examination of the MDFS data, we primarily limit ourselves to **moral particularism and moral impartiality** in order to gain further insights into our research question. In this context, we are able to present surprising evidence that points to more complex relationships than found in the relevant literature. Furthermore, our MDFS findings indicate that **moral deviance judgments and moral deviance relevance are not congruent**, which may promote the independence of both concepts. In addition, we supplement our previous insights with cross-cultural findings on the attribution of moral emotions. We conclude this chapter with a comprehensive discussion of our findings. In addition to clarifying our hypotheses, we also raise questions and **discuss the appropriateness of binary explanations in the context of the diversity of cultural realities**.

In *Chapter 7*, we will close the thesis and present a **summarizing discussion**. We focus on the key insights gained from our investigations of the human moral mind to address our research question conclusively. In addition, open questions and future possibilities for moral research as well as central limitations of our work will be discussed. Finally, we will conclude this thesis by highlighting the contributions our work has been able to make.

Overall, this project aims to explore essential questions regarding the interrelationship between **culture**, **selfhood** and **morality**. We hold that morality is part of the self and hypothesize universally shared moral domains and yet cross-cultural differences in the configuration of moral systems. Taken together, the aim of our project is to contribute to the understanding of the universal human mind and its cultural variations in order to shed light on a piece of the puzzle of human nature. Furthermore, we aim to contribute to the development of cross-culturally applicable measurement tools to advance further research on one of the core traits of our species, morality.

1.2 Culture

Humans are cultural beings. When we have time to think about something, we think in terms, and so in language. Now you are reading a text in English, but in what language do you think and interact with others in your everyday life? One thing is certain, language is a cultural element and exists in many ways. *Culture* is fundamental to us as human beings. It can likely be expected that an individual's particular cultural environment radiates into all aspects of the human psyche. Culture essentially affects our construal of selfhood, dominates the constitution of our social as well as material world, and affects even some of our basic cognitive processes (Kühnen & Hannover, 2003; Berger & Luckmann, 2013; Apicella et al., 2020). Studies support that the same stimuli shown to people from different cultures is capable to elicit differential activity in the neurocognitive processing of these stimuli (Han & Humphreys, 2016). It is hardly an exaggeration to say that culture causes people to see the world with different eyes. Also, such basic processes as visual perception is found to be influenced by culture, as demonstrated (e.g.) by the Müller-Lyer illusion (Henrich et al., 2010a). The illusion is largely about judging the length of two manipulated lines, whose endpoints

Figure 1: Müller-Lyer illusion*

Lines a and b are actually of the same length, but many people mistakenly perceive line b as being longer than a.
*The figure is taken from: (Henrich et al., 2010a, p. 64).

are encased by arrowheads pointing at opposed directions (see: ***Figure 1***). Many people tend to perceive the lines at first to be different in length although they are in fact not. However, the effect of misperceiving the lines as being different in length is no pan cultural phenomenon. People from some cultural groups, the south African San for instance, do not fall into the trap of the illusion. Albeit different explanations for the effect exist, one line forwards the notion of visual habituation to geometric shapes. Given that geometric shapes are largely absent in one´s immediate environment the effect of falling for the illusion is argued to be mitigated. However, the effect of this illusion is largely found in people from Western societies. This fact is not surprising if one accepts the argument of exposure to geometric shapes, as geometric architectural design is almost ubiquitous in Western countries, especially in urban areas. More than visual perception, our particular culture even affects our bodies and genes (Kitayama & Salvador, 2024). Changes of agricultural subsistence styles in the Holocene have not only shaped to a vast extend cultivated landscapes (e.g., rice terraces) but likely also our genes. A prominent example for such a cultural effect on genes can be found in relation to milk digestion

in adulthood (Brown et al., 2022). Most people from African and western Eurasian societies, that historically relied more on livestock farming and animal milk consumption, are capable of producing lactose even in adulthood. Whereas the LTC gene down-regulates lactose production after the nursing period in most parts of the world were societies historically relied more on paddy rice farming and other subsistence styles different from animal milk consumption (Richerson et al., 2010; Reilly, 2013). Apart from the field of genes, cultural influences can strongly affect our behavior, even in at first glance seemingly contra-intuitive ways. Let us make aware of the fact that tasting spiciness is actually nothing else than perceiving pain. Humans have usually a natural aversion against pain, more than likely a survival related tendency. Following this reasoning, thus, spicy food should be found very rarely on our species diet. However, this is not actually true around the world. Plenty of historically grown cultural cuisines are rich of cooking receipts in which lots of herbs and spices are used in preparing meals. More so, during socialization our cultural environments may affect us in developing also a taste for spicy food, even in so far that we like it regardless (or because of?) the pain (Boyd et al., 2011). So, how does it come, that tanginess found its way onto our plates? The reason for culturally divergent preferences in food spiciness is likely to be found in environmental pathogen prevalence, risks of meat deterioration and food borne diseases overall. "A clue to the ultimate reason for spice use may lie in the protective effects of phytochemicals against plants' biotic enemies" (Billing & Sherman, 1999, p. 455), i.e., bacteria and fungi. In fact, it was shown, that some herbs and spices, used alone or in combination, are effective in erasing germs, bacteria and the like, and can thus be used to prevent food poisoning. So, no wonder that especially in warm climatic areas, where meat is easy target to germs, traditional cooking books reveal the highest use of (antibacterial) herbs and spices (Sherman & Billing, 1999; Murray & Schaller, 2010).

So, we have given some examples of the rich and deep influence that culture has on the whole of our human being. But what do we actually mean by culture? That is a difficult question that has not yet been answered comprehensively in all aspects. However, if we want to get an idea of the meaning of the word culture, in order to get a sufficient understanding in terms of a working definition, the following elements are certainly of central importance. Human *culture* can broadly be conceived as a cumulative system of symbols and behavioral configurations within a social group. It is information that is stored in social narratives, traditions, institutions and the like, and gets intergenerationally transmitted. Cultural information results in material tools and products as well as in (e.g.,) values, beliefs and norms. Moreover, culture creates a shared meaning and symbolic understanding from which a shared reality among the members

of the respective cultural group emerges. In general, culture can be seen as the biological niche of humans, as it encompasses the realization of the interaction between humans and their (natural and cultural) environment, as well as the elements necessary for the survival of our species (Triandis, 2001; Markus & Kitayama, 2010; Esser, 2010; Fuchs-Heinritz et al., 2011; Smith, 2014; Kaasa & Minkov, 2020; Brown et al., 2022; Rippl & Seipel, 2022).

Only recently have researchers recognized that some other species may also exhibit aspects of culture (Whiten, 2017; Mesoudi & Thornton, 2018). However, human culture is special because it is an ever-growing corpus of transgenerational knowledge. Our culture has evolved and the framework of natural selection and adaptation underlies cultural evolution (Henrich et al, 2010a; Boyd et al, 2011; Pietraszewski, 2016; Whiten et al, 2017; Muthukrishna et al, 2021). The evolution of human culture can be conceived, at least in part, as (increasing) adaptation(s) to different environments and survival-related challenges that our species has faced during our distal and proximal history (Lehman et al, 2004; Henrich & McElreath, 2007; Gelfand et al, 2017). Human culture is thus subject to the process of cumulative evolution (Henrich & McElreath, 2007; Creanza et al., 2017; Mesoudi, 2017). Cumulative culture can be understood as an ever-growing body of (adaptive) cultural features that grow over time and generations (Heyes & Moore, 2021). Several core and extended criteria characterize cumulative cultural evolution (Mesoudi & Thornton, 2018). At the heart of cumulative cultural evolution is a behavioral novelty or behavioral change that is passed on through social learning. This new behavioral variant improves an existing behavior, i.e., it increases genetic and/or cultural fitness or its proxy, and is passed on repeatedly over time. The latter leads to sequentially accumulated behavioral changes known as the cultural ratchet effect (Tomasello, 2017). In addition, human cumulative cultural evolution may also include the following extended elements: new behaviors build on previous improvements, which can lead to functionally dependent chains of improvement where further innovations/modifications are functionally dependent on previous ones. In addition, diversification into multiple lineages of parallel behavioral improvements — bow and spear use as an example of diversified behavioral improvements in game hunting —, recombination by (e.g.) integrating tools from different contexts to functionally work in another, and cultural exaptation implying a functional change (use of one tool in different functional domains) can be mentioned as further elements. Finally, the construction of (cultural) niches, a typically human characteristic, is an element of our growing cultural corpus: culture is able to change our natural environment and shape it into a cultural ecology (Brown et al., 2022).

Essentially, over millennia, our species has altered its natural environment and created cultural ecologies as primary human environments, which in turn have resulted in

corresponding selection pressures and adaptive requirements. Considering divergent initial natural ecologies, cumulative cultural evolution and resulting specific cultural ecologies, which in turn imply independent historical path dependencies, the following becomes clear: on the one hand, culture is one of the universal and fundamental characteristics of humanity (Henrich et al., 2008; Richerson et al., 2010; Whiten et al., 2017). However, on the other hand, it is human culture that also produces variant forms of human living as well as cross-cultural differences in psychological tendencies (Henrich et al., 2010a; Henrich, 2020). Culture is thus a distinctively human and universal, yet diverse feature of our species. Furthermore, it follows from the logic of the growing corpus of adaptation that culturally constituted psychological tendencies are inherently meaningful in relation to their respective contexts of emergence. In this context, ethnocentrism, which should be rejected on principle, also loses any non-normative basis for argumentation (Toobey & Cosmides, 1998; Bizumic & Duckitt, 2008; Buzimic et al., 2009).

To produce, transmit and acquire culture, our species evolved a specialized brain and psychological apparatus (Henrich, 2016; 2020; Heyes & Moore, 2021). Among our evolved specializations for culture the following take prominent roles: learning- and norm-psychology, neuroplasticity, theory of mind, shared intentionality and skills for imitation as well as coalitional psychology, language abilities and our ability for interindividual cooperation (Kurzban et al., 2001; Gintis, 2003; Henrich et al., 2008; Chudek & Henrich, 2011; Muthukrishna et al., 2016; Tomasello, 2017; Hare, 2017; Gavrilets & Richerson, 2017; Shilton et al., 2020). Let's highlight some of the aspects of our evolved mind that help us in the processes of acquiring cultural information and socio-cultural adaptation. Essentially, our evolved learning- and norm-psychology are pivotal in this regard (Chudek et al., 2016). "We are adaptive learners who, even as infants, carefully select when, what, and from whom to learn" (Henrich, 2016, p. 4). Learning can be individual (asocial), social or cultural. While the former type of learning rests mainly on the individual observation of and interaction with the environment, the latter two types rely on observation and imitation of social others. So, regarding culture we are mainly concerned with the latter two. Our evolved learning psychology equips us with a range of mechanisms to effectively acquire (relevant) information from our social world (Muthukrishna et al., 2016; Creanza et al., 2017). This component of our mind leads us to prefer and focus automatically on information shared by specific social learning models. Characteristics such as same gender and ethnicity, as well as skill, success and prestige, serve as social cues that direct our attention. These cues make it particularly likely that we will imitate and adopt information from social learning models that have one or several of these characteristics. Hence, we are equipped with cognitive biases for selective attention towards

adaptive learning contents and models (social others). Our cultural psychological apparatus also encompasses a tendency to acquire (cultural) information via mentalizing and by drawing social inferences about others goals, beliefs, norms etc. (Henrich & McElreath, 2007; Henrich et al., 2008; Henrich & Muthukrishna, 2021). We have a theory of mind and make sense of other people´s behavior by holding that they behave due to intentions, beliefs, desires and knowledge. One component of our theory of mind is mentalizing or the ability to read the minds of others. This ability is pivotal in acquiring "knowledge about other peoples' beliefs and desires" (Frith & Frith, 2005, p. R645). Beyond that, our norm-psychology can be regarded as "a suite of psychological adaptations for inferring, encoding in memory, adhering to, enforcing and redressing violations of the shared behavioral standards of one's community" (Chudek & Henrich, 2011, p. 218). The tendency of our evolved mind to align with the behavioral norms prevalent in our group life requires that we pay close attention to the behavioral standards of our social environment. This means that we pay close attention to behavior that is acceptable and rewarded or unacceptable and punished (Ormel et al., 1999). Our evolved norm-psychology is thus not only relevant in reproducing and upholding social standards but also in individual sociocultural attunement.

Aside from our specialization in culture, however, one may ask how it is that we come to rely so much on culture in the first place. In terms of the evolutionary background to this question, we can first note that the capacity for social learning is likely to have evolved genetically as an adaptation "for surviving in environments in which individually acquiring information is costly" (Muthukrishna et al., 2016, p. 10). In this line, simulation models suggest that under conditions of intermediate environmental change-rates, evolution appears to be too slow to track changes via genetic adaptations but not via cultural ones (Richerson et al., 2010; Mesoudi, 2017; Henrich & Muthukrishna, 2021). Under these conditions natural selection favors cultural learning and increased our species' reliance on cultural information. The ability to produce cultural knowledge in order to adapt to a changing environment, in addition to the ability to transmit and acquire this knowledge, represents a fundamental fitness advantage for our species. In conclusion, research suggests that our species' heavy reliance on culture and our sophisticated learning psychology likely evolved and intensified because "[c]ulture is adaptive when asocial learning is hard and environments fluctuate" (Chudek et al., 2016, p. 753).

Overall, the extent to which we developed and rely on culture is unique in the animal world of our planet (Henrich, 2016). Thus, for the human species culture is conceivable as, besides genes, a second system of inheritance (Henrich & McElreath, 2007; Boyd et al., 2011) that encompasses "all of the information that individuals acquire from others by a variety of

social learning processes" (Richerson et al., 2010, p. 8985). Conclusively, despite the lack of a comprehensive definition, we know that culture evolved, is at the essence of the living of humankind, varies globally (to functionally promote living in context), and substantively affects our psychological tendencies (Henrich, 2020; Kitayama & Salvador, 2024). As Chudek et al., (2016) have put it: "[u]nlike all other species, we are addicted to culture" (p. 750).

1.2.1 A Glimpse at Cultural Dimensions

Our world is rich in cultures and research has identified many *relative* differences and similarities between cultural entities (e.g.: Murray & Schaller, 2010; Cross et al., 2011; Thomson et al., 2018; Schulz et al., 2019; Haerpfer et al., 2022). If the idea is to compare cultural entities with one another, i.e., to assess commonalities and differences, this must be done using cultural dimensions. Comparison gains meaning only if we have a common dimension upon which we can place the entities and which gives us a sense of the relative difference (or commonality) of the entities to be compared (on that dimension). Cultural dimensions, then, can be thought of as ideal-type variables that describe relative social and psychological patterns that emerge for the particular groups under investigation (Layes, 2003; Barmeyer, 2010). In our project, we are primarily concerned with the cultural dimensions of morality and self-construal. Nevertheless, to give a glimpse at the idea of cultural dimensions and the richness of cultural characteristics, we will touch on a few of them in the following.

The way information is shared can be regarded a cultural dimension. Some cultures show a preference for high context communication while others show a low context communication style. Information in the latter is stated rather explicitly context unbound. In contrast, in cultures that are high in context bound communication prior contextual knowledge plays an essential role. In entities marked by this characteristic there is lesser need for explicit codes as a common store of knowledge already contributes to meaning and understanding (Barmeyer, 2010). Also, power distance, i.e., the degree of accepting hierarchical relations and deferring to authorities, or the degree of marked differences in gender roles (cultural masculinity-femineity) pose, for example, other cultural dimensions (Smith, 2014; Żemojtel-Piotrowska, & Piotrowski, 2023).

Communication styles, power distance and cultural masculinity-femineity give a hunch on the variety of characteristics that reveal similarities and differences between cultural entities. However, certain cultural dimensions seem particularly relevant in the context of interpersonal cooperation, which is governed by our moral mind. As we will prominently highlight

throughout the course of this text, we grasp the culturally contingent modes of selfhood to correspondent to cross-cultural variance(s) in cooperation (and vice versa). However, we will touch self-construal more deeply below and focus for the moment on some other cultural dimensions associated with cooperation. Among the many cultural dimensions that could be described as relevant to cooperation are tightness/looseness, cultural logics (honor, face and dignity) and individualism-collectivism, which we would like to briefly highlight.

Cultures may vary in their degree of norm adherence, a dimension know as (normative) tightness/looseness (Gelfand et al., 2017). Historical threats, be it natural hazards, pathogen prevalence in the environment or intergroup conflicts, are known to affect a culture´s adherence to norms (Roos et al., 2015). On the one hand, those cultural entities historically more affected by exposure to threats are also found to be tighter in norms which goes hand in hand with stronger sanctioning of norm violation. The social rules governing interaction are being uphold strongly in tight cultures. On the other hand, cultural entities less affected by historical exposure to threats are evidently looser, and thus have a greater permissiveness to norm violations (Gelfand et al., 2011).

Beyond norms, cultures may also be compared on the ground of multiple (cultural) logics. A cultural logic can be considered an overarching frame of reference and guidance for people "to effectively engage with the contingencies of their social environment" (Uskul et al., 2023, p. 6). Cultures are found to vary in the logics of honor, face and dignity (Leung & Cohen, 2011). Each of these logics emphasizes a social orientation relevant to cooperation and deals with individual self-worth, but in different ways. Honor cultures are among several aspects characterized by strong norms of reciprocity, for the value of a person, it´s honor, has not just internal but also external sources, and can be contested (Uskul et al., 2019). "Honor must be claimed, and honor must be paid by others. A person who claims honor but is not paid honor does not in fact have honor" (Leung & Cohen, 2011, p. 3). Different from honor, and even to a higher extend do face cultures direct the focus of a person's worth to external sources. Deference to hierarchies, role fulfilment and not losing face in front of others (as not making them to lose face) are aspects of this cultural logic that is essentially characterized by interdependent social relations. Dignity cultures, in contrast to the already mentioned cultural logics, are themselves distinguished by an orientation towards a person´s inherent worth. Individuals are expected to possess internal standards that give consistently guidance across social situations. The person's worth is rather detached from external sources and one's dignity can't be lost which is signifying overall an independent social orientation in dignity cultures (Leung & Cohen, 2011).

Relationship (group) orientation or individual orientation is also a distinctive feature of the cultural dimension collectivism-individualism (Triandis et al., 1990; Triandis & Gelfand, 1998; Triandis, 2001; Żemojtel-Piotrowska & Piotrowski, 2023). In collectivist cultures people´s primary orientation is towards their (in-)group: behavior is tight to group norms; the goals of the group are paramount and the focus is on maintaining one´s relationship with the in-group and its members. In stark contrast, the culture level orientation of individualism fosters independent, autonomous behavior, personal goal pursuit and reliance on intra-individual attitudes rather than in-group norms (Henrich, 2020). As will come apparent later on, it is of no wonder, that predominance of the independent self is found in cultural entities that tend towards individualism while the interdependent selfhood is mostly the prevalent mode of self-construal in collectivist cultures (Markus & Kitayama, 2024). It should, however, be noted, that individualist and collectivist cultures are themselves heterogenous and the specific configuration of the self is more complex than being merely determined by collectivism and individualism (Krys et al., 2022). Beyond that, evidence supports the notion, that moral tendencies may differ between collectivist and individualist cultures (Miller et al., 1990). Whereas choice seems to matter in regard to helping behavior for people from individualist cultures (Markus & Schwartz, 2010), helping the in-group is rather seen as a duty (without choice) in collectivist cultures. More so, there are indications, that the scope (expansiveness) of pro-sociality varies across individualist and collectivist cultures. As for the latter morality is expected to be "not applicable to all but only to some members of one's social environment" (Triandis, 2001, p. 917), i.e., to members of one´s in-group. Morality, in the realm of cross-cultural commonalities and differences, will be our major concern of this text.

We will elaborate on our morals in general and also in relation to culture more deeply further below. From our short take on culture and cultural dimensions, we will yet for the moment lead over to the other central focus of our work, i.e., the individual level variable of the self and the culture contingent modes of self-construal.

1.3. The Human Self

1.3.1 Independent and Interdependent Self-Construal

What is the first thing that comes to your mind when you are asked to fill in the sentence "I am ..." five times?[3] Depending on where you grew up, this question was, is, and probably will be answered differently, because our self is fundamentally socially constructed and inextricably linked to the cultural environment of the person. All humans are bearers of a self, yet defining aspects of the self vary across cultural entities. The way we define our self in relation to others, i.e., our *self-construal* is fundamentally guided by our cultural surrounding (Markus & Kitayama, 1991; 1998; 2010; Cross et al., 2011). Furthermore, as will become evident throughout our focus on the self, the relationship between a respective cultural ecology and the configuration of the self is not arbitrary, but rather functional.

Culture and self are inextricably linked, they are essential constituents of each other (Markus & Kitayama, 2010). Accordingly, as cultures vary it is unsurprising that empirical evidence also suggest differences in human self-construal across cultures (Cross et al., 2011). Based on the distinction between more collectivist and more individualist cultures (Layes, 2003; Barmeyer, 2010; Minkov & Kaasa, 2022; Żemojtel-Piotrowska & Piotrowski, 2023), two divergent, overarching modes of selfhood are proposed theoretically and are also evident in empirical research (Henrich, 2020). At first glance, these different modes of self-construal seem to overlap almost perfectly with the cultural syndromes of individualism and collectivism. This overlap could be due to the fact that the different modes of self-construal differ in their focus on autonomy or on relationship(s), which also characterizes the culture-level dimension individualism-collectivism (Kitayama & Salvador, 2024). Motivated by these findings scholars of cultural psychology begin thus to differentiate between the ***independent self-construal*** and the ***interdependent self-construal*** (Markus & Kitayama, 1991; Dimaggio & Markus, 2010). It is the interdependent mode of self-construal that has its focal point on relationships with relevant others, while the defining aspect of the independent way of selfhood is on being an autonomous, independent person (Markus & Kitayama, 1991; Cross et al., 2011). The cognitive availability of seeing one's self as essentially connected to others is likely to reinforce the tendency to respond to the "I am question" with something like: I am a loving mother/father; I am a person who strives to live in harmony with others; I am similar to my dear friends, and I am a football player of team XY. Furthermore, as a tendency it is only later that a response

---

[3] Asking the "I am…?" question is one of the classic approaches to measure self-construal (Cross et al., 2011).

option such as "I am XY years old" pushes itself into the cognitive foreground of people who define themselves as interdependent, i.e., principally relationship oriented and interconnected with others. In contrast, people who see themselves rather as a single, autonomous unit of being are more inclined to answer the "I am question" with: I am XY years old; I am always the same; I am a unique individual; I am an open-minded person, for instance. Here, too, possibilities such as "I am a football player of team XY" tend to sneak into the foreground of consciousness as answers only later for they are less available to people that define themselves as independent of others. What we just described are cognitive differences in the availability of self-related attributes that diverge in tendency due to the different social orientations that are at the core of being interdependent or independent in self-construal. However, the way we define ourselves in relation to others does not only influence our cognition, it is more so the whole psychology of a person, its thoughts, feelings, motivations and actions, that are subject to culture relative ways of selfhood (Dimaggio & Markus, 2010). In line the *independent self-construal* is found to be associated with a cognitive tendency to differentiate information (Nisbett et al., 2001). Also, expressing and experiencing more socially disengaging (ego-focused) emotions and pursuing more personal goals is characteristic of the independent self-construal. Nearly opposed to these tendencies, the mode of the *interdependent self-construal* is empirically found to possess more socially oriented motives, other focused emotions and a tendency to connect information (Markus & Kitayama, 1991; Cross et al., 2011; Morf & Koole, 2014; Park et al., 2016). Beyond that, studies from the field of neuroscience demonstrate that one and the same stimulus can elicit deferential brain activity when comparing people holding an independent or interdependent self-construal. Han and Humphreys (2016) describe for instance that self-construal primes affect empathetic neural responses to stranger's physical pain: empathetic reactions were "decreased by interdependent self-construal priming among Chinese (…) but increased by independent self-construal priming among Westerners" (p. 12).

Having witnessed these substantial differences in psychological tendencies that are due to the different ways of self-construal, one may wonder where these differences originate from. An explanation widely shared among scholars is provided by Kitayama and Imada (2010). They suggest that the different tendencies and the distinct frames of social orientation of the self are due to a process of individual adaptation to one´s immediate social environment: "in adapting to their own cultural context, people are motivated to be independent or interdependent in accordance with the overarching imperative of the culture" (p. 179). We will discuss this point in more detail below.

As has already been emphasized, it is the cultural context that promotes the particular constitution of the self. This view, however, does not entail absolute differences in selfhood across cultures. To avoid misunderstandings, it needs to be pointed out that there is only a *relative* difference between cultures in regard to self-construal, as all types of self-construal likely exist to some degree across cultural entities. Nevertheless, although being relative differences, these differences in selfhood are potent in affecting psychological tendencies in particular directions, as described above (Brewer & Gardner, 1996; Markus & Kitayama, 1998; Gardner et al., 1999; Utz, 2004; Flinkenflogel et al., 2017). The differentiation between the independent and the interdependent self was a groundbreaking theoretical contribution to the understanding of cross-cultural commonalities and differences among humans. In the course of further research, however, it has become clearer that the relationship between culture and self is more complex than the mere distinction between two ways of selfhood. Recent findings support the view of individual differences in the acquisition of culturally patterned self-syndromes as well as the existence of multiple ways of being independent and interdependent (Park et al., 2016; San Martin et al., 2018; Uskul et al., 2023).

1.3.2 Intra-Cultural Variation in Selfhood

Contexts differ and afford variable menu-sets of available means to engage in culturally constituted live (Habermas, 1995; Esser, 2002a; Lehman et al., 2004; Ridgeway, 2006; Berger & Luckmann, 2013; Oyserman, 2016). Plenty of research shows profound differences in self-construal across cultures (Myers et al., 2010; Cross et al., 2011). Yet ways in which the self is construed also vary within cultural entities. Likely only a few view people will oppose the statement, that from the beginning of individual existence itself, each individual inhabits its very own space in time. Under this self-evident assumption, it seems inevitable that there must also be individual ways to acquire the respective cultural imperatives and mandates, because individual contexts differ in their nature. In other words, although culturally patterned syndromes exist, there is also individual and intra-cultural variation in acquisition and expression of the respective syndrome.

The cultural tasks approach by Kitayama and colleagues advocates this position and delivers an explanation needed to hold it (Kitayama et al., 2009; Kitayama & Imada, 2010; Park et al., 2016). Cultural mandates are understood as "ideals or general goal states that are strongly sanctioned and encouraged by a given cultural group" (Kitayama et al., 2009, p. 237). These mandates become prevalent in social interaction which is itself organized in culturally

shared ideas, norms, institutions, traditions and the like. Further, our evolved mind, encompassing a learning psychology and mentalizing abilities, supports the individual acquisition of culture (Chudek & Henrich, 2011; Henrich, 2016; Heyes & Moore, 2021). Since each individual inhabits its own space in time, assume now that also the individual´s context differs in opportunity structures, i.e., in the respective affordances of cultural tasks (means) to achieve a cultural mandate. If so, it follows that individuals likely engage in various ways to acquire the dominant cultural mandate and imperative(s) of the sociocultural ecology they find themselves embedded. In this line it is argued that certain psychological tendencies get "gradually fostered, acquired, and internalized through repeated engagement in and performance of pertinent cultural tasks" (Kitayama & Imada, 2010, p. 178). Eventually, this process of individual cultural task engagement leads to intra-cultural variation, which means variation of a cultural syndrome at the individual level. Empirical evidence on self-construal is supporting this view (Kitayama & Imada, 2010; Park et al., 2016, San martin et al., 2018).

1.3.3. The Multitude of Being Independent and Interdependent

Several studies point to associations between ways of selfhood and individualistic and collectivistic tendencies of cultural entities (Henrich, 2020; Kitayama & Salvador, 2024). However, the binary distinction between cultures with more collectivist or more individualist aspects is generally, although valuable, better understood as a heuristic, which also applies to its relation to self-construal. We need to pay more attention to both intra- and intercultural diversity and not assume that collectivist or individualist cultures are inherently homogeneous. With respect to the self, an illustrative example of this diversity shows that even more collectivistic cultures can promote the dominance of the independent mode of being (Vignoles et al., 2016; Krys et al., 2022). In general, a great deal of cultural diversity can be observed between and within cultural entities (Henrich et al., 2010a; Apicella, 2020).

In this light, even though absolutely valid for its parsimonious explanatory power — science must not only aim to add complexity via novel findings but must also aim to decrease it by the use of parsimonious models of explanation —, the two-type model of self-construal does not seem to adequately represent the configuration of human selfhood across the world (Vignoles et al., 2016). Thus, our culturally constituted ways of being were often times laid out too simplistic and appear to be more diverse in fact.

Refinements to the two-pillar model have already been made by differentiating the interdependent self into relational and collective interdependence (Brewer & Gardner,

1996). The relational mode is considered to be more inclined with a focus on dyadic relationships, while collective interdependence is more concerned with a focus on the (in-)group. Another differentiation was suggested by empirical findings of San Martin and colleagues (2018) who discovered a new category of selfhood prevalent in Arab honor culture named self-assertive interdependence. This mode of selfhood shows likewise classical attributes of both, the independent and the interdependent self, and is therefore, in respect to the previously known modes self-construal regarded as a mixed category. Besides these few examples, an attempt to expand the search for the possibility of more variations in selfhood was for long complicated due to unreliable (explicit/self-report) measurement instruments (Singelis, 1994; Triandis & Gelfand, 1998; Cross et al., 2011; Park et al., 2016). But Vignoles and colleagues succeeded in 2016 to develop a new tool to empirically capture the self, based on a refined multi-dimensional approach to independence and interdependence (Vignoles et al., 2016; Yang, 2018; Krys et al., 2022; Uskul et al., 2023). Without abandoning the value of the original two-pillar approach to selfhood, these authors argue that the human world offers many ways of being independent and interdependent. Moreover, the link between the self and contextual adaptation, already referred to above, is advocated by the authors too and additionally used as an explanation for multiple differences in selfhood: "we view cultural differences—including models of selfhood—in part as adaptations to differing ecological and sociopolitical circumstances" (Vignoles et al., 2016, p. 971). This refined model suggests *8 individual and cultural ways of being independent* and *interdependent*. The model includes differentiations between the following dimensions (with the independent mode in parentheses):

> "1. *Similarity (vs. Difference)*; 2. *Connection to Others (vs. Self-Containment)*; 3. *Receptiveness to Influence (vs. Self-Direction)*; 4. *Dependence on Others (vs. Self-Reliance)*; 5. *Variability (vs. Consistency)*; 6. *Harmony (vs. Self-Expression)*; 7. *Commitment to Others (vs. Self-Interest)*; 8. *Contextualized (vs. De-contextualized) Self*" (Uskul et al., 2023, p. 14).[4]

Empirical evidence, based on research conducted with the novel self-report scale, supports the multifacetedness perspective of independence and interdependence in selfhood (Vignoles et al., 2016).

Altogether, this recent approach to self-construal holds the following: self-construal is multifaceted. The 8 dimensions of selfhood are contrasted by two opposed poles indicating either a preference for a certain way of being independent or interdependent. More

---

[4] The emphasis by italicization is not from the original and was added by the author.

so, the contrasting poles of each dimension are mutually exclusive. Consequently, per dimension a clear preference of being either independent or interdependent in self-construal emerges from this model. This is, however, not to say that people must prefer 8 uniform ways of being interdependent (or vice versa). On the contrary, the individual potpourri of selfhood may reveal various arrangements of preference patters for being independent and interdependent. Presumably more realistic and complex structures of self-construal can be captured via this model as individuals may hold, for example, three preferences for *independent* dimensions of selfhood (i.e., being: different from others, self-contained and self-directed) while also holding five preferences for being *interdependent* in self-construal (i.e., being dependent on others and flexible in self-adjustment to context, preferring harmony over self-expression and commitment to others over self-interest, and holding a focus on contextual information to define one´s self-construal). Thus, each dimension of the 8-factor self-construal model "represents a choice for the individual about whether to think/feel/act in a relatively independent or interdependent manner within a given domain of personal and social functioning" (Vignoles et al., 2016, p. 976). To illustrate the 8-dimensional model of self-construal at one glance we adopted a table by Vignoles and colleagues (2016). The synopsis of the multifaceted approach to self-construal can be found below in ***Table 1***.

Table 1: The 8-Dimensional Model of Self-Construal*

Domain of functioning	*Independent* **way of being**		*Interdependent* **way of being**
Defining the self	Difference	↔	Similarity
Experiencing the self	Self-containment	↔	Connection to others
Making decisions	Self-direction	↔	Receptiveness to influence
Looking after oneself	Self-reliance	↔	Dependence on others
Moving between contexts	Consistency	↔	Variability
Communicating with others	Self-expression	↔	Harmony
Dealing with conflicting interests	Self-interest	↔	Commitment to others
Integrating contextual information to understand the self **	De-contextualized	↔	Contextualized

* The table is taken from: (Vignoles et al., 2016, p. 976). ** The De-contextualized vs. Contextualized dimension was integrated to the model by: (Yang, 2018). We have added this dimension to Table 1.

1.3.4. The Self´s Evolved Peculiarities

There are many ways to grasp and deal with the self (Abrams, 1999; Hogg et al., 2004; Myers et al., 2010; Morf & Koole, 2014). However, here we would like to draw attention to *the self as an evolved functional component of the human mind*, that is supporting survival in given ecology (Baumeister, 2022). The human mind evolved by a process of adaptation to the requirements of survival and reproduction. In the course of the evolution of the human mind we increasingly became capable to adapt to changing conditions and varying situations. Based on our naturally and culturally coevolved psychological apparatus (Henrich & McElreath, 2007; Henrich, 2016; Chudek et al., 2016; Tomasello, 2017; Hare, 2017; Heyes & Moore, 2021; Brown et al., 2022) we are equipped with the ability to adapt relatively flexible to context. A characteristic of our mind that makes us capable to fit in into a huge variety of natural and cultural ecologies. Fitting in into context means being able to orient, navigate and act within one´s environment in meaningful ways. In this process of adaptation, the *human self* plays an important role.

We can conceive the human self as a flexible mental process that processes, integrates, and actualizes incoming (self-relevant) information. These functions enable us to attune ourselves to the immediate sociocultural context (Abrams, 1999; Turner & Reynolds, 2012; Oyserman, 2016). More so, and simultaneously, our self is an enduring self-system of knowledge: less chaotic than structured it is built upon a repertoire of self-schemas at work, of which our self-construal is one. Self-schemas are apprehensible as mental structures that direct our perception and encode, i.e., organize and guide past experiences as well as newly incoming self-relevant information (Kühnen & Hannover, 2003; Gazzaniga et al., 2017). Further, these schemas play a potent role in evaluating information that is relevant to our self. Apparently, our self is omnipresent in processing and structuring our experience. In this line Markus and Kitayama (1991) note, that whenever an event in our world is relevant to our self "the ensuing processes and consequences are likely to be influenced by the nature of the self-system" (p. 230). This nature of our self-system is always culture-dependent (Oyserman, 2016). Or, to put it another way, our self is a constant process in flux that connects us to the sociocultural ecology that is surrounding each of us (Baumeister, 2022). Thus, having a self is essentially functional because it supports our ability to adapt, a capacity that in turn is vital in promoting reproduction and survival (Nettle et al., 2013). Consequently, the way our selfhood is construed cannot be viewed as something divorced from context: the cultural constitution of selfhood influences our cognition, affect, motivation, and behavior, bringing them into alignment with context and

enabling us to successfully navigate and interact in the particular sociocultural sphere that surrounds us (Markus & Kitayama, 2010; Cross et al., 2011). Evidently, as cultural ecologies differ in the demands that they pose at individuals and in the affordances that they offer to individuals it becomes clearer that the human selfhood is contingent on them. Nonetheless, there is more to the individual self as being (passive) subject to notably constituting sociocultural influences. As human beings we are agentic actors navigating in, producing and reproducing the social world of ours (Baumeister, 2010).

1.4.1. The Self Reflects Past, Current and Future Sociocultural Constitution — Integrating the Self into the Model of Sociological Explanation

We have illustrated that the self is fundamentally social and conceivable as an individual entity in which the respective cultural systematics are reflected. This finding is yet descriptive of only one path of the interrelation between (cultural) context and self. To form our argumentation, we need to integrate both pathways of the mutuality between self and sociocultural context; our self is a culturally constituted entity of being and simultaneously a constituting, agentic entity capable of action in context (Myers et al., 2010; Baumeister, 2010; Kühnen & Kitayama, 2020). Relating this to culture, it means that the individual is product and producer of its sociocultural reality (Ridgeway, 2006; Berger & Luckmann, 2013), and the self is shaped shaper of its sociocultural environment (Jaspers, 1956; Markus & Kitayama, 2010, p. 421). Thus, the relation between sociocultural context and self is two-sided: the self is subject to sociocultural structuredness yet also forming actor of its sociocultural world (Baumeister, 2010; Morf & Koole, 2014). Dimaggio and Markus (2010) illuminated this two-sided relation and parse it in the following statement: "cultures and selves (…) constitute each other in a cycle of mutual constitution" (p. 351). We follow this view. Moreover, in doing so we conclude that the culturally embedded individual entity of the self is conceivable as a reflective moderator of past, current, and future sociocultural constitution (Dimaggio & Markus, 2010). Since this notion is essential for the framework of our argumentation, we will spell it out more deeply in the following by introducing the Model of Sociological Explanation.

1.4.2. The Model of Sociological Explanation (MSE)

Hartmut Esser's ***Model of Sociological Explanation*** (**MSE**) focuses on social situations of (inter-)action and on actors as carriers of social structures (Esser, 2010). The model is intended to explain and understand the meaningful but often also automatic, unconscious actions of actors as well as social dynamics, i.e., persistent and/or changing social structures. The MSE serves as a heuristic: it combines basic assumptions about the object of a social science analysis (what is) with a (causal) explanation of it (why is it) (Greshoff, 2008). As already outlined with regard to the self, the human being is here also conceived of as the product and producer of the social, as will be explained in more detail. In addition, the MSE advocates methodological individualism. Emergent social structures can therefore only be understood by referring to the individual and their actions. Moreover, the MSE rests on basic evolutionary assumptions and is readily compatible with the overarching gene-culture coevolutionary theoretical framework of our endeavor. We will elaborate on this below. Furthermore, by drawing on the MSE, we will shed light on the outcomes of actions with regard to the preservation of the respective social system and the reproduction of an individual organism in general. We will also touch upon outcomes of social action in the context of self-construal and morality in the further course. A socially dynamic explanation of action and action consequences can be afforded via the MSE. Hence, we strive to integrate the self and morality into the Model of Sociological Explanation (MSE). The overarching functional sense of an action can be regarded as the causal impact of the action for upholding a social system and reproducing the individual organism.[5] It should, however, be mentioned that the function or functional sense of an action is not to be understood as a telos inherent in action. Rather, the organism and system-maintaining function may appear as emergent phenomenon resulting from the interplay of actions that comprise also the possibility of other outcomes. More so, the contribution to the functioning of social system preservation and organism reproduction, which can emanate from actions, may not be conscious to the acting actors themselves. Thus, the functional outcome of an action may be an unintended, unconscious consequence that results from the actions of the actors and goes beyond the conception of the individual (Esser, 2010).

---

[5] Nonetheless, it should be noted that we are not building on respectively referring to a metaphysical functionalism. On the contrary, we are merely advocating methodological functionalism. In other words, we refrain from a metaphysical overall explanation of the human mind and only want to assume relationships that could be generalized from the logic of induction if a large number of empirical findings were available to confirm them (Brüntrup, 2004). This does not mean, however, that we do not also make deductive assumptions in this work. We use deductive inference as an important part of our reasoning, which first precedes empirical testing and then awaits empirical falsification/confirmation.

We apply the MSE in this work to derive our hypotheses from the model and the fundamental human tendency to adapt. Furthermore, we use the MSE to explain why we assume our hypotheses. In what follows we will first outline the main features of the MSE. In doing so, we will integrate the self-construal approach into the model. Building on this, we will present an initial model (***Model 1***). This model will be further differentiated within a later section on the universal yet culturally contingent factor of morality (***Model 2***) and leads over to our hypotheses. Eventually, the differentiated model is transformed into an adapted research model (***Model 3***; ***Chapter 2***), which forms the backbone of the hypotheses examined in the studies we conducted (***Chapters 3–6***). But before coming to this part, let us turn to the ***Model of Sociological Explanation*** (***MSE***) and its basic components. Here we will only provide a clearly abbreviated introduction that serves our purpose of argumentation. Furthermore, our outline of the MSE presented below draws heavily on the texts by Esser (2010) and Greshoff (2008). Interested readers are therefore referred to Esser and Greshoff for more detailed information on the MSE (see also: Ormel et al., 1999; Esser, 1999; 2002a; 2002b; 2002c; Greshoff et al., 2011; Diekmann, 2013; Tutić, 2023).[6]

The basic assumption of the Model of Sociological Explanation (MSE) is the production and distribution of material and immaterial **resources** that the human individual needs to live and survive, i.e., to reproduce oneself and the social of the human organism in general. The background to this assumption is the Theory of the Social Production Function. Furthermore, it is assumed that the individual experiences what is useful for survival through the functioning of their actions in a given situation. In this context, action is regarded as social action, because at least two actors and actions directed towards each other are assumed for a social situation. The purpose of action itself is referred to as ***utility***. Here, the further basic assumption is made that individuals act (subjectively) in a ***utility-maximizing*** way. However, this is introduced within the framework of bounded rationality and the inclusion (and emphasis) of unconscious, often habitualized actions and potentially unintended consequences. The utility-maximizing assumption is therefore not to be confused with the classic assumptions of homo economicus. The maximization of utility is understood as an evolutionarily stable strategy for guiding action and forms the basic rule of action in the MSE. Moreover, utility maximization is related to two basic variables: the (external) environmental conditions and the (internal) functional conditions. The former forms the basis for ***expectations***, i.e., the probability that a consequence will occur

---

[6] We have largely refrained from further references in the outline of the MSE, as all the necessary information can be found in the literature already quoted. However, it should be mentioned once again that we explicitly point out that we are tracing the MSE in our text based on Esser (2010) and Greshoff (2008).

as a result of an action if the action is selected under given external conditions. The second forms the basis for *evaluations* of selected actions. Evaluations are important because they relate to the basic rule of utility maximization, since actions can have different returns. Utility maximization of action for the preservation of the organism thus takes place against the background of expectations towards and evaluations of the consequences of one's own actions, which in turn is evolutionarily beneficial and conducive to the reproduction of the organism.

Furthermore, it is assumed that resources are distributed in such a way that the individual does not readily have all the resources required for the reproduction of one´s organism. Therefore, the satisfaction of basic needs (i.e., social appreciation and physical well-being) and what is conducive to survival cannot be provided, for the most part, by the single individual. This is where the motive for social action lies: actors enter into reciprocal social relationships to acquire and exchange resources. Social situations therefore have a material-strategic structure in which resources are at the center. This *social structure*, in turn, is both a product and a producer of social action and unfolds in three types of socially valid expectations: *first*, based on the reciprocal social actions of actors, opportunity structures, i.e., expectations of the material interests and opportunities of the social situation, emerge as an emergent structure. *Second*, social rules of action, which are anchored in institutions, emerge as an emergent structure. *Third*, *cultural frames* emerge as an emergent structure. These frames are mental models of the socially recognized interpretation of the social situation. We will discuss frames also in more detail later and emphasize that this is where we locate the role of the self, with the respective foci of social independence or interdependence, in the MSE. First of all, however, it should be mentioned that cultural frames form a frame of reference that provides actors with orientation for action. Decisive for the activation of a specific and not another cultural frame are *significant symbols* anchored in the social situation. These symbols are furthermore linked to particular codes and scripts. If the symbol of a specific *code* is found in the social situation, then the description of the overarching goal of the situation emanates from this. Hence the situation receives its typical meaning through the code. So, what the situation is about and what interests and values make up the situation emerge from the code of the cultural frame. Codes are therefore also the basis for *evaluating* the results of the action(s) in the situation. **Scripts**, in turn, reveal the action models of the situation. They translate the evaluations contained in the codes into action and create *expectations* about the effectiveness of the action and about the typical (expected and socially approved) action in the situation. Expectations of action, valid rules, typical patterns of action, action sequences and consequences of action, therefore, emerge from the respective script. In summary, the (respective) cultural frame serves the actors to

identify what is socially valid in the situation on the one hand and the typical choice of action based on the code in a given social situation on the other. This is the general background of the MSE.

In the following, ***three logics*** are decisive in order to gain an understanding of actions, to grasp causal effects, and to build a bridge between the (analytically separated) levels of micro and macro. The micro-level refers to the individual-level of the actors whereas the macro-level refers to the level of social structures. To transition to this point, let us first shed some light on the latter. Social situations are located at the macro-level and include action-related social structures. Social structures in turn, according to the MSE, emerge from aggregated sequences of action at the micro-level of actors. The actions in a situation are therefore caused actions, and the structure of the situation is in turn an effect of the actions. Consequently, ***micro- and macro-levels*** are mutually dependent on each other. Assuming a temporal horizon, it is the actions on the micro-level at time ***T1*** that are conditioned by the structure of the situation already given from ***T0***.[7] Again, the structures of the social situation itself emerge from the (emergent) aggregation of sequences of individual actions at ***T1***, and become effective in turn in a future social situation, i.e., they structure a social situation at a further point in time ***T2*** (etc.). This process of mutual structuring can now in principle be extended to further loops as time progresses. ***Figure 2*** illustrates this process of mutual conditioning.

Figure 2: Mutual Micro- and Macro-Level Dependence in the MSE

---
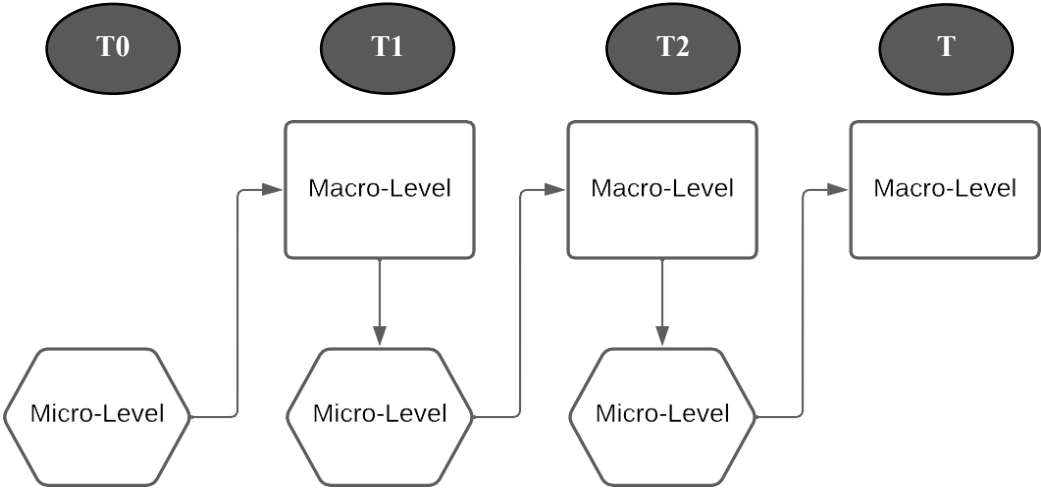
[7] Of course, ***T0*** did not emerge from nothing. In the context of what we have already explained, we are taking cultural evolution and cumulative cultural evolution as the principal starting point for the structuring of the theoretical unit ***T0***.

The interplay and mutual conditionality of social structure at the macro-level and action at the micro-level gives rise to a causal relationship that can explain both enduring social structures and changes in social systems. Let us take a look at two ideal-type illustrations. On the one hand, we find an example of enduring social structures explained by the MSE when the actors in the situation act in accordance with the structural (pre-)condition of the situation. This case is an example of structural reproduction. On the other hand, an example of changing social systems can be found. Such a scenario would be the case when the actors' actions in the situation take place in a deviation (e.g., through a creative or non-routine interpretation of the situational structure) from the structural (pre-)condition of the situation. This deviation flows into the future structure of the situation and holds a potential to bring about social (structural) change. As simply outlined here, the MSE is therefore able to explain socio-structural endurance and change. We will elaborate on this further below when we have brought more substance to the skeleton of the MSE. Furthermore, it becomes clear that only the actors located at the micro-level create social structures. It is the selection of action among opportunities and the subsequent actions of the actors that create social structures and the orientation-giving property of social structures in a future social situation. However, the actions of the actors are in turn conditioned by the structure of the situation. This mutual social process is now more closely specified on the basis of *three logics*: the logic of the situation, the logic of selection (including the selection of a cultural frame) and the logic of aggregation.

The MSE first requires a description of the social situation in which the actors are embedded. This description serves to derive the *logic of the situation*. A social situation is (pre-)structured by its socio-historical constitution. This structural constitution means that the actors in the situation routinely act in a certain way and not in another. The situation becomes effective for the actors' actions when it evokes an idea, a mental representation, which is formed by the actors' interpretation of the situation. This creates a *definition of the situation*. Furthermore, the situation is interpreted by the actors in a certain way due to the presence of significant symbols, which gives the situation a particular (cultural) frame. As already mentioned, the frame, and the codes and scripts implied by the frame, provide an action-guiding structure within social situations. This in turn becomes effective for the actor and leads to a logic of selection (of action(s)). The macro-level thus influences the actor's actions at the micro-level via the definition of the situation.

The *logic of selection* of actions touches on two aspects. First, the selection of the cultural frame of reference itself. Second, the subsequent selection of the action under the precondition of the respective frame. At this point an action-theoretical addition to the MSE

comes into play, i.e., the Theory of Frame Selection.[8] ***Frame selection*** involves the selection of the frame of reference on the one hand and the activation of a certain mode of information processing on the other. To begin with, let's return explicitly to frame selection: here we first assume the presence of ***significant symbols*** in the situation. These symbols are characteristic elements of the respective situation and distinguish it from others. Furthermore, we assume cultural socialization. Actors acquired specific associations (mental models) with (typical cultural) symbols through being socialized into a cultural body of knowledge and meaning; we consider actors as cultural subjects. Furthermore, through engagement in re-occurring social

---

[8] Tutić (2023) highlights core ideas as well as deficits of a line of evolutionary psychology and also of the dual-process perspective in an overarching action-theoretical focus, and also shows their relationship to counterparts in sociological theorizing. In fact, Tutić (2023) points out to the core elements of particular sociological and psychological theories and shows that "goal-framing theory should be interpreted as a sociological version of evolutionary psychology (…) [whereas the] model of frame-selection can be considered as a sociological descendant of the dual-process perspective" (p. 22). Essentially, he presents a school of evolutionary psychology that deals with domain-specific evolved psychological mechanisms (information processors) and advocates the massive modularity of our evolved minds. In addition, light is shed on the dual-process perspective. This perspective essentially comprises two types of cognitive processes, Type 1 and Type 2 (Kahneman, 2011). Furthermore, Tutić (2023) argues that dual-process and evolutionary psychology are in large parts compatible. On the one hand, Type 1 processes are said to be evolutionary old and shared with other animals. These processes involve automatic, unconscious, uncontrollable and low energy-consuming cognitive processes. The second type (Type 2), on the other hand, is comparably evolutionary more recent, distinctively human and involves time- and energy-consuming, slow, controlled and conscious cognitive processes. Type 1 processes can be understood to be similar to domain specific intuitive gut reactions (Haidt & Joseph, 2007), while "[th]e application of normative calculi such as logic, stochastics, and decision theory generally occur via Type 2 processes" (Tutić, 2023, p. 12). In the course of theoretical refinements, Type 1 is now called "TASS", i.e., the autonomous set of systems, since the individual parts (domain-specific modules) can be formed to a set but are not per se integrable to a single system. In contrast, Type 2 is not regarded as a collection of individual processors, but as a domain general system that integrates respective information processors under its umbrella. Beyond that and furthermore, it is argued that Type 1 processors take effect in routine tasks whereas Type 2 becomes active when confronted with novelty that cannot be solved by routinized ways of problem solving. Both of these statements must be seen under a perspective that tries to explain these cognitive types in terms of their evolutionary emergence — Type 1 processes may have emerged under reoccurring (domain specific) adaptive challenges whereas Type 2 processes (as a cognitive capacity) may have evolved as adaptive response to environmental fluctuations (be they natural or sociocultural) that induced novelty and the need for our mind to search for solutions beyond routine paths. The implication is given that Type 1 processes are may activated by familiar cues whereas processes of Type 2 are may activated under unfamiliar cues and under sufficient motivation and cognitive resources to process the stimuli at hand. Notably, it is emphasized that Type 2 can override Type 1 processes (Tutić, 2023). This is, at least in large parts, in line with an intuitionist model on moral judgment that we will deal later with (Haidt, 2001). For the time being, however, we will content ourselves with having pointed out, with reference to Tutić (2023), that the MSE, and here the action-theoretical component of frame selection, is largely compatible with and can be linked to theorizing from the dual-process perspective and evolutionary psychology. In a special case of MSE, which we assume for our modeling and which grounds on a perfect match between symbols in the social situation and mental models of the actors, our focus is on routine alike automatic Type 1 processes as will be explained. Albeit that we draw on psychology and the social sciences more broadly in this project, working with evolutionary, social and cultural psychology as well as sociological concepts and economic and anthropological influences, we are nevertheless confident that most of the theoretical positions we work with are not competing but rather complementary, as they broadly represent similar to the same core ideas and share major theoretical positions. For this reason, we believe that we can use the theoretical and empirical concepts we cite in our writing in exactly this respect, i.e., as complementary in most parts and only conflicting in the detours that are the focus of interest for the specialists of each approach.

situations actors are assumed to have these associations mentally available. Under these conditions, it can furthermore be expected that the presence of specific symbols in the situation causes the activation of a pattern of interpretation, i.e., a definition of the situation. This situation definition in turn indicates appropriate action in the situation. Situations can contain a better or worse matching of the existing symbols and mental models. Roughly speaking, this means that situations can be familiar or unfamiliar to the actors. This is where the information processing mode comes into play. According to the MSE, if there is a **perfect match**[9] between the elicited mental model and the significant signs (symbols) of the situation, this is followed by unreflecting, automatic frame selection and routine actions that are culturally habitualized. The latter also occur as an automatic (non-calculating) selection from the opportunity structure of the situation. This information processing mode is called the automatic-spontaneous mode (*as-mode*). If cultural framing takes effect, actions are conditioned according to code (the superordinate goal of the situation) and script (a model of appropriate action associated with the code; the typical action in the situation). Moreover, the principle of utility maximization becomes prevalent here, as subjectively proven actions show themselves against the background of cognitive thriftiness (low effort) in the unreflecting-automatic selection of actions in the situation. If, on the other hand, the situation is unknown, i.e., there is a **mismatch** between the symbol(s) present in the situation and the mental model(s), and if the costs of misjudgments are also high, a deliberate, calculating mode of action selection (i.e., the reflective-calculating mode; *rc-mode*) comes into play according to the MSE. Subject to assumptions not detailed here, this mode of information processing means that the decision to act is made more consciously and in a calculated manner on the basis of (a modified) Subjective-Expected Utility Theory. In simple terms, the choice of action in the rational mode is summarized as follows: given that the action alternatives are assessed, taking into account evaluations and expectations, it is the action alternative that promises the greatest utility that will be selected.

Before we now close the circle between the micro- and macro-levels through the logic of aggregation, we would like to make a brief digression and *integrate the self-construal approach into the Model of Sociological Explanation*. To localize the self in the MSE we assume a case in which there is a perfect match between symbol and mental model in a social situation. In this case, the definition of the situation and the choice of action are guided by cultural schemata that have been learned through socialization and that are available to actors

---

[9] "Perfect" match and "perfect" mismatch between symbol(s) and mental model(s) are conceivable as theoretical ideal states.

through everyday exposure to re-occurring social situations (see: Esser, 2010, pp. 320-324). The nature of our selfhood is a cultural schema itself. If we now continue to assume the presence of self-relevant information in the social situation, it can be concluded, as already mentioned above in relation to the self's functions, that the subsequent process and the consequences are conditioned by the configuration of our self-system, i.e., our culturally constituted **self-construal** (Markus & Kitayama, 1991; 1998; 2010). Our self as a mental structure guides information processing, brings the respective cultural focal point — that is social **independence** or **interdependence** — to the forefront of individual cognition, emotion, motivation and action tendency, and may thus be understood **as a (specific) cultural frame**. Under salience of the frame of independence or interdependence of the self, corresponding **codes** and **scripts** become prevalent. The codes are expressed in the focus on relationships (in-group) or autonomy as the superordinate goal in the situation. The scrips indicate code congruent socially appropriate routines of action in the situation. If we therefore assume the logic of a specific situation — which we set here e.g. in simplified, heuristic terms as collectivism and individualism —, the presence of significant self-relevant symbols in the situation, and a perfect match between mental model and symbols, then we can conclude that the **definition of the situation** (the interpretative frame of reference) is **guided by** an individual´s respective **self-construal**. It follows that the selection of actions is also based on the respective way of self-construal; the unreflecting, automatic-spontaneous mode of information processing determines the execution of habitualized, self-congruent actions. Thus, we **localize the self in the MSE** as specific cultural frame given particular symbols of self-relevant information are present in the social situation and given a perfect match between symbols and actors mental model(s). Self-construal as cultural schema can therefore be easily integrated into the theoretical framework of the MSE and the model of frame selection (if not already included in the latter as a special case, assuming a perfect match).

Now finally, we come to the **logic of aggregation**. This deals with the consequences of individual actions for the macro-level and its social structures. Once again, it should be emphasized that the consequences of individual actions for the macro-level are largely unconscious and go beyond the intentions of the actors. We will now return to the examples from above and expand on them. Assuming the two previous logics, we can highlight two macro effects by aggregating individual actions and the consequences of these actions by way of example and simplification. In **case 1**, the actors act in a social situation on the basis of a cultural frame and no deviations or disturbances occur. In this case we assume consequently a perfect match between significant symbols and mental models of actors. Here, the actors at the micro-

level, at a point in time T1, reproduce the already given structure of the situation at the macro-level through their actions for a future point in time T2. The structure of the situation retains its validity and becomes thus also effective in T2 as an orientation for action. Accordingly, if a respective self-construal forms the frame of reference, no change in selfhood would be necessary in case 1. With the re-production of social structures, the self-construal is also re-produced. In contrast to case 1 we outline an example of structural change in *case 2*. In this case we assume a social situation at time T1 in which there is a (perfect) mismatch. This mismatch of symbols and mental model requires alternatives for the otherwise selected cultural frame. Furthermore, we assume that the mismatch is always present when the same social situation exists and actors are always prompted to pursue alternative but not routine actions in the situation. For the sake of simplicity, let us also assume that the same alternative action always promises the greatest utility and is therefore selected by the actors in the situation. The alternative action is perceived by the other actors in the situation when it is implemented and thus modifies (i.e., re-frames) the (initial) social situation. The previously appropriate and valid action selection of the cultural frame can be called into question as a result of the new alternative action entering the situation. New evaluations and expectations of the typical action in the situation now gradually emerge from the consequences of the action (alternative). At time T2, the structure of the situation of T1 has then changed and the typical action of T1 no longer promises the same utility in the situation at time T2. Due to the change in utility, the formerly typical action is therefore, according to the rule of utility maximization, no longer selected. Thus, in case 2 the structure of the social situation in T2 has changed. By taking a look at the self in this scenario, we would witness also change in our individual unit of being as long as self-relevant information are affected by the structural change and code and script have altered too. Markus and Kitayama, (2010) state: "[a]s cultural content changes, the mediating self and psychological functioning change in turn" (p. 423). According to the MSE, social structures (at the macro-level) therefore determine actions at the micro-level, but the latter aggregates the social structures in the first place and thus also determines their persistence or change. In sum, structural persistence and change, and causal relations between micro- and macro-level are explained by the MSE via a recourse on the individual level of actors, a general rule for behavior (utility maximization) and the application of the three logics outlined (Greshoff, 2008; Esser, 2010).

From the outline of the MSE and the integration of self-construal into the MSE, the relation between cultural context and selfhood becomes clearer. For the hypotheses to be presented later, we refer to the MSE and the integration of the self-construal approach into the

model. ***We focus on an MSE special case*** in which significant symbols and self-relevant information are present in the social situation and available to the actors as cultural subjects.[10] Furthermore, we assume a perfect match between the actors' mental models and the significant symbols of the situation. The background to these assumptions is cultural socialization, which in turn is based on the evolutionary evolved human capacity for cultural learning that enables us to adapt to our sociocultural ecology. The MSE mode of social change will not be discussed further at this point, but it is also a component of the theoretical background.

The current constitution of one's own cultural ecology is reflected in the momentary culmination of cultural experiences of the individual self, which finds its expression in the current actions (and cognitions, emotions and motivations) of actors. So, the selection and execution of actions in a given situation are guided by the respective frame of reference encoded in our self-system, that itself is activated by significant symbols as well as self-relevant information in the situation. Based on the logic of situation and the logic of selection, the logic of aggregation comes into play. It can be deduced from the MSE that certain options for action in social situations of everyday interaction are chosen more frequently by cultural subjects (actors) than other options. This is due to the presence of significant symbols in the situation and the subsequent frame selection (i.e., activation of self-construal as frame of reference). We assume that significant symbols are present in everyday cultural situations that condition or

---

[10] Note however, the special case of perfect match, that is in the language of the MSE activating the automatic-spontaneous mode (as-mode), is only one route of possibilities in social situations. We are well aware, that also deviations eliciting Type 2 processes of reflective thinking and action based on it might be possible in the everyday (inter-)action of cultural subjects. The latter would refer to the reflective-calculating mode (rc-mode) in MSE language. This possibility is not called into question by the model that we will propose. Nevertheless, we base our heuristic model on the special case of perfect match between symbols in the social situation and actors' mental models due to two assumptions. These are: A) the actors are cultural subjects, i.e., socialization into and internalization of cultural information turn mental models into models of fit to the recurring, typically present symbols of the everyday, socio-culturally structured situation of interaction between cultural subjects. These models should be chronically available to actors due to the constant re-engagement of actors in social situations with similar overall social orientation and structure. B) We assume social (inter-)action between cultural subjects that represent a typical, everyday interaction situation. In doing so, we leave the meaning of the term "typical" deliberately open and closed at the same time. This means that the socio-culturally structured everyday interaction situation is open in order to include the rich plethora of cultural differences. Whatever is understood by typical is an inherent part of the specific context at hand. We regard the everyday interaction situation as closed, however, in the sense that the situation of (inter-)action that we assume for the actors is nevertheless a distinct constant that recurs for the cultural subjects of a given socio-cultural entity. So, by *typical* we mean two things: the situation we are talking about encompasses the many possibilities of socio-cultural structuring of the many cultural entities that exist, but at the same time the specific situation of actors is clearly characterized as culturally familiar for the cultural subject in the situation. By the way, the word subject comes from the Latin "sub" and "jacere" (https://www.merriam-webster.com/dictionary/subject) and can be understood semantically as an individual entity that is subjugated to something; the word means literally to "throw under". When we speak of a cultural subject, we are therefore referring to the semantical essence of the word and mean an individual human being that is subjugated to cultural influences. In this sense, we regard the actors as cultural subjects.

activate a particular cultural frame of selfhood. Our self is formed, i.e., socialized, depending on its sociocultural environment and shaped by it. It is therefore aligned with cultural imperatives that are expressed in preferences for *independence* or *interdependence*, as we have shown above. It further follows that the actions of cultural subjects in typical cultural situations manifest themselves in social patterns under the precondition of an activated self-construal frame. Due to recurrences of everyday social situations, it can be assumed that social patterns of action (in the given social situation) are repeated over time, since one and the same action is socially expected, has proven to be socially valid and also promises the greatest utility. Guided by the active cultural frame of self-construal, the chosen patterns of action themselves become aggregated social structures over time. These structures in turn define or (re-)produce the opportunities and significant symbols within social situations of the same type. This means that they (re-)construct the respective logic of the situation. In the special case of MSE that we assume, it follows that "cultures and selves define and build upon each other in an ongoing cycle of mutual constitution" (Markus & Kitayama, 2010, p. 420). Thus, by incorporating cultural modes of selfhood — i.e., the independent or the interdependent self-construal — into the Model of Sociological Explanation (MSE), and by accepting the assumptions we have made, it can be deduced that *the self reflects the past, present and future constitution of the sociocultural ecology in which it is embedded*. What we have outlined is now summarized in an initial model (*Figure 3*; *Model 1*), which we will later expand to include morality.

Figure 3: Model 1 — Integrating Self-Construal into the Model of Sociological Explanation
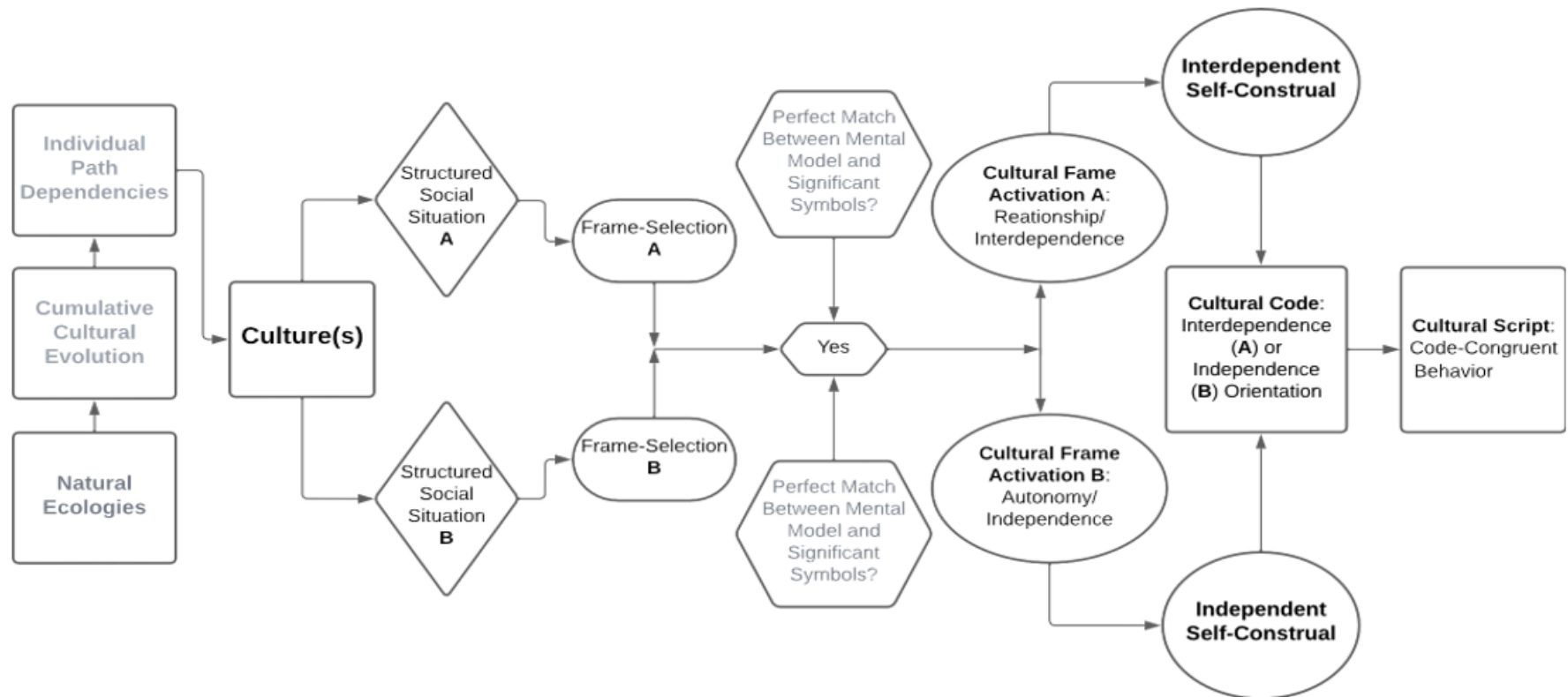
Model 1 shows the case of an ideal type frame selection for cultural entities in which different modes of selfhood become prevalent as frames of reference that provide orientation in social situations. *Figure 3* illustrates our integration of the self-construal approach for the Model of Sociological Explanation. Two structured social situations are shown which, given the presence of self-relevant information and a perfect match between (situationally present) significant symbols and mental models of the actors, lead to the activation of two different, self-construal congruent cultural frames. **Frame A** follows the logic of the situation that we have described in the text exemplarily and in simplified terms as **collectivism**. This is where the frame of an **interdependent self-construal** is selected. The code of the situation specifies relationship orientation as the superordinate goal of the situation; the script indicates corresponding, socially approved actions. **Frame B** follows the logic of the situation, which we have exemplarily and in simplified terms described in the text as **individualism**. This is where the frame of an **independent self-construal** is selected. The code of the situation indicates autonomy orientation as the superordinate goal of the situation; the script indicates corresponding, socially approved actions. Assuming a temporal horizon, Model 1 could be extended to the case of the reproduction of social structures described in the text. For the sake of clarity, we have, however, refrained from such an extension here. The background of the model is formed by different natural ecologies, cumulative cultural evolution and individual path dependencies that condition the emergence of different cultures.

1.5. Culture mediated Correspondence Between Self-Construal and Morality

At the beginning of the section on the self, we asked you to complete the question "I am ..." five times, and then proceeded to describe a tendency of cognitive differences in the availability of how people see their self in relation to others. Throughout this section, it should have become evident that differences in the way how people define their "I am…" result from internalized and incorporated sociocultural ecologies that are the basis for the construal of selfhood (Markus & Kitayama, 1991; Henrich et al., 2010a; Berger & Luckmann, 2013; Gavrilets & Richerson, 2017; Kitayama & Salvador, 2024). ***The human self*** is the individual model that mirrors ecological embeddedness; the self and its respective cultural context are intertwined and functionally related for they emerge as mutual source of existence, and endurance or change.

We hypothesize that there is a systematic correspondence between selfhood and our configuration of morality, and will briefly sketch the outlines of our argument here. Before we look in detail at our moral mind from the next section onwards, let us roughly anticipate what we mean by morality: ***morality*** is a universal feature of our species, and can be seen as a collection of ideas and social standards that define what is considered socially acceptable or unacceptable, right or wrong. From an evolutionary perspective, the human moral mind is inherently functional as it has evolved to constrain selfish drives, uphold social order and solve interaction problems of human cooperation (Haidt & Joseph, 2007; Tomasello & Vaish, 2013; Kurzban et al., 2015; Curry, 2016; Ellemers et al., 2019; Henrich & Muthukrishna, 2021). As such, our morality promotes and regulates basic elements of everyday human social life.

Baumeister (2022) builds a bridge between the self and morality by viewing morality as a long-term survival strategy in groups and by emphasizing the self as the bearer of (moral) reputation. In turn, it is the respective (moral) reputation that signals to people whether an individual is a potential cooperation partner or not. Baumeister thus comes to the conclusion that "[m]orality is a centrally important aspect of the self" (Baumeister, 2022, p. 104). We follow the view that our ***morality is an aspect of the self***, and will look at the self and morality from a (cross-)cultural perspective.

In doing so, we assume that the human self and morality serve specific purposes with regard to respective socio-historically evolved social systems (i.e., large groups such as societies). Both self-construal and morality are capable of fulfilling a dual purpose related to the re-production and survival of individual organisms and the re-production of social systems,

which ultimately also promote the survival of their members. In this respect, we expect the human self and morality to be adapted to our second inheritance system, i.e., to the prevailing culture in which the individual is embedded (Henrich & McElreath, 2007). The way we define our self in relation to others, i.e., our *self-construal* serves to align the individual with the prevailing social structures. It also serves to maintain (i.e., re-produce) social structures. As we have explained in detail, the respective characteristics of the human self work to reproduce the individual organism in context. The way the self is construed expresses individual's socio-cultural adaptation on the one hand and plays a potent part in (re-)producing the (respective) cultural environment of our species on the other. Human *morality*, in turn, serves to regulate human cooperation and the maintenance of social order. The regulation of human cooperation is relevant with regard to the distribution of resources. From the perspective of the MSE resources (material and immaterial) are an essential part of satisfying basic human needs and providing the necessities of survival (Greshoff, 2008; Esser, 2010). In the next sections we will highlight several domains associated with material and immaterial resources that are governed by our moral mind. In addition, our evolved morality regulates the social interaction of actors by evoking particular gut feelings in response to the perception of specific actions in the social environment. Eventually, our moral mind enables us to recognize and classify perceived actions as right and wrong behavior. Thus, based on a sense for right and wrong, morality serves to constitute social order, which in turn promotes the reproduction of the elements of social systems (i.e., individuals and groups) and so the upkeep of social systems themselves.

The need for human cooperation and social order appears to be universal to our species (Tomasello & Vaish, 2013; Curry et al., 2019a; Henrich, 2020). In this line Skitka and Conway (2019) note: "[s]ome aspects of morality are seemingly universal and necessary for social functioning" (p. 2). We will look at these and other aspects in more detail further below. Nonetheless, for the moment we would like to point out one of our central assumptions: *we assume that cultures differ in their respective requirements for cooperation and social order, and thus in the relevance they attribute to different moral domains*. This view is consistent with two prominent moral theories (Haidt & Jospeh, 2007; Curry, 2016) and we also base it on what we have explained in the section on culture. Consequently, we assume that cultural differences in human morality are due to the different histories of cumulative cultural evolution and the associated path dependencies of cultural entities.

We premise that morality is part of the human self. Since the way our self is construed mirrors the individual adaptation to socio-cultural environments, we further expect that our moral systems are configured in a form that corresponds to the prevailing cultural self-construal.

As an enduring self-system of knowledge our self harbors the sum of beliefs that we possess about ourself, and works as a collection of self-relevant schemata that regulate intra- and inter-personnel processes in context (Markus & Kitayama, 1991; Morf & Koole, 2014). In this respect, the "properties of the typical self are based on what works best in a person's social system" (Baumeister, 2022, p. 41). As humans, we are a species that is pro-socially cooperative in orientation, but also egoistic in our drives and actions. Therefore, we have both selfish and other-regarding tendencies and preferences (Cooper & Kagel, 2015; Tomasello, 2017; Crocker et al., 2017; Brown et al., 2022). Morality is concerned with (self-)regulation, as it serves to solve problems of human cooperation and the challenge to uphold social order under the prerequisite that we are antagonistic yet also pro-social cooperators. Thus, we find the interface of self and morality in (self-)regulation (Bandura et al., 1996; Baumeister, 2010). Essentially, we view the human self-system and moral mind to be only analytically separated; ***morality is an aspect of the human self***. The self is inextricably linked to the sociocultural context and is a social agent in the world. In this view, the interface between morality and the self becomes particularly evident when we focus on the regulation of the executive functions of the self.

If the theoretical association just stated should hold true, we further expect that also the particular cultural configuration of our moral system is aligned to what works best in a person's social system. We have highlighted two focal points of the culturally constituted selfhood: cultures seem to differ in their focus on the individual or the (in-)group. In accordance with these cultural foci, our self is formed to be independent or interdependent. So, given that morality is an aspect of the self and that the self is formed in accordance with contextual requirements, we predict cross-cultural differences in moral systems along the line of a superordinate individual (independence) or group (interdependence) orientation of the self.

Furthermore, we argue that independence and interdependence in self-construal, as well as corresponding moral systems, are expressions of what works best in given natural and socio-historically grown cultural context (Henrich, 2020). We hypothesize that certain cumulatively evolved sociocultural structures require(d) a predominant social focus on (in-)group life, while others rather require(d) a social focus on individual life and more interaction with strangers. Accordingly, on the one hand, we assume cultural contexts that are constituted by elements that render a focus on group life meaningful for survival and thus functional. Such elements include a general orientation towards social interdependence and a predominance of rather particularistic, group-based cooperation. Assuming such a cultural context, the typically prevailing situational logic of the actors' everyday interaction should consist of the above-mentioned contextual elements. This is to say that the respective situational logic should

comprise significant symbols indicating interdependence and overall group-orientation. On the other hand, we assume also cultural contexts that are constituted by elements which render a focus on individual life, an overall social orientation towards independence and a predominance of individual-based cooperation meaningful for survival and thus functional. Again, assuming such a cultural context, the typically prevailing logic of the situation of the actors' everyday interaction should consist of the aforementioned contextual elements and respective symbols. In other words, we assume different cultural contexts consisting of either elements that emphasize group-oriented or individual-oriented life, and hypothesize that these elements render the particular configuration of self-construal and moral system to be adapted to the contextual requirements.

So let us summarize: morality is universal to humanity. Nonetheless, we hypothesize that our self-construal and thus also our moral systems differ from culture to culture due to different contextual requirements that have emerged in the course of (cumulative) cultural evolution and subsequent path dependencies. We reason that the configuration of self and morality is linked via the respective cultural context: our self as well as our morals work best for the cultural subject as well as for (the re-production of) socio-cultural structures when they are adapted to the requirements of the respective cultural context. Ergo, the theoretical conclusion of correspondence between the configuration of the self and the configuration of moral systems is theoretically consistent. Thus, we hypothesize that the configuration of self and morality correspond to each other, as both phenomena reflect the requirements of the respective cultural entity.[11] To develop this argument further, we turn in detail to human morality in the next section.

---

[11] For the sake of clarity, we consider our world to be multicausal and take into account that not all requirements of a cultural entity need to be functionally related to selfhood and morality. However, as we will argue below, we see a glaring similarity between the overarching modes of selfhood (i.e., interdependence and independence in self-construal) and higher order moral constructs of binding and individualizing morality. We expect this theoretical similarity to hold also empirically true, i.e., we hypothesize culture contingent correspondence between ways of selfhood and moral relevance. Further details on the logic of our argumentation follow in the next section.

1.6. Morality

Humanity faces tragedies of dehumanization, killings and genocide in its history and still today (Hewstone & Cairns 2001; Houghton, 2009; Cottam et al, 2010; Lindblom et al., 2015). The atrocities that we humans can inflict on each other are vast in their extent and scope. Within the framework of man-made suffering morality takes in a prominent place. Several scholars propose that morality plays a crucial role in war and affects side taking in conflicts (Tooby & Cosmides, 2010; Kurzban & DeScioli, 2016). Moral transgressions and parochialism are also potent to affect and amplify violent interactions between human groups (Bandura et al., 1996; Choi & Bowles, 2007; Moore et al., 2012, Rusch, 2014; Moore, 2015). With regard to moral disengagement, it is further known that certain situational factors can render our moral based self-regulation ineffective. Research suggests that bilateral and multilateral market situations, for instance, can contribute to a diffusion of responsibility and consequently undermine our ethical standards (Falk & Szech, 2013; 2015). Despite terrible man-made horrors, however, it must be noted that human interaction is usually peaceful and often even ***cooperative***. The positive power of morality does not lie in the possibility of transgressing it, but rather in the (self-)regulatory function of morality, which enables human interaction that leads to mutual benefits (Curry, 2016; Carlo et al., 2016; Hare, 2017; Wrangham, 2018; Shilton et al., 2020). Recent research suggests that we are currently living in even the most peaceful time in human history, and points to the role of morality in this respect (Waytz et al., 2019; Kirkland et al., 2023).

Although morality plays a vital role in our everyday lives (Hofmann et al., 2014), findings support that people often do not have a concrete, consciously available definition of what morality means to them. Nonetheless, humans seem to recognize what is moral when they encounter it in certain phenomena in the social world (Skitka & Conway, 2019). Apart from different approaches, it is generally difficult to find a widely recognized term for morality. Even in science there is still no comprehensive and definitive definition of what morality is, and there is some confusion in the use of related terms (Pfattheicher et al., 2022). As a result, the field of moral psychology "comes with competing perspectives rather than with theoretical or empirical consensus" (Skitka & Conway, 2019, p. 35). This leaves room for contributions.

With the present research endeavor, we would like to make a small theoretical, a practical, and also an empirical contribution to the understanding of morality. Before proposing our approach to morality below, however, we would first like to present a fairly broad working definition to indicate the direction that we will take here. Carlo et al., (2016) state, that morality

centers around "common and differing ideas (…) surrounding what is deemed acceptable or unacceptable (…) [and functions to] support and organize interactions among individuals and groups" (p. 54). Human morality encompasses a wide range of phenomena. To name just a few: our capacity for empathy, other-regarding preferences, and certain forms of condemnatory or affirmative judgments, but also helping behavior and particular emotional impulses are elements of our morality (Haidt, 2003; De Waal, 2008; Cooper & Kagel, 2015; Tomasello, 2017; Crocker et al., 2017). Certainly, the human moral mind is universal in our species, yet also variation in morality can be found across cultures (Miller et al., 1990; Kurzban et al., 2015; Enke, 2019; Henrich, 2020). Overall, morality is at the heart of human cooperation, social order, societal coexistence and civilization (Graham et al., 2016; Curry et al., 2019a; Ellemers, 2019; Henrich & Muthukrishna, 2021; Muthukrishna, 2021; Baumeister, 2022).

In the following, we will further introduce to the topic of morality by first looking briefly at ways that shed light on the development of our moral minds. We then look at how morality is effective in regulating ourselves and our fellow human beings. Thereafter the focus will be set on two modern theories of moral pluralism by discussing the Moral Foundations Theory (MFT) and the Morality as Cooperation Theory (MaC). Building on this, we will present a perspective that aims to integrate MFT and MaC. In this section, we will outline our theoretical understanding of morality. In addition, we will sketch three new research tools emerging from our perspective that seek to enable researchers to study various aspects of human morality across cultures. Finally, we will not just present illustrative examples of cross-cultural differences in morality, but also derive specific ideal-type logics of the situation that allow us to draw hypotheses to be tested empirically.

1.6.1. Evolved Human Morality

The development of moral aspects reaches far into the history of natural evolution (Brosnan & De Waal, 2003; De Waal, 2008). However, the formation of the human moral mind goes back not only to natural processes; culture also has a significant influence on the shaping of our morality. Human morality encompasses "instincts, intuitions, inventions, and institutions" (Curry et al., 2019a, p. 48) and is rooted in processes of gene-culture co-evolution (Haidt, 2008; Henrich & Muthukrishna, 2021; Brown et al., 2022). The evolution of morality in humankind is of inherent adaptive utility. In this line several scholars argue that natural selection favored the evolution of morality in humans for it solved recurrent adaptive challenges that our ancestors have faced (Haidt & Joseph, 2007; Kurzban et al., 2015; Curry, 2016). By its nature,

morality is thus an evolved universal facet of the human mind. As Brown and colleagues (2022) have put it: "hominins, in particular, have been subject to selection promoting other-regarding preferences" (p. 476). Within the human development of progressively increasing pro-sociality the aspects of decreased and altered (re-active to pro-active) aggression, that came along with a concomitant domestication syndrome, are suggested to play a role (Hare, 2017; Wrangham, 2018; 2021; Shilton et al., 2020). The *Human Self-Domestication Hypothesis* proposes "that selection for reduced aggressiveness (…) led to physiological, psychological, and behavioral changes, specifically to social tolerance" (Sánchez-Villagra & van Schaik, 2019, p. 136), and overall, to heightened abilities of cooperation. Altruism, especially kin-ship altruism, can be considered the cornerstone in the evolution of morality (Kurzban et al., 2015; Carlo et al., 2016; Muthukrishna, 2021). Nonetheless, the evolution of cultural evolution further promoted pro-social tendencies in humans and expanded the scope of morality beyond kin (Diekmann & Lindenberg, 2015; Purzycki et al., 2018; Lang et al., 2019). In terms of cultural evolutionary mechanisms, certain social means have been shown to occupy a crucial place in human moral development: punishment, signaling, and reputation are capable of supporting any adaptive or maladaptive form of individually costly behavior (Axelrod, 1986; Chudek & Henrich, 2011; Henrich & Muthukrishna, 2021). Conclusively, the moral mind has evolved on the basis of a set of socio-cognitive, emotional, and self-regulatory capacities over the course of distal and proximal processes inherent in the natural and cultural evolution of humans (Henrich, 2016; 2020; Shilton et al., 2020; Heyes & Moore, 2021).

Since we will argue for the centrality of cooperation in human morality later on, we would like to highlight one approach to the development of our moral mind at this point. Tomasello and Vaish (2013) hold that, from an evolutionary perspective, morality is about cooperation. They see the origins of our evolved moral mind in two evolutionary steps that together distinguish the psychology of our moral mind from that of our animal relatives the great apes. Tomasello and Vaish (2013) state:

> "[o]ur assumption is that the two key steps in the evolution of human cooperation, and thus morality, took place before the advent of agriculture and cities, and law and organized religion, as humans first became obligate collaborative foragers and second created cultural groups that competed with one another" (p. 240).

The argument for their notion is as follows: over the course of human history, our environment changed so much that we turned into cooperative foragers. Cooperation became essential, because without it the individual would no longer survive in the changed

environment. This idea is known as the *Interdependence Hypothesis*. In this first evolutionary step, humans developed abilities for joint intentionality (Tomasello, 2017), and mutual cooperation became essential for survival. Humans come to be increasingly dependent on each other, which fostered prosocial drives, as helping a collaborative comrade not only supports the survival of the other, but also the survival of the self (Henrich, 2020). Humans began to care deeply about each other (Tangney et al., 2007; De Waal, 2008). Consequently, "[i]ndividuals became interdependent with one another, such that each individual had a direct interest in the well-being of others as partners" (Tomasello & Vaish, 2013, p. 239). The human groups, tribes and bands, comprised more than two individuals, which implied the possibility of choosing partners for joint ventures. This possibility in turn led to a social regulation of individual behavior. Cooperation is necessary for survival. Individuals evaluate cooperative behavior and choose partners for cooperation based on their evaluations. In this way, thus, all individuals are obliged to cooperate so that they are also selected as cooperation partners in the future. As a result of this dynamic, a joint morality has evolved that induces people to help those on whom mutual cooperation depends (Tomasello & Vaish, 2013). Above that, signals of one´s cooperativeness and cues of social reputation that identifies an individual as a good or bad cooperation partner became increasingly important (Henrich & Muthukrishna, 2021; Baumeister, 2022). The second evolutionary step describes the development of human morality that extends to group life as a whole. Tomasello and Vaish (2013) see group life structuring norms and competition between human groups as the basis for this development. In response to threats from other groups, group life as a whole develops into a sphere of cooperation, for which the same mechanisms as in the previous evolutionary step now apply as collective mechanisms. Competition between groups has thus set a process in motion through which the joint morality of interdependent individuals has developed into a psychology of our morality that is calibrated to track and recognize how the individual serves or harms the common good of the group (Tomasello & Vaish, 2013). In larger groups consisting of several subgroups, the coordination of social interaction is primarily determined by group membership. Group membership, in turn, can be indicated by norm-conforming behavior (Hogg & Reid, 2006; Henrich & McElreath, 2007; Henrich & Chudek, 2011). As a result, norm conformity became central, as did the social enforcement of norms through the punishment of violators and the intra-individual enforcement of norms through self-punishing moral sentiments (Haidt, 2003; Chudek et al., 2016). In this context Henrich (2020) describes that the overall selection pressure to adhere to the collective mode of interdependence was powerful: "When (…) sanctions failed to bring violators into line, hunter-gatherers escalated to ostracism, beatings, and even

executions" (p. 80). Taken together, these developments led to the decisive factor of the emergence of impersonal, actor-neutral norm-based psychology which extended the social scope of our moral mind from a joint to a collective morality that is governing human cooperation (Tomasello & Vaish, 2013). From this brief introduction to the origins of our evolved moral mind, we will now move on to take a closer look at how our morality supports self-regulation and pro-social behavior.

1.6.2. What Makes us to Regulate Ourselves and Motivates Moral Behavior — From Reason, Intuition, Emotion and Social Means

The evolved moral mind of humans is part of the answer to the question of what makes us most often to self-regulate and interact on the basis of cooperation rather than exploiting and harming each other. Our morality is promoted and effective within ourselves as well as in our social environment. It comprises moral reasoning, moral intuitions and emotions together with social means of regulating behavior.

One of the most famous derivations of what is moral comes from Immanuel Kant, as already emphasized in the introduction of this writing (see: Kant, 1788/2011, p. 738). According to this philosophical-rationalist view, moral action is action that can be elevated to a general law for all people at any time. This deductive derivation of morality implies that the potential to be selfish is regulated by conscious deliberation, since moral action must fundamentally take into account not just one´s own but also the position, needs and reasons of all others. Lawrence Kohlberg's *Theory of Cognitive Moral Development* can be seen as the cornerstone and psychological turn of the rationalist approach to morality (Skitka & Conway, 2019). The theory is strongly influenced by the cognitive revolution, the dominant paradigm at the time Kohlberg's approach was developed (Haidt, 2001). In consequence, the focus of this approach to morality is on reasoning to arrive at moral judgments that set in motion subsequent morally guided action. At the empirical center of Kohlberg's approach are the analysis of moral dilemmas and the reasons for how people resolve the contradictory claims of given scenarios. Another focus of this rationalist moral theory is on stages of cognitive development, which in turn are thought to imply an increasing sophistication of moral reasoning. Kohlberg and Hersh (1977) argue that moral development "represents the transformations that occur in a person's form or structure of though" (p. 54). In line with this idea, the theory predicts a sequence of six panhuman stages of

moral development.[12] Furthermore, these stages in turn form the structures of moral thinking as well as the basis for moral judgments. In the first two *pre-conventional* stages, the individual (i.e., the child) is oriented towards punishment and obedience (stage 1) in moral reasoning and judgment. Later the child is oriented towards an instrumental-relativistic orientation of need satisfaction (stage 2). The so-called *conventional* level follows in development and comes with two further stages: moral thinking at stage 3 is oriented towards the expectations of relevant others and interpersonal concord. The subsequent stage 4 is characterized by moral reasoning that is oriented towards rules and authorities as well as law and order in general. In addition, the *post-conventional* stage is according to Kohlberg distinguished by a social contract orientation that places utilitarian thinking and a legal perspective at the center of moral reasoning and judgment (stage 5). Finally, the sixth stage expects that people are guided by universal ethical principles that correspond to Kant's categorical imperative. At this stage abstract ethical reasoning defines what is right, and justice, reciprocity, equality and the dignity of the person are at the core of ethical principles (Kohlberg, 1973). In the Theory of Cognitive Moral Development, therefore, it is moral reasoning and the different orientations at each stage that bridles egoistic drives and regulates ourselves.

Kohlberg's approach is generally normative oriented, which is one of the criticisms of the theory. Furthermore, his theory emphasizes consciousness as the fundamental component that guides our moral judgments and the behavior based on it. However, in recent decades, newer trends in psychology have emerged that rather emphasize the unconscious, intuitive part of our moral mind (Haidt & Joseph, 2007; Ellemers et al., 2019). The intuitionist approach to morality highlights the importance of intuition and moral emotions. Essentially, it argues that moral judgments are in most cases based on (intuitive) automatic gut feelings rather than conscious reasoning and deliberation. In contrast to Kohlberg's approach, the intuitionist view puts forward that moral reasoning occupies the opposite place in the hierarchy of importance of moral judgment and action. Moral intuitions, which include moral emotions, come first, and reasoning only plays a potential role later in the process of forming a judgment. However, this is not to say that the intuitive approach denies reasoning. Rather, this view suggests that moral reasoning is a social product that only emerges after the initial gut feelings and serves primarily to justify one´s moral intuitions to others as well as to oneself.

The *Social Intuitionist Model of Moral Judgment* was developed by Jonathan Haidt (2001) and refers to the distinction between explicit and implicit cognition (Kahneman, 2011;

---

[12] One could criticize Kohlberg's approach precisely because of this notion, since the moral stages and the proposed order of the stages might reflect a WEIRD-perspective (Henrich, 2020) rather than a pan-human phenomenon.

Greenwald & Lai, 2020; Tutić, 2023). Later the model was expanded and specified via the integration of moral pluralism under the headline of Moral Foundations Theory (Haidt & Joseph, 2007; Graham et al., 2013). We will discuss the latter theoretical account in more detail below, but for now we are interested in the role of intuition and emotion in humans morally guided self-regulation. At the core of the Social Intuitionist Model of Moral Judgment are sudden and quick, automatic, effortless and uncontrollable gut feelings of moral intuition. These intuitions are distinguished from slow, effortful and controllable reasoning that may follows initial intuitive impulses. In Haidt's model, reasoning takes the place of affectively charged ex post facto rationalization of moral intuitions. Accordingly, in contrast to reasoning, moral intuition is defined "as the sudden appearance in consciousness of a moral judgment, including an affective valence (…), without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion" (Haidt, 2001, p. 1029). Moral intuitions are thus about flashes of positive or negative feelings that lead to evaluations that instantly pop up in one´s mind when one is confronted with moral content in the social sphere of ours (Haidt, 2003). In addition, the model includes six links that are proposed as pathways that form moral judgment (Haidt, 2001). The first four links build essentially the core of the model. To begin with, the *intuitive judgment link* states that moral judgments arise in our consciousness as the result of moral intuitions (link 1). Second, the *post hoc reasoning link* comprises the potential of individual post hoc reasoning that works in support for one´s initial gut reaction (link 2). Third the *reasoned persuasion link* puts forward that actors verbally communicate their initial gut reactions to justify their moral judgment to others (link 3). Eventually, the fourth link is called the *social persuasion link*. This link is built upon our proneness to follow group norms and proposes that actors' moral judgments are shaped by the moral judgements in their immediate social environment regardless of the use of persuasion. Hence, it is argued that the mere presence of moral judgments from (relevant) others brings about direct social influence on actors (link 4). This is where the social component of the model comes into play, as reasoning primarily becomes relevant to an actor's moral judgment when it is the reasoning of other people (i.e., links 3 and 4) (Haidt, 2001). Moreover, link 5, the *reasoned judgment link*, and the *private reflection link* (link 6) are proposed as slightly more outlying parts of the model. These links account for the possibility that some highly experienced individuals (e.g., skilled philosophers) may, on rare occasions, base moral judgments on logic that that surpasses the initial gut reaction (link 5). Link 6 further suggests that when reflecting on a situation, other intuitions may be activated that contradict the initial gut feeling. This in turn may lead to a moral judgment guided either by the strongest intuition or even by a

conscious choice between alternatives of the moral intuition. The latter two links by implication reflect the integration of the rational approach to morality into the social intuitionist model. Nonetheless, the model is primarily concerned with the first four links, which emphasize emotionally charged intuitions (gut feelings) and the social component in an individual's moral judgment. In essence, this model puts forward that our evolved moral mind is composed of fast and uncontrollable gut feelings that make us to adhere to the good and to instantly condemn the bad without necessarily the need of a single conscious thought. Thus, from this view it is rather the unconscious and emotional side that brings in the regulatory force of our morals.

There are further indications of the primary importance of moral emotions in regard to moral actions. Moral reasoning does not necessarily imply that morally guided actions follow. In contrast, the sense of unjust treatment of others or the suffering of fellow human beings reveals a more direct path to moral action. Various studies have shown that moral reasoning is only weakly correlated with moral behavior, but that moral emotions are strong moderators of our moral actions (Haidt, 2001). Moral emotions can be roughly understood as intra-individual systems of punishment and reward that are linked to motivations that either discourage or encourage certain behaviors (Tomasello & Vaish, 2013). The former reflects the negative emotional side and is caused by deviance or violations of moral standards. The latter, in contrast, reflects the positive side and is caused by morally conform behavior. It is the actions of ourselves or those of other actors that can be morally conforming or deviant. Consequently, action of both ourselves and other social actors can be identified as sources triggering moral emotions. According to Tangney and colleagues (2007), it is moral emotions that "provide the motivational force—the power and energy—to do good and to avoid doing bad" (p. 347). As such, moral emotions are of central importance to moral behavior.

Haidt (2003) believes that any emotion can actually be considered a moral emotion to some degree. In his view, the extent to which an emotion can be regarded as moral depends on two factors. As long as the source that triggers the emotion is disinterested in the sense that it primarily concerns the general social welfare and not personal interests, and as long as the action tendency triggered by the emotion is pro-social, it can be considered moral. Although any emotion can be considered moral according to this idea, research has identified particular emotions that are known to be linked to our morality in a specific sense. Haidt (2003) groups these emotions into superordinate categories, in his terminology "emotion families", and distinguishes, among other things, between the other-condemning emotions, the self-conscious emotions, and the other-suffering emotions.

At the core of *other-condemning emotions* are anger, disgust and contempt that comprise negative feelings about the action or character of another person. When we are witness that someone else or ourselves is treated unfairly without just reason a sudden negative impulse of arousal may arise in us. Unjustified moral wrongness of others is potent in causing an instant feeling of anger and subsequent inclination for revenge, i.e., a motivation to punish the wrongdoer. Disgust, the emotion that motivates avoidance, is not just caused by food-related phenomena, as rotten meat for instance, but is argued to be also caused by moral violations to cultural rules of how to use and treat the body right (Nietzsche, 1887/1991; Sherman & Billing, 1999; Murray & Schaller, 2010; Atari et al., 2022b). Furthermore, the emotion of contempt may be caused by people who do not live up to respective social standards of hierarchy given in one´s socio-cultural environment. The feeling involves an inclination of social-cognitive change in the appraisal of others towards which we feel this emotion (Haidt, 2003). Up-ward and down-ward contempt can be felt as response to violations of social ranks, disrespect towards authorities or unfulfillment of role duties. Overall, when we perceive moral deviance, our other-condemning emotions motivate subsequent action tendencies of punishment and generally negatively altered social relationships. The development of these emotions is therefore seen as adaptive, as they in turn motivate people to self-regulate their behavior, which ultimately contributes to the maintenance of social order. The self-regulatory mechanism of moral emotions becomes clearer when we turn to self-conscious emotions.

Shame and guilt belong to the family of *self-conscious emotions* as they are tangled to conscious or unconscious self-reflection and self-evaluation (Tangney et al., 2007; Wong & Tsai, 2007). These negatively valanced emotions deal primarily with reactions to the self or with anticipated reactions to the self and its behavior. Self-conscious emotions are thought to be beneficial to the group life of the individual by guiding the individual's behavior so that it does not violate moral standards. This in turn is adaptive and serves the individual in the group, as it prevents individual's behavior from provoking reactions of other-condemning emotions, thereby averting reputational damage and non-selection as a cooperation partner (Tomasello & Vaish, 2013; Baumeister, 2022). Shame and guilt are tight to the cultural constitution of selfhood (Haidt, 2003; Leung & Cohen, 2011). In societies promoting pre-dominantly independent self-construal embarrassment and shame exist as discrete emotions. Whereas shame and embarrassment form a single emotion in sociocultural entities characterized by hierarchy and interdependent self-construal predominance. Apart from cross-cultural differences both moral emotions work to regulate our self. The feelings of shame and guilt are evoked in us by the act of committing moral transgressions. We experience shame and guilt as

aversive, painful emotions. As such these emotions provide instant feedback in form of self-punishment when we violate to moral standards. Furthermore, we even anticipate shame- or guilt-laden behavior and the associated emotional costs and often act accordingly to avoid them. Essentially, when we experience shame, it is the entire self that is at stake and subject to negative evaluation. Guilt, in contrast, is associated with negative evaluations of the particular behavior of the moral breach. Certain behaviors, compared to a person's whole self, are easier and therefore more likely to change. Empirically it has been shown that feelings of guilt in particular elicit prosocial consequences, behavioral changes, and other-oriented reparative actions. Shame works also to decrease the likelihood to commit transgressive shame-inducing behaviors. However, shame is also associated with defensive reactions and more so with self-oriented distress minimization (Tangney et al., 2007). Nonetheless, the regulatory power of shame and guilt is truly profound and remarkable. Not only do these emotions effect self-regulation, they also signal to social others that the moral transgressor is already suffering from self-punishment. This signal, in turn, curbs the punitive feelings of others and transmits that the transgressor cares about and commits to moral standards, and is thus still a sufficiently reliable cooperation partner (Tomasello & Vaish, 2013). Self-conscious emotions also comprise a positive side of pleasant sense. The positive counterpart to shame and guilt is the emotion of pride. Adhering to moral standards can evoke pride in us, which eventually reinforces the behavior we have committed. Overall, self-conscious emotions are directly linked to the self and the intra-psychological structure of instant punishment and reward. Consequently, we carefully monitor our own behavior and regulate ourselves to avoid the emotional pain caused by shame and guilt, and are motivated to follow morality in order to experience the pleasurable sensation of pride (Haidt, 2003; Tangney et al., 2007).

Another family of moral emotions, according to Haidt (2003), is the class of *other-suffering emotions*, which includes sympathy, compassion and generally our ability to feel what others feel, i.e., empathy. We sense when others suffer, and especially when they are relevant and close others such as kin or peers, we are motivated to erase or at least alleviate the emotional pain they feel. Empathy might be better understood as an emotional process rather than a single emotion since it involves emotional and cognitive elements (Frith & Frith, 2005; Tangney et al., 2007). Commonly, this process is known as perspective taking in conjunction with emotional engagement. It involves putting ourselves in the other person's shoes and thus being able to be infected by the other person's emotional state, appraise the cause of the other person's emotions, and finally adopt the perspective of our social counterpart. Frans de Waal (2008), one of the world's most famous primatologists and behavioral scientists, provides ample evidence

for the evolutionary history of empathy in humans. For example, several signs of empathy are found in our animal relatives (great apes), but not in monkeys. He argues that empathy is a central mechanism in human behavior that serves to alleviate the needs or pain of others. From this perspective, empathy is therefore crucial to support directed altruism. Tangney et al., (2007) emphasize the role that empathy plays in the moral emotion system for three simple reasons: first, because we feel the suffering of others, we also feel concern for them. Second, because of our concern for others, we are motivated to engage in pro-social behaviors such as helping behavior. Third, as we are able to feel others harm and sense their suffering, we are less likely to engage in aggressive behavior and hurt others. Apparently, empathy is of fundamental importance for everyday human social life. In general, the evolved human capacity for empathy allows us "to quickly and automatically relate to the emotional states of others, which is essential for the regulation of social interactions, coordinated activity, and cooperation toward shared goals" (De Waal, 2008, p. 282).

Beyond empathy and as already indicated in relation to pride, there is also a bright side to moral emotions. Pleasant moral emotions are, for example, gratitude, awe and elevation (Haidt, 2003; Crocker et al., 2017). These emotions are triggered internally, i.e., in the individual due to a person´s own good deeds, and externally by the good deeds of others. Positive moral emotions form the counterpart to the negatively valanced, punishment-like emotions and reflect emotional rewards when we ourselves or others conform to morality. Research shows that feelings of gratitude or elevation are associated with increased pro-sociality, helpful behavior and positive affect towards social others who trigger the emotion. So, in addition to punishment-like negative emotions and our empathy, it is also the rewarding effect of positive emotions that can promote and reinforce moral behavior in ourselves and in social others.

That said, our social behavior is not only guided by reasoning, intuitions, and emotions within the individual. Our social environment also has means at its disposal and is permeated by a general incentive system that motivates us to behave morally. As humans, we are active norm enforcers, as has already become evident with regard to the other-condemning emotions. Normative and moral transgressions are actively punished by social actors (Henrich & Chudek, 2011; Chudek et al., 2016). As it can be found in the general theory about norms (Axelrod 1986; Abrams et al. 2000: Esser 2002a; Horne & Cutlip 2002; Fehr & Fischbacher 2004; Lapinski & Rimal 2005; Hogg & Reid 2006; Frings et al., 2012; Tomasello, 2017), certain social behaviors get socially incentivized, i.e., negatively sanctioned in case of deviance and positively rewarded in case of compliance. The reputation of actors plays an important role here, as it marks

individuals as reliable or unreliable cooperation partners (Baumeister, 2022). Actors can increase their reputation when they perform good deeds and decrease it when they commit moral violations (Henrich & Muthukrishna, 2021). So, reputation can be understood as a kind of social barometer that, depending on which way it swings, makes the individual more or less attractive for joint ventures to achieve shared goals. Reputation thus mediates cooperativeness and acts as a social gatekeeper for access to resources, whether tangible or intangible. The social relationship between actors can be affected by transgressions to such an extent that the punishment for transgressive behavior is ostracism from a social group, or worse and fatal (Henrich, 2020; Wrangham, 2021). In evolutionary terms, the social group is necessary for the survival of the individual and also nowadays our comrades are pivotal for our well-being as for social resources generally (Ormel et al., 1999; Esser, 2010; Crocker et al., 2017; Lippke et al., 2021). The extent and scope of the punishments for moral transgressions becomes apparent by implication; the individual consequences of punishment can be far-reaching. The disposition to cooperate and to punish moral violating non-cooperators can be argued to be hard-wired in humans, as evidently portrayed by an example from Henrich and Chudek (2011): "cooperating and punishing in locally normative ways activates the brain's rewards or reward anticipation circuits in the same manner as does obtaining a direct cash payment" (p. 224). We thus feel it self-rewarding when we punish transgressors. Moreover, research shows that even as unaffected third-parties, we carry within us a tendency to punish when others commit moral breaches. This can be an expression of altruistic punishment and the internalization of moral standards (Gavrilets & Richerson, 2017). Another explanatory approach sees in third-party punishment the tendency for the punishing actor to signal that they adhere to the standards of the group and thus display themselves to be a reliable cooperation partner. Third-party punishment therefore may serve to enhance the reputation of the punishing actor (Henrich & Muthukrishna, 2021). The tendency to punish moral transgressions occurs early in human ontogeny and is evident in young children. Already infants as young as 3 years old show active third-party punishment (Tomasello & Vaish, 2013; Tomasello, 2017). Beyond that, institutions, such as the justice system, are also an expression of social means that demand for self-control. Social institutions help to ensure the transgenerational transmission of knowledge of social, moral and justiciable standards as well as the punishments that follow as socially legitimate in the event of transgression (Berger & Luckmann, 2013; Tomasello & Vaish, 2013; Diekman & Lindenberg, 2015; Muthukrishna, 2021). Conclusively, in addition to the rather intrapsychic modes of punishment and reward, moral behavior is also actively demanded by our social environment. The social means of punishment, signaling, and reputation are able to uphold moral standards

and incentivize the individual self of actors to conform to group norms as well as general moral standards (Henrich & Muthukrishna, 2021).

So, what makes us to regulate ourselves? Evidently, an important part to the answer of this question can be found in our evolved moral mind. We are able to think in such a way that we consciously take care of others. Reasoning, reflected in institutionalized proclamations, society-specific laws (ius positivum) and universal human rights (ius naturae), are excellent testimonies to this capacity. In the plethora of our everyday interactions, however, it seems more likely that most of the time immediate flashes of intuition, emotion and empathy guide our behavior and shape it in ways that are consonant with moral standards. These emotion-driven internal incentive structures can be seen as "adaptive syndromes shaped by evolution to make people liable to "normative governance," that is, the pull of rules and moral discourse" (Haidt, 2003, p. 865). Morality functions within the individual to regulate us by emotional punishment or reward as consequences to moral deviance or moral conformity committed by ourselves or actors around us. Our moral sentiment is further guided by our evolved empathy that makes us capable to feel the harm of others and motivates us to care about each other. Beyond that, grounded in the social environment of us the means of punishment, signaling and reputation emerge as strong incentives to act in accordance with moral standards. Punishing transgressive behavior is so deeply ingrained in us that we have a tendency to follow our punitive disposition even as a non-involved third-party, although at first glance this creates costs for the punishing individual. However, by punishing others for their transgressions, we signal that we are reliable cooperation partners. This can have positive effects on our reputation and in turn motivate others to cooperate with us. Through our social credit system of reputation, we are therefore not only encouraged to curb possible selfish drives, but also gain access to resources, be they material, immaterial or social. Morality is essentially an adaptive response to the requirements of human life. It supports the maintenance of social order, which is important for the survival of social groups. Further, morality encourages individuals to take care of their reputation in order to participate in social life through mutually beneficial actions, i.e., through cooperation. But what actually happens when our moral capacity is absent or impaired? Psychopaths are not known for diminished intelligence or cognitive ability, but for a greatly reduced or even absent emotional side, which we have highlighted here as particularly important for moral self-regulation. Referring to psychopaths, Haidt (2001) states: "[w]ith no moral sentiments to motivate and constrain them, they simply do not care about the pain they cause and the lives they ruin" (p. 1038). We witness moral transgressions time and time again not just in psychopaths but in wars, murders and the most diverse atrocities that people can inflict on one

another. Imagine a world in which everyday life would be fully determined by transgressions of these kinds. Such a world would leave no room for increasing social organization, nor for the life that many of us are lucky enough to know. However, as mentioned before, despite the crises we still find all over the world in modern times, we are yet probably living in the most peaceful times in our human history and our morality plays a vital role in that (Waytz et al., 2019; Kirkland et al., 2023). Morality permeates people's everyday social life. It is readily apparent that our inner apparatus of pro-sociality is a central asset of all human coexistence. Morality helps to regulate selfish impulses and is part of why we are a social species living in large groups — it makes us to see pro-social behavior of interindividual care as good and harm-inducing action as bad. As humans, we are two-sided beings: we are selfish, antagonistic cooperators and we are also prosocial, other-oriented cooperators (Tomasello & Vaish, 2013; Crocker et al., 2017). With regard to this duality, means to promote our prosocial side have emerged in the course of evolution and cultural evolution. The means to promote self-regulation can be found both internally, within the individual, and externally, in the individual's social environment. Morality, then, is part of what causes us to self-regulate, and works via internal and external forces that include reason, intuition, emotion, and the social means of punishment, signaling, and reputation. In the everyday practices and interactions of our lives, "moral concerns and moral decisions arise from situational realities, characterized by people's *experiences*" (Ellemers et al., 2019, p. 336) and these experiences are in turn guided by the intra-psychological and resulting social means of our evolved moral minds.

1.7. Moral Foundations Theory and Morality as Cooperation Theory: A Coalescent Perspective

We are a species that is pro-socially oriented yet also selfish. Morality has the function of regulating individual egoism, which in turn facilitates altruism, individual and group cooperation, and generally other-regarding preferences and pro-social behavior (Kurzban et al, 2015; Blake et al, 2015; Cooper & Kagel, 2015; Carlo et al, 2016; Crocker et al, 2017; Tomasello, 2017; Summerville & Enright, 2018). *Moral Foundations Theory* (*MFT*) and *Morality as Cooperation Theory* (*MaC*) are two relatively new approaches to morality based on a functionalist perspective and an overall cultural evolutionary framework. After an introduction to these theories, we will propose a theoretical perspective that synthesizes MFT and MaC into a joint stance. Based on this coalescent perspective, we will then present three new research tools for the study of morality. This section is followed by a discussion of morality

and cross-cultural differences. Finally, in the following section, we present the research gaps we have identified, a theory-based research model, and a list of hypotheses to be tested empirically.

1.7.1. Moral Foundations Theory

In the introduction to this article, we have asked you if you could recall a situation where you have been treated unfairly and if you have felt an instant flash of negative sentiment towards those who have treated you like this. In the subsequent we will expound key elements of the Moral Foundations Theory (MFT), which is a framework that is conceptualized to clarify the nature of intuitive reactions towards particular patterns of behavior in the social world of ours.

In the center of the **_Moral Foundations Theory_** (**_MFT_**) lies a functionalistic approach to and an intuitionist perception of morality, as well as the idea of universal moral pluralism and cross-cultural variation in moral domain endorsement (Haidt & Joseph, 2007; Graham et al., 2013; Atari et al., 2020a; 2022b). MFT suggests that our moral mind is designed by natural selection to be inherently functional for it bridles human egoism: moral systems "suppress or regulate selfishness and make social life possible" (Graham et al., 2011, p. 4). Moreover, as an intuitive approach to morality, MFT emphasizes rapid, effortless, and uncontrollable gut feeling like responses as a consequence of perceiving moral content (Haidt, 2001; Haidt & Joseph, 2007). Without denying deliberate and time-consuming moral cognition, MFT is primarily placed on the opposite side of it: moral intuition represents a position more resembling implicit/Type-1 processes (Kahneman, 2011; Graham et al., 2013; Greenwald & Lai, 2020; Tutić, 2023). The principles of the theory are based on evolutionary thinking. Against this background, MFT argues that our ancestors faced multiple recurrent adaptive challenges that led to the development of plural moral foundations. Thus, it proposes that our moral minds consist of not just one but plural moral foundations to which we respond in a content, domain-specific fashion. Without claiming to present a comprehensive list of moral domains, MFT proposes the domains _harm/care_, _fairness/reciprocity_, _ingroup/loyalty_, _authority/respect_, and _sanctity/purity_ as the plural foundations of our moral mind (Haidt & Joseph, 2007; Graham et al., 2013). Furthermore, MFT was essentially developed to explain the evolution of moral domains (foundations) and cross-cultural variation of morality. In explaining their theory, Haidt and Joseph (2007) use the very appropriate metaphor of a first and second draft of the moral mind. Following a nativist position that emphasizes the universality of morality, MFT argues that the human species is evolutionarily endowed with a first draft of the moral mind. This first

draft consists of several moral intuitions that are formed and organized in advance of our experience. However, as indicated earlier, MFT is also a culturally sensitive approach. Haidt and Joseph (2007) state: "moral maturity is a matter of achieving a comprehensive attunement to the world" (p. 387). In this sense, it is further argued that the particular cultural environment edits the first draft of the moral mind to adapt the second draft to the context. Thus, although MFT holds that moral foundations are universally shared, the theory also claims that the particular emphasis on these foundations is context dependent.

Apart from the great contribution that MFT´s moral pluralism has made to the field, Haidt (2008) has also reflected on another facet of morality and proposed two diverging, higher order moral constructs — a ***binding approach to morality*** and an ***individualizing approach to morality***. *Binding* morality is suggested to be composed of *authority*, *loyalty* and *sanctity*. Furthermore, people endorsing binding over individualizing morality are theorized to promote the respective binding domains relatively more. The focal point of binding morality lies within groups (kin and in-group). Haidt (2008) and colleagues suggest in this respect that binding morality is functional in the sense that it restricts selfishness by uniting and binding "individuals into tightly knit collectives" (Mooijman et al., 2017, p. 2). Opposed to binding the focal point of the *individualizing* approach to morality lies in the protection of the individual. Characteristic of the individualizing morality is a relatively higher endorsement of the foundations *care* and *fairness*, for these moral domains "are all that are needed to support the individual-focused contractual approaches to society" (Graham et al., 2011, p. 6). Consequently, within the *individualizing* approach to morality "individuals are the fundamental unit of moral value [whereas within the] *binding* approach to morality (…) the group (…) [is] the fundamental source of moral value" (Haidt, 2008, p. 70).[13] Apparently, it seems as if the concepts of binding and individualizing morality reflect the essential features of an interdependent and independent way of selfhood (Markus & Kitayama, 1991; 2010; Cross et al., 2011; Vignoles et al., 2016). Apart from its theoretical value, the idea of binding and individualizing morality is also appealing because it combines moral pluralism, at least five moral foundations, with a parsimonious two-type model. However, empirical findings regarding binding and individualizing are mixed at best, and one of the most recent MFT studies on moral judgments found no cross-culturally valid pattern supporting the idea of these higher-order moral constructs (Graham et al, 2011; Mooijman et al, 2017; Enke, 2019; Atari et al, 2022a).

Overall, the Moral Foundations Theory (MFT) broadened the view on morality substantively and has been found to be applied tremendously successful in empirical studies

---

[13] Italicization adopted from the original source. Words in parentheses were added by the author.

that deal with the forecast and explanation of differences in political ideology (especially in the USA) (Haidt et al., 2009; Graham et al., 2009; Iyer et al., 2012). Also, several measurement instruments (e.g., two self-report scales and moral vignettes; Graham et al., 2011; Clifford et al., 2015; Atari et al., 2022a) to capture judgment and relevance of multiple moral domains were suggested under the MFT framework and belong today to the leading instruments in empirical moral research. However, also theoretical criticism and critiques on the proposed measures have become louder in recent years (Curry et al., 2019b; Skitka & Conway, 2019; Iurino & Saucier, 2020).

1.7.2. Morality as Cooperation Theory

Tomasello and Vaish (2013) state that from "an evolutionary perspective, morality is a form of cooperation" (p. 231). One line of the criticism of MFT stems from the Morality as Cooperation Theory (MaC) (Curry, 2016). This approach is conceptualized to overcome criticized shortcomings of MFT. The *Morality as Cooperation Theory* follows also a general cultural evolutionary framework, advocates moral pluralism and is functionalistic in nature (Curry et al., 2019a). At the core of MaC is the idea that the central function of morality lies within the promotion of human cooperation. Basically, it is argued that our moral mind evolved as adaptive response to the multiple challenges of cooperation. Hence, it is argued that "morality consists of a collection of biological and cultural solutions to the problems of cooperation recurrent in human social life" (Curry et al., 2019b, p. 107). Further, Morality as Cooperation Theory embeds the definition of cooperation in a game theoretical framework (Diekmann, 2013), and treats cooperation consequently as non-zero-sum interaction between human individuals and groups.[14] According to MaC natural selection has thus designed our evolved moral mind to enable human interaction to result in mutual gain. This means that morally guided cooperation leads to outcomes of non-zero sum, or in other words, to a win-win situation for the actors involved (Curry, 2016). In respect to moral pluralism MaC suggests an even more nuanced picture than MFT and proposes at least 7 moral domains (Curry et al., 2019b). MaC advocates

---

[14] As has been emphasized repeatedly, we humans are predisposed to be both selfish and pro-social. A (selfish) zero-sum game exists when actors pursue opposing interests and the gain of one actor always implies a loss for the other actor. From a game theory perspective, war is a prime example of a zero-sum game. The opposite, the (pro-social) non-zero-sum game, assuming that actors cooperate with each other, implies that the actors mutually benefit from each other (Esser, 2002c; Diekmann, 2013). Our evolved moral mind, which comprises both intrapsychic and social means of self-regulation, makes us principally an ultrasocial species (Haidt, 2003; Tangney et al., 2007; De Waal, 2008; Tomasello & Vaish, 2013; Henrich & Muthukrishna, 2021). The evolved pro-social drives in us cause us to have both a disposition to cooperate and a disposition to punish violations to the social cooperation regime. We have already outlined this aspect in detail.

that *fairness*, *reciprocity*, *property*, *family*, *in-group*, *deference*, and *heroism* are domains of cooperation that are regulated by our evolved moral mind. ***Table 2*** is adopted from Curry et al.,

Table 2: Overview of Morality as Cooperation Theory

	Label	Problem/Opportunity		Solution	Virtues	Vices	Epithet
1	Family	Kin selection		Kin Altruism	Duty of care, special kin obligations	Incest, neglect	*Blood is thicker than water*
2	Group	Coordination		Mutualism	Loyalty, unity, solidarity, conformity	Betrayal, treason	*United we stand, divided we fall*
3	Reciprocity	Social Dilemma		Reciprocal Altruism	Reciprocity, trustworthiness, forgiveness	Cheating, ingratitude	*One good turn deserves another*
4	Heroism	Conflict (Contest)	Resolution	Hawkish Displays	Bravery, fortitude, largesse	Cowardice, miserliness	*With great power comes great responsibility*
5	Deference	Conflict (Contest)	Resolution	Dove-ish Displays	Respect, obedience, humility	Disrespect, hubris	*Blessed are the meek*
6	Fairness	Conflict (Bargaining)	Resolution	Division	Fairness, impartiality, equality	Unfairness, favouritism	*Let's meet in the middle*
7	Property	Conflict (Possession)	Resolution	Ownership	Respect for property, property rights	Theft, trespass	*Possession in nine-tenths of the law*

The table is adopted from: (Curry et al., 2019b, p. 108)

(2019b) and gives an overview of the Morality as Cooperation Theory. Apparently, although highlighting especially the value for human cooperation, several domains of MaC reflect also the domains proposed by MFT.

Regarding universality and diversity MaC favors, alike to MFT, a mixed stance. Arguing that the recurrent challenges that have led to the evolution of our moral mind where universally the same for our species, it follows that the composition of our moral mind is universal too. Consequently, MaC predicts that the proposed 7 domains of "cooperative behaviors will be considered morally good in every human culture, at all times and in all places" (Curry, 2016, p. 40). A study across 60 societies partly supports this notion and was able to demonstrate that the domains proposed by MaC are cross-culturally considered to be good (Curry et al., 2019a). However, MaC further predicts that the extent of recurrent challenges differs across people and cultures, leading to varying priorities in respect to moral domains. Hence, variation in the relevance of moral domains, due to diverging sociocultural needs and affordances, is part of morality too. Taken together the aforementioned gives rise to the universalist yet also culture sensitive perspective of MaC (Curry, 2016).

Scholars of the Morality as Cooperation Theory, though, criticized not only parts of the theoretical MFT framework, but also the self-report scale proposed by the Moral Foundations Theory (Graham et al., 2011; Curry et al., 2019b; Iurino & Saucier, 2020). Pointing to psychometrical weak points of the Moral Foundations Questionnaire 1 (MFQ-1), Curry and colleagues (2019b) developed a new self-report instrument designed to assess the 7 domains

proposed by MaC. In addition, they were able to empirically demonstrate the superiority of their scale compared to the MFQ-1.

Overall, MaC highlights the functional utility of our co-evolutionary developed moral mind for human cooperation. The approach takes up many positions of MFT. However, due to its reference to game theory (Esser, 2002c; Diekmann, 2013; Diekmann & Lindenberg, 2015), MaC has a stronger theoretical foundation than MFT. Above that, MaC came up with a self-report scale that is, compared to the first self-report instrument proposed by MFT, broader in domain coverage and superior in terms of psychometric properties.

1.7.3. Coalescing MFT and MaC

Taking a look at the theoretical pillars upon which MFT and MaC are built, it is apparently visible that both theories share more similarities than discrepancies. MFT and MaC are conceptualized against the background of a gene-culture co-evolutionary framework and suggest moral pluralism instead of monism. Also, both theoretical accounts treat morality not from a normative but from a functional perspective. More so, they share the idea of a universal moral mind and simultaneously context sensitive adaptations of domain endorsement to the needs and affordances of respective sociocultural ecology. Conclusively, coalescing MFT and MaC to a joint theoretical stance appears tempting and possible.

In this respect, we propose a stance that is combining MFT and MaC. First and foremost, we agree on the major positions both theories share and follow generally a co-evolutionary framework (Haidt, 2001; 2003; 2008; Haidt & Joseph, 2007; Henrich & McElreath, 2007; Richerson et al., 2010; Boyd et al., 2011; Chudek et al., 2016; Curry, 2016; Brown et al., 2022). Also, we consider morality to be plural and yet traversed by a single guiding principle. Further, we regard human cooperation as the focal point of the regulatory effect of morality and grasp our moral mind as inherently functional and universal but still shaped by the respective sociocultural ecology. As working ***definition of morality***, that is grounded in our perspective, we suggest the following: ***moral systems have the function of regulating and identifying egoism, which in turn enables cooperation (i.e., non-zero-sum interaction) between individuals and (within/between) groups, and fosters social life between people and the evolving of human social organization.***

1.7.4. MFT and MaC Coalescent — Moral Pluralism

The joint approach that we suggest considers *moral domains* as social domains of human cooperation that are themselves regulated by our moral mind. Without claiming to give a comprehensive list, we further propose that our moral mind comprises at least the following 8 moral domains: *fairness*, *reciprocity*, *property*, *family*, *in-group*, *deference*, *heroism* and *trustworthiness*.

Apart from trustworthiness, all other domains are adopted from the original suggestion by MaC (Curry, 2016). However, we think that these domains also capture what MFT has proposed: fairness, in-group, deference/authority appear to be apparently equivalent between MaC and MFT. As can be seen, we have not included the harm/care dimension directly, and will come to this in detail further below. Moreover, we did not integrated MFT´s purity domain for the moment. Purity is considered for long as a moral category (Nietzsche, 1887/1991) and we can also envision the potential of this domain from a cooperation perspective. We have, however, for now decided against the inclusion of purity, because we follow the view that the adaptive challenges to which this domain has evolved in response are primarily natural rather than social in nature, which is contrary to all other MFT domains (Haidt, 2001; Haidt & Joseph, 2007; Graham et al., 2013; Atari et al., 2022b). More theoretical work is needed to grasp purity from a primarily cooperation-based perspective on morality.

We have furthermore extended MaC´s list of moral domains by proposing *trustworthiness*. Previous research has related trust and trustworthiness with morality and demonstrated its cross-cultural variability that at least partly results from differences in sociocultural ecologies (Henrich et al., 2005; Henrich et al. 2010b; Thomson et al., 2018; Cohn et al., 2019; Enke, 2019; Schulz et al., 2019; Curtin et al., 2020; Muthukrishna et al., 2020b; Henrich & Muthukrishna, 2021; Kirkland et al., 2023).[15] We have elaborated more detailed on trustworthiness as a moral domain elsewhere (Jessen et al., 2023)[16] and will concentrate here on a short description of our position. In a nutshell, we propose that trustworthiness is a normatively preserved, collectively (at large scale) shared expectation onto other people's behavior that is attuned to the social context, and individually updated based on diagnostic behavioral signals from potential partners of cooperation (Hollander, 1958; Henrich, 2009; Carter & Weber, 2010; Diekmann & Lindenberg, 2015; Van Lange, 2015; Muthukrishna, 2021; Bjørnskov, 2021; Baumeister, 2022). In terms of behavior keeping commitments, promises and

---

[15] Variations in trust across different personality types are also known (Engelmann et al., 2019).

[16] See the pre-registered research plan available online at: https://doi.org/10.23668/psycharchives.13059. More information on different research plans follows in the course of this thesis.

secrets — in general, showing one´s *reliability* — are signals of an actor's trustworthiness and core elements of this moral domain. Thus, different to MaC (Curry, 2016) we suggest to disentangle reciprocity and trustworthiness. Reciprocity best unfolds its pro-social effects as a consequence of repeated cooperative interaction (tit-for-tat) (Maus, 1968; Fehr & Fischbacher, 2004; Moebius, 2006). In contrast, we believe that trustworthiness as a moral domain primarily solves coordination problems related to the question of with whom to cooperate in the first place (without being taken advantage of). In essence, we argue that trustworthiness precedes cooperative interactions, especially the initial ones, and solves coordination problems related to whether or not to engage in a cooperative opportunity. Therefore, trustworthiness is functional for it helps us to identify and regulate egoism in regard to human reliability in cooperative interaction. We further presume on the one hand that trustworthiness is especially relevant in social systems with high interaction rates among strangers. On the other hand, we expect that reciprocity should be particularly prominent and relevant in social ecologies with high (historical) kinship intensity and low relational mobility (Thomson et al., 2018; Enke, 2019; Schulz et al., 2019; Talhelm, 2022).

1.7.5. A Guiding Principle of Morality — Right and Wrong Reconsidered: Moral Conformity/Care and Moral Deviance/Harm

The human mind has evolved to distinguish between *conformity* and *deviance* in a variety of social areas. We propose that this principle most likely operates in the human moral mind as well. For our coalescing perspective of MFT and MaC we suggest that human morality operates via a guiding principle that permeates the functioning of morality regardless of the domain. Essentially, we argue that the guiding principle of our moral mind enables us to distinguish between and to recognize a difference of *moral conformity* and *moral deviance*. Further, we pose that this principle triggers appropriate affective and cognitive responses to both perceptions. We suggest to grasp the evolution of the primal behavioral evaluation of right and wrong to stem from the realization of gain and the infliction of cost that are related to moral conformity and deviance in potentially cooperative interactions. Based on a view of morality that places cooperation at the center, we treat moral deviance as equivalent to losses and costs. Likewise, we view moral conformity as equivalent to (mutual) cooperative gain(s) of whatever kind of resource(s). Hereof moral deviance and harm, and moral conformity and care blend to form wrong and right as the evaluative ends of the continuum of moral behavior. Hence, at the core of the guiding principle of morality is the capacity to differentiate and recognize *moral*

conformity — cooperative gain realizing action that is tangled to positive evaluations and cognitive categorization into morally right/good —, and *moral deviance* — non-cooperative loss realizing action that is tangled to negative evaluations and cognitive categorization into morally wrong/bad.

1.7.6. Deviance and Conformity as Bad and Good in the Human Mind

Regarding group dynamics, in-group members compared to out-group members are evaluated less favorably and sanctioned more harshly when they show (in-group) norm-deviant behavior (Abrams et al., 2000; 2002; Marques et al., 2001; Frings et al., 2012). Generally deviant behaviors to social standards or subjective political beliefs are associated with changes in brain activity. The perception of deviance in this respect leads to higher activation of the amygdala that itself is associated with the processing of negative emotions such as threat and fear (Stephan et al., 2009; Schreiber & Iacoboni, 2012; Amodio, 2014; Kaplan et al., 2016). Not only has the Prospect Theory (Kahneman & Tversky, 1979) taught us that potential losses are of higher importance to humans than potential gains. A review by Baumeister et al., (2001), covering a wide range of psychological disciplines, further supports that our minds generally tend to distinguish between bad and good and to associate these states with loss and gain. They found evidence indicating a general psychological principle: bad is stronger than good. Rooted in evolutionary thinking Baumeister and colleagues (2001) argue for the survival relevant adaptive power to experience (and value) bad stronger than good:

> "it is evolutionarily adaptive for bad to be stronger than good (…) throughout our evolutionary history, organisms that were better attuned to bad things would have been more likely to survive threats and, consequently, would have increased probability of passing along their genes" (Baumeister et al., 2001, p. 325).

The evolutionary notion by Baumeister et al., (2001) receives additional support by Chudek and Henrich (2011). They pose that humans possess an evolved norm-psychology. Essentially, it is argued that natural selection operating on gene-culture co-evolution has equipped us with a neurocognitive apparatus that is specialized to detect social patterns of deviance and conformity (Chudek et al., 2016; Henrich, 2020). Given the crucial importance of interdependency for the survival and reproduction of our ancestors, this constitutes a fitness-relevant capacity (Tomasello & Vaish, 2013). Above that, another evolutionary grounded

approach (Kurzban et al., 2001; Pietraszewski et al., 2014) suggests that we are cognitively equipped with "a number of evolved inferential elements for alliance mapping" (Tooby & Cosmides, 2010, p. 208) that support our mind's ability to zoom in into social patterns of conformity and deviance.

Overall, the human mind has evolved to be able to distinguish between conformity and deviance and to further evaluate the particular behavior that falls under one of the two behavioral categories. We conjecture that our moral mind shares this property. Given a cooperation perspective on morality, we posit that the human mind tends to evaluate moral deviance as bad because it realizes loss(es), and moral conformity as good because it realizes gain(s). We also believe that this principle operates in our moral mind regardless of the moral domain in question. Moreover, we hold that all moral behavior is permeated by the dualism of deviance and conformity, which in turn is equivalent to the dualism of harm and care.

1.7.7. Moral Deviance as Harm and Moral Conformity as Care

The latest MFT definition of the moral domain care/harm stems from Atari et al., (2022a) who define care as "[i]ntuitions about avoiding emotional and physical damage to another individual" (p. 12). Moreover, it should be noted that MFT considers harm to be equivalent to the opposite of care. We do not follow the Moral Foundations Theory view on harm/care directly. Instead, we argue that *care* and *harm* are a component of the underlying core principle of morality. Essentially, we grasp them as a reflection of what we have differentiated before under the terminology of moral *conformity* and *deviance*. Viewing harm as the essence of morality is not a new position in moral research, as demonstrated by proponents of the Dyadic Model of Morality (Skitka & Conway, 2019). We, however, follow a pluralistic view on morality and likewise the idea that all moral domains are traversed by a single guiding principle. Essentially, our approach unites moral monism — a single guiding principle of morality — and moral pluralism — several moral domains, that are traversed by the guiding principle of morality. In the light of this and coming from a game theoretical cooperation-based position on morality we want to elaborate on our argument in the subsequent.

We posit that moral deviance, irrespective of the moral domain, results in harm, at least for one party of an initially cooperative interaction. Consider briefly an unfair share of the profits after working on a joint project, or the failure to receive something in return when reciprocity fails. So, on the one hand, moral deviance represents (at least) the denial of a (mutual) gain and inflicts (at most) direct harm in situations that were initially of a potential

cooperative nature. In other words, moral deviance in interactions of potential cooperation implies real effort and no gain, i.e., harm due to loss of whatever type of resource. From this perspective, it is moral deviance that is consistent with the MFT understanding of harm and for which we have an intuition of avoidance and detection. On the other hand, we see moral conformity from a game theoretical perspective as a non-zero-sum interaction. From what we suggest *moral conformity* means to *care* regardless of the moral domain. The English language allows an exemplary play of words to illustrate our point of view in regard to the 8 moral domains that we propose: *Fairness* — I care that you are treated the same as others (e.g.); *Trustworthiness* — I care that I keep my promise to you (e.g.); *Property* — I treat your property with care (e.g.); *Family* — I care that my family members are well (e.g.); *Deference* — I take care of your instructions (e.g.); *Reciprocity* — today you paid, but I will take care of it next time (e.g.); *Heroism* — I take care to protect my loved ones (e.g.); *(In-)Group* — I take care to preserve our customs (e.g.). When we approach morality from the perspective of cooperation, it becomes apparent why morality is so crucial to the human social: there is only a nonzero-sum game in the interaction between individuals when the actors of the interaction conform (at least to some degree) to morality. Thus, the human moral mind effectively regulates selfish tendencies and pro-sociality effects realize when we empathize with others and care about a person other than ourselves. Eventually, instead of following MFT directly we propose that to care is not about the avoidance of harm, but rather about an intuition to conform to morals, which means that it is about the realization of cooperative gain(s) irrespective of the single moral domain that is focal to the interaction.

Overall, we posit that a guiding principle traverses our moral mind. Regardless of the moral domain in question, our moral mind is designed to recognize patterns of *moral conformity* and *moral deviance*, and to trigger appropriate responses to the particular perception. We propose that reactions to moral conformity/deviance are expressed in a positive/negative affective response (Haidt, 2001; 2003; Tangney et al., 2007; De Waal, 2008), a following evaluation, and a subsequent cognitive classification into right/wrong. In this context, we predict that moral deviance leads overall to stronger reactions than moral conformity for bad is stronger than good (Baumeister et al., 2001). The capacity to distinguish moral conformity from moral deviance, and to recognize this difference is of inherent functional value. Recognizing morally conforming behavior and evaluating it as good is adaptive because it makes mutually beneficial interactions appealing to us. Recognizing morally deviant behavior and condemning it as bad is likewise adaptive, because it makes interactions of potential exploitation and realization of losses repulsive to us. An evolutionary evolved mechanism that equips us with

the capacity to differentiate moral conformity and deviance, to recognize it in behavior, and to relate this behavioral observation to appropriate, emotion-laden evaluations that are followed by distinct cognitive categorizations (right or wrong), could therefore be considered beneficial for the reproduction of the organism of our species.

Our re-interpretation of care/harm may not totally hit what Haidt and Joseph (2007) originally had in mind when they came up with the Moral Foundations Theory. However, as we suggest that care/harm is at work in the guiding principle of morality that is underlying all moral domains, we hope to meet the initial MFT conceptualization in our coalescence of The Moral Foundations Theory and the Morality as Cooperation Theory at least indirectly. We also readily admit that our position presented is still in its infancy. In our view, it requires further theoretical refinement as well as empirical underpinning of our perspective. However, we see for moral pluralism, which is at the same time permeated by a single guiding principle, explanatory and synthesizing strengths that should justify the further pursuit of this idea.

1.7.8. A Tripartite of Higher Order Moral Constructs? Binding, Individualizing and a General Disposition of Cooperation

The Moral Foundations Theory proposes an individualizing approach to morality, which focuses on the individual as the fundamental source of moral values, and a binding approach, in which groups are the focus of morality (Haidt, 2008). Empirical evidence for the existence of cross-cultural patterns of binding and individualizing morality has been mixed (Graham et al., 2011; Enke, 2019; Atari et al., 2022a). However, we consider the parsimonious explanatory potential of the idea of higher-order moral constructs to be very valuable. Moreover, we believe that the mixed evidence regarding binding and individualizing is likely due to a lack of appropriate measurement instruments, a domain mapping (to higher-order moral factors) that is in need of revision and a little demand of higher order construct re-conceptualization.

As discussed in more detail later in this thesis, we propose the *Morality as Cooperation-Deviance Relevance Scale* (MaC-DRS) as a self-report instrument to measure the relevance of 8 moral domains. This scale is based on our MFT and MaC coalescent approach and is developed to capture relevance ratings for deviant behavior in the following moral domains: fairness, trustworthiness, property, family, deference, in-group, reciprocity, and heroism. We postulate that these moral domains can be assigned to higher order moral factors theoretically and also empirically.

Basically, we argue that **the moral domains of family, deference, and in-group tend to reflect a group-oriented function of morality, i.e., a *binding* approach to morality.** These domains bind individuals to groups in order to regulate egoism so that predominantly group cooperation can flourish. Conformity to the family, deference, and in-group domains means caring for group cooperation. Furthermore, we presume that these group-oriented moral domains are relatively *particularistic* with respect to the scope (expansiveness) of pro-sociality. We assume for this reason an inner and outer realm one´s group(s), as supported by the Social Identity Theory and numerous social and evolutionary psychological research (e.g.: Tajfel & Turner, 1986; Abrams & Hogg, 2004; Schaller & Neuberg, 2008; Schaller et al., 2010; Turner & Reynolds, 2012; Huddy et al., 2015). Based on the distinction of group membership (in-group / out-group), we assume a relative social boundary for the selfishness-regulating effect of the respective binding moral domains. Consequently, we suppose that binding morality is rather particularistic than impartial; it unfolds selfishness-inhibiting effects especially in cooperative interactions with in-group members.

In addition, we posit that **fairness, property, and trustworthiness tend to reflect the functional utility of morality in terms of protecting individuals, i.e., an *individualizing* approach to morality.** Conforming behavior with respect to these moral domains predominantly means to take care of cooperation with individuals, regardless of the individuals' social affiliation (i.e., group belonging). Accordingly, it follows that we view the individualizing domains of morality as relatively universal and *impartial* in terms of the scope of their pro-social effects.

Moreover, we have a hunch, but are not yet entirely clear, whether the domains of reciprocity and heroism fall under either binding or individualizing morality, or whether they represent another higher-order moral construct corresponding to a ***general disposition of cooperation***. Theoretically, reciprocity and heroism can regulate egoism to serve the individual and/or the group. There does not appear to be an overarching focus of cooperation with regard to these moral domains. For example, it is difficult to imagine the functioning of a social group without reciprocity, but reciprocity can also serve to protect an individual's initial cooperative offer from exploitation by a free rider. Similarly, showing civil courage to protect an individual from harm, or fighting for one's country of upbringing against intruders, does not indicate a predominantly group or individual orientation of cooperation. We are thus reluctant to attribute reciprocity and heroism a priory and primarily to either binding or individualizing. Nonetheless, as indicated, we have a hunch and presume that these domains may correspond to a general disposition of cooperation. Romano and colleagues (2022) argue that we are "equipped with

complex *reciprocity psychology* that evolved to evaluate, enforce, and condition (…) social behavior on present and/or future opportunities to gain either direct or indirect personal benefits" (p. 254). In addition, they summarize state-of-the-art evidence that speaks in support for the notion that reciprocity can explain pro-social behavior among groups and individuals. Essentially, they assume that reciprocity is associated with three psychological mechanisms, namely concern for reputation, expectations of other people's (prosocial) behavior and anticipation (of future interactions). These mechanisms in turn are potent in explaining direct and indirect reciprocity. The former form of reciprocity describes a tendency (e.g.) for an actor to help those who have helped them. Indirect reciprocity, in contrast, is divided in down-stream and up-stream reciprocity. In regard to helping behavior down-stream reciprocity describes a tendency to help those who have helped another person, an in-group member for example, in the past, while upstream reciprocity refers to a tendency for an actor who has received pro-social behavior (e.g. help) to be more inclined to help someone else. Reciprocity and its associated prosocial psychological mechanisms are also thought to affect group membership, gossip, and third-party punishment, which are known to be strong determinants of prosocial behavior "among individuals and groups" (Romano et al., 2022, p. 255). Thus, based on the current state of science, we presume that reciprocity as a moral domain effectively regulates selfishness to promote group and individual cooperation (Axelrod, 1986; Henrich & Muthukrishna, 2021; Muthukrishna, 2021; Romano et al., 2022). Therefore, reciprocity may best be understood as a general disposition for cooperation and against selfish drives. Furthermore, it can be argued that *heroism* is in turn related to our general capacity for empathy (De Waal, 2008). In terms of empathy, we can argue that this human capacity enables us to empathize with the actual or anticipated suffering of individuals as well as the suffering of social entities such as groups. Empathy leads us to mirror the harm or suffering of either entity in ourselves, which may motivate a tendency to act heroically to alleviate suffering. Although there is a lack of comprehensive reviews in the literature, older and more recent findings show that heroism can be both individual- and group-oriented. In general, evidence of self-sacrifice in war, organ donation to strangers, or high-stakes altruism in everyday emergencies points to the role of empathy, lower risk aversion, and resilience to psychological distress as factors supporting these heroic deeds (Rusch, 2022). Based on this line of thought, we presume that heroism is most likely best understood as a tendency directed toward both the actual and anticipated alleviation of suffering for an individual or a social group. Therefore, we think it is reasonable to assume that heroism may fall under a general disposition for cooperation factor rather than binding or individualizing morality. However, our reflections on reciprocity and

heroism should not be interpreted as unambiguous, theoretically validated convictions on our part. Rather, the point here is to share our doubts about the clear assignability of heroism and reciprocity to higher-order moral constructs.[17] In addition to the possibility of a third higher-order moral construct, it would also be possible to adopt a further position. For example, it is conceivable that binding and individualizing morality are formed from the same basic dimensions across cultures — domains of binding morality are family, in-group and deference, while individualizing morality is consistently based on fairness, trustworthiness and property. However, in addition to these universal compositional elements, the higher order moral constructs could also contain a culture-specific compositional element. The latter element could mean for instance the context-mediated assignment of reciprocity and heroism to either binding or individualizing, or both domains to only one of the higher order moral constructs. In such a case, binding and individualizing would be consistently based on their inherent three dimensions, but reciprocity and/or heroism could be added as a context-mediated, culturally specific element to the respective overarching moral construct. The theoretical space of possibilities is open at this point. Consequently, empirical insights are needed to provide an evidence-based direction to the debate on the assignment of heroism and reciprocity to higher-order moral constructs. In summary, while we suspect a third higher-order factor besides binding and individualizing, we are cautious overall and suggest an exploratory approach to give this possibility an evidence-based direction of reasoning.

Taken together, our MFT and MaC coalescing approach adopts Haidt´s idea (2008) of higher order moral constructs. Although we adopt this idea, we propose a slightly modified composition of higher-order moral constructs in our approach. We suggest to treat family, deference, and in-group as core dimensions of *binding morality*, and fairness, property, and trustworthiness as core dimensions of *individualizing morality*. Regarding to the scope of pro-sociality we furthermore postulate a rather impartial/universal realm for individualizing and a rather particularistic realm for binding. With respect to *heroism* and *reciprocity*, we are undecided at the moment. We await further empirical evidence to draw a picture that allows us to assign these domains to either binding, individualizing, to both, or to an overarching construct of morality that resembles a *general disposition of cooperation*. Lastly, we hold that in every society there are both individual- and group-oriented efforts to cooperate. Consequently, a predominant focus on binding and individualizing may only be relative. We

---

[17] This discussion can certainly be extended to all moral domains to a certain degree, and a fundamental distinction between the group and the individual focus is to be avoided in this regard. However, we believe that the assignment conflict between either group or individual focus is most pronounced in the domains of heroism and reciprocity.

suspect yet that there are cross-cultural differences in the prioritization of binding or individualizing morality — the evolved and historically developed requirements of a society can differ and thus also the social needs and affordances for binding or individualizing in different cultures.

1.7.9. Theoretical Contributions

So, in a nutshell, what is our theoretical contribution? We first propose a broader coverage of moral domains than MFT and MaC. By including trustworthiness and proposing a total of 8 moral domains as a non-exhaustive list, our approach embraces and expands moral pluralism. Furthermore, we center morality around conformity and deviance, and refer in this way to care and harm. Thus, unlike MFT, we do not see care/harm as an independent moral domain, but as the guiding principle that permeates our entire moral mind, and therefore resonates in every moral domain. For what does our evolved moral mind focus on? It is focused on the actual or imagined behavior of people, including ourselves, and at the same time on the direction in which the pendulum of valuation of moral action swings. Self-regulation and external social regulation of cooperative action is the consequence of this moral valuation. This is where the intrapsychic, for example the moral emotional, and the social means, punishment, reputation, and signaling, which we have described above, come into play. We have also outlined the background to this valuation: In a social world crossed by cooperative interdependence, in which this very cooperative interdependence is fitness relevant (Tomasello & Vaish, 2013; Kurzban et al., 2015; Henrich, 2020; Henrich & Muthukrishna, 2021), morally deviant behavior means harm for at least one party in an initially cooperative venture. Given that the success of cooperation has a positive effect on the fitness of the actors, deviance and harm result in costs (of reduced fitness) for the harmed party. Here we bring the findings of Baumeister and colleagues (2001) into play — bad is stronger than good —, and postulate that our moral mind has evolved in such a way that acts of moral deviance have a special significance that outweighs that of moral conformity. According to our proposition, morally deviant behavior causes the pendulum of moral valuations to swing more strongly than its counterpart. However, we do not lose sight of moral conformity as the other side of the coin of the guiding principle of morality that we propose. Given the aforementioned cooperative interdependence, moral conformity results from our perspective in care; successful cooperation, which is understood as non-zero-sum interaction, is a gain of any kind (of tangible or intangible resource) for the morally conforming actors. Our selfish side finds itself regulated and we act to care for each other when

we conform to morals. Further above we have referred to the linguistic bridge to the various moral domains in the context of the word care as an example to highlight this point. Eventually, given our perspective, one might think that our moral approach builds a bridge between the Moral Foundations Theory and the Morality as Cooperation Theory. In our view, however, it is likely more correct to emphasize that we are building a small roof on top of the already existing bridge between these two theoretical approaches. What we are trying to say here metaphorically is the following: When we refer to the theoretical pillars of MFT and MaC, we already see a fundamental intertwining between these approaches. We use this theoretical foundation that both theories share (the bridge) — cumulative cultural evolution (Mesoudi & Thornton, 2018) is so aptly reflected in science, as new buildings of thought rest on the shoulders of giants — to build on it by placing deviance/conformity, harm/care at the center of morality (the roof). In the tradition of Jonathan Haidt (2008), we also consider the possibility of higher-order moral constructs, because we see the parsimonious potential of this idea. At the same time, however, it remains empirically open whether this potential is actually present, or whether constructs such as binding and individualizing morality may oversimplify cultural realities. Finally, we would like to emphasize once again that MFT and MaC are representatives of the universalistic moral perspective that can be derived from a gene-culture-coevolutionary theory, while at the same time emphasizing cultural variability. As explained, we build on MFT and MaC. Consequently, our approach is also anchored in the tradition of the perspective that assumes moral universalism while recognizing cultural variability in the extent of adherence to and importance of individual moral domains. Taken together, our theoretical contribution thus comprises four main aspects and is as follows: **1)** coalescing MFT and MaC, also with reference to higher-order moral constructs; **2)** expanding the (non-exhaustive list of) moral domains; **3)** emphasizing a guiding principle of morality: deviance/conformity; harm/care; **4)** accentuating the weighting of deviance/conformity in the moral mind, with a higher weight on the side of deviance. Next, we turn to our thoughts on how we might empirically capture moral deviance.

1.8. MFT and MaC Coalescent — Measures of Moral Deviance

At the very beginning of this article, we described several moral breaches and formulated the question of whether these morally deviant acts are considered equally *relevant* regardless of (cultural) context. Against the proposed theoretical background, we developed several instruments that shall allow researchers to investigate the above as well as other questions about the human moral mind. In what follows, we will briefly introduce three research instruments:

the *Morality as Cooperation—Deviance Relevance Scale* (MaC-DRS) and the *Moral Deviance Factorial Survey* (MDFS) are presented before we discuss what distinguishes these measures from each other and how they nevertheless complement one another. In addition, we will also introduce a series of 9 *moral dilemma scenarios* that we developed. In the further course of this thesis, all three instruments are used in four cross-cultural investigations and the psychometric properties of the scale we propose are thoroughly examined.

1.8.1. Morality as Cooperation—Deviance Relevance Scale (MaC-DRS)

A number of instruments for measuring moral pluralism arose from the frameworks of MFT and MaC (Graham et al., 2011; Clifford et al., 2015; Curry et al., 2019b; Atari et al., 2022a). The self-report instruments of MFT and MaC, in particular the Moral Foundations Questionnaire 1 (MFQ-1; Graham et al., 2011), have become the leading and most frequently used scales in empirical research on moral pluralism over the last decade. However, MFQ-1 suffers from serious psychometric difficulties (Curry et al., 2019b; Iurino & Saucier, 2020) and the recently improved version (MFQ-2) does not capture moral *relevance* anymore (Atari et al., 2022a). In contrast to MFQ-1, the Morality as Cooperation Questionnaire (MaC-Q; Curry et al., 2019b) comes with good psychometric properties, and measures moral judgment and relevance. Nevertheless, it should be noted that moral constructs of a higher-order have neither been theoretically postulated by MaC nor can they be justified on the basis of empirical evidence so far. Although our theoretical approach follows the main positions and clearly stands on the shoulders of MFT and MaC, we have also provided appropriate conceptual criticism of both theories, particularly with regard to the lack of distinction between *moral deviance* and *conformity*. These lines of criticism apply to the respective measurement instruments of the Moral Foundations Theory and the Morality as Cooperation Theory as well.

The MFT and MaC coalescing perspective that we suggest in this thesis postulates a single guiding principle of morality (i.e., the dualism of moral conformity/care and moral deviance/harm), a broader moral domain coverage then existed before (8 moral domains), and a re-conceptualization of higher order moral constructs (i.e., binding morality, individualizing morality, and likely a general disposition of cooperation). Against this background we developed the *Morality as Cooperation—Deviance Relevance Scale* (MaC-DRS) as a new instrument to capture moral pluralism. The self-report instrument is designed to examine individual *relevance* valuations of *morally deviant behavior*. Assessing the relevance of moral

conformity is thus beyond the scope of this scale. We developed the new instrument to cover valuations of deviant acts in eight moral domains. Moral tendencies in the domains of *fairness*, *trustworthiness*, *property*, *family*, *(in-)group*, *deference*, *reciprocity*, and *heroism* shall be made empirically measurable by MaC-DRS. It will further be an empirical question whether our scale is able to capture aggregated higher-order moral constructs (second-order factors) that are based on the 8 domains proposed. What is more, MaC-DRS is designed as a decontextualized and generalized measure of moral (deviance relevance) tendencies. In other words, we have not set up the instrument to capture *specific* valuations of individual acts of moral deviance, but rather domain-specific and, with regard to the respective moral domain, *general* tendencies in deviance valuation. For the reason that we focus on accurately distinguishing between moral domains while still capturing general tendencies, the construct breadth of MaC-DRS domains was given minor priority in the construction process of the scale. Accordingly, while our scale is designed to be precise, it is not designed to measure the full breadth of behavioral elements that fall under each moral domain. Our coalescence perspective rests on moral universalism, but postulates the possibility of cross-group differences in moral tendencies. To address this, MaC-DRS is overall meant to provide researchers with a means of empirically capturing similarities and differences in moral tendencies within and between cultural entities.

In our opinion it is generally to question, if self-report instruments are able to examine moral intuitions purely and directly. We hold the view that the *pure* measurement of moral intuitions is not possible with self-report instruments — moral intuitions should rather be measured with much "faster" methods such as the implicit association test (Karpinski & Steinman, 2006; Nosek et al., 2007; Sriram & Greenwald, 2009; Greenwald & Lai, 2020). However, MaC-DRS is not alone in not being able to measure pure intuitions, as the same problem applies to all self-report instruments assessing morality. In contrast to MFQ-1 and MaC-Q, however, our scale provides an explanatory definition of moral relevance in advance to all items in order to avoid increased reflection on the meaning of morality. Additionally, our scale instructs respondents from the outset to base their answers on their gut feeling(s). While MaC-DRS is still likely to be influenced from deliberate moral cognition, we assume that respondents' answers are nevertheless at least initiated by their moral intuitions. A position that is in line with Haidt´s intuitionist conception (Haidt, 2001). In contrast to other instruments, however, we designed MaC-DRS to foreground the tendency of quick, effortless, and uncontrollable initial response behavior. Overall, we do not believe that MaC-DRS captures pure intuitions, but we designed our scale to likely measure the space between moral intuition

and conscious moral cognition and, in the best case, moral intuitions to a greater extent.[18] In a later chapter, we will subject MAC-DRS to a variety of psychometric tests. In this context, we will also go into more detail about the scale, the items, and the process of constructing the scale. Here, we simply wanted to present the core aspects of the scale that we propose, as well as the rational for developing a new moral scale in the first place. In the following, we will move on to a supplementary instrument and introduce the Moral Deviance Factorial Survey.

1.8.2. The Moral Deviance Factorial Survey (MDFS)

You see someone damaging your property. Would you *judge* your observation in the same way if the person causing the damage was a family member of yours or a stranger? Or would the behavior in question be equally *relevant* to you, regardless of whether it was a stranger or a family member? Furthermore, try to put yourself in the following situation: someone publicly insults a respected person. Now please evaluate the following: should the person who insulted someone feel equally *ashamed* or *guilty* if the deviant act damages their own reputation or that of their friends? In addition, please ask yourself whether your evaluation would be the same if the insult had come from a woman or a man. Apart from individual evaluations, the question could also be raised as to whether there are different evaluation patterns for such scenarios in different cultures. MaC-DRS clearly has a shortcoming that is also inherent in the related scales stemming from MFT and MaC: it is not possible without further ado to differentiate ratings of moral judgment, relevance, shame or guilt for different social relationships, different gender and for different reputational consequences with scales. In order to meet these limitations, we have invented the *Moral Deviance Factorial Survey* (MDFS) as accompanying instrument to MaC-DRS.

A *factorial survey* is a special survey method in which one can vary every single detail of an item to causally follow the variations effect on respondents' answers (Hughes & Huby, 2004; Auspurg et al., 2009; Auspurg & Hinz, 2014). Items in a factorial survey design are alike to short stimuli scenarios and are called vignettes. Under ceteris paribus a vignette of a factorial survey only varies in one detail, i.e., the vignette dimension and/or the level of dimension, to all other vignettes of the respective vignette universe. When comparing the ratings among vignettes, the single variation of a vignette component, which is the only element of difference to all other vignettes, makes it possible to assess the causal effect of this variation

---

[18] As it is known that "people's self-reported dispositions and stated intentions may not accurately indicate or predict the moral behavior they display" (Ellemers et al., 2019, p. 335), we strongly encourage future research to examine if MaC-DRS holds a potential in predicting actual moral behavior.

on the particular rating. Different versions of how to create a factorial survey exist (Atzmüller & Steiner, 2010; Knutson et al., 2010; Auspurg & Hinz, 2015; Skilling & Stylianides, 2020). Here we propose with the MDFS a semi-experimental factorial survey design with random allocation of vignettes (without replacement) among respondents. Apparently, our vignettes are exemplary, hypothetical situations.

In line with our coalescence perspective, we created a factorial survey for morally *deviant* behavior, hence the name of the instrument. The vignettes that we created comprise four *dimensions* that each come with varying levels. Vignette-dimensions are: *gender*, *social relationship*, *reputational damage*, and *domain of moral deviance*. Each vignette describes briefly a scenario in which a person behaviorally deviates from one of seven moral domains.[19] Moral deviance towards property, fairness, heroism, deference, reciprocity, loyalty and trustworthiness is assessed by our tool (dimension: *domain of moral deviance*; seven levels). The behavior in question is either committed by a female or a male (dimension: *gender*; two levels). In addition, the passive person in the vignette-scenario(s) who is harmed by the deviant action varies: either a family member, an in-group member, or a stranger is harmed by the morally deviant act (dimension: *social relationship*; three levels). Lastly, four variations of reputational damage caused by the deviant action are part of the scenarios: each vignette comes either with no reputation damage, or damaged family, in-group or own reputation due to the deviant act (dimension: *reputational damage*; four levels). By varying the dimension-levels in the individual scenarios, the factorial survey adds variation of (hypothetical) situations to the measurement of moral deviance. Summing up all possible combinations from the dimensions of our factorial survey, the vignette universe comprises of 168 different vignette-scenarios.[20] ***Table 3***, shown on the next gage, provides a summarizing overview.

---

[19] Because we also vary social relationships (stranger, in-group and family) in the vignettes, the number of moral domains in the factorial survey differs from the number of moral domains in the MaC-DRS.

[20] The vignette-universe of our factorial survey design does not contain any completely implausible combinations.

Table 3: Overview — Moral Deviance Factorial Survey (MDFS)

Vignette Dimensions	Number of Dimension-Levels		Dimension-Levels
Domain of Moral Deviance	7	-	Property; Fairness; Heroism; Deference; Reciprocity; Loyalty; Trustworthiness
Gender	2	-	Female; Male
Social Relationship	3	-	Family member; In-group member; Stranger
Reputational Damage	4	-	No reputation damage; Damaged family reputation; Damaged in-group reputation; Own reputation damaged
Vignette Universe	7 x 2 x 3 x 4	-	Total number of vignettes N = 168

With regard to each vignette, four variables relevant to morality (Haidt, 2001; 2003; Tangney et al., 2007; Tomasello & Vaish, 2013; Henrich & Muthukrishna, 2021; Baumeister, 2022) are surveyed via our factorial design. Respondents are asked to evaluate the respective deviant behavior of each scenario in regard to the degree of *relevance*, *judgment* (wrong to right), and magnitude of self-conscious moral emotion (*shame* and *guilt*) attribution.

We will also discuss the *Moral Deviance Factorial Survey* (MDFS) in more detail in the course of this thesis. As in the context of MaC-DRS, we will then also provide concrete examples of the factorial survey and delve deeper into the structure of the instrument, both from a theoretical and empirical point of view. For now, we wanted only to present the MDFS with its core aspects. In our view, the scale we propose and the factorial survey represent complementary research instruments that shall enable researchers to examine the moral mind of the human being in more detail. In the following, we will discuss the specifics of interpreting both instruments.

1.8.3. Two Complementary Tools — On the Interpretation of MaC-DRS and MDFS

We suggest to regard the Moral Deviance Factorial Survey as complementary to MaC-DRS. The factorial survey is the contextualized and definite counterpart to the measurement of general tendencies captured by our scale. As such, the factorial design integrates a significant component by including situational variations in the measurement of moral deviance (Ellemers et al., 2019). Based on the differences between the two instruments, we assume that MaC-DRS and MDFS measure slightly different aspects of morality.

Why do we propose that our scale and the factorial survey are slightly different and yet complementary? The argument behind our claim goes as follows: first and foremost,

we theorize that each moral domain comprises a domain specific universe of deviant acts, i.e., an inherent set of deviant actions that are related to the particular domain. As we postulate that several deviant actions exist (in each moral domain), we further posit that these actions are themselves not a priori to be assumed as equal in terms of the respective *degree of deviance severity*. To give a somewhat more far-fetched but illustrative example: the stealing of a piece of chocolate is unlikely to be classified as equally severe as the stealing of a purse (full of money and other items), even though both acts constitute breaches of the property moral domain. The scale and the factorial survey that we propose are differently affected from the stated proposition.

On the one hand, *MaC-DRS* captures a general and decontextualized tendency of moral domain-specific deviance relevance across 8 domains. In this context, we would like to emphasize that the domains of the scale are composed of several items, which for their part represent various acts of moral deviance. These items are summed for the domains and formed into a domain-specific index. Accordingly, MaC-DRS does not capture differences between specific acts of deviance, but rather generalized tendencies that emerge from the sum of their parts. The *Moral Deviance Factorial Survey* (MDFS), on the other hand, measures a contextualized and specific deviant action in relation to the respective moral domain of the vignette scenario at hand. It follows that, in contrast to MaC-DRS, that MDFS does not measure a general tendency, but rather captures a specific act of morally deviant behavior for each respective moral domain. Important consequences for the interpretation of the two instruments follow from this logic.

MaC-DRS is designed to capture information about a general tendency of deviance relevance in 8 moral domains. These general relevance tendencies can be meaningfully compared across domains, and also across individuals and social entities. Opposed to this, *MDFS* is designed to collect information on a specific deviant behavior for each moral domain. No meaningful general and direct comparisons of the respective evaluations across moral domains can be made. Consequently, no meaningful hierarchy of the dependent variables of MDFS, i.e., relevance, judgment, shame, and guilt, can be drawn *across* moral domains. However, evaluations for one particular morally deviant action can be examined and meaningfully compared across individuals and social entities. In other words, the particular behavior shown in the respective vignette can be evaluated and compared for the same domain across groups. The Moral Deviance Factorial Survey allows researchers, for instance, to compare how people from group *A)* and people from group *B*) evaluate a male who is stealing property from a stranger with the consequence of reputational damage for his family. Here we

can test if different social groups hold different valuations for one particular action of property deviance. Provided that one tests also valuations towards other moral domains one can, however, not directly infer if people from group *A)* and people from group *B*) hold a specific relevance ranking (hierarchy) of moral domains. The latter is not possible because we cannot a priori assume that different acts of deviance are equal in terms of severity across moral domains. Therefore, since we do not know whether the respective action that we present in the scenario for one particular domain is equally severe as a different deviant action within the scenario of another domain, we cannot meaningfully interpret the differences in the valuation of deviant acts across domains. Nonetheless, a comparison across moral domains can be meaningful in relation to the other dimensions of MDFS, i.e., in relation to gender, reputational damage and social relationship. We can aggregate the individual breaches in the moral domains and compare them in relation to gender, for example. Here we could ask whether moral violations committed by a female or a male are e.g. judged to be equally wrong/good across the different moral domains. A comparison with the other dimensions of our factorial survey and across the moral domains would be possible in this case, since the same actions, and so the same degrees of severity, would be compared across groups.

To summarize: We developed *MaC-DRS* to capture general moral tendencies in different moral domains that, due to their general nature, can be compared across domains in terms of magnitude. A relevance hierarchy across moral domains can be built based on data collected with MaC-DRS. Complementary to this tool, *MDFS* captures valuations of one particular moral breach for each of seven moral domains examined with this tool. We propose that each moral domain comes with an inherent set of domain specific deviant actions that vary in magnitude of severity.[21] Hence, as long as no extensive knowledge about a severity hierarchy of deviant acts within moral domains exists, MDFS offers solely a sensible comparison across groups for one and the same moral breach in one and the same domain. It follows furthermore, that also no generalized claims beyond the respective deviant act portrayed in a vignette should be drawn from data gathered by our factorial survey. However, and in addition, MDFS allows the aggregation of moral domains, and comparisons across groups can be made in relation to the other dimensions of the factorial survey. Taking these limitations into account does, though, not

---

[21] An interesting research gap arises from the argument presented; to our knowledge, a typology of morally deviant (and conform) acts, that is hierarchically ordered by degree of severity is absent in the relevant literature. Furthermore, this gap can be applied to the following questions: Are there severity hierarchies of deviant behavior within moral domains, and if so, are these hierarchies universal or culturally construed? One might also examine, for example, whether the types of deviant behavior and the corresponding severity are relatively independent of the general tendency of the relevance of moral deviance. Future research addressing these questions might reveal interesting patterns within and across moral domains, and thus contribute to our understanding of human morality.

mean that the factorial survey design we have presented is of minor utility; on the contrary! The factorial survey integrates the impact of different social relationships into the assessment of moral evaluations. Thereby this tool gives researchers a way to examine hypotheses in regard to the scope (expansiveness) of morality. One may ask, for instance, is moral deviance judged as equally wrong when the person that is harmed by the act of deviance is a stranger or an in-group member. Also, MDFS allows investigations on effects of gender or the impact of different reputational damages in regard to moral deviance valuations. MDFS additionally allows researchers to capture and compare cognitive (relevance and judgment) and emotional (shame and guilt) moral valuations across groups. What is more, based on the vignette scenarios, which may include several of the factorial survey dimensions, and the fact that four items (relevance, judgment, shame, and guilt) are to be answered for each vignette, we believe that MDFS primarily captures *deliberate* rather than intuitive moral tendencies. In this respect, the instrument also differs from MaC-DRS, which we designed to primarily capture *intuitive* tendencies. Overall, we view MDFS and MaC-DRS as complementary, as the two instruments together can partially offset the limitations of each individual instrument and thus supplement each other. The two instruments discussed in the present section both work with 7-point response formats to allow sufficient fine-tuning of respondent's answers. However, this can also lead to challenges, such as response styles, which do not occur as such in dichotomous either/or response formats. In the following, we supplement our set of instruments for examining the human moral mind with another research tool that we have designed for the subsequent investigations.

1.8.4. Forcing to Decide — Binding vs. Individualizing Moral Dilemma Scenarios

Imagine you are in a social situation and you must decide between two mutually exclusive options. Without providing you further context, would you rather stay at the side of your group or respect the property of a stranger? Imagine another situation in which you again have to choose between alternative options: in a social situation where the consequences depend on your decision, would you rather behave fairly or favor a family member? In general, it is apparent that the structure of some social situations involves an inherent conflict between opportunities if these are mutually exclusive. Such situations are called *dilemma scenarios* and have a long tradition in moral research (Skitka & Conway, 2019; Ellemers et al., 2019). Dilemma situations, such as the trolley problem or Kohlberg's famous Heinz-Dilemma

(Kohlberg, 1973; Kohlberg & Hersh, 1977; Awad et al., 2020), usually describe a specific (hypothetical) scenario and ask respondents to choose between two competing options. To move away from general moral tendencies and towards the assessment of concrete social situations, we went a step further than the factorial design and created a set of (hypothetical) forced-choice moral dilemma scenarios.

With the *moral dilemma scenarios*, we essentially wanted to create another, less abstract and more situationally embedded way to measure deliberate individual preferences for binding or individualizing morality. Additionally, unlike MaC-DRS and the factorial survey, the response options for the scenarios are mutually exclusive and follow a forced choice — choose either option A or option B — instead of a rating scale. We included only moral domains that we consider to belong clearly to the two higher-order moral constructs and *excluded* the domains of *heroism* and *reciprocity* from the construction of this instrument. This resulted in a 3 x 3 design comprising a total of 9 dilemma scenarios, each with a force-choice question at the end. ***Table 4*** illustrates the essential conflict between moral domains in each scenario.

Table 4: Individualizing vs. Binding Morality Dilemma Scenarios

Fairness vs. Family	*Fairness vs. In-group*	*Fairness vs. Deference*
Property vs. Family	*Property vs. In-group*	*Property vs. Deference*
Trustworthiness vs. Family	*Trustworthiness vs. In-group*	*Trustworthiness vs. Deference*

Our scenarios should, as far as possible, resonate with the social world of the respondents in a certain way, because we believe that dilemmas like the trolley dilemma are rarely encountered in reality and for that reason alone are abstract. Hence, we have tried to design the scenarios as realistically as possible (Skilling & Stylianides, 2020) and have not opted for extreme social situations such as the trolley dilemma or other vignettes (Clifford et al., 2015). In addition, we paid attention to the unambiguity of the scenarios and tried to create situations in which only the respective competing moral domains are represented and in which the language is, at best, gender-neutral.

Nonetheless, we must acknowledge some limitations. We were unable to completely filter out possible distorting effects of social relationships and gender, as this would only be possible with grammatical difficulties and a considerable loss of realism in the scenarios. It should also be noted that the dilemma situations are similar to the factorial survey in that the scenarios describe social situations of *specific* moral actions. We have pointed out that it cannot be expected a priori that moral actions within and between domains are equally "severe". The dilemma scenarios therefore suffer from the possibility that the respective

binding/individualizing options themselves may not be equally severe. Furthermore, as with MDFS, we also believe that the dilemma scenarios primarily capture *deliberate* moral tendencies. This assumption is based on the nature of the competing options in the dilemmas and the fact that a decision in favor of conformity with regard to one moral domain always goes hand in hand with moral failure and deviance with regard to the other (Haidt, 2001; Tessman, 2014). In line, we hold that people need a certain involvement of deliberate calculus to reach at a trade-off that solves the inherent dilemma of the scenarios.

So, how do the dilemma scenarios complement MaC-DRS and MDFS? We believe the answer lies mainly in the *forced choice* nature of the instrument, but also in the concrete, non-abstract *situational embeddedness* of the dilemma scenarios. First, unlike MaC-DRS and MDFS, the dilemma scenarios do not provide an opportunity for neutrality; respondents are forced to decide and must choose only one from two competing options. The forced choice nature of the dilemma scenarios is designed to ultimately reveal a clear preference order for one moral domain over the other in given situation. Second, the scenarios deal with concrete social situations, which we have designed in such a way that they bring certain moral intuitions into a conflictual confrontation. Haidt's Social Intuitionist Model of Moral Judgment (2001) argues that in a case of conflicting moral intuitions, private reflection (link 6) can occur. Here the moral judgment is made on the basis of either the intuition that is experienced most strongly. Or, alternatively, on the basis of the deliberate choice of an intuition from the experienced set of intuitions. Insofar as the judgment in the dilemma scenarios is not made from the outset on the basis of the strongest available intuition and respondents reflect on the scenarios, which we assume to be the most likely case, we argue that a respective cultural framing is nevertheless evident in the forced choice. The intuitive response to the dilemma scenarios should reflect the cultural framing, whereas the deliberate choice from the dichotomous response option should be shaped by the moral normative orientation of the social environment. In any case, the social directness of the description and the situational contextualization, which are immanent in the scenarios, should elicit a socio-culturally embedded reaction within the respondents. The situational contextualization of the dilemma scenarios is even more pronounced than in MDFS. We are therefore moving from a decontextualized measurement instrument (MaC-DRS), to a more contextualized and varied measurement instrument (MDFS), and finally to a highly contextualized and very specific measurement instrument, the dilemma scenarios. Overall, we believe that the rich and concrete social embedding, together with the forced choice nature of the design, adds another facet to the other research tools we have proposed. In sum, the moral

dilemma scenarios provide a complementary means of capturing potential cross-group/cultural differences in the importance of endorsing binding over individualizing morality and vice versa.

The aim of the thesis is, in addition to the theoretical discourse, the empirical investigation of the human moral mind across cultures. In the light of our MFT and MaC coalescent perspective we propose three research instruments — the *Morality as Cooperation—Deviance Relevance Scale* (MaC-DRS), the *Moral Deviance Factorial Survey* (MDFS), and a set of 9 *moral dilemma scenarios* — as means to capture different facets of human morality. We readily admit that each instrument comes with limitations, yet hold at the same time that these tools can work in particular as mutually supporting means of moral research. Altogether, we are confident that the proposed instruments will enable us to gain deeper insights into the universal moral mind of humankind and its cultural makeup. From the tour of the instruments that we propose for the investigation of the human moral mind, we now turn away from the research tools and discuss in the following concrete examples of cross-cultural variations in human morality before turning to research gaps and the hypotheses to be investigated.

1.9. Morality and Culture

For hundred thousand of years, human groups have spread across the earth, eventually colonizing all the natural habitats this world offers (Henrich, 2016). As a response to different natural conditions, different cultural ecologies formed as strategy of adaptation. Culture is the biological niche of humans, that provides the necessary means for human survival and reproduction. Over the course of human history, natural ecologies were shaped by culture; cultural ecologies, in turn, exert further selection pressure and cause human adaptation to cultural environments (Whiten et al., 2017; Mesoudi & Thornton, 2018; Brown et al., 2022). Every human sociocultural system is originally confronted with the fundamentally same challenges of survival and reproduction. The universalism of the psychological concepts of the self and the moral mind illustrates this exemplarily for our species. However, variant ecological demands, corresponding cultural responses, and subsequent historical path dependencies also play an important role in the understanding of the particular configuration of societal structures and the human psychological apparatus (Henrich et al., 2008; Esser, 2010; Chudek et al., 2016; Henrich, 2020; Muthukrishna et al., 2021; Kitayama & Salvador, 2024). In relation to the self, this has already been presented, and in the following we turn to cross-cultural variances in human morality. It has been shown that at least fairness, property, reciprocity, heroism, family, (in)-group and deference, as proposed by MaC, are perceived as good across cultures (Curry et

al., 2019a). Nevertheless, this does not mean that these domains are also cross-culturally of the same *relevance* (Curry, 2016). Our approach that integrates MFT and MaC has its theoretical roots in the idea that the characteristic configuration of moral systems emerges on the basis of the requirements of the respective sociocultural environment. Moreover, we expect the existence of cross-cultural patterns of individualizing and binding morality, which might be reflected in rather particularistic or rather impartial (universalistic) moral tendencies. Against this backdrop, the following we will mainly highlight cultural dimensions that deal with *cooperation* and differences in *moral expansiveness*.

To receive sense and meaning from the comparison of cultural entities, one must compare them on a common ground. Cultural dimensions provide such a common basis (Barmeyer, 2010). We have already mentioned elsewhere that the cultural dimension of *collectivism-individualism* is found to have an influence on the shaping of our moral tendencies (Triandis, 2001; Żemojtel-Piotrowska, & Piotrowski, 2023). However, another important cultural dimension of relative differences that is tied to cooperation is *kin-ship intensity* (Henrich, 2020). This dimension comprises having either extended kinship ties and institutions on the one end of the continuum or loose family ties and nuclear family supporting structures on the other (Henrich & Muthukrishna, 2021). It is established that the emergence of religion and differences in religious belief systems (nonpunitive vs. punitive deity(ies)) affect particularistic and impartial tendencies in moral behavior (Purzycki et al., 2018; Lang et al., 2019). However, religious doctrines also affected given kinship structures; cross-cultural differences in kinship intensity resulted from, among other things, different religious family policies (Bahrami-Rad et al., 2022). Historical exposure to Roman Catholic Church family policies significantly reduced kinship intensity while supporting the nuclear family (Schulz et al., 2019), a process underscoring the impact of cultural path dependencies on contemporary psychological configurations (Mahoney, 2000; Henrich, 2020; Muthukrishna et al., 2021). Structures of tight kinship relations "incentivize the cultivation of greater conformity, obedience, nepotism, deference to elders, holistic-relational awareness, and in-group loyalty but discourage individualism, independence, and analytical thinking" (Schulz et al., 2019, p. 1). In addition, kinship intensity has been shown to be associated with strict liability in moral judgment across cultures. Evidence suggests that mental states (intentions) play a lesser role in the judgment of harm in cultures with strong extended kinship ties, whereas WEIRD cultures[22] with lower kinship intensity place greater importance on intentions in moral judgment (Curtin

---

[22] Remember: WEIRD is an anacronym coined by Joseph Henrich and colleagues (2010a) and stands for Western, Educated, Industrialized, Rich and Democratic countries (Henrich, 2020).

et al., 2020). This is explained by the argument that strict liability can reduce the potential for conflict in societies with strong and hard-to-break social ties. Thus, intensive kinship promotes people's psychology to attune to rather collectivistic demands of their social ecology (Schulz et al., 2019), while decreased or loose kin-ship intensity seems to affect the moral pendulum to swing into the direction of rather impartial, individualizing morality (Haidt, 2008; Enke, 2019).

Furthermore, historical *exposure to pathogens* does not only affect our cuisine and food preferences, but also our contemporary morality to promote, among others, higher in-group loyalty (Sherman & Billing, 1999; Murray & Schaller, 2010; Atari et al., 2022b). More so, pathogen exposure, exposure to natural and social threats, and differences in subsistence styles (herding vs. farming; wheat vs. rice farming) impacted on cultural entities relational mobility, and likely on impartial and particularistic tendencies. Cultures differ in regard to how stable and fixed and how open and pervasive networks and structures of social relations are, i.e., they differ in low vs. high *relational mobility*. This cultural dimension is further associated with diverging effects on psychological tendencies: high relational mobility is for instance positively correlated with trust in strangers and political rights, and negatively with hierarchy (Thomson et al., 2018). Also, direct effects of relational mobility on decision making in moral dilemmas are recorded; inspecting three types of trolley dilemmas across 70 countries Awad and colleagues (2020) found that "low relational mobility (…) is strongly associated with the rejection of sacrifices for the greater good "(p. 2332).

Apart from their impact on relational mobility, it is known that historically evolved *subsistence styles* have also influenced and shaped social organization and psychological tendencies over millennia, especially those related to cooperation (Henrich, 2020). In particular, the period after the agricultural revolution, i.e., approximately the past 10 thousand years, has had a remarkable impact on societal organization and structuring (Kitayama & Salvador, 2024). Different subsistence styles required different forms of cooperation to ensure sufficient food supply and thus survival. Compared to wheat cultivation and pastoralism, rice farming required significantly more sharing of labor in addition to coordination of irrigation. In essence, the cultivation of paddy rice welded people together into tight social networks characterized by interdependence. Evidence suggests that people from (historical) rice *farming* groups tend towards interdependence, tighter social norms, more holistic cognition, respect to parents, pronounced loyalty to friends, and a marked distinction between friends and strangers (Talhelm, 2022).

Furthermore, a recent theoretical and measurement (MFQ-2) refinement of Moral Foundations Theory (Atari et al., 2022a) was able to reveal cross-country differences in *moral*

judgments in six posited moral dimensions. This updated MFT version shows cultural as well as inter-religious (Christianity, Islam, and nonreligious) differences in the centrality of agreeing to statements related to caring, equality, proportionality, loyalty, authority, and purity.[23] Regarding moral judgments a cross-cultural pattern of the binding vs. individualizing approach to morality could not be replicated by Atari and colleagues. Hence, for the moment Atari et al., (2022a) propose that "the individualizing-binding distinction may not be how moral foundations are organized universally; rather the inter-relations between the foundations should be considered culture-dependent" (p. 38). Since we take a different position, we see reason for further research to give this discussion a consistent direction.

In a field experiment, Cohn et al., (2019) intentionally "lost" wallets with varying amounts of money in them and examined cross-cultural return rates. They found that return rates vary substantially across countries, and that they increase with the amount of money in the wallet. In addition, behavioral measures of wallet returns are associated with differences in *generalized trust* across cultures as measured by surveys (Bjørnskov, 2021). Although Cohn and colleagues (2019) interpret their results as differences in honesty and altruistic concern, the results can also be viewed more broadly as demonstrating morality at work and illustrating cross-cultural differences in the moral domains of property and trustworthiness.

Regarding institutions, the dimension *market integration* is known to be influential when it comes to variations in our moral behavior. Institutions supporting higher breath and intensity of market exchange are found to support differences in impartiality, fairness and trust in strangers (Henrich et al., 2005; 2010b). So, increased market integration, compared to lower market integration, seems to promote general impartial tendencies and moral domains associated with individualizing morality (and vice versa) (Henrich, 2020).

Research on *values* further points to cross-country differences that can be associated with diverging priorities of various moral areas (Minkov & Kaasa, 2022). From a theoretical perspective, variations in values can be linked to country-specific differences in socioeconomic development.[24] Cross-country variation in values can be found, among others, in the importance

---

[23] The MFQ-2 self-report instrument (Atari et al., 2022a) represents a long-awaited improvement to the frequently used, leading MFQ-1 scale (Graham et al., 2011). However, the MFQ-2 measures only moral judgments and omits the measurement of moral *relevance*, a circumstance touched on elsewhere in this thesis in the context of the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS). Nevertheless, at this point it should be noted that our MFT and MaC coalescing perspective on morality, which places cooperation at the center of morality, proposes (at least) 8 moral domains and a guiding principle of morality, rests partly on the shoulders of MFT but also takes different theoretical paths in parts, as has been explained.

[24] https://www.worldvaluessurvey.org/WVSContents.jsp?CMSID=findings&CMSID=findings

of family, equal gender rights, trust in strangers, the justifiability of stealing property and obedience.[25]

However, there is not only *variation* between countries, but differences in morality exist also *within cultural entities* (Iyer et al., 2012; Graham et al., 2016). Several scholars examined *moral expansiveness* (Waytz et al., 2019), a phenomenon understood as the "metaphorical boundary drawn around the entities we believe do and do not deserve our moral concern" (Kirkland et al., 2023, p. 305). Moral expansiveness thus deals with how far reaching, in social terms, our morality is, and variation in regard to this variable was found within and between countries. Among other findings Kirkland et al., (2023) were able to demonstrate an "indirect effect of perceived wealth gap (between-countries) via generalized trust on moral expansiveness" (p. 310). Also, Waytz and colleagues (2019) were able to show differences in moral expansiveness in regard to political ideology: conservatives tend to show a rather particularistic (tight) moral realm, while liberals tend to show a rather (loose) universalistic moral realm.

Overall, moral differences can be observed within, yet also substantially between cultural entities.[26] The empirical evidence referred to shows that patterns of moral variation are related to different requirements, i.e., needs and affordances of the respective sociocultural context. These requirements themselves are likely to be associated with recurring challenges that societies have faced throughout history (Curry, 2016). Thus, as posited and advocated by researchers coming from the stance of cultural evolutionary theory, a functionalist difference appears to underlie these cross-cultural variations (Kitayama & Salvador, 2024). Moreover, the available evidence seems to point to cross-cultural differences in the endorsement of a relatively *particularistic* or a relatively *impartial* (universalistic) moral system. Furthermore, the latest MFT research contains data indicating to interpret the relations between moral domains as a rather intra-cultural phenomenon — cross-cultural patterns of binding and individualizing morality could not be replicated by Atari et al., (2022a). However, we believe that the discussion of cross-cultural patterns of binding and individualizing tendencies should not yet be abandoned. We base our suggestion for the revival of binding/individualizing on our outlined theory and on the presented findings about the cross-cultural differences in group-oriented and individual-oriented tendencies in morality. Moreover, we see a striking theoretical similarity between culturally distinct forms of interdependent/independent self-construal and

---

[25] Within the **Appendix** we show various descriptive analyses of the World Value Survey data (Haerpfer et al., 2022) that support our statement.

[26] The differences we have listed are by no means an exhaustive list; no such claim is made here.

binding/individualizing morality. In the following, we will elaborate on the relationship between selfhood and morality and, by drawing on the Model of Sociological Explanation, offer an extension to Model 1 that we presented under the section of the self. Eventually, we will lead over to research gaps and hypotheses.

1.10. Research Model, Research Gaps, and Hypotheses

1.10.1. Research Model — Culture, Self-Construal and Morality in the MSE

In this thesis we posit an MFT and MaC coalescing approach to morality. This approach puts cooperation at the center of morality, proposes (at least) 8 moral domains, an even more nuanced view of moral pluralism than existed before. Additionally, with the dualism of moral conformity and deviance we pose also a single guiding principle of morality that traverses all moral domains. Note, we essentially suggest to view moral conformity as equivalent to caring about others, while we view opposed to that moral deviance as a reflection of harming others. Apparently, we suggest that morality is inherently social; caring about others realizes mutual gain in cooperative interaction and consequently conforming to morals harbors an empathetic concern at the core (Haidt, 2003; De Waal, 2008; Tomasello & Vaish, 2013). Above that, our approach further aims to revive the idea of higher order moral constructs (Haidt, 2008) and posits a partly revised binding and individualizing morality understanding and composition. As for binding and individualizing morality, we see the latter as expressing rather impartial, individual oriented moral tendencies and the former as expressing rather particularistic, group oriented moral tendencies. Both approaches to morality are fundamentally social, albeit with different locus of functioning. On the one hand, binding morality regulates egoism by binding people into tight social networks to promote predominantly group focused cooperation. On the other hand, individualizing morality promotes mainly inter-individual cooperation by taming egoistic drives due to individual concern for others. In addition, we introduce the possibility of a general disposition of cooperation that fills the space between binding and individualizing morality and is linked to both.

"The self as moral agent has its roots in the requirements of culture as a large, organized system of cooperation" (Baumeister, 2022, p. 112). Important to our research project is the idea of a functional relationship between the particular sociocultural ecology and the configuration of self-construal and moral system. Notice: the word function stems from Latin "functio", i.e.,

to perform/execute.[27] Viewing a phenomenon from a functionalist perspective therefore implies that the phenomenon in question must fulfill a purpose (that is to be performed in respective sociocultural context). We argue that morality and selfhood serve a double purpose that is ultimately related to individual survival and reproduction yet also to the re-production of individual´s sociocultural ecology. Specifically, this means that our morality as part of the self comes to the fore in the regulation of human cooperation, which in turn is necessary for social life between people, the individual reproduction of the organism (acquisition and exchange of resources), the development of human social organization and the upkeep of social order. In turn, the doing of actors enters into social situations re-producing the sociocultural structures of human´s cultural ecology.

Different requirements, needs and affordances, are inherent in different sociocultural entities for cultures are diverse and encompass different natural (climate, geography e.g.) and sociocultural (cultural dimensions e.g.) conditions. Cumulative cultural evolution and the respective path dependencies of cultural entities are an expression of these conditions. To promote life in context, the particular (dominant) mode of selfhood and the particular (dominant) moral system must be adapted to the prevailing requirements of the sociocultural ecology. Social situations do not happen in a vacuum; rather, they are products of past (inter-)action of actors under given sociocultural requirements and thus socio-culturally (pre-)structured. We have emphasized this notion with the **Model of Sociological Explanation (MSE)** (Esser, 1999; 2002a; 2002b; 2010; Greshoff et al., 2008), which we would like to take up again here. Based on the socio-cultural history of cultural entities, we reason that specific, significant symbols are present in sociocultural structured situations. Social situations are situations in which at least two actors are present and socially (inter-)act, i.e., direct their behavior towards each other, in order to acquire or exchange resources necessary for the satisfaction of basic needs and survival/reproduction ultimately. The MSE assumes that actors', often unconscious and automatic, selection of action in social situations is guided by subjective expected utility and follows the principle of utility maximization. We build on this and assume furthermore that actors are cultural subjects: Individual actors are born into cultural systems of meaning and are socialized into them in the course of their lives. Actors are hence able to process the content of significant symbols in familiar sociocultural situations based on what the social experience of their life-course has imprinted in their culturally constituted mind (Henrich, 2020). This means that individual actors have mental models available that match the symbols present in social situations of everyday (inter-)action. If there is a (perfect) match between

---

[27] See: (https://www.oxfordlearnersdictionaries.com/definition/english/function_1).

situationally present symbols and mental models, the social situation receives a certain meaning for actors; a definition of the situation emerges and actors interpret the situation and act in the situation based on the active framing of the situation. This cultural frame in turn entails specific codes, i.e., superordinate goals of the situation and behavioral scrips, i.e., typical, socially valid actions patterns. We argue that certain logics of the situation entail significant symbols that activate a particular ***self-construal as frame of reference***. It is our self that aligns our cognitions, feelings, motivations and actions to work in concert with the respective context. Furthermore, we advocate that morality has to be considered as an aspect of the human self. Consequentially, we grasp the moral mind as only in analytical terms separated from human self and thus configured in correspondence to our self-construal. Our self and moral system are functional for they align the individual with the surrounding social context. As such they work to support survival in given context and function furthermore to re-produce the social elements needed to uphold human´s sociocultural ecology. Therefore, we expect the configuration of self and morality to be culture-dependent. Thus, contingent on culture, we hypothesize a correspondence between modes of selfhood — i.e., being independent or interdependent — and the endorsement of moral domains — i.e., the relevance of deviance towards moral domains as measured by MaC-DRS.

As mentioned earlier, our functional stance is not to be confused with metaphysical functionalism and the problems associated with it (Brüntrup, 2004). Rather, the functional value of actions that have the potential to maintain a social system and ultimately contribute to the survival/reproduction of human organisms may or may not be a product of actor´s conscious deliberation. In most cases, however, it is the result of automatic, unreflective action by individuals who are unaware of the (functional) consequences of their behavior. We have made this point explicit by introducing the Model of Sociological Explanation (MSE).

Furthermore, to the end of our section on the self, we have integrated the culture contingent construal of selfhood into the MSE and provided a first model. Model 1 illustrates the mutual relation between culture and self-construal, and assumed a special case of the MSE. We would now like to expand this model and integrate morality into it. ***Model 2*** builds on the first model and deals accordingly also with the MSE special case of Model 1. We pick up what we stated above and summarize the MSE special case of our model as follows: actors in a social situation are cultural subjects. Significant symbols and self-relevant information are present in the social situation and available to the actors due to their cultural socialization. As a simplifying heuristic we assume a perfect match between the mental models of actors and present significant symbols. Consequently, action selection and action itself are guided by actor´s automatic,

unreflecting mode of information processing. The latter is in line with the intuitive approach to morality advocated by the Moral Foundations Theory (Haidt, 2001; Haidt & Joseph, 2007; Tutić, 2023). Derived from the significant symbols present in the situation, the respective (culturally constituted) self-construal frame of reference is thus activated and defines the situation. The situation-specific code (the overarching goal of the situation, i.e., independence or interdependence) and the script of action (socially accepted and situationally valid behavior that is aligned to the prevailing code) also emerge from present significant symbols and the active self-construal frame. The MSE now requires a description of the situation in order to eventually derive the logic of the situation and the subsequent logic of selection and aggregation. By describing the situation and deriving the three subsequent logics, we now have a model that allows us to understand what is by explaining why it is (Greshoff, 2008). In our project, the objects of analysis are the correspondence between selfhood and moral relevance (independence - individualizing; interdependence - binding), cultural differences in moral relevance (binding, individualizing, mixed morality), differences in moral impartiality and particularism, and moral universalism (we await to find the structure of 8 moral domains across cultures). In terms of why, we refer to cumulative cultural evolution, path dependencies and the reproduction of social structures through the mutual constitution of micro- and macro-levels as demonstrated by the MSE. On the basis of what we have elaborated so far, we are now in a position to approach culture-specific logics of the situation and to formulate conditions that characterize a cultural context as conducive to either a binding or individualizing approach in morality.

Several cultural dimensions take effect in the social focus on in-group life and in-group interaction, these are: collectivism, honor/face cultural logic, high kinship intensity, historical rice farming subsistence style, low relational mobility, high environmental pathogen prevalence and low (historical) market integration (e.g.,).[28] We regard the world as multi-causal, yet

---

[28] We do not claim to present a comprehensive list of cultural dimensions that promote cultural subjects' orientation towards interdependence or independence. Moreover, we use our theoretical considerations as ideal-type heuristics to derive hypotheses. This is not to say, however, that we wish to undermine the cultural diversity at hand. The world is multicausal, and this also applies to cultural configurations such as social constructions of reality. Cultures are diverse in many ways, and working with heuristics is merely an attempt to reduce complexity in the study of our socio-cultural world in order to gain insight into potentially existing patterns of the social. Furthermore, we do not argue that cultural changes do not occur. On the contrary. Such an argument would anyway be contra to the gene-culture co-evolutionary framework that we advocate here. Essentially, without cultural change there would be no cumulative cultural evolution. Nonetheless, "there may be a period of stasis before further modification or improvement" (Mesoudi & Thornton, 2018, p. 3). Admittedly, sociocultural change and reproduction probably overlap in countless ways in a multi-causal world. But in order to be able to grasp and capture something from and in the constant stream of a (causal) process, we are engaging in analysis here and separating the phenomenon of our interest from the simultaneity of other phenomena by means of our description. So, it is precisely the state of reproduction (period of stasis) and not that of change that concerns us here; our heuristic model is only concerned with the re-production of sociocultural structured situation(s). However, as it is evident from the recourse on the

provided the predominance of cultural dimensions that foster in-group focus and ***overall social group orientation***, we expect for respective cultural entities prevalence of the interdependent mode of selfhood and the corresponding binding approach to morality. In other words: we expect a logic of the situation, under given cultural dimensions, in which significant symbols are present that elicit frame activation of an interdependent self-construal and corresponding binding morality. However, we found also several cultural dimensions that take effect in the social focus on individual life and rather impartial and loose interaction among people that are more or less strangers to each other. These dimensions are (e.g.,): individualism, dignity cultural logic, low kinship intensity, historical wheat farming and herding subsistence style, high relational mobility, low environmental pathogen prevalence and high (historical) market integration. Provided, in turn, the predominance of cultural dimensions fostering a focus on ***overall individual-centered social orientation***, we expect for respective cultural entities prevalence of the independent mode of selfhood and the corresponding individualizing approach to morality. In other words: we expect a logic of the situation, under given cultural dimensions, in which significant symbols are present that elicit frame activation of an independent self-construal and corresponding individualizing morality. Based on these *ideal type* different logics of the situation, we summarize our thoughts in ***Model 2*** (***Figure 4***). Furthermore, the group-centered situational logic and the individual-centered situational logic serve as the starting point to infer our hypotheses.

It should further be noted in regard to the model that the investigations that will be described below only aim to directly examine the influence of culture on the configuration of self-construal and moral system, and the links between them. In other words, we do not directly examine that different cultural entities are associated with different cultural requirements. Rather, prior to our data collection, we will try to work out the *empirical* logic of the situation in the countries we include in our investigations by integrating findings from previous studies. The latter are then used as *indirect* associations in the context of the question which moral system guides cooperation in different cultures.

MSE, the model could also be modified to focus on cultural change and the case of cultural ratcheting that is at the core of cumulative cultural evolution. As we yet have set a different focus with our project, we leave out the case of cultural change. Once again, we focus on a model that centers around different logics of the situation, and offers a way to understand and explain cross-cultural differences in selfhood and moral domain relevance against the background of cumulative cultural evolution, path dependencies and adaptation to different (natural and cultural) ecologies.

Figure 4: Model 2 — Self-Construal, Cultural Frame and Frame Corresponding Moral Relevance
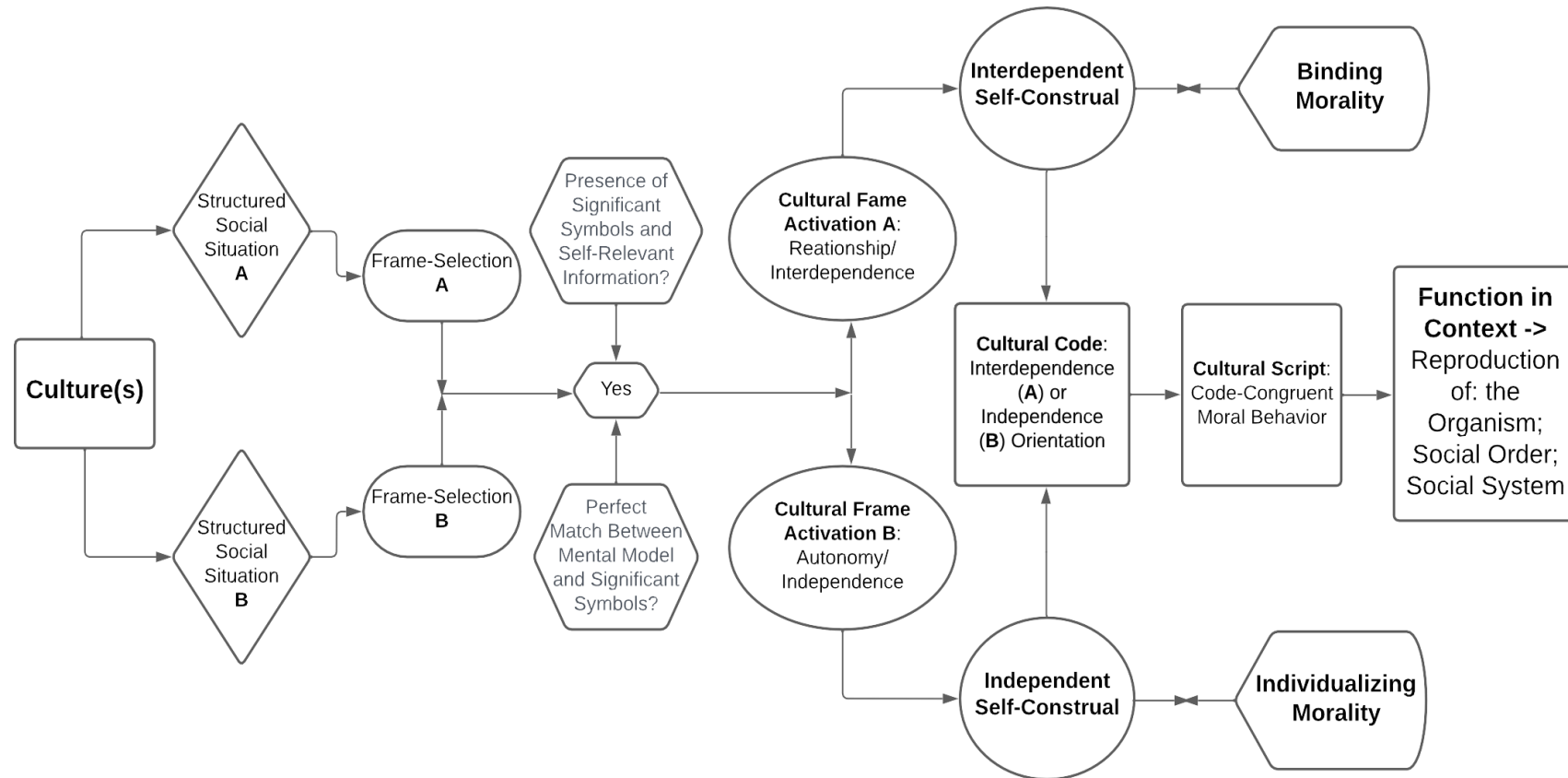
Figure 4 shows **Model 2** and illustrates our integration of self-construal and morality into the Model of Sociological Explanation. **Model 2** builds upon our first model and shows ideal cases of frame selection for cultural subjects provided specific logics of the situation, the presence of significant symbols, cognitive availability of (actor´s) mental models, and perfect match between symbols and mental models. Two structured **social situations** (**A** and **B**) are shown that lead to the activation of two different, self-construal congruent cultural frames. The **logic of the situation** (i.e., the structuredness of the social situation) in **case A** is determined by prevalence of the following cultural dimensions: collectivism, honor/face cultural logic, high kinship intensity, historical rice farming subsistence style, low relational mobility, high environmental pathogen prevalence and low (historical) market integration. The **logic of the situation** in **case B** is determined by prevalence of the following cultural dimensions: individualism, dignity cultural logic, low kinship intensity, historical wheat farming and herding subsistence style, high relational mobility, low environmental pathogen prevalence and high (historical) market integration. These are heuristic, *ideal-type* logics of the situation. A particular logic of the situation of the everyday actions of cultural subjects may be determined by the combination of all the dimensions listed here, as well as by other constellations. The respective logic of the situation must be derived empirically in addition to theoretical assumptions, depending on the case at hand.

According to our theorizing different modes of selfhood become prevalent as (cultural) frames of reference in the sociocultural structured situations A and B. Based on the presence of significant symbols the **cultural frame** of an **interdependent self-construal** is selected in **situation A**, and the **cultural frame** of an **independent self-construal** is selected in **situation B**. In addition, we theorize that the activation of a respective self-construal frame elicits frame congruent patterns of moral relevance. We theorize that an **interdependent self-construal frame** leads to a relatively higher activation of **binding morality** whereas an **independent self-construal frame** leads to a relatively higher activation of **individualizing morality**. Frame A and B further entail particular **codes**, i.e., superordinate goals of the situation, and **scrips**, i.e., code congruent socially approved behaviors. As we focus on morality, we argue that the code of the social situation also specifies what is socially approved in regard to moral behavior. We argue that it is primarily **group-oriented cooperation** that is specified by the active cultural frame in **situation A**. In turn, we argue that in **situation B** it is primarily **individual oriented cooperation** that is specified by the active cultural frame. For the sake of simplicity, we will leave out our notion of a general disposition of cooperation and assume that it is activated (to a certain extent) by both cultural frames. Since we assume a special case of the MSE, Model 2 does not focus on social change. Rather, the model deals with the reproduction of the individual organism within the respective context as well as the reproduction of the social structures of the respective context. However, the model can be expanded to include the aspect of social change, as shown in the section on MSE. But this is not our primary goal, which is also reflected in our research designs, as we shall see in Chapter 2.

1.10.2. Research Gaps — Culture, Selfhood and Morality

On the basis of what we have elaborated we argue that empirical findings on the particular cultural relevance of respective moral domains and also the relative prioritization of either a binding, individualizing or a mixed approach to morality are still missing in the field. To be precise, we do not talk about moral *judgments* (Graham et al., 2009; Atari et al., 2022a) but about moral *relevance* and a lack of cross-cultural findings in this regard. Haidt (2001) defines moral judgments broadly as "evaluations (good vs. bad) of the actions or character of a person that are made with respect to a set of virtues held to be obligatory by a culture or subculture" (p. 1028). Our view is that we can judge certain behaviors as right or wrong, but this does not mean that these behaviors must be relevant to the same extent. Someone may favor themselves in the distribution of goods, while another person may not defend the views of their in-group to outsiders. Both behaviors can be seen as moral violations of different domains, but are these violations cross-culturally of equal relevance? In other words, do these violations immediately strike people as morally wrong, and if so, do people tend to perceive them as very or extremely relevant or rather irrelevant or even less relevant? Fairness and in-group are universally regarded as moral domains (Curry et al., 2019a). However, this does not necessarily mean that they have the same degree of *relevance* across cultures (Curry, 2016), and this is precisely what we want to investigate. Research suggests that the human mind is attuned to its sociocultural ecology (Henrich, 2020). We share this position and therefore one aim of this project is to investigate *cross-cultural differences* of the human moral mind.

Our overarching research question — **which moral system guides cooperation across cultures?** — is a question about the *relative* relevance of moral domains. To our knowledge, research has not yet answered this question. Based on our proposed theoretical approach, asking "*which moral system*" also leads to the question of whether the respective moral system under investigation is a binding, individualizing, or mixed system, or whether we need more precise terms to describe the respective moral system.

Furthermore, we expect that the lack of empirical findings on moral relevance in different cultures is presumably due to a lack of conceptual clarity on the one hand, but primarily to the *absence of a reliable self-report instrument* on the other. For this reason, we propose novel research tools, the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS) together with two supplementary tools, to expand the existing research repertoire in the context of empirical studies on morality.

All human cultures dependent to some degree on inter-individual cooperation. Therefore, albeit looking for cross-cultural differences on the one hand, we are also concerned with *moral universalism* on the other. We predict to find the 8-dimensional structure of moral domains proposed via the MaC-DRS instrument across cultures. Being able to empirically demonstrate the eight moral domains of the Morality as Cooperation—Deviance Relevance Scale across highly heterogeneous cultural entities will provide a strong indication of the domain's universality and hence further insight into the constitution of the human moral mind. Moreover, a cross-cultural validation of our scale would also contribute to researchers' repertoires by filling the gap of a reliable self-report instrument that is designed to assess moral (deviance) *relevance*.

In addition, we see a striking theoretical similarity between the independent self-construal and individualizing morality, and between the interdependent self-construal and binding morality. More generally, we have identified various cultural dimensions that point to either a group-centered focus of social orientation or an individual-centered focus of social orientation. Independence and interdependence in selfhood and individualizing, as well as binding morality, seem to correspond to these overarching social orientations at their core. In line with our theorizing, we expect the *theoretical association between sociocultural contexts, ways of selfhood and moral domain endorsement*, as we laid out in **Model 2** via a recourse on the MSE, to hold also empirically true.

Beyond that, we ask whether cultural entities that promote relatively more binding or individualizing morality may differ from one another in terms of *moral particularism and impartiality*. We hope that these investigations will provide us with additional insights into the cultural constitution of moral systems and a better understanding of whether theoretical binary logics such as binding or individualizing adequately reflect cultural realities. The investigations on moral particularism and impartiality shall also contribute to a deeper comprehension of the sociocultural requirements for human cooperation in different contexts.

Overall, building on the gaps we have identified we suggest a systematic analysis of the relation between culture, self-construal and the relevance of moral domains. We propose several *hypotheses* that are intended to help us to shed light on the research gap(s) we have identified.

1.10.3 Hypotheses

Morality has evolved to promote human cooperation and thus survival and reproduction of humankind in general. Although we prescribe to the view that morality is a universal facet of the human mind, the emphasis on moral domains likely varies depending on contextual requirements and in correspondence with culture contingent self-construal. Cultural systems of morality have a universal core, but they also differ across cultures, as we assume. With regard to the question *Which moral system guides cooperation in different cultures?* we want to examine the broad picture — how does moral relevance unfold across cultures and across different moral domains, and how do intuitive and deliberate tendencies of the cultural valuation of moral breaches look like —, while also taking a closer look at tendencies of moral particularism and impartiality. In addition, we are interested in the relationship between ways of independent and interdependent self-construal and morality. Our model draws on a recourse of the MSE and builds a bridge between culture, self, and moral system. In the following (*Table 5*), we list the hypotheses that we want to test empirically in several consecutive investigations. In these investigations, we will examine and apply the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS), the moral dilemma scenarios, and the Moral Deviance Factorial Survey (MDFS). Our measurement instruments are presented again and discussed in more detail in the corresponding sections of our research. Overall, we will test our hypotheses in four empirical investigations, each of which will be the subject of a chapter in its own right. In the following, we will present our hypothesis before we discuss three underlying data collections, with a focus on a cross-cultural data collection, in more detail in the next chapter. In addition to the data basis, the following chapter will primarily address aspects of cross-cultural research that must be considered before data collection, but also require enormous attention post hoc each data collection. The following chapter is therefore primarily devoted to methodological issues before we can turn to the substantial empirical investigations thereafter. But to see how we will address the identified research gaps, let us now first focus on our research hypotheses.

Table 5: Cross-Cultural Investigations of the Human Moral Mind — Hypotheses

Universality of the human moral mind: MaC-DRS Hypotheses*

A1 - Main hypothesis	**We expect a universal pattern of morality and hypothesize to find 8 MaC-DRS factors of first order across cultures.** In other words: we hypothesize that MaC-DRS consists across cultures of an 8-dimensional factor structure comprising fairness, trustworthiness, property, family, in-group, deference, reciprocity and heroism as moral domains. This hypothesis is tested via EFA, CFA and invariance testing across four cultural groups.
A2 - Main hypothesis	**We hypothesize that higher-order moral constructs exist consistently across cultures.** In other words: moral relevance patterns of binding and individualizing, possibly also a general disposition of cooperation factor, exist universally across cultures and are cross-culturally construed based on the same moral domains. This hypothesis is tested via Confirmatory Factor Analysis.
A3 – Counter hypothesis	**Counter-hypothesis to A2) based on Atari et al., (2022a):** Higher order moral constructs are formed in a culture-dependent way, i.e., binding and individualizing, possibly also a general disposition of cooperation, do not exist consistently across cultures and are construed differently depending on the respective cultural context. This hypothesis is tested via Confirmatory Factor Analysis.

Sub-Hypotheses to A1 and A2

1)	The moral domains of fairness, trustworthiness and property build the dimensions of the higher-order *individualizing* moral construct across cultures .
2)	The moral domains of family, in-group and deference build the dimensions of the higher-order *binding* moral construct across cultures.
3)	The moral domains of heroism and reciprocity built a *general disposition of cooperation* higher-order moral construct across cultures.
4)	The moral domains of heroism and reciprocity fall in between binding and individualizing morality, i.e., they are highly correlated with domains of both higher-order moral constructs.

Hypotheses in the context of cross-cultural moral differences: Moral Intuitions (*MaC-DRS*)

Main Hypothesis Cultural Differences (CD)	Although we predict universalism of the 8 moral domains proposed by MaC-DRS, we also *hypothesize significant*

	differences in (intuitive) moral domain relevance across cultures.
	This hypothesis is tested via pairwise comparison (between samples) of predicted MaC-DRS scores (average marginal effects) after we fit a respective OLS regression for each moral domain examined.
Self-Construal	We expect cross cultural differences and…
Sub-hypothesis: CD 1	*…hypothesize that cultural entities that foster relatively more interdependent ways of selfhood also foster relatively more binding morality* (i.e., they have higher relevance ratings of the family, deference and in-group moral domains).
Sub-hypothesis: CD 2	*… hypothesize that cultural entities that foster relatively more independent ways of selfhood also foster relatively more individualizing morality* (i.e., they have higher relevance of the fairness, trustworthiness and property moral domains).
Cultural logics	
Sub-hypothesis: CD 3a	We hypothesize that cultures of *honor* and *face* are significantly higher in *binding morality* than cultures of dignity.
Sub-hypothesis: CD 3b	Due to prevailing honor logic and self-assertive interdependence in self-construal, we predict that Egypt, however, scores higher on individualizing domains than Japan.
Sub-hypothesis: CD 4	We hypothesize that cultures of *dignity* are significantly higher in *individualizing morality* than cultures of cultures of honor and face.

Hypotheses in the context of cross-cultural moral differences: Deliberate Moral Cognition (*Moral Dilemma Scenarios*)

Hypothesis A	We hypothesize that an *interdependent*, *group-focused orientation* is associated with a greater deliberate importance of *binding morality*. In our study, in particular the **JP-sample** and **EG-sample** should be characterized by an *interdependent*, *group-focused orientation* overall social orientation. Hence, we expect a tendency of deliberate choices towards *binding* morality (as measured via moral dilemma scenarios) for these cultural entities.
Hypothesis B	We hypothesize that an *independent*, *individual-focused orientation* is associated with a greater deliberate importance of *individualizing morality*. In our study, in particular the **GER-sample** and **US-sample** should be characterized by an *independent*, *individual-focused orientation* overall social orientation. Hence, we expect a tendency of deliberate choices towards *individualizing* morality (as measured via moral dilemma scenarios) for these cultural entities.

Both hypotheses are tested via between sample pairwise comparison of predicted MaC-DRS scores (average marginal effects) after we fit a respective logistic regression for each moral domain tested. |

The Relevance/Judgment Hypothesis and the Impartiality/Particularism Hypothesis (*Moral Deviance Factorial Survey*)	
Relevance/Judgment **Hypothesis**	We predict that the extent of the *relevance* of specific acts of moral deviance and the extent of the *judgment* about specific acts of moral deviance do not (necessarily) coincide. This hypothesis is tested via graphical analyses of country/sample specific average marginal effects for moral deviance relevance and moral deviance judgment across all moral domains examined.
Impartiality/Particularism **Hypothesis**	We hypothesize that cultures favoring *binding* moral domains (family, deference, and in-group) over *individualizing* moral domains (fairness, trustworthiness, and property) tend to rate moral deviance that harms a stranger as less severe (relevance, judgment, shame, and guilt) than deviance towards a member of one's in-group or family (and vice versa). This hypothesis is tested via the inspection of country/sample specific interaction effects (average marginal effects after we fit respective OLS regression models) for the three levels of the Moral Deviance Factorial Survey dimension social relationship (levels: family, in-group, stranger) on the dependent variables: deviance relevance, deviance judgment, deviance shame attribution and deviance guilt attribution. The hypothesis will be tested across all moral domains examined.

* We would like to mention at this point that we have indeed decided in favor of a hypothesis-testing approach for the most part, but we are still open to what may emerge from the data in an exploratory way.

Chapter 2: Setting the Stage to Cross-Cultural Investigations of the Human Moral Mind

2.1. Three Primary Data Collections

In order to pursue our research goals, we have conducted three primary data collections. In the course of this work, we will deal with the first two of them solely in the context of the development of the new morality scale (MaC-DRS), which we propose. **Data collection 1)** and **2)** are based exclusively on independent German samples, which is why they are not directly relevant for the overarching cross-cultural research project that we address mainly.

In the further course of this text, we will therefore focus primarily on **data collection 3)**, in which we collected data across four different countries that serve as proxies for different cultures (Smith, 2014; Minkov et al., 2021).[29] The cross-cultural design of the third data collection allows us to use the corresponding data to pursue our research goals, which are largely comparative in nature. On the basis of the data that we have been able to collect in four cultural entities, we will conduct analyses of the psychometric properties of the new morality scale, make various cross-cultural comparisons of intuitive moral tendencies, look at deliberate decisions in moral dilemma scenarios, and finally, on the basis of the Moral Deviance Factorial Survey, also examine cultural tendencies in regard to moral particularism and impartiality.

Before we touch on the background of the third data collection in more detail and, in this context, discuss our case selection strategy and the methodological prerequisites for cross-cultural research, among other things, we would like to provide a brief overview of all three primary data collections conducted for this project. *Table 6* displays the key aspects of our three primary data collections. Furthermore, we will address the individual data collections with their empirical sample properties in due course. However, as announced, we will *primarily focus on the third data collection* within the text at hand. For more in-depth insights into the first two data collections, we therefore refer to three preregistered research plans, which can be accessed online and whose access links can be found in the following table.[30]

---

[29] For a critical view on this approach see: (Rippl & Seipel, 2022).

[30] Prior to any discussion of the samples in our cross-cultural research, we want to make a general comment: *none of our data collections captured representative samples*. Consequently, generalizations about the cultural entities we study are hardly robust and difficult to make, and the results we present in the course should always be interpreted in light of this fact. In many places in this text, we explicitly point out that we are primarily referring to cultural samples and not to the respective cultural entities in a representative way. In other places, however, it is also important to put linguistic precision behind the flow of reading. In the latter cases, we kindly ask readers to bear in mind that our samples are *NOT* representative and thank them for their understanding.

Table 6: Key aspects of the three primary data collections conducted

	Data Collection 1*	Data Collection 2	Data Collection 3
Study	MaC-DRS Development Study	MaC-DRS Validation Study	Cross-Cultural Study
Design	Semi-Experimental Online-Survey	Online-Survey	Semi-Experimental Online-Survey
Data Collection Mode	Convenient sampling: Distribution of questionnaire link to universities / student sample	Commissioning of a company specializing in data collection: quota sample/access panel/ non-student sample	Commissioning of a company specializing in data collection: access panel / non-student sample
Achieved Sample Size	N = 792	N = 2,326	N = 2,982**
Areas of Data Collection	Germany	Germany	Egypt Germany Japan USA
Ethics Committee Approval	Approved	Approved	Approved by independent institutions in all four countries of the data collection
Preregistration	DOI: http://dx.doi.org/10.23668/psycharchives.12183	DOI: https://doi.org/10.23668/psycharchives.13059	DOI: http://dx.doi.org/10.23668/psycharchives.14630

* Detailed insights into the structure of the studies, the research instruments used for data collection, and the specific theoretical considerations for data collection can be found, among other things, in the preregistrations, which are available online. The complete questionnaires for each of the three data collections can also be viewed in the preregistered version of the data collection-specific research plan. The preregistered research plan for the third data collection also contains four language versions of the translated questionnaire. The complete questionnaire is hence available in Arabic, English, German and Japanese in the corresponding research plan, which is available online. ** In order to improve the robustness of the substantive analyses, we adjusted the sample from data collection 3 for cases with poor data quality. This fact will be explained in the course of the text and can be retraced with all the sample adjustment steps in the **Appendix**. The final (*adjusted*) sample, on the basis of which we conducted the substantive analyses that go beyond psychometric tests, comprises N = 2,360 cases.

2.2. Investigations of the Human Moral Mind

Our concern is the cross-cultural investigation of the human moral mind. Within the framework of this project, we will conduct a total of four larger empirical investigations. Although the investigations each serve their own purpose, they are linked by the common thread of the propositions we have formulated within our MFT and MaC coalescent perspective, and the overarching questions we seek to address. Our investigations aim to gain empirical insights into moral universals and the cultural calibration of people's moral mind. At this point, we would like to give the reader a first impression of what can be expected in the further course of this work by highlighting the most important aspects of our investigations, which will be presented and discussed in detail in subsequent chapters.

Investigation of the human moral mind I:

> As mentioned at the beginning, in the empirical part of this work we will primarily deal with the data from the third, cross-cultural data collection and only use the data from the first two collections in the context of developing the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS). Our first empirical investigation will deal with the development and empirical validation of our moral scale. On the basis of three primary data collections and associated preliminary investigations, we will guide you through the development process of MaC-DRS and subsequently examine the psychometric properties of the scale. In this section, we will provide empirical evidence demonstrating the suitability of our instrument for the cross-cultural study of the human moral mind. Moreover, in this part of our research, we take the first hypotheses into account and approach the assumptions put forward on moral universalism.

Investigation of the human moral mind II:

> In this section of our research, we will first take a closer look at the data from our third data collection so that we are familiar with the basis of our further investigations. Following on from this, we will briefly consider the aspect of culturally variant response styles, which will need to be taken into account in subsequent analyses. Once we have built our foundation in this way, we will move on to the analysis of cross-cultural similarities and differences in the context of moral intuition. We focus extensively on cross-cultural comparative analyses based on MAC-DRS. In addition to the examination

of the cultural differences hypotheses, we constantly keep in mind our research question: Which moral system guides cooperation in different cultural contexts? In this section, we will not only demonstrate the massive influence of culture on the calibration of the human moral mind, but we will also show that, contrary to what we expected, individualizing morality has a cross-cultural significance and, in this context, point to processes of social change. This section forms the core of our cross-cultural investigations of the human moral mind.

Investigation of the human moral mind III:

In our third investigation, we address (binding and individualizing) moral dilemma scenarios. We supplement the MaC-DRS findings targeting moral intuition with cross-cultural findings on deliberate decisions in moral dilemmas. The relevant results provide additional insight into our overarching research question. Furthermore, the empirical findings presented in this section largely support the view that individualizing morality is indeed important in WEIRD populations, but also beyond. Additionally, our analyses support the view of several researchers who assume that moral intuition and deliberate moral cognition do not per se show tendencies that must align with one another.

Investigation of the human moral mind IV:

Finally, in our last investigation, we focus on aspects of the Moral Deviance Factorial Survey (MDFS) that we developed. The main aim of this cross-cultural investigation is to test the impartiality/particularism hypothesis. However, our research tool also allows us to examine whether specific acts of moral deviance lead to a similar valuation of the relevance of deviance and the judgment of deviance. The insights gained in this section suggest that cultural realities are intrinsically complex, which is why a binary logic for investigating and explaining moral impartiality/particularism is unlikely to do justice to cultural entities. This section's analyses also contribute to answering our overarching research question. Together with the results of the previous studies, our investigations lead to empirical findings based on a variety of methods that allow us to better understand which moral system guides cooperation in the various cultural entities we study.

As already mentioned, we will primarily draw in our investigations on the data obtained from four different cultural entities, i.e., **data collection 3**. Cross-cultural research is not only appealing because it allows researchers to approach human universals and to illuminate the

specifics of cultural realities, but it also comes with special methodological challenges that must be taken into account. These challenges include not only the culturally sensitive translation of measurement instruments, but also the statistically based necessity of ensuring that the same phenomenon is measured across cultures in order for comparisons to be meaningful. Furthermore, it should also be noted that the cultural programming of the mind includes, among other things, language, culture-specific modes of communication, and a phenomenon called response style (Barmeyer, 2010; Bogner & Landrock, 2016; Żemojtel-Piotrowska & Piotrowski, 2023). There are pitfalls lurking here that cross-cultural researcher needs to take into account. Against this background, we see the need to take a closer look at various aspects of cross-cultural research so that they can be given appropriate consideration in the further course of the empirical examinations in this work.

In the following, we will discuss several aspects of cross-cultural research that are relevant to our third data collection and the analyses based on it. In this context, we will first address the *case selection* and then recognize *the team* of supporting researchers that were essential to conducting the cross-cultural data collection. Thereafter, we describe the instruments comprised by the *questionnaire* of the third data collection. From these introductory remarks we will take off to a methodological section starting with *power analysis* and sample size. As *measurement invariance* of MaC-DRS can be regarded a pre-requisite for the investigations that we aim for, we will focus lengthier on *equivalence* and *bias* in cross-cultural research. Beginning with design components that shall minimize bias in cross-cultural projects, we will lead over to discuss measurement invariance. This part is followed by a section on statistical strategies to be applied when confronted with non-equivalence. After this detailed discussion on important methodological elements in cross-cultural research and the background to our third data collection, we have laid the foundation for proceeding and will be able to embark on the cross-cultural investigation of human moral mind. In the next section, we begin with the rationale behind the selection of cases for our cross-cultural empirical endeavor.

2.3. Case Selection: Empirical Evidence Supporting the Assumption of Particular Logics of the Situation

Research has identified many relative differences and similarities between cultural entities (e.g., Triandis, 2001; Murray & Schaller, 2010; Cross et al., 2011; Leung & Cohen, 2011; Gelfand et al., 2011; Thomson et al., 2018; Schulz et al., 2019; Haerpfer et al., 2022; Kitayama & Salvador, 2024). When it comes to comparing cultural entities, this is best be done using cultural

dimensions for they give a common ground allowing researchers to interpret meaningful (relative) differences (Barmeyer, 2010).

With the aim of finding out which moral system guides cooperation in different cultures, we tacitly assume cross-cultural differences in moral domain relevance. As stated before, we base this assumption on previous empirical findings and the proposition of culturally diverging requirements in, among others, the regard to cooperation. Furthermore, we derived two ideal-type sociocultural structured logics of the situation (*Model 2*; *Chapter 1*). Based on them we reason that these different logics would promote overall social orientation on group life or individual life, and subsequent interdependent self-construal and binding morality, or independent self-construal and individualizing morality. In this respect and in order to test our hypotheses, it is advisable to opt for a *purposive sampling strategy*.

Our sampling strategy shall ensure that the cases we will select for our subsequent investigations have characteristics that allow us to classify them, on the one hand, as *relatively different* from one another on several cultural dimensions and accordingly also in sociocultural requirements. And on the other hand, it shall allow us to classify them as different in terms of a social focus that is either on *group* life and interaction with the group, or on *individual* life and a multitude of interactions between (autonomous) strangers. Consequently, we decided for a contrastive selection strategy. We base our case selection on the criterion of *relative difference* on cultural dimensions that are meaningful in regard to assume cultural contexts that differ with respect to an overall individual- or group-centered social orientation. Moreover, research in psychology still suffers from a WEIRD bias (Henrich et al., 2010; Ellemers et al., 2019; Apicella, 2020; Krys et al., 2022; Uskul et al., 2023). Therefore, to contribute to our understanding of the human psyche, it is also desirable not to focus exclusively on Western countries.

For our **sampling strategy**, we decided to consider several **cultural dimensions** that have at least some of the following characteristics: the dimension is at least somewhat related to an evolutionary reasoning. The dimension is known to take effect on the cultural constitution of the human moral mind and/or selfhood. Additionally, the dimension demonstrates proximal and/or distal cause(s) of cultural variation, and lastly takes effect in an overall individual- or group-centered social orientation. Based on these characteristics we decided to look at: self-construal (Markus et al., 2010; Vignoles et al., 2016; San Martin et al., 2018; Kitayama & Salvador, 2024); cultural logics of face, honor and dignity (Leung and Cohen, 2011; Uskul et al., 2019; Uskul et al., 2023); religious heritage; (historical) kinship intensity (Schulz et al., 2019; Henrich, 2020; Bahrami-Rad et al., 2022); (normative) tightness and looseness (Gelfand

et al., 2011; Roos et al., 2015; Gelfand et al., 2017); (historical) subsistence style (Talhelm, 2022); disease prevalence (Murray & Schaller, 2010); values (Haerpfer et al., 2022); relational mobility (Thomson et al., 2018), and Hofstede dimensions (including the dimension collectivism-individualism) (Minkov & Kaasa, 2022; Żemojtel-Piotrowska & Piotrowski, 2023).[31]

In addition, for we favor to include rather diverse cases into our study, we also looked at a relatively recent method for calculating cultural distances. In examining cultural distances, we used the method of Muthukrishna and coworkers (2020a) and calculated overall differences and differences on multiple dimensions for a set of countries. Muthukrishna et al., (2020a) state:

> the "cultural-distance metric [is] based on the F_{ST} technique from population genetics, [that is] applied to the WVS (…). CF_{ST} is a theoretically defensible and robust method of measuring cultural distance, grounded in evolutionary theory. It considers differences between distributions of cultural traits rather than point estimates or arbitrary dimensions" (p. 698).[32]

Based on our case-selection investigation, we have decided for ***Egypt***, ***Germany***, ***Japan*** and the ***United States of America*** as proxies of cultural entities in our study. Accordingly, we plan to collect data in four countries that differ with respect to various cultural dimensions. The following tables (***Table 7 – Table 10***) summarize main findings of cultural differences in respect to the cases we selected for our cross-cultural investigations. Additionally, more far-reaching insights can be found in the relevant literature (as cited above) and in the **Appendix**.

Table 7: Overall Cultural Distance

[33]	Egypt 2005-2014	Germany 2005-2014	Japan 2005-2014	United States 2005-2014
Egypt 2005-2014		0.243	0.307	0.24
Germany 2005-2014	0.243		0.075	0.079
Japan 2005-2014	0.307	0.075		0.118
United States 2005-2014	0.24	0.079	0.118	

---

[31] See also: https://www.hofstede-insights.com/fi/product/compare-countries/
[32] Italics are taken from the original source; WVS is the abbreviation for World Value Survey; CF_{ST} is the term for cultural distance; the words in square brackets were added by the author for grammatical reasons.
[33] Calculated at: http://www.culturaldistance.com/; results are based on data from 2005-2014. Note: scores can range between – 1 and + 1, whereby a distance of 0.1 is already considered meaningful (Muthukrishna et al., 2020).

Table 8: Hofstede Dimensions

Countries[34]	*Power Distance*	*Individualism*	*Masculinity*	*Uncertainty Avoidance*	*Long Term Orientation*	*Indulgence*
WEIRD:						
USA	40	91	62	46	26	68
Germany	35	67	66	65	83	40
Asia:						
Japan	54	46	95	92	88	42
Middle-East/ Africa						
Egypt	80	37	55	55	42	0

Table 9: Disease prevalence and reliance on paddy rice farming

	Disease prevalence (9 items) *	*Disease prevalence (7 items)*	*Historical reliance on paddy rice farming ***
Egypt	0.44	0.76	…
Germany	-0.87	-0.78	Very low
Japan	0.43	0.25	Very high
United States	-0.89	-0.64	Low

* Murray and Schaller (2010) note in regard to the disease index, that "for each overall index (whether based on (…) seven, or nine items), the mean is approximately 0; positive scores indicate disease prevalence that is higher than the mean, and negative scores indicate disease prevalence that is lower than the mean" (p. 101). ** The categorization (i.e., very low; low; very high) is based on insights from (Talhelm, 2022).

Table 10: Country variations on several cultural dimensions

Countries[35]	*Tightness score**	Countries	*Dominant Self-Construal***	Countries	*Relational Mobility****	Countries	*Kin-ship intensity*****
WEIRD:		**WEIRD:**		**WEIRD:**		**WEIRD:**	
USA	5.1	USA	Independent	USA	0.182	USA	Low
Germany (Former East)	7.5						
Germany (former West)	6.5	Germany	Independent	Germany	-0.01	Germany	Low
Asia:		**Asia:**		**Asia:**		**Asia:**	
Japan	8.6	Japan	Interdependent	Japan	−0.414	Japan	Low
Middle-East/Africa		**Middle-East:**		**Middle-East:**		**Middle-East:**	
Egypt	-	Egypt	Self-Assertive Interdependent	Egypt	−0.194	Egypt	High

Tightness/Looseness* indications are based on (Gelfand et al., 2011); ** *Self-construal* categorization is based on (San Martin et al., 2018; Uskul et al., 2023); * *Relational mobility* indications are based on (Thomson et al., 2018); **** *Kinship intensity* categorization is based on (Schulz et al., 2019).

According to the dimensions we examined, we can heuristically assume the following contexts and respective situational logics on the basis of empirical data: We suggest that the cultural contexts of **Egypt and Japan** are characterized, among other things, by ***group-centered (interdependent) overall social orientation***. Accordingly, we assume that the everyday logic of the situation of actors from the respective countries is shaped by present significant symbols that condition the activation of an interdependent frame of reference and corresponding codes and scripts. Since we consider morality to be part of the self, we consequently also expect a group-oriented, binding moral prevalence. Given that actors socialized in the respective cultural contexts are confronted with self-relevant stimuli, i.e., the items of our questionnaire, we assume that a particular response behavior will be expressed according to the active self-construal frame. This response behavior serves as the basis for the data and thus forms the basis for testing our hypotheses. Furthermore, based on the evidence of the dimensions that we inspected we heuristically assume, in contrast to Egypt and Japan, that Germany and the United States are characterized by the following contexts and respective situational logics: We suggest that ***Germany and the Unted States*** are characterized, among other things, by contexts of ***individual-centered (independent) overall social orientation***. Again, we assume that the everyday logic of the situation of actors from the respective countries is shaped by present significant symbols that condition the activation of an independent frame of reference and corresponding codes and scripts. Thus, in line with our reasoning, we also expect individual-oriented individualizing morality prevalence and frame congruent response behavior. We summarize our considerations in the context of our case selection in ***Model 3*** (***Figure 5***) which can be found below.

As mentioned elsewhere in this text, we would like to emphasize that cultures are diverse. So, the notion of diversity applies of course also to the four cases in our project, even though we heuristically classify them.[36] As shown, for example, by the scores for individualism (USA: 91; Germany: 67) (see also: Minkov & Kaasa, 2022), the cases with a common predominant overall social orientation should not be misunderstood as homogeneous despite our classification. Rather, the cultural entities in our study differ along our continuum-based dimensional classification as well as on many dimensions not considered here. Therefore, we expect overall effects of response behavior that are consistent with our classification and theorizing, but also variation(s) between the cases that we grouped together (i.e., Egypt and

---

[36] Our classifications (i.e., "overall group-oriented" respectively "overall individual-oriented") are best understood as an ideal-typical categorizations (Weber, 2013) or as a heuristic that we approach from varying angles of cultural dimensions.

Figure 5: Model 3 — Countries, Self-Construal Frame and Frame Corresponding Moral System
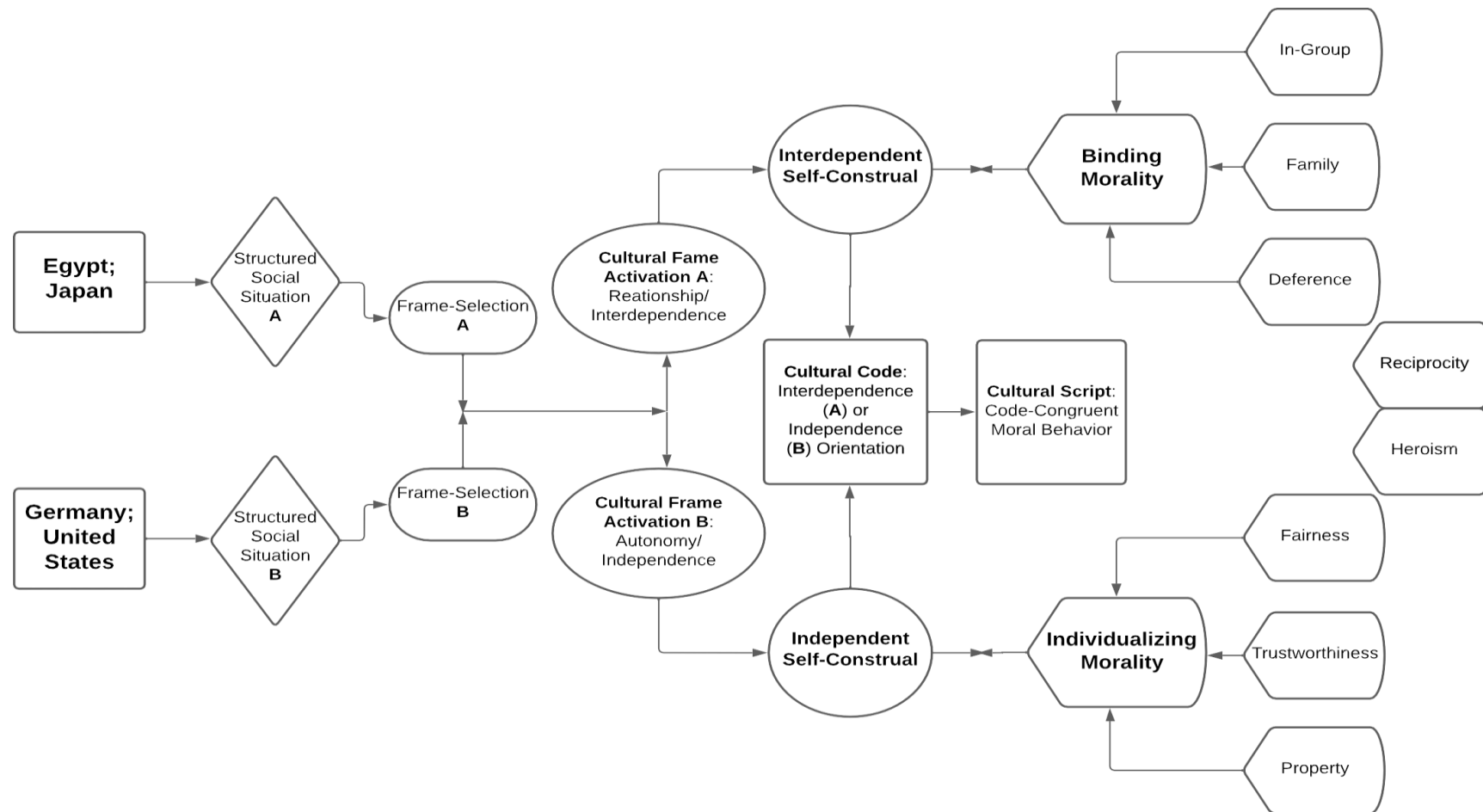
Figure 5 (*Model 3*) corresponds to an adaptation of Model 2 (Figure 4) to the case selection of our cross-cultural study just described. Based on empirical data we consider it justified to *heuristically* assume the (ideal-type) *logic of the situation* of an *overall group-centered (interdependent) social orientation* for *Egypt* and *Japan*. Furthermore, we consider it also justified to *heuristically* assume the (ideal-type) *logic of the situation* of an *overall individual-centered (independent) social orientation* for *Germany* and the *United States*. *Model 3* is based on the MSE (and the first two models that we have outlined) and shows the theoretically expected course of the respective frame selection, including the activation of self-construal and moral system in given socially structured situations. As we have explained theoretically, we further assume that reciprocity and heroism likely correspond to a general disposition of cooperation that occupies the space between binding and individualizing morality.

Japan; Germany and the United States). Since the scales we use to measure morality and self-construal are each 8-dimensional, we can capture the assumed heterogeneity of the cases in our study, provided that the scales show measurement invariance across our selected research areas. More information on our questionnaire and the instruments we use can be found below. Furthermore, it should also be emphasized once again that we are assuming relative cultural differences. Therefore, endorsement of all the moral domains we propose should be found in the cultures that we study. The latter, however, to varying degrees. Without the support of our cooperating partners, the realization of this cross-cultural project would not have been possible. We would therefore like to briefly introduce the team behind this project.

2.3.1. Cooperating Partners

When doing social science research on humans, it is important to do research with the people rather than just about people. In our view, this is true from a general epistemological perspective, but even more so when a research project aims at cross-cultural inquiries. Since we decided, as part of our selection strategy, to focus on comparisons between Egypt, Germany, Japan, and the United States of America in this project, we have accordingly tried to unite partners from the target countries of this study on the project. We are grateful and pleased to announce that we have found partners in all of the countries studied and beyond. Our project includes a diverse research team of collaborators from four continents, including six cooperating universities, six professors, two postdoctoral researchers, and three doctoral students. The research team is composed of the following individuals: Aya Murayama (Kindai University Osaka, Japan), Paul Teas, Linda Skitka (both University of Chicago, Illinois, United States), Rania Miniesy, Dina Rabie and Sahra Ahmed (all tree British University of Egypt), Vivian Vignoles (University of Sussex), Ulrich Kühnen and Klaus Boehnke (both Constructor University Bremen, Germany), Johannes Huinink and Pay Laurin Jessen (both University of Bremen, Germany). The last-named author is the principal investigator of this project. Among other support, our partners in the target countries took care of the application for approval of the ethics committee for our joint project at their respective home institutions; in all four countries, our data collection has received the approval of the respective ethics committees. Even though the outstanding contribution of this team does not receive the full attention it deserves within these pages, the remarkable support of all partners in the realization of this project should be emphasized. Without them, this thesis would not exist in its present form. On that note, the principal investigator of this project would like to express his sincere gratitude to

everyone involved. After providing insights into the case selection and the research team, we would now like to briefly discuss the instruments we used for the third data collection.

2.4. Data Collection in Four Cultures: The Questionnaire

As mentioned above, we collect the data in the four target countries exclusively online. Respondents of the third data collection are provided with a questionnaire in four languages (i.e., Arabic, English, German and Japanese) in a form suitable for online studies. In order to pursue our research objectives, our questionnaire comprises in total 111 items.[37] It should be noted at this point that we will not consider all the constructs covered by our questionnaire in the following. Various items are part of future studies. However, we would like to briefly present the complete questionnaire below so that a first overview of our cross-cultural data collection can be obtained.

We have integrated the three instruments developed against the background of our MFT and MaC coalescence perspective into the questionnaire. Accordingly, the *Morality as Cooperation—Deviance Relevance Scale* (MaC-DRS; 7-point response format, 32 items), the *Moral Deviance Factorial Survey* (four randomly allocated vignettes per respondent; vignette-pool = 168; four items per vignette each with a 7-point response scale),[38] and the *moral dilemma scenarios* (9 scenarios; forced choice response options) will be used to capture moral tendencies across cultures. S*elf-construal* will be assessed via the latest version of an 8-dimensional self-report instrument that comes with 48 items and a 9-point response scale (CIRN Self-Construal Scale Version 3; Vignoles et al., 2016; Uskul et al., 2023). Additionally, we constructed and integrated two items to assess *intention vs. consequence in moral judgment* (response format is a slider ranging from 0-100). We based the items on insights from Curtin and colleagues (2020). However, we will leave these items out in the following investigations. To capture potential

---

[37] Several items in the study explicitly address *deviant* behavior in relation to specific moral domains or combinations of moral domains. Although moral deviance is addressed in the study, extreme moral transgressions (e.g., murder or rape) do not occur in the items. In addition, the MaC-DRS items used in the study were tested in qualitative pretests. None of the participants in the pretests expressed any type of psychological discomfort, anxiety, distress, or the like related to the items. Also, the other items are either known not to cause distress, were tested in a pretest that showed they were not associated with distress, or touch on a topic (e.g., the number of herbs in traditional meat dishes) that is not known to cause distress. In addition, as also discussed further below, because participants are recruited by a professional company, vulnerable groups or children are not expected to participate in the study. Furthermore, our partners in the respective target countries of this study ensured that the items are culturally sensitive and appropriate for the context. All in all, no potential negative effects of the instruments used in this study are expected.

[38] We would like to express our gratitude to Prof. Dr. K. Auspurg for her immensely valuable advice on our factorial survey.

effects of *pathogen prevalence* directly we created three new items asking respondents to indicate the number of spices/herbs in a traditional meat dish from the country of their upbringing (response options are 0-30 spices/herbs). The items were based on insights from previous studies (Sherman & Billing, 1999; Ostfeld et al., 2005; Murray & Schaller, 2010; Atari et al., 2022b). To address one of the most important issues that humanity faces today, we also want to explore a possible link between morality and *attitudes to climate change* across cultures. Accordingly, we have integrated two corresponding items that each come with a 5-point response format (Lenzner et al., 2022). Studies on morality and climate change are part of future analyses and will not be addressed further in this text. Morality is known to be associated with political ideology (Graham et al., 2009; Iyer et al., 2012). However, the classic left/right scale of political ideology is not suitable for every cultural context (Zuell & Scholz, 2019) and primarily reflects a pattern of political ideology in WEIRD countries. To avoid this shortcoming of applicability and still incorporate a political measure, we integrated an item on the *probability to vote in the next official election* (four-point scale ranging from "very unlikely" to "very likely"; Boehnke et al., 2022b).[39] Also this item will be addressed in future research. Further, to be able to compare potential intra-cultural differences, we included two binary items on *growing up* and on *living in a city or on the countryside*. An additional item on *individual (residential) mobility* was included in the questionnaire in order to directly record individual experiences of changing social contexts (ISSP Research Group, 2018). The item comes with four response options ranging from "*I have lived in different countries*" to "*I have always lived in the same neighborhood*". Another social aspect that we want to examine in a future cross-cultural study with regard to morality is *relative (individual) deprivation* (Zick et al., 2011). We included one corresponding item to measure this concept (5-point scale). Lastly, to gain insight into the (sample) characteristics of the four cultural groups, the questionnaire contains a relatively extensive set of additional *sociodemographic* items (ISSP Religion IV, 2018; Wasmer & Baumann, 2019; SOEP, 2019) (10 items). The questionnaire captures *age*;[40] *gender*;[41] *religious denomination* and level of *religiosity*; *citizenship* (of respondent and respondents'

---

[39] The item on the probability to vote in the next official election could not be asked in Egypt and was therefore not surveyed in this sample.
[40] The countries in our study differ considerably in their average age. Reflections on the age of the samples and our data collection strategy in this context can be found in the **Appendix**.
[41] For the Egyptian region, it would be difficult and culturally inappropriate to use the gender item to ask about diverse (non-binary) in addition to female and male. We have therefore omitted the third response option for the gender item in the Arabic questionnaire.

parents); *education* (i.e., years in school[42] and highest educational degree attained)[43]; and net *earnings*. Furthermore, a welcoming and debriefing as well as transition texts explaining the respective instruments are included in the questionnaire. In order to assume sufficient power of statistical tests, the size of the samples to be analyzed is important. The following section is therefore devoted to the a priori power analyses that we conducted prior to the third data collection.

2.4.1. Power and Sample Size

In our cross-cultural investigations we will compare four countries, use them as proxies for cultural entities which come with certain characteristics, and look for, among other things, (cross-cultural) differences in relevance of moral domains. In this line we hold as the basic hypotheses for our empirical testing: H0 = there are no significant group differences; H1 = there are significant group differences. To determine a sufficient size of our sample we will consider *a priori power analysis* (Lakens, 2022). In this respect four statistical quantities are related to one another: the significance level α, which serves to control type-1 error; the effect size, which indicates the relative size of an effect and, besides the significance, the (relative) meaningfulness of the effect; the power of the test, which is calculated as $1 - \beta$ whereby β corresponds to type-2 error, and finally the sample size, i.e., the N of the study, where N represents the total size of the sample.

First, let us assume a simple (two-sided) T-test as the method for testing for differences in means between groups (i.e., cultural entities here). A T-test allows us to calculate whether a (potential) difference between two independent groups is significant or not (Jann, 2005).

Secondly, we turn to the effect size. "One problem with planning the sample size on the basis of the size of an effect (…) is that the effect size is precisely the information that the researcher is trying to uncover (…). As a consequence, there is always some uncertainty regarding the required sample size needed" (Lakens, 2014, 702). Our project suffers also from having no insight on the actual effect size. Furthermore, since we do not have an empirical basis

---

[42] It became apparent in team discussions that the item "*How many years did you attend school?*" was not understood the same way across the countries of our study. We have therefore added the following part, which appears in brackets, to the item: "*How many years did you attend school? (Our question refers only to the years you spent in school (i.e., primary and secondary education) before possibly attending another educational institution for further education (e.g. vocational training and/or university))*."
[43] The education systems of various countries may differ. Therefore, cooperating partners from each target country have defined the response options for the item "What is the highest educational level that you have attained?" individually adapted to the context. In order to achieve comparability of the response options across countries, we decided to re-code the responses of the education item according to ISCED 2011 (UNESCO, 2012).

for determining the effect size, we decided to proceed in the power analysis with a conservatively small effect size. We justify this decision (Lakens, 2022) based on a generally cautious stance in respect to effect size assumptions on the one hand and a predominantly theoretical rational on the other. Our main dependent variable (*dv*) is moral domain endorsement (moral relevance), and we hypothesize that (significant) differences between cultural entities exist in regard to this *dv*. We expect cross-cultural differences in relevance ratings of moral domains due to variant needs and affordances of cultural entities. The background to these variant needs and affordances of cultural entities is probably the circumstance that different cultural entities have had to solve different (reoccurring) social problems in the course of their existence (Curry, 2016). Behind this hypothesis is the theoretical consideration that cultural entities are themselves a product of diverse natural ecologies and the subsequent development of cultural ecologies, which for their part are a response and an element of change to their initial natural conditions. Moreover, cultural ecologies follow path dependencies of intergenerational transmission and persistence, which eventually leads to variations on different dimensions between cultural entities (Tomasello, 2017; Creanza et al., 2017; Mesoudi & Thornton, 2018; Muthukrishna et al., 2021; Brown et al., 2022). Nevertheless, although we expect to find cross-cultural differences on our *dv* between countries, we cautiously assume these differences to be rather small in effect size. We base our reasoning on the following: *A)* all human societies depend to a certain extent on cooperation (Graham et al., 2016; Henrich, 2020; Henrich & Muthukrishna, 2021; Muthukrishna, 2021). Our evolved moral mind incentivizes cooperation and is considered a human universal (Haidt & Joseph, 2007; Kurzban et al., 2015; Hare, 2017; Curry et al., 2019a; Brown et al., 2022). *B)* the instrument that we use to measure moral domain endorsement (i.e., the *Morality as Cooperation — Deviance Relevance Scale*) comes with items having a clear polarity (by asking respondents about the relevance of *deviant* behavior towards several moral domains). Further, as bad is stronger than good (Baumeister et al., 2001), we assume this polarity to heighten a general response tendency towards the relevance realm of the 7-point scale (ranging from 1 "Extremely irrelevant" to 7 "Extremely relevant"). In other words, although we expect cross-cultural differences in endorsement of moral domains, we take into consideration that the domains captured by our scale are more likely to be rated as relevant across cultural contexts because the functioning of societies in general is based, to some extent and without neglecting cultural particularities, on (universal) principles of cooperation that are regulated by our moral mind. Therefore, we cautiously assume effect sizes to be small and consider a small effect size as already meaningful. Admittedly, we do not want to neglect that the possibility of larger effect

sizes may also holds true, especially as we felt the case selection decision based on a purposive sampling strategy. However, regarding the uncertainty of actual effect size we prefer to be sensitive to the above outlined possibility of small effect sizes and rather have an overpowered than underpowered sample size. Consequently, for the mean differences, we assume that the effect size Cohen's d, in our investigations this is e.g. the effect of "culture" on the relevance of different moral domains, is small (d = 0.2).

Third, to control for type-1 error— "i.e., the probability of incorrectly rejecting H0 when is in fact true" (Mayr et al., 2007, p. 52) —, let us turn to the significance level. Let us assume here that the (error) probability is at the conventional level of 0.05. We therefore base our test at first on a significance level of 0.05. However, since we will compare four groups (i.e., cultural entities) in search for differences in moral domain endorsement, we have a multiple hypothesis testing design. Hence, due to multiple testing we need to consider Bonferroni correction and adjust the α-level to counteract alpha (type-1) error cumulation (Lakens, 2022): by inspecting potential difference between four groups via a T-test we have to perform three tests, and by setting the type-1 error rate to 5% we eventually obtain $(0.05 \div 3 =)$ 0.1666 as corrected α-level.

Fourth and furthermore, let us come to the power of the test. The power, "the probability of rejecting false null hypotheses" (Mayr et al., 2007, p. 51), is given as 1 - β, i.e., 1 minus the (error) probability of type-2 error (which corresponds to the probability of maintaining H0, although H0 is actually false). In other words, the power of a test is the probability that the test correctly detects actual significant differences between groups, as far as they exist.

If we assume at *fifth*, that the sizes of the groups to be compared are identical, then the program *G-Power* (Mayr et al., 2007) can be used to perform an a *priori power analysis* to determine the sample size needed for a sufficiently powered study (Lakens, 2022). Mayr and colleagues (2007) state: "[a]n a priori analysis is used to determine the necessary sample size N of a test given a desired α level, a desired power level (1 - β), and the size of the effect to be detected (i.e., a measure of the difference between the H0 and the H1)" (p. 52). Thus, if we assume values of **a)** 0.80 (sufficient), **b)** 0.90 (excellent), and **c)** 0.95 (conservative) for the power (1 - β), we have everything together to calculate the N for our study based on different power restrictions. So, following these three power-values we will be able to obtain three sufficiently powered sample sizes.

Based on what we have elaborated, we performed a priori power analyses prior to the cross-cultural data collection.[44] The results of the analyses suggest that the following sample

---

[44] All calculations are based on a priori power analyses (two tailed T-test, difference between two independent groups, d = 0.2, allocation ratio between groups = 1) using G Power (Mayr et al., 2007).

sizes per sample (i.e., per target country) should be aimed for in our international study under the given power restrictions:

A priori power analyses results:

- a) critical t value = 1.963, DF = 786, n = 394 (per target group), n = 788 (for country wise t-test comparison), N = 1576 (for four target countries) and actual power = 0.80059;

- b) critical t value = 1.962, DF = 1052, n = 527 (per target group), n = 1054 (for country wise t-test comparison), N = 2108 (for four target countries) and actual power = 0.90036;

- c) critical t value = 1.961, DF = 1300, n = 651 (per target group), n = 1302 (for country wise t-test comparison), N = 2604 (for four target countries) and actual power = 0.95086.

A priori power analysis results based on Bonferroni corrected alpha-level:

- a) critical t value = 2.397, DF = 1048, n = 525 (per target group), n = 1050 (for country wise t-test comparison), N = 2100 (for four target countries) and actual power = 0.80;

- **b) critical t value = 2.396, DF = 1352, n = 677 (per target group), n = 1354 (for country wise t-test comparison), N = 2708 (for four target countries) and actual power = 0.90;**

- c) critical t value = 2.396, DF = 1634, n = 818 (per target group), n = 1636 (for country wise t-test comparison), N = 3272 (for four target countries) and actual power = 0.95.

Out of economic reasons on the one hand and reasons of sufficient power on the other, we decided before data collection for a sample size of n = 677 per country (actual power = 0.90) which equals N = 2708 respondents in total. However, cross-cultural research requires more than the consideration of power and sample size. Therefore, in the following, we will discuss the preparations for our cross-cultural data collection in more detail and take a closer look at the topics of equivalence and bias.

2.5. Preparing for Cross-Cultural Data Collection — Equivalence and Bias

Cross-cultural research requires precise preparation of the data collection and the statistical validation of the usability of measurement instruments. Hence, to be able to follow the objectives of our planned investigations several study design and statistical hurdles need to be taken into account. The following will deal with *equivalence* and *bias* in cross-cultural research (He & van de Vijver, 2012). We begin this section with a general introduction to the issue (van de Vijver & Leung, 2011). Thereafter, obstacles in cross-cultural study design and our attempts to overcome them will be discussed (Smith, 2014). In this section we focus on questionnaire standardization (Fischer & Milfont, 2010). General issues of method bias, i.e., sample,

instrument, administration and item bias are also addressed. The latter is leading over to the complexity of and our approach to translation (Behr et al., 2016; Boehnke et al., 2022a). Equivalence is then discussed in more detail. Measurement invariance (Davidov et al., 2014; Leitgöb et a., 2023), a procedure to test levels of (non-)invariance (Milfont & Fischer, 2010; Gäde et al., 2020a), and strategies to be applied when invariance testing fails (Welkenhuysen-Gybels et al., 2003; Chan, 2003; Fischer, 2004; Morgan & Winship, 2015; Putnick & Bonstein, 2016; Cieciuch et al., 2019) are scrutinized in this section. In the following methodological part, we will not only address various challenges that need to be considered and overcome in cross-cultural research, but we will also discuss the strategies we *have* applied in the study design of our cross-cultural data collection. In addition, we will discuss the strategies we *will* apply after the data collection to check whether standards for analyses, e.g. mean comparisons between groups, are actually permissible. After our methodological explanations, we delve into the first of our cross-cultural investigations of the human moral mind in the next chapter.

2.5.1. Equivalence and Bias

Central to the study of the question ***Which moral system guides cooperation in different cultures?*** is the issue of *comparability*. Generally speaking, in order to be able to compare the subjects being analyzed across cultures one needs to design studies that allow for meaningful comparisons in the first place. Additionally, one needs to empirically test whether instruments used to collect data fulfill certain theoretical and empirical properties. Only if instruments are (sufficiently) free of bias, and measurement invariance is supported, researchers can draw conclusions from data that are said to be substantial rather than research artefacts.

In cross-cultural research the term *equivalence* "refers to the level of comparability of scores across cultures" (He & van de Vijver, 2012, p. 3). It needs to be noted that equivalence is not a property of instruments applied to collect data. On the contrary, researchers need to take care of equivalence by designing studies that promote comparability before collecting data and by attenuating several sources of bias also post hoc to data collection. The term *bias* is the diametral counterpart to equivalence. Bias stands for a number of (systematic) errors that may distort empirical (cross-cultural) analysis (Davidov et al., 2014). From this viewpoint bias "refers to the presence of nuisance factors" (van de Vijver & Leung, 2011, p. 22) that may lead to over- or under estimation of group differences, spurious results or measurement artefacts. So, biased data may pose unreliable information and lead to misguided interpretation that

potentially overlooks actual differences and highlights artificial relations. Therefore, bias can be regarded as an integral threat to every empirical research aiming at comparability.

2.5.2. Method Bias and Study Design Features

Let us first address the design component of creating cross-cultural studies that try to account for sources of bias prior to the data collection (Leitgöb et al., 2023). We will begin by turning to *method bias*, which "is a generic term for nuisance that derives from the sampling, structural features of the instrument, or administration processes" (He & van de Vijver, 2012, p. 5). Cross-cultural investigations aiming at comparability must ideally work with samples that have similar to the same characteristics, with the exception of the variation of interest (i.e. cultural belonging). Furthermore, familiarity with the instrument(s) used to collect data should be sufficiently same across samples. Also, culturally different response styles must be taken into account, and the administrational process of data collection should be the same across areas of data collection. Lasty, translations should be produced so that they convey the meaning of instruments across research groups while being at the same time sensitive to potentially specific (cultural) group requirements.

Comparability also stems from the characteristics of the samples. To not end up with spurious results, that are distorted by *sample bias*, the groups being compared should possess approximately the same properties, apart from group belonging. Experimental trials take care of comparability by random group assignment and the logic of the law of large numbers — with increasing sample size and random distribution to experimental groups, the characteristics of the study subjects in the groups approximate each other. Thus, operating with random group allocation and with sufficiently large sample sizes safeguards comparability in experimental designs (Fischer & Formann, 2007; Oehlert, 2010; Myers, 2012). Even with designs that do not fully follow the experimental tradition, meaningful comparisons can be made, although samples are initially unbalanced. In quasi-experimental designs statistical matching procedures can be applied (post hoc) to enhance comparability across groups under investigation (Morgan & Winship, 2015). Beyond that, other statistical approaches such as ANCOVA (Völkle & Erdfelder, 2010; Liu, 2012) or Average Marginal Effects (Williams, 2012; Wooldridge, 2016) can be used to hold certain variables constant across groups. By holding variables constant across groups, the effect of the constant variables on the relationship of interest is partialized out and comparability is enhanced. However, the nature of cross-cultural studies is different to the nature of experimental designs. In fact, cross-cultural studies "are always threatened by bias

and inequivalence (van de Vijver & Leung, 2011, p. 17) precisely because participants in these kind of study designs are not randomly allocated to cultures (i.e., the groups under investigation).

We apply several **in- and exclusion** (i.e., study participation eligibility) **criteria** to the four cultural **samples** being collected in order to enhance comparability. For each country we aim for a sample that is heterogeneous in terms of age (from 18 – 85) and balanced in terms of (dominant) gender (50/50 male/female ratio).[45] All samples shall also include respondents from urban and rural areas (heterogeneous distribution of approx. 30% countryside and 70% urban), and the samples shall be heterogeneous in terms of education. We must assume, however, country-specific differences in relation to certain variables. Specific differences in the populations cannot simply be balanced across samples using our inclusion and exclusion criteria, as it, for example, may be very difficult to find respondents over the age of 80 in some countries.[46] Therefore, we also directly control for several sample characteristics via a set of sociodemographic items. Among other variables our questionnaire assesses age, gender, education (years in school and highest degree obtained), citizenship, country of upbringing mother/father, and net income. By directly collecting data on sociodemographic properties, we gain on the one hand a means for deeper analysis and on the other hand a means to assess and control for sample similarity.

Another source potent in distorting analysis in cross-cultural research is *instrument bias* that deals with stimulus familiarity and response styles (van de Vijver & Leung, 2011; He & van de Vijver, 2012). Participants from different groups, e.g. from different cultural backgrounds, may not have the same degree of *familiarity* with measurement instruments and their ways of assessment (e.g., computer-based questionnaires). Differences in familiarity in turn may systematically influence scoring on scales and thus produce biased results.

In our study, we address the familiarity issue through our **participant recruitment strategy**. Participants in all four target countries will be recruited by a company that is specialized in recruiting survey participants. Outsourcing the recruitment process has two important advantages in the case of our study.[47] First, it ensures that the desired sampling criteria are achieved. Second, since the participants recruited by the company belong to a pool whose members are used to participate in studies (Access Panel), we can expect at least a certain

---

[45] We also capture whether a respondent has a non-binary gender conception of themselves and allocate these cases alternately in the male/female categories in order to obtain samples that are gender-balanced.

[46] See our considerations on the age variable in the **Appendix** for further information.

[47] Respondents (Access Panel) were compensated for their participation in our study according to the criteria of the respective companies commissioned to collect the data.

degree of instrument familiarity. Although we cannot completely rule out differences in familiarity, we can expect an overall sufficient degree of familiarity with our measurement materials due to our recruitment strategy. This is all the more the case as our (quasi-experimental) survey design is relatively simple (no pictures and videos etc.) and does not include any tasks to be carried out other than answering.

Different *response styles* pose another source of instrument bias. Generally, response styles "refer to a systematic tendency to use certain categories of the answering scale on some basis other than the target construct" (He & Van de Vijver, 2012, p. 6). The central problem with response styles (RS) is indicated by the title of an article by He et al., (2021) named "People use scales differently". However, the mere fact that scales are being used differently is in itself not enough to describe the challenges RS pose to cross-cultural research. If response styles were randomly "distributed" among all people, they would actually pose no problem. In fact, this is though not the case. Response styles are culturally patterned which introduces a systematic difference component to the analysis of questionnaire responses across cultures. Thus, from this perspective, RS are a source of nuisance in cross-cultural research that researchers must carefully pay attention to (Cheung & Chan, 2002; Fischer & Milfont, 2010).

Several response styles exist, e.g., social desirability, midpoint RS, acquiescent RS and extreme RS (Bogner & Landrock, 2016). These styles can be seen as different habitual ways of responding to a scale regardless of the content of the respective instrument (Cheung & Rensvold, 2000; Smith, 2014; He et al., 2021). Social desirability is concerned with impression management (Mummendey & Bolton, 1993). This style illustrates respondents desire to portray themselves in a favorable way. Furthermore, social desirability is known to be especially influential when sensitive topics are touched in studies. Midpoint response style was associated by some scholars with item ambiguity and satisficing. This bias describes respondents' tendency to choose the middle category of a scale irrespective of the scales content. Item ambiguity and educational level may also affect acquiescence responding (Watson, 1992). Acquiescence bias is also called "yay-saying" (opposite is "nay-saying") and describes a tendency to agree to statements regardless of content. Another category of RS is extreme responding. As the name of this style implies, this bias describes a tendency to choose endpoints of scales regardless of the content of the instrument.

The relevant literature portrays evidence of cross-cultural differences in the use of response styles (Davidov et al., 2014; Thomas et al., 2014). So, from a (traditional) perspective, RS are seen as bias that distorts the comparison of samples. The existence of response style bias in cross-cultural studies may affect relationships between variables and the direction of

associations among variables. Response styles may also inflate correlations and eventually may distort the validity, reliability and invariance of scales across contexts (Watson, 1992; Cheung & Rensvold, 2000). This highlights the need to account for RS in cross-cultural studies. Viewing response style as bias goes along with the claim to control and potentially correct for these nuisance factor(s). Approaches to account for RS focus either on the design prior to data collection and/or on statistical post data collection procedures (He et al., 2021). With regard to post-hoc methods, simulation studies and real applications show their usefulness. Various methods allow RS to be controlled and approximately filtered out of the data (Davidov et al., 2014; Savalei & Falk, 2014). We will focus on these methods, in particular on score standardization procedures (within-subject standardization using ipsative scores) and a SEM approach working with a response style factor (Welkenhuysen-Gybels et al., 2003; Fischer, 2004), further below.

The portrayed perspective treats RS as nuisance. However, response styles can also be seen as significant cultural differences that indicate group characteristics such as cultural communication styles or cultural norms (Smith, 2004). Furthermore, as "there is no objective indicator of knowing the 'true' level of a psychological construct" (Fischer & Milfont, 2010, p. 91) one should be very cautious about removing effects interpreted as response style distortion blindly. Cautiousness is warranted for what is being removed may in fact reflect actual individual or cultural differences and/or "a combination of response styles and genuine differences" (He & van de Vijver, 2012, p. 7). Hence, the issue of response styles is complex and a decision for post hoc procedures to adjust for RS must provide a decent justification before application.

Lastly, another form of method bias is *administration bias*. This bias occurs under conditions of ambiguous instructions, reactive data collection techniques (interviewer effect) and communications problems (He & van de Vijver, 2012). We tried to counter this form of bias in our design by several means. First, several instructions used in our questionnaire have already been established in other studies. In addition, we subjected the new instruments to cognitive interviews and a quantitative pretest. Above that, also the current questionnaire was again quantitatively pre-tested, yet admittedly only in one country (Germany). Second, we choose to collect data in a rather non-reactive fashion. Our study is therefore designed as an online questionnaire and contains no verbal component. Consequently, we do not expect interviewer effects, or effects of speaking tone to occur. Furthermore, as we explicitly point out to the anonymity of the data collection mode, we also do not expect high levels of social desirability elicited by our design. Although coming with advantages, online research poses also a potential

source of administration bias as there is neither insight in nor control over the environment of respondents (e.g., third party effect) (Kraut et al., 2004; Bogner & Landrock, 2016). Yet, under considerations of economic constraints and the given advantages we believe that collecting data online is a reasonable choice for our project. Third, we standardized the order of the questionnaire; all materials, except the quasi-experimental component of the factorial survey, appear in identical order across the four target countries (Fischer & Milfont, 2010). In the factorial design part of our study four vignettes out of the vignette pool are randomly allocated to participants. Fourth, regarding potential communication problems, we choose to rely on simple language and simple item wording. We therefore avoided colloquial language, gave clear instructions and short transition-texts between different measures and avoided negations in the questionnaire wherever possible. Additionally, we avoided the use of metaphorical language. Moreover, as researchers from each of the target countries of our study inspected the items, we are confident that no taboo topics are being touched by the items of the questionnaire.

2.5.3. Item Bias and Translation

The matter of language leads over to *item bias and translation*. Items are said to be biased if there essential meaning is different across cultural contexts (He & van de Vijver, 2012). The detrimental consequences for cross-cultural comparisons due to item bias are easily imaginable. Scores coming from biased items do not deliver meaningful results and pose a fundamental issue to interpretation, as they imply that different concepts have been assessed across research areas. In other words, biased items may lead to data that captures not a specific concept across groups but rather apples and oranges. An issue that is making the task of comparing constructs across groups/contexts impractical.

Insufficient translation, spanning around the issues of potential communication problems pose a sources of item bias. Translating questionnaires for the purpose of collecting quantitative data cross-culturally is in itself a challenge. The task involves numerous hurdles and subtle differences in meaning. Translation becomes even more difficult when the content to be translated deals not with hard facts but with the world of psychological and social phenomena, which are themselves subjects of cultural variability. However, "[a] good translation (…) can largely avoid item bias" (Davidov et al., 2014, p. 61).

The debate of translating measurement instruments for cross-cultural research is inherently difficult and holds diverging perspectives (He & van de Vijver, 2012; Boehnke, 2022a). Within our project, we decided to apply an extended version of the classic forward and

backward translation approach (Brislin, 1970; Smith, 2014; Behr et al., 2016). In the process of the translations, we paid utmost importance to preserve the essential meaning (semantic equivalence) of all material (items, instructions etc.). Though, we aim also to produce a culture sensitive and culturally appropriate translation. Therefore, we made equal efforts to ensure equivalence in meaning, and questionnaire appropriateness and applicability in context. We followed an approach of four control steps in the translation phase: *first*, independent bilingual student assistants received guidelines and instructions, and produced a forward translation. Within this phase an initial translation of original source material into a respective target language was produced. *Second*, a backward translation was produced by different bilingual student assistants. In this process the initial forward translation was used as source for another translation back into the original language. In the *third* step, the backward translation and the original source version were compared with each other to identify potential (semantical and cultural) pitfalls. This step was also used to control if the essence of meaning of constructs and instructions is preserved in the translation process. Furthermore, the third step included a joint discussion of the project leaders and the student assistants who did the translations. The discussion concentrated on accuracy in wording and the aspect of item applicability in the target contexts of our study. Finally, the *fourth* step focused on a final check and editing before going into the field. The final review in the control process was done by our native speaking partners of the respective target countries. This phase was essential for achieving an adequate level of translation, as it zoomed in on linguistic nuances. Additionally, this phase was concerned with checking the applicability of the items in each context, and it ensured the final linguistic polish of all the instruments in the questionnaire.

2.5.4. Equivalence — Testing Measurement Invariance

From design-based approaches to enhance equivalence we are now turning to approaches to test levels of *measurement invariance*. The terms equivalence and invariance are used interchangeably throughout the following. Diverse tools and methods exist in the realm of assessing and testing equivalence. Our analysis strategy rests on the classical test theory and the most commonly used model-based approaches to examine equivalence. That is, in investigating equivalence we use partly exploratory factor analysis (EFA) yet rely mainly on (multigroup) confirmatory factor analysis (MGCFA) (Bagozzi & Yi, 2012; Moosbrugger et al., 2020; Schermelleh-Engel & Gäde, 2020; Brandt, 2020; Gäde et al., 2020a, Leitgöb et al., 2023). These approaches ground on the idea of latent variables, i.e., theoretical concepts that are not

directly but indirectly measurable via observable manifest indicators (i.e., items) that are assumed to reflect the latent construct (i.e., factor). Davidov et al., (2014) state, that the key idea of the "latent variable approach is that measurement equivalence can be tested by comparing empirical relations between the latent variable and the indicators across populations. Similarity of these relationships (…) is taken as evidence supporting the hypothesis of measurement equivalence" (p. 62).

2.5.5. Testing Invariance Using Factor Analysis — EFA, CFA and MGCFA

Exploratory factor analysis (EFA), as the name already indicates, is an exploratory, rather data driven approach to search for patterns of statistical relations between indicators and assumed underlying latent dimensions (factors) (Brandt, 2020). This approach is particularly valuable for theory development and hypothesis formation. Though it can also be used to refine theories when theoretical relations fail to appear in the data or are non-invariant across contexts (Brown et al., 2017). Hence, working with EFA can be a fallback strategy when measurement invariance testing fails completely.

> Within the analysis of our cross-cultural data, we will apply EFA with varimax and oblique (oblimin) rotation in each of the four countries separately to inspect whether the 8-dimensional factor structure of MaC-DRS appears from the data.[48]

The *confirmatory factor analysis (CFA)* belongs to the structural equation modelling (SEM) framework and is the theory-driven counterpart to EFA (Gäde et al., 2020a). This approach is used to test whether a theoretically informed and explicitly specified (measurement) model fits to empirical data. Generally, CFA offers means for hypotheses, construct validity (dimensionality), reliability and invariance testing. Testable models can be uni- or multi-factorial, with only first or several hierarchically higher order factors (e.g. second-, third-order factors). By multiple means of fit indices and cutoff values the actual fit of a theoretical model to the data can be evaluated in CFA. An overall test of model fit is the Chi-square test (χ^2-test). Since this test is however sensitive to sample size, several descriptive fit indices are usually used in addition to the Chi-square test. Therefore, in order to evaluate the goodness of model fit and in addition to the χ^2-test, the *Root Mean Square Error of Approximation* (RMSEA), *Comparative Fit Index* (CFI), *Tucker Lewis Index* (TLI), and the *Standardized Root Mean*

---

[48] This analysis will thereafter be safeguarded by separate CFA´s and MGCFA.

Square Residual (SRMR) should be taken into account. Cutoff conventions (Milfont & Fischer, 2010; Bagozzi & Yi, 2012; Gäde et al., 2020a) for these indices are as follows:

> Acceptable model fit is indicated by: $\chi^2 /df \leq 3$ (Chi-square to degrees of freedom ratio test), RMSEA ≤ 0.08, CFI ≥ 0.90, TLI ≥ 0.90, and SRMR ≤ 0.10.

> Good model fit is indicated by: $\chi^2 /df \leq 2$, RMSEA ≤ 0.05, CFI ≥ 0.95, TLI ≥ 0.95, and SRMR ≤ 0.05.

> After applying EFA we will run country separate CFA´s to test the factor structure (8 dimensions of first-order and three higher-order factors) of our new moral scale (MaC-DRS) in each of the target cultural context.

Demonstrating acceptable model fit via CFA is however not sufficient when aiming for comparisons of scores (e.g., means) across groups. Groups may have a different meaning of a construct, so that not all dimensions of a construct found in one group are a priori assumable to be found in another group. Also, topics that are readily accessible in one group may be taboo in another, or differing response styles may preclude score comparisons across groups. As stated before, measurement invariance is not a property of an instrument but must be empirically tested. Hence, to do so the CFA approach must be applied to test for invariance across groups. This can be done using *multigroup confirmatory factor analysis* (MGCFA) (Leitgöb et a., 2023). "In MGCFA, the theoretical model is compared with the observed structure in two or more samples" (Milfont & Fischer, 2010, p. 113).

2.5.6. The Procedure of Invariance Testing

A distinction is made between different levels of invariance of measurement instruments across groups. Depending on which level of invariance is reached specific means of analysis, i.e., unstandardized associations or comparison of (latent) mean values, can be used (Cieciuch et al., 2019; Leitgöb et al., 2023). We will turn to the levels of measurement invariance below. First, however, we come to the procedure of invariance testing. Different levels of invariance are tested step by step by introducing restrictions to the model. The restrictions being introduced signify that certain model parameters are constrained, i.e., equated. Model restrictions include equating factor structure, factor loadings, intercepts and error variance to test for *measurement invariance*, and equating factor variance, factor covariance and factor means to test for *structural invariance* (Milfont & Fischer, 2010). In the procedure of stepwise invariance testing, the less constrained model is compared to a model with more constraints. If the model with more constraints reflects the data as well as the model with fewer constraints, the model with

more constraints is preferred. This logic of model selection is determined by the principle of parsimony, as more constrained models require fewer parameter estimates and are therefore more parsimonious.

To establish whether the more restricted model fits the data as well as the less restricted model, again fit indices are used. The chi-square difference test (Gäde et al., 2020a) is known to react sensitively to the sample size and is significant even for small deviations from invariance. Therefore, the alternative fit indices CFI, RMSEA and SRMR should also be taken into account. When comparing the less constrained model with the more constrained model and when considering the alternative fit indices, the following conventions apply:

The CFI value must not decrease by more than 0.01 units.
The RMSEA value must not increase by more than 0.015 units.
The SRMR must not increase by more than 0.01 units.

As long as the conventions for fit indices hold, it is demonstrated that both models reflect the data sufficiently well (Schwab & Helm, 2015; Putnick & Bornstein, 2016; Cieciuch et al., 2019).[49] Now, provided the conventions hold, the more parsimonious (i.e., more constrained) model will be selected and thus, depending on the specific model restriction, a certain level of measurement invariance is demonstrated across groups.

2.5.7. Testing For Different Levels of Invariance

We now come to the different *levels of measurement invariance*. In doing so we concentrate exclusively on three levels of measurement invariance, which are a necessarily to be tested in order to be able to compare mean scores across (cultural) groups. But for the sake of completeness, it should be mentioned that basically four levels of measurement invariance are differentiated. These hierarchically nested levels are *configural invariance* (structural equivalence), *metric invariance* (weak equivalence), *scalar invariance* (strong equivalence), and *residual invariance* (strict or full invariance) (Milfont & Fischer, 2010; Davidov et al., 2014; Gäde et al., 2020a; Leitgöb et al., 2023). As indicated before the psychometric assessment of MaC-DRS across the four target countries of our study focuses on *configural*, *metric* and

---

[49] Strictly speaking, the fit measures for metric and scalar invariance testing are slightly different with respect to the SRMR: *metric invariance* requires that the constrained (equal factor loadings) model, compared to the less constrained (configural) model, does not deviate more than 0.03 in the SRMR, while the model comparison for *scalar invariance* (equal factor loadings) compared to the metric model requires less than 0.015 change in the SRMR (cutoff value) (Cieciuch et al., 2019).

scalar invariance testing. Essentially, we follow a bottom-up approach to invariance testing from the least constrained to the most constrained (scalar) model (Rudnev et al., 2018).

Items that measure the same concept across contexts (groups) are interpretable as being free of item bias. Measuring the same concept across target contexts/groups is called *configural invariance* or structural equivalence. This is the most basic form of measurement equivalence. It is demonstrated when: "[a]n instrument administered in different cultural groups (…) measures the same construct(s) in all these groups" (van de Vijver & Leung, 2011, p. 20). Testing for structural equivalence across groups is testing for equality of the underlying dimensionality of a construct. Hence, to test for configural invariance we must constrain the model to have the same factor structure across groups. In other words: if the same factor structure of an instrument, where the same indicators are related to the same factors, is assessed across groups, configural invariance is demonstrated. This form of invariance is also called the baseline model as it is the least restricted model and the pre-condition for all further tests of measurement invariance.

> We base the new moral scale that we propose (MaC-DRS) on Moral Foundations Theory and Morality as Cooperation Theory. In the tradition of these theoretical approaches, we predict universalism of the 8 moral domains proposed. Furthermore, theoretically all societies depend on some degree of cooperation (Henrich, 2020). Also, empirical endeavors support the universalist notion of the MaC-DRS domains (Curry et al., 2019a; Cohn et al., 2019; Bjørnskov, 2021; Atari et al., 2022a). On this ground we expect to find structural equivalence — 8 MaC-DRS factors of first-order and presumably two to three factors of second-order — across the four target countries of our investigations. Albeit reason exists to expect universalism of the 8 MaC-DRS domains, we will empirically assess configural invariance of our construct using the method-based CFA approach. In the testing procedure we will run CFA´s separately for each country and inspect fit indices according to conventions. Thereafter, we will inspect equivalence of factor structure across countries via MGCFA.

Only if configural invariance is demonstrated we can proceed and move on to test the next level of invariance, that is *metric* or *weak invariance*. For this stage of invariance, all factor loadings are set equal as restriction, and the less constrained model is compared to the factor loading constrained model across groups. Factor loadings are indicative of the strength of the relations between manifest indicators (items) and latent constructs (factors). Thus, establishing metric invariance basically means "that each item contributes to the latent construct to a similar degree across groups" (Putnick & Bornstein, 2016, p. 75).

Metric invariance is demonstrated when the constrained model (same factor loadings) reflects the data equally well as the less restricted model. This is empirically shown by a comparison between the unrestricted baseline model and the factor loading restricted (metric)

model, and via a look at model fit indices. In case of an insignificant chi-square difference test and low CFI (decrease of less than 0.01), RMSEA (increase of less than 0.015), and SRMR (increase of less than 0.03) departure between models, metric invariance is established. As the chi-square test is unreliable due to sample size sensitivity it is recommended to rely more on the alternative fit indices (CFI, RMSEA and SRMR) in assessing invariance.

In the case we have been successful in demonstrating weak (metric) invariance, we can move on and test for *scalar* or *strong invariance*. To test for scalar invariance in addition to the equal factor loading constraint also all intercepts are constrained to be equal across groups. The intercept reflects the difficulty of an item and "refers to the expected value of the observed indicator when the score on the latent factor equals zero" (Davidov et al., 2014, p. 63). Again, the lesser constrained metric model (equal factor loadings) is compared to the scalar model imposed with more constrains (equal factor loadings and intercepts) across groups. A χ^2 difference test as well as alternative fit indices are inspected in regard to invariance conventions. To establish scalar invariance changes in alternative fit indices must be smaller than 0.01 for CFI, 0.015 for RMSEA, and 0.01 for SRMR when comparing the models. If and only if scalar invariance can be demonstrated we can meaningfully compare means across (cultural) groups.

> To examine MaC-DRS invariance we will step-wise test for configural (same factor structure), metric (equal factor loadings) and scalar invariance (equal intercepts) across Egypt, Germany, Japan and the United States using MGCFA. In doing so, we will follow the procedure described above. In the investigations beyond psychometric properties we aim, in the best-case scenario, to compare means of constructs across four countries. Scalar invariance is the prerequisite to be able to meaningfully compare mean values across groups. Hence, we will only assess the first three levels of measurement invariance and neither test strict (residual) invariance nor structural invariance types concerned only with latent factors.

2.5.8. Setting the Scale, Second-Order Factor Invariance and Within Country Invariance Test

Since latent factors are not by themselves parameterized, we need to set the scale for the metric and scalar invariance models. In doing so, we follow common practice: we set a manifest indicator per factor to 1 and the intercept of the same indicator to 0. This method is applied across all groups so that the factor corresponds to the scaling of the (marker) item (Putnick & Bornstein, 2016; Gäde et al., 2020a).

Moreover, we pose two to three higher-order factors for MaC-DRS and invariance for these factors must also be demonstrated. Therefore, we follow the advice by Rudnev et al., (2018) and test stepwise from configurational to scalar invariance for the hierarchical factor

levels of MaC-DRS. We begin with first-order invariance tests for the respective level of equivalence. This procedure is followed by a second-order invariance test for the same equivalence level, before moving on to the next comparison with more model constraints.[50] However, it is also possible that, for example, scalar invariance is not achieved. How to deal with problems of non-invariance and what strategies we apply if we encounter this problem is described in the following section.

2.5.9. Strategies and Approaches in The Light of Non-Invariance

Evidence of various empirical studies with comparative aims indicates that a higher level of invariance is rare or at least very difficult to achieve (Cieciuch et al., 2019). As a consequence, voices were raised arguing that the conditions for the invariance tests are too strict. In the light of this argument new conventions and methods are suggested as alternatives for classical (exact) invariance testing (Davidov et al., 2014). Though, also already commonly used approaches exist, partial invariance for instance, as means to handle non-invariance. In order to account for the empirical fact that studies are often not able to demonstrate e.g., full scalar invariance, implying that mean comparisons between groups are not possible, we will focus below on strategies when confronted with *non-invariance*. First, we touch on partial invariance. Second, we will shortly tackle the rather new approach of alignment optimization (Leitgöb et al., 2023). Third, as it is demonstrated that response styles are systematically related to culture and able to distort invariance testing (Cheung & Rensvold, 2000; Smith, 2004; He et al., 2021), we will also discuss post hoc ways to account for this sort of bias. Fourth, we suggest matching, item omission, sub-group comparisons, and a pan cultural explorative approach as strategies when only some items/groups demonstrate invariance or invariance testing completely fails.

There are several examples in the literature that show either a complete failure of invariance or the achievement of only a certain level of (full) invariance. Non-invariance on measures of values, gender role attitudes, or moral judgement for instance may be found across groups within cultures (e.g., between urban and rural inhabitants), across cultural groups, or across measures taken at different points in time (Davidov et al., 2014; Thomas et al., 2014; Schwab & Helm, 2015; Brown et al., 2017; Lomazzi, 2018; Cieciuch et al., 2019; Iurino & Saucier, 2020). Although metric or scalar invariance may not be supported by a respective

---

[50] We are also collecting data in urban and rural areas in each country of the cross-cultural study. Studies point to the fact, that measurement invariance must also be tested *within* cultures for the urban-rural distinction (Thomas et al., 2014). Should we identify problems with the factor structure or measurement invariance for one of the countries, we will investigate within country measurement invariance. In this case, we will also e.g. for within culture (urban vs. rural area) invariance of MaC-DRS. This approach may also help to identify sources of bias.

model, researchers have several options available to deal with the data, of which *partial invariance* is one (Putnick & Bornstein, 2016; Gäde et al., 2020a; Leitgöb et al., 2023). Partial invariance is a method in which parameter constrains are sequentially relaxed (backward approach) or added (forward approach) in order to reach at a partially invariant result. The approach is applicable for metric, scalar or higher invariance tests as well as for testing the invariance of first- or higher-order factors (Rudnev et al., 2018). Basically, partial invariance rests on the idea to work with a set of indicators that are invariant and some indicators which are non-invariant. For the latter indicators, the respective metric (factor loadings) or scalar (intercepts) equality constraint is abolished and the invariance test is rerun for restricted and unrestricted indicators. The literature recommends either having at least two invariant indicators, or that at least half of the indicators used for partial invariance models do not violate the assumptions (i.e. are invariant). The same suggestion applies for higher order factors (Rudnev et al., 2018). Furthermore, the relevant literature indicates that the minimum number of indicators for factors of first order is three. Consequently, working with the minimum number of indicators in partial invariance testing implies that at least two out of three indicators must be invariant. There are also indications that partial invariance at the metric level induces less bias than partial invariance at the scalar level. So, applying partial invariance should be less problematic when examining metric invariance. Partial invariance offers a possibility of not completely abandoning an analysis due to invariant results. However, the consequences of not fully invariant indicators are still being discussed and are not yet fully understood (Putnick & Bornstein, 2016; Cieciuch et al., 2019). Nevertheless, a recently conducted simulation study "concluded that partial invariance is sufficient under various condition and performs as well as other more recent approaches" (Leitgöb et al., 2023, p. 6).

> When we are confronted with non-invariance at the metric or scalar level in our MaC-DRS analysis, we will apply forward/additive partial invariance testing. Regarding the stepwise process, we follow the invariance testing procedure suggested by Milfont and Fischer (2010, p. 116). The application of the partial invariance method may be combined with the item omission strategy, which is discussed further below.

Alignment optimization is a rather new method of invariance testing that appears to be especially promising when confronted with non-invariance and large numbers of (e.g.) countries to be compared (Lomazzi, 2018; Leitgöb et al., 2023). This approach is different from classical invariance testing for it suggests that *approximate* rather than *exact* invariance might be sufficient. The alignment approach works with an optimization algorithm whose logic is comparable to factor rotation in EFA. This optimization algorithm is applied to a configural

model in order to search for "the most optimal measurement invariance pattern, in which a relatively small number of large noninvariant parameters—and many approximately invariant parameters—are present" (Cieciuch et al., 2019, pp. 164-165). In addition to approximate invariance optimization also non-invariant indicators can be identified within this method (Lomazzi, 2018). Furthermore, as a rule of thumb, a cutoff criterion for model evaluation is proposed. This criterion suggests that mean comparisons after alignment optimization are reliable when ≤ 20% of factor loadings and intercepts are invariant (Leitgöb et al., 2023). Alignment optimization and approximate invariance seem to be a promising path for future studies, in particular when the aim is to compare large number of cases. Nonetheless, as this approach is relatively new, its application in the literature is still rare. Also, the question of what degree of approximation to (non-)invariance enables reliable results remains an open debate to this day (Cieciuch et al., 2019).

> We consider applying the frequentist alignment optimization method under the condition that the (exact) partial invariance method does not provide MaC-DRS results that can be deemed invariant at the scalar level. When applied, the cutoff criterion of maximally 20-25% parameter non-invariance is used to evaluate if the alignment results allow reliable mean comparisons across groups. Alignment optimization will be potentially combined with the item omission strategy that is discussed below.

Response styles are known to affect invariance of instruments across groups (Cheung & Rensvold, 2000; He et al., 2021).[51] Beyond that, chances for the risk of scores being influenced by response styles increase with the cultural distance between the groups being compared (Thomas et al., 2014). Thus, as we selected the cases for our study based on the criterion of cultural differences on various dimensions, we must foresee the possibility of response style distorted scores. In what follows we will shortly and non-technically describe two approaches able to counter nuisance due to response style effects. The former approach deals with ipsative scores while the latter is about introducing a response style factor into one´s SEM invariance testing model.

Ipsatization is a process to standardize scores and suggested as a method to remove response style biases (Cheung & Chan, 2002; Fischer, 2004). Several forms of score standardization exist, yet here we are concerned with scores resulting from within-subject standardization, i.e., *ipsative scores* (Fischer & Milfont, 2010). To account for acquiescence bias in invariance testing it is suggested to subtract the individual mean from raw score data in order to obtain

---

[51] Response styles can also distort analyses that go beyond the examination of psychometric properties. To control for distorting RS effect, once justified reasons for the existence of response style biases are statistically given, RS variables can be created and included in appropriate models of analysis (Baumgartner & Weijters, 2015).

bias free ipsative data. Cheung and Chan (2002) give an example of how ipsatization works in regard to the elimination of acquiescent bias. We are taking up their example here.[52] Imagine we collected scale responses from two participants (participant A and B). Participant B´s responses are biased by a tendency to acquiesce and A´s responses are free of this bias. Imagine furthermore, that A´s scores on five items are 16, 19, 20, and 23, while B´s scores on the same items are consistently 3 units higher (i.e., 19, 22, 23, and 26) due to B´s tendency to acquiesce. Apparently, the score comparisons would be response style biased and the means of the two respondents on the same items are different (A: Ø = 20; B: Ø = 23). If we now transform the data, however, to ipsative scores via subtracting the individual mean from the individual raw scores, we obtain results that are freed from the acquiescence effect. After within-subject standardization both, A and B, would have identical scores of -4, -1, 0, 2 and 3. Additionally, if one also aims to partial out nuisance from extreme responding one may further divide the ipsative scores by the individual standard deviation across variables.

Overall, the use of ipsatization has increased in cross-cultural research and is, among other methods, demonstrated to be effective in removing certain response style biases from data (Fischer, 2004; Savalei & Falk, 2014). Albeit its usefulness, major drawbacks must be noted when working with ipsative scores. First one needs to account for the ipsative data structure, which is different from e.g., normative data. To obtain again data that can be tested for invariance and analyzed via (MG)CFA one must recover the factor structure by a complex procedure (Cheung & Chan, 2002; Chan, 2003). Second, due to the ipsatization of data the interpretation changes, as ipsative scores are reflective of "the relative position of the individual on a variable in relation to the other scores" (Fischer & Milfont, 2010, p. 92). So, one must pay additional attention to draw precise and adequate interpretations from ipsative data. Finally, we don't know the true score of a variable, and (cultural) groups may differ in communication styles and norms which in turn might affect responses. Thus, we cannot rest assured that the procedure of ipsatization only removes bias but not substantive differences across (cultural) groups (Welkenhuysen-Gybels et al., 2003; Smith, 2004; He & van de Vijver, 2012). Indeed, working with ipsative scores poses a way to deal with response style biases in cross-cultural research. However, the method comes with limitations making its application potentially costly. When working with ipsative scores, therefore, a solid justification for the choice of method is needed.

---

[52] The example is taken from Cheung and Chan (2002) and can be found in their article with a technical and non-technical explanation and demonstration of how ipsatization works as a means to reduce uniform response bias.

Considering the drawbacks, we consider the potential application of ipsatization in testing MaC-DRS for invariance only in two cases: a) when several other strategies fail and the analysis indicates a strong bias due to response style effects; b) in the case of another potential study of future investigations addressing the methodological interest of comparing several approaches to be used in the presence of non-invariance.

A rather direct approach to control and reduce acquiescence bias is to integrate a response *style factor* into a structural equation model (SEM) for invariance testing. This approach is based on the assumption that acquiescence is a personality trait associated with behavioral consistency. Furthermore, acquiescent responding is seen as a form- but not content-related response style bias. As a result of these assumptions, it is possible to correct the covariation in (content relevant) items caused by acquiescence using another (style) factor that has the same form (but is different in content). Watson (1992) applied this approach for an unbalanced scale. She used items from two constructs other than the content factor of interest to create an acquiescence (style) factor and controlled that the style factor effectively assesses acquiescent responding. Thereafter, she integrated the style factor in the invariance testing SEM model in order to capture "covariation among the items that is due to acquiescence" (p. 52). Welkenhuysen-Gybels et al., (2003) demonstrated the effectiveness of the style factor approach for a balanced set of items. Furthermore, Billiet and McClendon (2000) specified the conditions that must be met to be rest assured to have identified the style factor: ideally, the style factor should have a non-zero but lower variance than the content factor. In addition, the style factor should be based on at least two balanced and independent scales in order to capture only bias caused by acquiescence. Furthermore, it must also be empirically verified that the style factor measures acquiescence, as demonstrated by a high correlation between style factor and an acquiescence index.

The style factor approach comprises a component (i.e., the style factor) that is intended to capture covariance due to bias within the invariance test model. Compared to the ipsative score approach, which works with a data transformation, the SEM style factor approach therefore appears more direct. Additionally, once the style factor has been identified, the approach provides greater certainty than ipsatization that only group differences due to nuisance, but not substantial differences, are removed. Overall, the style factor approach has been shown to be able to successfully eliminate acquiescence bias from invariance tests. However, the application requires certain means (at least two independent sets of balanced items) to form and identify the style factor. In addition, several other model assumptions must be met, and the approach relies on the theoretical assumption that acquiescence is a behaviorally consistent personal trait,

a proposition that is still debated in the literature (Billiet & McClendon, 2000; Savalei & Falk, 2014; Bogner & Landrock, 2016).

> The SEM style factor approach appears tempting, yet comes with the requirement of two independent, balanced sets of items. Given economic constraints, limited questionnaire length, and the limited cognitive capacity of study participants, we were unable to include additional items on which a style factor could be based into our questionnaire. We will potentially examine yet whether we can build a meaningful style factor based on the balanced set of self-construal items that are part of our questionnaire. The latter considerations only become relevant, however, when we are confronted with RS problems and non-invariance in the examination of MaC-DRS.

Significant imbalances on measured sample characteristics may cause bias in group comparisons in general and likely also in invariance testing across groups.[53] The latter holds, provided that imbalanced variables between samples affect the scores on the instrument to be tested for invariance. Under the premise of significant sample differences (on sociodemographic variables for instance) one may enhance comparability of samples via the application of *matching procedures*. Admittedly, matching is seldomly seen in cross-cultural studies testing for invariance. Nonetheless, the procedures of propensity score matching (PSM) and doubly robust estimation, stemming from the field of causal analysis, are known to be able effectively reduce (overt) bias between samples (Harder et al., 2010; Funk et al., 2011; Morgan & Winship, 2015). Matching usually involves several steps: defining a measure of difference/similarity, the matching procedure itself, the evaluation of the matching model, and the estimation of effects. In regard to measurement invariance testing, one may return to MGCFA after matching and test for invariance again using the matched samples that are freed of measured sample imbalance. A drawback of this approach can be seen in the fact that the analysis based on matched samples is no longer based on "real" individuals, but on matched cases. Also, imbalances due to hidden bias may still distort comparisons, yet the approach of Rosenbaum bounds offers a means of sensitivity analysis to inspect the potential extent of this problem (DiPrete & Gangl, 2004; Rosenbaum, 2005; 2010).

> We may consider the application of PSM, doubly robust estimation and Rosenbaum bounds given we find significant imbalances between the samples.

---

[53] When going beyond invariance testing and significant sample differences have been identified, the application of (conditional) average marginal effects can be helpful to "control" for a set of covariates on the one hand and thus increase comparability between samples with regard to the effect of interest on the other (Williams, 2012; Wooldridge, 2016; Kohler & Kreuter, 2017).

Another and non-technical approach when confronted with non-invariance is to *omit items* (Putnick & Bornstein, 2016). The idea is to first identify items that are cause to the non-invariance, which can be done e.g., via analysis of variance and graphical plotting as demonstrated by van de Vijver and Leung (2011). Once disturbing items are identified, they can be removed from the model to increase the chances of measurement equivalence when the model is run again. However, a drawback of this approach is that the breadth of the construct being measured may decrease as items are dropped. Therefore, the decision to omit items must also be theory-driven, and a sufficient number of items (at least three per factor) must remain.

> In the cross-cultural data collection, we use a version of MaC-DRS that is designed to capture each theoretical dimension using 4 items. However, only 3 items are necessary to capture a factor. If MaC-DRS is tested non-invariant, we consider omitting one item per dimension and re-run MGCFA to test for full or partial invariance. We base the application of the item omission strategy on the following conditions: we identify items responsible for non-invariance and half of the items (the concept consists of 3 items per factor after omitting one item per factor), i.e., at least 2 out of 3 items are invariant.

Rather than modifying the data structure (ipsatization) or omitting indicators when confronted with non-invariance one optionally may proceed with *sub-group comparisons* (Welkenhuysen-Gybels et al., 2003). This approach is relatively straight forward: only those groups for which invariance is proven are compared with one another. Though, this approach may become more difficult when the number of groups to be compared increases. For the latter case sub-group detection algorithms already exist and can be applied. Apparently, the sub-group approach entails loss of information as some groups are left out from the analysis. This must be considered a limitation of this approach.

> We consider sub-group comparisons in concert with other approaches as an additional strategy when we fail to establish (partial) invariance across the four (cultural) groups of our study.

Lastly, if invariance tests completely fail one may take a step back and use the data to refine the theory and measurement instrument. A purely *data driven pan cultural approach* using EFA might be the choice of method when confronted with this situation (Brown et al., 2017; Brandt, 2020). Given the illustrated situation one may: a) split the cultural samples into two samples (of equal number), i.e., sample X and Sample Y, b) merge all cultural samples (for Sample X and Y), c) run an EFA in Sample X to search for the factor structure that (may) appears across cultures, and d) run a CFA in sample Y to inspect the pan-cultural structure coming from EFA models regarding model fit. Based on this explorative approach one may generate new hypotheses to be tested in further studies.

The exploratory approach will be applied by us when we fail to demonstrate MaC-DRS invariance by the classical testing procedure and under the application of the various other approaches discussed above.

After these methodological remarks, we will in the next section briefly look back at the chapter at hand and provide a summary.

2.6. Summary: Setting the Stage to Cross-Cultural Investigations of the Human moral Mind

As we have described, three primary data collections form the basis of our empirical analyses, with the first two being referred to exclusively in the context of the development of the moral scale that we propose. Although we will discuss some aspects of each of the three data collections in the further course, we cannot provide detailed insights due to space constraints in this text. For this reason, we refer interested readers to the preregistered research plans for more detailed information. The research plans for each data collection are available online. Starting in the following chapter, the four empirical investigations of human moral mind, that we have already broadly sketched in this section, will be examined in detail. Furthermore, we have explained the case selection for the third, cross-cultural data collection, which stands at the center of our empirical research. Based on a contrastive case selection strategy, which, in an ideal-typical way, aims to capture cultural entities with either a predominantly group-centered social orientation or a predominantly individual-centered social orientation, we have selected Egypt, Germany, Japan, and the United States of America as cultural samples for our third data collection and subsequent analyses. We then integrated the countries to be examined into our theoretical model and transformed it into a research model. Moreover, we have seen that cross-cultural data collection cannot be approached lightly. Even the design of such a data collection requires the researchers to think carefully and plan extensively before the field phase. Various pitfalls lie in wait and can compromise the data collection itself, leading to inaccurate sources of information. Artifacts, but not insights, can result from inaccurate data. The aspect of translating measurement instruments is certainly one of the most important aspects to consider. Not only can emic and etic perspectives be distinguished in this context (Boehnke, 2022a), but in addition to the semantic accuracy of translations, the perspective of the cultural entities under investigation must always be taken into account — for example, whether an item can be asked in the same way in, say, culture A and culture B is mediated by the context, because sociocultural contexts are normatively structured in themselves. This becomes apparent, for

example, from the fact that we ask the item of gender in Egypt with two response categories, but in the other cultural entities of our study with three response categories. In the Egyptian context, the question of non-binary gender identification could be perceived as an affront and may result in an increased number of participants dropping out of the study. It is not for us to evaluate this, but it is for us to ensure that potential triggers of selective non-response (Legewie, 2012; Morgan & Winship, 2015) are kept to a minimum. The diverse background of our supporting research team enabled us to create culturally sensitive translations of the applied measurement instruments by drawing on their culture-specific knowledge. In this chapter, we have not only briefly presented our three primary data collections, emphasizing in particular the cross-cultural data collection, but also highlighted pitfalls in data collection across multiple cultural contexts. In addition to these descriptive components, we also presented our strategies, which we hope will help us to collect sufficiently reliable data in our cross-cultural endeavor. As we will see in the further course of this work, despite our careful planning, other factors are also relevant for data quality itself; the subjects in the samples are the source of the information to be analyzed and also a factor that can come with unpredictabilities that can only be revealed by analyzing the data. In addition to the planning-related aspects, we have also emphasized the importance of empirically testing the equivalence of measurement instruments across different groups. The comparative perspective that we have built theoretically through our hypotheses and which we will put into practice in the empirical analysis of the data is the subject of scientific debate and different positions. Here, we take a psychometric position that does not presuppose equivalence, but rather makes equivalence itself the analysandum before any further investigation. We see ourselves rooted in the methodological tradition of cross-cultural research and therefore feel compelled not only to try to minimize sources of bias as much as possible in the study design, but also to test empirically whether sufficient conditions are met to allow us to make certain empirical comparisons between the cultural entities we are analyzing. The latter is all the more important because we propose a new scale, the *Morality as Cooperation–Deviance Relevance Scale* (MaC-DRS), which is not established and still has to be validated. Therefore, after our explanations here, we will discuss the development and detailed psychometric testing of the new morality scale in the following chapter. In the present section, we have presented the methodological equipment that we need to address the psychometric properties of MaC-DRS not only in intra-cultural but also in cross-cultural research. Following the chronology, after our excursion into methodological realms, we will subsequently test the empirical properties of the Morality as Cooperation-Deviance Relevance Scale before turning to the primarily substantive investigations thereafter.

Chapter 3: Investigations of the Human Moral Mind I

3.1. MaC-DRS — Developing and Testing a Scale on Moral Deviance Relevance

For most people, supporting their family is certainly relevant. But is it more relevant than fair behavior? Is it intuitively more important for all people across cultures to be loyal to their group of friends, or to treat another person's property with care? Among others, *family*, *fairness*, *in-group* and *property* are domains of cooperation that are regulated by our evolved moral mind, as proposed by the *Moral Foundation Theory (MFT)* (Haidt, 2001; Haidt & Joseph, 2007; Graham et al., 2013; Atari et al., 2022a) and the *Morality as Cooperation Theory (MaC)* (Curry, 2016; Curry et al., 2019a; 2019b). MFT and MaC are two theories in the field of moral psychology that advocate moral pluralism. These theories are based on a gene-culture co-evolutionary background (Henrich & McElreath, 2007; Richerson et al, 2010; Boyd et al, 2011; Chudek et al, 2016; Brown et al, 2022), consider morality as a *universal* facet of our species and yet hold that *cross-cultural differences* in moral domain endorsement exist.

MFT and MaC scholars have developed self-report instruments to measure *moral judgement* and *moral relevance* (Graham et al., 2011; Curry et al., 2019b; Atari et al., 2022a). In particular, the Moral Foundations Questionnaire-1 (MFQ-1) has been widely used over the last decade and can be considered one of the leading scales in moral research (Ellemers et al., 2019). However, MFT has come under criticism, as has the scale that emerged from this approach (Skitka & Conway, 2019; Curry et al., 2019a; 2019b; Iurino & Saucier, 2020). Scholars from MFT responded to some of the criticisms, refined the theory and developed a new scale, the MFQ-2 (Atari et al., 2022a). This new instrument, however, no longer measures moral relevance, but is limited exclusively to moral judgement. Turning to MaC, we find that this approach overcomes some of the theoretical hurdles of MFT (Curry, 2016; Curry et al., 2019b). Nevertheless, there are also problems with MaC, both in theory and to some extent in measurement, as we will demonstrate in the further course. Thus, there is a gap in moral research when it comes to measuring the *relevance* of different moral domains.

But why focus on *relevance*? We argue that while we can evaluate something as wrong, this does not per se mean that it is also subjectively evaluated as relevant. What is more, both the Moral Foundations Theory and the Morality as Cooperation Theory are approaches that explicitly address the dyadic link between the *universal* and the *culturally specific*. We see ourselves in the tradition of these approaches. In the context of our project, we will initially

address one side of the coin, moral universalism. In a later chapter, we will turn the coin over to the other side and also address cultural differences in relation to human morality. In the context of moral universalism and cultural differences, the aspect of *relevance* itself becomes relevant. The background to this is the assumption that our morality has evolved because human populations have faced recurring social problems throughout their history (Haidt & Joseph, 2007; Pietraszewski, 2016). These recurring challenges, which are related to human cooperation (Tomasello & Vaish, 2013; Henrich & Muthukrishna, 2021), are indirectly the focus of our research. Because these problems were so fundamental to the survival and reproduction of our species, they also provide the starting point for the assumption of the universalism of the human moral mind. But and this part is also important, as human development progresses and societies are forged, different societies can be expected to have faced, and likely continue to face, varying degrees of problems of cooperation (Curry, 2016). The MaC approach starts from an explicit assumption here: culture-specific prioritization of moral domains should correspond to the extent to which societies have been or are confronted with specific, recurring problems of cooperation. In terms of relevance, it can be deduced that different moral domains should be differently relevant in different cultures, depending on the extent to which domain-specific problems have played a role in the history of a particular cultural entity. Indeed, moral judgment can also be taken into account in this context, since we can also make gradations about the extent to which an action is good or bad. However, as we will elaborate below, we predict that although an (im)moral action can be judged as bad/good, this does not per se or fully capture the degree to which the action is valuated as subjectively relevant. However, according to our argument, subjective relevance comes closer to what Oliver Scott Curry (2016) predicts in the context of the Morality as Cooperation Theory, that is, cross-cultural differences through culture-specific *prioritization* of moral domains. According to this argument, a moral domain should be subjectively more relevant the more it is culturally prioritized, and it should be culturally prioritized the more there was a need for regulation of the moral domain throughout the history of the cultural entity due to recurring problems related to this domain. Furthermore, we argue that an action can be judged as morally wrong, for example, but at the same time be regarded as subjectively irrelevant. We therefore believe that the aspect of moral relevance could add a facet to the study of the human moral mind that cannot be fully captured by focusing exclusively on judgments — moral relevance, from our understanding, comes closer to what corresponds to the core of society's involvement with a recurring problem of cooperation. In the following, we will refer to the aspect of moral relevance at various points in this chapter and the course of this text.

Furthermore, we have identified a gap in the theory of both MFT and MaC, which is also reflected in the corresponding instruments of these approaches: both approaches do not distinguish between moral *conformity* and moral *deviance*. However, we argue that such a distinction is integral and should also be taken into account in the measurement of morality (Kahneman & Tversky, 1979; Baumeister et al., 2001; Kurzban et al., 2001; Chudek & Henrich, 2011).

We aim to close the addressed gaps by proposing a new self-report instrument — the **Morality as Cooperation—Deviance relevance Scale** (**MaC-DRS**). Our scale is based on a combination and partial extension of MFT and MaC. We place human cooperation at the center of morality (Tomasello & Vaish, 2013; Tomasello, 2017) and assume in line with the former theories moral pluralism. Different to MFT and MaC we argue that all moral domains are traversed by a single *guiding principle*. This principle is concerned with moral *harm* and moral *care* regardless of the domain in question, and is expressed in either moral *deviance* or moral *conformity*.

We assume moral pluralism. Against this background, and following MFT and MaC, we propose that our evolved moral mind perceives and recognizes at least 8 different moral domains as such. Our approach is explicitly anchored in the context of a gene-culture co-evolutionary framework. Thus, we hold that morality is a universally evolved component of the human mind. As part of this, we predict that the 8 moral domains we propose have evolved across cultures. However, since we also explicitly refer to cultural influences from a theoretical perspective, we further expect a culturally contingent calibrated moral mind in addition to the *universalism hypothesis*.

In this chapter, the steps taken to develop MaC-DRS are presented and further deepened through comprehensive psychometric analyses. We will focus on three studies. In these studies, we will develop and psychometrically analyze MaC-DRS. Furthermore, we will also compare the new morality instrument that we propose with the established MFT and MaC scales (Graham et al., 2011; Curry et al., 2019b). Finally, we will broaden the perspective on MaC-DRS by analyzing the psychometric properties of this scale across cultures. What is more, in the third study we will approach the universalism hypothesis by investigating whether we find evidence of an 8-dimensional moral structure across cultures. The main background features of the studies conducted are listed below.

Study 1 (data collection 1): Online questionnaire design, German student sample, N = 792 — MaC-DRS development study.

Study 2 (data collection 2): Online questionnaire design, nationwide non-student sample in Germany, N = 2,326 — MaC-DRS validation study.

Study 3 (data collection 3): Semi-experimental online questionnaire design, cross-cultural non-student sample comprising data from Egypt, Germany, Japan, and the Unted States of America, N = 2,982 — cross-cultural substantiation study.

In addition to MaC-DRS, the psychometric properties of the CIRN-Self-Construal Scale-3 (Vignoles et al., 2016; Yang, 2018; Uskul et al., 2023) are also examined across cultures (*Study 3*). We are testing this scale alongside MaC-DRS in order to prepare the further analyses that we will carry out in another chapter below. After commenting on the results that we found in relation to the self-construal scale, the present section concludes with a comprehensive discussion highlighting the relevant findings of all three studies conducted, pointing out the limitations of our studies, and making suggestions for future areas of research.

With the findings presented in the further course, we will demonstrate that MaC-DRS is able to measure deviance relevance of 8 moral domains — a broader coverage of moral domains than previously possible — validly and reliably within as well as across different cultures. Accordingly, we will provide strong indications in support of the universalism hypothesis of human morality. In contrast to the MFT and MaC scales, there is further evidence that MaC-DRS is also able to capture higher-order moral constructs (Haidt, 2008). Furthermore, an overview of the analyses that we will present shows that the Morality as Cooperation—Deviance Relevance Scale performs better than the established MFT and MaC instruments in all psychometric properties tested. In the light of empirical evidence, we therefore propose MaC-DRS as a new self-report instrument that enables *deviance relevance* to be measured accurately across different moral domains and cultures. The insights to be presented are important both for the remainder of the present project and for the field of research on human morality in general.

3.2. Integrating and Expanding MFT and MaC — The Theoretical Background of MaC-DRS

In order to avoid too much redundancy, we will only recall the key points of MaC-DRS's theoretical background and refer the reader to **Chapter 1** for a detailed explanation of our theoretical position.

MFT and MaC are pluralistic moral theories based on a cultural-evolutionary framework (Haidt, 2001; 2003; 2008; Graham et al., 2009; 2011; 2013; 2016; Curry, 2016; Mooijman et al., 2017; Enke, 2019; Curry et al., 2019a; 2019b; Atari et al., 2022a). They share the notion of (universal) moral foundations/domains and are yet sensitive to the effects of different socio-cultural context. Furthermore, they emphasize moral intuition without denying deliberate moral cognition. Overall, and despite some differences, the Moral Foundations Theory (MFT) and the Morality as Cooperation Theory (MaC) reveal more similarities than discrepancies. MaC-DRS builds upon, practically integrates and partially expands these theoretical strands into a new measurement instrument.

The moral approach that we are advocating sees itself mainly in the tradition of Haidt's (2001) intuitionist model, but recognizes that deliberate moral cognition can also play an important role (Ellemers et al., 2019). In line with MaC, our theoretical position places *cooperation* at the center of morality and follows the game-theoretical definition of cooperation as a non-zero-sum (win-win) interaction between actors (Esser, 2002; Diekmann, 2013; Tomasello & Vaish, 2013; Curry, 2016).

Moreover, based on insights from other studies (see e.g.: Kahneman & Tversky, 1979; Baumeister et al., 2001; Kurzban et al., 2001; Tooby & Cosmides, 2010; Chudek & Henrich, 2011; Schreiber & Iacoboni, 2012; Kaplan et al., 2016), we claim that the human moral mind is designed by evolution to pay particular attention to moral breaches i.e., to acts of moral *deviance*. On the basis of a moral approach that focuses on cooperation, we treat *moral deviance* as equivalent to losses, costs and ultimately *harm*. In contrast, we understand *moral conformity* as the realization of (mutual) cooperative gain of any kind of resources and ultimately as equivalent to *care*. Based on these assumptions, our position thus focuses on a theoretical supplement to MFT and MaC by emphasizing the difference between moral conformity and moral deviance and assuming a special calibration of the human moral mind with regard to the latter. In the light of these considerations, MaC-DRS will focus solely on moral *deviance* and leaves out the aspect of moral conformity.

In regard to moral domains, we suggest to treat *fairness*, *property*, *heroism*, *reciprocity*, *family*, *in-group*, *deference* and *trustworthiness* as distinct domains of cooperation that are regulated by our moral mind. This domain mapping embraces not only key ideas of MaC (Curry, 2016; Curry et al., 2019a; 2019b) but in large parts also the foundations suggested by MFT (Haidt & Joseph, 2007; Graham et al., 2013; Atari et al., 2022a).[54] Furthermore, we derived the *trustworthiness* domain from other studies (see e.g.: Henrich, 2009; Carter & Weber, 2010; Tomasello & Vaish, 2013; Diekmann & Lindenberg, 2015; Van Lange, 2015; Cohn et al, 2019; Enke, 2019; Bjørnskov, 2021; Muthukrishna et al, 2020; Kirkland et al, 2023) and propose this domain as a new and previously under-researched moral domain. Taken together, we predict that the human moral mind universally embraces these 8 domains as common elements of morality. In light of the considerations on moral universality, we will test several *hypotheses* in the cross-cultural study presented below. Furthermore, we would like to emphasize that we neither claim to provide a comprehensive list nor that we consider the list of proposed moral domains to be exhaustive.

Eventually, we also follow the idea of Haidt (2008) who suggested a *binding* approach to morality and an *individualizing* approach to morality. Theoretically we hold that fairness, trustworthiness and property correspondent to individualizing morality and center around individual-focused cooperation. Additionally, we suggest that family, in-group and deference form the domains of the higher-order binding approach to morality and promote mainly group-focused cooperation. As far as reciprocity and heroism are concerned, we suspect that they move in the space between binding and individualizing and serve the core aspects of both higher-order approaches to morality. All in all, we will design MaC-DRS to capture *deviance relevance* towards 8 moral domains and at least two higher-order moral approaches.

3.3. Self-Report Measures of Moral Relevance

There are several ways to capture moral domains empirically (Graham et al., 2011; Clifford et al., 2015; Curry et al., 2019b; Ellemers et al., 2019; Atari et al., 2022a). However, although frequently used, the MFT and MaC self-report measures suffer, among other problematic areas, from non-reproducible factor structures, a lack of measurement invariance across cultures and problematic (ambiguous and context-bounded) item-formulations (Curry et al., 2019b; Iurino

---

[54] As mentioned elsewhere, we do not adopt the MFT foundation of *purity*. Nonetheless, we are open to integrate purity and other moral domains in the future as long as we can envision their potential from a perspective of morality that centers around cooperation.

& Saucier, 2020). And most importantly, neither the theoretical pillars on which the scales are based nor their operationalization differentiate between *moral conformity* and *moral deviance*. Given these limitations, we intend to close the gap of theoretically sound, valid and reliable self-report instruments in moral relevance research and propose the **Morality as Cooperation—Deviance Relevance Scale** (MaC-DRS) to overcome this shortcoming.

3.3.1. Moral Relevance is Not Moral Judgment

We argue, that to capture the *relevance* of moral domains is not to be confused with the assessment of *moral judgment*. Moral judgments are broadly defined as "evaluations (good vs. bad) of the actions or character of a person that are made with respect to a set of virtues held to be obligatory by a culture or subculture" (Haidt, 2001, p. 1028; for a recent overview on moral judgment see: Malle, 2021). In contrast, we propose to understand *moral relevance* as a person's intuitive preference order in relation to various moral domains held to be obligatory by a culture or subculture. Morally *relevant* actions are therefore actions that are of immediate and intuitive importance to a person; the higher the subjective relevance, the higher the corresponding moral domain is prioritized. These actions also differ from morally *irrelevant* actions due to their intuitive character of importance, as in the case of the latter, the perceived behavior is morally charged as relevant only with difficulty and under the influence of reflection and cognitive load, or not at all. In the case of irrelevance, the regulation of the respective moral domain would then not be given much priority by a culture or subculture, and hence also not be given much individual priority. MFT also deals with moral relevance, and their conceptualization comes close to our idea in part: moral relevance is about "individual differences in the range of concerns that people consider morally relevant" (Graham et al., 2011, p. 6). However, MFT further argues that assessing moral relevance is about assessing abstract self-theories that may not fit well with moral judgment. We, by contrast, believe that moral relevance also involves a strong intuitive element, but adopt the assumption that relevance and judgment do not per se have to coincide.

We predict that actions can be judged as morally wrong by people, but this does not mean that they are equally relevant for everyone. On reflection, we can probably morally charge most of the behaviors we perceive in some way (Haidt, 2003). However, a spontaneous reaction, an automatic and uncontrollable feeling of perceiving a moral confirmation or violation, may not be shared to the same extent across people because different cultural groups may not hold and enforce certain moral values and imperatives as much as others. What we are saying is that

while we may judge certain behaviors to be wrong, this does not necessarily mean that the behavior in question has the same intuitive relevance for people. Although moral judgement and moral relevance have much in common, they are not the same (Graham et al., 2009). Consequently, we argue that moral relevance and moral judgment may touch on partly different aspects of morality and are not necessarily equivalent in terms of the extent attributed to a perceived moral action. That is, the perception of a moral action can lead to a degree of relevance attribution for this action that may differ from the degree of judgment of the corresponding action as right or wrong.

3.3.2. Re-Designing a Scale to Measure the Relevance of Multiple Moral Domains

The *Morality as Cooperation—Deviance Relevance Scale* (MaC-DRS) rests on our theoretical propositions. Our operationalizations of the scale´s constructs followed the state of the art (Lenzner & Menold, 2015; Moosbrugger & Brandt, 2020) and attempted to capture and combine what MFT and MaC have already proposed with their scales (Graham et al., 2011; Curry et al., 2019b, Atari et al., 2022a). To measure *moral relevance*, we set the response format to a 7-point scale, ranging from (1) "*extremely irrelevant*" to (7) "*extremely relevant*", with a neutral response option (4) ("*neither irrelevant nor relevant*") in the middle (Menold & Bogner, 2015).[55]

We designed MaC-DRS to measure 8 moral domains. First, we adopted the items of the 7 domains from the MaC scale (Curry et al., 2019b) and reformulated their operationalizations. We then integrated the MFT *loyalty* foundation into the *family* and *(in-)group* items, and the *authority* foundation into our operationalizations of *deference*. Additionally, by incorporating *equity* and *proportionality* (Atari et al., 2022a) into the *fairness* domain, we considered two relatively new MFT propositions yet without treating them separately. Our scale also adopts *heroism* from MaC. To operationalize *heroism,* we distinguished between a) overcoming natural (adaptive) fear and b) heroic behavior, i.e., behavior that seeks to avert harm to others (family, in-group, strangers) and at the same time poses a risk to the heroic person themself. We designed the heroism items to capture only the latter type of behavior.

Considering the direction of moral behavior in the operationalizations, we focus with MaC-DRS exclusively on *morally deviant behavior*. Developing a scale that measures the relevance of moral *conformity* is thus beyond the scope of the present study. The items of our

---

[55] Note: All gradations of the response format are labeled.

150

scale are constructed in a decontextualized way to avoid context effects. We further tried to make the domain specific items as disjunctive as possible, and to keep them gender-neutral. Cognitive load (Paas et al., 2003) and comprehensibility were also taken into account: we formulated the items as simply and clearly as possible. Eventually, we came up with a preliminary set of 48 items, consisting of 6 items per domain. Based on empirical EFA and CFA results, we will later reduce this number of items to a more economical scale encompassing 3 items per domain.

The items went through several rounds of revision, face-validity checks, and discussions among project collaborators. Above that, cognitive pre-tests (think aloud method) were conducted, and a GESIS psychometric expert on item-construction was consulted for advice (Prüfer & Rexroth, 2005; Lenzner & Menold, 2015; Lenzner et al., 2015; Höfling & Moosbrugger, 2020; Brandt & Moosbrugger, 2020; Konrad, 2020).[56] The cognitive pre-tests revealed that morality itself appears to the respondents as an abstract concept that is difficult to grasp without further ado (Skitka & Conway, 2019). Accordingly, it became evident that respondents would benefit from a definition of moral relevance and moral irrelevance. Thus, we have developed a more detailed introductory text and definitions of moral ir-/relevance, and suggest showing them to respondents before measuring MaC-DRS.

The pre-tests indicated furthermore that a specific item-order would be beneficial. The *fairness* items seemed especially accessible for respondents. This was unsurprising since various studies support that fairness tends to be a universal human principle deeply rooted in our evolutionary history (Brosnan & De Waal, 2003; Henrich et al., 2005; Blake et al., 2015; Sommerville & Enright, 2018). We suggest to capture fairness items first before measuring the other MaC-DRS domains suggested. Moreover, the *family* and *in-group* domains were given specific stimulus texts to explain them to respondents. These items should be placed at the end of any MaC-DRS measure so as not to confuse respondents with the different social relationships. In the next step, we subjected our items to a first quantitative test to obtain empirical information on their validity and reliability (Moosbrugger & Kelava, 2020). But before we get to Study 1, we would like to give exemplary insights into the MaC-DRS items that we have developed. Respective insights can be obtained from ***Table 11***.

---

[56] We would like to express our gratitude to Dr. Ranjit K. Singh of the GESIS Leibniz Institute for the Social Sciences for his extremely valuable comments and advices.

Table 11: Exemplary insights into MaC-DRS

Introductory Text MaC-DRS	Based on our own **morality** we judge whether our own actions or those of other people are right or wrong. The subsequent questions ask you to what extent you consider certain **actions** to be **morally relevant**. **Morally relevant** actions are actions for which you would make a moral judgment quite spontaneously. In other words, actions that immediately appear morally *right* or *wrong* to you. **Morally irrelevant** actions, on the other hand, are actions for which thoughts or feelings about whether the action is morally right or wrong do not even occur to you. In other words, actions for which no moral evaluation comes to your mind. To indicate the extent to which an action is morally relevant to you, please use the scale that ranges from "Extremely irrelevant" (left end of the scale) to "Extremely relevant" (right end of the scale). This is about your personal impression. Answer spontaneously.

Fairness	**Morally relevant** actions are actions that immediately appear morally *right* or *wrong* to you. **Morally irrelevant** actions, in comparison, are actions for which thoughts or feelings about whether the action is morally right or wrong do not even occur to you.
	
	To what extent do you find the following actions morally relevant?
	
	   – Someone does not reward other people based on their performance.
	   – Someone takes a larger share of a jointly earned profit than others.

Trustworthiness	– Someone betrays others. – Someone does not keep their word.
Reciprocity	– Someone does not respond to help from others. – Someone enjoys favors from others without reciprocating.
Deference	– Someone defies a widely-respected person. – Someone does not behave according to their social position.
Property	– Someone handles another person's property carelessly. – Someone damages another person's property without replacing it.
Heroism	– Someone does not stand up for the physical and psychological integrity of strangers out of fear. – Someone does not protect their own family from harm out of fear.

Family	Now think about your circle of **family** and **relatives**. To what extent do you find the following actions morally relevant?
	
	   – Someone does not take care of their own family and kin.
	   – Someone ignores the wishes of their own family and kin.

In-Group	Please think now of **groups** of people with whom you are not related by family but to whom you belong or feel you belong (e.g., circles of friends, clubs, fellowships). To what extent do you find the following actions morally relevant?
	
	   – Someone does not defend the views of their group to outsiders.
	   – Someone does not give preferential treatment to members of their group over outsiders.

3.4. Study 1 — Development Study

In *Study 1* we decided for an online-questionnaire design and convenient sampling by reaching out to German university students. Our questionnaire measured all of the 48 initial MaC-DRS items, sociodemographic variables and some other constructs not detailed here. In total we were able to gather data from N = 792 respondents that answered all MaC-DRS items.[57] We have developed our new scale to measure 8 distinct moral domains as first-order factors. Additionally, we aim to be able to capture binding and individualizing as higher-order constructs with MaC-DRS. The overarching goal of *Study 1* is thus to examine validity and reliability of our new scale using model-based methods of exploratory (EFA) and confirmatory factor analysis (CFA) (Bagozzi & Yi, 2012; Brandt, 2020; Gäde et al., 2020a; Gäde et al., 2020b; Schermelleh-Engel & Gäde, 2020). As another goal we aim to reduce the preliminary item set.

3.4.1. Participants

Our sample comprises N = 792 German university students that have answered all MaC-DRS items.[58] The mean *age* of respondents is μ = 23.93 (Median = 23). In terms of *gender* females make up to 68.63% of the sample. The sample is heavily biased towards *higher education*, due to our target group: 98.57% of the sample have a university entrance qualification (i.e., German "Abitur/Hochschulreife").[59]

3.4.2. Exploratory Factor Analysis (EFA) — Results Study 1

Using Exploratory Factor Analysis (EFA) we first want to examine whether the 8-dimensional factor structure of our instrument appears from the data gathered. For reasons of robustness, we decided to carry out the analyses for four different sample sizes: 1) N = 792, i.e., the full number of cases that completed all MaC-DRS items; 2) n = 628, i.e., full number of cases that completed

---

[57] We would like to express our sincere thanks to Mr. D. Schulte am Hülse, who took care of the online setup of our data collection. Also, we would like to thank all those who have supported the data collection and are especially thankful to Prof. Dr. B. Bleyer and Prof. Dr. Jens Luedtke for helpful feedback.

[58] Note: N = 792 responded to all MaC-DRS items. However, the descriptive statistics are given for n = 628, i.e., for those cases that completed the entire study, which also includes the measurement of other variables. Additionally, response rates to single socio-demographic items can vary because participants received, after viewing a socio- demographic item for five seconds, the option to skip the item.

[59] We used the statistic program Stata13 for all analysis in Study 1.

the entire questionnaire of our study; 3) n = 574, i.e., an adjusted sample of cases that completed the entire questionnaire; and 4) n = 287, i.e., a split half sample.[60]

We tested MaC-DRS items at first for multivariate normal distribution. The results showed a skewed distribution which impacted our EFA decision. Thus, we refrained from an EFA using Maximum-Likelihood (ML-EFA) and worked in all our EFA models with principal factor analysis (PFA) (Brandt, 2020).

We went on to examine the correlation between items that we considered to belong to a common dimension. All correlations are significant at the 0.000 level. Subsequently we inspected if MaC-DRS items are unidimensional and if they have a sufficient sampling adequacy (Kaiser, 1974). KMO results are (overall) = 0.9308, which is marvelous by criteria. Thereafter, we proceeded with model-based tests and eventually inspected the data using EFA (Bagozzi & Yi, 2012). Altogether, we conducted multiple EFA models with different rotation algorithms (varimax, quartimax and direct oblique oblimin) and the four different sample sizes described above. The EFA results for the 48 item-set support across models an eight-factorial solution. Thus, eight factors have an Eigenvalue of ≥ 1.000, an item-count per factor of above three, and sufficient factor loadings (≥ 0.40) for most items.

Based on these findings we discarded in the next step those items with lowest factor loading, and semantic redundancy to reduce the item-set from 6 to four items per domain (i.e., 32 in total). Subsequently, we conducted the exploratory analyses again with the reduced item set. The EFA results support again an eight-dimensional solution, an item count of above three per factor, and sufficient factor loadings of all 32 manifest indicators. The results hold across rotation algorithms and sample sizes. Based on these initial findings, we went on to test MaC-DRS in the next step using confirmatory factor analysis (CFA).[61]

3.4.3. Confirmatory Factor Analysis (CFA) — Results Study 1

Jonathan Haidt proposed within the MFT-framework the idea of two different (higher-order) approaches to morality which he called *binding* and *individualizing* morality (Haidt, 2008). We value the idea for their parsimoniousness and wanted to inspect if MaC-DRS can assess and differentiate between at least two higher order moral constructs. To do so we inspected at first

---

[60] Tables for all EFA´s can be found in the **Appendix**. Note: in the adjusted sample (n = 574), we have removed cases with completely implausible response behavior. In addition, we have split the adjusted sample into two independent samples for reasons of robustness, so that we have an independent sample of n = 287 for each of the model-based EFA´s and CFA´s.

[61] We have conducted an extensive exploratory analysis of the MaC-DRS data from data collection 1. The corresponding EFA models can be found in the **Appendix.**

the correlational patter of the 8 factors that we obtained via EFA. This analysis suggests that the domains closest to MFT´s original binding morality conception — i.e., *deference*, *family* and *group* — tend to cluster. Contrasting to the binding domains, *fairness*, *trustworthiness* and *property* seem also to cluster and correspond from a theoretical perspective to individualizing morality. *Reciprocity* and *heroism* show a slight tendency to fall between binding and individualizing, as they correlate moderately positive with all other moral domains (range of correlation: 0.18 – 0.40). This seems plausible from a theoretical point of view, as both domains can be thought of to serve group-oriented, binding cooperation as well as individual-oriented, individualizing cooperation (Mauss, 1968; Moebius, 2006; Leung & Cohen, 2011; Uskul et al., 2023). Nevertheless, reciprocity and heroism show a slight tendency towards binding morality in this first study.

Next, we conducted three different hierarchical CFA-models.[62] *Model 1* assumes binding and individualizing as second-order factors and assigns factors of first-order to them based on the correlational pattern discussed above. Thus, we assigned fairness, trustworthiness and property to the individualizing second-order factor and the other dimensions to the higher-order construct binding morality. *Model 2* mirrors Model 1 except for the first-order factor reciprocity which we assume here, in line with MFT (Graham et al., 2011), to load onto individualizing morality. For *Model 3* we assume again 8 factors of first-order, yet only one common factor of second-order. We will evaluate and compare the theoretical models in terms of goodness of fit indices following standard criteria: TLI and CFI $\geq$ 0.90/0.95; SRMR $\leq$ 0.05/0.10 and RMSEA $\leq$ 0.05/0.08, to check whether they fit to the empirical data (Bagozzi & Yi, 2012; Gäde et al., 2020a). Lastly, we decided to use *modification indices* in all models for reasons of fit improvement. ***Table 12*** illustrates our findings across models and sample sizes.

---

[62] Further insights can be found in the **Appendix** and in the pre-registered research plans, the online availability of which has already been mentioned elsewhere.

Table 12: CFA Results Across Models and Samples Sizes

Hierarchical CFA	N	χ² (p > chi2)	df	CFI	TLI	RMSEA	SRMR	Coefficient of determination
Model 1*	792	0.000	438	0.947	0.940	0.047	0.054	0.942
Model 2**	792	0.000	439	0.945	0.938	0.048	0.055	0.930
Model 3***	792	0.000	440	0.944	0.937	0.048	0.056	0.912
Model 1	628	0.000	438	0.945	0.937	0.048	0.058	0.956
Model 2	628	0.000	439	0.943	0.935	0.049	0.059	0.939
Model 3	628	0.000	440	0.942	0.935	0.049	0.060	0.929
Model 1	574	0.000	438	0.946	0.938	0.047	0.055	0.944
Model 2	574	0.000	439	0.945	0.938	0.047	0.056	0.919
Model 3	574	0.000	440	0.945	0.938	0.047	0.057	0.905
Model 1	287	0.000	438	0.932	0.923	0.055	0.064	0.966
Model 2	287	0.000	439	0.939	0.931	0.052	0.063	0.934
Model 3	287	0.000	440	0.938	0.930	0.052	0.064	0.918

* *Model 1* is a hierarchical CFA with *binding* and *individualizing* morality as factors of second-order, and fairness trustworthiness and property (*individualizing*), and group, family, deference, reciprocity, and heroism (*binding*) as factors of first-order; **** *Model 2*** is a hierarchical CFA with *binding* and *individualizing* morality as factors of second-order, and fairness, trustworthiness, property and reciprocity (*individualizing*), and group, family, deference, and heroism (*binding*) as factors of first-order; ***** *Model 3*** is a hierarchical CFA comprising the 8 first-order factors already mentioned and only one common higher-order factor.

Basically, all models are defendable in terms of goodness of fit. *Model 1* has, however, the best properties among models. The CFA results support the initial EFA findings: MaC-DRS captures 8 factors of first-order. Also, more parsimonious second-order moral constructs can be captured and formed using MaC-DRS. Additionally, all factor loadings of manifest indicators are above 0.50 and the coefficient of determination (range: 0.905 - 0.966) points to the fact that the factors explain a high amount of variance. To proceed with reliability analysis, we calculated McDonald´s Omega for all MaC-DRS dimensions (Gäde et al., 2020b; Schermelleh-Engel & Gäde, 2020). Our results support overall sufficient to excellent qualities of the MaC-DRS dimensions in terms of reliability, as demonstrated by a range of McDonald´s Omega scores of 0.7951 to 0.9115. A respective table summarizing the reliability results from Study 1 can be found in the **Appendix**.

3.4.4. Discussion Study 1

Initial empirical evidence suggests that MaC-DRS is a valid and reliable instrument consisting of 32 items. The EFA´s and CFA´s conducted demonstrate that our scale captures eight moral domains as first-order factors. Also, more parsimonious second-order factors can be construed

using MaC-DRS. However, we limited ourselves in this first study to a student sample. Hence, we likely worked with a sample of outliers (Henrich et al., 2010a). Another limitation is, that we cannot assume a pure measurement of the 32 MaC-DRS items that we retained for the CFA´s due to the presence of items in the questionnaire that we discarded after the initial EFA´s. The composition of the binding second-order factor also remains partly to be discussed, in particular with regard to the dimensions of heroism and reciprocity. We additionally worked with modification indices, which can be considered a limitation. Finally, in order to establish MaC-DRS, proof must also be provided that our scale can compete with the MFT and MaC scales from a psychometric point of view in a direct comparison. Given these limitations, we have designed a second study that allows us to address these limitations and open issues.

3.5. Study 2 — Validation Study

In *Study 2* we decided again for an online-questionnaire design, yet wanted to move beyond a student sample. We aim for a gender-balanced sample that is heterogeneous in terms of education and age groups (age range: 18-85), and proportionally distributed across all federal states in Germany. The sample shall comprise at least $N = 2000$ respondents. In order to reach out to respondents we worked together with a company who took over the data collection for us and ensured that our sample requirements were met.

3.5.1. Questionnaire

Our questionnaire captures the Moral Foundation Questionnaire-1 (*MFQ-1*; relevance and judgment scale; Graham et al., 2011), Morality as Cooperation Questionnaire (*MaC-Q*; relevance scale only; Curry et al., 2019b), Morality as Cooperation—Deviance Relevance Scale (*MaC-DRS*), socio-demographic variables (Allbus, 2006; Beckmann et al., 2016; SOEP, 2019) and other constructs for another study not detailed here. In order to capture each moral scale as unbiased as possible and without distorting order effects, we have designed *four different versions of the questionnaire*, which differ in the order of the moral scales to be measured. Thus, we made sure that MaC-DRS, MFQ-1 and MaC-Q are measured at least once first in different versions of our questionnaire, i.e., before any other morality scale is measured. Further

information about the questionnaire, the study design and our measures to control for order effects can be found in the pre-registered research plan of data collection 2. [63]

3.5.2. Participants

We gathered data from N = 2,326 respondents all across Germany. The case distribution between questionnaire versions can be seen in **Table 13**. We decided to combine the questionnaire versions in which MaC-DRS was assessed first, i.e., version A1 and A2 (n = 1,162). The sample criteria were approximately met: age μ = 50.21 (median = 50; range: 18-85); percentage of females in the sample 48.50% (males 51.07%; 0.43% non-binary); diverse composition regarding school education and vocational training; and approximately proportional distribution across German federal states. With regard to the samples of the different questionnaire versions, no significant differences were found in relation to socio-demographic variables.

Table 13: Case Distributions Across Questionnaire Versions

Questionnaire versions	Freq.	Percent	Cum.
A1 - MaC-DRS & MaC-Q	637	27.39	27.39
A2 - MaC-DRS & MFQ-1	525	22.57	49.96
B1 - MaC-Q & MaC-DRS	653	28.07	78.03
B2 - MFQ-1 & MaC-DRS	511	21.97	100.00
A1 + A2	1,162	49.96	
Total	2,326*	100.00	

* Note: We removed n = 36 cases from the total sample of N = 2,362 that had an impossible response pattern by design and/or violated the age criterion of ≥ 18 years.

3.5.3. Moral Scales in Comparison — EFA Results Study 2

In the exploratory analyses, we proceeded as in *Study 1*. This time, however, we only used the *oblique factor rotation*, as a rotation algorithm. This rotation algorithm allows for correlations between the factors and corresponds more to the nature of assessing psychological constructs. The aim of these analyses is to check whether the factor structure of MFQ-1 (5 first-order

---

[63] The German MFQ-1 items were taken from: https://moralfoundations-org.translate.goog/questionnaires/?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=sc; The German MaC-Q items were taken from: https://tinyurl.com/3hc7p7w2; In addition to the aim of the present study, we are also collecting data for another study not detailed here. The study was preregistered and received ethics committee´s approval. More detailed information on data collection 2, comprising the full questionnaire, can be found online in the pre-registration: (https://doi.org/10.23668/psycharchives.13059).

factors), MaC-Q (7 first-order factors), and MaC-DRS (8 first-order factors) can be replicated in a data driven fashion based on a diverse non-student sample in Germany.

We were able to replicate the 7-dimensional factor-structure of *MaC-Q* (KMO =0.9225) and the 8-dimensional factor-structure of *MaC-DRS* (KMO = 0.9441).[64] All factor loadings of these scales are sufficient for explorative analyses (≥ 0.38). The *MFQ-1 Relevance* (KMO = 0.8947) and *Judgment scale* (KMO = 0.8470) EFA´s support however only a 3-dimensional factor-structure. Thus, the theoretically targeted 5-dimensional structure of the *MFQ-1* could not be replicated. This finding is consistent with previous studies that criticized the *MFQ-1* for problems in configural invariance (Curry et al., 2019b; Iurino & Saucier, 2020; Leitgöb et al., 2023). The EFA-Tables can be found in the **Appendix**.

3.5.4. Moral Scales in Comparison — CFA Results Study 2

We conducted CFA´s for all three moral scales using MPlus version 8, and worked with Maximum-Likelihood as estimation method. In contrast to Study 1, this time we did not use any modification indices in the CFA models. At least two models are presented for each of the scales: the **a) models** test the validity of the respective 5-; 7-; or 8-dimensional factor structure of the scales. The **b) models** build on the a) models and test additional second-order factors. With regard to the latter, we examined whether the scales can be used to capture *binding* and *individualizing* higher-order moral constructs (Haidt, 2008). Moreover, we have a hunch, but are not yet entirely clear, whether *reciprocity* and *heroism* fall under either binding or individualizing morality, or whether they represent another higher-order moral construct corresponding to a deeply rooted ***general disposition of cooperation***. Theoretically, reciprocity and heroism can regulate egoism to serve the individual and/or the group (Romano et al., 2022; Rusch, 2022). There does not appear to be an overarching focus of cooperation with regard to these moral domains. For example, it is difficult to imagine a group functioning without reciprocity, but reciprocity can also serve to protect individual cooperative offers from exploitation by free riders. Similarly, showing civil courage to protect an individual from harm, or fighting for one's country of upbringing against intruders, neither indicates a predominant group or individual orientation of heroism. We are thus reluctant to attribute reciprocity and heroism a priory in Study 2 to either binding or individualizing. To investigate our hunch, we

---

[64] On the basis of semantic considerations and the psychometric properties determined in our analyses, we reduced the set of 32 MaC-DRS items (**long version**) to a **short scale** comprising only 24 items, i.e., 3 items per factor of first-order. The analyses here refer exclusively to the short scale. Further analyses of the long scale can be found in the **Appendix**.

have tested various theoretically justifiable second-order models for MaC-DRS. In the results below, however, we will limit ourselves to the best fitting models. Overall, we will evaluate and compare the different CFA-models of the three morality scales using the goodness of fit (*gof*) indices RMSEA, CFI, TLI and SRMR. ***Table 14*** summarizes our findings.

Essentially, the CFA results reflect the EFA findings briefly discussed above. The 5-dimensional **MFQ-1** model does not fit the data well. This applies when the Relevance and Judgment sub-scales are tested individually or in a joint model. Both MFQ-1 sub-scales also suffer from problematic manifest indicators in some cases, as shown by unsatisfactory low factor loadings of (e.g.) Relevance scale = 0.388 and Judgment scale = 0.221. The CFA for the MFQ-1 Relevance scale also shows a correlation of greater/equal to one between latent factors, which means that the model must be regarded as inadmissible/overparameterized. Furthermore, higher-order factor models are either not possible with MFQ-1 or, as in the case of the Judgment scale, do not have a good fit between data and model. Taken together, we were not able to replicate Graham et al., (2011). Based on the sample examined, our results indicate that the use of the MFQ-1 is problematic for Germany, as the items do not validly capture the theoretical MFT model.

In contrast, the findings of the **MaC-Q** look quite different. The *MaC a) model* with 7 first-order factors produces satisfactory factor loadings and has an acceptable fit of data and model. We also attempted to test a higher-order factor model based on MaC-Q: findings indicate mixed results as can be seen in the *MaC b) model*. The hierarchical CFA with the best fitting values that we could found has two second-order factors, acceptable CFI, TLI and SRMR, but is slightly above the threshold for RMSEA. In summary, results demonstrate that 7 moral dimensions can be validly measured using the MaC-Q relevance scale. In this sense, we were able to replicate Curry et al., (2019b) based on a diverse sample of respondents in Germany.

Turning to **MaC-DRS**, we see that the 8-dimensional first-order factor model shows acceptable fit values. Both, the longer 32-item scale and the 24-item short scale have adequate fit statistics, although the short scale performs slightly better. The two hierarchical factor models for MaC-DRS, i.e., *model b1)* and *b2)* have acceptable fit values and factor loadings. The second-order factors in *model b1)* are: *binding*, which is composed of the first-order factors *family*, *in-group* and *deference*, *individualizing* (consisting of: *fairness*, *trustworthiness* and *property*), and a third higher-order factor which we refer to as *general disposition of cooperation*. The latter is composed of *reciprocity* and *heroism*. Note, however, that the software we used for analyses issues a warning for the MaC-DRS *b1) model*.

Table 14: CFA Results Across Moral Scales and Models

First- and second-order CFA	χ^2 Test of Model Fit	df	CFI	TLI	RMSEA	SRMR	Coefficient of determination	Number of items	Factor loadings (manifest indicators on factors of first-order)	Number of second-order factors	Factor loadings (first- on second-order)	Number of Observations
MFT R a) *	0.000	80	0.783	0.715	0.127	0.096	†	15	0.388 –	0	/	n = 511
MFT R b)	…	…	…	…	…	…	…	15	…	2	/	n = 511
MFT J a)	0.000	80	0.882	0.845	0.070	0.061	0.880	15	0.221 – 0.779	0	/	n = 511
MFT J b)	0.000	84	0.882	0.852	0.069	0.061	††	15	0.221 – 0.777	2	/	n = 511
MFT R + J a)	0.000	360	0.798	0.756	0.079	0.081	†††	30	0.253 – 0.755	0	/	n = 511
MFT R+ J b)	…	…	…	…	…	…	…	30	…	2	/	n = 511
MaC a) **	0.000	168	0.948	0.934	0.077	0.059	0.966	21	0.670 – 0.952	0	/	n = 653
MaC b)	0.000	181	0.936	0.926	0.082	0.066	0.958	21	0.670 – 0.953	2	0.732 – 0.941	n = 653
DRS a) ***	0.000	436	0.942	0.934	0.069	0.058	0.972	32	0.696 – 0.957	0	/	n = 1162
DRS b1)	0.000	453	0.936	0.930	0.072	0.072	0.969	32	0.697 – 0.958	3	0.764 – 0.934	n = 1162
DRS b2)	0.000	452	0.936	0.930	0.071	0.070	0.968	32	0.700 - 0.958	2	0.759 – 0.935	n = 1162
DRS a)	0.000	224	0.968	0.961	0.059	0.041	0.968	24	0.740 – 0.963	0	/	n = 1162
DRS b1)	0.000	241	0.961	0.955	0.064	0.059	0.971	24	0.740 – 0.963	3	0.743 – 0.924	n = 1162
DRS b2)	0.000	240	0.961	0.955	0.063	0.057	0.963	24	0.742 – 0.963	2	0.737 – 0.925	n = 1162

* **MFT models are: MFT R a)** 1-order CFA MFQ-1 Relevance Scale; **MFT R b**) 2-order CFA MFQ-1 Relevance Scale; **MFT J a)** 1-order CFA MFQ-1 Judgment Scale; **MFT J b)** 2-order CFA MFQ-1 Judgment Scale; **MFT R + J a)** 1-order CFA MFQ-1 Relevance and Judgment Scale; **MFT R + J b)** 2-order CFA MFQ-1 Relevance and Judgment Scale. ** **MaC models are: MaC a)** 1-order CFA MaC-Q Relevance Scale; **MaC b)** 2-order CFA MaC-Q Relevance Scale (best fitting model). Second-order factor *binding* is composed of *family*, *in-group*, *reciprocity* and *heroism*, and the *individualizing* 2-order factor is composed of *fairness* and *property*. To our knowledge, this is the first time, that findings on 2-order factors are reported for MaC-Q; *** **MaC-DRS models are: DRS a)** 1-order CFA MaC-DRS; **DRS b1)** 2-order CFA MaC-DRS with three higher-order factors. Second-order factors are: *binding* (*family*, *in-group* and *deference*), *individualizing* (*fairness*, *trustworthiness* and *property*) and *general disposition of cooperation* (*reciprocity* and *heroism*). Mplus gives a warning for model b1) which is potentially due to the high intercorrelation between *reciprocity* and *heroism* with all other 1-order factors and consequently also the very high correlation between *general disposition of cooperation* and *binding/individualizing*. **DRS b2)** mirrors model b1), however, in this model the 1-order factors *reciprocity* and *heroism* are not assigned to a higher-order factor. In contrast to model b1) model b2) does not give a warning in Mplus. MaC-DRS models are given for a long (32 item) and short (24 item) version. Note: All CFA MaC-DRS models shown for n = 1162 can also be replicated with the individual samples A1 (n = 637) and A2 (n = 525); † Model **MFT R a)** suffers from a correlation greater 1.000 between latent factors and is thus inadmissible/overparametrized. As the model is inadmissible, we have refrained from running a 2-order model. †† Model **MFT J b)** suffers from negative residual variance and a correlation greater 1.000 between latent factors and is thus inadmissible. ††† **MFT R + J a)** suffers from a correlation greater 1.000 between latent factors and is thus inadmissible/overparametrized. As the model is inadmissible, we have refrained from running a 2-order model.

This warning is most likely due to the high intercorrelation between reciprocity and heroism with all other first-order factors (correlation range with other first-order factors: 0.50 - 0.71) or the very high correlation between the second-order factors: correlation *general disposition of cooperation* with *binding* = 0.965; with *individualizing* = 0.896. This very high correlation between second-order latent factors could point to a linear dependency among them. We found neither a negative variance/residual variance nor a correlation $\geq$ 1.000 between latent factors. To guard against possible over-parameterization or linear dependence between higher-order factors, we have calculated another hierarchical CFA-model for MaC-DRS. *Model b2)* mirrors *model b1)* except that we have not included a higher-order factor for reciprocity and heroism in this model. Inspecting this model, we found no warning. Additionally, the model has acceptable fit statistics. To summarize, our results show that MaC-DRS is empirically proven to be able to capture 8 distinct moral domains. The long and short scale have been shown to possess adequate psychometric properties. Standing in the tradition of MFT, the results show that MaC-DRS is able to capture higher-order moral constructs such as binding and individualizing morality. The MaC-DRS higher-order models also possess sufficient fit values, which are proven to be the best fit values compared to the other moral scales tested.

We also compared the three scales in regard to reliability.[65] Looking at the reliability scores the picture obtained by the previous psychometric analysis receives further solidification: the *MFQ-1* performs not all too well, as indicated by the range of McDonald´s Omega scores from 0.5620 to 0.7130 (Relevance scale) and 0.5235 to 0.6991 (Judgment scale). *MaC-Q* (McDonald´s Omega range: 0.8779 to 0.9433) and *MaC-DRS* (McDonald´s Omega range: 0.8717 to 0.9552), in contrast, come with good to excellent reliability scores.

3.5.5. Discussion Study 2

One of the most frequently used morality scales did not withstand the psychometric tests in our study: the **MFQ-1** fails in our EFA and CFA analyses, has inadequate factor loadings in some cases and exhibits poor reliability scores in others. Our results therefore indicate that the MFQ-1 has psychometric deficiencies in German-speaking countries/samples. However, if the research interest is not in moral *relevance* but in moral *judgements*, the MFQ-2 (Atari et al., 2022a) should be used instead. According to our analyses, if one though aims to investigate moral relevance, other scales may appear more suitable than MFQ-1. Our findings replicate that the **MaC-Q** has good overall psychometric properties. This instrument can be used to validly

---

[65] A corresponding table listing all reliability results can be found in the **Appendix**.

and reliably measure 7 distinct moral domains. We argue, nonetheless, that MaC-Q suffers from several conceptual problems (as does MFQ-2), of which the most serious in our view is the failure to distinguish between moral conformity and deviance. **MaC-DRS** integrates MFT and MaC into a single approach encompassing the distinction between moral conformity and deviance both theoretically and in its operationalization. A comparison between the first- and second-order CFA-models of the three morality scales in our study shows that MaC-DRS has the best psychometric properties among the scales tested. The empirical results on reliability yield additional credence supporting thus this statement. In our view, the proposed coalescent perspective on MFT and MaC not only makes a theoretical contribution: We were able to show empirically that our scale has a broader coverage of moral domains than MaC-Q and MFQ-1. With *trustworthiness*, we have introduced a previously unrecognized moral domain, both theoretically and operationalized. Additionally, MaC-DRS is also more parsimonious than the MFQ-1 (and partly MaC-Q). The findings demonstrate that MaC-DRS allows for higher-order factor models that are empirically justified by sufficient fit between model and data when strictly applying the rule (of thumb) of: CFI/TLI ≥ 0.90, RMSEA ≤ 0.08, and SRMR ≤ 0.10. Thus, MaC-DRS is able to measure *binding* and *individualizing* higher-order moral constructs. Furthermore, we found initial results suggesting that *reciprocity* and *heroism* may form an independent, general factor of cooperation. Nonetheless, a significant limitation of our empirical efforts must still be noted at this point: We have so far based our empirical MaC-DRS examination exclusively on data from Germany. Statements about the cross-cultural applicability of our scale — validity, reliability and measurement invariance — cannot be made on the basis of our studies conducted to this point. However, as we are in the tradition of MFT and MaC interested in both the universal nature of our moral mind and the culture-specific calibrations of intuitive moral tendencies, there is still an essential body of psychometric analysis to be clarified. Consequently, in a third study, we will examine MaC-DRS across a set of four heterogeneous cultures in terms of relevant psychometric properties. In Study 3, we will also address specific hypotheses and test them accordingly. In addition, we will also examine the psychometric properties of a multidimensional self-construal scale (Vignoles et al., 2016) in the following study.

3.6. Study 3 — Cross-Cultural Study

In Study 3, we again relied on an online design, this time collecting the data using a semi-experimental questionnaire. As discussed in more detail elsewhere, we have opted for a case selection of four heterogeneous cultural entities. Our decision fell on the following cultural entities: Egypt (EG-sample); Germany (GER-sample); Japan (JP-sample) and the United States of America (US-sample). Based on a priori power analyses, we strive for a sample size of at least n = 525 per cultural sample in order to be able to ensure sufficient power. Consequently, we are aiming for a minimum number of N > 2100 cases for the overarching sample.

We commissioned a company to collect data in the respective target countries. Hence, our data collection is relying on an *access panel*. To enhance interpretability/comparability between samples and additionally due to theoretical considerations, we set several eligibility criteria and aim in the data collection for following: balanced *gender* (male/female); heterogeneous composition in terms of *age* and *education*, and approximately 70% *city dwellers* and 30% *village dwellers*. As far as *age* is concerned, a minimum age of 18 years is a prerequisite for participation in the study. We have also taken into account the fact that the average age in the cultural entities that we study varies greatly in some cases, which is why we are only aiming for a heterogeneous distribution in this case. The company we employed was able to fulfill our sample composition criteria for all samples except the Egyptian one. We will address descriptive statistics of our cross-cultural study in more detail in the next chapter. Further information can also be found in the **Appendix**.

3.6.1. Questionnaire

Our questionnaire comprises a number of different variables and constructs. We will deal with some of the relevant variables in the course of this paper. However, we have also collected other concepts for studies not covered in detail in the scope of the present work. The following *Table 15* lists the constructs we collected.

Table 15: Constructs assessed in the cross-cultural study (data collection 3)

Construct	Number of items
Morality as Cooperation—Deviance Relevance Scale (MaC-DRS) *	32 items
Moral Dilemma Scenarios	9 binary choice scenarios
Moral Deviance Factorial Survey	4 vignettes, each with 4 items
CIRN Self-Construal Scale Version	48 items
Intention vs. consequence in moral judgment	2 items
Pathogen prevalence	3 items
Attitudes on climate change	2 items
Probability to vote in the next official election	1 item
Residential mobility	1 item
Relative deprivation	1 item
Sociodemographic variables	12 items

* Note: We have highlighted the concepts used here and in the following studies in black, while the concepts that are part of various other studies are highlighted in grey. The latter are not discussed in the further course of this thesis.

The measured concepts were either available in the various languages or were translated by us and a team of colleagues into the target languages on the basis of an extended forward-backward translation approach, as described in a previous section (**Chapter 2**) (Brislin, 1970; He & van de Vijver, 2012; Smith, 2014; Behr et al., 2016; Boehnke, 2022a). In total, we were able to collect data from N = 2,982 respondents who have completed the entire study. The case distribution across the four study groups is as follows: GER-sample, n = 751; JP-sample, n = 740; US-sample, n = 745; EC-sample, n = 746.[66]

3.6.2. The Aims of the Cross-Cultural MaC-DRS Investigation — Hypotheses

We aim overall to test the psychometric properties of **MaC-DRS** across the heterogenous set of four cultural entities. In the event that we can also demonstrate sufficient cross-cultural psychometric properties of MaC-DRS, we will find substantial empirical evidence supporting the usability of this new tool for (cross-cultural) research in the field of morality. Next to EFA and CFA inspections this time, though, also *measurement invariance* is a central property on which we focus (Milfont & Fischer, 2010; Davidov et al., 2014; Rudnev et al., 2018; Gäde et al., 2020a; Leitgöb et al., 2023). So, our investigations will also center around whether MaC-DRS possesses *scalar measurement invariance* across the four groups that we examine. Given we find evidence suggesting that scalar invariance is indeed established, we can proceed in later segments of this work with mean comparisons of moral deviance relevance evaluations across

---

[66] Study 3 (data collection 3) was pre-registered and received ethics committee approval. The questionnaire in four different language versions and further detailed information on the data collection of study 3 can be found online in the pre-registered research plan: (http://dx.doi.org/10.23668/psycharchives.14630).

the four cultural entities that we examine. Hence, establishing measurement invariance is not only integral from a perspective that strives to demonstrate the usefulness of MaC-DRS but also of direct practical concern for later investigations into the cultural calibration of the human moral mind.

Apart from the methodological research goals related to MaC-DRS, we also pursue a central theoretical concern with this cross-cultural study and the testing of the psychometric properties of MaC-DRS. This concern is related to the question of the *universality* of the human moral mind and is substantially tangled to the two moral approaches on which we mainly build, i.e., MFT and MaC (Haidt & Joseph, 2007; Curry, 2016). This universalistic notion further finds expression in our hypotheses first formulated in the theoretical part of this thesis. In the following we would like to recall these hypotheses and display them again in ***Table 16.*** Please note that we have three main hypotheses, out of which one is a counter hypothesis, and also several sub-hypotheses.

Table 16: Universality of the human moral mind — MaC-DRS Hypotheses

A1 - Main hypothesis	**We expect a universal pattern of morality and hypothesize to find 8 MaC-DRS factors of first order across cultures.** In other words: we hypothesize that MaC-DRS consists across cultures of an 8-dimensional factor structure comprising fairness, trustworthiness, property, family, in-group, deference, reciprocity and heroism as moral domains.
A2 - Main hypothesis	**We hypothesize that higher-order moral constructs exist consistently across cultures.** In other words: moral relevance patterns of binding and individualizing, possibly also a general disposition of morality factor, exist universally across cultures and are cross-culturally construed based on the same moral domains.
A3 – Counter hypothesis	**Counter-hypothesis to A2) based on Atari et al., (2022a):** Higher order moral constructs are formed in a culture-dependent way, i.e., binding and individualizing, possibly also a general disposition of cooperation, do not exist consistently across cultures and are construed differently depending on the respective cultural context.

Sub-Hypotheses to A1 and A2

5)	The moral domains of fairness, trustworthiness and property build the dimensions of the higher-order *individualizing* moral construct across cultures.
6)	The moral domains of family, in-group and deference build the dimensions of the higher-order *binding* moral construct across cultures.
7)	The moral domains of heroism and reciprocity built a *general disposition of cooperation* higher-order moral construct across cultures. *
8)	The moral domains of heroism and reciprocity fall in between binding and individualizing morality, i.e., they are highly correlated with domains of both higher-order moral constructs.

* Note: As we have found already indications that a third higher-order factor build upon heroism and reciprocity (general disposition of cooperation) may lead to overparameterization of the respective model, we have marked this hypothesis grey, for we conceive it already *partly* falsified.

Apart from MaC-DRS another central feature of the psychometric test of constructs is the examination of the **CIRN-Self-Construal-Scale-3** (Vignoles et al., 2016; Yang, 2018; Uskul et al., 2023). As elaborated elsewhere (see: **Chapter 1**), we build several direct and indirect hypotheses about the cultural calibration of morality and the correspondence between selfhood and moral deviance relevance. The selection of the four target cultures of our study was not random against the background of these hypotheses. In fact, on the one hand we pursued a contrastive case selection strategy, but on the other hand we also had a systematic approach in mind: the four cultural entities of the third study differ (among other things) in terms of *self-construal* (independent, interdependent and self-assertive interdependence), cultural level *collectivism-individualism*, and the three cultural logics of *honor*, *face* and *dignity*. Thus, if we can demonstrate hypotheses-congruent findings in the cultural calibration of the moral mind, the case selection alone provides *indirect evidence* for systematic moral tendencies associated with individual level self-construal, cultural level collectivism-individualism, and the three cultural logics.

Nonetheless, we are also interested in *direct associations*. We want to investigate whether moral domain specific deviance relevance correlates systematically across cultures with 8 dimensions of selfhood captured by the CIRN-Self-Construal-Scale-3. In order to carry out such tests, the self-construal scale needs to be shown to have appropriate psychometric properties. Consequently, in addition to sufficient goodness of fit values in confirmatory factor analysis, there should also be at least metric measurement invariance for the four study groups. In accordance with these requirements, we will examine the psychometric properties of the CIRN Self-Construal Scale-3 in addition to the MaC-DRS tests.

Overall, in the light of these research goals, we are going to apply pan-cultural models as well as culture specific models for both, MaC-DRS and the self-construal scale from Vignoles and colleagues (2016).

3.6.3. MaC-DRS Across Cultures — Psychometric Insights

In the cross-cultural study, we also start with explorative analyses (EFA). We will then supplement the initial insights with theory-guided model-based analyses (CFA), and insights on the cross-cultural reliability of MaC-DRS´s factors of first-order. In a third step, we will test for measurement invariance and highlight thereafter pan-cultural correlative insights. Eventually, the results are taken up in a discussion that focuses on the hypotheses and addresses

all previous steps. Our MaC-DRS tests are followed by the psychometric review of the self-construal scale.

3.6.4. EFA Insights Study 3

Using exploratory factor analysis (EFA),[67] we want to start with the psychometric investigation of MaC-DRS in different cultures and thus initially rely on a data-driven approach. In the course of the EFA, we will test whether the 8-dimensional factor structure of our instrument emerges from the data collected in all four cultural entities. However, it should be noted at this point that the theoretical model is also of crucial importance. Therefore, we will continue with the confirmatory factor analysis after the initial exploratory insights. In addition to the precise measurement of moral deviance relevance across 8 domains a further aim of this study is to obtain a scale that is economically thrifty from a research perspective. We have therefore psychometrically analyzed both the **MaC-DRS long version** with four items per domain (32 items in total) and the **short scale** with only 3 items per moral domain (24 items in total). In order to streamline the analysis section, we will only mention the respective main findings below and elaborate a bit further on the pan-cultural EFA results. Detailed information on all exploratory factor analyses for the four cultural samples are provided in the **Appendix**.

The EFA results of the **German sample** (Study 3; n = 751) replicate the findings of our previous MaC-DRS studies. All in all, exploratory analyses provide further proof of the usability of our newly developed research tool in the German context. This applies to both, the 24-item short and the 32-item long version of MaC-DRS. In the **Japanese sample (**n = 740), the EFA results also support the 8-dimensional MaC-DRS structure. However, one item in the analysis of the MaC-DRS long version does not have sufficient factor loadings. The analysis of the 24-item short version, however, shows consistently satisfactory psychometric properties. Overall, we were therefore able to replicate the results found in Germany for the Japanese context and the MaC-DRS items translated into Japanese. Let us turn further to the **US-American sample** (n = 745) and the MaC-DRS items translated into English. Here, too, we find supportive exploratory evidence for the 8-dimensional structure of our morality scale. The overall picture of the EFA results for the *long version* (32-items) of the MaC-DRS is satisfactory and we can largely replicate the results from Germany on the basis of the US-sample and the translated items. The findings from the *short version* (24-items) of the MaC-DRS also broadly

---

[67] Note: Once again, we conducted the EFAs using the statistical software Stata13; for the CFAs and measurement invariance tests we used MPlus Version 8.

replicate the previous findings. However, the factor loadings of the family domain items fall somewhat and require a closer examination of the accuracy of fit in the subsequent confirmatory analyses. Before turning to the pan-cultural analysis, we look at the EFA MaC-DRS results for the **Egyptian sample** (n = 746). Based on the exploratory analyses conducted, we can only assume the 8-dimensional MaC-DRS structure for the Egyptian sample to a very limited extent by relying on the long version of the scale. This is mainly due to two poorly loading items in the deference domain, but we also found other items that lacked good factor loadings altogether. Overall, the exploratory results for the 24-item short version of MaC-DRS also indicate that we cannot readily assume a replication of the results of our previous studies. Rather, the results suggest that problems with the 8-dimensional structure of MaC-DRS can be identified in the Egyptian sample if we restrict ourselves exclusively to a data-driven approach. However, there is no question that theory-based insights are also needed, and as we have seen in our theoretical chapter, we have well-founded assumptions for the 8-dimensional MaC-DRS structure. Therefore, we complement the EFA results with insights from the CFA to provide more clarity.

Finally, before we move on to the theory-based counterpart of the EFA, we come to the **pan-cultural sample** (N = 2,982) and the corresponding exploratory MaC-DRS investigation. We start with a glance at the **long MaC-DRS version** (32 items): The KOM (overall) = 0.982 looks promising and indicates that our items are well suited for factor analyses. We obtain 11 factors with an eigenvalue of ≥ 1, but only 8 (first-order) factors emerge from the data and have items with sufficient factor loadings. All factors except one comprise each four items with decent factor loadings. The deference item "*Someone disregards general rules*" though performs not all too well (factor loading = 0.273), as it may is too abstract. Excluding the respective item the factor loading range of all other manifest indicators of MaC-DRS is: 0.478 to 0.922. Overall, the results of the analysis of the pan-cultural sample suggest that we can indeed assume an 8-dimensional factor structure for MaC-DRS (long version). Moreover, all 8 dimensions comprise the minimum number of ≥ 3 items per factor, which in turn have sufficient factor loadings. The KMO value of the short **version of MaC-DRS** (24 items) also demonstrates that the corresponding items are well suited for factor analyses (KOM (overall) = 0.974). Furthermore, the principal factors EFA with oblique (oblimin) rotation yields 8 factors of first-order with an Eigenvalue of ≥ 1. Each of these factors comprises three items that for their part display good loadings, as indicated by the factor loading range: 0.530 to 0.976. In summary, the EFA findings based on the pan-cultural sample suggest altogether that we can assume an 8-dimensional structure for the short version of our moral scale. In addition, all items have demonstrably sufficient factor loadings. As already emphasized before, the assessment of

MaC-DRS's psychometric properties requires additional insights from the analyses of our theoretical model. We will therefore proceed directly to the **confirmatory factor analyses** without a summarizing discussion of EFA insights, and only later take a comprehensive look at the model-based results.

3.6.5. CFA and Reliability Insights Study 3

Table 17, to be found further below, presents the CFA results of our third study for all samples. At the top of the table, we have placed the analysis of the pan-cultural sample, which we have abbreviated as "**pan**". This is followed by the sample-specific CFA models. We have labeled these individual sample-specific models with the following abbreviations: Germany (**GER**); Japan (**JP**); United States of America (**US**); Egypt (**EG**). Furthermore, for each sample we have tested the 32-item long version of MaC-DRS (denoted by "**L**") and the 24-item short version (denoted by "**S**"). All in all, we depict three models per sample and per MaC-DRS version (long and short) in the table. The **first model** tests exclusively the 8 first-order factors. In other words, the moral domains *fairness*, *trustworthiness*, *property*, *reciprocity*, *heroism*, *family*, *in-group*, and *deference* are examined here for their psychometric properties. *Table 17* displays the first models marked with the "**1order**" label. In the subsequent **second model**, we then specified three additional higher-order factors (the respective models are marked by "**2order**"). These second-order factors are binding, individualizing and general disposition of cooperation. As we already know, the second-order *binding* factor is made up of the following moral domains (first-order factors): family, in-group, and deference. The dimensions of fairness, trustworthiness, and property in turn form the dimensions that make up the second-order *individualizing* factor. Finally, reciprocity and heroism remain. These moral domains form the dimensions that we have assigned to the *general disposition of cooperation* second-order factor. However, since we already gained insights from the previous study suggesting that a third higher-order factor could lead to overparametrization of the model, we also added and tested a **third model**. We refer to this third model as the alternative model (labeled "**A**" in *Table 17*). In the alternative model, we again tested the two higher-order moral approaches that we are particularly interested in, i.e., *binding* and *individualizing* morality. However, instead of specifying a third higher-order factor, we have not assigned *heroism* and *reciprocity* to a second-order factor in this model. Rather, these two domains are only correlated with binding and individualizing in the third model. Reciprocity and heroism therefore remain solely as first-order factors in the alternative model

that we present in **Table 17**. All in all, we proceeded in the CFA´s as in the previous studies, applying the same method and procedures.

The following can be inferred from **Table 17** in the context of the **MaC-DRS long version** ("**L**"): for the pan-cultural-, the GER-, the US-, and the EG-sample, the models with 8 first-order factors ("**1order**") as well as the alternative models with the binding and individualizing second-order factors ("**2order A**") show satisfactory goodness of fit (*gof*) values. So, here we find a good fit between theoretical model and data. However, the same cannot be said for the long version and the Japanese sample. This is evident by taking a look at the *gof* values in the JP-sample, for these are partly below respectively above the limits for acceptable model fit, which applies to all three models.

If we now glance at the *gof* values for the **MaC-DRS short version** ("**S**") in **Table 17**, satisfactory values can be found across all samples for the model with 8 factors of first-order ("**1order**") and for the alternative model with the higher-order factors binding and individualizing ("**2order A**"). These findings are supported by consistently satisfactory factor loadings across all samples and the respective models. The lowest factor loading of a manifest indicator across all samples is 0.702, while the lowest factor loading for second-order factors is 0.651, providing overall acceptable results for these models. The MaC-DRS short version thus proves to be satisfactory with regard to the results of the confirmatory factor analysis across all four cultural groups examined. Nonetheless, it should further be emphasized that the model with three higher-order factors ("2order") in different samples and for both the long and the short version leads to overparametrization. Accordingly, the empirical results suggest refraining from such a model and instead favoring the alternative model with two higher-order factors and the domains of heroism and reciprocity as first-order factors.

In summary, in the context of the cross-cultural analysis, we can report only affirmative results for the short MaC-DRS version. In addition, we were able to form two higher-order factors — **binding** and **individualizing** morality — that were *consistently* based on the same dimensions *across* four heterogeneous groups of cultural entities. As indicated by our analysis, the respective models comprising these higher-order moral constructs exhibit satisfactory psychometric properties in the CFAs conducted. However, the use of the MaC-DRS long version should better be avoided in view of our results for the Japanese sample, as can be deduced from the corresponding *gof* values in **Table 17**.

Table 17: CFA results for MaC-DRS (long/short) across cultural samples

First- and second-order CFA	χ² Test of Model Fit	df	CFI	TLI	RMSEA	SRMR	Coefficient of determination	Number of items	Factor loadings (manifest indicators on factors of first-order)	Number of second-order factors	Factor loadings (first- on second-order)	Number of Observations
Pan 1order **L**†	0.000	436	0.962	0.956	0.064	0.036	0.981	32	0.707– 0.969	0	/	N = 2982
Pan 2order **L**	0.000	453	0.954	0.950	0.068	0.045	0.982	32	0.720 – 0.970	3*	0.867 - 0.958	N = 2982
Pan 2order **L A ***	0.000	452	0.955	0.951	0.068	0.044	0.980	32	0.720 – 0.970	2	0.868 – 0.955	N = 2982
Pan 1order **S**	0.000	224	0.981	0.976	0.053	0.026	0.977	24	0.763 - 0.971	0	/	N = 2982
Pan 2order **S**	0.000	241	0.973	0.970	0.060	0.034	0.978	24	0.753 - 0.971	3*	0.864 - 0.954	N = 2982
Pan 2order **S A ***	0.000	240	0.974	0.970	0.059	0.033	0.976	24	0.752 – 0.971	2	0.865 – 0.952	N = 2982
GER 1order **L**	0.000	436	0.943	0.935	0.071	0.059	0.975	32	0.708– 0.969	0	/	n = 751
GER 2order **L**	0.000	453	0.937	0.931	0.074	0.069	0.973	32	0.711 – 0.970	3*	0.825 – 0.935	n = 751
GER 2order **L A ***	0.000	452	0.937	0.931	0.074	0.069	0.971	32	0.712 - 0.957	2	0.827 – 0.934	n = 751
GER 1order **S**	0.000	224	0.973	0.967	0.057	0.036	0.971	24	0.742– 0.973	0	/	n = 751
GER 2order **S**	0.000	241	0.967	0.962	0.061	0.047	0.969	24	0.743 – 0.974	3*	0.799 – 0.931	n = 751
GER 2order **S A ***	0.000	240	0.967	0.962	0.061	0.047	0.968	24	0.744 – 0.974	2	0.798 – 0.930	n = 751
JP 1order **L**	0.000	436	0.904	*0.895*††	*0.081*	0.092	0.964	32	0.597 – 0.940	0	/	n = 740
JP 2order **L**	0.000	453	0.903	*0.894*	*0.082*	0.092	0.960	32	0.551 - 0.939	3*	0.660 – 0.898	n = 740
JP 2order **L A ***	0.000	452	0.904	*0.895*	*0.081*	0.092	0.959	32	0.547 – 0.939	2	0.665 – 0.911	n = 740
JP 1order **S**	0.000	224	0.952	0.941	0.067	0.047	0.959	24	0.731 – 0.953	0	/	n = 740
JP 2order **S**	0.000	241	0.937	*0.928*	0.074	0.066	0.954	24	0.730 - 0.956	3*	0.649 – 0.890	n = 740
JP 2order **S A ***	0.000	240	0.940	0.931	0.073	0.064	0.953	24	0.727 – 0.957	2	0.651 – 0.907	n = 740

US 1order L	0.000	436	0.953	0.946	0.066	0.036	0.977	32	0.705 – 0.952	0	/	n = 745
US 2order L	0.000	453	0.941	0.936	0.072	0.047	**	32	0.706 – 0.950	**	/	n = 745
US 2order L A ***	0.000	452	0.942	0.937	0.071	0.046	0.976	32	0.706 – 0.950	2	0.789 – 0.958	n = 745
US 1order S	0.000	224	0.968	0.961	0.062	0.030	0.971	24	0.743 – 0.950	0	/	n = 745
US 2order S	0.000	241	0.955	0.949	0.071	0.041	**	24	0.733 – 0.950	**	/	n = 745
US 2order S A ***	0.000	240	0.956	0.950	0.071	0.041	0.970	24	0.732 – 0.950	2	0.810 – 0.960	n = 745
EG 1order L	0.000	436	0.959	0.953	0.074	0.025	0.988	32	0.689 – 0.982	0	/	n = 746
EG 2order L	0.000	453	0.952	0.948	0.078	0.029	**	32	0.688 – 0.982	**	/	n = 746
EG 2order L A ***	0.000	452	0.953	0.948	0.078	0.029	0.988	32	0.688 – 0.982	2	0.879 – 0.982	n = 746
EG 1order S	0.000	224	0.969	0.962	0.076	0.023	0.985	24	0.706 – 0.984	0	/	n = 746
EG 2order S	0.000	241	0.962	0.956	*0.081*	0.027	**	24	0.702 - 0.984	**	/	n = 746
EG 2order S A ***	0.000	240	0.963	0.957	0.080	0.026	0.985	24	0.702 - 0.984	2	0.884 – 0.979	n = 746

* Note: The statistical software Mplus issues a warning of PSI (latent variable covariance matrix) being not positive definite. This warning results in our analysis from a correlation near one by second-order factors *binding* (correlation = 0.992) and *individualizing* (correlation = 0.981) with second-order factor *general disposition of cooperation*. No negative variance or residual variance was found; **** Note:** There is a correlation greater 1.000 between second-order factors *binding* and *general disposition of cooperation* (MaC-DRS 32 item long version; this applies also to the MaC-DRS 24 item short version). This indicates that three factors of second-order would cause overparametrization of the MaC-DRS long/short version second-order CFA-model in the respective sample. Note, though, that the correlation of *general disposition of cooperation* with *individualizing* morality is also near to 1.000, which can be interpreted as if *general disposition of cooperation* fills in the space between binding and individualizing, although this factor turns admittedly slightly more to binding morality; ***** Note:** The Alternative model (labeled by "A") works with *binding* and *individualizing* as second order factors. *Binding* is composed of the dimensions (factors of first-order): family, in-group and deference; *individualizing* is composed of the dimensions: fairness, trustworthiness and property. First-order factors *reciprocity* and *heroism* are not assigned to a higher-order factors in this model; **†Note:** The following abbreviations are used to label different models and samples: **Pan** = pan-cultural sample; **GER** = German sample; **JP** = Japanese sample; **US** = United States of America Sample; **EG** = Egyptian sample; **L** = long version MaC-DRS (32 items); **S** = short version MaC-DRS (24 items); A = Alternative model. **††Note:** Numbers written in italics are below the common threshold for the respective goodness of fit value. Conventions suggest the following goodness of fit thresholds for acceptable model fit: CFI/TLI ≥ 0.90; RMSEA ≤ 0.08 and SRMR ≤ 0.10.

Taking also a look at the **reliability** across the four cultural samples, we find promising evidence: The lowest McDonald´s Omega score found across samples is 0.8217 (JP-sample, deference domain), which is still good. Thus, across all samples the MaC-DRS factors of first-order range between good to excellent in terms of reliability.[68]

3.6.6. MaC-DRS Measurement Invariance

The EFA results that we found for MaC-DRS in the four cultural samples are partly promising and partly challenging. The subsequent confirmatory factor analyses, however, confirm the results of the previous MaC-DRS studies and suggest good psychometric properties for the short version of our morality scale in the cultural entities tested. In addition, the short scale exhibits good reliability, which underpins the cross-cultural usability of the 24-item version of MaC-DRS. What is still missing, though, are insights that are partly based on the factor structure and reliability, but also go beyond this, namely insights into **measurement invariance**. We therefore finally turn to the testing of measurement invariance across the cultural samples in our study. In doing so, we will test all 8 first-order factors as well as the two *binding* and *individualizing* factors of the second-order (Rudnev et al., 2018). Based on the previous results, the idea of a third second-order factor (general disposition of cooperation) is not pursued further in the context of the measurement invariance tests.

Our approach is successive: we start with the least restricted model and test for *configural measurement invariance*.[69] We then restrict the factor loadings and test for *metric measurement invariance*, and finally we arrive at the most restricted model, the *scalar model*, and constrain in addition to factor loadings also the intercepts. We proceed in such a way that we perform these steps both for the first-order factors and separately for the binding and individualizing second-order factors. In each of the three steps, we inspect and compare the alternative fit indices (RMSEA, CFI and SRMR) between the models. In doing so, we follow established rules (of thumb): In the model comparison of the less and the more restricted model, the CFI value must not fall by more than 0.01 units and the RMSEA value must not rise by more than 0.015 units. With respect to the SRMR, testing for metric invariance requires that the constrained (equal factor loadings) model, compared to the less constrained (configural) model, does not deviate more than 0.03 unites, while the model comparison for scalar invariance (equal factor loadings and intercepts) compared to the metric model requires less than 0.015 change

---

[68] A detailed table with all McDonald´s Omega scores for each sample is provided in the **Appendix**.
[69] The alternative *gof* thresholds (RMSEA, CFI and SRMR), which are also used in the CFA, apply to the valuation of the configuration model.

in the SRMR (cut-off value) (Cieciuch et al., 2019). Furthermore, it should be noted that we start our analyses with tests for *full exact measurement invariance* and only consider further steps (such as partial invariance or alignment optimization) if we miss the corresponding cut-off values of the respective measurement invariance test for the first-order factors. In the context of the second-order factors, by contrast, we would refrain from further testing if no measurement invariance is found. We were indeed able to show by CFA that binding and individualizing can be formed consistently from the same factors across a set of four highly heterogeneous cultural entities. This supports our theoretical idea that we can find or at least assume binding and individualizing across cultures. However, as already emphasized, we examine four truly diverse cultural entities. Against this background and taking into account that we do not assume that we have presented an exhaustive list of existing moral domains, there are ultimately a large number of cultural influencing variables, so that we do not consider further statistical processing of the second-order measurement invariance results to be appropriate. Rather, a corresponding circumstance would require further and more in-depth research, which we cannot provide here due to the overarching focus of our work and limited resources. Now that we have prepared the ground for our approach and our analysis, we turn to them. The results of the measurement invariance tests conducted can be found in ***Table 18***. Results are presented for both the MaC-DRS long and short versions.[70]

Table 18: Testing measurement invariance for MaC-DRS across four cultural entities

	MaC-DRS long version (32 items)		
Factors of *first-order*	**Configural**	**Metric**	**Scalar**
RMSEA *	0.072	0.074 (increase = 0.002)	0.078 (increase = 0.004)
CFI	0.945	0.941 (decrease = 0.004)	0.932 (decrease = 0.009)
SRMR	0.052	0.056 (increase = 0.004)	0.058 (increase = 0.002)
Factors of *second-order*			
RMSEA	0.147	0.138 (decrease = 0.009)	0.137 (decrease = 0.001)
CFI	0.954	0.951 (decrease = 0.003)	0.943 (decrease = 0.008)
SRMR	0.038	0.056 (increase = 0.018)	0.056 (no change)
	MaC-DRS short version (24 items)		
Factors of *first-order*	Configural	Metric	Scalar
RMSEA	0.066	0.067 (increase = 0.001)	0.072 (increase = 0.005)
CFI	0.967	0.964 (decrease = 0.003)	0.956 (decrease = 0.008)
SRMR	0.035	0.043 (increase = 0.008)	0.044 (increase = 0.001)
Factors of *second-order*			
RMSEA	0.135	0.127 (decrease = 0.008)	0.126 (decrease = 0.001)
CFI	0.959	0.956 (decrease = 0.003)	0.948 (decrease = 0.008)
SRMR	0.035	0.053 (increase = 0.018)	0.054 (increase = 0.002)

* Note: The green coloring indicates where the respective cut-off values were complied with. The red coloring, in contrast, indicates where the cut-off values were exceeded and measurement invariance does not apply at the respective level.

---

[70] We used the statistical software Mplus for the measurement invariance testing. The respective case numbers of the four cultural groups correspond to the sample sizes already indicated.

As can be seen from the results in ***Table 18***, there are two contrasting findings: *firstly*, we can indeed demonstrate full exact scalar measurement invariance for the 8 first-order factors for both the short and the long version of MaC-DRS. In our view, this is a remarkable finding that empirically demonstrates the cross-cultural applicability of this new morality scale. *Secondly*, as far as the two higher-order factors are concerned, both the short and the long version show problematic findings. Although these problems do not result from changes in the alternative fit indices across the configural, metric and scalar models, they do concern the model fit in the configural model. While we also note a slight violation of the SRMR from the metric model onwards (if applying strict rules), the RMSEA in particular, with values of 0.147 (long version) and 0.135 (short version), clearly exceeds the respective threshold values of ≤ 0.05 and ≤ 0.08 for more relaxed fit assumptions. Consequently, we cannot demonstrate (full exact) measurement invariance for the second-order factors. Conversely, the result thus indicates that, despite the demonstrably cross-culturally consistent domain assignment in the CFA models, we are actually dealing with a partially culture-specific phenomenon when approaching the constitution of the higher-order *binding* and *individualizing* moral factors.

To conclude the cross-cultural psychometric inspection of MaC-DRS, we now turn briefly and ancillary to the **correlation pattern**.[71] Inspecting the correlational pattern between the moral domains and higher-order moral constructs, we obtain a picture suggesting that these are indeed to be regarded as belonging together. Even if there are slight differences between the cultural samples in terms of the correlative pattern, the overall result is relatively clear: all MaC-DRS factors correlate high and significantly. Fairness, trustworthiness and property tend to cluster correlatively, as do the family, in-group and deference domains. The higher-order constructs of *individualizing* and *binding* are composed of these domains. Between the two higher order constructs, we place the domains reciprocity and heroism, which correlate strongly with both individualizing and binding. No dominant direction can be identified for these domains, neither towards predominantly group-centered morality (binding) nor towards predominantly individual-centered morality (individualizing). All in all, we conclude that both the domains and the higher-order constructs (binding and individualizing) are components of our evolved moral mind, which in turn are anything but orthogonal to one another.

---

[71] A table summarizing the correlative analysis can be found in the **Appendix**.

6.6.7. Indications of Universality: MaC-DRS Hypotheses

Across a set of four heterogeneous cultural entities, we found by and large that the **24-item short version of MaC-DRS** is a new measurement tool in moral research that is suitable for both intra- and above all cross-cultural analyses. Admittedly, we found some challenging results in the data-driven exploratory analyses. However, results of the theory-driven CFA in addition to the measurement invariance findings clearly demonstrate that MaC-DRS is able to validly and reliably capture 8 moral domains *across* the groups we studied. The theoretical model of the 8 moral domains, which we have derived in large parts from MFT and MaC, is therefore receiving empirical support. This fact is important from a methodological point of view, as it expands the toolbox of moral researchers with a scale that appears from initial evidence to be applicable not only in WEIRD cultures but also beyond. We can provide four MaC-DRS language versions that broadly allow for cross-cultural application and investigations of moral deviance relevance, although some fine-tuning is certainly still needed, as the EFAs suggest (e.g. in the Arabic version; EG-sample). As described in more detail elsewhere, we proceeded very carefully with the four different language versions: After an extended forward-backward translation (Brislin, 1970; Boehnke, 2022a), local partners from the target countries of our studies gave the operationalizations the final polish and ensured the culturally sensitive accuracy of the items. Our results are also meaningful from a theoretical point of view: Based on MaC-DRS findings across four cultural entities, that are truly diverse in terms of a variety of cultural dimensions (e.g. self-construal, cultural logics, pathogen prevalence, religion, geography, Hofstede dimensions etc.), we have identified a strong empirical indication that the moral domains investigated are indeed universal for our evolved moral mind. Evidence suggests that *fairness*, *trustworthiness*, *property*, *reciprocity*, *heroism*, *family*, *in-group*, and *deference* are cross-cultural domains of human morality. In addition, we were also able to demonstrate that parsimonious higher-order moral constructs can be formed consistently from one and the same moral domains *across* cultures. Our findings suggest that binding and individualizing represent a cross-cultural moral phenomenon. However, if we consider the results of the measurement invariance tests in this context, it also becomes clear that there appear to be *additional culture-specific components* that frame binding and individualizing in culturally variant ways. The higher-order moral constructs that we examine seem to contain both cross-culturally shared and culturally specific elements. Thus, although binding and individualizing suggest to comprise the same components across cultures, the results suggest also that we have not yet put the whole puzzle together and that culture-specific pieces are still missing. Against

the background of our findings, we now turn to a concise discussion of our hypotheses, which we seize again in *Table 19* edited with comments from what we have learned empirically.

Table 19: Universality of the human moral mind — Comments on the MaC-DRS Hypotheses

A1 - Main hypothesis	**We expect a universal pattern of morality and hypothesize to find 8 MaC-DRS factors of first order across cultures.** In other words: we hypothesize that MaC-DRS consists across cultures of an 8-dimensional factor structure comprising fairness, trustworthiness, property, family, in-group, deference, reciprocity and heroism as moral domains. **Confirmed** ✓ Of course, we could only approach this hypothesis approximately. However, our findings suggest that we have indeed obtained strong indications of the universality of the human moral mind. We base this interpretation on the fact that we examined four very heterogeneous cultural entities and consistently found the 8-dimensional MaC-DRS structure. Against this background, there is certainly still much room for future research, but we consider our hypothesis overall to be confirmed.
A2 - Main hypothesis	**We hypothesize that higher-order moral constructs exist consistently across cultures.** In other words: moral relevance patterns of binding and individualizing, possibly also a general disposition of cooperation factor, exist universally across cultures and are cross-culturally construed based on the same moral domains. **Partly confirmed/refuted** We found evidence that partially confirms and also partially refutes this hypothesis: higher-order constructs of binding and individualizing morality can indeed be formed across cultures by relying on the domain mapping we propose. In this respect, our hypothesis is confirmed. But in addition, it should be noted that we did not find (full exact) measurement invariance for binding and individualizing across the tested cultural samples. Even the configural models showed problems. This fact points to the influence of culture-specific inputs and influences on binding and individualizing morality.
A3 – Counter hypothesis	**Counter-hypothesis to A2) based on Atari et al., (2022a):** Higher order moral constructs are formed in a culture-dependent way, i.e., binding and individualizing, possibly also a general disposition of cooperation, do not exist consistently across cultures and are construed differently depending on the respective cultural context. **Partly confirmed/refuted**
Sub-Hypotheses to A1 and A2	
1)	The moral domains of fairness, trustworthiness and property build the dimensions of the higher order individualizing moral construct across cultures. **Confirmed** ✓ Note: This higher-order moral construct also seems to encompass other elements that may well be culturally specific.

2)	The moral domains of family, in-group and deference build the dimensions of the higher order binding moral construct across cultures. **Confirmed** ✓
	Note: This higher-order moral construct also seems to encompass other elements that may well be culturally specific.
3)	The moral domains of heroism and reciprocity built a general higher-order general disposition of cooperation moral construct across cultures.
	- In our view, this hypothesis remains an open research question. The reason for this is that we could not investigate the factor in question in more detail due to the risk of overparameterization. Moreover, this open position could be due to the argument that the 8 moral domains we proposed are most likely not an exhaustive list of all possible moral domains and that consequently other domains of morality could merge into a possible *general disposition of cooperation* higher-order factor, which could detach this factor from binding and individualizing morality.
4)	The moral domains of heroism and reciprocity fall in between binding or individualizing morality, i.e., they are highly correlated with domains of both higher order moral constructs. **Confirmed** ✓
	We seem to be confirmed that heroism and reciprocity do not primarily tend towards binding or individualizing, but serve both aspects. Nonetheless, further research is needed to substantiate our findings.

As can be seen in *Table 19* and the previous presentation of findings, the majority of our hypotheses have been confirmed or remain an open research question. What we would now like to emphasize is the following: We interpret the CFA findings as well as the scalar measurement invariance results as a strong indication that the 8 moral domains examined are indeed a *universal* part of our evolved moral mind. Although we only have data from four cultural entities, these are extremely heterogeneous in terms of a variety of cultural dimensions. At the very least, we can thus conclude that the 8 domains proposed are cross-cultural elements of the human moral mind. Consequently, our findings provide supportive evidence for the research of Curry et al, (2019a), which suggests that conformity to 7 of the 8 moral domains proposed is rated as morally good across cultures. From these promising results, we will now turn to another important aspect of this work from a psychometric perspective — the human self and its measurability.

3.7. CIRN-SCS-3: Psychometric Properties of the 8-Dimensional Self-Construal Scale Across Cultures

We will now briefly turn to another relevant scale for this project; In the following, we look at the psychometric properties of the **CIRN-Self-Construal Scale-3** (Vignoles et al., 2016; Yang, 2018; Uskul et al., 2023). In the theory section of this work (***Chapter 1***), we have already pointed out that the Self-Construal Scale (SCS) distinguishes between different aspects of being *independent* or *interdependent*. The scale comprises a total of 8 dimensions and extends the possibility of measuring a particular view of the self. This view on self-construal emphasizes that people do not constitute their self per se and exclusively independently *or* interdependently. People are not fundamentally one-dimensional in the sense of the word, and this idea is taken into account by the theoretical position of Vivian Vignoles and colleagues (2016), but also by the practical measure, the self-construal scale. Accordingly, people may well be independent in some aspects of the self (e.g. viewing themselves as unique and different from others), but interdependent in other aspects (e.g. being susceptible to influences when making decisions). An *either-or binary* view of the self does not do justice to the diverse manifestations of human selfhood, so that a more complex perspective seems altogether also more realistic. We fully subscribe to this view. Next, we will now turn to a practical examination of SCS in the context of the Egyptian, US American, Japanese and German samples that we examine.

To anticipate it at this point already, we will ultimately decide, on the basis of empirical evidence, not to carry out any detailed and substantive analyses with the self-construal scale in the further course of the current work. In principle, this is due to the fact that, on the one hand, we find promising indications for the complex 8-dimensional model of the self-construal scale — and indeed, instead of rigidly relying on psychometric rules (of thumb), it is also important to consider the improvements compared to previous scales when evaluating a research instrument (Cross et al., 2011; Vignoles et al., 2016). On the other hand, we also found challenging results in parts that require further attention and that, given the psychometric perspective adopted here, do not allow us to directly enter into cross-cultural analyses with the scale. The main reason for our decision is, after all, the practical constraints we face: Unfortunately, due to time constraints and the scope of our further investigations, we cannot afford to invest the effort that would be required to delve deeper into the scale within the scope of this work. However, this in-depth investigation will be part of future research with the self-construal scale, as we expect it to yield a wealth of further enriching analyses and findings.

From these introductory words let us turn to our empirical analyses. We did not conduct exploratory analyses for the SCS, as other experts have been entrusted with the development of this scale and have already shown promising evidence for the 8-dimensional structure of the scale (Vignoles et al., 2016; Yang, 2018; Krys et al., 2021; Uskul et al., 2023). Instead, we limited ourselves to the inspection of three psychometric tests in the broadest sense and conducted confirmatory factor analyses (**Table 20**), reliability tests (**Table 21**), and measurement invariance tests.[72] The two tables below show the results of the first two analyses for all four cultural sub-samples and a superordinate pan-cultural sample.

If we first turn to the **CFA results**, it can be seen that the CFI/TLI thresholds are mainly undershot. The RMSEA and SRMR, however, are in good ranges and a large part of the variance is also explained, as can be seen from the coefficient of determination. Against this background, it can be argued that the scale is in fact an improvement on other explicit self-construal scales (Cross et al., 2011) and that the violation of some goodness of fit indices should therefore be accepted. However, if we then look at the lowest values in the specified range of factor loadings, we find partly factor loadings that are not satisfactory. This suggests that further in-depth analysis and possibly a few modifications to the scale operationalizations could be beneficial. We will give a few more detailed descriptions to strengthen our point.

The **Egyptian sample** has the highest value for factor loadings at 0.699, which is perfectly acceptable. However, we find also two manifest indicators with factor loadings that suggest to treat these items with caution. The item "*You usually ask your family for approval before making a decision*" (Receptiveness to influence vs. Self-direction) is too low in the association with the latent dimension as indicated by a factor loading of -0.156. [73] The same applies to the item "*You try not to express disagreement with members of your family*" (Harmony vs. Self-expression) with a factor loading of -0.252. Furthermore, we identified 8 additional manifest indicators for the EG-sample SCS CFA whose factor loadings are < 0.40, which at least gives us an indication that there is still potential for improvement on this front. The analysis thus suggests that at least two items should be excluded from the scale, and also points to further manifest indicators of the SCS that appear to show potential for improvement in the case of Egypt. In the **Japanese sample**, the item "*Someone could understand who you are without needing to know about your social standing*" (Contextualized vs. De-contextualized) proves to be unsatisfactory with a factor loading of 0.191. In Addition, further

---

[72] Once again, we conducted our analyses using Stata13 and Mplus.
[73] The negative loading results from the fact that we have not yet recoded the items in the CFA and consistently polarized them in one direction. The coding of the item thus expresses a polarization to the interdependent pole of the factor.

6 manifest indicators were found to yield factor loadings of < 0.40. The analyses for the **German sample** suggest that the item "*Being able to depend on others is very important to you*" (Dependence on others vs. Self-reliance; factor loading = -0.114), and the item "*You see yourself as unique and different from others*" (Similarity vs. Difference; factor loading = 0.098) should best be excluded from the scale to validly assess self-construal in the GER-sample. We have also identified further cases that indicate potential for modification and improvement: four other manifest indicators yield a factor loading of < 0.30, while 7 further items yield a factor loading of < 0.40. Manifest indicators that one may regards as unsatisfactory are also found for the **US American sample**: the item "*Your happiness is independent from the happiness of your family*" (Connection to others vs. Self-containment) only has a factor loading of 0.241, and the item "S*omeone could understand who you are without needing to know about your social standing*" (Contextualized vs. De-contextualized) yields a factor loading of 0.297. Additionally, 9 further manifest indicators have factor loadings of < 0.40, indicating room for modification and improvement. All in all, given the factor loading results, the SCS partly requires more in-depth-analyses, modifications and adjustments in order to capture self-construal validly across the four cultural samples of our study. We consider such an endeavor to be an absolutely worthwhile task, but for future work, as it would go well beyond the scope of what the present project is able to accomplish. For this reason, we will limit ourselves at this point to presenting only the results mentioned and to pointing out the corresponding possibilities for future research.

Eventually, also our SCS testing for **measurement invariance** yielded not the results hoped for. Starting from the *configural model* we can already identify problems with the alternative fit indices' cut-off values: CFI = 0.912; TLI = 0.854; RMSEA = 0.128; SRMR = 0.046.[74] We were therefore unable to demonstrate full exact measurement invariance for the CIRN-Self-Construal Scale-3 in our analyses. In addition, we refrained from applying partial invariance tests or the approximate method of alignment optimization due to the above-mentioned, partially challenging SCS items.

---

[74] The *metric model* provides the following results: CFI = 0.906 (change = 0.006); TLI = 0.890 (change = 0.036); RMSEA = 0.111 (change = 0.017); SRMR = 0.053 (change 0.007). The results for the *scalar model* are as follows: CFI = 0.880 (change = 0.026); TLI = 0.891 (change = 0.001); RMSEA = 0.110 (change = 0.001); SRMR = 0.064 (change 0.011). Note: The results reported refer to the test of measurement invariance with a model that includes all 8 factors in one model. However, we also tested individual factor measurement invariance models. The latter perform slightly better in some cases, but yield the same results overall.

Table 20: CFA results CIRN-Self-Construal Scale-3 across cultural samples

CFA * **	χ² Test of Model Fit	df	CFI	TLI	RMSEA	SRMR	Coefficient of determination	Number of items	Factor loadings	Method factor (Correlational range with factors of first-order) **	Coefficient of determination (Method factor)
Pan-Cultural	0.000	1043	0.868	0.857	0.057	0.049	0.896	48	0.290 – 0.673	-0.051– 0.237	0.981
GER-Sample	0.000	1043	0.815	0.800	0.058	0.060	0.882	48	0.098 – 0.707	0.016 – 0.119	0.967
JP-Sample	0.000	1043	0.838	0.824	0.067	0.062	0.911	48	0.191 – 0.685	-0.014 – 0.170	0.983
US-Sample	0.000	1043	0.879	0.869	0.061	0.050	0.904	48	0.241 – 0.643	-0.061 – 0.216	0.987
EG-Sample	0.000	1043	0.826	0.812	0.059	0.066	0.885	48	0.156 – 0.699	-0.002 – 0.130	0.972

* **Note:** Our CFA approach follows (Vignoles et al., 2016; Uskul et al., 2023). We used raw, un-recoded SCS items for each respective CFA: 25 out of 48 Self-Construal Scale (SCS) items are reverse-coded and the corresponding factor loadings are initially marked with a minus sign. However, since this circumstance is negligible for the correlative analyses and can be eliminated by simple recoding, we have decided to display the factor loadings in the table solely as positive; ** **Note:** We used 50 random starts for all CFA of the five samples to find a more robust and reliable solution for the underlying maximum likelihood algorithm. *** **Note:** We included a method factor in all CFA models and allowed correlations with the individual first-order SCS factors (see: Vignoles et al., 2016, p. 983; Welkenhuysen-Gybels et al., 2003).

Table 21: Reliability analyses CIRN-Self-Construal Scale-3 across cultural samples

SCS factors of first order *	Pan-cultural	GER-sample	JP-Sample	US-Sample	EG-Sample
Similarity vs. Difference **	0.7375 (0.7355)	0.6277 (0.6178)	0.7102 (0.7035)	0.7811 (0.7799)	0.7550 (0.7514)
Connectedness vs. Self-containment	0.7879 (0.7863)	0.7001 (0.6963)	0.8249 (0.8232)	0.6926 (0.6857)	0.7523 (0.7465)
Receptiveness to influence vs. Self-direction	0.7223 (0.7206)	0.7426 (0.7401)	0.7905 (0.7859)	0.7543 (0.7517)	0.7543 (0.7517)
Dependence on others vs. Self-reliance	0.7971 (0.7914)	0.7681 (0.7503)	0.8118 (0.8072)	0.7707 (0.7686)	0.7976 (0.7969)
Harmony vs. Self-expression	0.7137 (0.7111)	0.6953 (0.6905)	0.7985 (0.7915)	0.7086 (0.7053)	0.5954 (0.5877)
Commitment to others vs. Self-interest	0.7185 (0.7165)	0.6902 (0.6852)	0.7134 (0.7071)	0.7161 (0.7146)	0.7516 (0.74979
Variability vs. Consistency	0.8001 (0.7953)	0.8058 (0.7982)	0.7987 (0.7941)	0.8175 (0.8146)	0.7664 (0.7631)
Contextualized self vs. De-contextualized self	0.7280 (0.7280)	0.6839 (0.6866)	0.7182 (0.7112)	0.7613 (0.7554)	0.7457 (0.7508)

* **Note:** The reliability analyses were conducted using the recoded SCS items (all items are polarized in the same direction, i.e., towards independent self-construal); in addition, only ipsative SCS items were used for the reliability analyses (see: Fischer & Milfont, 2010; Vignoles et al., 2016). ** **Note:** The table shows McDonald's Omega and Cronbach's Alpha for each SCS dimension. The scores for Cronbach's Alpha are displayed in brackets.

In summary: The SCS results show that the CFA values for the goodness of fit are undershot in some parts. This result applies to all samples tested. It has also been demonstrated that various manifest indicators leave room for modification and improvement. Turning to the *reliability* of the SCS dimensions (***Table 21***), we find acceptable results by and large. In this context, our findings range from questionable to good. With regard to *measurement invariance*, we were unable to provide any evidence that would allow correlative analyses or comparisons of means across the four samples examined. Taken together, from our view, some of the operationalizations should be further elaborated and more in-depth-analyses on the scale is needed in parts. Owing to time constraints in particular, we are unfortunately unable to meet these challenges within the scope of this work. Therefore, at the end of our SCS investigations, we mainly revealed a need for further research in future projects. Based on our empirical findings, we feel compelled to forego substantive analyses using the current version of the Self-Construal Scale in our study. This is truly an unfortunate circumstance and the testing of a *direct* correspondence between the cultural calibration of moral deviance relevance and the constitution of the self cannot be accomplished in the present work. Consequently, there is also a need for further research from a theoretical point of view regarding the association of the cultural constitution of the self and morality. However, since we have selected the cases (i.e., Egypt, Germany, Japan, and the United States of America) precisely according to the findings of other studies (see: **Chapter 2**), this work is nevertheless able to provide a substantial investigation, which will, though, draw exclusively on *indirect* evidence within the framework of self-construal and other cultural dimensions, as we shall see further below. Following the above, we would like to conclude our primarily methodological, but also substantial investigations with a comprehensive discussion.

3.8. Overall Discussion: Psychometric Results and Indications of Moral Universality

The **Moral Foundations Theory** (MFT) has broadened and deepened the field of moral psychology by highlighting moral intuitions and moral pluralism (Haidt, 2001; Haidt & Joseph, 2007; Graham et al., 2013). Partly an extension to MFT it is the **Morality as Cooperation Theory** (MaC) that centers morality around human cooperation, and builds on a more solid theoretical ground by drawing on game-theory (Tomasello & Vaish, 2013; Diekmann, 2013; Curry, 2016; Curry et al., 2019a). Although these theories are tremendously fruitful, they nonetheless lack a theoretically and empirically defendable *self-report measure of moral relevance*. In the studies presented we propose an approach that combines and partly extends MFT and MaC in terms of theory and measurement. We have adopted the moral domains from MaC, re-operationalized them and also integrated parts of the MFT foundations. In addition, our approach recognizes trustworthiness as a new domain and proposes at least 8 distinct moral domains. We hold that conformity and deviance are at the heart of moral behavior, as they realize either care or harm as a consequence of actions in potentially cooperative interaction situations. Building on this idea, we further suggest that our moral mind evolved to recognize and identify patterns of moral conformity and moral deviance, and triggers appropriate responses to both kind of perception. In this regard, we predict moral deviance to trigger stronger reactions than conformity, for bad is stronger than good (Baumeister et al., 2001).

Based on our approach to morality we developed the **Morality as Cooperation—Deviance Relevance Scale** (MaC-DRS). In **Study 1** we were able to reduce our starting material from 48 to a more economical set of 32 items, and demonstrate that our operationalizations possess good psychometric properties. **Study 2** further validates these initial findings. We collected data from more than two thousand respondents across Germany via three different morality scales in a single design. The findings demonstrate that MaC-DRS is superior to MFQ-1 and MaC-Q in all psychometric properties that we investigated. Our scale makes it possible to measure 8 moral domains in a valid and reliable way. In addition, MaC-DRS enables the measurement of higher-order constructs as *binding* and *individualizing* morality (Haidt, 2008). The moral domains of *family*, *in-group* and *deference* cluster to a higher-order factor that we understand as *binding* morality. We hold that these domains regulate egoism so that predominantly group cooperation can flourish. *Fairness*, *trustworthiness*, and *property*, in contrast, reflect an *individualizing* approach to morality. In our view, conforming behavior in relation to these domains primarily means caring about cooperation with other individuals,

regardless of individual social affiliation. Lastly, we found also indications suggesting the possibility of a third moral factor of higher-order, which we call *general disposition of cooperation*.

The latter second-order factor is formed by *reciprocity* and *heroism*. These moral domains can be linked to processes of the evolutionary development of our moral mind. Since the survival of our ancestors depended to a large extent on cooperation, reciprocity becomes relevant both in inter-individual and in contexts of group cooperation, provided that reputation and partner selection play a role (Mauss, 1968; Tomasello & Vaish, 2013; Henrich & Muthukrishna, 2021; Romano et al., 2022). Heroism is also relevant from the point of view of cooperation when it serves to protect cooperation partners on whom our ancestors were dependent. Within this framework, the development of actor-neutral empathy can also be placed in parts if organism-preserving cooperation beyond family structures and kinship altruism proved relevant (De Waal, 2008; Kurzban et al., 2015; Rusch, 2022). Our capability to emphasize enables us to recognize suffering of potential cooperation partners and motivates us to alleviate it. Heroism therefore not only serves to protect the cooperation partner (which also ensures own survival), but can also be associated with reputation effects and signals of one's own willingness to cooperate, which act as a fitness advantage if the reproduction of the organism depends on cooperation (Tomasello & Vish, 2013; Baumeister, 2022). It may be that reciprocity and heroism are thus to be understood as elementary moral dispositions that may are evolutionarily prior or distinct to the other domains proposed here. At the very least, our results of the second study suggest that reciprocity and heroism appear to be more generally endorsed as straightforwardly binding or individualizing domains, as they should be relevant regardless of a social predominance of group- or individual-centered cooperation. We hope that our findings stimulate further research, also in regard to what reciprocity and heroism have in common as in regard to their "place" in the evolution of human morality.

In the first two studies of the practical development process of MaC-DRS, we solely collected data in Germany. We then explicitly addressed this limitation in **Study 3**, which covers four cultural entities. After an elaborate forward-backward translation and with the involvement of the expertise of local partners, we created a semi-experimental design to collect data in Egypt, Japan, the USA and again in Germany. In each of the four samples and an additional pan-cultural sample, EFA, CFA, reliability analyses and tests for measurement invariance were then conducted for MaC-DRS. As part of this, we psychometrically analyzed both the 32-item *long version* and a 24-item *short version* of **MaC-DRS**. Overall, the short scale in particular proved to be valid and reliable across all the cultural entities we examined. However, the EFA findings

for the Egyptian sample represent a limitation to this statement. The exploratory analysis could not replicate the 8-dimensional MaC-DRS factor structure from the data. Future qualitative studies, such as cognitive interviews (Konrad, 2020), would be useful in this context to further investigate and better understand the exploratory MaC-DRS finding in Egypt. However, all *CFA models* showed a good fit between the theoretical idea and the data across all samples tested. The models with the 8 first-order MaC-DRS factors and also the models specifying binding and individualizing as second-order factors delivered consistently satisfactory results. In contrast, the model with a third higher-order factor pointed to the danger of over-parameterization, which is why we will ultimately discard it for the time being from further investigations in this study. Since we do not assume that we have presented an exhaustive list of all moral domains with the 8 proposed moral domains, we can still presume a third higher-order factor — we have called this factor *general disposition of cooperation* due to the lack of a term more suitable — which could be sufficiently distinguished from binding and individualizing morality by including further moral domains. All in all, we believe that further studies investigating the possibility of this additional higher-order factor remain a worthwhile endeavor for future research. Moreover, we were able to further substantiate the MaC-DRS CFA findings with mostly very good *reliability* results. The analyses of *measurement invariance* also support the empirically grounded impression that MaC-DRS is a promising instrument for cross-cultural research on morality: We can demonstrate full exact scalar measurement invariance for the 24-item short version across all four study groups. This finding therefore confirms that it is possible to conduct mean value comparisons on the basis of MaC-DRS across the four groups in our study.

We also examined the substantive aspect of the methodological findings. The analysis of this aspect took place in the context of testing our hypotheses on the *universality* of the 8 moral domains that we investigate. The empirical results support our assumption of the main hypothesis A1: ***The expectation of a universal moral pattern expressed in 8 MaC-DRS first-order factors across all tested cultures is confirmed on the basis of our data.*** We interpret the MaC-DRS findings of our third study as providing strong overall indications that *fairness*, *trustworthiness*, *property*, *reciprocity*, *heroism*, *family*, *in-group*, and *deference* are universal domains of the human moral mind. At the very least, they clearly represent cross-cultural domains of morality. Our results thus provide strong indications of central positions of MFT and MaC, which theoretically derive moral universalism from an evolutionary perspective (Haidt & Joseph, 2007; Graham, 2013; Curry, 2016; Curry et al., 2019a). Again, it should be emphasized that we can of course only approximate the universalism hypothesis on the basis of only four cultural entities studied. However, and this should also be emphasized, the four

samples differ in a variety of cultural dimensions and are indeed heterogeneous. In this light, the strong indication of the universality of the 8 moral domains is underpinned by the diversity of the groups that we examine.

Eventually, we examined also the psychometric properties of the **CIRN-Self-Construal Scale-3** in the third study. The results are in the field of tension between two arguments. Against the background of other explicit self-construal scales, the version we examined stands in a good light, not only because of its realistic basic assumption of multidimensionality, but also because of the psychometric findings. However, it also became apparent that we would have to elaborate considerably on the current version of the scale in order to be able to use it for our present study in the context of substantial analyses. The latter is primarily based on the findings that not all of the tested manifest indicators are associated with the corresponding latent dimensions in the same way across all four samples. In some cases, we found results that showed *factor loadings* that were too weak, suggesting that the corresponding items did not adequately capture the underlying concept. Further, more in-depth analyses and modifications of the scale would be necessary if we wanted to conduct substantial analyses based on sufficient psychometric parameters of the SCS in this study. However, as such an undertaking would clearly go beyond the scope of this project, ***we ultimately decided in the context of the SCS CFA and measurement invariance results not to investigate self-construal directly in the further course of this thesis***.

An important part of our study, which formulated a direct (and measured) correspondence between dimensions of the self and moral domains across cultures as a theoretical claim, therefore remains an **open research question**. Due to the problems outlined above, we are unable to conduct corresponding empirical analyses in this study. However, not only do we hope that an empirical examination of the correspondence between the cultural configuration of the self and the calibration of our moral mind will be possible in future cross-cultural studies, but we will adhere from a theoretical stance to the assumption of a systematic correspondence between selfhood and moral (deviance) relevance until there is evidence to the contrary. In the context of moving away from a direct empirical association of a connection between self-construal configuration and calibration of our moral mind, our **case selection** of the study samples should be emphasized once again. On the basis of our case selection, we can still investigate important hypotheses, as other studies show that *Egypt*, the *USA*, *Japan* and *Germany* differ with regard to self-construal, collectivism-individualism and cultural logics (see: **Chapter 2**). Our study will therefore be able to examine various moral hypotheses in the cultural comparison that we undertake below and allow us to pursue our research question —

i.e., the question of *which moral system guides cooperation in different cultures*? In the background, though, it should be noted that we will rely solely on **indirect evidence** in our interpretations of the following investigations, as we cannot make a direct association of moral deviance relevance with various dimensions of the self in this study for the reasons given.

Like every study, ours have their **limitations**: although we have been able to greatly improve on the student sample of *Study 1* with our German wide sample in *Study 2*, the latter sample is nonetheless not representative. The same holds true for *Study 3*, which took our research into the field of cross-cultural psychology. Also, our design is obviously cross-sectional, and we thus have not been able to assess potential changes in moral deviance relevance across time. Furthermore, we have so far neither associated MaC-DRS with behavioral or emotional measures nor with instruments that are better suited to assess Type-1 processes, as the implicit association test (IAT) for instance (Haidt, 2001; 2003; Karpinski & Steinman, 2006; Nosek et al., 2007; Tangney et al., 2007; Ellemers et al., 2019; Greenwald & Lai, 2020; Tutić, 2023). Finally, one of the most important contributions of MaC-DRS is that this instrument exclusively measures the relevance of moral *deviance*. Conversely, this focus on deviance also implies that a scale that measures the relevance of moral *conformity* does not yet exist. From the perspective of positive psychology, an endeavor focusing on the good side of morality (Haidt, 2003) could be a valuable field for future scale construction and research.

Limitations are also pathways for **future research**. Does MaC-DRS predict pro-social behavior; does the relevance of moral domains change over time (if so, what causes these changes), how fast is people′s moral deviance relevance evaluation, and can we associate specific patters of moral deviance relevance with specific moral emotions? Furthermore, also studies comprising significantly more cultural entities would be highly desirable in order to further substantiate the strong indications of the universality of the 8-dimensional MaC-DRS structure that we found. Also, as already mentioned, more in-depth investigations of the MaC-DRS EFA findings from the Egyptian sample are certainly desirable. We will also take a closer look at this aspect and emphasize it in the further course of this work. Ultimately, we hope to stimulate these and other possible avenues of future research with our new morality scale. Finally, we believe that we have expanded our research repertoire with the **Morality as Cooperation—Deviance Relevance Scale** which is an empirically justified supplement to the MFT and MaC scales. We hope that this new instrument will stimulate further investigations on one of the central features of our species and the social world in which we live — human morality.

Now that we have provided ample evidence of the development and cross-cultural applicability of MaC-DRS, and have also been able to examine initial hypotheses, we can finally turn to further *substantive* investigations of human morality across cultures. In the next chapter, we will deal with our research question and the cultural calibration of the human moral mind in detail and this time, after a preliminary phase, primarily in terms of substance. Our further analyses will be based on the third, cross-cultural study (data collection 3) that we conducted. We begin the transition to the main field of our empirical examination of cultural entities and moral deviance relevance by turning first to the descriptive statistics, i.e., the sample characteristics of the four cultural groups (Study 3) that we seek to examine in more depth. Furthermore, we would like to recall once again that in the following chapter, in addition to several hypotheses, we will pursue our overarching research question and empirically explore *which moral system guides cooperation in different cultures.*

Chapter 4: Investigations of the Human Moral Mind II

4.1. Cross-Cultural Investigations in Morality

Now that we have provided ample evidence of the development and cross-cultural applicability of MaC-DRS, and have also been able to examine initial hypotheses, we can turn to further substantive investigations of human morality across cultures. In this chapter, we will deal with our overarching research question and thus focus on the empirical assessment of the question *Which moral system guides cooperation in different cultures*. Hence, in the center of attention will be the examination of the cultural calibration of the human moral mind and this time, primarily in terms of substance. Our further investigations will be based on the cross-cultural study (**Study 3**) that we have conducted. We begin the transition to the main field of our empirical examination of cultural entities and moral deviance relevance by turning first to the descriptive statistics, i.e., the sample characteristics of the four cultural groups (Study 3) that we seek to examine in more depth.

But before we turn to the descriptive properties of the samples, we would like to discuss an adjustment of the data set. For various reasons of data quality, we felt compelled to carry out a sample adjustment. We excluded a total of 622 cases from the N = 2,982 cases of the data set, which will from now on be referred to as the ***full sample***, in order to obtain the data set which will from now on be referred to as the ***adjusted sample***. In the following analyses, we refer exclusively to the *adjusted sample* (N = 2,360). However, we will refer to the *full sample* data set at various points. For reasons of traceability and transparency of our research process, we have also carried out the majority of the MaC-DRS analyses discussed below for the *full sample*. All analyses for the full *sample* can be found in the **Appendix**. In addition to the descriptive analysis of the sample and analyses of response styles, the *full sample* analyses also include further analyses in which we attempt to compare the cultural samples in the context of MaC-DRS and also present the results of various models.

It should be noted, however, that due to the data quality underlying the *full sample*, as well as the theoretically justified consideration of analyzing only cases of the respective cultural mainstream of the countries of our study (i.e., cases with a single citizenship), we refrain from interpreting the corresponding findings. Instead, we argue that solely the *adjusted sample* and the analyses based on it should be interpreted in terms of substantive findings. By excluding a total of 622 cases, the adjusted *sample* provides us with what we consider to be the most valid data basis permitted by our cross-cultural data collection (Study 3). At the same time, it should

be emphasized that we have taken into account the corresponding results of our a priori power analyses for the adjusted *sample*: all of the four cultural (sub)samples have sufficient statistical power even after adjusting the sample. The corresponding arguments in favor of the adjusted *sample*, which result primarily from methodological perspectives but also from theoretically justifiable assumptions, can be found in the **Appendix**.

It should also be noted that the analyses that we presented in the previous chapter are largely unaffected by the corresponding lack of data quality in the *full sample*. Rather, it can be argued that the psychometric MaC-DRS results, which we were able to show despite "poorer" data quality, speak for the robustness of the results. Furthermore, we have also carried out CFA´s — the short MaC-DRS version shows acceptable *gof* statistics across all four countries examined — and tests of measurement invariance for MaC-DRS based on the *adjusted sample*. We will refer to the latter at the appropriate point within the text and can prove that scalar measurement invariance is present for all four cultural groups in the *adjusted sample*.

Taken together, in the following we will work exclusively with the ***adjusted sample***, which comprises N = 2,360 cases. Detailed justifications for the use of *the adjusted sample* as well as descriptive and substantive analyses based on the *full sample* can be found in the **Appendix**. We are now turning to the descriptive analysis of the *adjusted sample* and dive thereafter consecutively into the cross-cultural investigation of the human moral mind.

4.2. Descriptive Insights Adjusted Sample

The following provides a descriptive overview of the *adjusted sample* and the four cultural groups in our cross-cultural study. At this point, we primarily deal with socio-demographic characteristics.[75]

Table 22: Case distribution on cultural groups in the adjusted sample

	Freq.	Percent	Cum.
Germany	666	28.22	28.22
Japan	543	23.01	51.23
USA	569	24.11	75.34
Egypt	582	24.66	100.00
Total	2,360	100.00	

[75] Note: In addition to data quality considerations and a theoretical argument, we also have excluded 13 further cases from the sample. A total of 8 cases stated non-binary when asked about *gender*. Regrettably, this represents a group that is too small for meaningful quantitative analyses, which is why we excluded these cases. A further 5 cases in the US-sample responded to the *education variable* with early childhood education (ISCED category 0). As this answer is very implausible with the other characteristics of these cases, we removed them from the sample.

The *age* distribution across the four cultural groups is as follows: GER-sample, Ø = 51.560 (median = 53); JP-sample, Ø = 51.981 (median = 54); US-sample, Ø = 49.001 (median = 50); EG-sample, Ø = 37.508 (median = 35).[76] After the adjustments to the sample and the resulting exclusions of cases, we still find the desired balanced distribution on the variable *gender* (***Table 23***). Overall, almost an equal distribution of the female/male ratio across the four groups can be identified.

Table 23: Case distribution across gender and cultural groups (adjusted sample)

Gender	GER-sample	JP-sample	US-sample	EG-sample	Total
Female	333	263	290	284	1,170
Male	333	280	279	298	1,190
Total	666	543	569	582	2,630

As far as the variable *place of upbringing* is concerned, it can be seen that 33,47% of our total sample grew up in a village (on the countryside) and 66,53% in a city. What stands out is the distribution in the JP-sample, were nearly half of the respondents grew up in a village. Opposed to that is the pattern in the EG-sample were a major share of respondents grew up in a city, as can be seen in ***Table 24***.

Table 24: Case distribution place of upbringing across cultural groups (adjusted sample)

Place of upbringing	GER-sample	JP-sample	US-sample	EG-sample	Total
Village/Countryside	258	245	177	110	790
City	408	298	392	472	1,570
Total	666	543	569	582	2,360

All in all, the *adjusted sample* comprises predominantly urban socialized people and in regard to the *place of living* an even larger share of respondents indicated to be a city-dweller. The latter statement, however, holds not true for the US-sample. Compared to the *place of upbringing*, there are more country residents (village-dwellers) in this sample, as can be inferred from ***Table 25*** to be found on the next page.

---

[76] As far as the variable of *age* in the four samples is concerned, we are mostly slightly yet also to some part strongly above the average age of the respective countries for each of the groups examined. See in this regard "Reflections on the Age Item" in the **Appendix**. This circumstance must also be taken into account with regard to possible attempts to generalize our findings presented below to respective societies, as such a conclusion is also made more difficult by the findings in regard to the age variable and should be avoided.

Table 25: Case distribution place of living across cultural groups (adjusted sample)

Place of living	GER-sample	JP-sample	US-sample	EG-sample	Total
Village/Countryside	205	173	188	70	636
City	461	370	381	512	1,724
Total	666	543	569	582	2,360

With regard to the continuous variable *years in school*, there are considerable group differences, which are primarily due to the outlier of the Egyptian sample: GER-sample, Ø = 11.288 (median = 12); JP-sample, Ø = 12.821 (median = 12); US-sample, Ø = 12.059 (median = 12); EG-sample, Ø = 14.254 (median = 15). This fact is also reflected in the variable *education* (ISCED), as can be seen in the following table.

Table 26: Case distribution across education and cultural groups (adjusted sample)

Education (ISCED)	GER-sample	JP-sample	US-sample	EG-sample	Total
Primary education	6	0	102	0	108
Lower secondary education	212	18	29	5	264
Upper secondary education	75	182	109	69	435
Post-secondary/non-tertiary education	128	5	46	28	207
Short cycle tertiary education	51	87	41	0	179
Bachelor´s or equivalent	72	115	137	440	764
Master´s or equivalent	101	128	49	29	307
Doctoral or equivalent	22	8	16	0	45
No response	0	0	67	11	51
Total	666	543	569	582	2,360

Note also, that the indication of the category "Primary education" in the US-sample does not fit to the responses on the variable *years in school*, where the average for the n = 102 cases that indicated primary education is Ø = 11.029 (median = 12). This shows an unwillingness to answer the corresponding *education* item in the US-sample.

When we look at the variable *income*, we find that more than 39% of the N = 2,360 cases of the *adjusted sample* indicated not to answer the item or to have no own income. For the cases that responded to the *net income* item, however, we obtain the following values: GER-sample (n = 466), Ø ≈ 2427€ (median = 2200 Euro); JP-sample (n = 287), Ø ≈ 256050¥ (median

= 200000 Yen); US-sample (n = 279), Ø ≈ 9804$ (median = 2560 US Dollar); EG-sample (n = 414), Ø ≈ 10832£ (median = 7000 Egyptian Pound).[77]

With regard to the variable *denomination* and *level of religiosity*, the following descriptive findings are available for the *adjusted sample*. The GER-sample is mainly Christian or consisting of people with no religion or denomination. The sample comprises 24.47% "Roman Catholic Church" respondents, 22.22% "Protestant Church", and 43.24% who indicated "No religion or denomination". The remaining percentage is distributed among other religions or the response category "Can´t choose / Not specified" and accounts for only a small proportion of the sample. With regard to the variable *level of religiosity* (7-point scale, higher values indicate a higher *level of religiosity*) we found the following for the German sample: Ø = 2.864 (median = 3).[78] In terms of religious beliefs, the Japanese sample mainly responded with "No religion or denomination" (50.64%) and "Can´t choose / Not specified". Overall, 62.61% of respondents in this sample answered the question about *religious denomination* with these response options. Besides these answers, the respondents who identified themselves as "Buddhists" form the largest religious group in the JP-sample (31.31%),[79] and the remaining percent scatter on the other categories of the *denomination* item. With regard to the *level of religiosity* the Japanese sample takes in comparison to the other cultural groups of our investigation the lowest values (Ø = 2.368; median = 2). Note, choosing the 2 on the *level of religiosity* item is equivalent with the response option "Very non-religious". The US-sample is dominated by Christians, as can be inferred from the following numbers: "Roman Catholic Church" 17.40%; "Protestant Church" 18.63%; "Protestant/Evangelical free Church" 7.21%; "Orthodox Church" 0.88%; Another Christian religious community 14.94%. Also, a non-insignificant share of respondents answered with "No religion or denomination" (26.89%) while 5.98% selected the response category "Can´t choose / Not specified". The remaining parts of the sample only form minor groups in the context of the *denomination*. The *level of religiosity* is in comparison relatively high in the US-sample: Ø = 4.328 (median = 5). Responding to the respective item with 5 is equivalent with choosing the response option "Somewhat religious". Furthermore, the Egyptian group examined mainly comprises Muslims: 90,73% of the sample responded with " Sunni Islam (Sunni)" (86.43%), "Shiite Islam (Shiite)" (0.69%), or "Another

---

[77] We have excluded cases with an unrealistic response to the net income item (e.g. several trillion Yen (¥) net monthly income) from the information provided here. We have taken this reduction in the number of cases into account when stating the respective sample size on the *net earnings* item.

[78] Remember: We have coded the cases that responded with "Can't choose / Not specified" as 0 in order not to lose any cases and at the same time to maintain the continuous structure of the variable.

[79] As for the denomination aspect, our sample from Japan seems to have relatively balanced characteristics, as the corresponding distribution can also be found in other data sets: (https://de.statista.com/statistik/daten/studie/166887/umfrage/religionen-in-japan/).

Islam religious community" (3.61%). While also 4.64% responded with "Orthodox Church", the remaining few percent only form small groups in the context of the *denomination* (e.g. "No religion or denomination" 1.20%). In terms of religiosity, the EG-sample has the highest values among the groups studied, but does not differ substantially from the responses of the US-sample: level of *religiosity* Egyptian sample Ø = 5.187; median = 5.

Overall, expected differences in *age*, *income* and *religiosity* can be identified in the *adjusted sample*. As far as the *education* variable is concerned, the EG-sample is an outlier, which should be considered and will be discussed against the background of the generalizability of our findings. The Egyptian sample also stands out from the other study groups in terms of the proportion of village and city dwellers. Finally, with regard to the variable *gender*, there is balanced coverage across all four samples. From these insights we will now briefly touch on response style differences in the cultural groups of the *adjusted sample*.

4.3. Measures to Control for Response Style Effects

In examining the distributions on each of the 32 MaC-DRS items across the four cultural groups in our study, we found that different *response styles* were likely to be prevalent in the cultural samples. In addition to the view that response styles do not exist, two other prominent perspectives on response styles can be found among cross-cultural researcher. The substantive perspective on response styles emphasizes peculiarities of cultural communication, while the methodological perspective accentuates the view of response styles as a nuisance in the data that can distort comparative analyses (Smith, 2004; van de Vijver & Leung, 2011; He & van de Vijver, 2012; He et al., 2021). We adopt a perspective that encompasses both aspects and believe that response styles are cultural ways of communication *and* yet may pose a threat to the comparison of data collected across different cultural groups.

Based on our inspections of the data across samples and given our cross-cultural design, we reason that it will proof useful and important to control for possible response style distortions with appropriate measures (Baumgartner & Weijters, 2015; Smith et al., 2016; He et al., 2021). Hence, we created measures for *middle category response style* (*MRS*), *agreement response style* (*ARS*), and *disagreement response style* (*DARS*) based on a set of measured items (Bogner & Landrock, 2016). The measures *ARS*, *DARS* and *MRS* were constructed following the suggestions of Baumgartner and Weijters (2015). The construction processes of the response style variables were as follows: In order to build the variables, we used 60 items from different

scales that are polarized in the same direction.[80] The identification of one and the same response style across different scales is important in order to check whether the response style in question is actually determined by the content of the construct or whether it is determined irrespective of the construct's content. The latter is taken as an indication that a response style in the classical sense is present (He & van de Vijver, 2012). Regarding *ARS* and *DARS* we created first weighted measures for each by coding all respective end-point answers as 2, answers on the next following category as 1 and all other answers on the item's response formats as 0. After we obtained single *ARS* and *DARS* measures we further proceeded by creating a *net acquiescence response style* (*NARS*) measure for which we subtracted *DARS* scores from *ARS* scores (Baumgartner & Weijters, 2015).[81] We did not create the *MRS* measure as a weighted measurement instrument. For the *middle category response style* measure, we simply coded the responses to the middle category as 1 and all other responses as 0 across all the 60 items that we have also used for *NARS*.

Now that we have formed the measures for the response styles, we want to investigate empirically whether we can confirm our suspicions stated above across the cultural groups of our study. Hence, we conducted simple one-way ANOVAs (Bonferroni corrected) (Völkle & Erdfelder, 2010; Aden et al., 2021) to test for significant differences in *NARS* and *MRS* across the four cultures of our study. Should we find significant differences between the groups, we plan to include *NARS* and *MRS* as covariates in the models of our subsequent data analysis.

The ANOVA results (N = 2360) for the net *acquiescence response style* (NARS) reveal in parts highly significant differences (Prob F > 0.000) between groups. However, no significant difference between the GER- and US-sample was found. The effect is by far the strongest in the Egyptian sample, and the Japanese sample exhibits the lowest *NARS* values. These findings are reflected by the culture specific means and the effect direction of *NARS*: GER-sample Ø = 10.752; US-sample Ø = 14.441; EG-sample Ø = -18.259, and JP-sample Ø = 2.491. The Egyptian sample has thus a profound tendency to choose the lower end-point of scales. In

---

[80] We used the following items to create our response style measures: MaC-DRS, 32 items; attitude on climate change, 2 items; level of religiosity, 1 item; role of intention vs. consequence in moral judgment, 2 items; CIRN-SCS-3, 23 items (only positively coded items were used). Note: Including items from scales of different content is important to measure response styles as response styles are ways of responding irrespective of the constructs content.

[81] Admittedly, the items we used to create NARS are not consistently semantically equivalent in their response format to the "disagree/agree" Likert scale format for which the ARS/DARS measures were originally developed. However, we believe that the underlying logic of ARS/DARS can also be applied to the item response formats that we used to create our response style measures. Thus, we take a position shared by Smith and colleagues (2016). These authors state: "Our position is that response style will vary between samples no matter what type of measurement is employed—hence, for cross-cultural studies to approach valid measurement, effects of response style must be estimated" (p. 454).

contrast, the GER- and US-samples possess a somewhat milder tendency towards the opposite direction. To understand the scores from the JP-sample, we turn to the *middle category response style* (*MRS*). A one-way ANOVA (Bonferroni corrected; N = 2360) for *MRS* reveals highly significant differences (Prob > F 0.000) across all four groups. Group means are: GER-sample Ø = 10.557; US-sample Ø = 8.741; EG-sample Ø = 6.223, and JP-sample Ø = 14.255. Apparently, the Japanese sample has the highest tendency to choose the middle category as measured by *MRS* and is followed by the GER- and then the US-sample. The Egyptian sample, by contrast, has the lowest *middle category response style* tendency, which is not surprising considering the *NARS* results.

To summarize, we can conclude that differences in response styles between cultural groups are very likely present in the *adjusted sample*. The empirical evidence therefore suggests that it is advisable to include the *NARS* and *MRS* measures as covariates in subsequent models of analysis.[82]

4.4. Which Moral Systems Guides Cooperation in Different Cultures? Moral Deviance Relevance from a Cross-Cultural Perspective

In the following we will draw on our investigations of *MaC-DRS* across cultures and focus on the examination of our overarching **research question** — i.e., **which moral system guides cooperation in different cultures**. In the context of this overarching research question, we would like to recall our hypotheses that we derived from the theory section for the four cultural samples of our study. Remember that previous research has characterized the cultural entities of Germany and the United States of America by prevailing cultural level *individualism* and *independence* in individuals' self-construal. In contrast, the cultural entities of Japan and Egypt have been characterized by prevailing cultural level *collectivism* and (self-assertive) *interdependence* in self-construal at the individual level (Minkov & Kaasa, 2022; Kitayama & Salvador, 2024). Furthermore, individuals' ways of selfhood, their cognitions and social orientations are embedded in cultural logics (Leung & Cohen, 2011; Uskul et al., 2023). In regard to cultural logics, Egypt has been characterized by prevailing *honor logic*, Japan by prevailing *face logic* and the Unted States as well as Germany mainly by prevailing *dignity logic*. Drawing on collectivism-individualism, differences in selfhood, varying cultural logics

---

[82] Response style analyses based on the *full sample* and additional interpretations can be found in the **Appendix**.

and also on several other cultural dimensions, we have derived different (ideal-type) ***overall social orientations*** for the cultural groups that we study. We would now like to emphasize these social orientations once again:

> We suggest that the cultural contexts of **Egypt** and **Japan** are characterized by **group-centered (interdependent) overall social orientation**. By drawing on findings from other studies, we assume that the everyday logic of the situation of actors from the respective countries is shaped by present significant symbols that condition the activation of an interdependent frame of reference and corresponding codes and scripts. We expect hence, that group-oriented ***binding*** **morality** and frame congruent (response) behavior is especially important in Egypt and Japan.

> We suggest that **Germany** and the **Unted States of America** are characterized by **individual-centered (independent) overall social orientation**. Taking the findings from other studies into account, we assume that the everyday logic of the situation of actors from the respective countries is shaped by present significant symbols that condition the activation of an independent frame of reference and corresponding codes and scripts. Therefore, we expect that individual-oriented ***individualizing*** **morality** takes in special importance in Germany and the United States and is apparent in frame congruent (response) behavior.

Against this background, we have built our **cultural difference (CD) hypotheses** as heuristic means to approach the moral systems prevalent in the cultural entities of our study. Overall, we aim to investigate one *main hypothesis* in the context of cultural differences in morality. This main hypothesis is in turn further informed on the basis of four *sub-hypotheses*. The corresponding hypotheses are listed again in the ***Table 27***, to be found on the next page.

In the course of our cross-cultural investigations into human morality, we will empirically test the hypotheses on cultural differences (CD main hypothesis and sub-hypotheses 1 to 4). Our analyses and interpretations in the cross-cultural examination of MaC-DRS will focus on several aspects: We begin with investigations in cross-cultural commonalities and differences in moral deviance relevance and, accordingly, discuss the influence of *culture* on intuitive valuations of moral breaches. Next, we briefly turn to the moral domains that exhibit the highest cross-cultural deviance relevance. Finally, based on the MaC-DRS results, we approach a preliminary classification of the moral systems that guide cooperation in the cultural entities examined. The section on the results of Morality as Cooperation—Deviance Relevance Scale (MaC-DRS) concludes with an overarching discussion of the findings from the *adjusted sample*. Also, limitations of our study as well as promising avenues for future research will be addressed. In the next chapter, the cross-cultural analysis of morality is supplemented by a look at moral reasoning in the realm of *dilemma scenarios*. However, before moving on to the dilemma scenarios, the following parts examine *moral intuitions* and the research question

which moral systems guide cooperation in different cultures based on the *MaC-DRS* findings. To disclose our research process, we first turn to our models of estimation.

Table 27: Hypotheses in the context of cross-cultural moral differences

Main Hypothesis Cultural Differences (CD)	Although we predict universalism of the 8 moral domains proposed by MaC-DRS, we also *hypothesize significant differences in moral domain relevance across cultures.*
Self-Construal	We expect cross cultural differences and…
Sub-hypothesis: CD 1	…*hypothesize that cultural entities that foster relatively more interdependent ways of selfhood also foster relatively more binding morality* (i.e., they have higher relevance ratings of the family, deference and in-group moral domains).
Sub-hypothesis: CD 2	… *hypothesize that cultural entities that foster relatively more independent ways of selfhood also foster relatively more individualizing morality* (i.e., they have higher relevance of the fairness, trustworthiness and property moral domains).
Cultural logics	
Sub-hypothesis: CD 3a	We hypothesize that cultures of *honor* and *face* are significantly higher in *binding morality* than cultures of dignity.
Sub-hypothesis: CD 3b	Due to prevailing honor logic and self-assertive interdependence in self-construal, we expect that Egypt, however, scores higher on individualizing domains than Japan.
Sub-hypothesis: CD 4	We hypothesize that cultures of *dignity* are significantly higher in *individualizing morality* than cultures of cultures of honor and face.

4.4.1. The OLS Models: Covariates and Interaction Terms

To estimate cross-cultural commonalities and differences in moral deviance relevance, we take the 8 moral domains measured by MaC-DRS as consecutively tested **dependent variables**. Thus, we examine *fairness*, *trustworthiness*, *property*, *reciprocity*, *heroism*, *family*, *in-group*, and *deference* in terms of cross-cultural commonalities and differences by using independent multivariate OLS regression models for each moral domain.

As **covariates**, we incorporate the following measured variables into our models: first, we include the response style measures described above, i.e., *middle category response style* (*MRS*) and *net acquiescence response style* (*NARS*) as covariates in our models. Furthermore, we include a *pathogen prevalence* measure. This variable consists of three items and is intended to indirectly capture how high the *pathogen prevalence* is in the country of the respective sample. We asked the respondents three times to provide the following information: "*Please think of a traditional meat dish from your country. How many herbs/spices are used in total for this dish? Please give us a number*". The average number resulting from the three answers build

our pathogen prevalence variable. Background to the items is that meat dishes spoil more quickly in warmer natural environments, as germs can infest a meat dish more quickly under such conditions. In this context, it was also found that the use of herbs/spices in meat dishes is directly related to the presence of pathogens in the environment. The reason for this underlying relationship is that herbs and spices contain natural antibacterial defenses and cuisines around the world take advantage of this property to protect their food from pathogens and germs (Sherman and Billing, 1999; Murray & Schaller, 2010). We therefore have an indirect measure of pathogen prevalence with our items. In addition, pathogens are associated with morality, as has been shown in other studies (Atari et al., 2022b). For this reason, we decided to include this variable as a covariate in our models.

As already mentioned in parts, there are significant imbalances between the samples with regard to the variable's *denomination*, *level of religiosity*, *place of upbringing* (village vs. city), *place of living* (village vs. city), and *education* (ISCED-2011). Since we can presume a theoretical relationship between these covariates and the dependent variables (moral domain deviance relevance), we have included them in the model.[83] In addition to the education variable, which solely asks about educational attainments in categories, we have also included the continuous variable *years in school* in the model.

From a theoretical perspective, we can assume an influence of the variable *residential mobility* on the relevance of moral deviance (Henrich, 2020). We included a corresponding item in our questionnaire in order to be able to capture possible intracultural, but also cross-cultural differences in the experience of different (residential) contexts. The variable re*sidential mobility* takes into account the possibility of socialization experiences outside the dominant cultural influences of the place of birth. Different residential contexts, within a country but also across national borders, may well be associated with different normative experiences, which in turn may affect the calibration of people´s moral mind. Consequently, we have included this variable as a covariate in the models The *residential mobility* variable is built upon the following question: "*We are interested in the extent to which you have moved from one kind of place to another. Which do you think is most like your experience of life?*". This item comprises four possible response options of increasing residential mobility: 1 "*I have always lived in the same neighborhood*", 2 "*I have lived in different neighborhoods in the same place*", 3 "*I have lived in different places in the same country*", and finally 4 "*I have lived in different countries*". We

---

[83] As already indicated elsewhere, we coded cases that indicated "No response" on the item *level of religiosity* with the value 0, so that the continuous structure of the variable is maintained and we still do not lose any cases. We used this coding for the variable *level of religiosity* throughout all further OLS models that follow.

also included *age* in years as a (continuous) covariate, as moral intuitions may change over the course of a person's life and experiences. Another reason is that the variable *age* diverges significantly between countries — the cultural groups in our study differ substantially on the *age* variable, as among other things socioeconomic differences between countries are also reflected in life expectancy.

The last covariate we included in the models is the *gender* of the respondents. One of our sample eligibility criteria was the balance of the respondents' gender, which is why we have almost the same number of females and males in the respective samples across the four groups of our study. However, we have slight differences in the relative sample sizes (adjusted sample: GER N = 666; JP = 543; US = 569; EG = 582), and furthermore excluded 8 non-binary cases from the analysis. Next to these reasons we are also interested whether gender differences in moral deviance relevance can be found. Hence, we decided to include *gender* (female vs. male)[84] as covariate into the models.

After presenting the covariates included in our models, the models are not yet complete.[85] The assumption that the covariates of our model have the same (additive) influence on the dependent variables (moral deviance relevance of the 8 MaC-DRS dimensions) in the four cultural groups cannot be regarded as a priori justified. This might be precisely the case because covariates may impact on one another, so as they are not assumed to be independent. Therefore, it could be important to include **interaction terms** in the OLS covariate models which we use to examine moral deviance relevance (Jann, 2005; Judd et al., 2014). Statisticians speak of *moderation* or *interaction* when the effect of a predictor variable x_1 (e.g. education in years) on the dependent variable y (e.g. income) varies because x_1 and thus the effect of x_1 on y is not independent of another predictor variable x_2 (e.g. parents' education) (Lohmann, 2010; Boehnke & Hadjar, 2015).[86] Based on theoretical considerations, we can assume interaction

---

[84] As indicated elsewhere, a total of 8 cases responded with non-binary on the *gender* item. As this is a too small group for meaningful quantitative analyses, we exclude these cases from all OLS models.

[85] We assume that the individual moral domains (factors) are interrelated and anything but orthogonal to each other (Brandt, 2020). On the one hand, this follows from the assumption that we understand the corresponding moral domains as interrelated facets that are part of the human moral mind. On the other hand, our position follows also from empirical results demonstrating MaC-DRS cross-domain correlation, and the fact that the domains can be partially aggregated to higher-order factors. The latter, as correlative analyses by us show, are also not orthogonal but are to be understood in strong relation to one another. This is yet not to say that we necessarily advocate that the moral domains cannot also be different modules (Haidt & Joseph, 2007), but merely that we assume that the respective moral domains are part of humans' internal moral system. Humans possess a moral mind. We are primarily interested in the influence of *culture* on the human moral mind and on the cultural calibration of individual moral domains. Since we consider the individual moral domains to be interrelated, and we are explicitly interested in the influence of culture on each of these interrelated domains, we refrain from including the moral domains not measured as dependent variable as covariates in the respective OLS models.

[86] Let us assume, for example, that years in *school* have a positive effect on *income* and that the binary variable *parental education* (low education = 0; high education = 1) is also positively associated with a higher income if parental education = 1, whereas the state parental education = 0 has no such effect. Now the interaction comes into

effects between several covariates of our OLS model, especially with regard to the variable *culture*.

For example, it could be assumed that the years spent in an educational institution (variable *years in school*) — we are talking here about secondary socialization in institutions (see e.g. Berger & Luckmann, 2013, pp. 148-157; Bourdieu, 2014) — influence the respondents' world view and their socio-economic positioning, and may therefore have an impact on moral tendencies. However, it can also be assumed that this effect of secondary socialization is already culturally variable, as the educational content in some areas (e.g. in history lessons) is likely to differ from culture to culture. The variables of *culture* and *years in school* are therefore possibly not independent but interact with each other, which could lead to culturally variable effects of education (*years in school*) on moral deviance relevance. The inclusion of an interaction term between *culture* and *years in school* in our model can therefore be regarded as theoretically permissible, as these variables are probably not independent of each other.

We consider it furthermore possible that people's *level of religiosity* impacts on moral deviance relevance evaluations. This effect is potentially also non-uniform across cultures. In other words, it is likely not independent from culture because cultures can largely vary in the respectively prevalent religious belief system. Examples for this are the predominantly Christian belief system in the US-sample and the predominantly Islamic belief system in the EG-sample. These different religious belief systems may be associated with different world views (Bar-Tal, 1998; Hunsberger & Jackson, 2005; Hogg et al., 2010) and in parts with different moral codices affecting eventually intuitions of moral deviance relevance. Hence, an interaction of *culture* and *level of religiosity* as well as the *level of religiosity* and religious belonging (*denomination*) can be assumed and should be integrated into our models.

In addition, it can be supposed that the *age* variable also releases culturally varying intuitive valuations of moral deviance relevance. Assuming that there is no individual history of migration, it can further be presumed that culturally specific experiences also increase with increasing age. To put it differently, the longer people interact in and with their immediate

---

play, because the effect of *years in school* on *income* is not independent of the *parents' education*. Or in other words: the slope of the regression lines of the effect of *years in school* on *income* are moderated by the variable *parental education* and are therefore no longer the same for people with parents who have a low level of education and for people with parents who have a high level of education — the effect of *years in school* on *income* is greater for the group of people whose parents have a high level of education. Therefore, if we had not included the interaction effect (moderation) between *years in school* and *parental education* in the estimation of the effect of *years in school* on *income*, we would have underestimated the effect of *years in school* for the group with highly educated parents.

socio-cultural environments, the more culturally influenced experiences they are likely to have gathered. Since cultural experiences can also shape our moral mind, it can be assumed that the influence of *age* on moral deviance relevance is not independent of *culture* and should be regarded with an interaction term in our models.

We have demonstrated further above that the cultural groups under investigation differ in the response style tendencies *middle category response style* (*MRS*) and *net acquiescence response style* (*NARS*). These findings reflect that ways of communication can't be treated as independent from culture. As response style differences are part of the assessment of intuitive moral deviance relevance, we consider it important to include interaction terms between *culture* and *response styles* in our OLS models.

Last but not least, we can hypothesize an interaction effect between our measure of *pathogen prevalence* and *culture* on the relevance of moral deviance. Different degrees of *pathogen prevalence* might have led to culture-specific adaptations such as norms, which ultimately also influence the calibration of the moral mind. However, as *pathogen prevalence* varies around the globe this effect might not be independent from culture. Therefore, it seems reasonable to also include an interaction term for *pathogen prevalence* and *culture* in our models.

Taken together, we cannot assume a priori that various measures of our set of covariates are independent of the variable *culture*. For this reason, we consider the inclusion of interaction terms in the OLS covariate models relevant to prevent potential misspecification and under- or overestimation of cultural similarities/differences when examining intuitive moral deviance relevance. Accordingly, we decided to include 7 interaction terms in our MaC-DRS estimation models in addition to the covariates already mentioned.[87] These interaction terms are: *Culture* x *Years in School*, *Culture* x *Level of Religiosity*, *Culture* x *Age*, *Culture* x *Response Style* (*MRS* and *NARS*), and *Culture* x *Pathogen Prevalence*. We also included an interaction term for *denomination* and *level of religiosity* in the models. Overall, we integrate into our multivariate models of estimation a set of 13 covariates and 7 interaction terms next to the variable *culture* in which we are primarily interested in.[88] Based on these models we will estimate and examine

---

[87] We have refrained from integrating interaction terms that would be between two categorical variables (e.g. *culture x education/ISCED*), as otherwise the identification of the *average marginal effects*, in which we are interested, would have become statistically difficult.

[88] Remember: we have also asked respondents about their *net income* and gave them the option of not answering this item. Since a considerable percentage in each sample answered "No response" or "No own income" to the *income* item, we decided to drop this variable from all models. For an insight into the net incomes of the four samples examined, please refer to the descriptive statistics presented further above. Further insights on the *net income* variable in the *full sample* can be found in the **Appendix**.

moral deviance relevance margins (Average Marginal Effects; Williams, 2012; Wooldridge, 2016; Kohler & Kreuter, 2017) across the four cultural groups of our study.

4.4.2. Moral Deviance Relevance — The Covariate Models: Main and Interaction Effects

One major goal of our cross-cultural study is to examine the research question **which moral system guides cooperation in different cultures**? This question translates in major parts to the examination of moral deviance relevance, i.e., which moral domain(s) are of primary concern in cultural entities and should be safeguarded the most against breaches. Now that we have introduced the covariate model(s), we will be able to approach our research question and the cultural difference hypotheses (*CD hypotheses*). In doing so, we will draw on the ***adjusted sample*** as the basis for MaC-DRS analyses. As described above, we excluded (in total 622) cases with poor data quality in addition to cases likely not exclusively socialized in the cultural mainstream of our study countries from the *full sample* to obtain the *adjusted sample*.[89] According to our argumentation, the *adjusted sample* provides us with the most valid database to conduct our analyses. Please note, although the *adjusted sample* with N = 2360 has significantly fewer cases than the *full sample*, it still has sufficient power. In other words: each of the cultural sub-samples comprises more than n = 525 cases, which translates into statistical power of > 0.80 as a priori power analysis that we conduced revealed.

Before focusing in detail on average marginal effects and cross-cultural commonalities and differences, we discuss briefly each of the moral deviance relevance models by taking a look at the **main** and **interaction effects** identified. So, in addition to identifying possible main effects, the following analysis also serves to investigate whether significant culture-specific effects exist. This step will also help us to empirically inform the subsequent analysis and interpretation of the moral deviance relevance margins across cultures. We start by taking a look at *Table 28* in which we find the coefficients and standard errors for significant main and interaction effects displayed.[90] Thereafter we turn briefly to the individual models and their effects, graphically illustrate the interaction effects found, and conclude with a summarized discussion before focusing and elaborating on the moral deviance relevance margins.

---

[89] As already mentioned, detailed MaC-DRS analyses based on the *full sample* can be found in the **Appendix**.
[90] In the interests of clarity, we have refrained from presenting variables in *Table 28* that proved to be non-significant across all models. A list of these variables can be found in the *note* below the table.

Table 28: OLS covariate models — main and interaction effects in the realm of moral deviance relevance

GER-Sample n = 666
JP-Sample n = 543
US-Sample n =569
EG-Sample n = 582
Total N = 2,360

	Fairness Coef. (Std. Err.)	Trustworthiness Coef. (Std. Err.)	Property Coef. (Std. Err.)	Reciprocity Coef. (Std. Err.)	Heroism Coef. (Std. Err.)	Family Coef. (Std. Err.)	In-Group Coef. (Std. Err.)	Deference Coef. (Std. Err.)
Culture †								
- USA (base) ††	n.s. †††	††††				n.s.	n.s.	n.s.
- Japan		1.486 (.395) ***	1.932(.387) ***	1.410 (.345) ***	1.297 (.362) ***			
- Germany		-.035 (.460) n.s.	1.381 (.450) n.s.	.982 (.402) n.s.	1.053 (.422) n.s.			
- Egypt		-.895 (.606) n.s.	.296 (.593) n.s.	-.315 (.530) n.s.	.523 (.556) n.s.			
Culture x LOR Δ								
- USA (base)	n.s.	n.s.	n.s.	n.s.	n.s.			n.s.
- Japan						-.048 (.050) n.s.	-.040 (.049) n.s.	
- Germany						-.040 (.035) n.s.	-.106 (.034) *	
- Egypt						-.317 (.094) *	-.329 (.092) ***	
Age	n.s.	n.s.	n.s.	n.s.	.012 (.002) ***	.008 (.003) *	n.s.	n.s.
MRS	n.s.	.021 (.006) *	n.s.	.020 (.005) ***	n.s.	n.s.	n.s.	n.s.
Culture x MRS								
- USA (base)	n.s.		n.s.		n.s.	n.s.	n.s.	n.s.
- Japan		-.033 (.008) ***		-.024 (.007) *				
- Germany		-.003 (.009) n.s.		-.017 (.008) n.s.				
- Egypt		-.010 (.011) n.s.		-.006 (.009) n.s.				
NARS	.038 (.001) ***	.045 (.001) ***	.048 (.001) ***	.036 (.001) ***	.041 (.001) ***	.044 (.001) ***	.032 (.001) ***	.035 (.001) ***
Culture x NARS								
- USA (base)			n.s.		n.s.	n.s.		n.s.
- Japan	-.003 (.003) n.s.	-.004 (.002) n.s.		.000 (.002) n.s.			-.001 (.002) n.s.	
- Germany	.009 (002) ***	.003 (.002) n.s.		.003 (.001) n.s.			.000 (.002) n.s.	
- Egypt	.005 (.002) n.s.	.006 (.001) *		.007 (.001) ***			.006 (.001) ***	

ISCED	*n.s.*	*n.s.*		*n.s.*	*n.s.*	*n.s.*	*n.s.*
- Primary education (base)							
- Bachelor or equivalent							-.366 (.127) *
- Master or equivalent							-.457 (.135) **

†**Note:** main effects are highlighted bold; interaction effects are highlighted by italics; ††**Note:** we have chosen the US-sample as reference category as it is probably the cultural group best examined (see: Henrich et al., 2010a). ††† **Note:** several variables proved non-significant (Holm-Bonferroni corrected) throughout all models and had consequently no effect on moral deviance relevance. These variables are: *Gender*; *Years in School*; *Denomination*, and interaction: *Culture x Age*; *Denomination x Level of Religiosity; Residential Mobility; Place of Upbringing*; *Place of Living*; *Pathogen Prevalence*, and *Culture x Pathogen Prevalence*. Those variables that proved consistently non-significant are not displayed in the table for reasons of clarity. †††† **Note:** All significance indications are Holm-Bonferroni corrected by number of tests per test family (e.g. the test family for *culture* comprises 27 tests). The asterisks indicate (Holm-Bonferroni corrected) significance: * p <0.05, ** p <0.01, *** p <0.001. Non-significant (*n.s.*) p-values are highlighted in italics. Δ **Note:** we have abbreviated the variable L*evel of Religiosity* as LOR.

Let us briefly inspect the results of the OLS covariate and interaction term models in ***Table 28***. In doing so, we will look at all 8 moral domains one after the other and start with the **fairness** deviance relevance model. Comprising main effects and interactions terms the model contains a total of 27 tests that relate to the variable *culture* and accordingly belong to a test family. The variable *culture* has four characteristics, with the US-sample representing the *reference group* in our OLS models. We therefore have three main effects for *culture*. Other main effects, 6 in total, are: *years in school*, *level of religiosity*, *age*, *pathogen prevalence*, and the response style measures of *NARS* and *MRS*. For these variables, we also included interaction terms with the variable *culture* in our model. This results in a total of 18 interaction terms, as we have one interaction term in the model for each of the three cultural samples tested (remember, the US-sample is the reference group). Taken together, this results in a total test family of 27 tests for the variable *culture*. We are thus in the realm of multiple testing, and correct accordingly for possible alpha-error cumulation by using the Holm-Bonferroni method to adjust the significance level.[91] Please note, that the same number of 27 tests within the test family relating to the variable *culture* can of course also be found in the following models, which do not examine fairness deviance relevance but the other dependent variables.

Without taking the Holm-Bonferroni correction into account, 8 p-values would be significant in the fairness deviance relevance model. After the correction, however, only a positive main effect for *NARS* (p < 0.001), and an interaction effect of *culture* and *NARS* for fairness deviance relevance remain significant. The latter applies exclusively to the German sample and is highly significant (p < 0.001) and positive. All other p-values of the model relating to the variable *culture* are found to be not significant. Since we identified a significant and positive interaction effect for *culture* (GER-sample) *x NARS*, in addition to the positive main effect of *NARS*, we see an even stronger positive influence on the fairness deviance regression slope for this particular sample. However, as indicated, this double positive effect does not apply to the other cultural groups, as no such interaction effect could be found for them. The interaction effect finding is illustrated in the following ***Figure 6***. From ***Figure 6*** it can be inferred that the regression slope for the *German* sample, which is colored in *green*, is steeper than the slope for the other groups. This graphically demonstrates the double positive effect of *NARS* that solely applies to the GER-sample.

---

[91] Note: For the categorical variables *denomination*, *education* and *residential mobility*, we correct the respective significance levels according to the number of their characteristics for multiple testing with the Holm-Bonferroni correction.

The fairness deviance relevance margins, which we will look at below, are therefore in parts attributable to the *culture x NARS* interaction effect for Germany, as the effect of *culture* (GER-sample) on fairness deviance relevance is partially moderated (amplified) by the interaction with the *net acquiescence response style*

Figure 6: Margins NARS
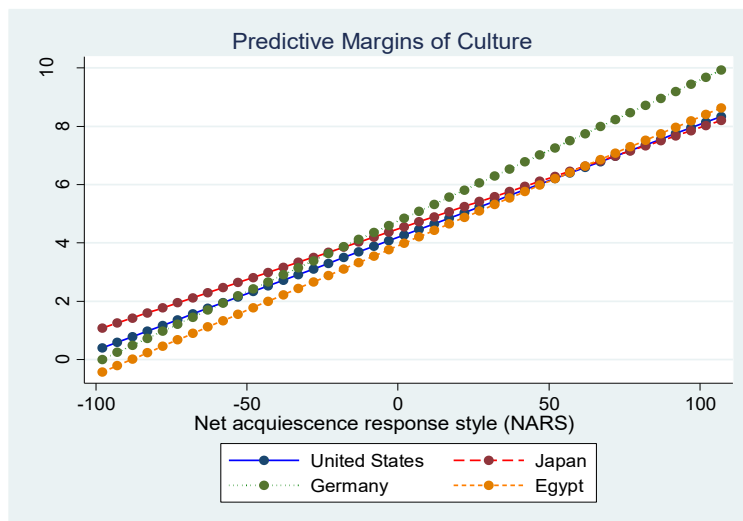
variable (*NARS*). It should further be noted that after the correction for alpha-error cumulation, no other effects besides those described proved to be significant.

Turning to the results of the **trustworthiness** OLS model, we find again 8 significant effects in the test family of the variable *culture* before applying Holm-Bonferroni correction. After the significance level correction, 5 effects still prove to be significant: Firstly, there is a positive and significant main effect of *culture* for the JP-sample ($p < 0.001$). Note, this main effect is found for the comparison with the reference group of the US-sample. Also, for the response style measures positive main effects were found (*NARS*: $p < 0.001$; *MRS*: $p = 0.024$). In addition to the main effects, results yield evidence for two significant interaction effects. The first is negative and applies for *culture* (JP-sample) *x MRS* ($p < 0.001$) on trustworthiness deviance relevance. Thus, the positive main effect of *MRS* reverses for the JP-sample.

The second interaction effect, in contrast, is positive ($p = 0.024$) and is found for *culture* (EG-sample) x *NARS*. The main effect of *NARS* that is already present is hence amplified for the EG-sample. Both interaction effects, i.e., *culture x MRS* and *culture x NARS*, can be found graphically displayed in *Figure 7* and *Figure 8*. Please note that the regression slope for the

Figure 7: Margins MRS
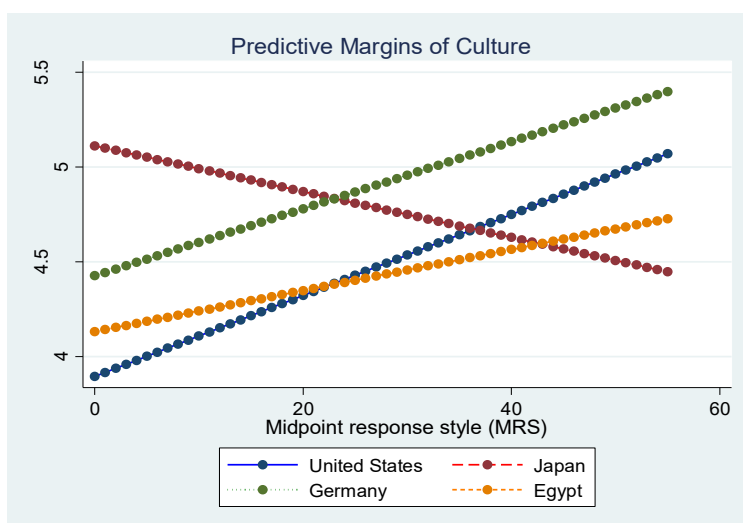
Japanese sample is colored in *red*, and the slope of the *Egyptian* sample is colored in *orange*. Apparently, the interaction effects found should be considered when interpreting the trustworthiness deviance relevance margins (see below). In addition to the effects stated, no further significant results were obtained.

Figure 8: Margins NARS
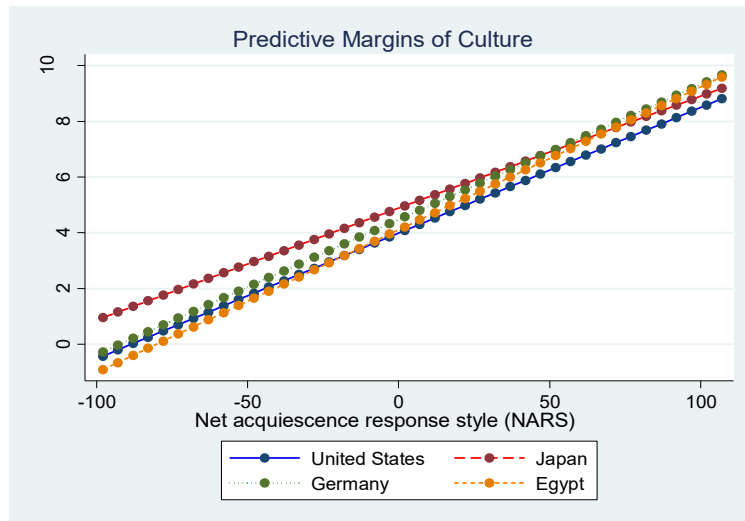
If we look at the **property** deviance relevance results from the respective OLS model, we find 10 significant p-values before the Holm-Bonferroni correction. After correcting for alpha-error cumulation, there are however only 2 significant effects left within the 27 tests that involve the variable *culture*.[92] As can be seen in **Table 28**, we found a positive and significant (p < 0.001) main effect for *culture* (JP-sample) having the US-sample as reference group. Also, a positive main effect of *NARS* was identified (p < 0.001). Moreover, although several effects from the test family related to *culture* showed different directions in the groups examined, no further effects below the threshold of p < 0.05 were found among them.

Next, we move on with the **reciprocity** deviance relevance findings. The respective uncorrected OLS model yields us 9 significant p-values that emerge from the test family with the variable *culture*. After we have also reacted here to alpha-error cumulation with adjusted significance levels (Holm-Bonferroni correction), a total of 5 significant values result. We observe a significant (p < 0.001) and positive main effect for *culture* (JP-sample). In addition, the two response style measures *NARS* and *MRS* also display a positive and significant (p < 0.001) main effect. Next to these main effects, however, also culture specific response style measure interactions are identified. These interaction effects can be found for Egypt (*culture*) x *NARS* (positive effect, p < 0.001; *orange* regression slope in **Figure 9**), and Japan (*culture*) x *MRS* (negative effect, p = 0.023; *red* regression slope in **Figure 9**).

---

[92] Note: The (positive) main effect of *culture* (GER-sample) is exactly at p = 0.05 (Coeff. = 1.381; Std. Err. = 0.450) after applying the Holm-Bonferroni correction, while a negative interaction effect for *culture* (GER-sample) and *age* on the relevance of property deviance is also exactly at p = 0.05 (Coeff. = -.011; Std. Err. = 0.003) after correcting for alpha error cumulation. We have marked these effects as non-significant in **Table 28**.

Figure 9: Margins NARS

Predictive Margins of Culture

Net acquiescence response style (NARS)

United States — Japan
Germany — Egypt
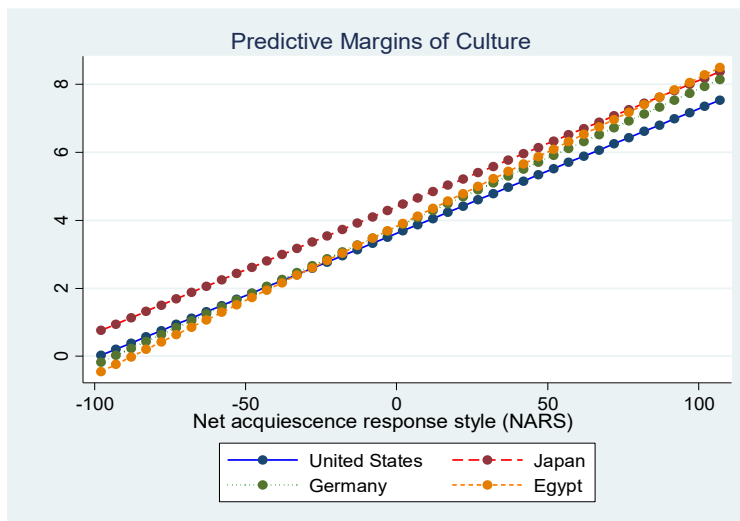
Hence, for the EG-sample, the positive main effect of *NARS* is further amplified by the additional culture specific interaction effect with *NARS*. In contrast, the positive main effect of *MRS* on reciprocity deviance relevance is reversed for the JP-sample due to the culture specific interaction with the response style variable.

Figure 10: Margins MRS

Predictive Margins of Culture

Midpoint response style (MRS)

United States — Japan
Germany — Egypt
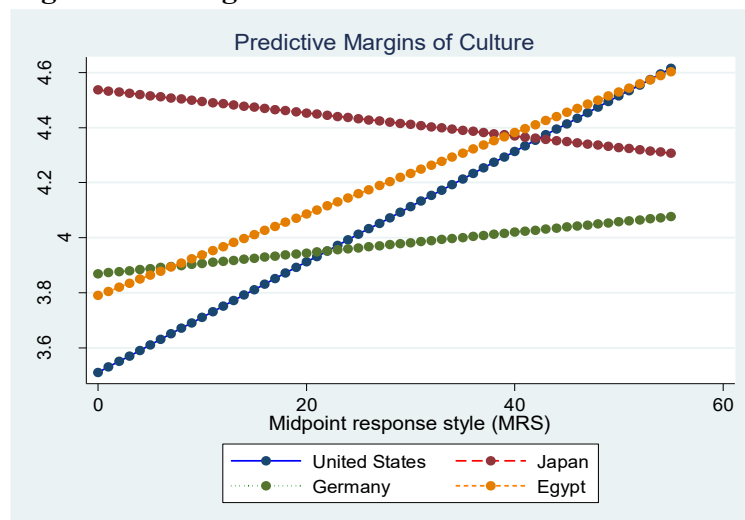
Please note furthermore that the *MRS x culture* interaction for the German sample (*green colored regression slope in **Figure 10***) is not significant, although the slope is not as steep as for the EG- and US-sample. Apart from the effects related to the test family of the variable *culture*, no further effects were found to be significant in the model testing reciprocity deviance relevance.

This brings us to the **heroism** domain. In the 27 tests concerning the *culture* variable, we obtain 9 significant p-values before the alpha-error correction, and 3 remain significant after applying the Holm-Bonferroni method. We found a positive main effect of *culture* (JP-sample; $p < 0.001$) on heroism deviance relevance (reference group US-sample). Additionally, a positive and significant main effect for *age* ($p < 0.001$) and an equally significant ($p < 0.001$) and positive main effect for *NARS* was identified. No significant interaction effects are present in

the model after the alpha-error correction. In addition, the other tests besides the test family relating to *culture* also proved to be non-significant (see: **Table 28**).

The uncorrected **family** deviance relevance model initially shows 8 significant p-values for the tests relating to the variable *culture*. After correcting significance levels two effects remain significant. We found again a positive and significant (p < 0.001) main effect for the *net acquiescence response style* variable (*NARS*). Furthermore, also a negative interaction effect between *culture* (EG-sample) and *level of religiosity* proved to be significant (p = 0.026) (see: **Figure 11**; the regression slope of the EG-sample is colored in *orange*).

The variable *level of religiosity* thus (partially) moderates the effect of *culture* on family deviance relevance for the Egyptian sample. After correcting for alpha-error cumulation, no further significant effects were found in the OLS model concerned with family deviance relevance.

Figure 11: Margins level of religiosity
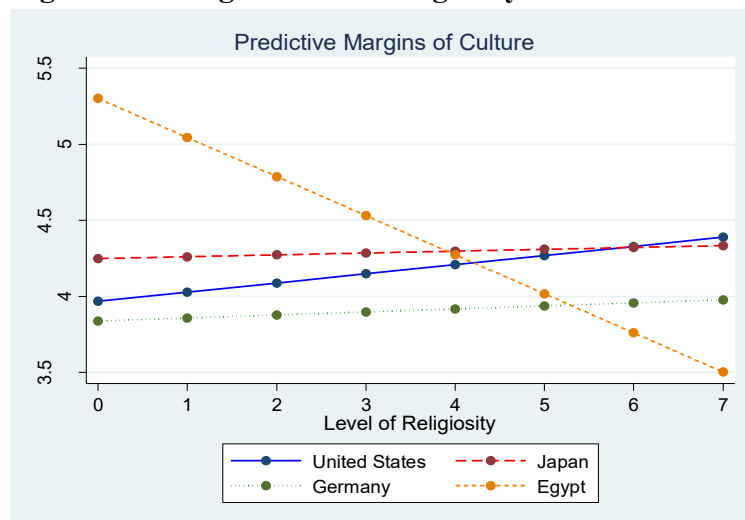
The **in-group** deviance relevance model shows uncorrected 7, and Holm-Bonferroni corrected 4 significant p-values for the tests concerning the variable *culture*. First of all, the results reveal a significant and negative interaction effect for *culture* and the *level of religiosity* variable. This effect is found for both the EG-sample (p < 0.001) and the GER-sample (p = 0.048). Next, we found again a significant and positive (p < 0.001) *NARS* main effect. Furthermore, this main effect is amplified for the EG-sample due to a positive and significant (p < 0.001) interaction for *culture* (EG-sample) and *NARS* on in-group deviance relevance. We graphically display the stated interaction effects in **Figure 12** and **Figure 13** to be found below.

Figure 12: Margins NARS

Predictive Margins of Culture

Net acquiescence response style (NARS)

United States — Japan
Germany — Egypt

Figure 13: Margins level of religiosity
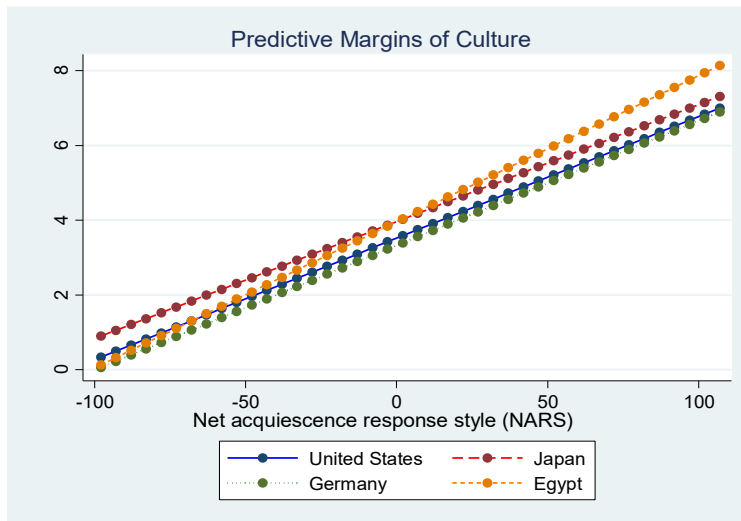
As *Figure 13* clearly shows, the interaction effect with the variable *level of religiosity* on in-group deviance relevance is profound for the Egyptian sample, and also significant for the GER-sample, but not as pronounced. None of the other variables tested revealed any further significant effects.

Predictive Margins of Culture

Level of Religiosity

United States — Japan
Germany — Egypt

Finally, we come to the last of our models and look at **deference** deviance relevance. Before we correct for alpha-error cumulation we find 3 significant effects among the 27 tests related to *culture*. After applying Holm-Bonferroni correction, though, only one effect remains significant: Again, we identified a positive main effect of *NARS* ($p < 0.001$) proving to be most consist predictor across models. What is more, the deference deviance relevance model displays for the variable *education* (ISCED-2011) also significant p-values after the Holm-Bonferroni correction. The category "*Primary education*" serves as a reference, and both the category "*Bachelor or equivalent*" ($p = 0.028$) and the category "*Master or equivalent*" ($p = 0.008$) demonstrate a significant and negative effect on deference deviance relevance. Apart from these findings, there are no further significant effects in the deference model.

4.4.2.1 Discussion — The Covariate Models: Main and Interaction Effects

First and foremost, several variables proved **non-significant** throughout all models. After correcting for multiple testing applying Holm-Bonferroni correction, we found neither a significant main nor a significant interaction effect for the following variables: *Gender*; *Years in School*; *Denomination*, and interaction: *Denomination x Level of Religiosity; Residential Mobility*; *Place of Upbringing*; *Place of Living*; *Pathogen Prevalence*, and *Culture x Pathogen Prevalence*. Even though some of these variables showed a different (positive/negative) direction of effect for the four groups in our study, none of these effects proved to be significant.

Furthermore, we identified several **main effects** impacting on moral deviance relevance across groups. Our results demonstrate significant main effects of the variable *culture* taking the US-sample as reference group. The JP-sample differs from the US-sample with significantly higher values on trustworthiness, property, reciprocity and heroism deviance relevance, when taking only the variable *culture* as main effect into account. Neither the Egyptian nor the German sample revealed any such main effect. Moreover, the variable *age* exerts significant, positive influence on heroism and family deviance relevance. That is to say that the higher the age the greater the relevance to help those in need regardless the own costs. Also, with rising age kinship altruism becomes more important likely revealing the social security function of the family. Another main effect demonstrates significant and attenuating influence of educational degrees attained (variable *education*/*ISCED*) on deference deviance relevance. This effect might be interpreted in the context of socioeconomic positioning. Higher educational attainment most often goes in line with higher socioeconomic positioning, making people more independent (Markus & Schwartz, 2010; Minkov & Kaasa, 2022; Kühnen & Kitayama, 2024), and potentially less obedient and deferring. Interestingly, the effect is found across samples and thus in both the more collectivist JP-sample and the more individualistic GER-sample. Looking at the effects on the different moral domains, it can be seen that the response style variables in the OLS models arguably have the most overarching influence on the different moral domains. *NARS* proved significant and positive across all 8 moral domains assessed, and also the *middle category response style* (*MRS*) is found significant and positive in the trustworthiness and reciprocity OLS models.

However, as we also found **interaction effects** between *culture* and response style measures, these effects are not per se uniform across groups in their influence on the moral deviance margins. Out of 8 interaction effects found in total, 6 are concerned with *culture*

variant effects of response style measures (*MRS* and *NARS*). Accordingly, in addition to the 10 main effects found in total for the *midpoint response style* (*MRS*) and the *net acquiescence response style* (*NARS*), non-uniform but culture-dependent interactions also exert an influence on moral deviance relevance. These non-uniform effects across the cultural groups studied can be seen in ***Figure 6*** to ***Figure 13*** that we have presented above. For example, although *MRS* exerts a positive main effect on trustworthiness deviance relevance (across all groups), this effect is attenuated in the Japanese sample, leading to lower scores at higher *MRS*. The positive effect of *MRS* on trustworthiness deviance relevance is therefore at a certain point reversed in the Japanese sample, indicating a culture-specific moderation.

Next to response style interaction effects, we found two negative interaction effects between *culture* and the variable *level of religiosity*: the higher the *level of religiosity* the lower the deviance relevance margins of the family and in-group domain for Egypt, and the lower the deviance relevance margins for breaches towards the in-group domain in Germany. This finding initially surprised us, as religions tend to emphasize the importance of the family and the religious (in-)group. However, with regard to the reduced family and in-group deviance relevance, the effects found could partly fit into the context of a hypothesis by Lang et al., (2019). These authors investigated the notion of whether beliefs in supernatural monitoring and punishing deities lead to reduced local (in-group) favoritism and were able to find supporting empirical evidence for this hypothesis. Lang and colleagues (2019) state:

> "findings support the idea that the cultural evolution of supernatural agents into punishing and monitoring gods who care about interpersonal, normative conduct may have played a role in the extension of the cooperative circle beyond kin-networks and local ingroup interests" (p. 8).

Although our results for Egypt and Germany are based on a completely different research design and different samples than the investigations by Lang et al., (2019), they at least point in the same direction. However, it remains to question why the effect found for the EG- and the GER-sample was not visible in the samples from the USA and Japan. Consequently, more research is needed to safely integrate the interaction effects for *culture x level of religiosity* found in our study into the hypothesis by Lang and colleagues (see in the same regard also: Purzycki et al., 2018; Henrich, 2020).

Overall, we found that all models that included the interaction terms relating to *culture* had a slightly higher adjusted R^2 and therefore more explained variance than the models without

taking into account these interactions. This finding already speaks in favor of the inclusion of interaction terms. In addition, and more generally, the results of the moderation analysis also suggest that it was largely important to include interaction terms in our OLS models. The analysis showed evidently that we did not only find additive (main) effects, which is why some group specific marginal values would have been under- or overestimated if we had not included the interaction terms. Finally, the analyses just presented suggest also that the effects identified should be taken into account when interpreting the moral deviance relevance margins for the four cultural groups in this study. Nevertheless, we would like to emphasize that although the interaction effects in part moderate the influence of *culture* on the relevance of moral misconduct, they must still be understood as essentially related to *culture*. Albeit the interaction effects allow us to better understand the cultural effects themselves, it would still be erroneous to interpret the moral deviance relevance margins only in terms of the moderating component of the interaction effects. On the contrary, it is the particular *culture* that interacts with certain variables in culturally specific ways. The results of the interaction effect therefore underline the influence of cultural ecology and sociocultural socialization on the human mind. We would like to share one more notion: if we refer to and interpret the moral deviance relevance margins below and do not explicitly mention the moderating components but emphasize *culture*, this implies conversely though that *culture* has also several culture-specific effects with other measured variables in our models, as discussed and demonstrated in the present section.

4.5. Moral Deviance Relevance Across Cultures — Margins on Display

After we inspected the OLS models main and culture specific (interaction) effects, we turn next to the margins and marginal effects (MEs) that we obtain for each of the cultural groups after we fit the respective regression having the 8 moral domains as dependent variables.

Margins can be understood as statistical values resulting from the predictions of a previously calculated (OLS) model(s) for a (set of) covariate(s) with fixed values and the integration of the average of the non-fixed covariates of the model. Moreover, marginal effects are able to express the change in the dependent variable *y* in one single summary value when a variable from the set of covariates changes by one unit (Williams, 2012; StataCorp. 2013; Kohler & Kreuter, 2017). In order to approach our overarching research question and to address the *cultural difference* (*CD*) *hypotheses* that we base on culturally different overall social orientations drawn from theory, we focus on the average marginal effects (AMEs) of *culture*

on moral deviance relevance, while controlling for the influence of the covariates that our OLS models encompass.

We use the margins to determine the deviance relevance values for the different cultural groups on all moral domains measured by MaC-DRS.[93] This is to say that we are interested in the predicted value on the dependent variable(s) — i.e., moral domain specific intuitive deviance relevance — for each category of *culture*, given that we keep all other covariates of our estimation model at their mean values. Both the main and interaction effects as well as the other covariates of our OLS models are taken into account in this regard.

In addition, we also use the *average marginal effects* to compare the four cultural groups in our study in a pairwise fashion. This is done to test for cross-cultural similarities and differences in the effect of *culture* on moral deviance relevance (while accounting for the influences of the covariates).[94] Wooldridge (2016) describes quite clearly how we can obtain the average marginal effect (AME):

> "Often, one wants a single value to describe the relationship between the dependent variable *y* and each explanatory variable. One popular summary measure is the average partial effect (APE), also called the *average marginal effect*. The idea behind the APE is simple (…). After computing the partial effect and plugging in the estimated parameters, we average the partial effects for each unit across the sample. (…) [However, as we] do not want to report this partial effect for each (…) [case] in our sample (…) we average these partial effects to obtain" (pp. 180-181) the (estimated) *average partial effect/average marginal effect* of the sample. In our case we estimate the average marginal effects for the four expressions of the variable *culture*.

Apparently, our most important predictor variable of moral deviance relevance is *culture*. In other words, we are mainly interested in the impact of people´s sociocultural environment on the human moral mind in order to pursue our research question and to examine

---

[93] We can demonstrate full exact scalar measurement invariance for the 8 first-order factors of MaC-DRS relying on the *adjusted sample*. The MaC-DRS measurement invariance test results (*adjusted sample*, N = 2360) for the alternative fit indices are as follows: *configural model*: RMSEA = 0.066, CFI = 0.965, SRMR = 0.038; *metric model*: RMSEA = 0.068 (increase = 0.002), CFI = 0.961 (decrease = 0.004), SRMR = 0.047 (increase = 0.009); *scalar model*: RMSEA = 0.074 (increase = 0.007), CFI = 0.951 (decrease = 0.010), SRMR = 0.050 (increase = 0.003). Consequently, mean comparisons of moral deviance relevance across the four cultural groups of our study are permissible using the *adjusted sample* as data basis.

[94] To obtain the AME for each expression of the variable *culture*, given the set of covariates that our OLS model comprises, we first fix the values on the variable *culture* "but for the other variables we use the observed values for each case. We then compute a predicted probability for each case with the fixed and observed values of variables, and then we average the predicted values" (Williams, 2012, p. 324).

the cultural difference (CD) hypotheses. The categorical variable *culture* has four expressions: Germany (*GER-sample*), Japan (*JP-sample*), the United States of America (*US-sample*), and Egypt (*EG-sample*). As already noted, we are interested in the predicted values of moral domain specific deviance relevance for each expression of the independent categorical variable *culture*, given that we keep all other covariates of our estimation model at their mean values. We use these predicted values (margins) to identify the respective culture-specific moral system prevalent in the groups we examine. Additionally, also to approach our CD hypotheses, we want to know whether the average marginal effects (AMEs) of moral domain specific deviance relevance differ between the cultural entities in our study.

To obtain the corresponding values, we proceed as follows. Our methodological approach is first to determine the AMEs for all four cultural samples in our study for each of the eight moral dimensions. We base this step on the OLS models and the same set of covariates presented and discussed in the section before. So, we first run the covariate OLS model independently for each moral domain assessed and predict thereafter the respective AMEs for each expression of the variable *culture*. Determining the AMEs results also in the predicted values for the expressions of *culture* on moral domain specific deviance relevance while controlling for the other covariates of our models. After obtaining the AMEs, we use them and test each of the 8 moral domains for cross-cultural similarities and differences between the four cultural entities under study. In doing so, we test the average marginal effects (AMEs) of each group against those of the other groups in a pairwise fashion. This procedure is resulting in 6 cross-group comparisons per moral domain.[95]

> "The ME for categorical variables shows how $P(Y = 1)$ changes as the categorical variable changes from 0 to 1, after controlling in some way for the other variables in the model. With a dichotomous independent variable, the ME is the difference in the adjusted predictions for the two groups" (Williams, 2012, p. 323).

This statement about binary categorical variables can be applied to our variable *culture*, for which we examine commonalities and differences between the four groups in pairwise comparisons. In our case, we only have to take into account significance level corrections (Holm-Bonferroni), as we are in the realm of multiple testing by comparing 6 AMEs per moral domain.

---

[95] The 6 comparisons of AMEs per moral domain are as follows: GER-sample vs. JP- sample; GER-sample vs. US-sample; GER-sample vs. EG-sample; JP-sample vs. US-sample; JP-sample vs. EG-sample; US-sample vs. EG-sample.

The following ***Table 29*** shows the marginal values (AME) obtained for each cultural group after running (independent) OLS regressions for the 8 moral domains assessed via MaC-DRS. Consequently, the table shows the average relevance of moral deviance across tested domains and cultures, taking into account the covariates and interaction terms of our models. In addition to the margins, information on the respective pairwise culture/sample comparison in the context of cross-cultural similarities and significant differences is provided for all 8 moral domains. We understand non-significant differences in these comparisons as an indication of cross-cultural similarities in the intuitive relevance of moral deviance. Furthermore, ***Table 29*** uses the following labeling to indicate (Holm-Bonferroni corrected) significance: $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ***Table 29*** displays in addition an ordered comparison of moral deviance relevance margins across the four cultural groups. In this respect, the highest domain specific deviance relevance, in the comparison of the four cultural sub-samples, is marked in *green*. The next highest is marked in *blue*, and followed by the second lowest marginal value in the comparison, which is displayed in *yellow*. Finally, the lowest domain specific deviance relevance margin in the cross-cultural comparison is highlighted in *red*.

Starting from ***Table 29***, we will first look at the results of the pairwise comparisons and discuss them accordingly. Cross-cultural similarities and differences, as well as the effect sizes of the marginal values of *culture* on moral deviance relevance are at the center of attention in this section. Then we return to ***Table 29*** and look at potential patterns of cross-cultural deviance relevance. Findings in this context are discussed in the realm of rising individualism (Cai et al., 2019). Afterwards, we examine and elaborate on our CD hypotheses and the question *which moral system guides cooperation in different cultures* based on the MaC-DRS findings. Eventually we conclude the MaC-DRS analyses with an overarching discussion.

Table 29: Margins: † Adjusted sample analysis of moral deviance relevance across moral domains and cultural groups

Germany n = 666	Samples Compared (pairwise)††	Japan n = 543	Samples Compared (pairwise)	USA n = 569	Samples Compared (pairwise)	Egypt n = 582
4.870 (0.083) **Fairness** †††	GER vs. JP *n.s.* GER vs. USA *** GER vs. EG ***	4.570 (0.126) **Fairness**	*JP vs. USA n.s.* *JP vs. EG n.s.*	4.291 (0.086) **Fairness**	*USA vs. EG n.s.*	4.015 (0.191) **Fairness**
4.601 (0.069) **Trustworthiness**	GER vs. JP ** GER vs. USA *** *GER vs. EG n.s.*	4.993 (0.104) **Trustworthiness**	JP vs. USA *** JP vs. EG ***	4.107 (0.071) **Trustworthiness**	*USA vs. EG n.s.*	4.239 (0.159) **Trustworthiness**
4.772 (0.068) **Property**	GER vs. JP ** GER vs. USA *** GER vs. EG *n.s.*	5.181 (0.102) **Property**	JP vs. USA *** JP vs. EG ***	4.336 (0.070) **Property**	*USA vs. EG n.s.*	4.477 (0.155) **Property**
3.906 (0.060) **Reciprocity**	GER vs. JP *** GER vs. USA * *GER vs. EG n.s.*	4.495 (0.091) **Reciprocity**	JP vs. USA *** JP vs. EG **	3.709 (0.062) **Reciprocity**	*USA vs. EG n.s.*	3.936 (0.139) **Reciprocity**
4.115 (0.063) **Heroism**	GER vs. JP *** GER vs. USA *n.s.* GER vs. EG *n.s.*	4.563 (0.096) **Heroism**	JP vs. USA *** JP vs. EG *	4.072 (0.065) **Heroism**	*USA vs. EG n.s.*	4.062 (0.145) **Heroism**
3.961 (0.068) **Family**	GER vs. JP ** GER vs. USA ** *GER vs. EG n.s.*	4.344 (0.102) **Family**	JP vs. USA *n.s.* *JP vs. EG n.s.*	4.239 (0.069) **Family**	*USA vs. EG n.s.*	4.407 (0.155) **Family**
3.412 (0.066) **In-Group**	GER vs. JP *** GER vs. USA * *GER vs. EG ***	4.042 (0.100) **In-Group**	JP vs. USA *** *JP vs. EG n.s.*	3.603 (0.068) **In-Group**	USA vs. EG *	4.058 (0.152) **In-Group**
3.374 (0.068) **Deference**	GER vs. JP *** GER vs. USA* GER vs. EG **	4.043 (0.103) **Deference**	JP vs. USA *** *JP vs. EG n.s.*	3.603 (0.070) **Deference**	*USA vs. EG n.s.*	3.960 (0.156) **Deference**

† **Note:** The margins presented are statistical values calculated from predictions of the respective previously fit OLS-model at fixed values of the covariate culture and the integration of the average of the non-fixed covariates of the model (Average Marginal Effect). The margins for moral deviance relevance (dependent variables) are shown for the cultural groups Germany (GER), Japan (JP), United States of America (US) and Egypt (EG). Each moral domain was tested in an independent OLS. The (delta method) standard errors are displayed in parentheses; †† **Note:** To test for significant differences in moral deviance relevance (margins), cultural groups were compared in a pairwise fashion. The asterisks indicate (Holm-Bonferroni corrected) significance: * p <0.05, ** p <0.01, *** p <0.001. Non-significant comparisons (*n.s.*) are highlighted in italics; ††† **Note:** colors indicate score ranking in country comparison (Green > Blue > Yellow > Red).

4.5.1. Pairwise Comparisons: A First Look at Commonalities and Differences in Moral Deviance Relevance

First, we turn to the cross-cultural *commonalities* and *differences* through **pairwise comparisons** of the cultural groups. In these pairwise comparisons, we refer exclusively to the deviance relevance marginal values obtained via the OLS models, as explained above. The results of the pairwise comparisons, reduced to the respective significance level, can be found in *Table 29*. Please note, that the significance level was adjusted for multiple comparisons (alpha-error cumulation) using the Holm-Bonferroni correction. Overall, we find more significant cross-cultural differences than commonalities with regard to the intuitive relevance of moral deviance. However, a considerable proportion of moral transgressions are also of intuitively similar relevance across the cultures studied. Of the total of 48 pairwise comparisons between the cultural entities, 28 show significant differences and, accordingly, 20 non-significant cross-cultural evaluations of the relevance of morally misconduct emerge. We interpret non-significant differences as cross-cultural similarities. These cross-cultural similarities in the relevance of moral deviance are not evenly distributed across the four cultural groups in our study, as we shall see.

With 8 moral domains and three comparison groups each, this results in a total of 24 pairwise comparisons per sample. The **GER-sample** shows a total of 17 significant differences in the relevance of moral deviance compared with the other groups. Accordingly, there are also 7 non-significant pairwise comparisons. Five of these non-significant findings are attributable to the comparison with the EG-sample and one non-significant comparison each to the JP- and the US-sample. With regard to the individualizing moral domains of fairness, trustworthiness and property, the margins for the GER- and JP-samples are similar in their overall direction despite significant differences in magnitude on two domains. However, as far as the binding domains family, in-group and deference are concerned, we see a clear and pronounced polarization in the comparison of these two cultural entities. Interestingly, there are also significant differences in 7 out of 8 comparisons of the GER- and the US-sample. This result already indicates that WEIRD-samples are not a priori to be assumed as homogeneous. Overall, most commonalities for the German sample are found with Egyptian sample.

As far as the intuitive evaluation of moral deviance relevance is concerned, the Japanese **(JP-) sample**, together with the GER-sample, shows the most differences among study groups. Our analysis yields a total of 17 significant differences for the JP-sample in the pairwise comparison of samples. This contrasts with 7 non-significant pairwise comparisons of moral

deviance relevance. In line with what has been stated above, cross-cultural similarities between Japan and Germany can only be found in relation to one moral domain namely fairness. The US-American sample (2 non-significant values) and the Egyptian sample (4 non-significant values) thus show more cross-cultural similarities with the JP-sample than the GER-sample does. It should be emphasized at this point that the Japanese and Egyptian samples do not differ significantly with regard to the domains of family, in-group and deference, and consequently show a similar deviance relevance for binding morality. The same cannot be said for the comparison with the GER-sample. Germany differs strongly from the Japanese sample in terms of binding morality, as demonstrated by significantly lower deviance relevance margins for all binding domains. Regarding binding morality, the US-sample, however, only shows significantly lower values than the JP-sample in the context of in-group and deference deviance relevance. Hence, the US-sample is not as different from the JP-sample as the German sample. Again, most commonalities can be found between the JP- and the EG-sample.

In the pairwise comparison of cross-cultural similarities and significant differences, the **US American sample** comprises a total of 14 significant differences and 10 non-significant comparisons with the other groups examined. What stands out is that of these non-significant comparisons, 7 are attributable to the comparison with the Egyptian sample. In other words: there is only one significant difference between the US American and the Egyptian sample in our study, namely that of in-group deviance relevance. In terms of the intuitive relevance attribution of moral violations in different domains, these samples appear to be quite similar. Our data suggests accordingly that the moral systems of the US- and EG-sample seem to be largely similar with regard to the relevance of moral breaches. However, as already mentioned above, the US- and GER-sample differ significantly in 7 out of 8 comparisons of deviance relevance margins. As there are only 6 significant differences with the JP-sample, the German sample shows the most differences to the US-sample in comparison to all groups. This finding is evidently pointing to the fact that WEIRD cultures are heterogeneous among themselves.

In the fourth and final group we examined, we found more cross-cultural similarities than differences across all pairwise comparisons. A total of 16 out of 24 pairwise comparisons across all measured moral domains are non-significant for the Egyptian (**EG-)sample**. Out of these, 7 comparisons with the US-, 5 with the GER-, and finally 4 with the JP-sample are not significant. As far as the relevance of moral breaches is concerned, the predicted values (margins) of the EG-sample are relatively between those of the other samples. However, there is also a tendency for the Egyptian sample to evaluate deviance in terms of binding morality similarly to the JP-sample, which partially polarizes it away from the GER-sample. The latter

evidence is in line with our heuristic hypothesizing and will be discussed in more detail below. All in all, when examining the relevance of moral deviance, the Egyptian sample shows the most similarities overall with the other cultural groups. In this light, however, we should not forget the composition of the EG-sample, which will be again addressed in more detail later.

4.5.1.1. Discussion: A First Look at Commonalities and Differences in Moral Deviance Relevance

Overall, we found more cultural differences than similarities in the intuitive evaluation of the relevance of moral deviance. Thus, our results already speak strongly in favor of our **main CD hypothesis** and suggest that *there are indeed significant differences in the relevance of moral misconduct across cultures.* Regarding cross-cultural *similarities*, i.e., non-significant differences in marginal values, the US- and EG-samples are partly positioned between the German and the Japanese samples. Egypt and US America show the most commonalities in terms of moral deviance relevance across cultures studied. Focusing in turn on the cross-cultural *differences*, we find a polarization of the moral deviance relevance margins: in terms of the number of significant differences and the size of these differences, the GER- and JP-sample in our study differ the most. However, we were also able to identify the same number of differences between the GER- and the US-sample. These findings taken together are interesting in the light of the fact that we are investigating two WEIRD (GER and US) (Henrich et al., 2010; Henrich, 2020) and two non-WEIRD (JP and EG) samples, for it illustrates that a simple binary classification of cultures is clearly too short-sighted, as also emphasized by Apicella and colleagues (2020). Both non-WEIRD and WEIRD samples are heterogeneous, at least as far as the aspect of moral deviance relevance is concerned. Accordingly, we therefore also share the position of the canon that emphasizes how important it is to refrain from overly simplistic classifications of cultures into WEIRD vs. non-WEIRD in cross-cultural analyses and theories.

As already mentioned, there are more differences than similarities in moral deviance relevance between the cultural groups in our study. However, this should not obscure the fact that in none of the cultural samples a complete irrelevance with regard to moral domain specific deviance can be observed. Cultural entities therefore differ *relatively* in their attribution of intuitive relevance when it comes to moral breaches. But moral deviance appears never completely irrelevant. This statement is true at least across the four cultural groups we study, and the eight domains covered by MaC-DRS. Our finding underpins hence the results of the MaC-DRS CFA (see: **Chapter 3**), and indicates that the human moral mind is calibrated across cultures to recognize moral deviance and to attribute a relative, culturally mediated relevance

to it. In this respect, the universalism of our pan-human moral mind is once again revealed. Nonetheless, although moral deviance is not irrelevant across cultures, the respective socio-cultural environment exerts a pivotal influence on *how* relevant moral breaches are intuitively. In order to shed more light on this aspect, we will now briefly focus on the strength of the effect of *culture* on moral deviance relevance.

4.5.2. The Effect of Culture on Moral Deviance Relevance

The MaC-DRS results not only reveal cross-cultural differences, but also that a respective culture does indeed have a profound influence on the calibration of people´s moral mind. To use Haidt and Joseph's (2007) metaphor in this context, our findings suggest that a person's cultural environment actually exert a force of (cultural) editing of the first (universal) draft of the moral mind. The moral domains, according to the MaC-DRS results for the four heterogeneous cultural groups of this study, are universal, but socialization into different cultural environments and moral imperatives causes a different prioritization of moral domains. This is evident in cross-cultural differences in moral deviance relevance between the four sampled groups. Culture thus exerts imprinting influence on our moral mind by calibrating the extent to which people intuitively evaluate deviant behavior as relevant with regard to different moral domains — this finding points also to differences in the configurations of moral systems across cultural entities.

What we have just described is clearly supported when we look at the effect size (Cohen's d) of the marginal value of *culture* on the dependent variable *deviance relevance* across all moral domains tested. ***Table 30*** portrays the effect size calculations of the pairwise comparisons of the marginal values for moral deviance relevance.[96] The respective significance level, together with the size of the effect, indicates the relative meaningfulness of an effect (Mayr et al., 2007; Lakens, 2022). In order to highlight the findings that by evidence must be considered meaningful, we have marked non-significant results in *gray* and highlighted significant results in *black* in the table below.

---

[96] To calculate the effect size *Cohen's d*, we use the margins of moral deviance relevance estimated with our OLS covariate model as mean values and their standard deviation. We have also taken into account the different sample sizes to calculate effect sizes. Corresponding values can be found in ***Table 29***. Cohen's d then results (for example) as: (margins fairness deviance relevance GER-sample - margins fairness deviance relevance JP-sample) / pooled standard deviation, and taking into account the respective sample size n. All effect sizes were calculated using the effect size calculator for comparisons of groups with different sample sizes to be found under: (https://www.psychometrica.de/effect_size.html).

Table 30: Effect size (Cohen´s d) of culture on moral deviance relevance (margins) across moral domains and cultural groups

Predictor Variable	Dependent Variable	Cohen´s d†	Cohen´s d	Cohen´s d
Culture on:	*Fairness deviance relevance*	GER vs. JP: d = -2.87	JP vs. US: d = -2.598	US vs. EG: d = -1.856
Culture on:	*Fairness deviance relevance*	GER vs. US: d = -6.861	JP vs. EG: d = -3.407	
Culture on:	*Fairness deviance relevance*	GER vs. EG: d = -5.944		
Culture on:	*Trustworthiness deviance relevance*	GER vs. JP: d = 4.532	JP vs. US: d = -9.993	US vs. EG: d = 1.068
Culture on:	*Trustworthiness deviance relevance*	GER vs. US: d = -7.064	JP vs. EG: d = -5.574	
Culture on:	*Trustworthiness deviance relevance*	GER vs. EG: d = -3.024		
Culture on:	*Property deviance relevance*	GER vs. JP: d = 4.814	JP vs. US: d = -9.701	US vs. EG: d = 1.143
Culture on:	*Property deviance relevance*	GER vs. US: d = -6.325	JP vs. EG: d = -5.233	
Culture on:	*Property deviance relevance*	GER vs. EG: d = -2.523		
Culture on:	*Reciprocity deviance relevance*	GER vs. JP: d = 7.8	JP vs. US: d = -10.138	US vs. EG: d = 2.101
Culture on:	*Reciprocity deviance relevance*	GER vs. US: d = -3.233	JP vs. EG: d = -4.726	
Culture on:	*Reciprocity deviance relevance*	GER vs. EG: d = 0.287		
Culture on:	*Heroism deviance relevance*	GER vs. JP: d = 5.633	JP vs. US: d = -6.016	US vs. EG: d = -0.089
Culture on:	*Heroism deviance relevance*	GER vs. US: d = -0.673	JP vs. EG: d = -4.047	
Culture on:	*Heroism deviance relevance*	GER vs. EG: d = -0.485		
Culture on:	*Family deviance relevance*	GER vs. JP: d = 4.508	JP vs. US: d = -1.211	US vs. EG: d = 1.395
Culture on:	*Family deviance relevance*	GER vs. US: d = 4.061	JP vs. EG: d = 0.477	
Culture on:	*Family deviance relevance*	GER vs. EG: d = 3.815		
Culture on:	*In-group deviance relevance*	GER vs. JP: d = 7.59	JP vs. US: d = -5.156	US vs. EG: d = 3.85
Culture on:	*In-group deviance relevance*	GER vs. US: d = 2.854	JP vs. EG: d = 0.124	
Culture on:	*In-group deviance relevance*	GER vs. EG: d = 5.645		
Culture on:	*Deference deviance relevance*	GER vs. JP: d = 7.904	JP vs. US: d = -5.018	US vs. EG: d = 2.942
Culture on:	*Deference deviance relevance*	GER vs. US: d = 3.374	JP vs. EG: d = -0.624	
Culture on:	*Deference deviance relevance*	GER vs. EG: d = 5.012		

† Note: We used the margins (obtained via the OLS covariate model) as means, pooled standard deviations of the margins, and the respective sample size of each cultural group (i.e., we took the different samples sizes into account in the comparisons) to calculate Cohen´s d for pairwise sample comparisons between the GER-, JP-, US-, and EG-Sample. *Table 29* further above displays the margins and standard deviations for the respective moral domains and also the sample size per cultural group. **†† Note:** The pairwise comparisons of the moral deviance relevance margins marked in *gray* are not significant whereas those marked in *black* are significant.

Our results demonstrate that not only the majority of cultural comparisons show significant differences. But also, that the effect of *culture* on moral deviance relevance is strong and substantial in most of the pairwise comparisons of the samples. As can be seen in ***Table 30***, some of the effect sizes even reach up to over 9 standard deviations, which illustrates clearly the extent of influence a particular culture can exert in calibrating the human moral mind. Although some cross-cultural findings are not significant, the impact of culture on moral deviance relevance is yet mostly large in the size of the effect. In our study only 7 of 48 pairwise comparisons in total show effect sizes of less than one standard deviation. Among these 7, there are only two comparisons (US-sample vs. EG-sample: heroism deviance relevance; JP-sample vs. EG-sample: in-group deviance relevance) that show negligible and insignificant effects of culture, and one comparison (GER-sample vs. EG-sample: reciprocity deviance relevance) that yields a small non-significant effect of culture. We found also 4 pairwise comparisons that display a medium to strong (below one standard deviation) yet non-significant effect of culture on moral deviance relevance (see: heroism, family and in-group domain results in ***Table 30***). With a total of 41 out of 48 cross-cultural comparisons, out of which 28 are significant, a major share of the comparisons between margins shows a meaningful and large effect of culture on deviance relevance across moral domains. Thus, we have found clear evidence that the sociocultural world we grow into not only shapes our reality (Berger & Luckmann, 2013), but also calibrates the intuitive moral tendencies that are deeply entrenched in the evolved neural network of the human moral mind. In summary, our findings reveal cultural differences in intuitive moral deviance relevance and show that the influence of culture is profound.

4.5.2.1. Discussion: The Effect of Culture on Moral Deviance Relevance

Our results clearly demonstrate that culture is potent to exert a considerable influence on intuitive evaluations of the relevance of moral misconduct. But why do cultural entities differ in moral deviance relevance if the moral domains themselves are universal to the moral mind of our species? The human specialization in culture developed as an evolutionary solution to the challenges of reproduction and allowed the colonization of diverse natural ecologies by members of our species (Henrich, 2016). Due to these diverse ecologies, various cumulatively growing cultures also developed as human habitats, which in turn follow their own path dependencies (Mesoudi & Thornton, 2018). Cultural environments are social environments. In their social environments, early humans were repeatedly confronted with specific social

challenges concerning their reproduction. As Tomasello and Vaish (2013) have pointed out, a significant part of fitness-relevant and recurring challenges includes that of human *cooperation*, which ultimately contributed to the development of our moral mind. Furthermore, due to independent path dependencies of cultural entities, cultures are confronted with different types of social challenges and thus also with different types of relevance of moral domains. What has occupied the space of an important challenge in one culture over the course of socio-cultural development may also be relevant in another culture, but the same degree of relevance cannot be assumed a priori. Against this background, Curry (2016), arguing from Morality as Cooperation Theory (MaC), states the following: "the prediction is that, to the extent that different people and different societies face different portfolios of problems, different domains of morality will loom larger—different cultures will prioritise different moral values" (p. 40).[97] The same assumption can also be found in the other theoretical pillar — Moral Foundations Theory (MFT) (Haidt & Joseph, 2007) — on which the moral approach we are advocating is partly based. Against this background, we are able to confirm both moral universalism and culture-specific organization of moral systems in our study.[98] Our CFA models for MaC-DRS show that the theoretically conceived 8-dimensional moral structure in the four cultural entities of our study fits well with the empirical data. In this respect, we can reasonably expect that the 8 moral domains also exist across the different cultural samples of our study (see: **Chapter 3**). However, it can be seen from the *Tables 29* and **30** that despite the cross-cultural existence of the 8 moral domains, moral deviance relevance differs in large parts and substantially across cultures. Thus, theoretical pillars of MFT and MaC are evident in our data: human morality is both universal and culturally specific. Without having referred to the CD sub-hypotheses so far, our results already suggest that we can confirm our **main cultural difference (CD) hypothesis**. Furthermore, our data point to cross-cultural trends in relation to moral deviance relevance. This topic will now be discussed below with reference to collectivism and individualism.

---

[97] Note: Just as it is with the self — "all cultures include a great variety of contexts that require and hence foster either independence or interdependence" (Kühnen & Kitayama, 2024, p. 7) — it is also with our moral mind. This is to say that culture creates *relative* differences in the calibration of our moral mind, because the basic challenges of human cooperation can be expected to be universal to our species. However, the latter is not to be equated with the idea that these challenges of cooperation had and have the same, uniform social relevance across cultural ecologies.

[98] Although we are aware that our study covers a very limited sample, the MaC-DRS findings nevertheless provide a strong indication of the universality of the eight moral domains covered. See *Chapter 3* further above.

4.5.3. Cross-Cultural Deviance Relevance of Moral Domains

When we compare the deviance relevance margins (*Table 29*) across cultural groups and among the 8 moral domains captured, the following becomes apparent: the **property** domain proves to be the most relevant. The respective margins obtained via the OLS-covariate model are as follows: GER = 4.772; JP = 5.181; US = 4.336; EG = 4.477. With the exception of the German sample, property deviance relevance consistently ranks highest in the four cultural samples. In the GER-sample, property takes second place behind fairness in the within-sample ranking of moral domains. Accordingly, moral breaches in the realm of (private) ownership thus appear to be of intuitive relevance across the modern societies compared. At this point, we emphasize the aspect of *modern* societies, which we will take up and elaborate on in the interpretative discussion of these findings further below.

Evidence also demonstrates that **trustworthiness** deviance is of intuitive relevance in all cultures studied, despite the differences between the cultural samples. The trustworthiness deviance relevance margins of the individual groups are as follows: GER = 4.601; JP = 4.993; US = 4.107; EG = 4.239. We expected a slightly different distribution of relevance in terms of trustworthiness deviance in the non-WEIRD samples, partly due to differences in historical kinship intensity (Schulz et al., 2019; Enke, 2019) and relational mobility (Thomson et al., 2018). However, according to the data, presenting oneself as a reliable cooperation partner — to keep promises, commitments and secrets — seems to be of greater relevance across cultures in modern societies.

In addition to property and trustworthiness, deviations from the third individualizing domain of **fairness** are overall also intuitively relevant across cultures. Analysis yields the following fairness deviance relevance margins: GER = 4.870; JP = 4.570; US = 4.291; EG = 4.015. The GER-sample stands out particularly in this context. In this sample the fairness domain shows the highest deviance relevance across cultures and likewise in the within-sample comparison of moral domains. However, fairness violations also rank second in the US American within-sample ranking of moral domains, and third in the Japanese within-sample ranking. In the Egyptian sample fairness deviance relevance takes the sixth place in the within-sample ranking, although admittedly only tiny difference exists to the margins of the fifth (heroism) and fourth (In-group) place. Altogether, though, it is primarily the two WEIRD samples, and also Japan, in which fairness deviance relevance is given relatively great weight among moral domains.

The margins of the **heroism** domain also suggest that this domain seems to be intuitively relevant across cultures. Although Japan stands out with higher margins, the general direction of deviance relevance across the four cultural samples is broadly similar: GER = 4.115; JP = 4.563; US = 4.072; EG = 4.062. By implication, these findings suggest that standing up for others in need is of cross-cultural relevance, even though such deeds may come with own costs and risks. This fact is particularly interesting against the background of the construction of the MaC-DRS heroism items, as they are designed to ask about deviance relevance in the context of protecting family members, friends and strangers under the risk of own sacrifice.

Finally, we would like to take a look at **individualizing** and **binding** approaches to morality. As already described, the data shows that deviance towards *individualizing morality* — i.e., property, trustworthiness and fairness — is in large parts significantly more relevant than deviance towards binding morality (especially in-group and deference). We find it remarkable that this result holds true for the most part for *all* four cultural groups in our study. Admittedly, there are limitations to the individualizing statement: the family domain is also of intuitive deviance relevance, as can be seen in the EG-, US- and JP-samples. Also, the EG- and JP-sample exhibit a slight tendency towards intuitive in-group deviance relevance. In Japan, we found furthermore a tendency of intuitive deference deviance relevance.[99] However, the in-group and deference domains generally rank at the lower end amongst the deviance relevance margins of the four cultural samples. Overall, our study reveals thus a general tendency suggesting that binding deviance seems to be less relevant than individualizing deviance in the cultural groups studied.

With regard to the binding domains of family, in-group and deference, another fact stands out that has already been touched on above and that is relevant to our hypotheses (see:

---

[99] At this point, we pick out the *deference* finding for Japan and the corresponding discussion of the finding: As far as deference is concerned, this moral domain seems to play an important role in collectivist, rather hierarchically organized social structures and forms of society that center around interdependence of self and have historically cultivated paddy rice (Markus & Kitayama, 1991; Triandis, 2001; Talhelm, 2022; Kitayama & Salvador, 2024). In such societies, in our study this is the JP-sample, deference is functional and supports the success of cooperative processes (Henrich, 2020). In individualistic societies, deference also has its value without question, but seems to be of less relevance, likely because in these societies individuals interact and cooperate primarily on the basis of their own, but not collective, goals. In short, we believe that deference can be seen as a sign of connectedness that signals a willingness to cooperate and to follow collectively shared norms in hierarchical collectivist cultures characterized by historical paddy rice cultivation and an interdependent self-construal. In contrast, deference may rather be considered a matter of choice (Markus & Schwartz, 2010) in more individualistic cultures marked by independent self-construal and historical reliance on wheat farming and herding. Given the latter sociohistorical conditions, deference can be granted as a matter of choice but only lesser intuitive relevance is attributed to this moral domain. This might help explain the relative high deviance relevance of the deference domain in the JP-sample compared to the relative low deviance relevance in the GER- and US-sample, having the EG-sample in between this polarization.

CD sub-hypotheses 1 and 3a): the JP- and EG-sample do not differ significantly from each other in any of these moral domains. As far as *binding morality* is concerned, these samples therefore show an overlapping tendency of moral deviance relevance. However, in comparison to the WEIRD GER- and US-samples, significant differences in relation to binding morality can be found. These differences are particularly evident when comparing the German and the Japanese sample which differ significantly ($p < 0.01$ to $p < 0.001$) in terms of the relevance attribution to breaches of binding morality. Taken together, in comparison to the JP- and EG-samples and excluding the family domain, the WEIRD samples of our study have a tendency to evaluate binding deviance as less relevant.

4.5.3.1. Discussion: Cross-Cultural Deviance Relevance of Moral Domains

Our data show that ***property*** deviance is overall most relevant across the moral domains assessed via MaC-DRS and the cultural groups under investigation. In our view, this fact can be explained in part by three considerations. First, respect for ownership is widespread in the animal kingdom and has also existed in humans for many thousands of years. A second consideration concerns the cross-cultural institutional protection of the property domain. Third, we argue that the domain of property is fundamental to the success of cooperative ventures in probably most *modern* societies, which are characterized by historical processes of resource accumulation, transgenerational property inheritance, and division of labor.

First, and as already indicated, there is evidence for property recognition even long before human modernity and also found among animals. Loss of possession may be associated with the cost of losing the resource in question, may impose a fitness disadvantage if antagonists gain access to the resource in question, and may increase perceived exploitability if the loss of possession is due to theft or other forms of property deviance (Asao & Buss, 2016, p. 17). From a Machiavellian perspective on morality, humans are hence motivated to protect property and punish deviance. Numerous game theoretical analysis, but also observations in the animal kingdom, support that a disposition to respect ownership develops under a wide range of conditions. Most often, but not absolute, recognition of ownership builds an evolutionary stable strategy (Sherratt & Mesterton-Gibbons, 2015). Signaling one´s will to defend property, the value of the property itself, advantages due to the property, coordinated defense of property, information about the value of property, experience gained by and costs invested in obtaining property, and many more aspects render respect for ownership to be in most cases (i.e., partially) the best strategy. In other words, recognizing ownership usually yields higher pay-offs than

contesting property claims (Sherratt & Mesterton-Gibbons, 2015). Furthermore, in the case of humans, the recognition of property must also be seen in the context of the fact that even before the Neolithic period, humans were partly able to produce surplus resources due to technological progress and stabilizing environmental conditions. As a result, surplus resources were no longer only used for supply, but also for exchange. It was probably in this context that the first forms of property rights emerged as conventions aimed at recognizing ownership among people (Huinink & Schnettler, 2024). There are therefore many reasons to believe that a disposition to respect property has developed evolutionarily, affecting both property owners and antagonistic contestants, and has a long history among humans. These evolutionary insights can partly explain the widespread importance of property and the cross-cultural relevance of transgressions for respect of ownership.

Second, the institutional consideration. Property, the gifting of property and competition for property can have a considerable influence on people's social status and even structure entire social systems (Maus, 1968). Given the relevance of property for people and their social systems, it is therefore not surprising that the domain of property finds manifestation in social institutions. The recognition of property is arguably anchored in most legal texts across human societies and violations to this (legal) domain are subject to institutionally organized punishments (Curry, 2016). The well-founded assumption of a cross-cultural institutional safeguarding of the property domain, which comes into effect in particular through socially organized punishments and also role bearers who enforce these punishments, could in part explain its cross-cultural deviance relevance. Furthermore, this institutional protection may also be an expression of the relevance of the property domain itself. If we follow this interpretation, the data suggest that it is of great social relevance to clarify property claims if we assume — and we must likely do so for most societies in the world — that ownership is not repeatedly set to zero in a tabula rasa fashion.

Third, in line with the Morality as Cooperation Theory (Curry, 2016) we treat property as a domain of cooperation that is regulated by our moral mind. Evolved human morality, which is part of our psyche and which finds its manifestation inter alia in institutions and institutional policies, guides inter-individual interaction in such a way that cooperation is possible. Social order arises from this function (Ellemers et al., 2019): not only do we have a disposition and, in part, an experiential knowledge of what we should not do to others, we also grow into a social world in which our fellow human beings have the same disposition and a similar experiential knowledge. As a result, our morality enables us to enter into and cooperate with others in a space of interaction that is significantly reduced in its complexity. A part of this complexity

reduction and cooperation opportuneness is brought about by the moral domain of property: certain dilemmas of cooperation can be transformed into manageable channels by the moral domain of property, as it regulates claims to tangible and intangible objects (in the broadest sense). This regulation not only reduces the complexity of action but also the potential for conflict about objects decreases. To put it in a nutshell and (overly) simplistic, we partly possess a disposition to recognize property because "deference towards owners can evolve as a convention to reduce costly fighting" (Sherratt & Mesterton-Gibbons, 2015, p. 1197). Property morality supports to solve cooperation dilemmas that can arise when (in the broadest sense) the use of goods is necessary for the success of a cooperative venture. In other words, and by way of example, we could say: you and I could hunt a prey in a joint venture that we could not catch on our own. However, we can only hunt this prey with the help of tools such as spears. In this example, you have been given hunting tools by your social allies or parents, but I have not — the relation between property and inequality is well documented for the last 15 thousand years of human history (Huinink & Schnettler, 2024). Thanks to our property morality, you can now provide me with these tools necessary for the success of our hunt, because you will get them back. This drastically reduces complexity and potential for conflict. Thus, our intrapsychic and socially anchored moral system ensures property recognition and so you will (most likely) get the tools back from me after our joint hunt. In this example, it is our respect for property that facilitates the likelihood of a cooperative enterprise that requires objects/goods/resources to succeed. The example given here illustrates relatively simply the obvious cross-cultural benefits that the property morality has even in comparatively simplistic, primal scenarios of cooperation. However, this example can easily be transferred to cooperative situations in modern societies in which, for example, means of production are used for the success of diverse cooperative ventures. Objects, in the broadest sense, can be provided by the owner in the context of cooperative ventures and yet remain in their possession or return to them, and this is safeguarded not only by law but also our evolved moral mind. The morality of property thus (still) plays a relevant role across modern societies in order to enable a variety of cooperative opportunities that, on the one hand, require a division of labor for successful cooperation and, on the other hand, start of from a basis of non-equally distributed means of production. If we now assume that division of labor and unequal distribution of the means of production characterize the status quo of most people´s daily working lives in modern societies, this helps also to explain the cross-cultural relevance of deviance in relation to property. Note: All this happens with reduced complexity and lower risks of conflict, because property morality

regulates our claims to objects, and compliance with it is considered good across cultures (Curry et al., 2019a).

All in all, respect for ownership proves to be an evolutionarily stable strategy under most conditions and has a long history among humans. Property recognition is important in the context of prestige and social positioning, enables cooperation under conditions of resource inequality in modern societies, and is manifested and secured in human institutions. These approaches taken together may partly explain the strong cross-cultural deviance relevance of the property domain that we found.

Besides property deviance relevance, the findings show also that breaches towards the **trustworthiness** domain are of cross-cultural relevance. Remember, we identified reliability as the core aspect of the trustworthiness domain, and consider trustworthiness as a disposition that grants reliability in the first place yet also updates initial granting based on experiential knowledge acquired in one´s life. At first, we thought that trustworthiness relevance applies more to cultural entities with high relational mobility (Thomson et al., 2018; Talhelm, 2022). However, the data demonstrates that reliability violations are of importance across societies characterized by different levels of relational mobility. Hence, we interpret this finding as follows: in both, relatively low relational mobility societies (JP- and EG-sample) and relatively high relational mobility societies (GER- and US-sample), it is important to keep selfish drives regarding one´s reliability in check. In the following we draw two complementary explanations of trustworthiness deviance relevance in the realm of relational mobility.

First, we argue that in societies with *low relational mobility*, it is particularly important to pay attention to reliability, in order to minimize problems of potential non-selection for cooperative ventures and to avoid partner-selection of uncooperative/selfish individuals. The reliability of individuals affects their reputation and thus marks potential cooperation partners as more or less suitable for joint ventures. In societies with low relational mobility, the focus is not primarily on the initial granting of trustworthiness, which also plays a role, but it is rather on the component of experiential knowledge. Since in societies with low relational mobility the relationship structures are more fixed rather than fluid, actors are encouraged to pay particular and intuitive attention to experiences of trustworthiness violations: it is better to build one's cooperation network with those that one can rely on, especially when relationship exodus is difficult and rarely possible. Assigning intuitively trustworthiness deviations with relevance is therefore functional in societies with low relational mobility for two reasons: it promotes, on the one hand people's tendency to be reliable, which contributes to them being chosen as (long-term cooperation) partners. On the other hand, trustworthiness deviance relevance promotes the

protection of one´s own network of cooperators against bad apples when choosing new partners is difficult. The findings of trustworthiness deviance relevance in the low relational mobility JP- and EG-samples at least allow for such an interpretation of the data.

Second, in societies with relatively *high relational mobility*, however, trustworthiness is also important as it encourages people to enter into initial cooperative relationships, even with strangers or casual acquaintances. Everyday opportunities to interact with large numbers of people, some of whom are likely to be strangers, are a characteristic of societies with high relational mobility (Thomson et al., 2018; Henrich, 2020). Such societies are accompanied by an opportunity structure of cooperation that presupposes intrapsychic and socially normative supports that make cooperation with strangers or at least only fleetingly known, fluctuating cooperation partners possible in the first place. The moral domain of trustworthiness offers such intrapsychic support. By disposition trustworthiness prompts us to grant potential cooperation partners, including strangers, reliability before acquiring experiential knowledge. This disposition enables cooperation with strangers in the first place. Nonetheless, this initial grants of a leap of faith are then adapted to the lived reality of the respective social environment by adding experiential knowledge. The extent to which and when initial cooperation with strangers can actually be entered into is fine-tuned in the course of life through experiencing the lived societal norm of reliability.[100] Thus, although the aspect of granting reliability in advance is of inherent value in societies with high relational mobility, the actors are anything but blind to experiences that mark others as (un)trustworthy. Rather people´s psych is calibrated to pay particular attention to the magnitude of reliability breaches so as to adapted one´s own granting of reliability to the lived reality of fellow human beings. Derived from this argumentation, and this is supported by our data, also societies characterized by relatively high relational mobility, as are the GER-and the US-sample, consider trustworthiness deviance as of relevance.

> Note: Due to the historically relatively recent development of online interaction, most people are confronted with a wider range of potential cooperation partners and a greater complexity of partner selection than in the past. Against this background, it may be important to signal one's own trustworthiness more strongly in order to be selected for cooperation and to protect one's own social network of cooperation contacts from migrating to other cooperative ventures. It is therefore theoretically conceivable that the relevance of trustworthiness deviations is increasing across cultures due to the historically recent development of online interaction. Future studies could aim to investigate whether and how online interaction influences trustworthiness deviance

---

[100] We see the basis of the fine-tuning in the oaths and actions that take place in the actors' social environment. In the course of their lives, actors accumulate experiential knowledge about the degree of initial reliability granting in their society, mediated by injunctive and, above all, descriptive norms, and adapt their disposition of reliability granting on this basis (see on norms: Reno et al, 1993; Cialdini & Trost, 1998; Lapinski & Rimal, 2005; Hogg & Reid, 2006; Schultz et al, 2007; Henrich & Chudek, 2011; Reynolds et al, 2015).

relevance cross-culturally. Apart from the online sphere, we also believe that a closer look at the aspect of market integration, i.e. the "breadth and intensity of market exchange" (Henrich et al., 2010b, p. 1480), as well as the study of small-scale societies will contribute to a deeper understanding of cross-cultural trustworthiness deviance relevance.

The **fairness** domain also exhibits cross-cultural deviance relevance. So, how can this cross-cultural fairness deviance relevance be interpreted? In 2008, Haidt assigned the fairness domain theoretically to the individualizing approach to morality, because the regulation of selfishly motivated deviance in relation to fairness primarily protects the individual from exploitation. This morality approach aimed at the individual, which according to our interpretation also includes property and trustworthiness in addition to fairness, is relevant across cultures: in the four cultural groups studied, morality is particularly important in protecting inter-individual cooperation from the danger of zero-sum games, which is expressed in corresponding deviance relevance margins (***Table 29***). Furthermore, the data shows that the latter statement applies both in predominantly individualistic (United States of America; Germany) and predominantly collectivist (Japan; Egypt) cultures. At this point, however, we want to go beyond collectivism-individualism. We designed the MaC-DRS fairness items primarily to capture the aspect of proportionality and less the fairness component of equity (Atari et al., 2022a). According to our data, proportional fairness now appears to play an important role in the regulation of cooperative ventures so that they do not end in zero sums. This finding is remarkable in light of the fact that our samples differ not only in terms of collectivism-individualism, but also in terms of power distance and their political systems (Barmeyer, 2010). Compared to the democratic systems in Germany, Japan and the US, the Egyptian government is currently more akin to an authoritarian regime (Armbruster, 2021). In addition, Egypt is characterized by a significantly higher power distance, as compared to the other groups under investigation.[101] So, although social participation and co-determination are structurally restricted, as is the case in authoritarian regimes (see: Armbruster, 2021, p. 49), and given a relatively high acceptance of power distance and thus also relative acceptance of inequality, fairness deviance relevance nevertheless plays an important role in Egypt.[102] The intuitive attribution of relevance for

---

[101] The country-specific scores for the cultural dimension *power distance* are as follows: power distance score GER = 35; power distance score JP = 54; power distance score US = 40; power distance score EG = 80, with higher values indicating more power distance. The values for the cultural dimension *power distance* can be accessed via the following website: https://www.hofstede-insights.com/country-comparison-tool

[102] Note: We would like to briefly highlight once again a circumstance that is also addressed in more detail in the **Appendix**. The EG-sample differs in its composition from the other cultural groups with regard to the *education* variable (ISCED), among other things. In fact, the Egyptian sample contains significantly more people with a higher level of education than the other samples. This becomes apparent when we look at the proportion of individuals with a bachelor's degree or higher in the four groups of our study (*adjusted sample*; N = 2,360): GER-sample = 29.13% (n = 666); JP-sample = 46.22% (n = 543); US-sample = 35.5% (n = 569); EG-sample = 80.58%

breaches of fairness thus appear to have a regulatory effect that matters across cultures, which lends credence to universalism assumptions regarding fairness (Brosnan & De Waal, 2003; Summerville & Enright, 2018; Curry et al., 2019a).

We have already touched on this above: the MaC-DRS **heroism** items we used are designed in such a way that they ask about heroism deviance relevance in general and independently of the social relation (family, in-group, stranger). It turns out that this general conception of heroism is relevant across cultures, as indicated by the margins in *Table 29*. We attribute this essentially to two components: First, the actual or potential suffering of others evokes our empathic capacity and, because we mirror the suffering of others within us, motivates us to alleviate this very suffering (Tangney et al., 2007; De Waal, 2008; Rusch, 2022). Second, a general tendency towards the relevance of heroism can also be seen from a cooperation perspective. Our heroism disposition enables people to stand by family, in-group and also strangers and to protect them, at least in part, in the face of danger. This keeps cooperation opportunities open and is also likely to be sanctioned with positive reputation. Since we have become increasingly dependent on cooperation partners for survival in the course of human development, there is also an inherent sense in risking something to protect those we depend on, because this also ultimately contributes to the reproduction of our own organism (Tomasello & Vaish, 2013; Henrich, 2020). Derived from the cooperation perspective and with the addition of our empathetic capacity, the cross-cultural intuition for deviance relevance of heroism becomes explainable for us.

> We would like to mention at this point, that the results of our study say nothing about what the relevance of heroism deviance would be if we were to consider *specific measures of heroism* in relation to different social relationships rather than a *general measure*. In other words, our results are not informative about how intuitive heroism deviance relevance configures across cultures if we would compare separate measures of heroism deviance in relation to family, in-group, and strangers. Future studies would therefore be desirable that develop separate instruments to measure the relevance of deviance towards the heroism domain for different social relations. Moreover,

---

(n = 582). Not only is it clear that the EG-sample is by no means representative. We also suspect that the composition of the EG-sample is associated with biases in the relevance margins of moral deviance when it comes to cross-cultural comparison between groups. In our view, this sample skewness toward higher education is related to overestimated/underestimated deviance relevance margins in the EG-sample. We investigated this suspicion by splitting the EG-sample into two samples according to educational level. So, we created an EG-sample containing only respondents with higher education (Bachelor's degree or higher; n = 469) and an EG-sample containing only respondents with lower education (education level below Bachelor's degree; n = 102). We then re-estimated our fairness OLS model for these two EG-samples and calculated the margins on this basis. The margins for these two samples are as follows: fairness deviance relevance higher education sample (n = 469) = 3.844; fairness deviance relevance lower education sample (n = 102) = 4.404. Although the lower education sample comprises in fact too few cases to make generalizable or reliable statements, the results suggest nevertheless that the fairness deviance relevance margin for Egypt is likely underestimated, as our EG-sample (n = 582) is biased towards higher education.

prospective studies that examine such social-relationally specific measures of heroism for similarities and differences across cultural entities would be worthwhile. Insights such as these would eventually allow us to better understand the extent to which our disposition to help others in need spans across different social relations.

Let us now turn our attention to the higher order approaches of **binding and individualizing morality**. Based on theory and findings from other studies, we have derived in the theory section a definition of the situation for the GER- and US-samples that can be described by individual-centered overall social orientation. On this basis, we hypothesized that the prevailing definition of the situation in these cultural contexts fosters individualizing morality. In particular, we focused on self-construal and cultural logics and formulated the following hypotheses:

> We expect cross cultural differences and *hypothesize that cultural entities that foster relatively more independent ways of selfhood also foster relatively more individualizing morality* (i.e., fairness, trustworthiness and property deviance relevance) (**CD sub-hypotheses 2**).

> We hypothesize that cultures of *dignity* are significantly higher in *individualizing morality* than cultures of cultures of honor and face (**CD sub-hypotheses 4**).

Contrasting to this, we derived an overall group-centered social orientation for the JP- and EG-samples from the literature, including the prevalence of a corresponding definition of the situation. Based on this, and drawing again on self-construal and cultural logics we hypothesized the following:

> We expect cross cultural differences and *hypothesize that cultural entities that foster relatively more interdependent ways of selfhood also foster relatively more binding morality* (i.e., family, in-group and deference deviance relevance) (**CD sub-hypotheses 1**).

> We hypothesize that cultures of *honor* and *face* are significantly higher in *binding morality* than cultures of dignity (**CD sub-hypotheses 3a**).

We will now take an initial, non-conclusive look at the hypotheses in the light of the evidence presented. In fact, we see the individualizing and binding hypotheses in the data partly contradicted and partly confirmed. We found evidence suggesting that a sociocultural context that fosters independence in self-construal does indeed also foster individualizing morality. This is demonstrated when inspecting the results from the GER- and US-sample and speaks initially in favor for the **CD sub-hypotheses 2**. Remarkably and to our surprise, the high deviance relevance of individualizing morality is also to a large extent present in the JP- and EG-sample! So, an overview of the MaC-DRS evidence would support the rejection of the

hypothesis that dignity contexts lead to a significantly higher deviance relevance in the domains of individualizing morality (**CD sub-hypothesis 4**). Furthermore, while binding morality is not dominant in the JP- and EG-samples, it is by and large more important in these cultural groups than in the WEIRD samples in our study. In other words, this finding suggests that our **CD sub-hypotheses 1** and **3a** can be partly confirmed: *Cultural entities that foster relatively more interdependent ways of selfhood*, in our study these are Japan and Egypt, *also foster relatively more binding morality* (**CD 1**). The same holds true in respect to the cultural logics of *face* and *honor* (**CD 3a**). On the basis of empirical evidence, these hypotheses appear to be fully valid with regard to in-group deviance relevance, and partially valid with regard to family and deference deviance relevance. We will not make definitive judgments about the hypotheses stated at this point, but will address them in more detail below when we discuss *which moral system guides collaboration within the cultural samples in our study*. Instead, in the context of higher-order moral approaches, we will now focus specifically on the finding that deviance towards *individualizing morality* is in part significantly more important across the cultures in our study than deviance towards binding moral domains. We believe that this finding is striking given the sociohistorical facts that Japan and Egypt, in contrast to Germany and the United States, are characterized by relatively more historical kinship intensity (Schulz et al., 2019; Curtin et al., 2020; Bahrami-Rad et al., 2022), disease prevalence (Murray & Schaller, 2010; Atari et al., 2022b), collectivism (Triandis, 2001; Barmeyer, 2010; Minkov & Kaasa, 2022; Żemojtel-Piotrowska & Piotrowski, 2023; Kitayama & Salvador, 2024),[103] and lower relational mobility (Thomson et al., 2018; Talhelm, 2022).

In order to interpret the initially surprising result of higher individualizing moral deviance relevance *across* cultures, we refer to the view that the cultural entities that we examine are *modern societies* and draw on trends in global sociocultural developments. Modern societies demand in particular the protection of the individual in cooperation, because in everyday life in modern forms of society a large part of potentially cooperative interaction takes place under market conditions outside the in-group and the family (Henrich, 2020). As a result of this everyday situational requirement, the regulation of *property* claims, the regulation of the reliability of cooperation partners (*trustworthiness*), and, the regulation of distributive (proportional) justice (*fairness*) is of great importance. Accordingly, our moral mind is calibrated to intuitively identify and evaluate violations of these moral domains as more relevant. So, we argue that the evidence we found, which is pointing to cross-culturally higher

---

[103] Compare in this regard also the respective scores on *individualism* across the four cultural entities of our study: https://www.hofstede-insights.com/country-comparison-tool

deviance relevance of *individualizing morality*, is a product of our moral mind that is adapted across a set of diverse cultures to prevailing sociocultural conditions of *modern* societies. This interpretative notion, which relates to social modernity, is paralleled and supported by findings on the rise of individualism throughout the world.

Research has identified that in more collectivist societies, the core tendency is that the group takes precedence over the individual, whereas in more individualistic societies, a key aspect is that the individual is at the center and takes precedence over the group (Triandis, 2001; Kitayama & Salvador, 2024). Apparently, the *individualizing* approach to morality mirrors the core aspect of the cultural dimension *individualism*, and the *binding* approach to morality essentially mirrors the key component of *collectivism* (Haidt, 2008). The same holds true, as we have hypothesized before, for the prevailing modes of selfhood — i.e., independent and interdependent self-construal — predominating to a large extent in either individualistic or collectivistic cultures (Markus & Kitayama, 1991; 1998; 2010; Dimaggio & Markus, 2010; but see also: Vignoles et al., 2016; Krys et al., 2022; Uskul et al., 2023).

Across the four cultures of our study, we found to a large extent a dominance of individualizing morality among moral domains, and in fact, recent studies support that in the last couple of decades individualism increases while collectivism is globally on decline (Hamamura, 2012; Santos et al., 2017; Kaasa & Minkov, 2020). Therefore, we propose to interpret the evidence for predominance of individualizing morality in Germany, Japan, the United States of America and Egypt in the wake of increasing individualism that seems to characterize modern societies.

An overview article by Cai and colleagues (2019) summarizes empirical findings that buttress the rise of individualism in: *social indicators* and *practices* (e.g. weakened family cohesion and increasing divorce rates, increase in first-person (I, my, me, mine) pronoun usage in books as in popular music), *values* and *attitudes* (e.g. increasing emphasis on choice, uniqueness and self-expression), *personality* (e.g. higher rates of narcissism, extraversion, and agentic self-evaluation), *religious* and *sexual attitudes* and *behavior* (e.g. more practice and tolerance of traditionally unacceptable sexual practices, and secular orientation), *childrearing practices* (e.g. fostered autonomy and competition), *emotions* (e.g. growing emphasis on internal factors as determinants of emotions), and also *cognition* (e.g. rising trends of private self-knowledge, and decontextualized cognition). Notably, the trend towards higher individualism around the globe can be found on the societal and individual level of analysis, in approaches collecting data inside the head (e.g. self-report) and outside the head (e.g. cultural products), and across societies already marked before by individualism or collectivism.

Core determinants of rising individualism are first and foremost socioeconomic developments indicated by growing national GDP, more white-collar than agricultural jobs, greater occupational prestige, higher income, and last but not least increasing educational attainment (Santos et al., 2017). We take up the latter point in more detail when we discuss the properties of the EG-sample in the context of moral deviance relevance further below. All of the just mentioned socioeconomic indicators signify components associated with the *modernization* of societies (Hamamura, 2012; Minkov et al., 2021).[104] Next to socioeconomic factors also socio-ecological factors affect the rise of individualism: decreasing pathogen prevalence, increasing disaster frequencies, and also harsher climatic conditions, as appearing in the course of climate change, foster individualism (Santos et al., 2017).

We interpret the dominance of individualizing over binding morality in intuitive deviance relevance attribution by referring to the antecedents of rising individualism in modern societies. Accordingly, although we rely on cross-sectional data, we propose to view the MaC-DRS results in ***Table 29*** as suggesting that modern societies — characterized by, among other things, high market integration and everyday interactions beyond kinship and in-groups — are associated with social conditions that require particular human self-regulation to protect the individual from the potential of failed cooperation (zero-sum games). The human moral mind thus adapts to the prevailing societal conditions of modernity by calibrating the relevance of moral transgressions so that they are particularly pronounced in the realm of individualizing moral domains. Accordingly, we suggest to understand the findings of predominance in deviance relevance towards property, trustworthiness and fairness across the entities under investigation as supporting evidence for the interpretation of a moral mind that is adapted to the societal conditions of *modernity*. Moreover, from our point of view and especially by referring to Japan and Egypt, the MaC-DRS findings contribute to studies indicating a global rise in individualism by showing that our moral mind appears to be affected by this trend in human social development.

Kühnen and Kitayama (2024) note: "cultural differences are not fixed and stable, but dynamically flexible in nature" (p. 7). In this line, we foresee a further increase in cross-cultural individualizing morality deviance relevance. However, we base this notion on the assumptions that the future will be an even more globalized and interconnected world that is (hopefully) characterized by further socioeconomic development around the globe and, at best, not by

---

[104] These factors of socioeconomic development apply also to emerging economies. A case in point is the absorption of the old Bedouin culture by modern market economy changes in Arab countries (Cole, 2003). See for instance in the case of our study also the GDP development in Egypt from 1989 with prognosis to 2029: https://www.statista.com/statistics/377349/gross-domestic-product-gdp-in-egypt/.

further catastrophes such as the corona pandemic. Although sociocultural developments are by no means to be seen as linear processes (Cai et al., 2019), changes may happen at different speed (Kaasa & Minkov, 2020), and backlashes may occur, while cultural heritage may persist unaffected alongside sociocultural changes (Esser, 2002a; 2010; Greshoff, 2008; Hamamura, 2012), our interpretation of the MaC-DRS data in the wake of increasing individualism nevertheless makes us to expect an increase in individualizing morality in modern societies around the globe in the years to come.

> Based on our results and interpretations, future studies may aim to investigate whether the relevance of deviance towards domains of binding and individualizing morality has changed over time, and whether such a change may even reveal a tendency towards partial moral alignment across cultures. Initial indications point in the direction of an increase in individualizing deviance relevance compared to binding deviance relevance (see: Enke, 2019) and bring into play, among other things, a changed cooperation situation in the course of the industrial revolution as an explanatory factor. In order to better understand the demands that societies pose on our psyche, we believe that it is important to investigate whether modern societies are undergoing a trend towards the development of a greater relevance of individualizing morality, and whether more traditional, small-scale societies may (still) place greater value on binding morality. To support our interpretation of the cross-cultural dominance of individualizing deviance relevance in modern societies, future research with small scale societies that still maintain a more traditional way of life would thus be very welcome. In this context, we predict that societies with a more traditional way of life will also endorse higher binding moral relevance as compared to modern societies. Furthermore, it will also be of interest to consider the development of the internet and the associated (global) social interconnectedness in future studies on morality. In this context, it could be asked whether the advent of the internet represents a catalyst for increased relevance of individualizing morality. The background to the research questions and areas described here is to gain a better overall understanding of the demands that modern societies place on the human moral psyche. With further research, we could try to shed more light on how our moral mind adapts to modern societies to enable cooperation and to maintain social order under current societal conditions.

Taken together, our cross-cultural findings initially show that none of the 8 moral domains is intuitively completely irrelevant in any of the groups studied. Nonetheless, violations of certain moral domains are more relevant across cultures: In particular, deviant behavior towards individualizing morality, i.e., especially property and trustworthiness, but to a large extent also fairness, is evaluated as relevant. We argue that our moral psyche adjusts the relevance of morally deviant behavior to the requirements of (modern) societies in order to enable non-zero-sum (i.e., mutually benefitting) interaction between people under current socio-structural conditions and requirements. This argument is based on the assumption that our first draft of the moral mind (see: Haidt & Joseph, 2007) is universally rooted in our species and yet undergoes cultural editing to remain effective by adapting to the contemporary nature of the

social world of the phenotype. The four cultural groups we study are marked by societal indications of modernity. In such societies, as we interpret our data, there appear to be social requirements for cooperative endeavors that require the human moral mind to pay particular attention to recognize and regulate individual selfishness in the domains of property, trustworthiness, and fairness. From this glimpse on cross-cultural commonalities in moral deviance relevance we will in the following take a closer look at the specific peculiarities of the moral system prevalent in Germany, Japan, the United States of America, and Egypt. In particular, we will examine which moral system guides cooperation *within* these cultural entities. In the course of these investigations, we will also finally clarify the status of our CD hypotheses.

4.5.4. Which Moral System Guides Cooperation in Different Cultures? MaC-DRS Findings

After the pairwise comparisons and the cross-cultural similarities, we now come to the analysis of the moral systems *within* the four cultural entities of our study. Our main aim is to pursue the research question of which moral system primarily guides cooperation in the cultural groups studied. In this context, as explained elsewhere, we hypothesized that the sociocultural constitution of the GER- and US-sample fosters moral systems that pay particular emphasis to individualizing morality (**CD sub-hypotheses 2 and 4**). For the JP- and EG-samples, we furthermore hypothesized sociocultural constitutions which foster moral systems paying particular emphasis to binding morality **(CD sub-hypotheses 1 and 3a; *see also Figure 5; Chapter 2*)**. It is also worth noting that our hypotheses take into account from the outset that cultures are diverse. Therefore, despite our binding and individualizing hypotheses, we do not expect a homogeneous picture of the relevance of moral deviance neither for the GER- and US-samples nor for the JP- and EG-samples. Our hypotheses are therefore meant primarily as informed heuristics to guide the analysis.

Let's take a look at **Table 31** below, in which we have sorted the different moral domains measured by MaC-DRS in descending order of *within-sample deviance relevance*. **Table 31** displays again the margins (AME) obtained for each cultural group after fitting the respective OLS covariate models. In addition to the deviance relevance order, we have colored and highlighted the three most important moral domains per sample in *green*. Furthermore, we left the two moral domains with medium deviance relevance in *black*, and colored the three moral domains that are least relevant within each sample in *red*. We now use the ranked deviance

relevance margins within the samples to investigate our research question and to finally clarify the **CD sub-hypotheses 1** to **4** in the context of the MaC-DRS results. We would like to emphasize that this ranking is yet only of heuristic value, as some of the moral domains differ within the samples only in slight and neglectable deviance relevance attribution. Even if only of heuristic value, the within-sample ranking of moral domain specific deviance relevance gives us a simple and easily understandable impression of how the moral system is organized in the respective cultural entity. Later, we will supplement the MaC-DRS findings discussed here with findings from the moral dilemma scenarios and findings from the factorial survey to give the examination of our research question even more data-supported depth. But for now, we turn to the final examination of the MaC-DRS findings, by taking a look at *Table 31*.

Table 31: Adjusted sample ranking (highest to lowest): order of moral deviance relevance (margins) within cultural groups

Germany n = 666	Japan n = 543	USA n = 569	Egypt n = 582
4.870 † Fairness	5.181 Property	4.336 Property	4.477 Property
4.772 Property	4.993 Trustworthiness	4.291 Fairness	4.407 Family
4.601 Trustworthiness	4.570 Fairness	4.239 Family	4.239 Trustworthiness
4.115 Heroism	4.563 Heroism	4.107 Trustworthiness	4.062 Heroism
3.961 Family	4.495 Reciprocity	4.072 Heroism	4.058 In-Group
3.906 Reciprocity	4.344 Family	3.709 Reciprocity	4.015 Fairness
3.442 In-Group	4.043 Deference	3.603 Deference/ In-Group	3.960 Deference
3.374 Deference	4.042 In-Group	3.603 In-Group/ Deference	3.936 Reciprocity

† On display are the within-sample ranked margins of moral deviance relevance for each sample after fitting the OLS-covariate model. The moral deviance relevance scores (margins) shown are equivalent to the scores in *Table 29*. The margins are ordered by magnitude (descending order: Green > Black > Red).

As the results in *Table 31* speak for themselves and have already been addressed in the sections above in a different form, we will dispense with a breakdown and go straight to the discussion of the results of the individual samples. We will thereafter bring the individual discussions of the findings together in a comprehensive discussion of all four samples.

4.5.4.1. Discussion: Which moral system guides cooperation in Germany? MaC-DRS findings

First, to the **GER-sample**. Looking at ***Table 31***, we can see quickly that the **CD sub-hypotheses 2** for this cultural group is correct.[105] The domains of individualizing morality, i.e., fairness, property and trustworthiness, are obviously the most relevant moral domains in the German within-sample ranking. It is interesting to note at this point that the fairness domain in particular has a high priority in the GER-sample. This relatively pronounced fairness deviance relevance score may originate in part from a philosophical history and tradition of Enlightenment shaped in the 18th century that manifested inter alia in the claim of inherent dignity of the person (see (e.g.): Kant, 1788/2011; Uskul et al., 2019). Dignity as a concept became also important in the German constitution where especially the first article emphasizes the inherent value of the person. The first article of the German constitution states: "Human dignity is inviolable. It is the duty of all state authority to respect and protect it".[106] Furthermore, according to Leung and Cohen (2011), interactions and exchanges in dignity cultures are characterized by "[c]ontract among equals" (p. 3), which highlights the relevance of fairness in cultural logics of dignity.

To interpret the relatively pronounced fairness deviance margin in Germany we can also draw on the findings of the OLS covariate model. Results of this model revealed a positive and significant ($p < 0.001$) interaction effect between *culture* (GER-sample) and *NARS*, contributing to higher margins. So, next to our substantive interpretation, the mentioned interaction effect certainly also impacted on the fairness deviance relevance margin displayed in ***Table 31***. Taken together, along with the high degree of individualistic orientation (Cai et al., 2019), and the prevalence of the independent self-construal (Kitayama et al., 2009) as the dominant way of selfhood, the relatively pronounced finding of fairness deviance relevance in Germany can be explained from our point of view by drawing on the prevalent dignity context, [107] the philosophical tradition of enlightenment, and also the interaction effect of the GER-sample with the net acquiescence response style measure in our model.

---

[105] Note: We are well aware that reality is multi-causal. Our following interpretations and the explanation of the effects found in the realm of moral deviance relevance certainly only cover a part of what has caused the findings themselves. We therefore do not claim to explain the findings in their entirety, but merely approach reality in part with our theorizing and interpretations.

[106] see: https://www.bundesregierung.de/breg-de/themen/75-jahre-grundgesetz/artikel-1-gg-2267756

[107] The differences between the GER-sample and the US-sample with regard to the deviance relevance of fairness may be, next to historical differences, also partly due to the fact that the USA is characterized not only by dignity but also by honor (Uskul et al., 2019), has stronger religious ties, and exbibits in some parts/regions also historical kinship intensity (Schulz et al., 2019).

Apart from individualizing domains, the deviance towards heroism also seems to be intuitively relevant in this sample, as in all others samples too, although the relevance margin is not as high as for fairness, trustworthiness and property. Furthermore, the binding domains family, in-group and deference are found in the bottom half of the deviance relevance ranking in the GER-sample. It should be emphasized that especially transgressions of the in-group domain and the deference domain are intuitively evaluated as significantly lower in relevance compared to all other moral domains tested. Breaches in these domains are clearly not as relevant in Germany as violations of individualizing moral domains.[108] Thus, the MaC-DRS findings for the German sample show above all that the moral system is primarily characterized by individualizing morality. We derived from other studies (see: **Chapter 1** and **2**), that the German sociocultural context can be described as low in historical kinship intensity, individualistic in terms of cultural orientation, independent in prevailing pursuit of self-construal, wheat farming and herding oriented in terms of historical subsistence style, low in pathogen prevalence, and marked by the logic of dignity. Given these sociocultural conditions and taking into account the MaC-DRS findings, we can therefore draw the following preliminary conclusion: ***Individualizing morality primarily guides cooperation and dominates the moral system in Germany***.[109] Accordingly, the individual-centered overall social orientation predicted for the GER-sample can be confirmed in the context of moral deviance relevance. For the German sample, we thus find evidence that affirms **CD sub-hypotheses 2**.[110] However, since we also found high individualizing morality deviance relevance values in the cultural entities characterized by face logic (Japan) and honor logic (Egypt), we cannot confirm **CD sub-hypothesis 4** for the investigation focused on Germany.

---

[108] For the GER-sample, we found an interaction effect with the variable level of religiosity in the in-group deviance relevance model. Considering oneself to be religious reduces the relevance of violations to the in-group domain in Germany. The findings on the within-sample ranking of moral margins in Germany should also be interpreted in this context.

[109] The raw mean values from data collection 2 (N = 2,356; Germany-wide non-student sample) support the conclusion that intuitively individualizing moral domains prevails over binding domains in Germany. See the **Appendix** for respective evidence.

[110] It should be noted that we discuss some of our findings in this section as if they would apply to the (entire) cultural groups and societies that we examine. This is done for the purpose of putting the respective socio-cultural world we are investigating at the center and not the technical term "sample". However, as our remarks in the course of this work should have indicated, these statements are best be treated with appropriate caution. This is because, as emphasized here and elsewhere, our samples are not representative and, in some cases, have significant limitations. We therefore ask our readers to consider our statements against the background of our data collection and the characteristics of the four samples and to classify them accordingly.

4.5.4.2. Discussion: Which Moral System Guides Cooperation in Japan? MaC-DRS Findings

Second, to the **JP-sample**. We find a similar deviance relevance ranking in the moral domains highlighted in green in the Japanese sample as in the GER-sample (see: ***Table 31***). However, in our opinion it would be short-sighted to interpret the results of the JP-sample with an exclusive prioritization of individualizing morality. Rather, it is striking to note that all other moral domains display margins of ≥ 4 and tend towards deviance relevance within the Japanese sample. Therefore, due to the relatively high deviance-relevance scores across all moral domains, we suggest to consider the dominant moral system in Japan as a comprehensive mixed moral system. In this Japanese comprehensive mixed moral system, besides emphasized individualizing morality, heroism, reciprocity, and also binding morality guides cooperation. We have two interpretations regarding these findings that draw mainly on cultural Tightness/Looseness Theory (Roos et al., 2015; Gelfand et al., 2017), and on culturally divergent ways of cognition (Nisbett et al., 2001; Kühnen & Hannover, 2003).

First of all, the data shows an overall tendency for higher deviance relevance scores in Japan across all moral domains. This applies also when we compare the JP-sample with the other cultural groups studied (see: ***Table 29***). We explain this by the fact that Japan is a relatively *tight culture* (Gelfand et al., 2011). Tightness/Looseness Theory is an approach explaining strength and intensity of norm adherence and enforcement within and across cultures. The theoretical approach is recurring on more distal ecological and historical factors as well as on more recent socio-political developments and institutions that blend together in affecting contemporary social and psychological processes related to norms. In this regard and referring to empirical findings, Gelfand and colleagues (2017) state: "tightness correlated with an extensive array of historical and ecological threats (…) tight societies had greater historical prevalence of natural disasters, food scarcity, population density, and territorial threats compared to loose societies" (p. 802). Supporting evidence is also provided by Talhelm (2022), whose findings show that societies that have historically relied on paddy rice cultivation, such as Japan, have developed strong norms that were necessary to effectively organize the cooperation required for this form of subsistence. Thus, drawing on the Tightness/Looseness Theory and history of subsistence, our findings on moral deviance relevance presented in ***Table 31*** suggest that the causes of normative tightness in Japan affected not only norms but also morality. Accordingly, the moral system in Japan appears to be influenced by the distal ecological and historical factors that produce a contemporary world of institutions and societal regulations that can be considered as relatively tight.

Furthermore, Japanese society can be seen as a society caught between the poles of tradition and modernity. However, tradition and modernity do not have to be mutually exclusive in Japan. Unlike in most Western cultural entities, innovation does not per se go hand in hand with the rejection of tradition in Japan. In their comments on cultural practices in Japan, Traphagan and Thompson (2006), for example, state the following:

> „But radical discarding of values and social structures of the past has not been a characteristic of the postwar era [in Japan]; instead, augmentation and improvisation on older themes, the invention of new approaches to social organization, and new ideas about what should be valued by members of society have combined to form the diversity of the current modern moment" (pp. 5-6).

This diversity seems also to be reflected in the configuration of Japanese moral system and relates, from our interpretation, to the *interdependent* construal of selfhood and the holistic system of thought prevailing in Japanese culture. The *independent* self strives for self-consistency in behavior and more generally for cognitive consistency and the avoidance of dissonance. However, these aspirations are not found to the same extent in people who understand their self as essentially connected and *interdependent* with others. Rather, situational variability of the self, which includes possible inconsistencies of the self's behavior and attitudes in different situations, is a sign of a mature, interdependent self that is able to adapt to the demands and needs of different social situations and actors within it (Markus & Kitayama, 1991; 2010; Wong & Tsai, 2007; Markus & Schwartz, 2010; Cross et al., 2011).

Nisbett and colleagues (2001) examined culturally diverging systems of thought (e.g. reasoning, perception and beliefs) in depth, and differentiate between *holistic* cognitive processes associated with the interdependent self, and *analytic* cognitive processes associated with the independent way of selfhood. These authors take recourse on ancient Greek and Chinese societies to explain contemporary cross-cultural differences in cognitive processes. Referring to the scientific and philosophical traditions (e.g. considerations about nature and causality, epistemologies, degrees of certainty, styles of discussion) and the different worldviews and habits associated with these traditions, Nisbett et al., (2001) note: "Greek civilization gave rise to European civilization and post-Columbian American civilization, and Chinese civilization gave rise to the civilizations of East Asia, including Japan and Korea, and also greatly influenced Southeast Asia (p. 292). In addition, these authors posit a link between social organization and cognitive processes. By doing so, they assume that the former influences basic cognitive processes through different emphases (inside, outside) in the

attention towards the environment and distinct normatively regulated patterns of communication style. Based on this idea, they present diverse evidence of cross-culturally divergent cognitive processes, which they trace back to ancient Greek and Chinese civilization and respective social organization. Nisbet and colleagues (2001) basic claim is that ancient Greek greatly influenced Westerners in the formation of an *analytical system of thought*, while ancient China had a great influence on the formation of a *holistic system of thought* prevalent in Eastern societies. Empirical findings demonstrate that analytic thought is (e.g.) characterized by detachment of object and context, the use of formal logic and rule-based inference about categories and their properties, reliance on abstract knowledge, and the *avoidance of contradictions*. In contrast, findings show that holistic thought is distinguished (e.g.) by an orientation to the context and relationships in the context, change and dialecticism, experience-based knowledge, and the *recognition of contradictions* (Nisbett et al., 2001; Myers et al., 2010, pp. 55-94).

Hence, instead of the abstract and analytical dissection of phenomena, we find in Japan a culturally predominant system of thought that allows the integration of tradition and modernity and the promotion of behaviors being socially accepted in certain situations but not in others. These cognitive integration efforts can be accompanied by contradictions without, however, being avoided. Conceivably it is precisely this holistic system of thought that also explains why a comprehensive mixed moral system prevails in Japan rather than *either* a binding *or* an individualizing one. [111] Thus, our findings suggest that the emphasis on individualizing morality *and* binding morality in Japan need not be seen as mutually exclusive, although this may seem contradictory to some (analytical thinkers).

This reasoning is in line with evidence found by Hamamura (2012). Although Japan is part of the global trend towards more individualism, this development does not mean that collectivist values and practices are disappearing in Japan. Collectivism and individualism are multidimensional concepts. Hamamura's (2012) insights suggest that while some individualistic aspects are indeed increasing in Japan, some collectivist ones also remain relevant: "findings (...) show that the importance of collectivistic living has continued in Japanese society even during a period of significant economic development" (p. 17) which has broad about increasing individualism in the first place. Thus, not only the influences of modern life, but also cultural

---

[111] Consistent with our explanation and the reasoning of Nisbett et al., (2001), we found a distinct middle category response style in the JP-sample, as detailed further above, which strongly points to the prevalence of a holistic system of thought among respondents in this sample. See the **Appendix** for related discussions.

heritage (Hamamura, 2012) and independent cultural path dependencies influence the shaping of culture specific moral systems.

In addition, the results of our OLS models revealed a significant and negative interaction effect between *culture* (JP-sample) and the *midpoint response style measure* (*MRS*). We found this effect in the models dealing with trustworthiness (p < 0.001) and reciprocity (p = 0.023). The same models also showed a positive and significant main effect of *MRS*. However, this effect is reversed for the JP-sample due to culture-specific moderation. We interpret this to suggest that the communication style of modesty and restraint in Japan prevented even higher trustworthiness and reciprocity margins for this sample. As the corresponding domains nevertheless prove to be relevant — the trustworthiness and reciprocity margins in the JP-sample are the highest in the comparison of groups examined —, there is no substantial change for us in the data interpretation that needs to be considered in more depth.

Overall, we found evidence suggesting that the **CD sub-hypotheses 1** and **3a** are largely confirmed for the Japanese sample. One exception in the context of the hypotheses, however, is the comparison with the US-sample, primarily with regard to family deviance relevance. We will address the US American moral system next, but for the moment we will stick to Japan. Here, then, it can be noted that a social context characterized inter alia by collectivism, interdependent self-construal and the logic of face seems indeed to foster relative deviance relevance towards family, in-group and deference. However, although we were correct with our hypotheses, it would (also) clearly be too short-sighted to distinguish the JP-sample primarily by referring to binding morality. Instead, a relatively comprehensive mixed moral system, broad in moral domain coverage and pronounced in deviance relevance, seems to prevail in Japanese contemporary society. As elaborated, we explain the finding of such a mixed system by integrating insights from the theory of Gelfand and colleagues (2011; 2017; Roos et al., 2015), findings within the framework of the self-construal approach (Vignoles et al., 2016), distinct systems of thought (Nisbett et al., 2001), and also by reference to the rise of individualism (Hamamura, 2012; Cai et al., 2019; Kaasa & Minkov, 2020; Minkov et al., 2021). Consequently, in a preliminary conclusion we refer to Japan as a ***tight moral culture pursuing a comprehensive mixed moral system***. Moreover, in regard to morality it seems to be an individual- and group-centered social orientation that we must assume for Japanese contemporary culture, and likely this social orientation varies across situational contexts.

4.5.4.3. Discussion: Which Moral System Guides Cooperation in United States of America? MaC-DRS Findings

Third, when we turn to the **US-sample**, the within-sample ranking shows that the predominant moral system is primarily characterized by deviance relevance of individualizing domains. In contrast to Germany, however, we find in the US case a different deviance relevance order and besides heroism, also the binding dimension family, which is distinguished by higher deviance relevance margins.[112] In the broadest sense, individualizing morality guides cooperation in the US-sample, but not exclusively — according to our MaC-DRS results, kinship altruism also has a non-negligible importance in the United States of America. Our **CD sub-hypothesis 2** is for the US-sample therefore mainly confirmed, but with an extension. Although individualizing morality primarily guides cooperation in this cultural entity, moments of binding morality (the family domain) seem also to play an important role.

As we have been correct on the importance of individualizing morality in the United States, we will focus in our interpretation on the deviance relevance of the family domain. Culture is not only found in the minds, practices and artifacts of people, but also in the institutions they create and in the policies that form the essential content of institutions (Henrich, 2020). One reasonable explanation for why we find deviance relevance in regard to the family domain in the USA but not in the other WEIRD-sample of Germany may lie in the different social security systems of these countries. Although the German and US models of social security are similar in parts, there are striking differences in the legal entitlement to social security between these countries. In Germany, there is a constitutional right to minimum social security provision, which is not found in the US social system (Seeleib-Kaiser, 2014), and there is history in this institutional circumstance.

Melinda Cooper (2020) traces the neoliberalist, partly neo-conservative policy development in the context of sometimes morally charged discourses and draft laws in the United States, which have led to the fact that the family in particular has always functioned as the main social safety net in the USA. The origins of the policy development that the family has the basic duty of care (and welfare) can be traced back to the year 1601 and the Elizabethan poor laws. Cooper (2020) describes how these views, which were developed in England and

---

[112] Hamamura (2012) found a decline in trust in the USA. This trend may also be reflected in our data, as trustworthiness ranks only fourth among moral domains in the US-sample. From our theoretical perspective, our moral mind adapts to the prevailing conditions of its sociocultural ecology and calibrates the initial grant of reliability based on experiential knowledge acquired over the course of a lifetime. In this sense, investigating the causes of potentially declining trustworthiness in the United States could prove to be an exciting area for future study. Longitudinal studies would be particularly valuable in this context.

contain the "principle of familial responsibility" (p. 99) at their core, found their way into state system in the USA early on:

> "The early American colonies imported the poor laws virtually verbatim and they were later incorporated into state legal systems during the early American Republic. These laws were continually reinvigorated and embellished to adapt to what we might call periodic episodes of sexual revolution" (p. 99).

With regard to these episodes of sexual revolution, it is fundamentally described that socio-historical processes are not linear. Efforts were made in the United States, albeit only for a short period of time from a structural point of view, to question the family as a fundamental institution of social security, and the formal marriage and economic dependence of women on their husbands that had long been associated with it. In fact, however, the principle of familial responsibility remained institutionalized in the US almost throughout, at least in part, and was largely reevoked by the US welfare reform of 1996 (Cooper, 2020).

In terms of the US social security system, the policies are likely embedded in narratives and imperatives of autonomy, freedom, and choice (Markus & Schwartz, 2010), but they entail obligations as consequences. Kinship altruism includes „obligation to kin (…) [and] duty of parental care" (Curry, 2016, p. 38). In the absence of a constitutional entitlement to a minimum income (e.g.), there is a social requirement for other institutions to step in to compensate for this deficiency. Thus, where in the event of loss of work or similar state-organized social security is not or only limited available, other security institutions such as the family (must) remain crucially relevant, as is also stipulated by the US welfare system (Cooper, 2020). In addition, recent research shows that the normative obligation of family support within a society strongly contributes to curbing the demand for public welfare services (Arévalo-Iglesias, 2024), which likely structurally reinforces social path dependencies once they have been established. An explanation for the intuitive relevance of deviance vis-à-vis the family domain could therefore lie in the historical welfare policies of the USA, which for their part go hand in hand with pronounced imperatives of independence and autonomy, and still exert an ongoing effect today also visible in the calibration of the human moral mind. Our MaC-DRS results for the US (*Table 31*) receive complementary support by findings on individualism-collectivism of Hamamura (2012), who states: The relationship with the family in the US has "been largely stable or (…) even strengthened over time, as indicated by (…) an increase in unconditional love and respect for parents" (p. 13).

Indeed, we were mainly correct with CD hypothesis 2 about the deviance relevance of fairness, trustworthiness and property in the US. But our findings also suggest that it would be too narrow to distinguish the US American moral system solely by reference to individualizing morality. Overall, we can trace a socio-historical context of social policy in the US that largely contributes to the strengthening of family dependency. Some regions in the US exhibit high historical kinship intensity (Schulz et al., 2019), and our US-sample also shows high levels of religiosity among respondents.[113] Furthermore, the US sociocultural context is essentially characterized by individualistic orientation, independence in selfhood and mainly by prevailing dignity. In our view, these factors together explain why the family domain, along with individualizing morality, is found high in deviance relevance in the US. In addition, the United States of America are also a rather loose culture (Gelfand et al., 2011), which may contribute to the overall picture of the more reserved moral deviance relevance margins in this sample compared to the other samples (see: *Table 29*). Based on the MaC-DRS findings, we can accordingly draw the preliminary conclusion to find an extended individualizing morality as prevailing in the United States of America. With regard to morality, we therefore interpret the data to suggest that there is an individual- and family-centered overall social orientation in the US-sample, which is reflected in an ***extended individualizing moral system*** that guides cooperation in contemporary US-American society. Finally, it should be noted that we find no confirmation of CD sub-hypothesis 4 in the United States or in Germany. We have explained above that we are investigating modern societies. We have placed the cross-cultural finding of a prevailing individualizing morality in this context as well as in the increasing individualism that accompanies it. Thus, it is not primarily a context of dignity that determines individualizing morality. We therefore close the **CD sub-hypothesis 4** and do not consider it to be confirmed in light of the MaC-DRS evidence.

---

[113] If we compare the US- and GER-sample of this study, we find that both samples are predominantly Christian, which also seems fitting from a historical perspective. However, when the variable *level of religiosity* is compared between the two samples (two-sided t-test; equal variance; obs. Ger-sample = 666, mean = 2.864; obs. US-sample = 569, mean = 4.328; df = 1233), a highly significant difference is found (t = -14.1385; Pr (|T| > |t|) = 0.000), which identifies the US-sample as clearly more religious. We mention this fact because Christian religions explicitly emphasize and define the family domain in their moral codes, such as to be found in the 10 Commandments. We therefore reason, that the level of religiosity and accompanying (unmeasured) covariates, may also impact on the family deviance finding in the US-sample. This statement, however, remains in the realm of a hypothesis.

4.5.4.4. Discussion: Which Moral System Guides Cooperation in Egypt? MaC-DRS Findings

Finally, we turn to the **Egyptian (EG-) sample**. If we take the 7-point response format of MaC-DRS as a basis, then the margins in the Egyptian sample reveal for all but deference and reciprocity values of ≥ 4.0, and so tendencies towards intuitive relevance. However, among the three domains most relevant we find property on top, followed by the family domain and the trustworthiness domain.

We believe that this particular pattern of deviance relevance in the EG-sample can be explained in part by two backgrounds: substantively by a) ancient socio-historical sources that have left traces of influence in Arab culture, and accompanying by b) the composition of the sample. First, let us take a look at the substantive interpretation. Specifically, we refer in this regard to historical kinship intensity (Schulz et al., 2019), the culture of the partly nomadic desert tribes, which are summarized under the term Bedouin (Cole, 2003), and the logic of honor as well as the social context of the emergence of this logic (Uskul et al., 2019).

Let's first take a quick look at kin-based institutions and kinship intensity. Based on the Ethnographic Atlas (a database containing anthropological data on several thousand ethno-linguistic groups), researchers have created a so-called kinship intensity index (KII) (Henrich, 2020). This index comprises data on historical family structure and decent systems across countries, expressed in a single number. Influencing variables within the KII represent data on the following socio-historical aspects of family: stronger marriage norms between relatives/cousins vs. weakened norms of the same kind; extended family vs. nuclear family; living together or not living together with the family group of one of the spouses; unilineal decent vs. bilateral decent; clan structure vs. no-clan structure (Enke, 2019; Henrich, 2020). Egypt is found to score high on the KII.

For centuries, the family policies of the Roman Catholic Church in particular have contributed to a lower kinship intensity in Europe and the European-influenced USA than in other societies around the world. Furthermore, historically formed kin-based institutions and associated norms in turn have a demonstrable influence on people's current psychology:

> "By constructing denser, tighter, and more interdependent social networks, (…) kin-based institutions intensified in-group loyalty, conformity, obedience to elders, and solidarity. For example, instead of favoring marriages to distant kin, cultural evolution often favored some form of cousin marriage, which tightened existing bonds among families" (Schulz et al., 2019, p. 1).

Different to regions under historical influence of the Roman Catholic Church, cousin marriage was not prevented by certain family policy measures on the Arabian Peninsula and African territory next by. On the contrary, extended kinship ties and cousin marriage became a widespread norm, especially among Bedouin tribes. Explanations for the higher rates of cousin marriages in Arabia and Arab-influenced Africa take recourse to an adaptive advantage that cousin marriages can bring. If dependence on (camel) milk consumption is high, as among the Bedouin, then marriage among those who can digest milk is advantageous to maintain this ability. Indeed, Reilly (2013) argues: for "Bedouins, who depended heavily on the lactase persistence (LP) allele for their subsistence, any marriage practice that maximized the frequency of this allele in the lineage would be favored over alternative marriage strategies" (p. 375). Above that, as Bedouins enjoyed a high status and prestige in traditional Arab societies (Reilly, 2013, p. 384), their marriage practices and norms are widespread in societies that have recognized the social status of Bedouins.[114] Hence, by recourse on historical kinship intensity, which is found to be high in Egypt (Schulz et al., 2019; Enke, 2019; Bahrami-Rad et al., 2022), we explain why especially among the EG-sample we find relatively high margins of family deviance relevance. Note: among cultural groups compared, the family deviance relevance margin of the EG-sample takes on the highest value.

Furthermore, honor logic is prevalent in Egyptian culture (Uskul et al., 2023). Regarding the ancient social context that gave rise to the logic of honor, Leung and Cohen (2011) state: "Cultures of honor tend to originate in "lawless" environments, where a weak (or nonexistent) state is unable to enforce contracts, protect individuals from predation, or punish the guilty" (p. 3). Under these conditions, positive and negative reciprocity, which blend into reputation, become important for the organization and regulation of social interactions. Following this reasoning it is people's reputation that indicates who is considered trustworthy and who is not (Leung & Cohen, 2011), with whom cooperative ventures can be entered into and with whom not.

In this context, the high trustworthiness deviance relevance found in the EG-sample is interesting and can be explained by drawing on the honor logic and nomadic lifestyle of the early Bedouin (Cole, 2003). As predominantly non-sedentary nomadic people, the early Bedouins not only practiced pastoralism, but also cross-regional trade (San Martin et al., 2018), which formed another important material basis for survival in the deprived life in the desert.

---

[114] In the context of the evolutionary mechanism of intergroup competition, Henrich (2020) lists various processes — i.e., war and raiding, differential migration, prestige-biased group transmission, differential group survival without conflict, and differential reproduction — which are also important with regard to the explanation of the adoption and spread of social norms (pp. 96-99).

Assuming that trade was an important basis for the supply with scarce resources and that social regulatory mechanisms of morality such as reputation and gossip were present (Henrich & Muthukrishna, 2021), it can be further argued that, from a historical perspective, a socio-cultural context existed in Egypt for a long time that required people to prove themselves reliable. Baumeister (2022) states: "humans talk and gossip, so a few selfish actions can turn off even people who were not directly affected. They know your reputation" (p. 108). Reputational damage could reduce (e.g.) future opportunities for the exchange of goods, i.e., opportunities for cooperation in the broadest sense, and thus pose a serious threat to groups partly relying on trade to gain access on scares resources. From this reasoning and under considerations of sociocultural path dependencies the relevance of trustworthiness deviance found within the Egyptian sample can be derived. In line with our findings of trustworthiness deviance relevance in the Egyptian sample Uskul et al., (2019) state: "The scoundrel, liar, or thief cannot be considered honorable. Instead, the honorable person is trustworthy, hospitable, honest, and true to his or her word" (p. 799). And indeed, also other studies found empirical support for the centrality of trustworthiness in honor codes and thus mirror our empirical findings (Uskul et al., 2019).

To an even greater extent than for family and trustworthiness, we found relevance of deviance towards property in the Egyptian sample of our study. This particular property deviance relevance can be interpreted in line with our previous argument. Remember that the honor logic prevalent in Egypt culture partly originates from historical conditions marked by the absence of contractual regulations of human interaction (i.e., law). Now drawing on nomadic Bedouin culture, San Martin and colleagues (2018) write: "much (…) property was portable, able to be stolen and thus necessary to be vigilantly protected" (p. 831). Thus, considering the historical context from which the logic of honor is derived, the importance of livestock as capital and the commodity trade as part of the Bedouin material base. If, in addition to these factors, the high social status of the Bedouin (in traditional Arab societies) and their influence on the dissemination of norms are taken into account, the high relevance for property deviance in Egypt becomes understandable. Besides, we found in the Egyptian within-sample ranking of moral domains, heroism, in-group and fairness as the next most relevant domains after property, family, and trustworthiness.

In addition to the substantive interpretation, the statistical effects that we found for the EG-sample must also be taken into account. Interestingly, in our OLS models for family and in-group deviance relevance, we found a negative interaction effect between *culture* (EG-sample) and *level of religiosity* (see: **Table 28**). Moreover, the latter variable is strongly

pronounced in the EG-sample. We have already pointed out elsewhere that findings suggest that the cultural emergence of religion is in part linked to more extensive cooperation with, among others, distant co-religionists (Lang et al., 2019). Although family and in-group deviance in particular are relevant in the EG-sample, the pronounced willingness to believe in a supernatural (monitoring and punishing) deity seems to promote the relevance of the other domains, which shifts the relationships and attenuates the level of deviance relevance for family and in-group.

Moreover, in addition to a positive main effect of *NARS* found across all 8 MaC-DRS domains, we also observed culture specific interactions with the EG-sample in this regard. Positive interaction effects between *culture* (EG-sample) and *NARS* on trustworthiness ($p = 0.024$), reciprocity ($p < 0.001$), and in-group deviance relevance ($p < 0.001$) amplified the response style variable effect on respective margins in a culturally specific way. The corresponding margins in ***Table 31*** should therefore also be interpreted by taking these culture specific interactions into account. It is theoretically plausible to read the response style effect of the Egyptian sample in connection with the logic of honor, among other things.[115] Against this background, it seems that self-confidence and assertiveness are particularly important in the context of deviance towards trustworthiness, reciprocity and in-group in the EG-sample.

Overall, with regard to the question of which moral system guides cooperation in Egypt, we now come to two preliminary conclusions. First, we find in parts an individualizing moral system in the Egyptian sample that is characterized by property and trustworthiness deviance relevance. Next to these domains, clearly, the binding domain of the family in the EG-sample also exhibits fairly high deviance relevance. In addition, the in-group domain in Egypt, as in the JP-sample, takes on higher deviance relevance margins compared to the WEIRD samples in our study. The latter is consistent with tendencies of collectivism and (self-assertive) interdependence in self-construal, that can be found for Egypt (San Martin et al., 2018; Minkov & Kaasa, 2022). Second, we would like to draw on findings that are explained in detail and empirically substantiated in the **Appendix**. We are convinced that the composition of the EG-sample is not nearly representative enough to accurately determine which moral system guides cooperation in contemporary Egyptian society. The main reason for this is that the composition of the EG-sample is heavily skewed towards higher education. Furthermore, we argue that attending institutions of higher education is associated with unobserved heterogeneity that affects the measurement of moral deviance relevance (Rosenbaum, 2005; 2010; Fink et al.,

---

[115] See in this context the discussion of the response style analyses based on the *full sample*. The respective analyses and discussion can be found in the **Appendix**.

2011; Morgan & Winship, 2015). The background to this assumption is that higher education is usually associated with higher socioeconomic status, which in turn may protect against non-severe consequences of failed cooperation (Markus & Schwartz, 2010; Sachdeva et al, 2011; Bourdieu, 1983; 2014; Santos et al, 2017; Kühnen & Kitayama, 2024). In this line, we presume that unmeasured heterogeneity that is connected to higher education leads in large parts to downward biased scores of MaC-DRS. Since in our study the Egyptian sample in particular is strongly biased towards higher education, we therefore hypothesize that it is also primarily the EG-sample that is affected by this effect of unobserved heterogeneity. We found empirical evidence that is in support of our hypothesis. Based on our considerations, we have built samples for the four cultural groups of our study that exclusively include respondents with higher education. Across all four cultural groups and all 8 MaC-DRS domains, these *higher education samples* display lower moral deviance relevance margins as compared to the results obtained via the analyses of the *adjusted sample*. We can therefore present empirical evidence that consistently demonstrates that higher education is associated with lower margins with regard to moral deviance relevance. Hence, we conclude two points. Firstly, the EG-sample of our study is by no means representative of Egyptian society, as it is (e.g.) heavily skewed towards higher education. Our moral deviance relevance findings should therefore only be applied to Egyptian society with due caution. Secondly, based on what we have elaborated, we assume that findings from the adjusted sample for Egypt underestimate the actual moral deviance relevance in Egypt's society due to the presence of a downward bias associated with higher education.

In light of evidence just discussed and considering the results of the *adjusted sample* (**Table 31**), we now address our preliminary conclusions for the EG-sample. So, have we been correct with the cultural difference (**CD**) **sub-hypotheses 1**, **3a** and **3b**? First of all, we must state that the **CD hypothesis 3b** is refuted: The deviance relevance margins of the Egyptian sample in the individualizing domains are not higher than the margins of the Japanese sample. In fact, the reverse is true, and the margins of the JP-sample in the domains of fairness, trustworthiness and property are higher than those of the EG-sample. This fact could, as we suggest, be explained by the moral tightness in Japan. Accordingly, the corresponding hypothesis, which we derived primarily with reference to the self-assertive interdependence in self-construal, must be rejected. Nevertheless, with regard to **CD sub-hypotheses 1 and 3a**, we found mainly evidence that supports our theorizing: Intuitive deviance relevance towards binding morality is indeed promoted in the Egyptian sociocultural context, as demonstrated by the MaC-DRS results. Here too, however, it should be borne in mind that the comparison with the US-sample deviates in

part from the CD sub-hypotheses, which is primarily due to the family deviance relevance margins, as explained in the section on the USA.

Overall, and again, a narrow focus on binding morality would not do justice to the configuration of the Egyptian moral system, at least based on what we have learned from the limited data in our EG-sample. In addition to property, family and trustworthiness, heroism, and in-group, but also deference[116] will be likely regarded as relevant guiding principles of cooperation in Egyptian society. Among these moral domains, property, family and trustworthiness appear, though, to be the main pillars promoting cooperation. Comprehensively, this would speak neither for a binding nor for an individualizing moral system, but for a limited mixed moral system with a distinct deviance relevance composition. Although CD sub-hypothesis 3b could not be confirmed, our findings on moral deviance relevance are nevertheless in line with what San Martin and colleagues (2018) have described in relation to self-construal in Egypt — the way the self is construed in Egypt encompasses both Western (independent) and non-Western (interdependent) aspects. The same seems to be the case with regard to individualizing and binding morality. Altogether, we explain the configuration of this moral system in the framework of the logic of honor (Uskul et al., 2023), by drawing on the heritage of Bedouin culture (Cole, 2003; Reilly, 2013), self-assertive interdependent self-construal, the rise of individualism (Santos et al., 2017; Cai et al., 2019), and with reference to historically high kinship intensity (Schulz et al., 2019).[117] Moreover, it is important to point out that there are reasonable doubts about the robustness of the EG-sample. We based our findings in Egypt on a sample which is heavily biased towards higher education and other sample issues, which among other things partly hinder comparability with the GER-, JP-, and US-sample, were also identified (He & van de Vijver, 2012). The results that we found should therefore be set against this background. Nevertheless, we draw a tentative conclusion based on our explanations and describe Egypt as a cultural entity comprising a ***limited, mixed moral system that emphasizes parts of individualizing and binding mechanisms to ensure cooperation***. In respect to morality, it seems as if the Egyptian cultural context prefers an overall social orientation that centers around individual (mainly property and trustworthiness) and group (mainly family but also in-group and likely deference) aspects of

---

[116] Remember: The EG-sample is heavily skewed towards higher education and we found a negative and significant main effect of educational degrees attaint (Bachelor´s degree and Master´s degree) in relation to the deference deviance relevance margins.

[117] However, it should be emphasized that our interpretations are of course to some extent conjectural, as we rely primarily on indirect evidence from other studies, rather than direct associations, to theoretically explain our results.

cooperation. However, it should be emphasized again that our conclusions should be understood in the context of the limitations of the EG-sample.

4.5.5. Comprehensive Discussion: Which Moral System Guides Cooperation in Different Cultures? MaC-DRS Findings

The aim of our cross-cultural study is to examine human morality in its universal nature and culturally variable characteristics. To this end, we examine a set of four heterogeneous cultures with regard to the intuitive relevance of moral deviance. Overall, we pursue the research question *which moral system guides cooperation in different cultural entities*. We also formulated specific hypotheses about cultural differences (*CD hypotheses*) between the four cultural samples that we examine. In order to analyze what we believe to be the most valid data from our cross-cultural study (data collection 3), we have adjusted our original sample, correcting for cases with poor data quality and further tailoring the dataset to include only cases with solely the citizenship of the countries under investigation. Eventually, based on the *adjusted sample*, we approached our research question and the corresponding hypotheses.[118]

Let us first turn to the overarching discourse of our **hypotheses** on cultural differences (**CD**). From the discourses in the theoretical part of this paper, we derived several overall social orientations for the cultural entities in our study. For **Egypt** and **Japan**, we derived the following: Against the background of (among others) collectivism, interdependence in selfhood, and cultural logics of honor and face, we hypothesized a *group-centered* (interdependent) *overall social orientation* that attaches particular importance to *binding morality*. In contrast, we have derived the following for the **United States of America** and **Germany**: Drawing inter alia on individualism, independence in selfhood, and the cultural logic of dignity, we have proposed to assume an *individual-centered* (independent) overall *social* orientation that is associated with a special importance of *individualizing morality*.

Based on the data obtained with *MaC-DRS* and the analysis of the *adjusted sample*, our heuristic hypotheses about cultural differences are partly confirmed and partly rejected. It also

---

[118] For reasons of accuracy, robustness considerations and to make our research process transparent, we have carried out detailed analyses based on the *full sample* and additional comparisons between the *adjusted sample* and the *full sample*. To view the results of these analyses and to inspect the corresponding tables, we refer once more to the **Appendix**. In the main text, we have analyzed in detail what we consider to be the most valid data basis, namely the *adjusted sample*. Nevertheless, we would like to point out that our samples are not representative and show weaknesses in parts, as can be observed in particular in the EG-sample. On the whole, our results in the GER-, JP- and US-sample are relatively robust. However, this cannot be said for the EG-sample. Overall, we believe it is important to keep these robustness considerations in mind when evaluating the MaC-DRS results — our analyses, findings, and interpretations should be evaluated against the background that we do not have representative samples as a data basis, and also with reference to our robustness analyses.

emerged that the cultures that we examine are indeed diverse in terms of their respective prevalent moral system. For this reason, it appears reasonable in the context of the hypotheses to also consider the individual pairwise comparisons of the deviance relevance margins. *Table 32* revisits our hypotheses once more and shows which are supported and which refuted on the basis of empirical evidence. Furthermore, by commenting on the hypotheses we provide additional information in order to better classify the hypotheses. Reviewing the hypotheses, a culture specific yet systematic binding/individualizing calibration of intuitive moral deviance relevance can be derived *in parts*. We place this systematic calibration of the moral mind primarily in the context of the culturally different emphasis on the group or the individual and the different cultural logics. However, in the course of analysis conducted it also became evident that our results go beyond a pure focus on binding and individualizing. After focusing on *Table 32* we will discuss these results in relation to our research question and hypotheses comprehensively.

Table 32: Hypotheses in the context of cross-cultural moral differences

Main Hypothesis Cultural Differences (CD)	Although we predict universalism of the 8 moral domains proposed by MaC-DRS, we also *hypothesize significant differences in moral domain relevance across cultures.* **This hypothesis is confirmed ✓** - **Note:** Overall, we found more cross-cultural differences than similarities among the groups examined. Moreover, analyses of the effect size revealed that the influence of culture is not only significant in 28 out of 48 pairwise comparisons, but also substantial.
Self-construal	**We expect cultural differences and…**
Sub-hypothesis: CD 1	*...hypothesize that cultural entities that foster relatively more interdependent ways of selfhood also foster relatively more binding morality* (i.e., they have higher relevance ratings of the family, deference and in-group moral domains). **This hypothesis is mainly confirmed ✓** - **Note:** As far as family deviance relevance is concerned, the US-sample is an exception to this hypothesis. Also, in a comparison of the US- and EG-sample, only the in-group deviance relevance proves to be significant.
Sub-hypothesis: CD 2	*...hypothesize that cultural entities that foster relatively more independent ways of selfhood also foster relatively more individualizing morality* (i.e., they have higher relevance of the fairness, trustworthiness and property moral domains). **This hypothesis is mainly confirmed ✓** - **Note:** This hypothesis is largely confirmed, but requires refinement: according to our interpretation, societal modernity and the global rise of individualism condition a cross-cultural fostering of individualizing morality. In our study, therefore, it is not only the German and US socio-cultural contexts that are characterized by a corresponding moral

	system. In the WEIRD samples, however, it can be seen that the corresponding moral systems primarily and in Germany entirely dispense a strong emphasis on binding domains and rely on individualizing morality.
Cultural Logics	
Sub-hypothesis: CD 3a	We hypothesize that cultures of *honor* and *face* are significantly higher in *binding morality* than cultures of dignity. **This hypothesis is mainly confirmed** ✓ - **Note:** Although we were largely correct with this hypothesis, limiting our findings to the logic of honor and face would probably be too simplistic and therefore too short-sighted. It appears much more important to look at culturally specific circumstances from a multi-layered perspective and also taking history into account in order to explain moral systems. One example of this is the policy development in the USA in the context of family deviance relevance. It should furthermore be noted that honor, for example, can also contribute to the emphasis on the individualizing morality, as described in the context of the EG-sample and trustworthiness for instance. Despite the confirmation of this hypothesis, it should be classified appropriately.
Sub-hypothesis: CD 3b	Due to prevailing honor logic and self-assertive interdependence in self-construal, we expect that Egypt, however, scores higher on individualizing domains than Japan. **This hypothesis is refuted** *X* - **Note:** We see moral tightness in Japan as a key factor in explaining the rejection of this hypothesis.
Sub-hypothesis: CD 4	We hypothesize that cultures of *dignity* are significantly higher in *individualizing morality* than cultures of cultures of honor and face. **This hypothesis is refuted** *X* - **Note**: The hypothesis could not be confirmed. We explain this mainly against the background of modernity and the global rise of individualism.

Apart from the limitations of the data basis itself, the cross-sectional design, the indirect evidence of our interpretations, and the self-report design of MaC-DRS, which clearly allows only a confined assessment of intuitive tendencies, our study has nevertheless revealed remarkable results. Across a heterogeneous set of cultures, we have found strong indications that human morality is universal for our species (***Chapter 3***), but that cultural ecologies form distinct moral systems in the course of their socio-historical development. This is evidently portrayed in the findings from the *adjusted sample* on moral deviance relevance across the four cultural samples of our study. We link these different moral systems to independent processes of socio-cultural and historical development, but can also provide consistent explanations in the context of other findings and theories. In our view, collectivism-individualism, (self-assertive)

interdependent and independent self-construal, and the three cultural logics of honor, face and dignity are particularly important in explaining the partly systematic yet by and large culturally specific calibrations of the human moral mind.

The MaC-DRS findings suggest comprehensively that **Germany** is characterized by an individualizing moral system. Furthermore, Germany is the only one of the samples examined in the cultural comparison that specifically emphasizes fairness deviance relevance. In explaining this particular finding, we refer besides an interaction effect found to the history of Enlightenment and the pronounced logic of dignity prevailing in Germany. Beyond that, the GER-sample displays also the lowest binding deviance relevance in comparison across cultural groups studied. Future studies examining other cultural entities characterized by individualism, independence in selfhood, a logic of dignity, a strong welfare state, and low power distance (e.g., Scandinavian countries) would be important for a deeper understanding of the conditions that shape and produce a moral system characterized in particular by individualizing morality and fairness.

In the **United States of America** cooperation is guided by an extended individualizing moral system that encompasses next to individualizing morality also the family domain. We attribute family deviance relevance in the US in part to a lack of welfare state provision and a history of institutional policies supporting familial responsibilities. Based on this interpretation, the intuitive relevance of moral deviance may also be shaped by the institutional worlds in which people are socialized. Hence, our interpretation of the findings suggests that, in addition to ecological sources of influence, also the histories of socio-cultural structures should be taken into account when investigating and explaining moral systems (Berger & Luckmann, 2013; Muthukrishna et al., 2021). Hence, for the future, further empirical comparisons with a broader group of cultural entities would be desirable, which examine institutional policies and their developments in the context of moral deviance relevance. Altogether, although our individualizing hypothesis is confirmed for the US, the results go beyond solely individualizing morality: intuitive property, trustworthiness, fairness and family deviance relevance indicate an extended individualizing moral system for the US American sample in our study.

The JP- and EG-samples each show mixed moral systems of distinct composition. **Japan** tends to a tight moral culture marked by a comprehensive mixed moral system: In this sample, deviant behavior is intuitively evaluated as relevant in regard to all moral domains measured. Moral tightness and a system of thought that tends to integrate rather than compartmentalize traditional and newer social developments, even if they may seem contradictory, account from our point of view to explain this comprehensive mixed moral

system in large parts. In addition, the rise of individualism also seems to be affecting the moral system in Japan, as we have theorized. Based on our results, we believe that it would be worthwhile to examine different systems of thought and their relation to moral systems in a systematic fashion. Two assumptions may guide such endeavors: On the one hand, we predict that an *analytical system of thought* goes hand in hand with a cultural tendency of a delimited moral system, falling mainly either under individualizing or binding. On the other hand, we expect that a *holistic system of thought* is associated with cultural moral systems being more comprehensively configured with regard to the relevance of different moral domains. Overall, future studies that investigate cultural configurations of moral systems and systematically compare several cultural entities characterized by analytical and holistic thought are desirable. Beyond that, studies that integrate normative Tightness/Looseness Theory (Gelfand et al., 2011) with morality would also be very welcome.

In **Egypt**, the results are not quite as clearly interpretable as the data from the other groups, which is among others due to the sample composition. The EG-sample is strongly biased towards higher education, which demonstrably results in lower deviance relevance scores (see: **Appendix**). Nevertheless, relying on the *adjusted sample*, 6 out of 8 moral domains measured prove to be intuitively relevant in this cultural group. Especially, property, family, and trustworthiness appear as the pillars of intuitive moral importance however. Furthermore, we hypothesize that the deference domain would also tend towards intuitive relevance in Egyptian samples that are not biased towards higher education. In order to investigate this notion, further studies in Egypt would be helpful, working with samples that are not predominated by highly educated respondents and that are generally more diverse in terms of socioeconomic positioning of respondents. Nonetheless, the challenge of the initially unsatisfactory sample composition also proved fruitful, raising inter alia a question about the role of socioeconomic status in the intuitive relevance of moral deviance across and within cultures. All in all, based on the MaC-DRS findings we refer to a limited mixed moral system in the case of Egypt. From our interpretation we attribute the configuration of this moral system in part to Bedouin cultural heritage, self-assertive interdependence in selfhood, historical kinship intensity and the logic of honor but also draw on a global increase in individualism.[119]

---

[119] Further and more in-depth investigations of our data are planned for the future. In doing so, we will involve our partners from the target countries more closely in the data interpretation in order to mitigate the WEIRD bias of the current interpretations and to bring the cultural expertise of our partners to the fore. As already mentioned elsewhere, it is important not only to do research on people but also with people, both in general and from an epistemological perspective. The latter means nothing more than including experts from the respective cultures under investigation in the research and interpretation process in cross-cultural studies. Due to time constraints, we were only able to involve our partners from the target countries of our study to a limited extent in the interpretations

With regard to our hypotheses for Japan and Egypt, we found by and large support that the respective cultural contexts foster binding morality relatively more as compared to the WEIRD cultures in our study. However, we have not found a predominance of *binding morality* in the EG- and JP-sample: neither in Japanese nor in Egyptian culture does binding morality seem to be the dominant source that primarily guides cooperation. Instead, our results suggest a general tendency towards *individualizing morality* in the four cultural entities of this study. Taking this into account, we wonder whether hypotheses should be formulated exclusively on the dominant components of a respective moral system, or whether relative differences and also the commonalities of different cultural entities should also be taken into account with regard to morality. We suggest that cross-cultural studies in the field of morality are well advised to look not only for dominant components of moral systems but for *relative* cross-cultural commonalities and differences in order not to neglect important analytical components. In our view, it is an integrative perspective on cultural commonalities and differences that makes a *relative* classification of cultural entities possible in the first place. The suggestions we have made for classifying the four identified moral systems should also be evaluated against this background.

More generally, it can be inferred from our analyses that *binding* and *individualizing* as analytical categories can certainly appear useful in order to guide cross-cultural analyses themselves. Nevertheless, the actual configuration of moral systems seems to go beyond mere binding and individualizing, and cultural moral systems appear to be in fact more diverse (Atari et al., 2022a). This empirical fact resembles the discussions on collectivism-individualism (Krys et al., 2022) and self-construal (Vignoles et al., 2016). Here, too, overarching binaries guide research and theory (Kitayama & Salvador, 2024), but at the same time the concepts are also described as multidimensional comprising many components. Emphasizing the multidimensional view may seem more appropriate for approaching reality. From our perspective the binding and individualizing morality template emerges as an appropriate research heuristic capable of yielding directed hypotheses. However, researchers should be wary of oversimplification and ideally also consider the nuances beyond the polarized pillars of binding and individualizing in order to capture the *actual* moral systems of different cultures. It could therefore prove to be a mistake to limit future cross-cultural research of morality *exclusively* to binding and individualizing, as actual cross-cultural commonalities and differences that would be revealed in more nuanced analyses might be overlooked in this way.

---

presented in this chapter, which is why our data interpretation can be accused of a certain degree of WEIRD bias. Overcoming this limitation is therefore on of the goals of our future work.

However, our study has demonstrated that both the overarching binaries of binding and individualizing as well as nuances of morality can be investigated simultaneously: With MaC-DRS we have a tool that allows for heuristically directed examinations along higher order moral factors and yet nuanced investigations of moral plurality.

Altogether, classifications into binding or individualizing moral systems are possible, but they do not fill the space of possibilities of how moral systems are shaped across cultures. Cultural entities are diverse and so appear their overall-social orientations, the respective significant symbols (codes and scripts) within social situations of everyday interaction, and corresponding moral systems. This is demonstrated by the MaC-DRS findings of an *individualizing moral system* (**Germany**), an *extended individualizing moral system* (**USA**), a *tight and comprehensive mixed moral system* (**Japan**), and a *limited mixed moral system* (**Egypt**).

4.6. Overarching Discussion: Cross-Cultural MaC-DRS Findings

New findings and their interpretation always contribute both to a better understanding of our world, and yet to further (research) questions raised by the findings themselves. This is also the case with the insights and interpretations we have presented in the context of *MaC-DRS* on intuitive deviance relevance of moral domains across cultural entities.

What shapes the moral mind? In the first place this question can be answered with reference to evolution. The panhuman design of our moral mind has its origins in evolutionary processes. In the words of Haidt and Joseph (2007), this is the *universal first draft of* the moral mind. But next we must take into account diverse cultural ecologies, cumulative cultural evolution, and distinct path dependencies that result in variant sociocultural constitutions and culturally contingent calibrations of our moral mind (Henrich & McElreath, 2007; Chudek et al., 2016; Mesoudi & Thornton, 2018; Henrich & Muthukrishna, 2021). The letter is accompanied by the emergence of variant configurations of societal order and cultural moral systems. This is reflected in what Haidt and Joseph (2007) call the *cultural editing* of the first draft of the human moral mind. As we interpret our data, we find evidence for both: across a truly diverse set of cultural entities the CFA models together with full exact measurement invariance of MaC-DRS (see: **Chapter 3**) suggest that the theoretical idea of 8 distinct moral domains is cross-culturally supported. In this evidence, we see a fertile indication of the universality of the 8 moral domains proposed. Indeed, also Curry and colleagues (2019a) found supporting evidence of the universality of 7 of the 8 proposed moral domains across an even

wider set of 60 different societies: across these different societies compliance to respective domains is considered as morally good. In addition to the evidence of universality, our results also clearly demonstrate the cultural editing of the human moral mind: Comparing the relevance margins for moral deviance between different domains and cultural groups, we found 28 significant differences. Our findings on the effect size of these differences underpins the importance of the influence of culture on people's intuitive responses to moral breaches. So, reviewing our empirical results, we conclude to observe evidence suggesting that our main hypothesis of cultural differences is supported: *Indeed, while we find strong indications for the universality of the 8 moral domains proposed by MaC-DRS, we also see evidence of significant differences in moral deviance relevance across cultures.* Importantly, these cultural differences do not appear to be random, but rather systematic to a certain extent: cultural entities that are primarily characterized by individualism, an independent self-construal and prevalence of dignity logic are distinguished by a focus on (extended) individualizing morality. In contrast, cultural entities characterized by collectivism, an interdependent self-construal and prevailing face/honor logic promote in addition to individualizing morality also binding morality. Interestingly, these indications of systematic differences in the cultural calibration of the human moral mind are to a certain extent similar to what Haidt and colleagues (Haidt et al., 2009; Graham et al., 2009; 2016; Iyer et al., 2012; Mooijman et al., 2017) have found (mainly intra-culturally) in relation to moral differences between different political camps. In a cross-cultural context, it is pivotal though to emphasize that our analyses have also revealed that a pure focus on binding/individualizing falls short and that moral systems are diverse in their configuration despite indications of systemic differences. Overall, the human moral mind appears thus to be truly universal to our species yet culture exerts deep influence calibrating intuitive attributions of moral deviance relevance to be adapted to the requirements of people´s immediate sociocultural environment. The latter should also be seen in the context of *modernity* and the accompanying societal demands.

Distal and more recent factors form to blend into the constitution of human morality – cultural heritage persistently effects moral deviance relevance but trends in human development also account for cross-cultural differences and commonalities in morality. Social practices, traditions, values and norms, as well as institutions, policies and cultural products, which include technological changes in human interaction (Dolata, 2011), cause the moral mind to calibrate the intuitive relevance attributed to transgressions in order to align them with the demands of the contemporary constitution of people's socio-cultural ecology. Of large influence in regard to the shaping of the moral mind appear from our interpretations several cultural

dimensions. People´s way of selfhood (independence, interdependence and self-assertive interdependence), cultural collectivism and individualism, institutional policies, as well as cultural logics (face, dignity, honor) and prevailing systems of thought enter into people´s mind adjusting and fine tuning our moral apparatus. These culturally variant factors, themselves formed by sociocultural evolutionary and historical processes, impact from our understanding on the calibration of the moral mind, leading to varying emphasis of intuitive moral deviance relevance. This is to say, across different societies, these constructs reflect, to varying degrees, the same recurring challenges of cooperation that cause our moral minds to prioritize different moral domains and their violations by differences in relevance (Curry, 2016). With regard to these cultural dimensions, a whole field of future research opens up which should strive to demonstrate *direct* empirical associations of our interpretations.

What is striking is that across the heterogeneous cultural groups examined, we found that *individualizing morality* is overall of higher priority than *binding morality*. Respective evidence can be seen in the within-sample ranking of the moral deviance relevance margins to be found in ***Table 31*** provided in the present chapter. This insight seems to parallel evidence from other studies indicating rising individualism across the globe (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019; Kaasa & Minkov, 2020). Accordingly, the overall dominance of the moral domains of fairness, trustworthiness and property raises the question of whether this pattern can be interpreted as an indication of a cross-cultural trend towards greater relevance of the individualizing morality? Do modern societies primarily demand certain domains of cooperation in everyday life, so that people prioritize individualizing morality in order to curb the part of mankind that tends towards egoistic exploitation of cooperation? Future longitudinal studies aimed at investigating trends in the development of the relevance of moral deviance, similar to the study by Kaasa and Minkov (2020), for example, would be desirable in order to deepen the understanding of the socio-cultural demands that modernity places on the human psyche. Such insights into the longitudinal development of morality would allow us to better understand the extent to which societal changes affect the calibration of people´s morality. Studies with small scale societies and societies that still lead a traditional way of life, compared to modern civilizations, would also be very welcome in order to understand and trace the effects of societal development on the respective fine-tuning of moral intuitions. Examinations of this kind likely make the adaptive process and cultural editing of intuitive moral reactions more visible, traceable and better relatable to prevailing societal conditions.

We investigated four cultural groups that are diverse with respect to several cultural dimensions. Our findings in these cultural entities suggest commonalities and differences in the

configuration of moral systems. In addition to main effects, we also found several culture-specific interaction effects in the analyses of the OLS models with the moral domains as the dependent variables. These findings not only illustrate the influences that affect the moral mind across cultures. They also show the importance of understanding and recognizing the culturally specific sources that shape people's moral apparatus in ways that are distinctive to a particular cultural ecology. Looking at the big picture, we were able to identify four different moral systems based on the empirical results. Resting on the MaC-DRS findings we draw the preliminary conclusion that **Germany** is dominated by an individualizing moral system, and the **United States of America** are characterized by an extended individualizing moral system. In the case of **Japan**, we found a tight and comprehensive mixed moral system, while the data for **Egypt** indicates that a limited mixed moral system seems to prevail. Clearly, neither the so-called WEIRD nor the non-WEIRD cultures can be regarded a priori as homogeneous. On the contrary, once more our findings suggest that cultures are diverse and so are corresponding moral systems. In the context of our ideal-typical model (see: **Chapter 1** and **2**), which we set up before the data collection and analysis, we have to conclude that, although we have found indications that support our model, the empirical findings as a whole are more complex and indicate that our model is clearly only of a heuristic value.

Although we belief that our empirical results revealed valuable insights, the small set of cultural entities, the cross-sectional design of our study and the way of data collection brings certainly limits to the study´s outcomes nevertheless. In this context, further cross-cultural investigations that comprise a larger, more diverse set of cultural samples — including e.g. China, India, countries from South America, other African and Middle Eastern countries as well as countries from Scandinavia and Eastern Europe —, as well as studies with longitudinal designs, are desirable. Such empirical enquiries will not only provide more insights into cross-cultural commonalities and differences in social order and the psychology of intuitive moral relevance. They will also likely reveal a glimpse into the demands of current and future social life and its societal organization. Hence, upcoming endeavors into moral deviance relevance are encouraged to expand the insights we found across a wider range of cultures then examined in this study. Evidence portrayed suggests that MaC-DRS offers a reliable and valid way for such endeavors. Thus, we not only call for the use of this instrument in future research, but also for further translations of the scale to make it applicable across the large bouquet of cultures and linguistic groups around the globe.

We should furthermore be aware that the data basis of the present study is certainly problematic in parts, but should at least be seen as a general limitation of our investigations.

The data we collected is not representative of the four cultural entities studied. Also, the effects of culturally specific response styles loom large and must be taken into account when interpreting our findings. Furthermore, we were able to identify a sample bias for the Egyptian sample, which makes it difficult to compare this particular group with the three other samples if we wouldn't take covariates in the models into account. We have demonstrated the latter in this thesis in general, but also in particular within the framework of descriptive analyses and extensive supplementary analyses, which can be found in the **Appendix**. Also, the other samples reveal limitations, as our samples are e.g. older than the average population of the countries we studied. And more generally, we should also pay attention to the likelihood of a selection bias in our data (Morgan & Winship, 2015). We obtained the data for this study from an *online access panel*. Do you regularly take part in online studies to earn money and have you signed up to a company's panel to receive invitations to studies? If, as expected, this does not apply to you, then you are probably not part of such online access panels. Access panels usually consist of registered individuals who have agreed to repeatedly participate in (online) studies. These individuals form the entire data pool of the panel and can then choose to participate in different studies to which they receive invitations. In other words, the individuals who make up the data pool of an access panel are likely to be characterized by a particular motivation, which may be exclusively monetary, that leads them to participate in an access panel and possibly results in a bias in the data based on such panels. This is known in the jargon as self-selection bias (Morgan & Winship, 2015), as the individuals who are part of the panel, as indicated, may be characterized by certain traits/motivations that led them to participate in the panel in the first place, while other individuals are not part of the data and are likely not to have these traits/motivations. What we ultimately want to say is relatively simple: it is important to consider the background of our data collection when drawing conclusions about the four societies under study. Overall, we have referred to this fact at various points, be it in the context of the descriptive findings, in the context of response style bias, or also in the context of the robustness of our analyses. Nevertheless, we also do not have a bad data basis and can certainly keep up with many other cross-cultural studies and the status quo in psychological research: the samples we studied are not limited to students, do not comprise exclusively Westerners, are largely diverse in terms of various socio-demographic characteristics and have sufficient statistical power. Also, due to case exclusions and sample restrictions, the *adjusted sample* provides us with a data basis that is suitable in many respects for comparative cross-cultural studies such as this one. All in all, given the limitations of our database, we should exert appropriate caution in respect to the findings portrayed. This applies in particular to the data

from the Egyptian sample, but also for all four samples in general we should consider the data in the context of their origin and evaluate results obtained against this background. Given the limitations of the data, we would like to encourage future studies with the same research interest to further explore the human moral mind across cultures and to determine how reliable the empirical results presented in this study actually are.[120]

To summarize, what has been shown so far in the context of the **Morality as Cooperation—Deviance Relevance Scale** (MaC-DRS) is the following: deviance towards the moral domains of fairness, trustworthiness, property, reciprocity, heroism, family, in-group and deference is relevant across a diverse set of cultures. This is a fertile indication of the universalism of these domains of morality. Furthermore, our empirical findings demonstrate that a particular culture has a strong influence on the calibration of the human moral mind. Evidence suggests partly systematic cross-cultural differences in the relevance of morally misconduct and yet diversity in the configuration of moral systems. Overall, the moral mind appears to be universal for our species and at the same time culturally contingent in its calibration. These findings coincide with core assumptions of the Moral Foundations Theory (Haidt & Joseph, 2007) and the Morality as Cooperation Theory (Curry, 2016) on which not only our scale but also our theoretical underpinnings are based. What about binding and individualizing morality from a cross-cultural perspective? As stated before, we found some indications of systematic differences between the cultural entities examined. Ultimately, however, a pure focus on binding and individualizing may be too narrow to capture the actual diversity of cultural moral systems. The same basically also applies to the focus of overall social orientations. When working with overarching concepts such as binding and individualizing, we must be careful not to lose sight of the nuances of the cultural constitution of moral systems and the corresponding diversity. After focusing on the *MaC-DRS* results, we turn to the binding vs. individualizing *moral dilemma scenarios* in the following chapter to gain further empirical insights into which moral system guides cooperation in the cultures we are studying. In contrast to MaC-DRS, the dilemma scenarios leave no room for neutrality and the participants are forced to decide for binding over individualizing, or vice versa. This way, we can tease out clear preferences, at least for the juxtaposition of specific moral actions, based on data informed by the dilemma scenarios.

---

[120] We have also found exploratory indications that people's social positioning (Kühnen & Kitayama, 2024) might have an impact on the calibration of the moral mind: At high socioeconomic status, non-drastic moral offenses may not be particularly relevant, as economic, social, and cultural resources might buffer the impact of non-zero-sum consequences. Relevant findings can again be found in the **Appendix**. Next to further investigations into the moral mind across cultures, this line of reasoning clearly calls for more attention and empirical examination in the future.

Chapter 5: Investigations of the Human Moral Mind III

5.1. Moral Dilemma Scenarios: Cross-Cultural Insights

In the previous chapter, we analyzed the four cultures studied with regard to moral deviance relevance and, against this background, focused on *moral intuitions* and culturally variable moral systems. We will now supplement these findings with a cross-cultural analysis of moral dilemma scenarios and a focus on *deliberate moral choice* in particular. Theoretical assumptions suggest that moral intentions and deliberate moral cognition do not necessarily have to lead to one and the same result. We therefore continue to follow our hypothesizing derived from the theory section for the cultural contexts we examine, but this time with a focus on deliberate moral cognition.

Let us briefly revisit our theoretical deductions in the context of the case selection of the four cultural entities that this study examines. We continue to examine Egypt, Germany, Japan and the United States of America. The four cases (i.e., cultural entities) were selected based on a theoretical rationale and partial empirical substantiation of this idea: The cultural contexts we examine differ (among other things) in their *overall social orientation* toward the *in-group* (*interdependence*) or the *individual* (*independence*). As several studies suggest, the former social orientation is characteristic of Egypt and Japan, while the latter is characteristic of Germany and the USA (See e.g.: Markus & Kitayama, 1991; 1998; 2010; Triandis, 2001; Kitayama et al., 2009; Markus & Schwartz; Cross et al., 2011; Leung & Cohen, 2011; Park et al., 2016; Vignoles et al., 2016; San Martin et al., 2018; Minkov & Kaasa, 2022; Uskul et al., 2023; Kitayama & Salvador, 2024). In the following, we focus on deliberate moral cognition and examine two main hypotheses based on a set of moral dilemma scenarios.

> **Hypothesis A):** Consistent with theory, we hypothesize that an *interdependent*, *group-focused orientation* is associated with a greater deliberate importance of *binding morality*. In our study, in particular the **JP-sample** and **EG-sample** should be characterized by an *interdependent*, *group-focused orientation* overall social orientation. Hence, we predict a tendency of deliberate choices towards *binding* morality (as measured via moral dilemma scenarios) for these cultural entities.

> **Hypothesis B):** In contrast, we hypothesize that an *independent*, *individual-focused orientation* is associated with a greater deliberate importance of *individualizing morality*. In our study, in particular the **GER-sample** and **US-sample** should be characterized by an *independent*, *individual-focused orientation* overall social orientation. Hence, we predict a tendency of deliberate choices towards *individualizing* morality (as measured via moral dilemma scenarios) for these cultural entities.

Having emphasized the focus of our subsequent investigations, we will now give a little more space to the dilemma scenarios and deliberate moral cognition. We have developed a total of **9 moral dilemma scenarios.** Each of these scenarios and the corresponding mutually exclusive response options are outlined below.[121] In these dilemma scenarios, two choices are contrasted, implying either binding *conformity* and individualizing *deviance* or vice versa. In order to solve the respective dilemma, respondents must therefore either choose the moral domains of family, in-group and deference, which are considered *binding domains* from a theoretical perspective, or the moral domains of fairness, trustworthiness and property, which are considered *individualizing domains* from a theoretical perspective. Regarding the moral domains, we have thus a 3 x 3 design of moral dilemma scenarios to contrast each binding domain with each individualizing domain.

As the name *dilemma* suggests, the basic structure is such that the two possible response options given for each scenario are mutually exclusive. This type of mutually exclusive choice is the core of the dilemma. Consequently, respondents are forced to decide for only one option per scenario: either a binding or an individualizing choice must be made, and unlike MaC-DRS, no gradations can be made in the choice. It should further be noted that the dilemma scenarios (henceforth abbreviated to DS) differ also in other aspects fundamentally from the MaC-DRS measurement tool. In MaC-DRS, it is general tendencies of *moral intuition* that are the focus of what we want to capture. However, in the moral DS, the assessment of intuition can hardly be assumed — choice between conflicting options rather lead to the activation of our slow thinking, which also seems to be the case in the context of morality, if we follow Haidt's (2001) argumentation. In his intuitionist model of moral judgment Haidt (2001) states:

> "A person comes to see an issue or dilemma from more than one side and thereby experiences multiple competing intuitions. The final judgment may be determined either by going with the strongest intuition or by allowing reason to choose among the alternatives on the basis of the conscious application of a rule or principle" (p. 819).

Lawrence Kohlberg (Kohlberg, 1973; Kohlberg & Hersh, 1977; Sachdeva et al., 2011; Ellemers et al., 2019; Skitka & Conway, 2019) also worked primarily with dilemmas and did not use them to investigate moral intuitions, but rather the reasons for choosing a decision in the dilemma. This means that moral dilemma scenarios were mainly applied in the context of *deliberate moral reasoning*. Moral intuitions certainly also play a role in the choices to be made

---

[121] More detailed insights into the four different language versions of the dilemma scenarios can be found in the pre-registered research plan that is online available: http://dx.doi.org/10.23668/psycharchives.14630.

within the DS that we examine, and this may be supported by the fact that we asked respondents to listen to their gut feeling when making a respective decision. So, the impact of intuitions is by no means to be disputed here. Intuitions are the first, quick reactions that run through our moral mind when we encounter moral phenomena in the social world (for a critical overview see: Skitka & Conway, 2019). Nonetheless, it is probably precisely because of the dilemma itself that people´s explicit, Type 2 cognition comes into play — the nature of a dilemma may imply that it cannot be solved unconsciously, quickly, and effortlessly by the use of routines and gut feelings without a 'second though' (Esser, 2002a; 2010; Greshoff, 2008; Kahneman, 2011; Greenwald & Lai, 2020). Moreover, and this is another integral difference to MaC-DRS, *specific actions* are contrasted in the dilemma scenarios. It is therefore not a question of general tendencies, but rather the opposite: in the dilemma scenarios, specific, individual actions are emphasized and juxtaposed as mutually exclusive options, but not general tendencies of moral intuition. Against this background, the following should be noted: the *dilemma scenarios* represent an independent measurement instrument and the moral content recorded using this instrument likely differs greatly from what *MaC-DRS* measures. In comparison to MaC-DRS, we are not dealing mainly with intuitions that are to be interpreted as general tendencies that transcend situations, but with very concrete situations and deliberate choice considerations of specific moral actions in the respective dilemma scenario.

So, why do we deal with the dilemma scenarios in our cross-cultural study? Quite simply: we want firstly to add another component into the cross-cultural examination of our moral mind by adding insights on (mainly) deliberate moral choice. This is done by examining our hypotheses. Secondly, we want to show that choices for individual moral actions determined by deliberate moral considerations can differ from the general tendencies of people´s moral intuitions. Our world presents an almost infinite variety of situations and general moral tendencies guide us across situations. In contrast, however, certain situations, such as dilemmas, also require a situation-specific reaction and solution — this is where perhaps our slow thinking comes into play, because a dilemma asks for non-routine but likely highly specific and novel solutions (Tutić, 2023). To summarize, we will look at the dilemma scenarios in order to shed light on another component in the cosmos of morality, which is examined here against the background of different entities comprising variant sociocultural ecologies with different social orientations. In the following, we will therefore gain an impression of choices in moral dilemma scenarios that are the result of predominantly deliberate processes. We examine dilemma scenarios. In these scenarios the choice of one of the mutually exclusive options implies *conformity* to a respective moral domain, but at the same time *deviance* to another moral

domain. Moral failure (Tessman, 2014) is thus an integral part of dilemma scenarios, and the question is which failure in relation to which of the juxtaposed domains one is prepared to accept in order to conform to the other moral domain. The dichotomous character of forced choice thus gives us thus an impression of whether binding or an individualizing morality would dominate over the other in the very specific contexts of the scenarios studied.

5.2. Logistic Regression Models to Inspect Moral Dilemma Scenarios

Apart from these introductory remarks, we use single, successively tested logistic regression models, with a set of covariates and interaction terms, to explore the dilemma scenarios. The background to the use of logistic regression (Jann, 2005; Giesselmann & Windzio, 2013) at this point is the binary structure of the moral DS variables (either binding = 0 or individualizing = 1). Our investigations will be based again on the *adjusted sample* (N = 2,360) and we omit any analysis of the *full sample*.

We have gone through a step-by-step process to identify a covariate model for each of the 9 dilemma scenarios that is as comprehensive and at the same time as parsimonious as possible. Initially, in our base model, **Model 1**, we included a large number of variables in the respective covariate models. The aim was then to check whether certain variables consistently proved to be non-significant. We then removed these non-significant variables in a next step with the aim of obtaining more sparse models. Thus, if a corresponding variable proved to be non-significant across all 9 of the logistic regression models, we removed it from the set of covariates to be included in the next model. However, if a corresponding variable was found to be significant in (only) one of 9 models, we retained it for the set of covariates.

In addition to the dependent variable of the respective DS and the variable of interest, i.e., *culture*, we estimated the following covariates in Model 1: all 8 MaC-DRS dimensions (i.e. *fairness*, *trustworthiness*, *property*, *heroism*, *reciprocity*, *family*, *in-group* and *deference*); the *pathogen prevalence measure*, *years in school*, *education* (ISCED), *level of religiosity*, *denomination*, *place of upbringing*, *place of living*, *residential mobility*, *age* and *gender*. We did not include the response style measures *NARS* and *MRS* from the outset due to the binary structure of the dependent variables. The following interaction terms were also taken into account in the initial models: *culture x MaC-DRS* (all 8 dimensions); *culture x age, culture x*

pathogen prevalence, culture x years in school, culture x level of religiosity, and *level of religiosity x denomination*.[122]

It should be mentioned here that although we included the MaC-DRS domains as potential covariates, we do not assume a priori that they are strongly correlated with DS, nor that they necessarily emerge as significant predictors for either choice in the dilemma scenarios. The background to this assumption was already indicated above and builds mainly on the difference in the nature of MaC-DRS (general measure of moral intuitions) and the dilemma scenarios (specific measure of deliberate moral choices). In short, we assume that moral intuition is not to be equated with decisions that are based predominantly on deliberate moral considerations.

Finally, the following variables from these first models were retained for **Model 2**, as they proved to be significant as main effects or in the course of an interaction effect, or were classified by us as being of theoretical interest: all 8 *MaC-DRS* dimensions, *age*, *gender* (female/male), *years in school*, *education* (ISCED), *place of upbringing* (village/countryside vs. city), and *level of religiosity*. These covariates form the basis of the second model. All 9 logistic DS models estimated on the basis of the listed variables and corresponding interaction terms reached convergence. However, when we were to run these models and to correct for alpha error cumulation, we would obtain a test family for all tests (main and interaction effects) related to the variable *culture* that would comprise 47 tests in total. Such a high number of tests would in turn make it very difficult to assess significant effects, given that likely not all effects are highly significant. For this reason, more frugal models are needed. Accordingly, we then opted for separate individual models for each of the 9 DS as the respective final model number three. This is to say, that we keep for **Model 3** the sociodemographic variables of *age*, *gender* (female/male), *years in school*, *education* (ISCED), *place of upbringing* (village/countryside vs. city), and *level of religiosity* for each model investigating the 9 different dilemma scenario choices as dependent variable. However, to reduce the number of tests per model we only include MaC-DRS moral domains that exhibit before Holm-Bonferroni correction significance of $p < 0.05$ in the respective model.[123] Furthermore, we decided again to use the US American sample as reference group in all logistic regression models. Each individual model, together with the corresponding covariate set, is discussed below.

---

[122] In the initial Model 1 we could not estimate the dilemma scenario No. 8 (family vs. fairness) because the model was over-specified and did not converge. All other DS models converged.

[123] For the Holm-Bonferroni correction we have again used the corresponding calculator on the following website: https://statistikguru.de/rechner/adjustierung-des-alphaniveaus.html

5.3. Cross-Cultural Insights on Moral Dilemmas

Before we turn to the respective models for all 9 moral DS, let us first make ourselves familiar with the overall picture. To do this, we inspect the descriptive case distribution onto the binary choice categories for each of the 9 dilemma scenarios. These insights can be found in the following *Table 33*.

Table 33: Descriptive case distribution on dilemma response options

	GER-Sample	JP-Sample	US-Sample	EG-Sample	Total
DS No. 1:					
- Family	427	4175	262	157	1,021
- Property	239	368	307	425	1,339
DS No. 2:					
- Deference	395	408	426	407	1,636
- Trustworthiness	271	135	143	175	724
DS No. 3:					
- In-Group	276	61	164	71	582
- Fairness	390	482	395	511	1,777
DS No. 4:					
- Family	307	240	277	190	1,014
- Trustworthiness	359	303	292	392	1,346
DS No. 5:					
- In-Group	85	95	104	40	324
- Property	581	448	465	542	2,036
DS No. 6:					
- Deference	63	62	141	76	342
- Fairness	603	481	428	506	2,018
DS No. 7:					
- In-Group	40	54	111	42	247
- Trustworthiness	626	489	458	540	2,113
DS No. 8:					
- Family	72	50	105	65	292
- Fairness	594	493	464	517	2,068
DS No. 9:					
- Deference	245	108	200	104	657
- Property	421	435	369	478	1,703
Total	666	543	569	582	2360

The big picture obtained by descriptive analyses is already informative for our hypotheses and shows that in most cases we find a clear preference for the *individualizing* choice option over the *binding* choice. This statement applies in 8 out of 9 dilemma scenarios and largely for *all* groups examined. Especially DS No. 5 to Ds No. 8 reveal strong predominance of individualizing choice. This first impression would speak against our hypotheses for the Egyptian and Japanese samples. However, this impression would yet speak in favor of what we have found as a result of the investigations of MaC-DRS: individualizing morality seems to

dominate by and large over biding morality across the four cultural groups tested. We have attributed this finding in MaC-DRS in the context of increasing individualism and in part to the notion that *modernity* requires primarily the protection of the individual from failed cooperation and zero-sum consequences. The same argument is emphasized here in the context of decisions in moral dilemmas based on deliberate moral considerations. Beyond this cross-cultural trend, however, as can be also inferred from looking at the merely descriptive statistics, there are differences in choice across dilemma scenarios and groups. For this reason, we will subsequently inspect the individual models of the dilemma scenarios one by one and then discuss them together afterwards. We begin with the dilemma in which the domains of family and property are juxtaposed.

5.3.1. Dilemma Scenario No. 1: Family vs. Property

Taking the logistic regression Model 2 as the initial model, we found that none of the 8 MaC-DRS domains prove to be significant before the Holm-Bonferroni correction. Consequently, our final model for DS No. 1 contains only the sociodemographic variables in addition to *culture* and the dependent variable of the dilemma scenario. The first DS requires respondents to make a decision between family conformity and property deviance or vice versa. The scenario with the respective response options is presented below.

Moral Dilemma Scenario No. 1: Family vs. Property

"You go on vacation with members of your extended family and stay in a comfortable hotel. When you leave, two family members tell you that they have taken some precious vases from the hotel. Shortly after you arrive back home, an employee of the hotel calls you and asks about the missing vases..."

Response option A):
You protect your family members and report that you do not know what happened to the vases.

Response option B):
You respect the property of others and report the theft of the vases.

In the final model for DS No. 1, we conducted a total of 15 tests for the test family with the variable *culture* (main effects: *culture*, *years in school*, *level of religiosity*, and *age*; interaction effects: *culture x years in school*; *culture x level of religiosity*; *culture x age*).[124] The model also includes the following variables: *education* (ISCED), *place of upbringing*, and *gender* (female/male).

---

[124] An example of multiple testing: For the main effect of *culture* itself, but also for the interaction effects with the variable *culture*, three tests each result: US-sample vs. GER-sample; US-sample vs. JP-sample; US-sample vs. EG-sample. This results in 12 tests plus three main effects for *years in school*, *level of religiosity*, and *age*. Hence, the test family for the variable culture comprises 15 tests, which is why we must correct for alpha-error cumulation due to multiple testing.

After applying Holm-Bonferroni correction we find three significant effects in our model for DS No. 1. A significant and positive main effect of the variable *culture* and the JP-sample (reference group is the US-sample) can be demonstrated (Coeff. = 2.303; Std. Err. = .714; P>|z| = .013 *). There is also a significant and positive main effect of the variable *age* (Coeff. = .037; Std. Err. = .005; P>|z| = 0.000 ***), as well as a significant and negative interaction effect of the variable *age* and *culture* (JP-sample) for DS No. 1 (Coeff. = -.049; Std. Err. = .008; P>|z| = 0.000 ***). Consequently, the positive main effect of *age* reverses in its direction for the Japanese sample with higher *age*.

Based on this model, let us now take a look at the margins (average marginal effects) for the four cultural groups in our study (***Table 34***). Note, the margins range from 0 (binding) and 1 (individualizing) due to the binary structure of the dependent variable DS No. 1.

Table 34: Margins across cultures: DS No. 1 Family vs. Property

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.511	.023	.465	.557
JP-sample	.704	.025	.653	.755
GER-sample	.371	.025	.322	.421
EG-sample	.726	.037	.653	.799

The culture specific margins show that in this dilemma scenario, the GER-sample tends to support the protection of the family and thus, in turn, decides mainly against the recognition of property. The US-sample is furthermore undecided between the two options, whereas the JP- and EG-sample tend towards property conformity and family deviance in this scenario.

Testing the margins of the four cultural groups in a pairwise comparison for significance, we find Holm-Bonferroni corrected highly significant differences (P>|z| = 0.000 ***) in all comparisons except the comparison of the Japanese and Egyptian sample. The latter shows no significant difference. The values compared between

Figure 14: Margins across cultures — DS No. 1
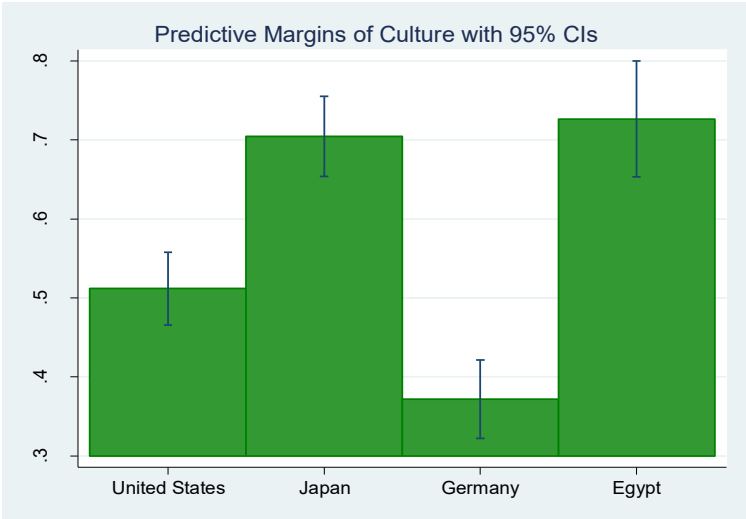
the four groups can also be seen in the following graph (*Figure 14*), which displays the predicted margins given our logistic regression model for all four groups as a bar chart.

Taken together, we not only find cross-cultural differences in DS No. 1, but also an exciting result that seems to support one of the assumptions we made above. Specific dilemma situations, in which only one action in the context of morality becomes prevalent and deliberate moral considerations are significantly involved, can deviate from general moral tendencies based on moral intuitions. Apart from this finding, however, **hypotheses A)** and **B)** are refuted in DS 1: In fact, the expected pattern is reversed, and the Japanese and Egyptian groups show a higher tendency towards individualizing of choice compared to the two WEIRD samples. In comparison to the EG- and JP-sample especially Germany, but also the US-sample tend towards binding choice in this scenario.

5.3.2. Dilemma Scenario No. 2: Deference vs. Trustworthiness

Also in DS No. 2, none of the MaC-DRS dimensions is found to be significant before the Holm-Bonferroni correction (logistic regression Model 2). Consequently, the respective final logistic regression model consists of the same variables as in DS No. 1. In the present dilemma scenario binding and individualizing morality are contrasted by juxtaposing deference vs. trustworthiness. The scenario reads as follows.

Moral Dilemma Scenario No. 2: Deference vs. Trustworthiness

"Your supervisors give you a new work assignment and make it clear that you are to complete it in full by the next morning. You realize that you would have to work longer to complete this assignment on time. However, some time ago you arranged to meet with colleagues in the early evening. On top of that, you have made a commitment to pick up your colleagues by car…"

Response option A):
You follow the orders of your supervisors.

Response option B):
You keep the commitment to your colleagues.

After we corrected for alpha-error cumulation (Holm-Bonferroni), we find neither a main nor a significant interaction effect for any of the tests related to *culture*. However, the model displays a significant and positive main effect for *gender* (binary variable: female = 0, male = 1; Coeff. =.365; Std. Err. = .094; P>|z| = 0.000 ***). This effect yields that males show a significantly higher tendency towards trustworthiness choice in DS No. 2 than females.

Taking now a look at *culture*, we obtain margins based on our model (*Table 35*) that reveal across the groups a clear tendency to choose binding (deference) in scenario No. 2.

Table 35: Margins across cultures: DS No. 2 Deference vs. Trustworthiness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.277	.022	.233	.321
JP-sample	.246	.025	.196	.296
GER-sample	.410	.025	.360	.461
EG-sample	.262	.034	.194	.330

This result could be driven by the potentially punitive or consequential nature of workplace deviance: We could interpret the results to mean that potential work-related consequences outweigh social consequences from colleagues.

When we now look at the pairwise comparisons of the margins for DS No. 2 across the four groups in our study, we come to the following conclusion: the comparisons of the margins between Egypt and Japan, for the JP-sample and the US-sample, as well as for the US- and EG-sample are not significant. The situation is different when we look at the comparison between the German sample and the Japanese and US American samples (Holm-Bonferroni corrected): both comparisons show highly significant differences in the margins (P>|z| = 0.000 ***). The pairwise comparison between the German and the Egyptian sample is also significant (P>|z| = 0.004 **). Although all samples (see: **Table 35**) tend towards deference, this tendency is less pronounced in the GER-sample than in the other cultural groups examined. This finding can also be derived from the following bar chart (**Figure 15**).
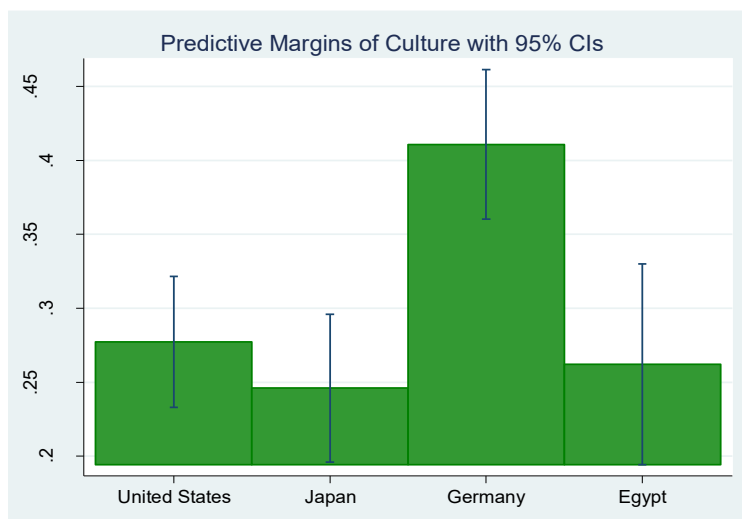
Figure 15: Margins across cultures — DS No. 2

Interestingly, we again found that the German sample is particularly different from the other groups studied. This time, although all groups tend to choose binding over individualizing in DS No. 2, the German sample differs from the other three samples in that the respective choice lies somewhat between the two available options. Even though this result is partly in line with what we found in MaC-DRS in regard to deference deviance relevance (see: **Table 31** further above; **Chapter 4**), none of the 8 moral domains assessed via MaC-DRS proved to be a relevant predictor of choices in DS. No. 2. As far as our **hypotheses** are concerned, we are correct for Egypt and Japan. However, the two WEIRD samples also tend towards binding choice in this scenario. Despite this cross-group tendency, though, we can also identify a distinction for the GER-sample that is congruent with the hypothesis, as this sample shows a higher tendency towards individualizing choice in comparison. The same does yet not apply to the US-sample, leading to a partial rejection of hypothesis B).

5.3.3. Dilemma Scenario No. 3: In-Group vs. Fairness

In the logistic regression model on the dilemma scenario that juxtaposes in-group vs. fairness we found again no effect (main or interaction) of MaC-DRS and hence, the final model for DS No. 3 equates the models from the dilemma scenarios discussed before. The content of the third DS can be viewed below.

Moral Dilemma Scenario No. 3: In-Group vs. Fairness

"You are at a shop with a group of friends. A friend who knows a lot about watches discovers two original watches of a famous brand. In order to push down the selling price, your friends together insistently talk to the shop employee to convince him that the watches are fakes. Then you are asked what your opinion is..."

Response option A):
You stand by the statement of your friends.

Response option B):
You behave fairly towards the trader.

We found after correcting for alpha-error cumulation a significant and positive main effect for *age* (Coeff. = .038; Std. Err. = .006; P>|z| = 0.000 ***). The *education variable* (ISCED) also proved to be significant initially before correction for alpha-error cumulation, but not significant afterwards. However, we found again an effect for *gender*. This effect is negative and significant (binary variable: female = 0, male = 1; Coeff. =-.367; Std. Err. = .106; P>|z| = 0.001 **), indicating a tendency for females to make more individualizing choices (i.e., fairness) in DS No. 3 than males.

Taking now against this background a look at the margins for *culture*, we see that the Japanese and the Egyptian sample tend towards a preference of choosing fairness. The same,

however, to a lesser degree, holds true for the US-sample and we find the German sample in between the chairs of the two options available in the scenario (***Table 36***).

Table 36: Margins across cultures: DS No. 3 In-Group vs. Fairness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.679	.022	.653	.741
JP-sample	.888	.017	.852	.923
GER-sample	.592	.025	.541	.642
EG-sample	.819	.036	.748	.890

Pairwise cross-cultural comparisons of the margins reveal that five out of six of these comparisons are significant. The Egyptian sample compared to the Japanese sample again shows non-significant differences when comparing the respective margins. However, the comparisons between Japanese and the GER- and US-samples show highly significant differences (P>|z| = 0.000 ***). Also, the margins between the EG- and GER-samples (P>|z| = 0.000 ***) are significantly different. In addition, the margins for the US American and German sample (P>|z| = 0.006 **), and the margins for the EG- and US-sample (P>|z| = 0.01 *) are empirically significantly different from each other in the model on DS No. 3. The findings described can also be derived from ***Figure 16***.

Once more we find the German sample to be partly an outlier when compared to the margins of the other cultural samples. Next to this undecided tendency between binding and individualizing found in the GER-sample, we see for DS. No. 3 margins that indicate preference for choosing fairness in the US-, the EG-, and the JP-sample. Accordingly, the results reveal that we must by and large reject our **hypotheses**: the highest individualizing choice tendencies
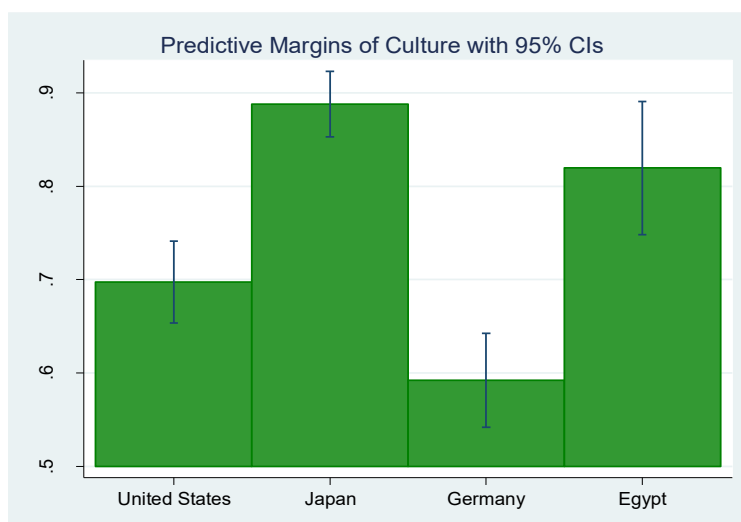
Figure 16: Margins across cultures — DS No. 3

in this scenario are found among the EG- and JP-sample. By contrast, the US American and the German sample show significantly lower individualizing choice tendencies as the non-WEIRD samples. The expected pattern is thus found to be in fact reversed.

5.3.4. Dilemma Scenario No. 4: Family vs. Trustworthiness

As we found again no significant effect of either a main or an interaction effect for all MaC-DRS dimensions, the respective final model for DS No. 4 remains the same as in the previous dilemma scenarios. Furthermore, after the Holm-Bonferroni correction, we were unable to identify any other effects in our model apart from a significant and negative effect for *gender* (binary variable: female = 0, male = 1; Coeff. =-.187; Std. Err. = .087; P>|z| = 0.032 *).[125] The dilemma scenario No. 4 asks respondents to make a choice between family and trustworthiness. The corresponding scenario is described below.

Moral Dilemma Scenario No. 4: Family vs. Trustworthiness

"Your family is getting together with all the relatives for a big family reunion, which you are supposed to attend. However, you remember that you promised to help out at an association as a volunteer on the very weekend the reunion will be held...

Response option A):
You fulfill the expectations of your family and go to the family reunion.

Response option B):
You keep your promise to the association.

The culture specific margins for the family (binding) vs. trustworthiness (individualizing) scenario are as follows (**Table 37**).

Table 37: Margins across cultures: DS No. 4 Family vs. Trustworthiness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.505	.024	.456	.553
JP-sample	.570	.029	.512	.628
GER-sample	.525	.026	.474	.577
EG-sample	.567	.043	.482	.652

Apparently, respondents seem to find it hard across groups to favor one option over the other. We find all margins with a slight tendency towards individualizing, but overall, with an undecided tendency between the mutually exclusive response options of the dilemma. Cross

---

[125] The significant gender effect that we found indicates that females are more likely than males to choose the individualizing domain of trustworthiness in the given dilemma scenario.

group comparisons of the group specific margins support this statement: no significant differences between groups can be observed (see: *Figure 17*).
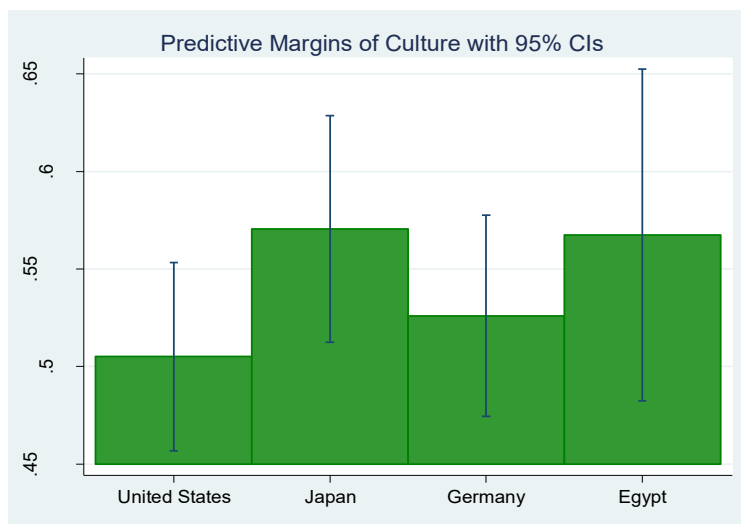
Figure 17: Margins across cultures — DS No. 4

With regard to DS No. 4, which contrasts a choice between family and trustworthiness conformity respectively deviance, we see a clear cross-cultural tendency: Neither individualizing nor binding is preferred and there remains a stalemate between the two options. In the light of this evidence, we see our **hypotheses** once more refuted.

5.3.5. Dilemma Scenario No. 5: In-Group vs. Property

The dilemma scenario No. 5 contrasts in-group vs. property. As we have found in the model uncorrected for alpha-error cumulation significant interaction effects associated with fairness and deference deviance relevance, our logistic model for DS No. 5 comprises these two MaC-DRS dimensions in addition to the sociodemographic variables that also the previous models comprised. After applying Holm-Bonferroni correction (23 tests of a test family) and adjusting significance levels, we find, however, only a positive main effect of the *age* variable to remain significant (Coeff. = .044; Std. Err. = .007; P>|z| = 0.000 ***). All other variables are proved to be insignificant in the model having DS No. 5 as dependent variable. The textual description of DS No. 5 can be viewed below.

Moral Dilemma Scenario No. 5: In-Group vs. Property

"You are out with your friends when a person loses a wallet in front of you. Someone in your group picks up the wallet. It contains just under $400 and your friends decide to spend the money together. Shortly thereafter, the person slightly ahead of you notices the loss of the wallet, turns around looking, approaches your group and asks if you have seen a wallet. One of your friends replies, "we didn't see a wallet or anything"…"

Response option A):
You stick to your group and report not having seen a wallet.

Response option B):
You report finding a wallet, since it is someone else's property.

The margins we obtain from our logistic model consistently point in one direction: the individualizing option (i.e., property) is preferred in across all cultural groups tested.

Table 38: Margins across cultures: DS No. 5 In-Group vs. Property

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.812	.019	.773	.851
JP-sample	.828	.023	.782	.875
GER-sample	.853	.022	.810	.896
EG-sample	.923	.024	.875	.970

As far as the cross-cultural pairwise comparisons of the margins are concerned, the following can be observed: In addition to four non-significant differences, we find a significant difference for Egypt and the United States of America (P>|z| = 0.000 ***), and also for the comparison between the EG- and JP-sample (P>|z| = 0.03 *). **_Figure 18_** graphically depicts our results of the culture-specific margins for DS No. 5 as bar chart.

Figure 18: Margins across cultures — DS No. 5
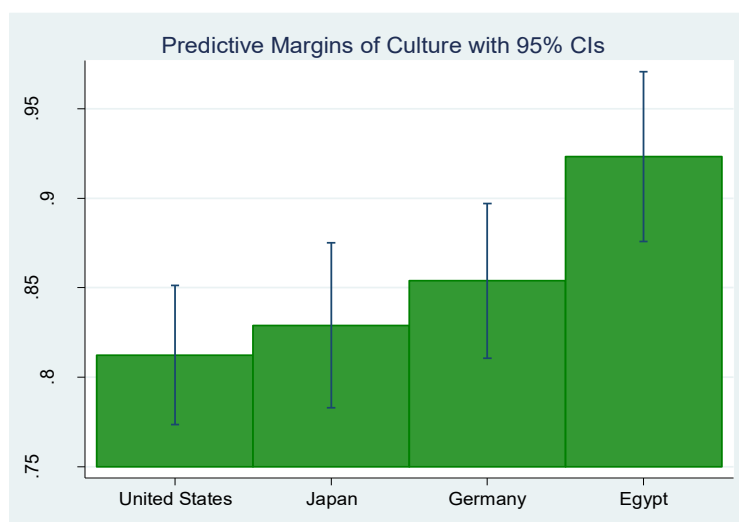
Taken together, we thus find a clear tendency towards the choice of individualizing morality in the dilemma scenario that contrasts choice for in-group deviance/conformity with choice for property deviance/conformity. This tendency extends across all four groups we examined, but is particularly pronounced in the EG-sample. In line with our results, **hypotheses A)** and **B)** are by and large rejected. We have found indeed arguments in favor of hypothesis B), but in DS No. 5 it is the Egyptian group in particular that tends towards the individualizing choice, while the American sample, although also exhibiting an overall individualizing tendency, shows significantly lower margins.

5.3.6. Dilemma Scenario No. 6: Deference vs. Fairness

In the uncorrected model for the dilemma choice between deference and fairness (DS No. 6) we find main and interaction effects for fairness deviance relevance. Consequently, we keep the MaC-DRS variable fairness next to the sociodemographic variables of *age*, *gender* (female/male), *years in school*, *education* (ISCED), *place of upbringing* (village/countryside vs. city), and *level of religiosity* in the respective logistic regression model. After correcting for alpha-error cumulation results yield three significant effects. The two main effects of *age* (Coeff. = .031; Std. Err. = .006; P>|z| = 0.000 ***) and *fairness deviance relevance* (Coeff. = .194; Std. Err. = .058; P>|z| = 0.000 ***) prove to be significant and positive. However, the letter effect is attenuated for the Egyptian sample, as shown by a significant and negative interaction effect between *culture* (EG-sample) and *fairness deviance relevance* (Coeff. = -.297; Std. Err. = .078; P>|z| = 0.000 ***) on DS No. 6. The sixth dilemma scenario that we assessed reads as follows.

Moral Dilemma Scenario No. 6: Deference vs. Fairness

"In a company you are responsible for supervising and testing interns. All interns are equally good, no one is more talented than the others. Before the exams, your boss talks to you. He instructs you to give the two interns whom your boss knows well the best grades in the exams, no matter what…"

Response option A):	**Response option B):**
You follow your boss's instructions and give preference to two interns in grading.	You grade all interns according to just criteria.

After running the respective logistic regression model for DS No. 6, we obtain the following margins (*Table 39*) for each cultural group.

Table 39: Margins across cultures: DS No. 6 Deference vs. Fairness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.749	.021	.707	.792
JP-sample	.890	.018	.855	.926
GER-sample	.877	.021	.835	.919
EG-sample	.822	.038	.747	.898

When comparing these margins in pairs between the groups studied, we furthermore encounter two significant differences between cultural entities. The Japanese and the German sample differ from the US American sample significantly by higher margins (P>|z| = 0.000 ***). Nevertheless, in general, a trend towards the individualizing pole can be recognized across

groups. *Figure 19* graphically displays the margins obtained for each group on DS No. 6 as a bar chart.
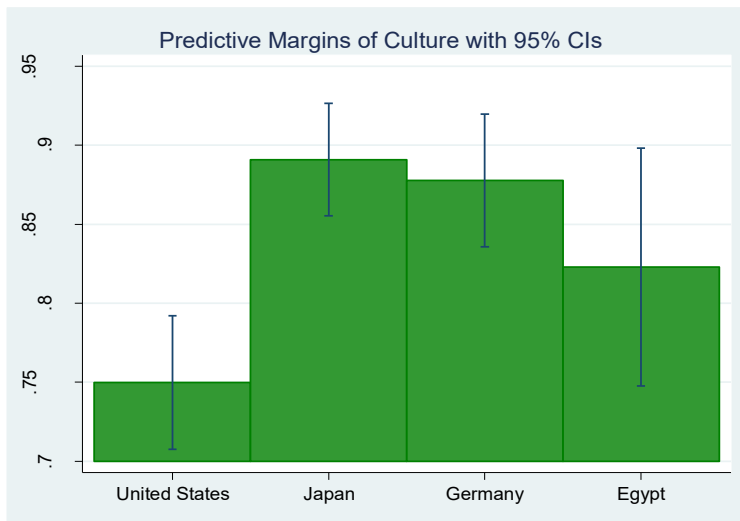
Figure 19: Margins across cultures — DS No. 6

Overall, next to the main effect of *age*, proving so far as the most relevant factor measured across models, we found this time a significant main and interaction effect of the MaC-DRS fairness domain impacting on DS No. 6. But this main effect is attenuated for the EG-sample, as a culture-specific interaction shows. Apart from these results, a trend to choose fairness over deference, although with cross-cultural differences, can be recognized across the cultural entities examined. With regard to our **hypotheses**, which predict a tendency of deliberate choices towards *binding* morality for the EG- and JP-samples and a tendency of deliberate choices towards *individualizing* morality for the US- and GER-samples, we are, by and large, once again refuted. Although there is indeed an *individualizing* choice tendency in the WEIRD samples, this can also be observed in the Japanese and Egyptian samples. Moreover, not only does the German sample show a pronounced individualizing tendency in DS No. 6, but the results for the Japanese group also reveal a significantly higher individualizing choice average marginal effect than found in the US-sample.

5.3.7. Dilemma Scenario No. 7: In-Group vs. Trustworthiness

Taking a look at the uncorrected model for DS No. 7, we find three MaC-DRS variables displaying a significant main and/or interaction effect. These variables are: fairness, in-group and property deviance relevance. Consequently, we keep these variables next to the sociodemographic variables in the logistic regression model for the dilemma choice between in-group (binding) and trustworthiness (individualizing). Comprising main and interaction

effects we obtain thus 27 tests related to *culture*, for which we must correct for multiple testing. After Holm-Bonferroni correction three effects of the *culture* variable test family remain significant. The model displays a positive main effect of *age* (Coeff. = .054; Std. Err. = .008; P>|z| = 0.000 ***), a negative main effect of *in-group deviance relevance* (Coeff. = -.325; Std. Err. = .099; P>|z| = 0.025 *), and a negative interaction effect between *culture* (JP-sample) and *age* for DS No. 7 (Coeff. = -.061; Std. Err. = .012; P>|z| = 0.000 ***). The positive main effect of *age* is thus reversed for the Japanese sample. Next to the effects related to the test family of *culture*, we also encounter a main effect for the variable *residential mobility*. Reference category is the response option "*I have always lived in the same neighborhood*", and increasing *residential mobility* also heightens individualizing choice in DS No. 7: "*I have lived in different neighborhoods in the same country*" (Coeff. = .776; Std. Err. = .218; P>|z| = 0.000 ***); "*I have lived in different places in the same country*" (Coeff. = .437; Std. Err. = .181; P>|z| = 0.032 *); "*I have lived in different countries*" (Coeff. = .724; Std. Err. = .311; P>|z| = 0.032 *).[126] No other effects were found for DS No. 7, which can be inspected below.

Moral Dilemma Scenario No. 7: In-Group vs. Trustworthiness

"An acquaintance of yours tells you something confidential in private and asks you to keep the content of the conversation to yourself. When you meet with your friends the next day, someone in the group asks about the conversation with your acquaintance. Among your friends, people always tell each other everything, and so this time, too, your friends expect to know the content of the conversation with your acquaintance..."

Response option A):

You follow the expectations of your group and tell your friends the content of the conversation.

Response option B):

You tell nothing of the conversation and keep the trust of your acquaintance.

The four cultural groups examined show a clear tendency towards the individualizing pole of choice options, as can be inferred from the group specific margins (***Table 40***).

Table 40: Margins across cultures: DS No. 7 In-Group vs. Trustworthiness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.794	.020	.753	.834
JP-sample	.911	.017	.876	.946
GER-sample	.884	.021	.843	.926
EG-sample	.928	.022	.885	.972

---

[126] The significance level is corrected for alpha-error cumulation by applying Holm-Bonferroni correction.

In terms of pairwise comparisons, we find significant differences between the US American sample and all other groups: US-sample vs. JP- and EG-sample (P>|z| = 0.000 ***); US-sample vs. GER-sample (P>|z| = 0.008 **). This difference can be found depicted in the **Figure 20**.

Figure 20: Margins across cultures — DS No. 7
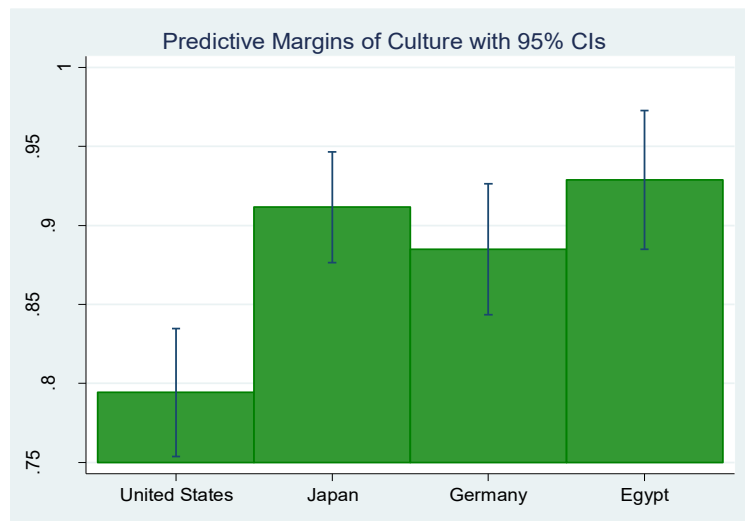
So, although we find clear evidence of a cross-cultural tendency towards individualizing in DS No. 7, we also see significant differences between the groups studied. The US-sample exhibits lower margins than the other groups, but overall trustworthiness outcompetes the in-group choice. While the results would at first glance support hypothesis B), the comparison of the four groups and the significantly lower margins of the US-sample lead to the rejection of both **hypotheses A)** and **B)**. Besides these findings also (intuitive) in-group deviance relevance impacts on the choice in this dilemma scenario.

5.3.8. Dilemma Scenario No. 8: Family vs. Fairness

The eighth dilemma scenario juxtaposes family and fairness as mutually exclusive options of choice in a given scenario. The uncorrected model results for DS No. 8 display significant effects for the MaC-DRS dimensions property, family, and in-group, which is why we kept these domains in the final regression model for this dilemma scenario. After correcting for alpha-error cumulation results yield three significant effects. The variable *age* exerts a positive main effect on choice in DS No. 8 (Coeff. = .063; Std. Err. = .008; P>|z| = 0.000 ***). However, this positive effect is found to be reduced for the Egyptian (Coeff. = -.050; Std. Err. = .014; P>|z| = 0.025 *) and the Japanese sample (Coeff. = -.063; Std. Err. = .012; P>|z| = 0.000 ***). The latter is indicated by a negative significant interaction effect between *culture* (EG- and JP-sample) and *age* for the dilemma choice between family (binding option) and fairness (individualizing option). Next to the tests related to *culture*, we found again an alpha-error corrected significant and positive main effect for *residential mobility*. The categories "*I have lived in different neighborhoods in the same country*" (Coeff. = .796; Std. Err. = .202; P>|z| =

0.000 ***), and "*I have lived in different places in the same country*" (Coeff. = .405; Std. Err. = .165; P>|z| = 0.028 *) display a significant, positive effect as compared to the response category "*I have always lived in the same neighborhood*". Higher residential mobility is thus in this dilemma scenario associated with increased individualizing choice tendency. No other effects in the model for DS No. 8, shown below, were proved to be significant.

Moral Dilemma Scenario No. 8: Family vs. Fairness

"You are a coach of a youth team in a sports club where a relative of yours has been training for a while. Your relative is not bad, but still has some catching up to do. Now there is another important competition coming up, for which you decide the team. Your family knows about this competition and wants you to nominate your relative for the team…"

Response option A):
You comply with your family's request and select your relative for the competition team.

Response option B):
You select the competition team according to the performance principle.

Looking at the margins for each of the cultural groups examined (*Table 41*), we see a clear trend towards choosing fairness (i.e., individualizing) over family (i.e., binding) in DS No. 8.

Table 41: Margins across cultures: DS No. 8 Family vs. Fairness

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.797	.020	.757	.836
JP-sample	.898	.022	.854	.941
GER-sample	.863	.018	.825	.900
EG-sample	.889	.026	.836	.942

The margins of the US-sample are slightly below those of the other groups, but only significantly different from the Japanese sample (P>|z| = 0.035 *). Consequently, evidence suggest a general tendency towards the individualizing choice for DS No. 8, as can also be seen from the bar chart in *Figure 21*.

Figure 21: Margins across cultures — DS No. 8
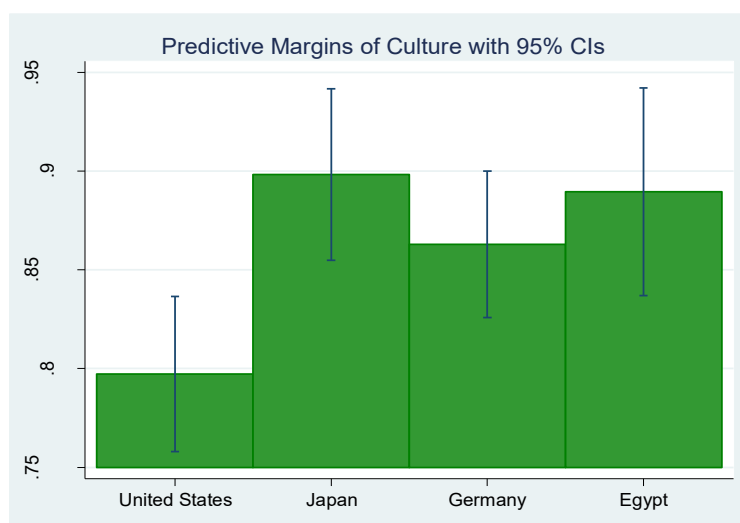
As indicated by the cultural group specific margins, the dilemma scenario depicted above is across cultures mainly solved by choosing the individualizing option of applying fair criteria. Considering the general individualizing tendency found in all groups, and also taking into account the significantly lower margins in the US-sample as compared to the Japanese group, we must again accept the rejection of **hypotheses A)** and **B)**.

5.3.9. Dilemma Scenario No. 9: Deference vs. Property

The last of the scenarios examined is a choice in a dilemma situation between mutually exclusive deference and property conformity respectively deviance options. Again, we inspected first the uncorrected model and found significant effects for the following MaC-DRS dimensions: in-group, family, trustworthiness and fairness deviance relevance. As in the previous models, we kept these variables in addition to sociodemographic variables in the final logistic regression model for DS No. 9. The test family related to culture comprises 31 single tests. After we corrected for alpha-error cumulation we found among these tests two remaining significant effects. The variable *age* was found again to exert a positive main effect (Coeff. = .028; Std. Err. = .006; P>|z| = 0.000 ***). Additionally, and to our surprise, we found a negative interaction effect between *culture* (EG-sample) and *fairness deviance relevance* (Coeff. = .262; Std. Err. = .114; P>|z| = 0.03 *) in the dilemma scenario that we designed to contrast deference and property as mutually exclusive options of choice. Next to these effects also respondents *gender* revealed to exert significant impact on the choice: being female is associated with a tendency for the individualizing choice in DS No. 9 (binary variable: female = 0, male = 1; Coeff. =-.409; Std. Err. = .100; P>|z| = 0.000 ***). The respective scenario of dilemma No. 9 can be inspected below.

Moral Dilemma Scenario No. 9: Deference vs. Property

"You visit a trade fair with your supervisors. After hearing another company's presentation, your supervisors decide to take up the product idea presented and instruct you to do so. In the process, you're not supposed to involve the actual idea developers, which essentially means stealing the other company's idea..."

Response option A):
You follow your supervisors' instructions and carry out the order.

Response option B):
You preserve the intellectual property of the other company.

Although we find again a general tendency to choose individualizing more often than binding, the margins (*Table 42*) yet already indicate cross-cultural differences in this choice.

Table 42: Margins across cultures: DS No. 9 Deference vs. Property

	Margin	Delta-method Std. Err.	[95% Conf. Interval]	
US-sample	.648	.024	.600	.696
JP-sample	.808	.026	.756	.860
GER-sample	.608	.027	.553	.662
EG-sample	.771	.039	.694	.848

Inspecting now the pairwise comparisons of margins obtained for each culture, we see us supported in the hunch stated above: the GER- and US-sample do not differ from each other and the same holds true for the comparison between JP and EG-sample. However, the German sample differs significantly from the Japanese (P>|z| = 0.000 ***) and the Egyptian (P>|z| = 0.004 **). Also, the US American sample holds significantly lower margins on DS No. 9 than the JP-sample (P>|z| = 0.000 ***) and the EG-sample (P>|z| = 0.027 *). These differences are graphically depicted in the bar chart of *Figure 22*.

Figure 22: Margins across cultures — DS No. 9
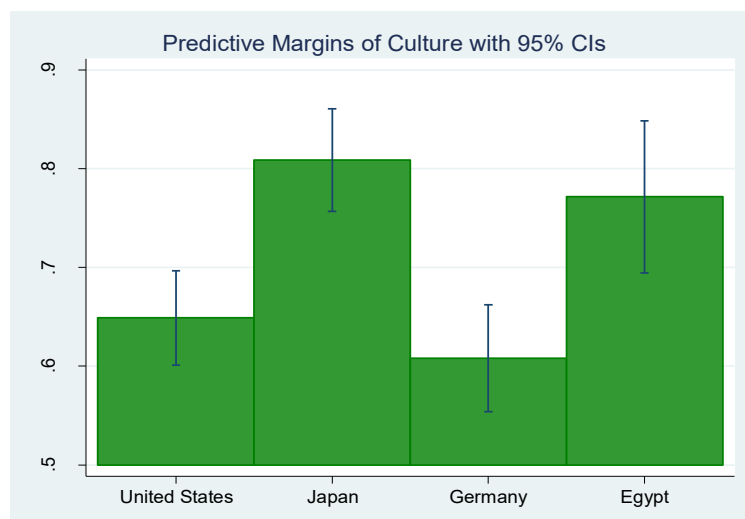
Taken together, we found again a cross-cultural tendency to choose individualizing over binding, but apart from the general tendency also cross-cultural differences were revealed. Compared to the GER- and US-sample the Japanese and the Egyptian sample exhibit a stronger tendency to place recognition of property over deference in DS No. 9. In addition to the cross-cultural tendency, differences between the samples can also be observed, which, if we proceed from our hypotheses, provide by and large evidence to the contrary. The results therefore clearly suggest that **hypotheses A)** and **B)** should be rejected when drawing on the findings from dilemma scenario that contrasts choice between deference and property.

Overall, in light of the present results and those of the previous dilemma scenarios, we found convincing evidence that intuitive moral tendencies captured by MaC-DRS (may) play a role

in the decision-making process in moral dilemmas, but not necessarily so. Our findings suggest that **moral intuitions** do not have a consistent influence on the decision-making process between two mutually exclusive options, but rather come into play in certain but not all dilemma situations and sometimes even in surprising ways. With regard to the **hypotheses** that binding options should be weighted more heavily in the Egyptian and Japanese contexts than in the WEIRD samples (and vice versa), we found clear evidence that is not supportive of this assumption. Apart from finding **cross-cultural tendencies in favor of deliberate individualizing choice** in the majority of cases, we also found evidence that these tendencies are even more pronounced in some dilemma scenarios in the EG- and JP-samples than in the German and US American samples. The picture we painted in our hypotheses is in fact even reversed in the context of these findings. On the whole, then, the analyses of the dilemma scenarios provide consistent evidence that **individualizing morality is indeed important even beyond WEIRD cultures**. The latter applies, as suggested by empirical evidence, not only to the sphere of moral intuitions (see the MaC-DRS analyses in **Chapter 4**), but also to the realm of **deliberate moral cognition** that becomes effective in solving dilemma scenarios.

5.4. Discussion: Deliberate Choices in Moral Dilemmas Across Cultures

The study of moral reasoning requires a cross-cultural perspective, as argued, among others and entirely rightly, by Sachdeva and colleagues (2011). Certainly, a lot has happened in this respect in the field of moral psychology over the last decade (See e.g.: Graham et al., 2016; Curry et al., 2019a; Apicella et al., 2020; Awad et al., 2020; Henrich & Muthukrishna, 2021). But more cross-cultural research is still needed to better illuminate the human moral mind. We aim to make a small contribution to the cross-cultural examination of moral reasoning with the present study on *binding vs. individualizing moral dilemmas*. Our investigation of the dilemma scenarios has yielded various results, but limitations must also be noted. Therefore, we will now turn to a general discussion of the findings and background to our investigations. In doing so, we focus on four main points.

First, we are turning to our **hypotheses**. Informed by other studies and derived from theoretical notions we assume that Egypt and Japan are marked by focal social orientation on the group, while US America and Germany hold the individual as focal point of social orientation. Accordingly, we have hypothesized that an *interdependent*, *group-focused orientation* is associated with a greater deliberate importance of *binding morality* (**Hypothesis**

A) and that, by contrast, an *independent, individual-focused orientation* is associated with a greater deliberate importance of *individualizing morality* (**Hypothesis B**). **The empirical findings drawn from 9 binding vs. individualizing moral dilemma scenarios are not in support of our hypotheses**. Conclusively, also when we assess processes more akin to deliberate moral considerations, we see a general, cross-cultural tendency that indicates dominance of individualizing choice over binding choice. This pattern is found beyond WEIRD cultures as it applies to the whole set of heterogenous cultural groups examined. In the context of deliberate moral cognition, which is effective in the resolution of the mutually exclusive choice between conformity to one moral domain and simultaneous deviation from another domain, it is *individualizing morality* that trumps *binding morality* in its weight across cultures. This is demonstrated by our moral dilemma scenario analyses. In this respect, the dilemma scenario findings are in line with what we have discovered in the inspection of the MaC-DRS results (***Chapter 4***). Thus, not only *moral intuitions* but also moral choices that hinge more on *deliberate cognition* seem to give precedence of individualizing morality over binding morality in the four groups studied. The results obtained from the 9 dilemma scenarios described above give thus further support to the interpretation of a cross-cultural trend towards individualizing morality. We place this result in the context of increasing individualism and modern societies, as already discussed in detail elsewhere (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019). However, as previously noted, we are working here with data from a cross-sectional design, and consequently our trend statement should be seen against this background. Along with the predominance of choosing individualizing morality in the dilemma scenarios, we observed nonetheless also cross-cultural differences in the tendencies of the choices, as demonstrated and touched on above. Overall, 28 out of 36 pairwise comparisons of the group specific margins that reflect choice tendencies in moral dilemmas proved to be significantly different across the cultural groups examined. These findings suggests that while we find a clear predominance of individualizing morality across the cultural entities studied, there are still significant differences in the decision tendencies of how moral binding vs. individualizing dilemmas are resolved. Importantly to note, the cross-cultural difference findings yield that the EG- and JP-sample partly show significantly higher tendencies of choosing individualizing morality than the WEIRD samples. These insights underline once again that our hypotheses in the cross-cultural examination of deliberate moral cognition have been refuted. Overall, the fact of significant differences between the entities examined relates to what we have also discussed in the context of the MaC-DRS results: Despite similar cross-cultural trends in the calibration of the human moral mind, culture-specific influences play yet a substantial role in fine tuning people´s moral

tendencies. The latter is evident not only in regard to general intuitive tendencies of culture specific moral deviance relevance, but also in the context of rather deliberate cognitive processes of choice in moral dilemma scenarios.

Second, (increasing) *residential mobility*, but mainly *gender* and above all *age* impacted on the decisions in the dilemma scenarios tested. In 7 out of 9 dilemma scenarios *age* exerts a positive and significant impact. This finding suggests that older age is cross-culturally associated with a higher likelihood of choosing individualizing over binding morality. Other studies cited before examined the cultural dimension collectivism-individualism, and found a trend of rising individualism across the globe (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019; Kaasa & Minkov, 2020). Drawing inter alia on modernization theory, the respective studies suggest that, among other things, a positive development of national GDP contributes to the trend of increasing individualism (With respect to the individual level of self-construal see also: Markus & Schwartz, 2010; Kühnen & Kitayama, 2024). Our results on the influence of *age* on decisions in favor of individualizing morality can presumably also be interpreted in this sense. Increasing national GDP is strongly associated with longer life expectancy (Jetter et al., 2019), and, as we found, older age is in turn associated with a greater likelihood of choosing individualizing morality over binding morality in dilemmas with mutually exclusive options. This line of reasoning once again supports the interpretation that places the cross-cultural individualizing morality dominance we found in the context of modern societies and in relation to the rise of individualism. Nevertheless, it should also be noted that we found some significant and negative interaction effects between *culture* and *age* for several DS. The reasons for these cultural peculiarities need to be investigated further in the future. Moreover, our results regarding the *residential mobility* variable are for DS No. 7 and 8 consistent with existing theory and findings from other studies: increased residential mobility is associated with an increased choice of the individualizing option in moral dilemma scenarios. Frequent interactions beyond established group ties require appropriately calibrated moral minds that prioritize individualizing morality over binding morality (Henrich, 2020). Eventually, we also found significant effects for the variable *gender* in 4 out of 9 DS. Gender differences in morality are already known (Atari et al., 2020a). With regard to the influence of the *gender* variable, we would though like to hold back at this point and await further, future results before a coherent explanatory logic can be presented that is consistent with our findings from the current study, the direction of the results and a above all a (cross-)cultural perspective.

Third, our results presented above suggest that *moral intuitions* captured via MaC-DRS may differ from rather *conscious moral decisions* made in the context of dilemma scenarios

that juxtapose individual actions related to different moral domains. This is consistent with the argument stated in the introduction, and which we derived largely from Haidt's (2001) intuitionist model. Also, other studies in the fields of psychology and philosophy support that more intuitive moral processes must not necessarily come to the same evaluative and judgmental results as rather deliberate moral processes. Lisa Tessman (2014) states in this regard:

> "Usually, automatic intuition and controlled reasoning work smoothly together. But given that intuitive and reasoning processes can be triggered by different stimuli, are underwritten by different kinds of affective responses, and involve different neural mechanisms, there is no reason to expect that an excellent controlled reasoning process and an excellent automatic, intuitive process would always yield the same verdicts" (p. 58).

As we have seen, moral intuitions can play a role in primarily deliberate processes involved in solving moral dilemmas with mutually exclusive choices (see: DS No. 6, 7, and 9). However, intuitions are not necessarily a dominant source of influence in these situations. Because dilemmas require reason, and perhaps cost/benefit calculus to resolve competing interests, processes distinct from intuitions but more akin to Type 2 processes of deliberate moral cognition may take over in people's moral mind to guide choice and subsequent action when unconscious, quick, and effortless routine solutions are not available or feasible. In 6 out of 9 dilemma scenarios, none of the moral intuitions that we measured with MaC-DRS played a significant role and the choices were accordingly guided by processes other than mere intuitive drives. Also, we found differences between intuitive moral tendencies and choices in the dilemma scenarios. This can be exemplarily illustrated by reference to the GER-sample. Comparing the four study groups, this sample had the lowest MaC-DRS values in the context of family deviance relevance. In the dilemma scenario that contrasts family and property as mutually exclusive options, it is however the GER-sample that tends towards the family choice, whereas the other groups tend towards the property choice. Consequently, general tendencies of moral intuition, and more situational and action specific conscious moral choice may differ, and this is likely due to contextualization given by the scenario, calculus needed to resolve competing interests, and the inclusion of further information within the process of choice to resolve the dilemma. What should not be ignored, however, despite the differences mentioned, is the fact that both our findings from the context of moral intuitions and the findings from the context of deliberate moral cognition as a whole point to the cross-cultural predominance of individualizing morality.

By highlighting intuitive and deliberate moral cognition we build on the tradition of dual process models (Haidt, 2001; Tessman, 2014; Tutić, 2023). However, although we refer here to the dual-process tradition of moral cognition, we do not want to ignore the fact that neurological as well as psychological studies argue for a greater emphasis on more complex and dynamic models in the context of the human moral mind. Not only is there a difference between hypothetical dilemma scenarios, such as those presented here, and real word moral dilemmas, but there are also a large number of different brain areas involved in the respective process of moral evaluation and decision making (Van Bavel et al., 2015). Therefore, the dual-process model (intuition vs. deliberate reasoning) could be too simplistic and also fail to adequately account for dynamic updates, as in the sense of Bayesian models. There is also a need for more clarity as to which class of moral judgments (evaluations, norm judgments, wrongness judgments, blame judgments) is produced by which process (intuitive / deliberate) or by which combination of cognitive processes (Malle, 2021). Our remarks on deliberate moral cognition in moral binding/individualizing dilemmas and intuitive moral evaluations captured via MaC-DRS should be seen against this background. Nevertheless, we will only touch on the dual-process debates here, as a detailed treatment of this complex of topics would go beyond the scope of the present research focus. All in all, the results of our dilemma scenarios suggest that we can indeed observe a similar trend — i.e., predominant choice of *individualizing morality* — as in the MaC-DRS results across cultures. However, it needs to be highlighted that tendencies driven by processes of either intuition or more deliberate moral cognition may also differ in parts and do not necessarily lead to the same results. Ellemers and colleagues (2019) note in this regard: it "seems to be that it is the *interplay* between deliberate thinking and intuitive knowing that shapes moral guidelines" (p. 335). Following this idea, we have found further evidence based on the dilemma scenarios that suggests a predominance of individualizing morality in the four cultural groups studied.

Fourth, we would like to point out limitations of our dilemma scenario study that we believe may also be relevant for future research. As already mentioned, we worked with a cross-sectional design. However, longitudinal studies are far more meaningful and necessary to uncover real trends. This also applies to a deeper investigation of the influence of age on moral decisions in the context of individualizing and binding dilemmas. A developmental and life course perspective on decisions in moral dilemmas will help to develop a more comprehensive understanding of how our moral mind is shaped over the course of life. This perspective will broaden our view onto a person's stable, unchanging moral principles and the principles that change throughout the course of a lifetime and the (sociocultural) experiences gathered.

Furthermore, we did not examine the *process* of decision making and thinking itself, but rather the outcome of this process, i.e., the respective choice in the dilemma scenarios, in a quantitative manner. Qualitative research, research that comprises time measures, and research on cognitive load in moral decision making could contribute to a better understanding of the process of moral reasoning and would be desirable against the background of the binding and individualizing dilemmas tested — especially from a cross-cultural perspective. The dilemma scenarios are also characterized by the fact that they are of course only hypothetical in nature. This fact must always be taken into account when interpreting the results presented here. Finally, in the context of the limitations, it should nevertheless be emphasized that our design has in most part good external validity, but less strong internal validity. In this context, future experiments (Fischer & Formann, 2007; Oehlert, 2010) would be both important and desirable to gain insights into moral dilemmas and especially moral behavior in addition to the findings from the results presented here. We did not conduct behavioral measurements, which is probably the most important limitation of our cross-cultural study. Future research examining moral dilemmas in the context of actual behavior would allow us to bridge the gap between survey research and behavior prediction in the realm of pro-sociality (Cohn et al., 2019; Ellemers et al., 2019; Bjørnskov, 2021). However, this applies not only to the dilemma situations we focus on in this section, but also to the intuitive tendencies we discussed above. Studies that associate actual behavior with MaC-DRS would certainly give us much deeper insights into the human moral mind and the individual social and societal outcomes brought about by people's actions. Furthermore, we would like to mention one more addition: If we assume that each moral domain comprises a domain-specific set of deviant and conforming actions, we do not yet know the *severity* of the respective conform/deviant act. As far the severity of an act in the context of a particular moral domain is concerned, we used the example earlier in the theoretical chapter of someone stealing a piece of chocolate from someone, and another person stealing someone else's car. Both are offenses against the moral domain of property, but are they equal in *severity*? We believe that they are not the same in terms of severity and that each moral domain includes a set of domain-specific actions that come with different levels of (conformity/deviance) severity. So, we may ask for example, whether (specific) deviance in relation to the protection of the family is to be equated with (specific) conformity in relation to the protection of the property of others in terms of severity (see: DS No. 1)? This cannot be assumed a priori, and so an element of uncertainty remains in our DS results. In other words: Only when we know the domain-specific hierarchies of deviance and conformity of individual actions, across cultures, will we be able to develop more targeted dilemma scenarios that compare similar actions in

terms of their severity in different moral domains. We believe that these notions about the severity of acts of moral conformity/deviance are far reaching and certainly a task for future theorizing and research.

      In conclusion, regardless of whether we look at relevance valuations based on moral intuitions or analyze deliberate decisions in dilemma scenarios, and leaving aside the fact that deliberative and intuitive processes may yield different outcomes, one fact becomes apparent: **In the four heterogeneous cultures studied, *individualizing morality* appears to surpass *binding morality* in its significance for the respective moral system**. In the following, we will use the data from the *Moral Deviance Factorial Survey* (MDFS) to supplement our previous insights with further cross-cultural findings on moral relevance and judgment, as well as shame and guilt attribution in situations where moral deviations are at the heart of the matter. In particular, we will address questions of moral particularism and moral impartiality. We will therefore sharpen our focus again to explore further aspects of the question of which moral system guides cooperation in different cultures.

Chapter 6: Investigations of the Human Moral Mind IV

6.1. Findings from the Moral Deviance Factorial Survey (MDFS)

Imagine that someone is harmed because another person steals from them, breaks a promise to them, or betrays them. In other words, someone is harmed because of specific acts of moral deviance. How would you react to such acts? Would you say that the act is *relevant to* your sense of morality when a stranger is harmed, or is the behavior perhaps even more *wrongful* when a friend is harmed? Moreover, does the act of moral misconduct evoke the same extent of relevance attribution and right/wrong judgment in you? Regardless of relevance and judgment, moral deviance can elicit intense emotions such as anger in the observer of the deviant action (Haidt, 2003). Other moral emotions such as shame and guilt are more likely to be imposed on the perpetrator, as we experience these feelings as aversive emotions that can motivate us to refrain from deviant actions in the future (Tangney et al., 2007). Would you attribute more shame and guilt if a family member were harmed by a morally deviant act than if a stranger were harmed, or are you impartial in this regard?

We wonder whether it matters *who* is harmed by morally deviant acts: Are the valuations of moral transgression the same or different across cultures when a *stranger*, a *family member*, or an *in-group member* is harmed by the act deviance? In the following we will examine considerations of moral *particularism* and *impartiality* to further inform our investigations into which moral system guides cooperation in different cultures. In addition, we will also investigate whether specific acts of moral deviance evoke the same extent of deviance relevance and deviance judgment. To do this and to conclude our current cross-cultural investigations of the human moral mind, we will take a complementary look at the data from the Factorial Survey. The **M**oral **D**eviance **F**actorial **S**urvey (**MDFS**)[127] we designed offers a further opportunity to pursue our research question and supplement our previous findings. This time, however, we will work with contextualized data to an even greater extent than we have done with the dilemma scenarios, as our analyses are based on a set of 168 different vignettes, which we will now discuss in more detail.

The vignette scenarios of the MDFS describe (specific) deviant actions by one person that harms another. Once again, **moral deviance** is the focus of our analyses and not moral conformity. Our theoretical perspective understands moral deviance as the failure of

---

[127] We will henceforth use the abbreviation **MDFS** for the Moral Deviance Factorial Survey.

cooperation, which results in costs and ultimately harm for at least one party of an initially cooperative endeavor. The Factorial Survey vignette design allows us to analyze the **dimensions of the vignettes** themselves. This puts us in the perspective of investigating (quasi) causal relations, as we can investigate the consequences of the variation of vignette dimensions (Hughes & Huby, 2004; Auspurg et al., 2009; Atzmüller & Steiner, 2010; Knutson et al., 2010; Auspurg & Hinz, 2014; 2015; Skilling & Stylianides, 2020).

Although already covered in the theoretical part of this thesis, let us briefly and succinctly recall the Moral Deviance Factorial Survey vignette design. The **vignettes** can be understood as scenarios that resemble items in a questionnaire. What is special about the vignettes, however, is that they consist of a kind of modular system and allow for variations in content. These variations can then be examined by the researcher. The design we have developed comprises various **dimensions**, which in turn have a certain number of **levels**. ***Table 43***, to be found below, shows the dimensions and the corresponding levels in a formalized depiction. We have, of course, thought about adapting the vignettes to the cultural contexts we are investigating, without, though, abolishing the basic structure (ceteris paribus) of the vignettes themselves.[128]

The vignette scenarios describe that a person performs a morally deviant action in one of several moral domains (dimension: ***domain of deviance***). We vary the corresponding moral domain in our vignettes: The Factorial Survey comprises a total of 7 different moral domains, as can be seen in ***Table 43***. It should be noted at this point that, similar to the dilemma scenarios and in contrast to MaC-DRS, only *specific* actions are captured on the basis of the vignettes. In other words: The scenarios described in the vignettes, portray only *single specific act of deviance* per domain. Consequently, the corresponding investigations are about the reaction to very specific types of deviance, but not about general tendencies. Furthermore, we vary the gender (dimension: ***gender*** — male/female) of the deviant person in half of all vignettes. The cultural contextualization of the gender dimension now takes place via the use of names that are popular within the respective cultural entity. For example, we used the names Sabine (Germany), Yui (Japan), Merna (Egypt) and Charlotte (United States) in the vignettes for the gender dimension to indicate that the deviant action described in the vignettes was performed by a woman. Our design further encompasses the dimension ***reputational damage***: a total of

---

four levels describe whether the deviant act leads to reputational damage and who is affected by the loss of reputation. If reputational damage is present in the vignettes, the deviant act can

Table 43: The MDFS building blocks (dimensions) and their expressions

Vignette	Expressions Levels	Expressions
Dimension: 1 — *Gender*	2	• D1a = Male/Name/His; • D1b = Female/Name/Her; Name Germany: D1a = Michael; D1b = Sabine; Name Japan: D1a = Haruto; D1b = Yui; Name Egypt: D1a = Ramy; D1b = Merna; Name USA: D1a = Oliver; D1b = Charlotte
Dimension: 2 — *Reputational damage*	4	• D2a = not present; • D2b = Through (her/his) behavior, D1 harms the reputation of (her/his) family; • D2c = Through (her/his) behavior, D1 harms the reputation of (her/his) friends; • D2d = Through (her/his) behavior, D1 harms her/his own reputation;
Dimension: 3 — *Social relation*	3	• D3a = a stranger (D3a); a stranger (D3b); a stranger (D3c); a stranger (D3d); of a stranger (D3e); a stranger (her) (D3f); a stranger (D3g); • D3b = a family member (D3a); a family member (D3b); a family member (D3c); a family member (D3d); a family member (D3e); a family member (him) (D3f); a family member (D3g); • D3c = a friend (D3a); a friend (D3b); a friend (D3c); a friend (D3d); of a friend (D3e); a friend (D3f); a friend (D3g);
Dimension: 4 — *Domain of deviance* **(MaC + MFT)**	7	• D4a = *Property*: D1 steals something from...; • D4b = *Fairness*: When it came to sharing something with ... (D3), D1 kept the best for D1 (himself/herself); • D4c = *Heroism*: When it came to protecting ... (D3) from danger, D1 behaved in a cowardly manner. • D4d = *Deference*: D1 behaves disrespectfully and publicly insults ... (D3); • D4e = *Reciprocity*: D1 behaves ungratefully and does not return the favor of a ... (D3); • D4f = *Loyalty*: D1 betrays ... (D3) and publicly undermines them; • D4g = *Trustworthiness*: D1 breaks his/her promise to ... (D3);

Note: *D* stands for dimension. The *numbers* (e.g. D1) and *letters* (e.g. D1b) are used to formally classify the building blocks of the vignettes.

either damage the reputation of the family, the reputation of the in-group, or the reputation of the person committing the moral offense in the vignette scenario. Finally, the Moral Deviance Factorial Survey allows us to vary another contextual element of the vignette scenarios, on which we will focus in the following analyses: We also vary the ***social relation*** to the person harmed by the act of moral deviance in the vignettes. The dimension social relation(ship)

comprises a total of 3 characteristics. As mentioned, these characteristics/levels are to be regarded as building blocks of the vignette scenario and vary whether the person harmed by the deviant act in the scenario is a *stranger*, a *family* member or an *in-group* member (a friend) (for a similar research approach, see: McKee et al., 2024).

Taken together, our MDFS vignette design can therefore be understood as a modular system. This modular system allows us to examine a large number of different scenarios with the dimension *gender* and two levels, the dimension *reputational damage* and four levels, the dimension *social relation* and three levels, and the dimension *domain of deviance* and 7 levels. Comprising all possible variations of the dimension levels, the **vignette universe** consists of (2 x 4 x 3 x 7 =) 168 different vignettes. As an example, we would like to present a few of these vignettes. The examples are taken from the English version of the Moral Deviance Factorial Survey, illustrate the possibility of varying vignette levels and deal with all 7 different levels of the dimension *domain of deviance* (**Table 44**).

Table 44: The Factorial Survey: Examples from the vignette universe (N = 168)

Dimensions and expressions (The building blocks of the scenarios)	Vignette examples
- Gender: *Male* - Reputational damage: *Family* - Social Relation: *Stranger* - Domain of deviance: ***Property***	Oliver steals something from a stranger. Through his behavior, Oliver damages the reputation of his family.
- Gender: *Male* - Reputational damage: *Not present* - Social Relation: *Family* - Domain of deviance: ***Fairness***	When it came to sharing something with a family member, Oliver kept the best for himself.
- Gender: *Female* - Reputational damage: *Not present* - Social Relation: *Stranger* Domain of deviance: ***Heroism***	When it came to protecting a stranger from danger, Charlotte behaved in a cowardly manner.
- Gender: *Male* - Reputational damage: *Own Reputation* - Social Relation: *Family* - Domain of deviance: ***Deference***	Oliver behaves disrespectfully and publicly insults someone of his own family. Through his behavior Oliver damages his own reputation.
- Gender: *Male* - Reputational damage: *In-Group (friends)* - Social Relation: *Family* - Domain of deviance: ***Reciprocity***	Oliver behaves ungratefully and does not return the favor of a family member. Through his behavior Oliver damages the reputation of his friends.
- Gender: *Female* - Reputational damage: *Family* - Social Relation: *In-Group (friend)* - Domain of deviance: ***Loyalty***	Charlotte betrays a friend and publicly undermines them. Through her behavior, Charlotte damages the reputation of her family.
- Gender: *Female* - Reputational damage: *Own Reputation* - Social Relation: *In-Group (friend)* - Domain of deviance: ***Trustworthiness***	Charlotte breaks her promise to a friend. Through her behavior, Charlotte damages her own reputation.

Note: All dimensions and their expressions together result in a so-called vignette universe of N = 168 different vignettes. The vignettes were randomly assigned to the respondents. Each respondent received a total of four different vignettes in the online survey (Study 3). For each vignette, the respondents were asked to respond to the extent of deviance *relevance*, deviance *judgment*, deviance *shame* attribution and deviance *guilt* attribution.

With the **MDFS**, we thus have a measurement instrument that allows us to examine certain *contextual* factors of the vignette *scenarios* that can affect the valuation of the deviant act in question. This leads us to the **dependent variables**. We integrated four different aspects of responses to the perception of moral deviance. As part of our data collection, we randomly assigned four different vignettes to each participant in our cross-cultural study (Study 3). Respondents were then asked to provide four responses on a 7-point scale to each vignette. Firstly, we asked respondents to indicate the extent to which the act of deviance was (ir-)relevant to their individual sense of morality (item: "Was the behavior by the person in the story relevant to your sense of morality?", **deviance relevance variable**). This question shall elucidate whether individuals attach a certain degree of relevance to various acts of (specific) moral breaches. Respondents were also asked to indicate the extent to which they considered the deviant action to be right or wrong (item: "In your opinion, did the person in the story behave morally right or morally wrong", **deviance judgment variable**) (Malle, 2021; Atari et al., 2022a). This item thus aims to capture the degree of individual moral judgment in relation to specific acts of moral deviance. In addition, we asked respondents to indicate the extent to which the person committing the moral offense in the scenarios should feel ashamed and guilty (item: "To what extent should the person in the story feel shame about his or her behavior?", **shame attribution variable**; item: "To what extent should the person in the story feel guilty about their behavior?", **deviance guilt attribution variable**) (Tangney et al., 2007). [129] Needless to say, much has been done in cross-cultural research in the last two decades (Henrich et al., 2010a; Apicella et al., 2020). This certainly also applies to research on moral emotions. Nevertheless, research on moral emotions is yet dominated by studies based on Western samples (Wong & Tsai, 2007). By including the variables of shame and guilt attribution, our study, which comprises in addition to two WEIRD samples also Japan and Egypt, makes also a small contribution to the field of cross-cultural research on moral emotions.

We will use the Moral Deviance Factorial Survey (MDFS) to supplement the insights we have already gained from the investigations on MaC-DRS and the binding/individualizing dilemma scenarios. At this point, we would like to emphasize in advance that we will by no means exhaust the full potential of the MDFS and the data we have collected with this

---

[129] One could ask where the evaluations collected via the Moral Deviance Factorial Survey are more likely to be placed: on the side of *moral intuitions* or on the side of *deliberate moral cognition*? We believe that the latter is the case. Since the vignettes of the Factorial Survey are highly contextualized and we ask four items per vignette, we assume that primarily deliberate processes guide the responses on the relevance, judgment, shame and guilt variables. This does not mean, as we have already emphasized several times, that moral intuitions do not play a role in this context, but we do believe that deliberate processes are primarily involved in the response behavior due to the rather rich scenic presentation of the vignettes and the answering of four items to each vignette.

instrument in the present study — there are a large number of findings and analysis opportunities that cannot be covered here due to their scope. Rather, the aim will be to examine supplementary, highly contextualized findings on the valuation of specific moral deviance across different domains and thus add another facet to further approach our **overarching research question**. In addition, we extend our previous findings by measuring moral judgment and the attribution of the moral emotions shame and guilt in the context of moral deviance. It should also be mentioned that we measure three moral domains in the MDFS, i.e., *property*, *fairness* and *trustworthiness*, which we assign to the superordinate concept of individualizing morality. Furthermore, with *deference* and *loyalty* deviance we only measure two binding moral domains. This distribution is based on the fact that we indirectly include the family and in-group domains in the context of the reputational damage dimension and the social relation dimension, but cannot record these as independent domains due to the design itself. In addition to the individualizing and binding domains, the moral domains *reciprocity* and *heroism* are also included in the Factorial Survey.

We further aim to show that the *relevance* of moral deviance and the *judgment* of moral deviance are not one and the same. We have proposed a focus on moral *deviance relevance* in the theoretical part of this thesis (see: ***Chapter 1***) and then translated this into two practical measurement instruments, i.e., MaC-DRS and MDFS. In fact, one of the differences between MaC-DRS and the Moral Deviance Factorial Survey is that the former measures general and decontextualized tendencies of moral intuitions, while the latter measures specific types of deviance and more contextualized, deliberate tendencies. However, the vignette design allows us to test whether the extent to which deviant behavior is evaluated in terms of **relevance** and right/wrong (**judgment**) is the same or different. Since we argue that a deviant action can be judged as wrong, but does not necessarily need to be perceived as morally relevant to the same extent, we propose that both concepts can differ in their scope of evaluation, although the evaluation is caused by the same stimulus. In order to further establish the investigation of moral *relevance*, it is therefore our goal to empirically investigate across cultures whether deviance relevance and deviance judgment exhibit the same valuation or tend to diverge in their extent. The corresponding **relevance/judgment hypothesis** is as follows:

> **Moral deviance relevance/judgment hypothesis**: We predict that the extent of the *relevance* of specific acts of moral deviance and the extent of the *judgment* about specific acts of moral deviance do not (necessarily) coincide.

In the following, we will provide initial insights into the research possibilities of the Moral Deviance Factorial Survey. As already mentioned, we will not be able to exploit the full potential of the data collected with this instrument, as this would go far beyond the scope of this thesis. Therefore, we focus on the following variables: *deviance relevance*, *deviance judgment*, *deviance shame attribution* and *deviance guilt attribution* in the context of the *7 different moral domains* captured by the vignette design (dimension: *domain of deviance*). Above all, however, we will analyze the **social relation(ship) dimension**, which serves as the main focus of the following analyses. We single out the social relation dimension for two reasons: first, we want to shed light on one aspect of the research possibilities of MDFS and, second, as already announced, we want to approach the question of whether it makes a difference *who* is harmed by (specific) acts of moral deviance. The idea behind this investigation is to empirically examine whether the cultural entities we study differ in terms of moral *particularism* and *impartiality* (Lang et a., 2019; Enke, 2019; Waytz et al., 2019; Kirkland et al., 2023). Particularism describes a socially narrow morality that applies primarily to particular groups or persons and thus does not unfold its pro-social effect universally or impartially (Henrich, 2020). In its most extreme form, moral particularism can be associated with parochial altruism (Choi & Bowles, 2007; McDonald et al., 2012; Rusch, 2014; Aldering & Böhm, 2020; De Dreu et al., 2022). With regard to populations with intensive kinship institutions and correspondingly calibrated psychology on the one hand, and populations characterized by the detachment from kinship institutions and correspondingly differently calibrated psychology on the other, Henrich (2020) discusses research results on particularism and impartiality and comes to the following conclusion: "every population breaks impartial rules; but, it turns out that some populations break such rules more than others" (p. 42). Evidence suggests that these population differences in impartiality and particularism are anything but random. Particularism is, inter alia, associated with low market integration, low relational mobility, historical subsistence styles as paddy rice farming, and above all a strong (historical) reliance on kinship institutions (Henrich et al., 2005; 2010b; Thomson et al., 2018; Schulz et al., 2019; Henrich, 2020; Talhelm, 2022). Typically, so-called WEIRD societies are characterized by properties that run counter to the triggers of particularistic tendencies just mentioned. WEIRD societies are therefore associated with moral impartiality (Enke, 2019; Henrich, 2020). Furthermore, groups imply boundaries of in-group and out-group, and the focal point of collectivism and the interdependent self-construal is the group. By contrast, independent self-construal and individualism focus on the individual and are not mainly concerned with the group (Hogg et al., 2004; Reicher et al., 2010; Cross et al., 2011; Turner &

Reynolds, 2012; Żemojtel-Piotrowska & Piotrowski, 2023; Kitayama & Salvador, 2024). In this regard and by reference to collectivism, cross-cultural psychologist Harry Triandis noted in 2001 the following: "In many collectivist cultures, morality consists of doing what the in-group expects. When interacting with the out-group, it is "moral" to exploit and deceive. In other words, morality is not applicable to all but only to some members of one's social environment" (p. 917). In addition to the characteristics already mentioned, WEIRD societies are characterized by cultural-level individualism and independence in individual-level self-construal (Henrich, 2020). To build our hypothesis we draw on the literature just cited. Furthermore, as also prominently emphasized throughout this thesis, we theorize a systematic relationship between ways of selfhood, corresponding cultural-level collectivism-individualism and the endorsement of binding and individualizing morality (Haidt, 2008). What is more, we theorize that binding morality is linked to particularism and individualizing morality is liked to impartiality. Against this background, we have derived the following **hypothesis** before data collection:

> **Initial impartiality/particularism hypothesis**: We hypothesize that cultures favoring *binding* moral domains (family, deference, and in-group) over *individualizing* moral domains (fairness, trustworthiness, and property) tend to rate moral deviance that harms a stranger as less severe (relevance, judgment, shame, and guilt) than deviance towards a member of one's in-group or family (and vice versa).

In the light of our research conducted using MaC-DRS and the moral dilemma scenarios it needs to be mentioned, that we cannot pursue this hypothesis in exactly the way that we have planned to do before data collection. The reason is straight forward: none of the four cultural entities that we examine favors binding moral domains over individualizing moral domains as indicated by the MaC-DRS findings (*Chapter 4* and *5*). Actually, as elaborated on extensively in the MaC-DRS section, it is individualizing morality that is intuitively relevant across the two WEIRD samples of Germany and the Unted States of America, but also across the Japanese and Egyptian samples. The findings from the moral dilemma scenarios point to the same direction. As mentioned in detail elsewhere, we have placed this circumstance in the context of modernization processes (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019; Henrich, 2020, Kaasa & Minkov, 2020). Nevertheless, we would like to remind the reader that the MaC-DRS analyses have shown that the moral domains of family, in-group and deference are yet intuitively more relevant in Japan and Egypt than in Germany and partly more relevant than in the USA. Japan is marked by historical paddy rice farming, low relational mobility and collectivism as well as interdependence in selfhood (Kitayama & Salvador, 2024). The case is different though for Egypt which is promoting interdependent and independent aspects of self-

construal (San Martin et al., 2018), and is yet largely shaped by collectivism (Minkov & Kaasa, 2022). Furthermore, and in contrast to the WEIRD samples in our study, the kinship institutions in Japan and Egypt have probably not been culturally weakened over a long period of time and are therefore likely to play a greater role than in the German and US cultural contexts. The considerations regarding kinship institutions apply particularly to Egypt (Cole, 2003; Reilly, 2013; Schulz et al., 2019; Enke, 2019; Henrich, 2020). Given the insights from the studies mentioned and our findings from the previous investigations, we modify our main hypothesis and expect the following:

> **Modified impartiality/particularism hypothesis**: We hypothesize that the JP- and EG-cultural samples tend towards moral *particularism*, i.e., we expect a tendency to rate moral deviance that harms a stranger as less severe (relevance, judgment, shame, and guilt) as compared to deviance towards a member of the in-group or family. In contrast, we expect that the GER- and US-cultural samples tend towards *impartiality*, i.e., we expect a tendency to rate moral deviance that harms a stranger as equally severe (relevance, judgment, shame, and guilt) as deviance towards a member of the in-group or family.

We consider *impartial* moral deviance evaluations to be given if we find no significant difference in the valuation of deviance between the levels of the social relationship dimension, i.e., stranger vs. family member respectively stranger vs. in-group (friend). If, by contrast, the evaluations of moral deviance are significantly weaker when a stranger becomes the victim of failed cooperation, then we have an indication of moral *particularism*. Altogether, by taking up the social relationship dimension, we are not only introducing ways to use the MDFS instrument, but we will also investigate another facet of the human moral mind — i.e., impartiality and particularism — from a cross-cultural perspective.

In the following, we will *first* briefly discuss descriptive and correlative analyses. *Secondly,* we will touch on the OLS model of the analyses. *Thirdly*, we turn to the main part of the empirical analyses and will look at each of the 7 levels of the *domain of deviance* dimension. In this context, we examine the four dependent variables *relevance*, *judgment*, *shame* and *guilt*, and estimate respective margins (average marginal effects) for each of the cultural samples. We will initially provide graphic evidence and supplement these analyses with pairwise comparisons between the samples testing for significant differences. Furthermore, the findings of the MDFS dimension *social relation* are then extracted from the OLS models and displayed in tabular form. Within each of the seven moral domains examined, we will address evidence in the light of the **moral deviance relevance/judgment hypothesis** and the (modified) **impartiality/particularism hypothesis**. The empirical analysis is followed by a comprehensive discussion. After the present chapter, we turn to the last section of this thesis

and consider the overall contribution of our theoretical considerations and empirical investigations.

6.2. Descriptive and Correlative Insights

We will carry out the MDFS analyses again on the basis of the *adjusted sample* from Study 3 (N = 2,360). Since our study asked each respondent to answer four items per vignette, we obtain a total of 4 x 2,360 = 9,440 responses to the vignette universe (N = 168).[130] Ignoring the different levels of the dimensions for a first overview, the dependent variables show an interesting distribution, which can be seen in *Table 45* below.

The descriptive findings from *Table 45* provide initial insights. First of foremost, the deviant actions in the vignettes are clearly identified as such. As far as relevance, judgment, shame and guilt are concerned, a large proportion of the respondents' answers are above the midpoint of the 7-point response format of the respective items. Consequently, a large proportion of respondents evaluates the actions as relatively relevant and wrong. The respondents are also largely in favor of attributing guilt and shame to the deviant actor — these responses reveal an emotional appeal to refrain from the act in question.

Another fact becomes apparent when we look at deviance relevance and pay more attention to the Egyptian sample in *Table 45*. We already had difficulties with the concept of deviance relevance and the EG-sample in our MaC-DRS analyses. Both the EFA findings in *Chapter 3* and the MaC-DRS analyses in *Chapter 4* indicated that, in addition to a notable response style effect, something else might be distorting the EG-sample analyses. This problem seems to be further qualified in the MDFS analyses. If we look at the lowest category of the relevance item in the Factorial Survey, i.e., "*Extremely irrelevant*", we find a total of 1,022 cases that indicated this response category. Of these 1,022 cases, 8.12% are attributable to the GER-sample, 7.14% to the JP-sample, 15.6% to the US-sample and finally 69.58% to the Egyptian sample. If we continue to focus only on the Egyptian sample, we can also see that the "*Extremely irrelevant*" category accounts for most of the responses in the EG-sample, with a total of 30.54%. What raises questions, however, is the fact that the opposite category, i.e., "*Extremely relevant*", accounts for the second most cases with a total of 28.65%.[131]

---

[130] The analysis of the Factorial Survey is accompanied by a data transformation. Details on this transformation can be found in the following publication: (Auspurg & Hinz, 2015).

[131] A total of 22% of the EG-sample stated "rather relevant" or "very relevant" for the relevance item, whereas only 12.24% stated "rather irrelevant" or "very irrelevant". "Neither irrelevant nor relevant" accounted for 6.57% of the responses in the EG-sample.

Table 45: Descriptive insights — Specific moral deviance distribution across relevance, judgment, shame and guilt

	United States	Japan	Germany	Egypt	Total
Deviance *Relevance*					
Extremely irrelevant	155	73	83	711	1,022
Very irrelevant	172	88	92	159	511
Somewhat irrelevant	143	261	171	126	701
Neither nor	293	524	350	153	1,320
Somewhat relevant	504	534	752	199	1,989
Very relevant	550	437	732	313	2,032
Extremely relevant	459	255	484	667	1,865
Total	2,276	2,172	2,664	2,328	9,440
Deviance *Judgment*					
Right behavior	41	16	17	16	90
Rather right behavior	38	37	30	44	149
Somewhat right behavior	70	89	69	60	288
Neither nor	309	506	334	111	1,260
Somewhat wrong behavior	482	411	551	174	1,618
Rather wrong behavior	509	421	736	322	1,988
Wrong behavior	827	692	927	1,601	4,047
Total	2,276	2,172	2,664	2,328	9,440
Deviance *Shame* (attribution)					
Not ashamed at all	35	15	13	16	79
Not ashamed	53	36	50	40	179
Rather not ashamed	80	97	106	61	344
Neither nor	297	451	348	104	1,200
Rather ashamed	564	549	785	216	2,114
Very ashamed	572	544	752	445	2,313
Extremely ashamed	675	480	610	1,446	3,211
Total	2,276	2,172	2,664	2,328	9,440
Deviance *Guilt* (attribution)					
Not guilty at all	40	18	21	17	96
Not guilty	54	42	49	43	188
Rather not guilty	74	96	96	49	315
Neither nor	273	447	392	107	1,219
Rather guilty	582	557	829	292	2,260
Very guilty	584	561	682	485	2,312
Extremely guilty	669	451	595	1,335	3,050
Total	2,276	2,172	2,664	2,328	9,440

Furthermore, a look at the distribution of the data for the judgment, shame and guilt items reveals that the polarization mentioned only exists in the EG-sample and only for deviance relevance. The data situation therefore raises questions as to whether there were problems in the EG-sample with regard to understanding the concept of deviance relevance, whether there is a strong polarization of moral relevance itself, or whether, for example, a certain proportion of respondents are also unwilling to answer the corresponding deviance relevance item. On the whole, even before any in-depth analysis, the descriptive findings suggest that the aspect of moral deviance relevance in the context of Egypt should be examined

in more detail in future studies. This statement can be further substantiated by the correlative analyses presented below in *Table 46*.

Table 46: MDFS: correlations between relevance, judgment, shame, and guilt within the cultural samples

	Relevance	Judgment	Shame	Guilt
US-Sample				
Relevance	1.0000			
Judgment	0.3004	1.0000		
Shame	0.3055	0.7341	1.0000	
Guilt	0.3002	0.7678	0.8366	1.0000
JP-Sample				
Relevance	1.0000			
Judgment	0.4015	1.0000		
Shame	0.4330	0.8406	1.0000	
Guilt	0.4237	0.8094	0.9286	1.0000
GER-Sample				
Relevance	1.0000			
Judgment	0.4994	1.0000		
Shame	0.5008	0.7607	1.0000	
Guilt	0.4923	0.7474	0.8866	1.0000
EG-Sample				
Relevance	1.0000			
Judgment	-0.0838	1.0000		
Shame	-0.0694	0.8212	1.0000	
Guilt	-0.0809	0.7971	0.8596	1.0000

Several findings emerge from the correlative *within* sample analyses between specific deviance relevance, judgment, shame and guilt. Let us first stay with the EG-sample. *Table 46* shows that specific moral deviance *relevance* in the EG-sample is weakly and above all *negatively* correlated with specific deviance judgment, shame and guilt. We only encounter this negative correlation in the EG-sample. Although specific moral deviance relevance is not as strongly correlated across all samples as judgment, shame and guilt, it is consistently positive with the exception of the EG-sample. The correlations of relevance, judgment, shame and guilt are also significant *across* all samples. To further examine the correlations in the EG-sample, we created a specific deviance relevance variable that does *not* include the 711 Egyptian cases that indicated "*Extremely irrelevant*". Relevance correlates in this test sample (n = 1,617) with a value of 0.3267 for judgment, 0.3646 for shame, and 0.3725 for the guilt variable. The negative correlation of relevance with the other variables in the "base sample" is therefore caused by the

proportion of cases in the EG-sample that are polarized towards the lower extreme of the response format. The descriptive and correlative analyses of the EG-sample do not allow any definitive conclusions to be drawn and call for further investigations. Nevertheless, they at least indicate once again that doubts about the robustness of our EG-sample in the context of moral deviance relevance are certainly justified.

Going beyond the EG-sample, the correlative findings across all four groups reveal initial insights in regard to the **moral deviance relevance/judgment hypothesis**: The correlative analyses indicate that deviance relevance and deviance judgment are in fact not one and the same aspect of morality. Moreover, the analyses show that, although all four variables are significantly correlated with each other, deviance judgment is primarily correlated with shame and guilt attribution. Additionally, considerable fluctuations in the correlation of relevance with judgment, shame and guilt are observable across the cultural samples. Furthermore, although there are also cross-cultural fluctuations in the correlation of the other constructs, these are nowhere near as strong as those relating to the variable specific deviance relevance. The correlative analyses already cast a shadow of further questions: Which aspects of morality exactly are measured with deviance relevance? Also, the extent to which deviance relevance could complement the measurement of, for example, moral judgments in various aspects such as correlation with other concepts and prediction of behavior must be determined in future studies. In our view, this opens up a wide field for future work.

Taken together, the descriptive and correlative findings indicate that the deviant actions portrayed in the vignettes are indeed regarded as such *across* cultures. This can initially be interpreted as a preliminary swing towards supporting the universalism thesis of the moral mind. Furthermore, the results once again indicate that the robustness of the EG-sample should at least be doubted in the context of moral deviance relevance. This calls for further research in order to gain e.g. a better understanding of the "*Extremely irrelevant* vs. *Extremely relevant*" polarization of the Egyptian sample in this study. Finally, initial correlative findings suggest that specific deviance relevance and specific deviance judgment are indeed correlated but distinct aspects of our morality. We thus see initial supporting evidence for our **moral deviance relevance/judgment hypothesis**. Following these introductory empirical insights, we will now briefly discuss the models we used to analyze the data.

6.3. Analyzing MDFS: Comments on the Baseline OLS Regression Model and Multiple Testing

In addition to *culture* and the dimension *domain of deviance*, we are primarily interested in the effect of the *social relation*(*ship*) dimension (levels: stranger; family; in-group/friend), and also the variables *deviance relevance* and *deviance judgment*. As indicated before, we cannot address the entire Moral Deviance Factorial Survey of our study with all dimensions in detail and with the necessary precision in the present paper, as this would be an excessive undertaking that would clearly go beyond the scope of this thesis. On this basis, we have decided not to examine the MDFS dimensions *gender* and *reputational damage* in this study.

The MDFS vignettes deal with *specific types of deviance* from different moral domains. We hold the theoretical position that the findings from the Factorial Survey, in contrast to the MaC-DRS analyses, cannot be compared with each other across the various moral domains and that such a comparison would not be meaningful either. The background to this argument is that the vignettes portray *specific*, domain-related behavior and we simply cannot assume a priori that the described behaviors are the same in their (deviance) *severity* across the domains. Accordingly, within-sample deviance relevance rankings (e.g.), as in the MaC-DRS analyses, are therefore neither intended nor considered expedient. Instead, we will analyze and discuss *contextualized* and, above all, *specific* deviance within 7 moral domains across cultures with regard to *relevance*, *judgment*, *shame attribution*, and *guilt attribution*.

It also follows from the above argument that the respective models for the 7 moral domains may differ from one another, as we are not aiming for a comparison across domains. Therefore, we will empirically derive a specific model of analyses adapted for each moral domain investigated.[132] By empirical derivation, we mean that we only retain (measured) variables that have proven to be significant in several rounds of preliminary analyses for the respective moral domain tested. This approach allows us to design our OLS analysis models to be more parsimonious and to reduce the complexity of the analyses to influential (measured) variables that carry the most weight. However, it is important to note that a constant model for the dependent variables is estimated for each moral domain and cultural group.

The **baseline model** for our analyses encompasses *deviance relevance*, *deviance judgment*, *deviance shame attribution* and *deviance guilt attribution* as dependent variables. In addition

---

[132] Please note that all covariates in the initial models were integrated into the models by us on the basis of theoretical considerations that associate them to some extent with morality and cross-cultural variations. Although we are mainly taking an empirical approach here to make our models more parsimonious, our approach is yet principally informed by theoretical considerations as well.

to the variable *cultur*e, the following variables are included in the baseline model: *years in school*, *level of religiosity*, *age*, *NARS*, *MRS*, *pathogen prevalence*, and the dimension *social relationship* (levels: strangers, family, in-group/friend) as a variable to test our hypothesis. The baseline model comprises interaction effects with *culture* for the latter variables. Further covariates are *residential mobility*, *place of living* (village vs. city), *gender* (female/male) and *denomination* as well as an interaction effect for *level of religiosity* and *denomination*. This baseline model forms the starting point for the MDFS analyses and will yet be adapted to each moral domain under investigation, as explained. On the basis of the domain-specific estimation model, we then carry out the respective analyses.

Since we examine 7 different moral domains for four dependent variables, this results in a total of 28 different OLS models for each of the four cultural entities under investigation (i.e., 112 OLS models in total). However, in order to focus on our research interests, we will only consider and discuss the effects of the variable *culture* and the MDFS *domain of deviance* and *social relation*(*ship*) dimensions in relation to the dependent variables.[133] All other variables and effects of the respective models are therefore not considered any further in the current study and are only included in the estimation of the culture-specific margins.[134]

With our models of analyses, we are in the realm of **multiple testing**. Although the Holm-Bonferroni correction is not as conservative as the Bonferroni method (Hemmerich, 2020), corresponding p-value adjustments are still restrictive when working with large test families. Our final OLS models include test families that comprise 27 to 35 individual tests and thus entail conservative p-value adjustments. In the context of the correction for alpha-error cumulation, we therefore consider it important to take into account potential losses of statistical significance. A procedure that takes into account both the *uncorrected results* and the *p-value adjusted findings* seems sensible to us in order to obtain initial indications of possible effects that could be suppressed by the strict correction and to still work statistically clean. We consider our approach to be all the more appropriate given that the Moral Deviance Factorial Survey is a newly developed research tool and that, in addition to our theoretical interest, we also want to demonstrate the applicability of this tool. For this reason, we have decided to refer to both the

---

[133] Since we will inspect interaction effects for family vs. stranger and in-group vs. stranger in the context of the Factorial Survey dimension social relationship per cultural sample on four dependent variables and over 7 moral domains, there will be 56 interaction effects per sample to be considered, resulting in a total of 224 interaction effects altogether.

[134] It becomes readily apparent that with our data collection based on the Moral Deviance Factorial Survey, we have tapped into a rich source of information for future studies. The data and the corresponding models used for analyses can be obtained from the author on request for further insight into the results not discussed here. In the **Appendix**, we have also described a more detailed derivation of our models of analysis. Moreover, out of methodological research interest, we also conducted analyses of *response style* effects in the context of MDFS. The corresponding results of these analyses can be found in the **Appendix**.

unadjusted and the Holm-Bonferroni *corrected* results in the context of the analyses of the *social relation* dimension. In the analyses of cross-cultural similarities and differences, we use the group-specific margins (AME; Wooldridge, 2016) and pairwise comparisons of the samples with each other. Since there are only 6 pairwise comparisons per sample and dependent variable, we do not see the danger of large test families and correspondingly restrictive p-value correction. Therefore, we consistently apply the Holm-Bonferroni correction for multiple testing in the context of these analyses. To obtain the adjusted p-values, we again used the online calculator from Hemmerich (2016). Based on what we have outlined, we now turn to the specific deviance models for the dependent variables and their analysis.

6.4.1. MDFS Analysis I: Property Deviance

The analysis of the baseline property OLS models reveals that the variables *level of religiosity*, *denomination*, *level of religiosity x denomination*,[135] *place of living*, and *gender* show no significant effect in all four models of the dependent variables.[136] Consequently, we remove these variables from our final model(s). Subsequently, we perform our analyses based on the adapted model for specific property deviance. ***Figure 23*** shows the margins (Average Marginal Effects, AMEs; Williams, 2012; Wooldridge, 2016) resulting from the estimation of our models for each of the four cultural groups as well as for the four dependent variables. For reasons of clarity, we have colored deviance relevance *blue*, deviance judgment *red*, deviance shame attribution *green*, and deviance guilt attribution *orange* in ***Figure 23*** and also used different symbols for the dependent variables (see: legend below the figure).

As can be seen in ***Figure 23***, a clear gap emerges particularly between specific deviance relevance and the other dependent variables. Hence, in the property models it is apparent that the extent to which moral deviance is *relevant* is not the same as the extent to which moral deviance is *judged* as wrong. A specific act of property deviance can therefore be judged as morally wrong, but this does not mean that it must also appear subjectively relevant to the same extent, which provides initial supporting evidence for our **relevance/judgment hypothesis**. The same seems to apply to the attribution of shame and guilt, since the imposition of moral emotions can differ from the extent of deviance relevance and, to some part, from the extent of

---

[135] We indicate interaction effects by using an *x* between two variables.

[136] To reduce complexity, we relied on OLS models that use the US-sample as the reference category for the analyses to determine the variables that exert a significant influence in the context of the property domain. We use the same approach in the further models that deal with the remaining 6 of the 7 moral domains examined. We corrected all p-values of the variables for multiple testing with appropriately adjusted significance levels using the Holm-Bonferroni method.

moral judgment. Furthermore, the results show that all four dependent variables are cross-culturally above the neutral midpoint (i.e., 4) of the response formats used.[137]

Figure 23: MDFS — Valuations of property deviance across cultures
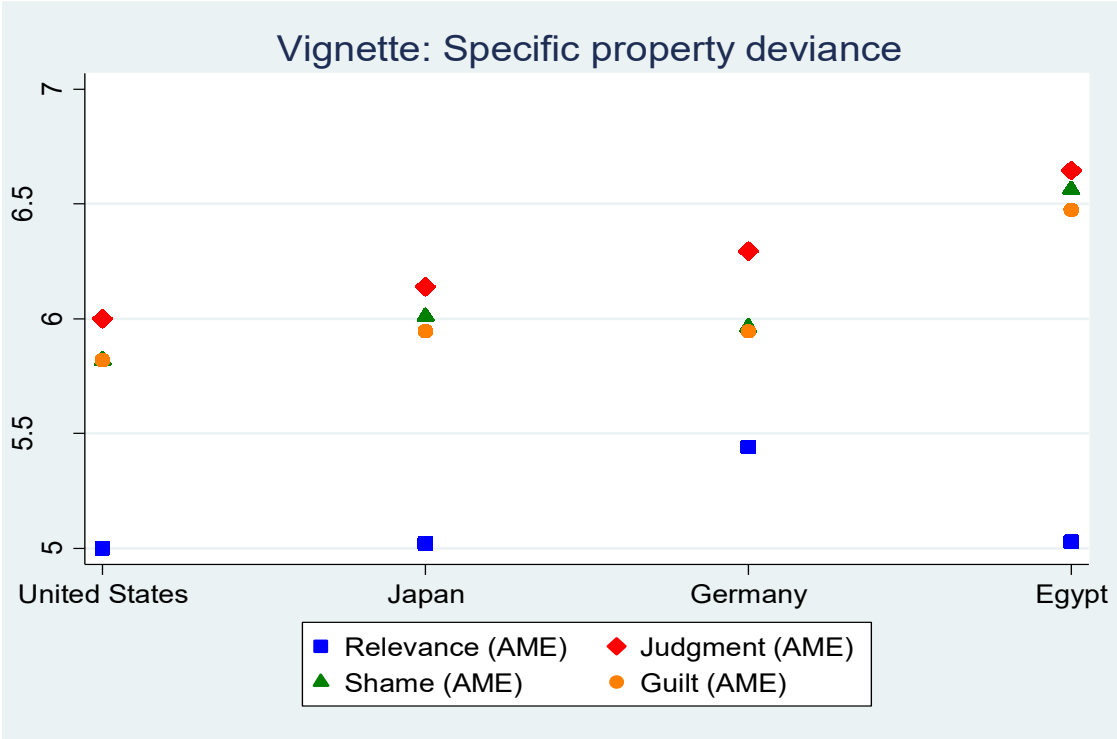
Our graphical analysis of the average marginal effects is complemented by the pairwise group comparisons carried out for the respective margins of the cultural samples. Corresponding findings can be taken from ***Table 47***. The only (Holm- Bonferroni corrected) significant difference between the US American and German sample was found in the context of deviance relevance, with the latter group ascribing significantly greater relevance to the stealing of an object as described in the respective property vignettes. For judgment, and shame and guilt attribution, though, the findings show a different picture. It can be seen from both ***Figure 23*** and ***Table 47*** that the EG-sample evaluates stealing an object as significantly more wrong, and attributes significantly more shame and guilt to such an act of moral deviance than the other samples. Furthermore, the GER-sample evaluates the specific property deviance vignettes as significantly more wrong than the US-sample. The other comparisons show no differences below $p < 0.05$, despite slightly variant margins. As a superordinate finding, however, we note that, despite group differences, the evaluation of the act of specific property deviance points in a similar direction across cultures. In addition to similar cross-cultural

---

[137] Note: The dependent variables are based on a 7-point scale where 7 corresponds to *extremely relevant*, *wrong*, (the person should feel) *extremely ashamed* and *extremely guilty*, respectively. The value 4 represents the neutral midpoint of each response format.

Table 47: Property deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific Deviance *Property*	USA n = 323	Samples Compared (pairwise) †	Japan n = 316	Samples Compared (pairwise)	Germany n = 392	Samples Compared (pairwise)	Egypt n = 333	N = 1364
Relevance	5.000 (.111)	US vs. JP *n.s.* US vs. GER * US vs. EG *n.s.*	5.020 (.129)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.441 (.121)	GER vs. EG *n.s.*	5.028 (.177)	
Judgment	5.998 (.072)	US vs. JP *n.s.* US vs. GER * US vs. EG ***	6.139 (.084)	JP vs. GER *n.s.* JP vs. EG ***	6.294 (.078)	GER vs. EG *	6.645 (.115)	
Shame	5.818 (.067)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG ***	6.010 (.078)	JP vs. GER *n.s.* JP vs. EG ***	5.963 (.073)	GER vs. EG ***	6.562 (.107)	
Guilt	5.820 (.069)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG ***	5.945 (.080)	JP vs. GER *n.s.* JP vs. EG ***	5.947 (.075)	GER vs. EG ***	6.474 (.110)	

† **Note:** All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

tendencies, the EG-sample sets itself, though, somewhat apart, as it attributes greater emotional costs to specific property deviance in particular.

We are primarily interested in the MDFS dimension *social relation* as we want to test the **impartiality/particularism hypothesis**. Hence, we wonder: Is there an effect of whether a *stranger*, a *family member* or an *in-group member* (a friend) is harmed by the deviant act? If we now turn to the respective OLS *for each culture*, the effects of *social relation* dimension, with their p-values (alpha error; before Holm-Bonferroni correction) and standard deviations in parenthesis, can be taken from the table to be found below. We will first discuss the country/sample-specific findings, as these are the focus of the hypothesis testing, and will only later mention the cross-country findings as a supplement. Across models we found (before Holm-Bonferroni correction) a total of 3 significant culture specific effects and further interesting indications with a p-value close to 0.10.

Looking first at the **US-sample**, we see besides impartial tendencies also that stealing from a *family* member is followed by higher attributions of guilt as compared to the reference category (i.e., stranger). We have highlighted the respective effect (p = 0.005), and all other significant effects, within **Table 48** in bold.[138] Next to the effect described, analyses yield also a further indication counter to our theorizing: Stealing from a family member seems to entail stronger judgments of wrongness in the US (p = 0.094) as compared to stealing from a stranger.

The results of the **German sample** show, in addition to impartial tendencies, that stealing from a *family* member is associated with a stronger attribution of shame (p = 0.045). This effect conflicts with the assumption of impartiality. The guilt variable also points to a hint of particularism in the *family* context (p = 0.088). Another particularistic indication is furthermore revealed when looking at judgment and the variable *in-group*, as a positive effect with p = 0.091 suggests a tendency that stealing from an in-group member is judged as more wrong than the same act of deviance when a stranger is harmed.

In the **Japanese sample**, we observe consistently impartial tendencies when examining the culture-specific relevance, judgment, shame, and guilt OLS models. No significant interaction effect for the JP-sample and the three social relationship levels were found.

Also, the effects found in the **EG-sample** are not in support of our hypothesis. In fact, evidence from the Egyptian sample demonstrates significantly decreased relevance when stealing from a *family* member as compared to stealing from a stranger (p = 0.006). This effect,

---

[138] For the remaining tables in the context of the test of the impartiality/particularity hypothesis, we will also highlight significant findings in bold.

Table 48: Who is harmed? The (non-adjusted) effect of different social relations on *property* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Property* Deviance	USA n = 323	95% Conf. Interval	Germany n = 392	95% Conf. Interval	Japan n = 316	95% Conf. Interval	Egypt n = 323 Δ	95% Conf. Interval
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	-.266 (p=.286; .249)	-.7553732 .2233431	.277 (p=.231; .231)	-.1763523 .7306985	.058 (p=.819; .255)	-.4427032 .5599927	-.692 (**p=.006**; .250)	-1.183788 -.2011704
- In-Group	-.296 (p=.239; .251)	-.7892472 .1970729	.375 (p=.102; .229)	-.0747187 .8257002	.259 (p=.307; .253)	-.2383078 .7570611	-.152 (p=.539; .247)	-.6382789 .3338842
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.271 (p=.094; .162)	-.0464371 .5894153	.137 (p=.360; .150)	-.1571508 .4321419	.037 (p=.824; .166)	-.2887101 .3627213	.237 (p=.145; .162)	-.0817309 .5566559
- In-Group	.015 (p=.923; .163)	-.3045722 .3362202	.252 (p=.091; .149)	-.0402336 .5447505	.133 (p=.418; .164)	-.1897947 .4568765	-.052 (p=.744; .160)	-.3682787 2633161
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.197 (p=.191; .152)	-.0988433 .4944355	.281 (**p=.045**; .140)	.0067782 .5566148	.041 (p=.789; .154)	-.2624235 .3453913	.200 (p=.187; .151)	-.097409 .4982346
- In-Group	-.069 (p=.650; .152)	-.3681642 .2297239	.154 (p=.268; .139)	-.1187546 .4270619	.192 (p=.212; .153)	-.1096335 .4937398	-.039 (p=.792; .150)	-.3342994 .255007
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	.432 (**p=.005**; .154)	.1290869 .7361282	.244 (p=.088; .143)	-.0363353 .526256	.053 (p=.736; .158)	-.2575972 .3643173	.182 (p=.240; .155)	-.1222207 .4872403
- In-Group	.026 (p=.864; .155)	-.2792452 .3325123	.187 (p=.188; .142)	-.0919534 .4665245	.211 (p=.178; .157)	-.0968242 .5205458	-.091 (p=.550; .153)	-.3932755 .2097012

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. Δ**Note:** The total sample size for the *fairness* vignettes, encompassing the four sub-samples, is N = 1,354.

next to impartial tendencies on judgment and the attribution of moral emotions, reveals a picture that runs counter to the particularism part of our hypothesis.

The sample-specific effects reveal in parts the opposite of what we have expected. Regarding the evidence on the attribution of moral emotions, the two WEIRD samples demonstrate counter to the theorizing particularistic tendencies. Japan furthermore exhibits impartial tendencies across all measured dependent variables, and the EG-sample goes even further, with property deviance in a family context being less relevant than stealing from a stranger. Drawing on the *uncorrected* p-value results the findings hence paint a more complex picture than that suggested by our binary impartiality/particularism hypothesis. Applying however the Holm-Bonferroni correction to adjust for multiple testing, the *corrected* findings reveal across cultures no difference in the valuation of specific property deviance depending on whether the act of deviance harms a stranger, a family member, or a member of the in-group.[139]

The *non-p-value-adjusted results* that we have presented in ***Table 48*** should therefore best be considered as *indications*, since our OLS models risk being distorted by alpha-error cumulation due to multiple testing. In the context of the *corrected p-values*, though, we should also consider that the adjustment of the significance level is rigorous and conservative due to the size of the test family.

Eventually we turn to the results of the cross-country/sample comparison. We observe that the GER-sample differs significantly and positively from the culture-specific *US-sample* interaction effect for relevance and the *in-group* variable (dimension social relationship; Coeff. = 0.671, Std. Err. = 0.340, p = 0.049). Compared to the US-sample, the German sample thus attributes greater relevance to the vignette scenarios in which a friend is the victim of theft. Furthermore, a significant and negative deviation from the *GER-sample* interaction effect is found when compared with the EG-sample (Coeff. = -0.969, Std. Err. = 0.341, p = 0.005). The latter effect suggests that the German sample, as compared to the Egyptian sample, places significantly more relevance to property deviance when a *family* member is harmed. In regard to the *JP-sample* interaction effects, we identified a significant and negative deviation in the relevance OLS model when compared to the EG-sample and looking at the *family* variable: Coeff. = -0.751, Std. Err. = 0.358, p = 0.036. When it comes to an act of specific property deviance, as described in the respective vignettes, the Japanese sample places greater relevance as the Egyptian sample to the act of deviance when a family member is harmed. No further significant deviations from the sample-specific interaction effects were found. Next, we turn to the analysis of the specific fairness deviance vignettes.

---

[139] The specific property deviance OLS models comprise 31 tests of a test family related to the variable *culture*.

6.4.2. MDFS Analysis II: Fairness Deviance

Based on the analyses conducted with the baseline model, we found that the variables *years in school*, *level of religiosity*, *denomination*, *place of living* and *gender* showed no significant effect across all four models of the dependent variables. Consequently, we excluded these variables from our final models of investigation.[140] In the context of the results obtained for the *specific fairness deviance vignettes*, we first look at the graphical analysis of the dependent variables across the four samples (***Figure 24***).

We find evidence suggesting that deviance relevance is not the same as deviance judgment (**relevance/judgment hypothesis**). These two concepts differ cross-culturally in the extent to which they are attributed to the act of specific fairness deviance. Furthermore, in three out of four samples, the attribution of shame seems to predominate over the attribution of guilt. The US-sample forms an exception though, and attributes slightly more guilt to the deviant actor. Findings reveal in addition that the extent to which moral emotions are attributed to deviant actions does not a priori correspond to the extent to which a deviant action is judged as wrong, as can be seen in ***Figure 24***. If we look e.g. at the *average marginal effects* (AMEs) for the German sample we not only find a divergence between relevance and judgment, but also between these two variables and the attribution of moral emotions. Additionally, ***Figure 24*** shows that shame and guilt attribution, despite slight differences, largely coincide across cultures in their respective strength of attribution.

Moreover, different margins can be found in parts for the samples examined. As far as relevance is concerned, a similar picture emerges across cultures, which is also confirmed by the fact that there are no significant differences between the group-specific margins. However, evidence suggests also cross-cultural differences in the valuation of the deviant act. It is true across cultures that people consider it wrong to take more for oneself when sharing something, as described in the specific fairness deviance vignettes. In the JP- and EG-samples, though, a corresponding action is seen as significantly more wrong than in the US-sample. Moreover, a pattern emerges with regard to the moral emotions of guilt and shame: Japan and Egypt attribute significantly stronger feelings of shame and guilt to the unfair actor in the vignettes than the two WEIRD samples. The corresponding findings can be found in ***Table 49*** (further below).

---

[140] The OLS regression models for the specific fairness deviance vignettes includes the following variables: *Culture*, *social relationship*, *age*, *MRS*, NARS, *pathogen prevalence*, *residential mobility*; interaction terms with *culture*: *social relationship*, *age*, *MRS*, NARS, *pathogen prevalence*.

Figure 24: MDFS — Valuations of fairness deviance across cultures
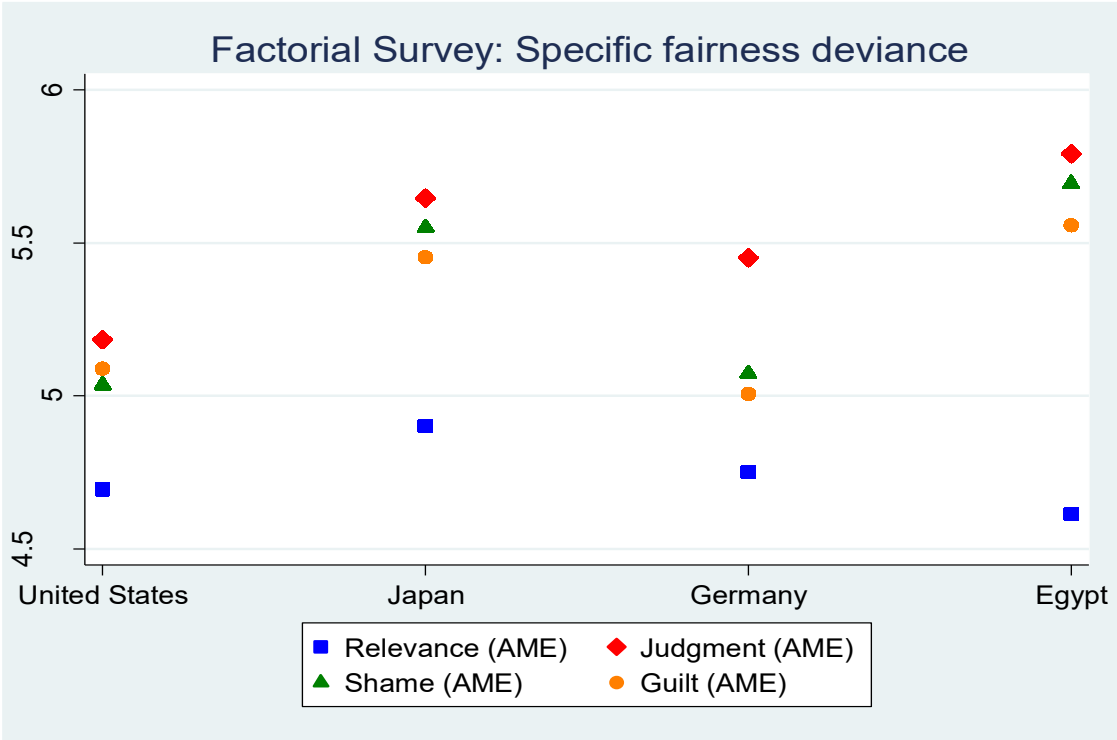
Turning to the **impartiality/particularism hypothesis** in the context of the specific fairness deviance vignettes we display the (*uncorrected*) main results of our OLS models summarized in ***Table 50*** to be found further below. We begin the analysis again with the country-specific *non-adjusted* p-value results for the **US-sample**. Across the four dependent variables a consistent pattern counter to our hypothesis is found. Valuations of specific fairness deviance are stronger in the US group given that a *family* member is harmed through the act of deviance (remember: the reference group is a *stranger*). Neither for the in-group nor for strangers the same effect is found revealing familial particularism in regard to the fairness domain.

Next, we look at the *non-adjusted* findings obtained from the OLS models of the **German**, the **Japanese**, and the **Egyptian sample** together. Results yield evidence for the impartiality part of our hypothesis when we focus on the GER-sample. However, the OLS results of the JP-, and EG-sample reveal the same tendency towards impartiality, leading us to reject our hypothesis for these groups. Across the three groups and all four dependent variables no special importance is placed on the in-group or the family when compared to a deviant act that harms a stranger. Specific fairness deviance is thus evaluated relatively impartial in the German, Japanese, and Egyptian group, which is demonstrated via the variables relevance, judgment, shame, and guilt, and holds across the three levels of the social relation dimension.

In summary, the *non-p-value adjusted results* are only (fully) consistent with our hypothesis for the GER-sample. In the context of specific fairness deviance, impartial tendencies could be demonstrated for the German, Japanese and Egyptian groups examined. By contrast, solely the US-sample shows a particularistic tendency and consistently evaluates specific fairness deviance more strongly when a family member is harmed by the act of deviance. The culture-specific OLS models each comprise a test family with 31 individual tests that refer to the variable *culture*. So, once again we are in the realm of multiple testing. Furthermore, due to the considerable number of tests in the test family, a conservative correction is obtained by applying the Holm-Bonferroni procedure to adjust the significance levels. After this p-value adjustment, we find no significant effect across the four samples and the different levels of the social relation(ship) dimension.

From a comparative perspective, we can identify several deviations from the sample-specific interaction effects on the social relationship dimension. In regard to the relevance valuation of specific fairness deviance the JP-sample departures significantly and negative from the *US-sample* interaction effect by exhibiting lower relevance when a *family* member is harmed: Coeff. = -0.729, Std. Err. = 0.340, p = 0.032. No such effect is found for the judgment variable. However, the US American shame interaction effect for the *family* expression of the social relationship dimension is significantly different and higher than the effect found in the German sample (Coeff. = -0.475, Std. Err. = 0.238, p = 0.046). A further strengthening of the importance of specific fairness conformity in familial contexts in the US-sample is identified by means of integrating another sample-specific deviation from the US interaction effect: The GER-sample, in comparison to the US-sample, exhibits significantly lower guilt attribution when a *family* member is harmed (Coeff. = -0.491, Std. Err. = 0.247, p = 0.048). Thus, we uncovered cross-cultural differences from the United States in the valuation of specific fairness deviance under consideration of different social relationships. No further significant deviations from the country-specific interaction effects could be identified between the samples. In the following, we turn to the vignettes in which a promise is broken and focus thus on specific trustworthiness deviance.

Table 49: Fairness deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific *Fairness* Deviance	USA n = 315	Samples Compared (pairwise) ††	Japan n = 314	Samples Compared (pairwise)	Germany n = 380	Samples Compared (pairwise)	Egypt n = 332	N = 1341
Relevance	4.694 (.102)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.901 (.123)	JP vs. GER *n.s.* JP vs. EG *n.s.*	4.751 (.094)	GER vs. EG *n.s.*	4.613 (.151)	
Judgment	5.183 (.078)	US vs. JP *** US vs. GER *n.s.* US vs. EG ***	5.645 (.094)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.450 (.072)	GER vs. EG *n.s.*	5.791 (.116)	
Shame	5.035 (.076)	US vs. JP *** US vs. GER *n.s.* US vs. EG ***	5.549 (.091)	JP vs. GER *** JP vs. EG *n.s.*	5.073 (.070)	GER vs. EG ***	5.693 (.112)	
Guilt	5.089 (.079)	US vs. JP * US vs. GER *n.s.* US vs. EG **	5.453 (.095)	JP vs. GER *** JP vs. EG *n.s.*	5.007 (.070)	GER vs. EG ***	5.557 (.117)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

Table 50: Who is harmed? The (non-adjusted) effect of different social relations on *fairness* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Fairness* Deviance	USA	95% Conf. Interval	Germany	95% Conf. Interval	Japan	95% Conf. Interval	Egypt	95% Conf. Interval
	n = 315		n = 380		n = 314		n = 332 Δ	
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	.726 (**p=.002**; .238)	.258458 1.195519	.266 (p=.211; .212)	-.1510278 .6836312	-.002 (p=.993; .242)	-.4776184 .4735854	.102 (p=.656; .229)	-.3478745 .5522905
- In-Group	.254 (p=.278; .234)	-.2058395 .7156824	.175 (p=.413; .214)	-.2454926 .5967553	.132 (p=.579; .239)	-.3370648 .602727	-.077 (p=.736; .230)	-.5297622 .3744378
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.437 (**p=.017**; .183)	.0771689 .7981046	.046 (p=.778; .163)	-.2748199 .3673323	.294 (p=.115; .186)	-.0714128 .6604041	.139 (p=.430; .176)	-.2069058 .485644
- In-Group	-.013 (p=.941; .180)	-.3679622 .3410186	.041 (p=.802; .165)	-.2826202 .3653706	.067 (p=.713; .184)	-.2937988 .4292382	-.013 (p=.940; .177)	-.3612543 .3343998
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.519 (**p=.004**; .177)	.1710073 .8686096	.043 (p=.782; .158)	-.2668572 .3545113	.210 (p=.243; .180)	-.1433256 .5648056	.251 (p=.140; .170)	-.0831686 .5869664
- In-Group	-.105 (p=.546; .174)	-.4484986 .2375357	-.046 (p=.771; .159)	-.3601267 .2668915	.064 (p=.719; .178)	-.2856255 .4140099	-.092 (p=.590; .171)	-.4289937 .2441452
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	.590 (**p=.001**; .185)	.2271407 .9532806	.098 (p=.550; .164)	-.2249004 .4218871	.123 (p=.510; .187)	-.2446274 .4924722	.108 (p=.543; .177)	-.2404922 .4570569
- In-Group	-.119 (p=.513; .182)	-.4760852 .2380136	-.035 (p=.832; .166)	-.3615606 .2911077	.005 (p=.978; .185)	-.3589871 .3692692	-.226 (p=.205; .178)	-.5768429 .1238329

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. **ΔNote:** The total sample size for the *fairness* vignettes, encompassing the four sub-samples, is N = 1,341.

6.4.3. MDFS Analysis III: Trustworthiness Deviance

In the analysis of the baseline model for the *specific trustworthiness deviance* vignettes, we were only able to determine that the variable *gender* has no significant effect. However, since an effect was also found for this variable at the 10% significance level, we decided to rely on the baseline model for the final analyses of the vignettes. Accordingly, the corresponding OLS regressions include the following variables: *culture*, *years in school*, *level of religiosity*, *MRS*, *NARS*, *pathogen prevalence*, *social relationship*, *denomination*, *place of living*, *residential mobility* and *gender*. In addition, several interactions terms are also part of the models: interaction terms *culture* with *years in school*; *level of religiosity*; *response styles*; *pathogen prevalence*; and *social relationship* (Factorial Survey); interaction term *level of religiosity* with *denomination*. Based on this set of variables we have estimated the culture specific margins, which are shown in ***Figure 25***.

A divergence in the extent of the valuation between relevance and judgment can also be found in the specific trustworthiness deviance models. This evidence once again highlights the difference between these constructs and is in support for our **relevance/judgment hypothesis**. Apart from this fact, the Egyptian sample stands out. In all samples except the Egyptian one, the attributions of the moral emotions coincide almost perfectly. In the Egyptian sample, however, we see a stronger attribution of shame for the act of specific trustworthiness deviance. Thus, guilt and shame are not universally attributed to the same extent across cultures. What is more, we can infer from ***Figure 25*** that the Egyptian sample is characterized by higher margins in deviance relevance, judgment and shame attribution. However, a more detailed analysis of the margins in the pairwise comparison of the samples reveals only a significant difference between the US- and the EG-sample for deviance judgment (see: ***Table 51*** further below). Thus, despite minor differences, a cross-cultural tendency emerges by and large: The valuation of specific violations in the domain of trustworthiness is largely similar across the cultural entities examined.

Figure 25: MDFS — Valuations of trustworthiness deviance across cultures

Factorial Survey: Specific trustworthiness deviance

Legend:
- Relevance (AME) — blue square
- Judgment (AME) — red diamond
- Shame (AME) — green triangle
- Guilt (AME) — orange circle
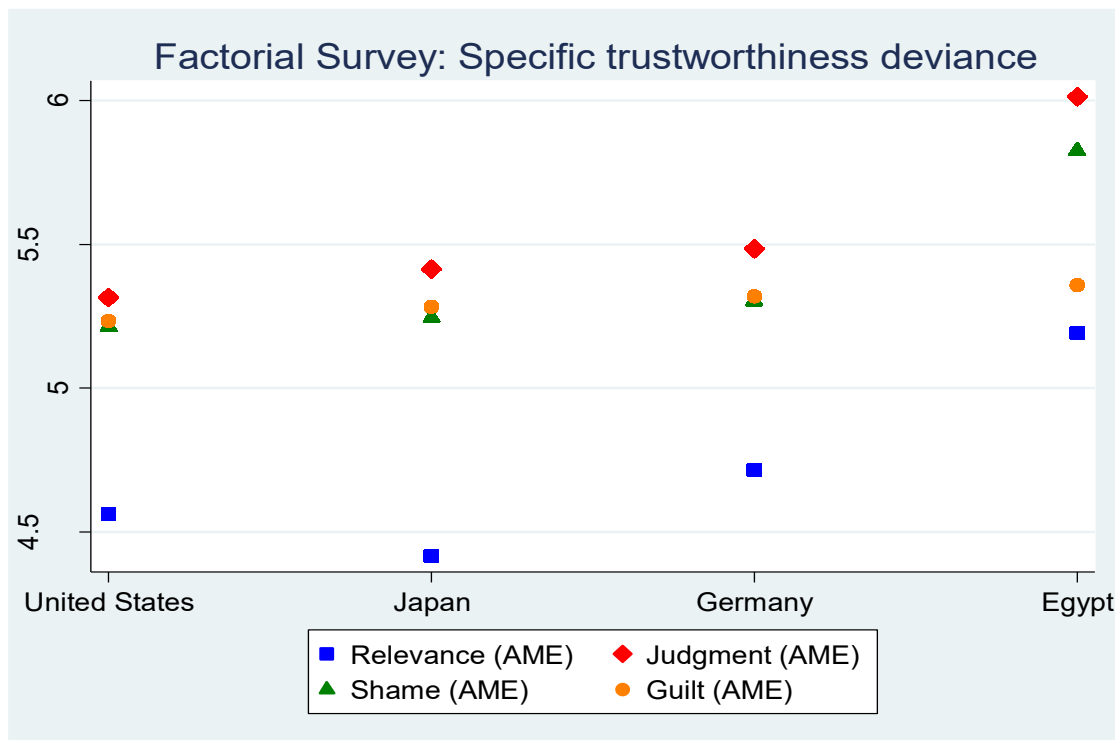
Subsequently we are again pursuing the question of whether there is an effect of *who* is harmed by the act of deviance and focus on our main research interest the **impartiality/particularism hypothesis**. We start by addressing the sample-specific effects and then mention the cross-country comparison findings further below. The results of the specific trustworthiness deviance OLS models for the **US-sample** lend only credence to our hypothesis when looking at the relevance variable and partly for the family level of the social relation dimension. In accordance, we find particularistic tendencies in the realm of deviance judgment as well as shame and guilt attribution. The findings reveal positive interaction effects for the *in-group*, highlighting the particular severity of breaking a promise to a friend. Also, a particularistic indication (p= 0.053) is revealed for guilt in the context of a harmed *family* member. Hence, the results by and large suggest that the impartiality part of our hypothesis has to be partially rejected in the US context.

In the **German sample**, we find evidence for our assumption that this group tends to be impartial. Regardless of whether the issue is relevance, judgment, or the attribution of moral emotions, the GER-sample shows no differences in the evaluation of specific trustworthiness deviance when a stranger, a family member, or a member of the in-group is harmed.

The OLS results for the **JP-sample** contradict the particularism part of our hypothesis for this group despite one exception. Analyses show only a particularistic tendency when judging the act of deviance: To break a promise to a friend (*in-group*) is evaluated as

Table 51: Trustworthiness deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific *Trustworthiness* Deviance	USA n = 321	Samples Compared (pairwise) ††	Japan n = 297	Samples Compared (pairwise)	Germany n = 373	Samples Compared (pairwise)	Egypt n = 349	N = 1340
Relevance	4.561 (.153)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.416 (.235)	JP vs. GER *n.s.* JP vs. EG *n.s.*	4.716 (.154)	GER vs. EG *n.s.*	5.192 (.284)	
Judgment	5.314 (.107)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *	5.413 (.163)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.484 (.107)	GER vs. EG *n.s.*	6.013 (.197)	
Shame	5.212 (.106)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	5.245 (.162)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.302 (.106)	GER vs. EG *n.s.*	5.825 (.196)	
Guilt	5.234 (.104)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	5.284 (.159)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.319 (.104)	GER vs. EG *n.s.*	5.358 (.192)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

Table 52: Who is harmed? The (non-adjusted) effect of different social relations on *trustworthiness* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Trust- worthiness* Deviance	USA n = 321	95% Conf. Interval	Germany n = 373	95% Conf. Interval	Japan n = 297	95% Conf. Interval	Egypt n = 349 Δ	95% Conf. Interval
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	.063 (p=.785; .231)	-.3906062 .5169746	.066 (p=.759; .218)	-.3614549 .4952183	.345 (p=.160; .245)	-.1360127 .8260728	-.126 (p=.578; .227)	-.5736504 .3203123
- In-Group	.112 (p=.640; .241)	-.3606411 .5860934	-.040 (p=.857; .226)	-.4848713 .4030492	.272 (p=.270; .247)	-.2126636 .7582705	-.132 (p=.558; .226)	-.5760046 .3111655
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.190 (p=.237; .160)	-.1252983 .5057741	.225 (p=.138; .151)	-.0722398 .5234348	.039 (p=.818; .170)	-.2953021 .3736694	-.267 (p=.091; .158)	-.57875 .0428534
- In-Group	.360 (**p=.032**; .167)	.0315918 .6898893	-.029 (p=.849; .157)	-.3386606 .2787414	.376 (**p=.029**; .172)	.0392901 .7144143	-.203 (p=.196; .157)	-.5119015 .1049787
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.152 (p=.339; .160)	-.1609611 .4668746	.103 (p=.494; .151)	-.1929144 .399705	-.016 (p=.924; .169)	-.3488919 .3166485	-.287 (p=.068; .157)	-.5965801 .0218351
- In-Group	.352 (**p=.035**; .166)	.0246704 .6795915	-.031 (p=.842; .156)	-.3382338 .2760016	.159 (p=.351; .171)	-.1760322 .4956294	-.091 (p=.560; .156)	-.3980339 .2156823
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	.304 (p=.053; .156)	-.0037193 .6118284	.081 (p=.581; .148)	-.2086681 .3723526	.028 (p=.863; .166)	-.297641 .3548734	-.262 (p=.089; .154)	-.5658148 .0404968
- In-Group	.441 (**p=.007**; .163)	.1206416 .7627446	-.011 (p=.940; .153)	-.3126654 .2895481	.202 (p=.229; .167)	-.1270907 .5314251	-.104 (p=.494; .153)	-.4057574 .1959472

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. Δ**Note:** The total sample size for the *trustworthiness* vignettes, encompassing the four sub-samples, is N = 1,340.

significantly more wrong as breaking a promise to a stranger. Next to this particularistic peak, the results across deviance relevance and the attribution of moral emotions show yet exclusively impartial tendencies.

The particularism hypothesis for the **EG-sample** is clearly refuted. Not only do we find no significant effect in favor of the social relationship levels family and in-group, but also all algebraic signs of the EG-sample interaction effects are consistently negative (see: ***Table 52***). Moreover, we obtain indications that specific trustworthiness deviance in the *family* context is judged (p = 0.091) as less wrong. Further indications suggest that also moral emotions are imposed to a lesser degree in the *family* (shame: p = 0.068; guilt: p = 0.089) and *in-group* (shame: p = 0.056) context. Thus, instead of a tendency towards particularism in the Egyptian sample we observe rather the opposite.

Drawing on the *non-adjusted* results, our **impartiality/particularism hypothesis** is only fully supported in Germany. Although the Japanese sample evaluates breaking a promise to an in-group member as more wrong as specific trustworthiness deviance towards a stranger, the JP-sample yet displays impartial tendencies in regard to the relevance variable, and the attribution of moral emotions. The Egyptian and US American results most strongly contradict our hypothesis. We observe particularistic tendencies in the US-sample, stressing in-group trustworthiness conformity. Furthermore, over and above our theorizing, results yield that the EG-sample exhibits a slight inclination to particularly emphasize keeping promises beyond family and friendship ties, in addition to a general tendency towards impartiality. Our analyses are based on OLS models with a test family of 34 individual tests referring to the variable *culture*. When applying the Holm-Bonferroni correction, no significant effect is found for the four samples and the three levels of the social relation dimension.

Turning to the cross-sample comparison no significant departure from the *US-sample* interaction effect is found for the other cultural samples when looking at the relevance variable. Nevertheless, significant and negative departures are identified in the deviance judgment (EG-sample: *family*, Coeff. = -0.458, Std. Err. = 0.225. p = 0.042; *in-group*, Coeff. = -0.564; Std. Err. = 0.230; p = 0.014) and deviance guilt attribution (GER-sample: *in-group*, Coeff. = -0.453, Std. Err. = 0.223, p = 0.043; EG-sample: *family*, Coeff. = -0.566, Std. Err. = 0.220, p = 0.010; *in-group*, Coeff. = -0.546, Std. Err. = 0.224, p = 0.015) OLS models. These effects reveal significantly stronger particularistic tendencies in the US-sample as compared to the EG- and

the GER-sample. A significant and negative departure from the *German sample* interaction effect of the judgment variable and the social relationships category *family* is observed when compared with the Egyptian sample (Coeff. = -0.493; Std. Err. = 0.219; p = 0.025). This effect demonstrates further cross-cultural differences in the valuation of moral deviance and suggests that, in the realm of the trustworthiness domain, harm towards a family member is more important in the GER-sample than in the EG-sample. Looking at deviations from the *JP-sample* interaction effects, we observe that the Egyptian samples does not fall for the tendency to judge specific trustworthiness deviance towards an *in-group* member as more wrong as towards a stranger, which is why the respective EG-sample effect is negative and significantly different from that of the JP-sample (Coeff. = -0.580, Std. Err. = 0.233, p = 0.013). The *EG-sample* results have already been mentioned and we found no further significant deviations from the country-specific interaction effects among the samples. In the following we deal with vignettes that portray cowardly behavior of a person although another person needs help and turn thus to the vignettes that portray an act of specific heroism deviance.

6.4.4. MDFS Analysis IV: Heroism Deviance

The baseline model for the *heroism vignettes* identifies the following variables as consistently non-significant: *years in school*, *denomination*, *place of living* and *gender*. Consequently, we removed these variables to obtain the adapted model of analysis.

First and foremost, although this is not as pronounced in the JP- and GER-sample as in the previous analyses, we still find further support for the **relevance/judgment hypothesis** overall. Evidence suggests that deviance relevance and deviance judgment differ in the extent to which they are attributed toward an act of specific heroism deviance. Moreover, as can be seen in ***Figure 26***, the US-sample stands out as the only sample of the four cultural groups examined attributing slightly more guilt than shame. In addition, we only observe minor differences in deviance relevance between the cultural samples. However, as far as moral judgment, shame and guilt are concerned, the JP-sample appears as an outlier with lower margins. What is interesting about this case is that, firstly, all four dependent variables differ only slightly from each other and, secondly, the attribution of moral emotions takes on a higher value than moral judgment for the first time across our models. ***Figure 26*** also shows that relevance, judgment, shame and guilt in the GER-sample, and partly in the US-sample, are largely in a similar cluster as the range of values in the Japanese sample. This cannot be said,

Figure 26: MDFS — Valuations of heroism deviance across cultures
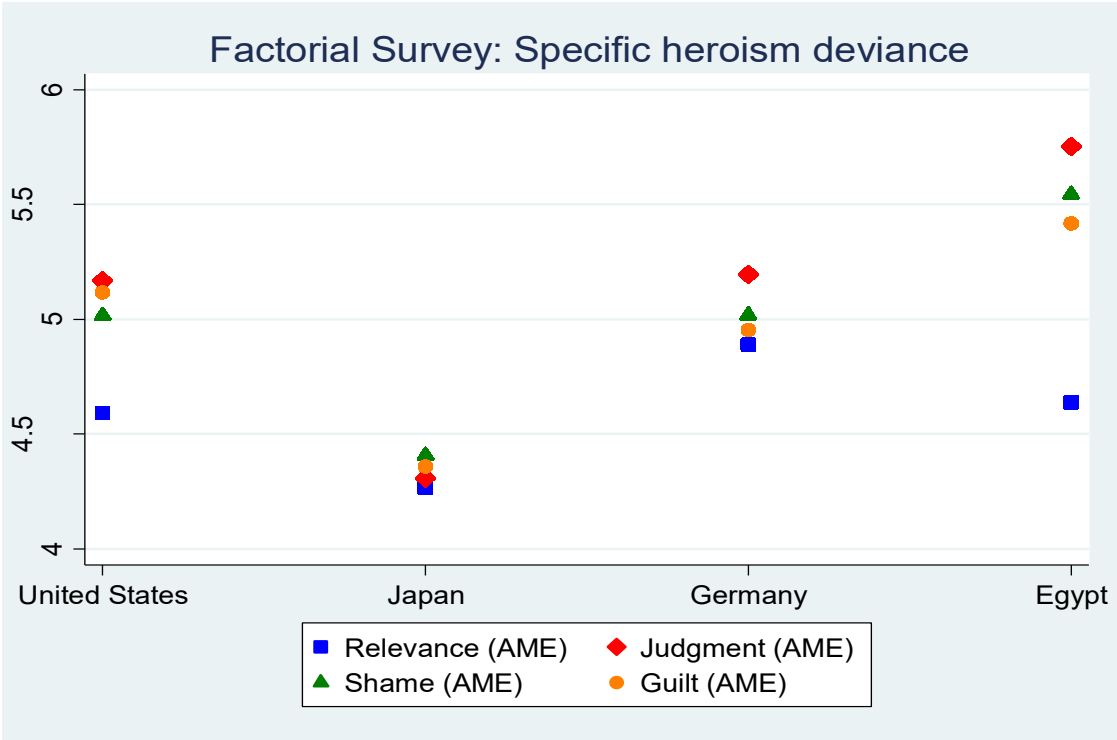
though, for the Egyptian group: A considerable gap exists within the range of values between deviance judgment and deviance relevance in particular within the EG-sample. The specific heroism deviance vignettes describe cowardly behavior although someone else is in distress. Against the background of these vignettes, the Egyptian sample seems to form the opposite pole to the JP-sample. In terms of shame and guilt attribution as well as moral judgment, the EG-sample is marked by significantly higher margins.

Overall, we can identify quite clearly cultural differences among our samples, especially with regard to judgment and the attribution of behavior-regulating emotions (shame and guilt). If we were to sort the four samples in ascending order, the Japanese sample would be significantly different from the other groups at the lower end, the GER- and US-sample would be in the middle without major differences to each other, and the Egyptian sample would be significantly different from the other groups at the upper end. The statement just made is supported by the results of the pairwise comparison of margins shown in *Table 53*.

Does it make a difference *who* is harmed by specific heroism deviance? Next, we are turning to the **impartiality/particularism hypothesis** to approach this question. The *non-adjusted* **US-sample** OLS results reveal particularistic tendencies in deviance judgment, shame and guilt attribution but not in deviance relevance. Significant effects are exclusively found for the *family* variable and so the results reveal kinship altruism by showing that the protection of the family is particularly important in the United States. Thus, when drawing on the results for

Table 53: Heroism deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific Deviance *Heroism*	USA n = 319	Samples Compared (pairwise) ††	Japan n = 320	Samples Compared (pairwise)	Germany n = 373	Samples Compared (pairwise)	Egypt n = 329	N = 1341
Relevance	4.593 (.106)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.265 (.147)	JP vs. GER ** JP vs. EG *n.s.*	4.889 (.120)	GER vs. EG *n.s.*	4.636 (.210)	
Judgment	5.167 (.083)	US vs. JP *** US vs. GER *n.s.* US vs. EG **	4.307 (.115)	JP vs. GER *** JP vs. EG ***	5.194 (.094)	GER vs. EG **	5.754 (.165)	
Shame	5.016 (.084)	US vs. JP *** US vs. GER *n.s.* US vs. EG *	4.407 (.117)	JP vs. GER *** JP vs. EG ***	5.017 (.095)	GER vs. EG *	5.543 (.167)	
Guilt	5.119 (.084)	US vs. JP *** US vs. GER *n.s.* US vs. EG *n.s.*	4.359 (.117)	JP vs. GER *** JP vs. EG ***	4.953 (.095)	GER vs. EG *	5.418 (.167)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

the family level and apart from the relevance variable, our theorizing of impartiality must be seen as largely refuted for the US.

Turning to the **GER-sample**, we see that the *non-adjusted* p-values reveal significant interaction effects suggesting higher shame and guilt attribution in the context of heroism deviance when an *in-group* member is harmed. Standing at the side of one´s in-group, although this may imply personal harm, appears to be particularly emotionally incentivized in Germany. Further particularistic indications can be found on the judgment variable, which indicates a special importance of the *in-group* (p = 0.064), and in the shame-model, in which the *family* variable (p = 0.075) reveals a tendency away from impartiality. Thus, especially the moral emotion-variables reveal evidence contra to our hypothesis for Germany.

The results of the **JP sample** contradict our hypothesis, as evidenced by impartial tendencies found on all dependent variables. Consequently, in this sample, it does not appear to make a difference who is harmed by specific heroism deviance.

Turning lastly to the *non-adjusted* **EG-sample** results, we see our particularism hypothesis mainly supported. Across the dependent variables judgment, shame, and guilt, it is the *family* category that elicits higher wrongness evaluations and more attributions of aversive emotions when a person deviates from the moral domain of heroism. We also find a positive effect for the *in-group* variable in the shame-model, demonstrating particularistic tendencies that comprise the in-group next to the family.

Taken together, the results of the *unadjusted p-values* of the heroism vignettes show that more complex relationships exist in the multitude of cultural entities and realities in our social world than those posited by our binary **hypothesis of impartiality/particularism** (see: *Table 54*). Not only were we wrong in the context of the WEIRD samples, but the Japanese sample also shows an inverted picture exhibiting a tendency towards impartiality instead of particularism. Moreover, the particularism effects observed are not homogeneous across cultures: in the United States it is heroism deviance towards the family, in Germany towards the in-group, and in Egypt towards the family and the in-group that is particularly important. Finally, it should be mentioned that our OLS models each include a test family of 34 individual tests that refer to the variable *culture*. Accordingly, the correction of the significance level, applying the Holm-Bonferroni method, is restrictive. After adjusting the p-values, no significant effects are found for the levels of the social relation dimension and the four dependent variables examined.

Table 54: Who is harmed? The (non-adjusted) effect of different social relations on *heroism* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Heroism* Deviance	USA	95% Conf. Interval	Germany	95% Conf. Interval	Japan	95% Conf. Interval	Egypt	95% Conf. Interval
	n = 319		n = 373		n = 320		n = 329 Δ	
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	.174 (p=.445; .228)	-.2742725 .623821	.207 (p=.329; .212)	-.2091485 .624268	-.095 (p=.679; .231)	-.5489887 .3577099	-.164 (p=.464; .224)	-.6049714 .2761734
- In-Group	-.046 (p=.842; .231)	-.5013152 .408869	.202 (p=.344; .214)	-.2178574 .6236581	.214 (p=.356; .232)	-.2413103 .6711355	.115 (p=.611; .226)	-.3288393 .5593188
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.424 (**p=.019**; .180)	.071239 .7778815	.096 (p=.563; .167)	-.2310725 .4246805	.192 (p=.289; .181)	-.163764 .5496492	.386 (**p=.029**; .176)	.0394677 .7327746
- In-Group	.156 (p=.391; .182)	-.2013171 .5148387	.313 (p=.064; .168)	-.0179426 .6441829	.275 (p=.132; .182)	-.0834226 .6345127	.151 (p=.396; .178)	-.1982568 .5005682
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.387 (**p=.034**; .182)	.0300334 .7450187	.300 (p=.075; .169)	-.0308646 .6326305	-.158 (p=.183; .390)	-.5191498 .2026862	.501 (**p=.005**; .178)	.1509306 .8524229
- In-Group	.138 (p=.459; .184)	-.2256424 .4989685	.401 (**p=.019**; .170)	.0662864 .7362292	.197 (p=.185; .285)	-.1653827 .5610288	.388 (**p=.031**; .180)	.0350941 .7421697
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	.385 (**p=.034**; .181)	.0287948 .7421442	.270 (p=.109; .168)	-.0603404 .6016366	-.049 (p=.787; .183)	-.4097822 .3104022	.405 (**p=.023**; .178)	.0557865 .7556736
- In-Group	.260 (p=.157; .184)	-.1005963 .6223567	.425 (**p=.013**; .170)	.0914275 .7598374	.221 (p=.231; .184)	-.1408866 .5838628	.273 (p=.129; .179)	-.079437 .6260207

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. **ΔNote:** The total sample size for the *heroism* vignettes, encompassing the four sub-samples, is N = 1,341.

The cross-country comparison reveals that no significant departure from the *US-sample* interaction effect is found on the relevance, judgment, and guilt variables when compared to the other samples. Nonetheless, the OLS shame model shows a significant and negative deviation from the US-sample interaction effect on the *family* variable when compared to the Japanese sample: Coeff. = -0.545, Std. Err. = 0.258, p = 0.035. The latter effect suggests that the attribution of shame in the specific heroism deviance vignettes is associated with stronger emotional promotion of moral conformity in the US-sample than in the Japanese group examined. Furthermore, significant departures from the *JP-sample* interaction effect in the OLS shame model are identified. Next to the significant difference between the US- and JP-sample mentioned above, we observe also a significant difference between the Egyptian and the Japanese sample. When it comes to the attribution of shame, given a *family* member is harmed by the act of heroism deviance, the EG-samples displays significantly higher emotional incentivizing than the JP-sample (Coeff. = 0.659, Std. Err. = 0.256, p = 0.010). Comparisons with the German sample reveal no further significant cross-cultural differences. The following vignettes deal with specific reciprocity deviance and their valuations across cultures. In these vignettes, a person behaves ungratefully and does not return a favor.

6.4.5. MDFS Analysis V: Reciprocity Deviance

Based on the insights from analyzing the baseline model, we removed the variables *denomination*, *place of living* and *gender* from our final *specific reciprocity deviance* models.[141] The resulting adapted model then forms the starting point for estimating the average marginal effects (AMEs) for the dependent variables and cultural groups. ***Figure 27*** provides a graphical representation of our findings.

First of all, ***Figure 27*** shows that a cross-cultural pattern emerges: The highest values are given for judgment — behaving ungratefully and not returning a favor, i.e., showing an act of specific reciprocity deviance is evaluated as relatively wrong across cultures. In the extent of the deviance valuation across cultures judgment is followed by shame, guilt and then the relevance margins appear at the lower end of the range of values. Accordingly, a look at ***Figure 27*** yields further evidence that deviance judgment and deviance relevance do not seem to capture one and the same concept, at least when referring to the extent of evaluating specific

---

[141] The adapted model for the reciprocity vignettes comprises the following variables: *Culture* and interaction terms with *culture* (*years of school*, level of religiosity, *age*, *NARS*, *MRS*, *pathogen prevalence* and *social relationship*); *years of school*, level of religiosity, *age*, *NARS*, *MRS*, *pathogen prevalence*, *social relationship* and *residential mobility*.

acts of moral deviance (**relevance/judgment hypothesis**). What is more, in regard to specific reciprocity deviance relevance, we did not find major differences across the samples. However, the Egyptian sample appears to be partly an outlier in terms of shame/guilt attribution and deviance judgment, with higher margins in the range of values than the other samples.

The pairwise comparison (***Table 55***) supports the previously derived tendency: As for the variables of shame, guilt and judgment, the EG-sample differs significantly from the other groups, revealing cross-cultural differences in the valuation of specific transgressions from the moral domain of reciprocity. Thus, violating reciprocity, by acting ungratefully and not returning a favor, is seen as more wrong in the Egyptian sociocultural context and is subject to a stronger push for conformity, which is reflected in higher attributions of shame and guilt. In addition, it can also be noted that the GER-sample, in comparison to the other groups, is characterized by significantly lower attributions of guilt and shame.

Figure 27: MDFS — Valuations of reciprocity deviance across cultures
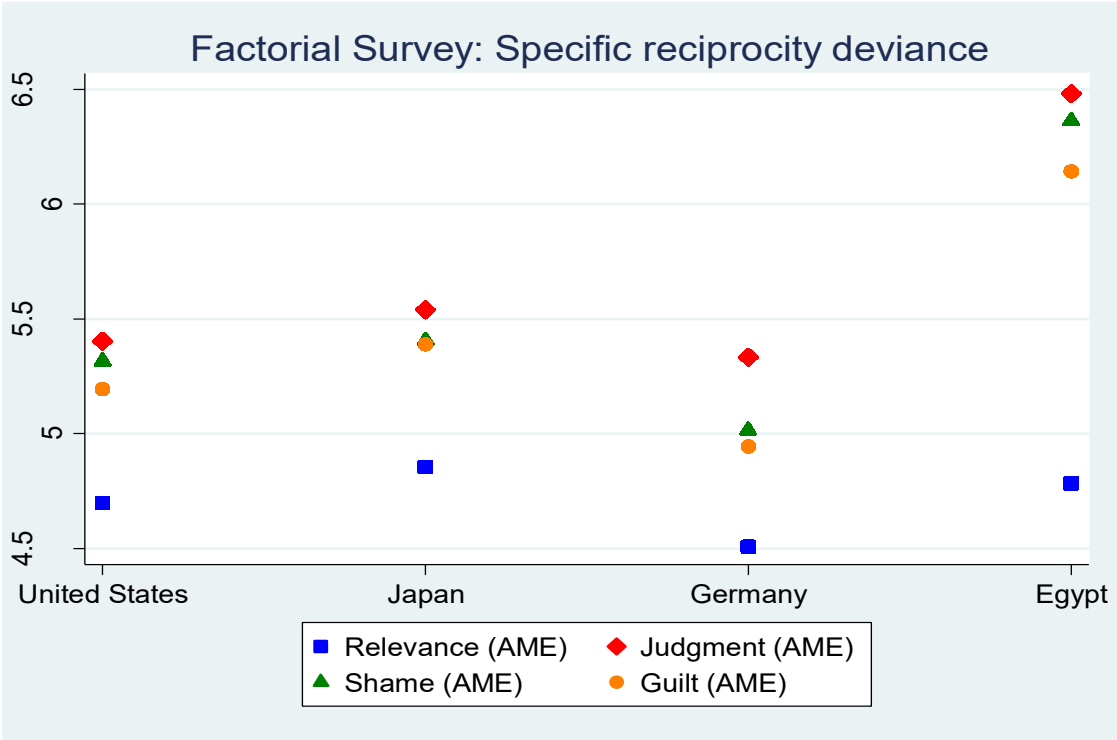
In the course of examining the **impartiality/particularism hypothesis** the following can be noted: The **US-sample** exhibits no tendencies of particularism. In fact, the relevance, shame, and guilt OLS models reveal impartial tendencies in line with our theorizing. However, in addition to these tendencies, we also observe an interaction effect in the (*unadjusted*) judgment model indicating that reciprocity outside the *family* context, i.e., in interactions with *strangers*, is judged to be more wrong. A further indication in the same direction can be found in the shame-model, in which the effect of the *family* variable has a negative algebraic sign and

Table 55: Reciprocity deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific *Reciprocity* Deviance	USA n = 338	Samples Compared (pairwise) ††	Japan n = 307	Samples Compared (pairwise)	Germany n = 386	Samples Compared (pairwise)	Egypt n = 323	N = 1354
Relevance	4.697 (.112)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.855 (.155)	JP vs. GER *n.s.* JP vs. EG *n.s.*	4.507 (.119)	GER vs. EG *n.s.*	4.784 (.200)	
Judgment	5.402 (.074)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG ***	5.539 (.103)	JP vs. GER *n.s.* JP vs. EG ***	5.333 (.079)	GER vs. EG ***	6.481 (.132)	
Shame	5.316 (.074)	US vs. JP *n.s.* US vs. GER * US vs. EG ***	5.406 (.102)	JP vs. GER ** JP vs. EG ***	5.014 (.078)	GER vs. EG ***	6.364 (.131)	
Guilt	5.194 (.074)	US vs. JP *n.s.* US vs. GER * US vs. EG ***	5.389 (.103)	JP vs. GER ** JP vs. EG ***	4.946 (.079)	GER vs. EG ***	6.143 (.132)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

Table 56: Who is harmed? The (non-adjusted) effect of different social relations on *reciprocity* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Reciprocity* Deviance	USA	95% Conf. Interval	Germany	95% Conf. Interval	Japan	95% Conf. Interval	Egypt	95% Conf. Interval
	n = 338		n = 386		n = 307		n = 323 Δ	
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	-.257 (p=.279; .237)	-.7236262 .2091188	.078 (p=.716; .215)	-.3441076 .5010933	-.181 (p=.458; .244)	-.6616937 .2986209	.022 (p=.923; .236)	-.4419071 .4878162
- In-Group	.066 (p=.777; .234)	-.3942227 .5275992	.285 (p=.193; .219)	-.1446196 .7158928	-.056 (p=.814; .240)	-.527931 .4150781	.043 (p=.854; .238)	-.4243135 .5119954
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	-.397 (**p=.012**; .157)	-.7073737 -.0885962	-.247 (p=.083; .142)	-.5281863 .0325149	-.023 (p=.885; .162)	-.3421241 .2949429	-.038 (p=.805; .157)	-.3472038 .269569
- In-Group	-.129 (p=.406; .155)	-.4352276 .1763036	-.056 (p=.698; .145)	-.3419934 .2288654	-.247 (p=.121; .159)	-.5601304 .0654562	.078 (p=.620; .158)	-.2319525 .3891892
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	-.258 (p=.099; .156)	-.5653331 .0485759	-.237 (p=.094; .141)	-.5159541 .0403355	-.048 (p=.761; .161)	-.364964 .26709056	.087 (p=.573; .155)	-.2180922 .3938279
- In-Group	-.031 (p=.841; .154)	-.3344334 .2722863	.042 (p=.769; .144)	-.2408505 .3255169	-.221 (p=.162; .158)	-.5317766 .0888879	.135 (p=.389; .157)	-.1727585 .4434961
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	-.169 (p=.284; .157)	-.4786588 .1405482	-.155 (p=.278; .143)	-.4358632 .1252272	-.056 (p=.729; .162)	-.3750655 .2624437	.076 (p=.629; .157)	-.2325853 .3846157
- In-Group	-.038 (p=.803; .155)	-.3449635 .2669922	.225 (p=.121; .145)	-.0596631 .511592	-.209 (p=.190; .159)	-.5221786 .1038423	.207 (p=.190; .158)	-.1029872 .5185857

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. **ΔNote:** The total sample size for the *reciprocity* vignettes, encompassing the four sub-samples, is N = 1,354.

a p-value of p = 0.099 (see: ***Table 56***). Taken together, the US-sample results are largely consistent with the impartiality part of our hypothesis and yet suggest also that conditional cooperation may be especially important beyond group relationships (see the almost consistently negative algebraic signs for the social relationship levels *family* and *in-group*: ***Table 56***).

The *non-adjusted* **GER-sample** results give further support to our theorizing. No inclinations of particularism are found and similar to the US-sample, the algebraic signs of the *family* variable are consistently negative. Also, further indications in the judgment-model (p = 0.083) and in the shame-model (p = 0.094) suggest that reciprocity conformity could play a particularly important role in interactions with strangers.

Looking at the **JP-sample** results, we encounter results that contradict the particularism assumption for this group. Neither for the in-group nor for the family a special importance can be observed in regard to specific reciprocity deviance. Additionally, all algebraic signs of the effects for the *family* and *in-group* variables are negative (see: ***Table 56***). In consequence, not only is our hypothesis for this group refuted, but we find consistent signs of impartiality across all four dependent variables in the Japanese sample.

Similarly, the analyses of the **EG-sample** show no particularistic tendencies in the context of specific reciprocity deviance. Although we do not find predominantly negative algebraic signs for the interaction effects on the dependent variables, on the whole an impartial tendency is yet revealed.

All in all, our findings suggest partly a cross-culturally similar valuation of specific reciprocity deviance, which seems to apply regardless of who is harmed by the act of deviance. Interestingly, though, reciprocity may seem to play a special role beyond group ties in the WEIRD samples of our study, as indicated e.g. by a negative interaction effect for judgment and the family category when drawing on the *non-adjusted* US-sample results. After applying the Holm-Bonferroni correction, we were no longer able to detect a significant effect for the different levels of the social relationship dimension.[142]

In the country comparison, we found no significant deviations from the sample-specific interaction effects, with one exception: The *JP-sample* shows significantly less guilt attribution than the GER-sample when a member of their *in-group* is harmed by the act of specific reciprocity deviance (Coeff. = -0.435, Std. Err. = 0.216, p = 0.044). In the following, we will look at the moral domain of loyalty, in which the focus is on the act of betraying a person.

---

[142] The corresponding OLS models comprise a test family with 35 individual tests that relate to the *culture* variable. Hence, the p-value adjustments using the Holm-Bonferroni method are correspondingly conservative.

6.4.6. MDFS Analysis VI: Loyalty Deviance

The results from analyzing the baseline model show that neither in the relevance, judgment, shame nor in the guilt model the variables *level of religiosity*, *denomination*, *place of living* and *gender* exert a significant effect. As a consequence, we removed these variables from the covariate set to obtain the adapted OLS model. Based on this model, we analyze the *specific loyalty deviance* vignettes in which a person publicly betrays and undermines another person.[143]

     Figure 28 quite clearly demonstrates a gap between deviance relevance and deviance judgment. This gap is particularly noticeable for the EG-sample, where judgment, shame and guilt cluster with the highest values in the comparison across samples, and relevance drops significantly and is in a similar range of values compared to the other samples. In fact, the four groups do not differ significantly in terms of deviance relevance (*Table 57*). So, by and large, all samples tend to cluster in terms of judgment, shame and guilt and display a gap to the

Figure 28: MDFS — Valuations of loyalty deviance across cultures
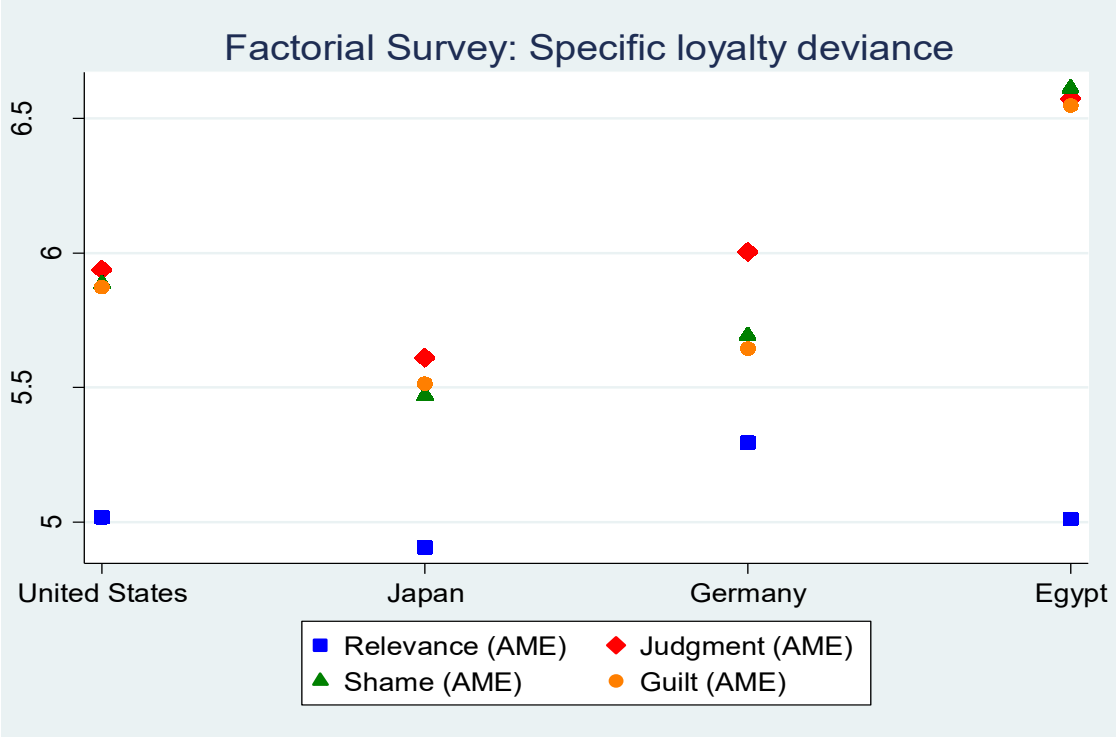
concept of deviance relevance. Accordingly, results yield evidence in support of the **relevance/judgment hypothesis**. We observe also cross-cultural differences. The JP-sample seems to form an antipole to the Egyptian sample in terms of its values (AMEs). While we find

---

[143] The *specific loyalty deviance* models are estimated on the basis of the following variables: *culture* (interaction terms with: *years in school*, *age*, *MRS*, *NARS*, *pathogen prevalence* and *social relationship*); *years in school*, *age*, *MRS*, *NARS*, *pathogen prevalence*, *social relationship* and *residential mobility*.

the highest judgement, shame and guilt AMEs for the EG-sample, the Japanese group displays consistently the lowest margins across the samples. Furthermore, even if it is only a minor difference, it should also be highlighted that a guilt value was found for the JP-sample that exceeds the margin of shame attribution. In the EG-sample, moreover, the attribution of shame has the highest value of the group-specific margins and thus also exceeds the deviance judgment value, which consistently exhibits the highest value of the margins for the other groups.

The graphical analysis is further substantiated by the pairwise comparisons of the sample-specific margins (*Table 57*). The US- and GER-sample only differ significantly in the attribution of guilt — the US-American sample shows a higher attribution of this moral emotion for specific loyalty deviance. Otherwise, and apart from relevance, the EG-sample scores significantly higher on judgment and the attribution of moral emotions than the three other samples. The following ranking of significant differences for judgment, shame and guilt can be identified across the four cultural samples: Egypt shows the highest values, in the middle are the two WEIRD samples, and the Japanese sample shows the lowest margins among the groups.

Next, we turn to the **impartiality/particularism hypothesis** and address first the sample specific findings. The results for the **US American group** consistently show impartial tendencies for all four dependent variables when evaluating specific loyalty deviance vignettes (see: *Table 58*). Consequently, our hypothesis is confirmed and betraying someone, i.e., the act of specific loyalty deviance portrayed in the vignettes, is evaluated with impartiality in the US.

Further supporting evidence for our hypothesis is found in **German sample**: This cultural entity consistently shows impartial tendencies and our analyses suggest that it makes no difference whether a stranger, a family member or a member of the in-group is affected negatively by a particular act of loyalty transgression.

Interestingly, the data show that the particularism part of our hypothesis has to be regarded as refuted for the **Japanese group**. For all dependent variables, we observe consistently impartial tendencies in the JP-sample.

Finally, the results from the **Egyptian sample** also refute our particularism hypothesis: The analyses reveal solely impartial tendencies in the context of the specific loyalty deviance.

Drawing on the findings from all four samples, we can identify a cross-cultural tendency: Betraying someone and publicly undermining them (i.e., specific loyalty deviance) is evaluated relatively similar across various cultures when focusing on different social relationships. Based on the three different levels of the social relationship dimension, we could not identify any significant effect. Although we were right about our hypothesis for the two WEIRD samples,

Table 57: Loyalty deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific Deviance *Loyalty*	USA n = 332	Samples Compared (pairwise) ††	Japan n = 299	Samples Compared (pairwise)	Germany n = 371	Samples Compared (pairwise)	Egypt n = 341	N = 1343
Relevance	5.019 (.114)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.906 (.130)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.294 (.120)	GER vs. EG *n.s.*	5.011 (.161)	
Judgment	5.937 (.072)	US vs. JP ** US vs. GER *n.s.* US vs. EG ***	5.609 (.082)	JP vs. GER ** JP vs. EG ***	6.003 (.076)	GER vs. EG ***	6.574 (.102)	
Shame	5.885 (.068)	US vs. JP *** US vs. GER *n.s.* US vs. EG ***	5.469 (.077)	JP vs. GER *n.s.* JP vs. EG ***	5.693 (.071)	GER vs. EG ***	6.612 (.095)	
Guilt	5.874 (.066)	US vs. JP *** US vs. GER * US vs. EG ***	5.514 (.076)	JP vs. GER *n.s.* JP vs. EG ***	5.645 (.070)	GER vs. EG ***	6.548 (.094)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

Table 58: Who is harmed? The (non-adjusted) effect of different social relations on *loyalty* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Loyalty* Deviance	USA	95% Conf. Interval	Germany	95% Conf. Interval	Japan	95% Conf. Interval	Egypt	95% Conf. Interval
	n = 332		n = 371		n = 299		n = 341 Δ	
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	.164 (p=.508; .248)	-.3234865 .6534765	-.061 (p=.795; .235)	-.5234394 .4012333	.164 (p=.508; .248)	-.3234865 .6534765	-.016 (p=.948; .251)	-.5101089 .4775565
- In-Group	.319 (p=.203; .251)	-.1727794 .8123337	.088 (p=.704; .233)	-.369526 .5467145	.319 (p=.203; .251)	-.1727794 .8123337	.327 (p=.181; .244)	-.1523161 .8073182
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.119 (p=.449; .158)	-.1908507 .430611	.014 (p=.922; .149)	-.279472 .308727	.119 (p=.449; .158)	-.1908507 .430611	.244 (p=.127; .160)	-.0693405 .5589292
- In-Group	.233 (p=.145; .159)	-.0802366 .5464095	.062 (p=.673; .148)	-.2287092 .3541259	.233 (p=.145; .159)	-.0802366 .5464095	.245 (p=.115; .155)	-.0598505 .5505882
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.203 (p=.170; .148)	-.0873857 .4944164	.072 (p=.603; .140)	-.2024052 .3482568	.203 (p=.170; .148)	-.0873857 .4944164	.129 (p=.388; .149)	-.1647596 .4234159
- In-Group	.172 (p=.249; .149)	-.1209289 .4657267	.043 (p=.755; .139)	-.2294039 .3162366	.172 (p=.249; .149)	-.1209289 .4657267	.108 (p=.456; .145)	-.1770642 .3944182
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	.232 (p=.110; .145)	-.0527106 .5179462	.003 (p=.979; .137)	-.2664377 .2736756	.232 (p=.110; .145)	-.0527106 .5179462	.067 (p=.645; .147)	-.2206126 .3562956
- In-Group	.223 (p=.128; .146)	-.0641637 .5112537	.119 (p=.380; .136)	-.1478971 .3872908	.223 (p=.128; .146)	-.0641637 .5112537	.148 (p=.300; .142)	-.1319908 .428544

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. **ΔNote:** The total sample size for the *loyalty* vignettes, encompassing the four sub-samples, is N = 1,343.

our particularism assumption is rejected for the Japanese and the Egyptian sample. The OLS models of the loyalty vignettes comprise a test family of 30 tests that relate to the *culture* variable. Since we could not find any significant effects within the context of the social relationship dimension, we do not have to correct for multiple testing on these effects and the cross-cultural tendency that has already been identified remains valid.

Furthermore, no significant deviations from the sample-specific interaction effects could be determined in the comparison of the four groups in our study. In the following, our focus is on vignettes in which a person shows a lack of respect for another person and publicly insults them, and so, we are focusing on an act of specific deference deviance.

6.4.7. MDFS Analysis VII: Deference Deviance

Finally, we come to the seventh and last moral domain which we will examine in the context of the Moral Deviance Factorial Survey. We now turn to *specific deference deviance* and, as before, first inspect the baseline model. Across the four models of the dependent variables, the findings yield that the *pathogen prevalence* variable, the *place of living* variable, the *denomination*, and the *age* variable consistently show no significant effect. We excluded these variables to obtain the adapted model on which we base the estimations of the margins.

In *Figure 29* we see an already familiar picture: judgment and relevance do not seem to be synonymous as far as the extent of the evaluation of moral deviance is concerned. This evidence lends further support the **relevance/judgment hypothesis**. Moreover, a clear pattern can be seen in *Figure 29* as deviance judgment takes the highest value in the comparison of the dependent variables across the four cultural entities. This value is followed, with the exception of the US-sample, by shame and guilt attribution. Finally, after a gap to the mentioned values, we find the deviance relevance margins with the comparatively lowest values. Interestingly, the US-sample again shows slightly higher values for guilt attribution than for shame attribution. Otherwise, the distribution of the margins for the four study groups as a whole shows a fairly homogeneous picture, apart from minor differences. Only the EG-sample seems to be a small outlier in the sample comparison in terms of judgment, shame and guilt.

The latter statement is supported by results from the pairwise sample comparisons (see: *Table 59* further below). What is also noteworthy is the empirical fact that the two WEIRD samples do not differ from each other, but have significantly lower margins than the EG-sample with regard to specific deference deviance judgment (partly) as well as shame and guilt attribution. Moreover, when comparing the samples with each other, the JP-sample is located

in the middle of the polarization just described: neither the EG-sample nor the US- or GER-sample exhibit a significant difference in comparison to the margins of the Japanese group.

Figure 29: MDFS — Valuations of deference deviance across cultures
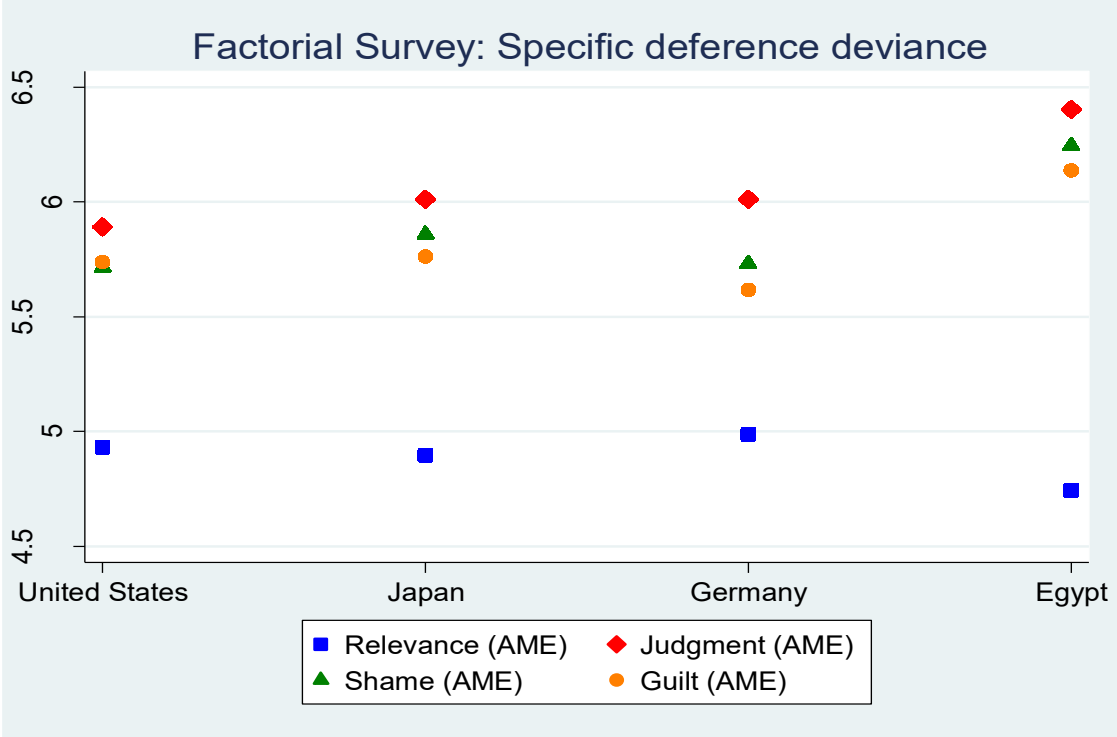
In approaching our main research interest, the **impartiality/particularism hypothesis**, we can note the following for the **US American sample**: exclusively impartial tendencies are found across the dependent variables (***Table 60***). Hence, evidence suggests that our hypothesis is confirmed for this sample.

Turning to the **GER-sample** we find impartial tendencies on the relevance and judgment variable and also impartial tendencies for the attribution of moral emotions in the in-group vs. stranger comparison. However, beyond these measures, significant and negative effects are identified for the shame and guilt variables when a *family* member is harmed by the act of deviance. These effects suggest that the imposition of aversive moral emotions in the face of an act of public disrespect and insult, i.e., an act of specific deviance from the moral domain of deference, applies even more strongly beyond familial group ties in Germany. A further indication in the same direction is provided by the judgment-model and the *family* variable, which not only has a negative algebraic sign but also a p-value of $p = 0.072$. Several effects observed in the GER-sample thus go somewhat further than our impartiality hypothesis.

Once more evidence contradicts the particularism hypothesis in the Japanese group. The **JP-sample** displays impartial tendencies on all four dependent variables. Thus, our hypothesis is refuted and instead of moral particularism we observe solely impartial tendencies —

Table 59: Deference deviance relevance, judgment, shame, and guilt margins across cultural groups

Vignette: Specific Deviance *Deference*	USA n = 328	Samples Compared (pairwise) ††	Japan n = 319	Samples Compared (pairwise)	Germany n = 389	Samples Compared (pairwise)	Egypt n = 321	N = 1357
Relevance	4.931 (.116)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *n.s.*	4.895 (.139)	JP vs. GER *n.s.* JP vs. EG *n.s.*	4.986 (.127)	GER vs. EG *n.s.*	4.744 (.199)	
Judgment	5.890 (.078)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG ***	6.013 (.094)	JP vs. GER *n.s.* JP vs. EG *n.s.*	6.011 (.086)	GER vs. EG *n.s.*	6.404 (.135)	
Shame	5.714 (.074)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG ***	5.858 (.088)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.730 (.081)	GER vs. EG **	6.244 (.127)	
Guilt	5.738 (.074)	US vs. JP *n.s.* US vs. GER *n.s.* US vs. EG *	5.762 (.089)	JP vs. GER *n.s.* JP vs. EG *n.s.*	5.618 (.081)	GER vs. EG **	6.139 (.127)	

Note: All significance levels are corrected for multiple testing (i.e. for a total of 6 group comparisons per dependent variable) using Holm-Bonferroni method. Standard errors are shown in parentheses.

Table 60: Who is harmed? The (non-adjusted) effect of different social relations on *deference* deviance relevance, judgment, shame, and guilt across cultural groups

Vignette: Specific *Deference* Deviance	USA	95% Conf. Interval	Germany	95% Conf. Interval	Japan	95% Conf. Interval	Egypt	95% Conf. Interval
	n = 328		n = 389		n = 319		n = 321 Δ	
Relevance								
Social Relation								
- Stranger	0 (base) †		0 (base)		0 (base)		0 (base)	
- Family	-.263 (p=.276; .241)	-.7375556 .2109599	-.239 (p=.280; .221)	-.6732672 .1946799	.221 (p=.378; .251)	-.2719141 .7154116	-.393 (p=.110; .245)	-.875292 .0887125
- In-Group	.060 (p=.803; .242)	-.4149735 .5362496	-.307 (p=.172; .225)	-.748942 .1339828	.012 (p=.959; .242)	-.4637069 .4884097	-.318 (p=.191; .243)	-.7966889 .1595686
Judgment								
Social Relation								
- Stranger	0 (base) ††		0 (base)		0 (base)		0 (base)	
- Family	.001 (p=.992; .163)	-.3195455 .3226492	-.269 (p=.072; .149)	-.5634833 .0241624	-.224 (p=.188; .170)	-.5589067 .1095646	-.006 (p=.969; .166)	-.3328913 .3197903
- In-Group	.124 (p=.449; .164)	-.1976643 .4463636	-.168 (p=.268; .152)	-.4678506 .1299358	-.008 (p=.958; .164)	-.3310347 .3135982	-.017 (p=.916; .165)	-.3410656 .3063709
Shame								
Social Relation								
- Stranger	0 (base) †††		0 (base)		0 (base)		0 (base)	
- Family	.119 (p=.438; .154)	-.1831843 .4224294	-.308 (**p=.029**; .141)	-.5859225 -.0317505	.002 (p=.987; .160)	-.3125399 .3178536	-.031 (p=.840; .156)	-.3394577 .2760456
- In-Group	.112 (p=.466; .154)	-.1908595 .416483	-.236 (p=.100; .143)	-.5180288 .0457062	.160 (p=.300; .154)	-.1434391 .4644739	-.080 (p=.606; .155)	-.3856157 .2249412
Guilt								
Social Relation								
- Stranger	0 (base)		0 (base)		0 (base)		0 (base)	
- Family	-.003 (p=.980; .154)	-.3075027 .2997495	-.391 (**p=.006**; .141)	-.669594 -.1139228	-.026 (p=.867; .161)	-.3429996 .2890993	.019 (p=.901; .157)	-.2889365 .3282319
- In-Group	-.049 (p=.751; .155)	-.3537291 .2552564	-.206 (p=.153; .144)	-.4888033 .0764568	.053 (p=.731; .155)	-.2513505 .3582071	-.007 (p=.959; .156)	-.3140471 .2981616

†**Note:** The term *base* denotes the reference category (stranger). ††**Note:** First, the coefficients of the sample-specific interaction effect are given. The (unadjusted) p-values are then given in the following parentheses, followed by the standard errors. Significant (unadjusted) p-values are marked in bold. †††**Note:** A total of *31* tests, related to the variable *culture*, belong to a test family in the respective property OLS models. After correcting for multiple testing using the Holm-Bonferroni method no effect remains under the classical significance level of p < 0.05. One should consider the size of the test family, the corresponding conservative p-value correction and the sample sizes of the individual sub-samples in order to contextualize the Holm-Bonferroni correction appropriately. **ΔNote:** The total sample size for the *deference* vignettes, encompassing the four sub-samples, is N = 1,357.

deference to others is valid in the Japanese social context, independently of social affiliations and group ties.

Also, the results found for the **Egyptian sample** yield no support for the particularism part of our hypothesis. In fact, we see again solely impartial tendencies across the dependent variables.

Overall, we observe by and large a cross-cultural tendency towards impartiality. Acts of disrespect are largely evaluated independently of *who* is harmed across cultures: Whether a stranger, a family member or a member of the in-group is treated disrespectfully and publicly insulted does not appear to play a major role in the valuation of the deviant act. Although our hypothesis was confirmed for the United States and partly for Germany, we were wrong about the Japanese and Egyptian groups examined. Next to tendencies of impartiality, an interesting pattern was found in Germany, suggesting that a violation of deference would be more harmful for cooperative relationships beyond familial ties. The OLS models comprise a test family of 27 tests that relate to the variable *culture*. When correcting for the possibility of alpha-error cumulation using the Holm-Bonferroni method, it must be noted that none of the social relation effects is significant anymore. In sum, both *the unadjusted* and *adjusted p-values* indicate mainly the aforementioned tendency towards cross-cultural impartiality in the specific deference vignettes. Our hypothesis as a whole, which is drawing heavily on self-construal differences, differences in collectivism-individualism, and differences in kinship institutions, thus seems to be too simplistic.

Taking in a comparative perspective we find a significant and negative departure from the *US-sample* interaction effect on the shame variable and the social relationship category *family* when compared to the German sample: Coeff. = -0.428; Std. Err. = 0.208; p = 0.041. As indicated by the negative algebraic sign, the effect suggests that the GER-sample imposes lower shame attribution than the US-sample in the case of specific deference deviance when a family member is harmed by the moral breach. Besides the significant departure between the GER- and US-sample interaction effects, no further effect with a p-value of < 0.05 was found in the comparisons across the four samples of our study.

6.5. Discussion: Insights from the Moral Deviance Factorial Survey (MDFS)

The analysis of the Moral Deviance Factorial Survey (MDFS) was able to provide additional insights to our previous investigations and revealed supporting as well as contradicting evidence in regard to our hypotheses. Especially the results related to the impartiality/particularism hypothesis call for interpretation and explanation. Nevertheless, also the limitations of our study should of course not be ignored and several unanswered questions remain. Hence, we will comprehensively discuss and classify the MDFS findings in the following.

First and foremost, we turn to the postulated difference between the judgment of moral deviance and the relevance of moral deviance, and address accordingly the results in regard to the **relevance/judgment hypothesis**. We found a gap between *relevance* and *judgment* in the extent to which specific, morally deviant actions are evaluated across all of the 7 moral domains examined. Our results suggest consistently that the extent to which an action is judged as morally wrong does not necessarily correspond to the same extent of deviance relevance. Rather, people can clearly identify an action as morally wrong and yet attribute a different weight of moral relevance to it. Correlative analysis and findings across all further models estimated support this statement and demonstrate that it holds true across cultures. Thus, we see ourselves confirmed in the theoretical position that separates *moral deviance relevance* from *moral deviance judgment*. Graham and colleagues (Graham et al., 2011; Atari et al., 2022a) argue that the concept of moral relevance is probably more intertwined with explicit processes of moral cognition and already anticipate a difference between judgment and relevance in the realm of morality: "[R]eports of moral relevance are best understood as self-theories about moral judgment, and they are likely to be concordant with explicit reasoning during moral arguments" (Graham et al., 2009, p. 1031). We propose that moral phenomena can also appear intuitively relevant to us (see: **Chapter 3 and 4**). This does not mean, however, that moral relevance is a concept that only concerns implicit processes of moral cognition. On the contrary, we further argue that it encompasses explicit and implicit cognition and depends on the way it is measured. The scale that we invented and introduced in a different chapter, i.e., MaC-DRS, was designed to approach mainly implicit moral tendencies. The Moral Deviance Factorial Survey, by way of contrast, was designed to capture mainly deliberate processes of moral cognition. Unlike MaC-DRS, the Factorial Survey instrument does not aim at general tendencies, but at concrete, contextualized acts of moral breaches. With this in mind, our results show that deliberate processes of judgment and relevance differ in the extent to which they are

attributed to specific acts of moral deviance. Wrongness judgments cannot be transferred one-to-one to the extent to which a phenomenon is perceived as relevant to a person's sense of morality. ***Thus, moral relevance is not moral judgment***. So, although we can identify a moral transgression as such, and judge it accordingly as wrong, this does not necessarily mean that we also experience this transgression to the same extent as subjectively relevant to our sense of morality. To illustrate our point, we will draw once again on an example that we have already used several times throughout this work: Imagine that someone steals a piece of your chocolate bar. Next, imagine that someone steals your wallet, which, in addition to cash, also contains a photo of your family, your ID card, your driver's license, your credit card and similar items. Both "scenarios" are violations of the property domain of morality, and stealing is morally wrong, but doesn't the wallet scenario seem more relevant to you? In fact, we can also make gradations about the extent to which we judge an action to be wrong, but we also make such gradations in the context of relevance, and the latter, as our data shows, does not have to be equivalent to the extent of the moral judgment. ***Overall, this study provides cross-cultural empirical evidence on the basis of correlative analyses and average marginal effects suggesting to treat moral deviance relevance as something different than moral deviance judgment.***

> **Moral deviance relevance/judgment hypothesis**: We predicted that the extent of the relevance of specific acts of moral deviance and the extent of the judgment about specific acts of moral deviance do not (necessarily) coincide. **In the light of the empirical evidence presented, we regard this hypothesis as confirmed.** ✓

In our view, the evidence presented is a step in the direction of giving more depth to and partly reviving the concept of *moral relevance*, which, for example, was dropped from the latest scale of the Moral Foundations Theory (Atari et al., 2022a). However, the contribution of the concept of moral relevance itself to investigations of the human moral mind is certainly still an open question. We will discuss this and other considerations regarding moral relevance further below.

What about **cross-cultural similarities and cultural differences** in the analyses of the MDFS? We believe that what is interesting is the empirical fact that our investigations of **specific deviance relevance** across 7 different moral domains revealed only 2 significant differences between the samples. It should be noted, though, that we carried out a total of 42 comparisons between the four samples in the context of deviance relevance across the 7 domains examined. As far as the relevance of the scenically contextualized vignettes limited to a single act of deviance is concerned, we thus find mostly cross-cultural similarities. In addition, also another general tendency can be noted across the moral domains and cultural groups

examined: The deviant acts depicted in the vignettes are evaluated with a slight tendency towards relevance. This conclusion can be drawn from the 7-point response scale of our dependent variable and from the AME value of 4.265, which is the lowest value among all relevance margins. In summary: Regardless of the specific act of deviance, transgressions in seven domains of cooperation are considered meaningful in the context of individual conceptions of morality – deviations from cooperative standards are morally relevant –, and this tendency applies across cultures.

In the context of **specific deviance judgment** there are no longer as many similarities as were found for relevance. The 42 pairwise comparisons between the four cultural samples, across the 7 moral domains examined, reveal 19 significant differences between the groups in the context of moral judgment. Despite cross-cultural differences, we can nevertheless also recognize a general tendency. The lowest AME value for deviance judgment is 4.307 (heroism), but all other judgment AMEs show values of > 5.1 on the 7-point response scale.[144] The findings thus suggest a cross-cultural inclination: Several acts of specific deviant behavior in different domains of cooperation are judged as wrong across cultures. These results complement the study by Curry and colleagues (2019a). Not only is compliance with different moral domains considered *good* across cultures, but violations of different moral domains are also considered as relatively *wrong* across a heterogenous set of cultures, as proven in this study.

We experience the moral emotions of **shame** and **guilt** (for the most part) as aversive feelings that motivate us to refrain from actions that evoke these feelings (Haidt, 2003; Tangney et al., 2007; for a cross-cultural perspective, see also: Wong & Tsai, 2007). In this respect, moral emotions have an inherent behavior-regulating element supporting human cooperation (Tomasello & Vaish, 2013). When we sense, think, and signal that someone should feel ashamed or guilty for their behavior, this can be interpreted as an attempt to impose costs on the person in question in order to ultimately motivate them to refrain from the respective behavior in the future (Horne & Cutlip, 2002; Fehr & Fischbacher, 2004; Chudek & Henrich, 2011). Moreover, it is also an expression of the fact that the behavior that provokes shame and guilt attribution represents a violation of socially accepted moral standards (Haidt, 2003; Wong & Tsai, 2007). We used MDFS to determine the extent to which the person committing a moral transgression in the vignette scenarios should feel ashamed and guilty. In other words, we measured the extent to which certain moral emotions were attributed as response of the deviant action depicted in the vignettes. Our results show that the four cultural samples in our study have much in common

---

[144] The response format of the *judgment* variable ranges from 1 (right behavior) to 7 (wrong behavior), with all intermediate levels labeled and value 4 representing the neutral midpoint of the response format.

and yet also differ in the attribution of guilt and shame. Overall, we found 23 significant differences for each of the two emotion attribution variables in the pairwise comparisons of the samples. The Egyptian sample in particular tends to attribute more shame and guilt to specific acts of moral deviance than the other samples in the cross-cultural comparison. Consequently, a large proportion of the cultural differences recorded in the present analyses are attributable to the pairwise comparison of the JP-, US- and GER-samples with the EG-sample. This result clearly demonstrates the importance of not limiting cross-cultural research to WEIRD samples or comparisons between the US and Japan, as we would otherwise miss much of the facets of human (cultural) realities and psychologies (Henrich et al., 2010a). We will come back to this point a bit further below.

We found minor indications that deviant behavior is attributed slightly more guilt than shame in the US-sample. In four out of seven models, the guilt margins are slightly higher than the shame attribution margins for the US American group. The opposite pole to the US is found in the EG-sample. In the Egyptian group, shame attribution dominates over guilt attribution. Corresponding margins are consistently higher across all measured moral domains and are even considerably more pronounced in the case of specific trustworthiness deviance. In both the Japanese and German samples, either shame attribution slightly predominates, or the attribution of both moral emotions coincides. Nevertheless, it should be noted that across all four cultural samples there are by and large and with a few exceptions no major differences between shame and guilt attribution *within* the groups themselves. In most cases, the extent of shame and guilt attribution as a reaction to moral deviance coincides almost perfectly. This empirical fact is important against the background of the self-construal prevailing in the respective cultural samples, as well as within the framework of the prevailing cultural logics. From the cosmos of research within the framework of the self-construal approach, assumptions emerged that associate the independent mode of selfhood primarily with guilt and link the interdependent mode of selfhood more with shame (Markus & Kitayama, 1991; Wong & Tsai, 2007). Furthermore, shame is associated more with honor and face cultures, and guilt more with dignity cultures (Leung & Cohen, 2011). Of course, the assumptions, and also some supporting evidence, go beyond mere associations and refer, among other things, to cross-culturally different functions, triggers and consequences of shame and guilt (see e.g.: Enke, 2019; Henrich, 2020; especially pp. 34-36; pp. 198-204). Nevertheless, our results show that when attributing feelings of shame and guilt to acts of moral deviance, these two moral emotions are mostly and across the cultural entities examined attributed to a similar extent. The findings across our models support this statement, although not absolutely, but to a large extent. In our

analyses, we found accordingly only minor indications of the association of shame and honor (see: EG-sample results) as well as of dignity and guilt (see: US-sample results). In line, our findings suggest overall a *relatively* similar attribution of shame and guilt within the four different cultural entities, although the samples that we investigated differ with regard to cultural logics (Uskul et al., 2019; 2023) and the construal of the selfhood (Markus & Kitayama, 1991; 1998; 2010).[145] Our results could thus indicate that we are dealing with a more complex phenomenon than previously assumed in the context of moral emotions. It may therefore be advisable to expand the question of whether guilt or shame "dominates" a cultural entity by asking in *which* (specific) context? This interpretation can be derived from our results in the field of specific moral deviances, which, to a large extent, suggest similar tendencies in the attribution of moral emotions across the four highly heterogeneous cultures examined.

Regarding the attribution of moral emotions, the AMEs of 4.407 (heroism) for shame and 4.359 (heroism) for guilt are the lowest values we were able to determine. If we disregard the outliers for heroism, we find that all other margins in the four samples are above the value of 5 for shame and all other margins are above 4.9 for guilt. The average marginal effects thus indicate a cross-cultural tendency to incentivize behavior emotionally as a consequence of the perception of moral deviance. In drawing on the findings from the pairwise comparisons, we conclude that there are indeed differences in the extent of guilt and shame attribution *across* the cultural entities in our study. At the same time, however, there also appears to be a general tendency indicating that shame and guilt are largely attributed to a similar extent *within* cultures. Moral transgressions thus trigger a tendency in the social environment to impose behavior-regulating and behavior-altering emotions on the deviant actor (Haidt, 2003). Our results suggest that people largely react to the perception of moral deviance with the same (emotional) means across cultures: We can reasonably interpret the data to suggest that an emotional cost is imposed on the deviant actor to adjust their behavior, refrain from the deviant act and behave morally compliant in the future (Tangney et al., 2007).

Henrich (2020) notes: "we have evolved genetically to learn adaptively in ways that calibrate our minds and behavior to the environments we encounter" (p. 63). We follow Henrich's view. With Japan, the USA, Egypt and Germany, we examined a set of four

---

[145] As described, this tendency is *not* absolute. Taking a look at the Egyptian sample, we see e.g. in the context of specific trustworthiness deviance that guilt and shame attribution can also differ intra-culturally in their extent. Thus, despite the suggested cross-cultural tendency to attribute guilt and shame to a relatively similar extend, there is also evidence, as outlined, that especially in Egypt (honor culture) shame seems to play a greater role than guilt. This finding should altogether, nonetheless, not obscure the fact that in most models only slight differences were found between the attribution of shame and guilt in the EG sample, which is why the established cross-cultural tendency undoubtedly has a value from our perspective.

heterogeneous cultural entities and take now a brief, sample-based look at the **cross-cultural differences** found in the evaluation of specific acts of moral deviance. In this context, it can be generally stated that the majority of cultural differences found are attributable to comparisons with the Egyptian sample. Taking into account all results for the dependent variables deviance relevance, judgment, shame, and guilt, the EG-sample shows 18 significant differences compared to the US-sample, 16 compared to the GER-sample, and 9 significant differences compared to the JP-sample. In addition, 9 significant differences in the evaluation of moral violations were also found in the pairwise comparison of the JP-sample with the two WEIRD samples. The latter in turn differ comparatively little from each other. The margins (AMEs) of the valuation of specific moral deviance differ significantly in only 5 of the pairwise comparisons between the US American and German sample. Since we consider margins as the basis for our results, it should not be forgotten, in view of the cross-cultural differences, that we have kept a considerable set of covariates constant and controlled for them, and yet we still find the cultural differences described. We can therefore conclude that culture does indeed influence not only implicit moral tendencies but also deliberate tendencies of moral cognition — the evidence presented highlights the fact that the human moral mind is calibrated in a culturally specific way.[146] Moreover, our findings demonstrate once again evidently how important it is to further overcome the (too) narrow focus on WEIRD samples in investigations on the human mind (Henrich et al., 2010a). In order to even begin to grasp the diversity of human culture, cultural similarities and cultural differences, it is still crucial to expand the sample base of cross-cultural studies beyond the focus on WEIRD samples (Apicella et al., 2020). Our results, e.g.

---

[146] An interesting sidenote arises when we compare the scale-based MaC-DRS results with those obtained from the analyses of the MDFS. Based on the various instruments that we have used in the last three chapters to measure morality, we can observe cross-cultural differences in moral tendencies in both implicit and explicit moral cognition. However, we wonder whether intuitive processes tend to reflect cultural differences in the context of moral deviance more than deliberate processes? Our findings reveal in the context of deliberate moral evaluation, i.e., the Factorial Survey, significantly more cross-cultural *similarities* than in the context of intuitive moral evaluations as measured via MaC-DRS. Perhaps this circumstance indicates that cross-cultural *differences* in the field of morality are primarily found among the automatic, intuitive reactions of our moral mind, and that deliberate processes of moral cognition are cross-culturally more similar. Let us briefly and exemplarily focus exclusively on the specific *relevance* variable of the Factorial Survey and compare these findings with those from MaC-DRS (see: ***Chapter 4***). First, we must note that the Factorial Survey analyses, which focus on *deliberate* processes of moral cognition in 7 distinct moral domains, yielded only two significant group differences. The MaC-DRS analyses, by contrast, compare 8 moral domains and focus on *intuitive* moral (deviance) relevance. Analyses with the latter instrument yielded a total of 28 significant group differences. A *speculative* consideration can be derived from this difference in the number of significant results: We wonder whether, in particular, the implicit cognitive processes and intuitive reactions (Haidt, 2001; Esser, 2002a; 2010; Haidt & Joseph, 2007; Graham et al., 2013; Greenwald & Lai, 2020) reflect the socially embedded and repeatedly positively and negatively sanctioned incorporation of cultural mandates (Kitayama et al., 2009; Kitayama & Imada, 2010; Park et al., 2016: Henrich, 2016; Heyes & Moore, 2021)? From our perspective, it could be a worthwhile task for future research to investigate whether reactions that are located in the intuitive realm of moral gut feelings reveal more cultural peculiarities and thus cross-cultural differences than studies that focus on deliberate moral processes. Should this *speculative* thesis be given empirical weight, an appropriate, explanatory theory would of course be needed.

regarding the Egyptian sample, clearly underscore this point and expand our insights of the human moral mind with data from an underrepresented cultural group.

After these conclusions, we will finally turn to our main concern, the general discussion of the **impartiality/particularism hypothesis** and the MDFS dimension **social relation** with the levels: *stranger* (reference category), *family*, and *in-group*. Recall, the hypothesis that we investigated is as follows:

> **Modified impartiality/particularism hypothesis**: We hypothesized that the JP- and EG-cultural samples tend towards moral *particularism*, i.e., we expected a tendency to rate moral deviance that harms a stranger as less severe (relevance, judgment, shame, and guilt) as compared to deviance towards a member of the in-group or family. In contrast, we predicted that the GER- and US-cultural samples tend towards *impartiality*, i.e., we expected a tendency to rate moral deviance that harms a stranger as equally severe (relevance, judgment, shame, and guilt) as deviance towards a member of the in-group or family.

We focused in our analyses on both the presentation of the uncorrected findings and the reference to the Holm-Bonferroni corrected p-values. Relying on the *non-p-value-adjusted findings*, the overall relationship between moral impartiality, moral particularism and culture is more complex than expected. Our hypothesis draws mainly on theory and findings in the realm of collectivism-individualism, self-construal, and kinship institutions and infers, in accordance with the theory, a binary assumption. This assumption is on the one hand concerned with impartial tendencies for the two WEIRD cultural entities characterized by individualism, independence in self-construal and low kinship intensity, i.e., the US American and the German sample, and on the other hand with particularistic tendencies for the two cultural entities characterized by collectivism, interdependence in self-construal and high kinship intensity, i.e., the Japanese and the Egyptian sample. Although we observed some supporting evidence, large parts of our findings run yet counter to our binary hypothesis, indicating that there is more to the story of impartiality/particularism than predicted. As an overarching conclusion, we interpret the uncorrected findings in such a way that we see a more complex field of associations suggesting that ***moral impartiality/particularism is likely culture dependent and moral domain as well as social relation specific***. However, when we consider the *results after alpha-error correction*, we can only conclude that we are exclusively observing impartial tendencies. The latter results would lead us to consider the (WEIRD) half of our hypothesis as confirmed and the EG- and JP-sample part of the hypothesis as refuted, indicating ultimately also a more complex pattern that cannot be easily derived from theories and findings on which we built our hypothesis. Hence, we see a lack of understanding and the ***need for further research and improvements in theory***. We will now discuss our insights and finally ask, based on the

empirical findings, whether something is perhaps afoot in the sociocultural world that our current theories may not fully grasp and explain — do our findings point in parts to sociocultural change? We will start by taking a closer look at the non-adjusted findings, as these are more complex. All results *not* corrected for multiple testing are summarized in ***Table 61***. After that, we will briefly touch on the p-value-adjusted findings.[147]

First and foremost, we find consistent impartial tendencies across all four dependent variables in the **US-sample** for *only three* of the seven moral domains examined. The results yield *full* empirical support of our impartiality hypothesis in the specific *loyalty* and *deference* domains, and suggest furthermore that *reciprocity* is either evaluated impartial (in-group vs. stranger) or with a special emphasis of tit-for-tat beyond the group focus (family vs. stranger). More generally, the in-group does not appear to be the main trigger of particularistic inclinations in the US. Looking at the US-sample in-group vs. stranger interaction effects, we find only hypothesis-consistent impartial tendencies, with the exception of the moral domain *trustworthiness*. The US evaluates moral violations towards a stranger or an in-group member thus largely impartial. Over and above impartiality, the findings demonstrate that a violation of *reciprocity* that harms a stranger carries more weight on the judgment and shame variables than the same type of deviance in interactions with the family. In the US-sample, therefore, particular importance is attached to compliance with the tit-for-tat principle outside the family, i.e., in *conditional* cooperative interactions (Axelrod, 1986; Kurzban et al., 2015; Henrich & Muthukrishna, 2021; Muthukrishna, 2021; Romano et al., 2022) with strangers. From our point of view, this effect can be well located theoretically, because it may suggest that reciprocity is especially important for building cooperative relationships with people with whom one is not deeply connected through family ties. We interpret the findings described as follows: The identified tendencies indicate that cooperation beyond established and intimate long-term relationships, such as those that prevail in the family, can be particularly important for tapping into and maintaining new networks of cooperative exchange relationships. Furthermore, our findings point to socio-cultural conditions in the US that seem to foster corresponding tendencies. The background of this psychological disposition is likely in parts attributable to an element that is particularly found in modern societies, i.e., high market integration (Henrich et al., 2005; 2010b):

---

[147] We are aware that our results may also be due to our measurement instrument respectively the operationalization used. We will refer to this below in the context of limitations.

> "Well-functioning impersonal markets, in which strangers freely engage in competitive exchange, demand (…) market norms. Market norms establish the standards for judging oneself and others in impersonal transactions and internalization of motivations for trust, fairness, and cooperation with strangers and anonymous others" (Henrich, 2020, p. 293).

In a context of market norms, conditional cooperation is integral for maintaining and promoting one's own reputation. Reputation, in turn, affects the social success of the individual and is important for promoting well-being and socioeconomic positioning (Ormel et al., 1999; Henrich, 2020, pp. 287-321; Henrich & Muthukrishna, 2021; Baumeister, 2022, pp. 104-113). If such market norms are present in the sociocultural ecology, as it is the case in the US, a corresponding calibration of the moral mind can be considered adaptive. Large parts of impartial moral tendencies, in addition to a special focus of tit-for-tat compliance in interactions with strangers, are the resulting consequence. However, *conditional* cooperation is apparently not demanded in the family. On the contrary, it should be noted that the (modern) family in WEIRD cultures can be regarded as a social institution in which people encounter and recognize each other as individuals, affirm their unique identity and interact beyond purpose-means relations (Huinink, 1995; Honneth, 2012). According to Johannes Huinink (1995), a special mode of action prevails in familial interactions of modern families, i.e., dialogical action (pp. 92-93). Furthermore, Huinink assumes a socially constituted basic need for individual self-founding, for subjective identity assurance, which cannot be provided by formal market relationships (pp. 87-98). What is important, however, is that family relationships, in which actors engage in dialogical action, can provide precisely this self-founding. Violations of reciprocity as such are not important in these modes of interaction (p. 103), because it is precisely by forgiving possible breaches of reciprocity that the actors signal the seriousness and durability of a personal relationship between unique subjects that are not interchangeable. If we follow this logic, market norms and conditional reciprocity apply primarily in impersonal interaction with strangers and to some part with in-group members. However, the family in modern WEIRD cultures (pp. 114-128) is a place (of cooperation) that organizes interaction without "you scratch my back and I will scratch yours", because here, by contrast, the principle "I do something for you because I am deeply connected to you as a person and recognize you as such" tends to prevail. Seen in this light, dialogical action can be understood as an integral process of individualistic subjectivation practices, which is why it seems largely fitting that we find the negative reciprocity interaction effects in the context of the family in the WEIRD sample of the USA.

Beyond the findings that are by and large in accordance with our hypothesis for the US-sample, we also found tendencies that partly refute our theorizing in four of the seven moral domains examined. As already mentioned, the trustworthiness domain is an outlier among the otherwise impartial effects in the US in-group vs. stranger comparisons. Interestingly, we found in the US-sample that moral transgressions against the in-group are only given special weight in the context of *trustworthiness*, which reveals a particularistic tendency for this moral domain. One may reason that it is the high degree of relational mobility in the United States that requires individuals to pay particular attention to reliable behavior when dealing with in-group members in order to reassure and solidify, and eventually, to maintain cooperative relationships with friends (Huinink, 1995, pp. 106-109). Thomson et al., (2018) note in this context: "as relational mobility increases, it becomes more adaptive to actively invest in building interpersonal relationships" (p. 7526). Regulating one's own behavior in order to prove oneself as reliable in interactions with in-group members can be interpreted as such an investment as it signals interest and effort put in the durability of the (cooperative) relationship.

The family can be considered the main relational component that drives particularistic tendencies in the United States of America, at least in several moral domains. The seven moral domains and the four dependent variables measured for each allow us to examine 28 interaction effects per family vs. stranger and in-group vs. stranger comparison. For the family variable, we were able to identify a total of 8 positive effects in the US-sample, which demonstrate that deviant behavior toward a family member is evaluated as more wrong and attributed higher emotional costs than the same act of deviance and the case of a harmed stranger (see: ***Table 61*** to be found below). In addition, two effects fall at the 10% significance level in the direction of particularistic tendencies, providing further indications that moral transgressions receive more weight when a family member is harmed. We find these tendencies of particular importance of moral conformity in family relations fully in the context of *fairness* deviance, mainly in the context of *heroism* deviance, also in the moral domain of *property*, and observe additionally an indication of this special importance in the *trustworthiness* domain. Sharing something in equal measure, standing up for each other in dangerous situations, and respecting each other's property therefore appears to be particularly important in the familial context of the US. A moral mind that is calibrated to attribute particular weight to violations in these moral domains within the context of long-term, intimate, family-based relationships can be understood as an expression of sociocultural requirements placed on the individual. The individual requires a corresponding psyche that allows them to navigate and act under given requirements of the social world without jeopardizing their relationship with the essential social safety net. In the

US, the family takes on the role of a social safety net (Seeleib-Kaiser, 2014; Cooper, 2020). In the context of the MaC-DRS analyses, we have already worked out that the institutional world of the US requires a moral mind calibrated to pay *intuitively* particular relevance to moral breaches in the family domain (see: ***Chapter 4***). With the Moral Deviance Factorial Survey (MDFS) findings, we can now supplement these insights by demonstrating that the moral psyche in the US is partly calibrated so that the moral circle (Waytz et al., 2019; Kirkland et al., 2023) prioritizes the family in specific acts of *fairness*, *heroism*, and *property* deviance, and thus cannot be considered fully impartial. When we consider the findings and corresponding p-values before Holm-Bonferroni correction, the significant effects together with the indications at the 10% significance level, we find particularistic tendencies in 13 out of a total of 56 (family vs. stranger; in-group vs. stranger) comparisons. Accordingly, our findings supplement Joseph Henrich's (2020) considerations on impersonal pro-sociality in WEIRD cultures, and refute our hypothesis for the US-sample in the respective comparisons (see: ***Table 61***). Overall, from our perspective, a complex picture emerges: The non-p-value adjusted findings in the United States suggest that ***moral impartiality/particularism is likely moral domain as well as social relation dependent***.

As in the USA, we also hypothesized an impartial moral psychology for the **GER-sample**. In parts, we encounter a similar and in parts a different picture of the empirical findings in our second WEIRD sample. For the similar results in the GER-sample, we refer to the same explanatory approach as in the US American sample. But we must also recognize differences between the samples in the context of our hypothesis, which call for a culture-specific interpretation. In five of the seven moral domains examined, our hypothesis was largely confirmed in the German sample, and we found impartial tendencies and thus evidence that speak against a special importance of self-regulation in support of group-oriented cooperation. Across all dependent variables, results yield either impartial tendencies in the evaluation of specific acts of moral deviance in the domains of *fairness*, *trustworthiness*, *reciprocity*, *loyalty* and *deference*, or evidence suggesting an even greater weight to corresponding moral violations when a stranger is harmed.

The GER-sample shows, as in the US-sample, indications that *reciprocity* violations in the family context are given less weight than in interactions with strangers. These indications point to *conditional cooperation* with strangers and *dialogical interaction* with family members, though, only supported at a 10% significance level. However, we find significant interaction effects in the context of the *deference* domain: Behaving disrespectfully towards a stranger is imposed with more emotional costs in Germany than the same act of deviance in

interaction with a family member. Respect and civility seem therefore particularly reserved for strangers and partly for interactions with the in-group. The special importance of strangers in the domain of deference may seem strange at first glance. However, the effect can be explained by recourse to the hierarchical social structure, soaked in formalities, which is found in German history not only in the corporate society (Elias, 1976; in particular: pp. 1-42; Schulze, 2008, pp. 1-200; in particular: p. 63), but also in more recent history (Piper, 2001; Houghton, 2009, pp. 46-56; Myers et al., 2010, pp. 220-259). Moreover, norms of civility apply in this context, which convey the respective relationship status of persons to one another. In this context let us draw on an example from German language. In German, a distinction is made between "Sie" as a polite form of address used respectfully for strangers and people of higher social status, and "Du", which is basically the same form of address but expresses a greater familiarity and closeness to the person being addressed. The symbolically conveyed intimacy of the relationship with another person can be expressed in Germany linguistically in this way but also drawing more generally on norms and practices of civility. If we relate this to our findings, the following can be stated: In the family, the tone can sometimes be rougher, deference must not always be followed, but this does (usually) not pose a major threat to the (cooperative) relationship. On the contrary, not beating around the bush, speaking one's mind, being frank, and maybe even a bit rude, can be a sign of an intimate relationship that expresses informality and closeness to the social other rather than a relationship characterized by distance, formalities, social stiffness, and *deference* (Huinink, 1995, pp. 97-102; Kotthof, 2003). Signals of intimacy through bluntness point to individualistic practices and the individualistic self-construal character in Germany (Markus & Kitayama, 1991; 1998; 2010; Kitayama et al., 2009). In this respect, the negative effect for the family identified in the GER-sample becomes understandable. In Germany, modesty and civility can in parts be interpreted as a symbol of social distance; they form the framework for formal, non-intimate interaction. These forms of interaction can certainly be a key to accessing further interaction and cooperation with strangers, because people treat each other with respect and deference, but in the WEIRD culture of Germany they are at the same time likely an expression of a non-intimate relationship status. Heightened attribution of shame and guilt when a stranger is harmed by the act of deference deviance indicate that especially the emotional forces of behavioral regulation are at work in this context (see: *Table 61*). Furthermore, Shame and guilt "function as an emotional moral barometer, providing immediate and salient feedback on our social and moral acceptability" (Tangney et al., 2007, p. 347). Therefore, it seems that in Germany it is neither socially nor

Table 61: Overview MDFS findings for the social relation dimension across cultural groups and moral domains

	Specific *Property* Deviance	Specific *Fairness* Deviance	Specific *Trustworthiness* Deviance	Specific *Heroism* Deviance	Specific *Reciprocity* Deviance	Specific *Loyalty* Deviance	Specific *Deference* Deviance
Dimension Social Relation †	Family/In-Group/Impartial	Fam./In-Gr./Impar.	Fam./In-Gr./Impar.	Fam./In-Gr./Impar.	Fam./In-Gr./Impar.	Fam./In-Gr./Impar.	Fam./In-Gr./Impar.
US-Sample							
- **Relevance**	Impartial ††	+ **Family** / Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
- **Judgment**	+ *Family* / Impartial	+ **Family** / Impartial	+ Impartial / **In-Group**	+ **Family** / Impartial	- **Family** / Impartial	Impartial	Impartial
- **Shame**	Impartial	+ **Family** / Impartial	+ Impartial / **In-Group**	+ **Family** / Impartial	- *Family* / Impartial	Impartial	Impartial
- **Guilt**	+ **Family** / Impartial	+ **Family** / Impartial	+ *Family* / **In-Group**	+ **Family** / Impartial	Impartial	Impartial	Impartial
GER-Sample							
- **Relevance**	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
- **Judgment**	Impartial / + *In-Group*	Impartial	Impartial	Impartial / + *In-Group*	- *Family* / Impartial	Impartial	- *Family* / Impartial
- **Shame**	+ **Family** / Impartial	Impartial	Impartial	+ *Family* / **In-Group**	- *Family* / Impartial	Impartial	- **Family** / Impartial
- **Guilt**	+ *Family* / Impartial	Impartial	Impartial	Impartial / + **In-Group**	Impartial	Impartial	- **Family** / Impartial
JP-Sample							
- **Relevance**	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
- **Judgment**	Impartial	Impartial	Impartial / + **In-Group**	Impartial	Impartial	Impartial	Impartial
- **Shame**	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
- **Guilt**	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
EG-Sample							
- **Relevance**	- **Family** / Impartial	Impartial	Impartial	Impartial	Impartial	Impartial	Impartial
- **Judgment**	Impartial	Impartial	- *Family* / Impartial	+ **Family** / Impartial	Impartial	Impartial	Impartial
- **Shame**	Impartial	Impartial	- *Family* / **In-Group**	+ **Family** / **In-Group**	Impartial	Impartial	Impartial
- **Guilt**	Impartial	Impartial	- *Family* / Impartial	+ **Family** / Impartial	Impartial	Impartial	Impartial

†**Note:** The labels *Family* and *In-Group* stand for the categories/levels of the Moral Deviance Factorial Survey dimension *social relation*(*ship*). The denotation *Impartial* indicates that there is no significant difference between the categories *Family* respectively *In-Group* and the reference category *Stranger*. If only *Impartial* is specified, this applies to both the *Family* vs. *Stranger* and the *In-Group* vs. *Stranger* effects. Where an algebraic sign is given before either *Family* or *In-Group*, this indicates the direction of the effect. The algebraic sign - indicates that the corresponding act of deviance is more important in interactions with a *Stranger*. As long as only one algebraic sign is given for *Family* and *In-Group*, this applies to both effects. ††**Note:** Effects written in bold indicate significant (p < 0.05) differences between the social relation categories *Family* or *In-Group* and *Stranger*. Effects written in italics denote indications at the 10% (p < 0.10) significance level.

morally acceptable to overstep the line of deference when dealing with strangers, as this would signal a social relationship not (yet) established and endangers future cooperative interaction.

In addition to results that are largely in line with our hypothesis, the unadjusted findings for the German sample also reveal particularistic moral tendencies. Similar to the US, we find particularistic tendencies in the domains of *property* and *heroism* in the German sample. Interestingly, despite similarities between the two WEIRD samples, we also observe cross-cultural differences. The GER-sample punishes property deviance in a family context with more shame attribution; in the US, emotional punishment is also imposed in this context, but with the attribution of guilt instead of shame. Furthermore, it turns out that it is primarily the in-group for which the GER-sample shows particularistic tendencies in regard to heroism. In the US-sample, by contrast, it is mainly kinship altruism (Kurzban et al., 2015; Curry, 2016), and thus the family that is crucial for the particularistic tendency. In the context of the latter, we have already pointed elsewhere to institutional differences in the social security systems of Germany and the United States (Seeleib-Kaiser, 2014).

Overall, drawing on the *non-Holm-Bonferroni-corrected results* in the German sample, we observe three significant effects that speak against our impartiality hypothesis, as well as four effects at the 10% significance level pointing in the same direction. The (full) impartiality hypothesis is therefore rejected in the GER-sample for the moral domains of *property* and *heroism*; in particular, particularistic self-regulation is promoted via the attribution of aversive emotions in these domains. In sum, a large number of results support our initial assumption of impartiality. Nevertheless, the uncorrected p-values also suggest that moral impartiality cannot be expected per se in the German context. Therefore, our final interpretation for the German sample, similar to the overall interpretation for the US-sample, is as follows: ***Moral impartiality/particularism depends to some extent on both, moral domains and social relationships. Moreover, taking into account the cross-cultural differences that we have discussed, we must also extend the latter statement to include the aspect of culture***.

Turning to the **JP- and EG-sample**, we find mainly evidence speaking clearly against the particularism part of our hypothesis — and thus also against parts of established theory. We would like to remind the reader that at the beginning of this chapter we picked out a statement by cross-cultural psychologist Harry Triandis (2001, p. 917) in which he points out that members of more collectivist cultural entities tend to be particularistic in their morality and apply moral guidelines and respective pro-social implications primarily in the context of (in-)group relations. Our findings are by and large not in support of this statement and run, as stated, in most parts also contra to our hypothesis. In fact, we find even more impartial tendencies

among the two cultural groups that are associated with collectivism and interdependence in self-construal (Cross et al., 2011; San Martin et al., 2018; Minkov & Kaasa, 2022; Kitayama & Salvador, 2024) than in the two individualistic samples already touched on.

We first focus on the *non-Holm-Bonferroni corrected* **JP-sample** results. As noted, comprising all family vs. stranger and in-group vs. stranger comparisons combined we have a total of 56 interaction effects that we examine for each cultural group in the context of the dimension *social relation*. Contrary to the hypothesis of moral particularism, the Japanese sample is the group displaying the most *impartial* tendencies among the cultural entities examined. Across the four dependent variables and on all moral domains investigated we find consistent impartial tendencies with only one exception. Solely in the (specific) *trustworthiness* deviance judgment model we observe a particularistic tendency in favor of the in-group: To break a promise and violating from standards of reliability is judged as significantly more wrong when someone from the in- group is harmed as compared to when a stranger is harmed by the same act of deviance. Reliable cooperation in groups is known to have played a major role in ensuring sufficient crops in societies that relied mainly on growing paddy rice in the past (Talhelm, 2022; Kitayama & Salvador, 2024). Historically, the Japanese population has relied heavily on rice cultivation for their livelihoods. Drawing on this historical subsistence style, a moral mind that is calibrated to judge especially trustworthiness violations by members of the in-group as wrong, can therefore be considered adaptive. Accordingly, the trustworthiness effect for the in-group in the JP-sample could be partly explained by the history of rice farming and the reliable collective cooperation required for subsistence. When we consider the trustworthiness effects found in the JP-sample, in addition to the results of the US-sample, and refer to the different explanatory approaches we have used to interpret the effects — i.e., (high) relational mobility in the US and a specific historical subsistence style in Japan —, we have to recognize that different paths may lead to the same results in calibrating people's moral mind. From our point of view, this is at least a possibility that should be considered.[148] No other evidence nor any indications at the 10% significance level reveal further support of the moral particularism theorizing for Japan. Thus, we must see our hypothesis to be refuted by empirical evidence in 55 of all 56 interaction effects investigated for this cultural group. Additionally, for Japan, our findings seem to contradict Harry Triandis (2001) theorizing clearly, and reveal furthermore impartial moral tendencies beyond WEIRD cultures (Henrich, 2020).

---

[148] It should be noted that among the four cultural groups examined, the USA in particular shows a relatively high degree of relational mobility. Germany can be considered neither particularly relationally mobile nor immobile, whereas Japan (strong) and Egypt (moderate) tend towards relational immobility (Thomson et al., 2018).

Consequently, our data show that it is not particularism but *impartiality* that appears to reign in Japan when people are asked to evaluate the *specific* moral breaches that we assessed in the context of different social relations between perpetrator and person harmed by the act of deviance. At this point, we will refrain from an ad hoc attempt at an explanation of the impartial effects found and will continue to address the hypothesis-contradicting effects of the Japanese sample together with findings from the Egyptian sample further below.

We also predicted particularistic moral tendencies in the Egyptian sample. However, the **EG-sample** results support the particularism part of our hypothesis only on one of the seven moral domains examined. We observe in the EG-sample particularistic tendencies in the evaluation of (specific) *heroism* deviance when a family member is harmed. This effect is found in moral judgment, as well as in the attribution of moral emotions. In addition to the effects for family, also more shame is attributed when a member of the in-group is harmed by an act that violates the domain of heroism. Note in this context: within honor cultures "good behavior comes from a desire to avoid shame" (Uskul et al., 2019, p. 795). Thus, different to the WEIRD samples, the results point to a particularistic tendency in Egypt that focuses more extensively on cooperation with one's group, be it the family or friends. Despite a self-construal that comprises both independent and interdependent aspects (San Martin et al., 2018), Egyptian culture is still strongly shaped by collectivistic elements (Minkov & Kaasa, 2022). When we draw on these collectivistic elements next to codes of honor (Leung & Cohen, 2011; Uskul et al., 2023), and also take prevailing self-assertive interdependence, intense (historical) reliance on kinship institutions (Cole, 2003; Reilly, 2013; Schulz et al., 2019; Curtin et al., 2020) and a moral reputation that is likely tangled to the reputation of one´s group (Uskul et al., 2019; Baumeister, 2022) into account, a cultural world becomes apparent in which a comprehensive and extensive group morality of heroism appears inherently meaningful.

Next to the particularistic tendency described, we also observe in the EG-sample (see: *Table 52*) indications ($p < 0.10$) suggesting that it might be more important to signal one´s reliability (i.e., the *trustworthiness* domain) in interactions with strangers. The respective effect becomes understandable by drawing again on the honor context prevalent in Egypt: Not only can hospitality be interpreted as a means to create ongoing circles of positive reciprocity and thus cooperative relationships in honor cultures, but, in addition, a person's reputation, their honor, also depends in part on a person's reliability. Honor logic expert Ayse Uskul and her coworkers (2019) state: "the honorable person is trustworthy, hospitable, honest, and true to his or her word" (p. 799). However, since a person's reputation can be seen as closely tied to their group (be it family or friends), and group members can be expected to avoid attempts to degrade

the collective reputation of their group, it becomes understandable why we find evidence in the EG-sample of a particular emphasis on keeping one's promise to strangers — after all, strangers can gossip and thus cause reputational damage (Henrich & Muthukrishna, 2021; Baumeister, 2022; Romano et al., 2022). Beyond the indications of a special importance of strangers in the *trustworthiness* vignettes, a significant effect in the same direction can be found for the Egyptian sample in the scenarios on specific *property* deviance. This effect could possibly be explained by recourse to the history of collectively owned goods, as found in intensive kinship institutions (Enke, 2019; Schulz et al., 2019; Curtin et al., 2020; Henrich, 2020; Bahrami-Rad et al., 2022). Against the background of a history of corporate ownership, the individual taking of a good by a group member is possibly not as relevant (and perhaps it is even not always classified as theft) as when a stranger is robbed and the collective reputation suffers or even worse consequences of negative reciprocity follow.[149] Although we believe that parts of the findings just mentioned could indeed be explained due to the logic outlined, we also hold, on the whole, that a more sophisticated explanation is needed, which we cannot provide at this point. We therefore see a need for further explanation of the effect that only the EG-sample in our study displays a significant negative effect (before Holm-Bonferroni correction) for the family level of the social relation dimension in the specific *property* deviance relevance model.

What ultimately stands out in the examination of the EG-sample and our impartiality/particularism hypothesis is the fact that our models show solely impartial tendencies in a clear majority of the family vs. stranger and in-group vs. stranger comparisons. Our hypothesis is to be considered refuted for 52 of a total of 56 interaction effects for the Egyptian sample. In summary, we see, similar to Japan, comparatively more impartial moral tendencies in Egypt than, for instance, in the US-sample of our study. Accordingly, the basic assumption of predominantly particularistic moral tendencies in collectivistic cultural entities (Triandis, 2001), such as Egypt (Minkov & Kaasa, 2022), can be called into question on the ground of our findings.[150]

We also refrain from an ad hoc explanation in the context of the impartial effects in the EG-sample. Instead, we now want to discuss the impartial JP- and EG-sample findings together from a perspective arguing in parts against binary conceptions of cultural phenomena. In our

---

[149] For deeper insights on the effects of positive and negative reciprocity in societies with high kinship intensity, see: (Maus, 1968).

[150] At this point, we would like to remind the reader of the sample characteristics of the EG-sample. Neither is the EG-sample of this study representative, just as the other samples of the further cultural entities are not representative, nor does it seem to reflect the Egyptian population in an unbiased way. The latter is especially true with regard to the variable education, as already emphasized elsewhere. Therefore, our results and statements should be evaluated in the context of the limitations of our samples, which applies in particular to the EG-sample.

view, the ***impartial results of the Japanese and Egyptian samples*** in particular raise the question of whether there is more — likely related to cultural processes of social change (Esser, 2002; Greshoff, 2008; Mahoney & Thelen, 2010; Greshoff et al., 2011; Hamamura, 2012; Cai et al., 2019; Kaasa & Minkov, 2020; Minkov et al., 2021) — than our current theories can fully grasp and explain. The findings on specific moral deviance in Japan and Egypt not only contradict large parts of our hypothesis, but also cannot be fully integrated into established theories. This circumstance may be due to our measurement instrument, but it can also be interpreted substantially. We suspect that our surprising results may have the potential to refocus parts of the theory of cross-cultural research: Binary explanations, such as those found e.g. in the context of the collectivism-individualism distinction, the individualizing-binding distinction, and the independent-interdependent distinction may not or no longer adequately reflect people's complex cultural realities. This leads us to three suggestions: 1) Established binary theoretical concepts from cross-cultural research, although they are practical heuristics, should perhaps be *revised in part* to better reflect the more complex relationships in reality.[151] Such research processes are already underway (see, e.g., Vignoles et al., 2016; San Martin et al., 2018; Krys et al., 2022) and, in our opinion, deserve more attention. 2) Theories and models of the social and the cultural should be *constantly checked empirically*, at appropriate intervals, of course. As Kaasa and Minkov (2020) state: "we need to update our models of national culture every decade" (p. 547). Regarding our knowledge of human beings compared to laws of nature, the crucial difference is that human quasi-laws are subject to change over time, i.e., they are subject to processes of social and cultural change. 3) Finally, our results suggest, in our view, a simple conclusion: Given the impact of processes of social change, but also of persistence, researchers concerned with the sociocultural and psychological constitution of the human mind should *continually explore* the social and cultural conditions that are so essential for the formation of human reality. In fact, we already know a lot — this thesis is just a tiny example of what it means to stand on the shoulders of giants — but there still seems to be more, at least in the cross-cultural context of moral impartiality/particularism, that should be taken into account not only empirically but also in the theoretical landscape. In summary, particularly based on the results of the JP- and EG-sample, we see a need for further research that addresses cultural tendencies towards moral particularism and impartiality, and we emphasize that the aspect of social change should be considered at best in both, empirical research designs but also in regard to established theories. A recent study by McKee and colleagues (2024) investigated moral impartiality and particularism across six countries drawing on the Moral Foundation

---

[151] Our statement should not be confused with a call to abandon parsimony in the theoretical explanations.

vignettes (Clifford et al., 2015). This study revealed the importance of the social affiliation of the perpetrator of a moral transgression as well as the victim being harmed by the act of deviance, highlighting overall the need to consider subjective group dynamics and social identity in cross-cultural research on impartiality/particularism (Abrams et al., 2000; 2002; Abrams & Hogg, 2004; Hogg & Reid, 2006; Pinto et al., 2010; Frings et al., 2012). The study by McKee et al., (2024) also found no effect for collectivism in the context of moral particularism and therefore emphasizes, similar to us, that an oversimplified interpretation of collectivism-individualism should be avoided in future studies of the same research interest (p. 8005) — it is important, though, to keep in mind that collectivism-individualism is a multidimensional concept. We presume that our results can be partly attributed to processes of socio-cultural change. Moreover, not only our findings from the two non-WEIRD samples in the Moral Deviance Factorial Survey are likely to be linked to processes of social change, but evidence from other studies (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019; Kaasa & Minkov, 2020; Minkov et al., 2021) points to socio-cultural and associated psychological changes around the world that call for recognition in established (e.g., cultural-psychological) theories. To be clear: this is not a plea to abandon established theories — such a plea based on a single study would reflect nothing but hubris anyway — but rather a plea to potentially revise some parts of existing theories, which are yet of tremendous value, to grasp a human constant, social change, and associated outcomes more accurately in our theorizing. We have hinted at potential for revisions and we see it, e.g., in the context moral impartiality/particularism and theoretical associations with collectivism-individualism. ***Overall, the fact that we found considerably more impartial than particularistic moral tendencies in Egypt and Japan calls thus for further investigations and reveals a need for theoretical explanations.***

Apart from the considerations just mentioned, it also becomes clear in the context of the JP- and EG-samples that moral impartiality/particularism is likely more complex than previously expected. In fact, the uncorrected Holm-Bonferroni results from all four samples suggest that ***moral impartiality/particularism is likely culture dependent and domain as well as social relation specific***, which we finally propose as a general conclusion from our investigations. Our conclusion is clearly underlined when we highlight two exemplary findings across all four cultural groups and focus on the moral domains of *property* and *heroism* (see: ***Table 61***). In the US effects show that moral violations in the *property* domain carry more weight in a familial context. The same applies to the German sample, but we also find indications here that, in addition to the family, property violations within the in-group are judged to be more wrong. In the Japanese sample, by contrast, we find only impartial

tendencies, and the Egyptian sample shows an effect indicating that specific property deviance is less relevant in a family context than when a stranger is harmed. Additionally, considering the results regarding *heroism* in all four cultural groups examined, our formulated overarching conclusion is clearly supported: A binary view such as collectivism vs. individualism (Triandis, 2001; Kitayama & Salvador, 2024), but also the already significantly more complex (and principally heterogeneous!)[152] WEIRD vs. other cultural entities perspective (Enke, 2019; Henrich, 2020; Apicella et al., 2020), as our data suggest, do not do full justice to the phenomenon of moral impartiality respectively moral particularism. In regard to heroism, the US-sample shows moral particularism for the family, the GER-sample shows moral particularism for the in-group, the JP-sample shows no moral particularism (but impartiality), and the EG-sample shows moral particularism for both family and in-group. As exemplarily illustrated, our findings point to complex relationships that go beyond binary conceptions as a means of explanation. Likewise, the findings on the other moral domains examined support the conclusion we have drawn.

As already mentioned, we will only briefly discuss the ***Holm-Bonferroni corrected results***, because they are statistically unambiguous. If we look exclusively at the p-value adjusted findings, we find nothing but impartial tendencies across all samples, dependent variables and moral domains examined. Consequently, the US- and GER-sample part of our hypothesis would be fully confirmed, while the JP- and EG-sample part of our hypothesis would be completely rejected. These findings hence indicate that there is more going on across the cultural worlds examined than we predicted. The OLS models we used for our analyses are, as has been emphasized repeatedly, complex and include large test families. Furthermore, our analyses revealed a multitude of significant influencing variables that likely not only affect the dependent variables but also moral impartiality/particularism.[153] It is clear from this that we are nowhere near the end of our investigations of our data set. Stepwise moderation analyses, more in-depth analyses of individual effects and the MDFS dimensions that have not been covered

---

[152] In principle, the researchers who introduced the WEIRD perspective into the scientific discourse (Henrich et al., 2010) do not suggest a simple dichotomy between WEIRD and non-WEIRD, but emphasize from the outset that both WEIRD cultures and those entities that do not fall under this acronym are heterogeneous (Apicella et al., 2020). In our Factorial Survey study, we find clear support for a perspective that emphasizes cultural diversity and heterogeneity in the context of moral impartiality/particularism. Nevertheless, our results also suggest that the WEIRD/non-WEIRD school, although it generally embraces heterogeneity, may still be partly too much of a heuristic in the context of moral impartiality/particularism (Henrich, 2020). At the same time, the differences in the data basis between our research and the cited studies should also be considered. At the very least, nonetheless, our results suggest that we apparently need to further investigate the phenomenon of moral impartiality/particularism, since our findings point to more complex relationships than previously assumed.

[153] Due to the sheer quantity of findings, we are not able to go into detail on them at this point, as already mentioned.

so far demand further attention, which we, though, cannot provide at this point. Therefore, we see a need of further research in the context of moral impartiality/particularism in general, but also on the basis of our data. To conclude our investigations on impartiality/particularism in this study, we now summarize our findings and considerations.

Modified impartiality/particularism hypothesis: We expected the US American and German cultural samples to tend towards *impartiality*. We did find result in support of this hypothesis but also deviations from the impartiality assumption. Regarding the samples from Japan and Egypt, we expected to find predominantly *particularistic* moral tendencies, but the results provided reveal mainly contrary evidence, indicating a large degree of impartiality rather than particularism. Overall, a binary impartiality/particularism hypothesis does not do justice to the complexity of valuations of moral deviance in WEIRD cultures and beyond. In fact, we found evidence suggesting that moral impartiality/particularism depends to some extent on the moral domain and the respective social relationship. For this reason, the *moral domain* to which the act of deviance relates and the question of *who* is harmed by the moral offense are likely to serve as antecedents of either impartial or particularistic evaluations. Since we also found cross-cultural differences between the samples, we need to expand what we have already established. **Overall, we conclude that moral impartiality/particularism is likely to be** *culture-dependent* **and both** *moral-domain-specific* **and** *social-relationship-specific*. To better understand our results, further research is needed at a global scale that takes into account processes of sociocultural change. In our view, processes of social change probably also affect established cultural psychological theories and should be reflected in them, because some binary concepts may no longer seem to correspond to the diversity of current cultural realities. Our data can be interpreted in part as suggesting that we should adapt some widespread and established theories to a possibly changing cultural world and that some binary cultural concepts, however heuristically practical they may be, should potentially be partially revised. **Taken together, we regard the modified impartiality/particularism hypothesis, as a whole, as refuted as more complex associations emerged from our analyses.** *X*

Finally, what do the MDFS analyses contribute to answering the **overall research question of which moral system guides cooperation in different cultures**? We have already compiled the most important information on this. First and foremost, morally deviant behavior tends to be judged as wrong *across* cultures. Moreover, aversive emotions are attributed to the actor of moral deviance *across* cultures, which sanction the person committing the moral offense. Hence, our investigations reveal a cross-cultural practice in which moral violations are punished by the actor's social environment with the attribution of shame and guilt. The perception of moral violations is therefore not without social consequences: *The human moral mind appears to be cross-culturally calibrated to recognize violations of moral principles as wrong and to punish them emotionally*. A certain degree of moral relevance is also attached to violations in various domains of cooperation *across* cultures. When focusing on the identified *cross-cultural tendencies*, the following becomes apparent: Not only specific *binding* deviance (deference and

loyalty) but also specific deviance towards *individualizing* morality (property, fairness and trustworthiness) is evaluated as relatively relevant and wrong, and attributed with shame and guilt *across* cultures. We found hence further evidence suggesting that individualizing moral domains are not only endorsed in the WEIRD cultures of our study, but also in the Japanese and Egyptian contexts. In this regard, our cross-cultural findings contribute to the Moral Foundations Theory framework (Haidt, 2008; Graham, 2011; 2016; Mooijman et al., 2017; Atari et al., 2022a). Furthermore, the Moral Deviance Factorial Survey findings seem to complement to the MaC-DRS results and the evidence obtained from the analysis of the moral dilemma scenarios. Altogether, *these results refer to the importance of individualizing morality not only in but also beyond WEIRD cultures*. In light of the variables deviance *relevance* and deviance *judgment*, as well as *shame* and *guilt* attribution, we believe that our analyses have revealed the aforementioned cross-cultural tendencies that contribute to the understanding of moral systems, despite the cultural differences that have also been identified. Moreover, we concluded that we can't find a systematic pattern of moral impartiality/particularism related to differences in e.g. collectivism-individualism, interdependence-independence in selfhood, and low/high (historical) kinship intensity. Instead, *moral impartiality/particularism is likely culture dependent and domain as well as social relation specific*. On the basis of our (unadjusted and p-value-corrected) findings, we suggest that both more individualistic and more collective cultural entities should not be expected per se to promote either moral impartiality or particularism. The background that sets the framework for whether people practice a universalistic morality that applies to everyone or whether people practice a morality that is reserved to benefit individual actors and groups is probably more complex than previously assumed — at least when examining specific acts of moral deviance and comparing *modern* societies, as we did.

As far as the **limitations** of our study and new avenues for future research are concerned, we have already given some indications: For Egypt in particular, we need better data and deeper insights. Furthermore, the number of cultural entities we studied is of course limited, and studies comparing a much larger number of cultural groups using MDFS would be desirable. Besides the limitation that we only examined four, albeit very heterogeneous, cultural entities, it should not be forgotten that we primarily represented modern people in our samples. This is already partly reflected in the way we collected data, which was due to an online survey method. Small-scale societies and populations with more traditional ways of life are largely underrepresented, if not absent in our study. Supplementary research that also includes people with more traditional ways of life is certainly relevant in the context of moral impartiality/particularism

research (Henrich, 2020). We have severely limited our analyses presented due to the sheer number of research possibilities. Not only the other MDFS dimensions, i.e., *gender* and *reputational damage* await in-depth investigation. Furthermore, it should also be emphasized that a large proportion of effects within the OLS models we used are not mentioned here and invite to conduct follow-up research. As mentioned earlier, we worked with quite complex models that included a considerably large set of covariates. A fundamental consideration in our covariate models was that we worked with the adjusted sample (Study 3: N = 2,360; MDFS: N = 9,440) and aimed to account for differences in sample composition in part by including the covariate set described in our models. Cross-cultural studies based on the Moral Deviance Factorial Survey instrument presented are therefore to be welcomed, which not only generally work with more cultural groups, but in the best-case balance them better than we were able to do against the background of different sample composition characteristics (He & van de Vijver, 2012). Furthermore, the respective sample size in our studies should not be ignored in our MDFS analyses, because these range between n = 297 and n = 392, and could have insufficient power for effects with a small size. Studies that pursue the same research interest as we do are therefore not only desirable in the context of expanding the cultural groups studied, but also in terms of expanding the sample size (Mayr et al., 2007; Lakens, 2022). There are also slight imbalances in the sample size between the groups we examined, which may well have an impact on the comparison of the groups with each other (Auspurg & Hinz, 2015). Future studies are therefore desirable that also work with balanced samples in terms of sample size. However, there is more to consider than just the limitations of our study design. In the context of the considerations and interpretations presented above it should not be forgotten that we have *operationalized* the moral constructs under investigation in a specific way, i.e., we have tried to make them measurable via the Moral Deviance Factorial Survey. We can never measure the construct of interest directly, but only via the respective operationalizations (Moosbrugger & Kelava, 2020; Moosbrugger et al., 2020). Furthermore, we have already presented a theoretical argument at the outset: We assume that each moral domain comprises a range of domain-specific actions that can be violated or adhered to. Moreover, specific acts of moral deviance may differ in their severity. This allows for the possibility of an intra-domain hierarchy of domain-specific acts of deviance. The consequence is that we have to assume a complex structure of a multitude of domain-specific acts. This inevitably raises the question of comparability of studies and study results: which moral aspect, which moral domain, which domain-specific act, which degree of deviance severity was operationalized and did the reference literature use a similar operationalization? In our study, we operationalized 7 specific

acts of moral deviance. The fact that some of our results do not fit into the canon of previous studies could also be related to (e.g.) different operationalizations, different domain-specific acts of deviance covered, and different degrees of severity in the moral violations that are not identical between our study and other investigations. Therefore, our results may also be due to the specifics of the operationalization we used (i.e., the vignettes) and not necessarily due to the moral impartiality/particularism construct itself. In other words, since our method of data collection was based on a new instrument and therefore not identical to the instruments used in the cited studies, it is possible that we were unable to replicate the results in question simply because we used a different instrument. In the context of our statements shared above, this uncertainty must be taken into account, as the operationalizations we used may also have contributed to the observed and interpreted results of the current study. However, in the light of our operationalizations, we have demonstrated evidence in support for the notion that moral deviance *relevance* is not exactly the same as moral deviance *judgment*. From this position, the question arises as to what extent the concept of *moral deviance relevance* contributes to the (cross-cultural) investigation of the human moral mind in a way that concepts such as *moral deviance judgment* cannot. We have already touched on it elsewhere, but would like to emphasize it again due to its importance: Our study is limited by the fact that we have not conducted direct behavioral measurements (Ellemers et al., 2019). In order to establish the concept of moral deviance relevance, however, it would be helpful to analyze to what extent moral *relevance* predicts pro-/anti-social behavior and to what extent these predictions, if they can be found, differ from what the concept of moral judgment can achieve. Investigating the predictive validity of moral deviance relevance in the context of behavioral measures is an area of future research (Moosbrugger & Kelava, 2020). In addition, it will certainly be interesting for future cross-cultural studies to examine how the attribution of guilt and shame in the context of moral deviance unfolds when costs arise for the sanctioning party (Horne & Cutlip, 2002). In the MDFS design, respondents are outside observers of *hypothetical* scenarios (vignettes), which is why a transfer of our results to more realistic designs that take actual behavior and sanctioning costs into account would be desirable. In order to establish research on moral relevance, correlative studies are also important to show how moral deviance relevance is connected with other concepts and whether the moral relevance aspect differs from the judgment aspect. Associations of moral deviance relevance with concepts such as the dark factor of personality (Moshagen et al., 2018), anti-social personality (Engelman et al., 2019), Social Dominance Orientation and Right-Wing Authoritarianism (Pratto et al., 1994; Altemeyer, 2004; 2006; Pratto et al., 2013), or concepts such as moral disengagement (Bandura

et al., 1996; Moore et al., 2012; Moore, 2015), to name just a few, would be beneficial in order to establish this approach more firmly. Of course, these statements do not only refer to the Moral Deviance Factorial Survey, but are also relevant in the context of MaC-DRS that we have introduced in a different chapter. Our study is further limited by the fact that we were unable to process physiological data, as our investigations are exclusively based on a semi-experimental questionnaire design. Neurological studies on the effects of deviant behavior are already known (see for example: Schreiber & Iacoboni, 2012; Amodio, 2014). A cross-cultural and physiological perspective, as it is provided in the context of the self-construal approach (Han & Humphreys, 2016), would also be desirable in the context of moral deviance and moral conformity in general. Furthermore, such studies could contribute to the understanding of moral deviance relevance and moral deviance judgment by looking at potential similarities and differences between these concepts from a physiological perspective. In our investigation of moral impartiality/particularism, further limitations were identified. These are partly related to the corpus of theory and could indicate that various theoretical approaches may require revision in parts, as already mentioned. We also found effects, particularly in the Japanese and Egyptian samples, that require further explanation. In our view, the limits of what we can explain have been reached in some aspects of our findings which calls for further research and theorizing. Finally, we have once again pointed out aspects of sociocultural change that may have had an impact on some of the results. Uncovering the current sociocultural constitutions of cultural entities could help us explain the results of moral impartiality/particularism that do not fit into the theoretical perspective adopted. Our study is a cross-sectional study. However, it has now become apparent that a longitudinal perspective is also desirable in the context of cross-cultural studies on morality. Overall, and despite limitations, we believe that our study on the Moral Deviance Factorial Survey (MDFS) contributed nevertheless to our understanding of the human moral mind across cultures. The MDFS instrument makes it possible to examine contextual factors of the vignette scenarios that cannot be collected relying exclusively on scales such as MaC-DRS, the MFQ 1 and 2, or MaC-Q (Graham et al., 2011; Curry et al., 2019b; Atari et al., 2022a). In this respect, we consider the MDFS instrument to be a further useful addition to the research repertoire for scholars in the field of morality. All in all, the limitations of our study provide starting points for future research using the Moral Deviance Factorial Survey presented in this chapter.

Let us **summarize** our findings and considerations. Moral deviance relevance and moral deviance judgment are not congruent in the extent of their valuation. Findings indicate that these concepts may touch on different aspects of morality. It will be a challenge for future

research to determine whether studies on moral relevance can make a contribution that cannot be covered by studies focusing on moral judgment. Also, more theorizing on moral deviance relevance is needed. Furthermore, our findings show that specific moral breaches are not only evaluated as relatively relevant across cultures, but that they are also cross-culturally viewed as wrong, and sanctioned via the attribution of guilt and shame. In regard to the MDFS social relationship dimension, our analyses suggest that moral impartiality/particularism is likely culture dependent as well as moral domain and social relation specific. Accordingly, we read the findings as to suggest that reality is more complex than our theory guided expectations and may therefore call for a partial refinement of the underlying theories. Finally, the limitations of our investigations were also highlighted. These limitations, however, also reveal potential for future investigations. Overall, with the Moral Deviance Factorial Survey (MDFS), we have added another facet to the cross-cultural investigations of the human moral mind and expanded the research repertoire to include a further instrument alongside MaC-DRS and the binding/individualizing dilemma scenarios.[154]

---

[154] Our supplementary analyses also suggest that when using the Moral Deviance Factorial Survey presented in this chapter in cross-cultural studies, it is important to control for the influence of response styles. The **Appendix** shows corresponding findings.

Chapter 7: General Discussion

7.1. Moral Deviance Across Cultures — Investigations of the Human Moral Mind

"You AND I are very unusual beasts" (Chudek et al., 2016, p. 749). But not only are we humans unusual because we rely on culture like no other species in the world to ensure our survival and reproduction. We are also unusual because we cooperate with each other to an extent that is probably unparalleled among self-reflective species. We harbor the inclination to *care* about each other. Our capacity for empathy may have its origins in kinship altruism (Frith & Frith, 2005; De Waal, 2008; Kurzban et al., 2015; Henrich, 2020). But this capacity goes far beyond that. And not only in the sense that we care about other people, but we also do so in various domains. Not that we misunderstand each other: in me as in you and probably in every other of our species, selfish drives also reside. We are not only pro-socially minded. On the contrary, taking care of ourselves is fundamental to our survival. But and this *but* is at the very least as important, without the social world that surrounds us and the regulation of our selfish drives in that social world, we are lost. Not only do we need relevant others, for example, who allow us to grow up and take care of ourselves. Moreover, it took hundreds of thousands of years of cumulative cultural evolution (Creanza et al., 2017; Mesoudi & Thornton, 2018) to enable our species not only to take a seat at what we might call the "sunny spot and richly laid table" in some parts of the world, but also to find a home in all of the Earth's natural environments. Our species has spread across the globe, developed complex cultures and shaped natural ecologies into cultural ones, allowing us to call a wide variety of habitats our own (Boyd et al., 2011; Henrich, 2016; Brown et al., 2022). At the core of cultural evolution is social learning from others and the ability to *cooperate* with others. The latter is the focus of this work, and by it we mean human morality (Curry, 2016; Henrich & Muthukrishna, 2021). The human development towards complex societies, as we call them today our social reality, is fundamentally based on our ability to share intentions with each other (Tomasello, 2017) and to put them into joint action. It is only through cooperation, through working together for our mutual benefit, in which you and I have a common interest in sustaining precisely because of this benefit, that the human being is able to rise above the limitations of the individual self. Cooperation has enabled us to produce complex cultural elements and highly differentiated social organization. Fundamental interdependency (Tomasello & Vaish, 2013; Henrich, 2020) is the starting point for a multitude of psychological developments of our mind, and can also be observed where culture has shaped

our reality into a cultural reality. In our everyday lives, for example, we encounter techniques that we use all the time but do not know how they work (Boyd et al., 2011). In our daily lives we act, in addition to the possibility of rational decision-making, primarily on the basis of traditions, norms and aligned with narratives that create identity (Bar-Tal, 1998; Esser, 2002a; Hogg & Reid, 2006; Chudek & Henrich, 2011; Phillips DeZalia, & Moeschberger, 2016; Keblusek et al., 2017). And yet we often do not know what the original core of knowledge is that guides our actions. The co-evolutionary trajectories and path dependencies that produced much of what we take for granted in our development often remain hidden from us (Henrich, 2020). However, these cultural paths shape our cultural realities, according to which our self and our psychological apparatus align themselves. Culture forms our biological niche and penetrates to the innermost part of us, our self (Brown et al., 2022). As a self-reflective species, we refer to ourselves day in and day out. For ourselves as well as for other fellow human beings, a respective self is the point of reference, as can be seen, for example, from the use of names. As humans, we all harbor a sense of self. However, our sense of self is fundamentally socially and culturally constituted (Baumeister, 2010; 2022; Markus & Kitayama, 1991; 1998; 2010). The individual self is unique as such, but we as humans share this uniqueness, so it is common to us that each of us is a self. We speak of ourselves and in doing so we use the words "I am", a statement of the reflective being of the self. However, since we can only grow up through the care of other people in such a way that we begin to recognize our self and others over time (Honneth, 2012), the need for the social is fundamentally inherent in us (Berger & Luckmann, 2013). Only through the care of others, which enables us to grow up, do we gradually become able to act for ourselves. In order to be able to grow up at all, humans need social support from other humans. This is where it comes to light what is meant by kinship altruism, and it becomes evident how fundamentally social each of us is. Our self and our morals intersect at the point where self-regulation becomes prevalent (Bandura et al., 1996; Baumeister, 2010; 2022; Cross et al., 2011). We understand morality as part of the self. Emerging from the survival- and reproduction-relevant factor of interdependency with other humans, encompassing the family but also going beyond it, a moral mind has evolved in the course of our species' history and under processes of gene-culture-co-evolution, prompting us to curb our selfish side so as to act in pro-social ways (Haidt & Joseph, 2007; Henrich, 2016; Curry, 2016; Brown et al., 2022). What about the sharing of food, for example (*fairness*)? What about obeying to others when we are partaking in a joint endeavor (*deference*)? What about reliability, and when can we count on each other (*trustworthiness*)? What about support for and from my social group (*in-group*)? How are all these and other vital domains of human cooperation socially regulated? The answer

lies largely in the fact that we are equipped with an evolved moral mind that causes us to self-regulate across plural domains of cooperation. Our moral mind prompts us to regulate ourselves, in that moral intuitions and deliberate thoughts, moral emotions and also social means of reputation, signaling and punishment repeatedly admonish us, remind us and ultimately regulate us in such a way that cooperation can arise in the form of mutual benefit and our selfish side falls behind the needs of the social. All human societies are based on the cooperation of their members and cooperation is fundamentally guided by our evolved moral mind. Morality is therefore a universal facet of the human mind. However, the extent to which each moral domain is important to people is itself a product of cultural path dependencies. In other words: The extent to which cultural entities were confronted with recurring problems of cooperation is reflected in the culture-specific calibration of the moral mind (Haidt & Joseph, 2007; Curry, 2016; Henrich, 2020). Our self is culturally constituted, self-regulation is the interface between morality and self, and so the psychological apparatus of self-regulation, our moral mind, which is universal to humans, is also a cultural product in its specific constitution. The fact that we have a moral mind, which is adapted to its socio-cultural environment in its specific constitution, is central to every human being. The moral mind helps individuals to navigate the given social world, to successfully fit into the social order, and to find well-being in one's social environment. What is more, our moral mind is important as a guiding compass to prompt the individual to re-produce the social order of the given socio-cultural environment through their actions (Ormel et al., 1999; Esser, 1999; 2002a; 2010; Greshoff, 2008). Against this background, the following research question arose as the core of the present project: **Which moral system guides cooperation in different cultures?** Our cross-cultural investigations of the human moral mind and the empirical findings presented in this project are located at the intersection of *culture*, *self* and *morality*, and aim to capture an element of what is universally human and yet culturally specific. Human morality and reactions to moral deviance from a cross-cultural perspective are at the center of our project. We found evidence for both: Our investigations have revealed indications of 8 cross-cultural moral domains, as well as the massive influence of culture on the calibration of our moral mind.

7.1.1. Theory

The investigations presented in this work are based on our moral approach. Against a gen-culture-co-evolutionary background (Henrich & McElreath, 2007; Richerson et al., 2010; Boyd et al., 2011; Henrich, 2016; Chudek et al., 2016; Brown et al., 2022), we have tried to integrate the Moral Foundations Theory (MFT) and the Morality as Cooperation Theory (MaC) (Haidt, 2001; 2003; 2008; Haidt & Joseph, 2007; Graham et al., 2013; Curry, 2016; Curry et al., 2019a; 2019b) to develop our theoretical perspective on morality. In addition to the aspect of *cooperation* we have placed moral *deviance* and moral *conformity* at the center of our moral mind. What social phenomena does our moral mind focus on? It focuses on the actions of ourselves and other social actors. Our actions can follow selfish drives, or they can be pro-socially oriented. By placing deviance and conformity at the center of our moral mind, we also place the identification, processing and valuation of human actions at the center of our moral apparatus. Human moral actions take shape as either moral *deviance*, which implies *harm* of others, or moral *conformity*, which in turn expresses *care* for others. Our theoretical perspective sees material or immaterial costs arising from moral deviance, at least for one party of the actors involved in a principally cooperative enterprise. Consequently, we consider moral breaches to be equivalent to harming someone. If we assume fundamental interdependence to ensure survival and reproduction, then moral deviance is associated with fitness costs for one party of a cooperative venture (Tomasello & Vaish, 2013; Kurzban et al., 2015; Henrich, 2020). Our moral mind is therefore, according to our proposal, which is in line with the findings of Baumeister and colleagues (2001; see also, for example, Kahneman & Tversky, 1979; Chudek & Henrich, 2011; Schreiber & Iacoboni, 2012), specifically designed to detect and evaluate moral deviance, since exposure to deviant behavior reduces the fitness of the individual harmed by the act of moral transgression. Moral conformity, by contrast, guides successful cooperation that is resulting in mutual benefit. The basis for the success of the cooperative venture is that we (also) care about someone other than ourselves. Rooted in the tradition of MFT and MaC, our approach takes a pluralistic moral perspective (Graham et al., 2013; Curry et al., 2019a). In the duality of moral deviance/conformity, which in practice expresses itself in harm/care, we see the guiding principle that permeates our entire moral mind and therefore also comes into play in all moral domains. On the basis of the theoretical pillars on which our moral approach is founded, we propose a non-exhaustive list of 8 different moral domains. Fairness, trustworthiness, property, heroism, reciprocity, family, in-group and deference are domains of cooperation regulated by our evolved moral mind. Our investigations are based on the

theoretical approach that we have outlined here in its basic tenets. The aim of this project is to examine questions of moral universalism and the culturally specific of human morals.

7.1.2. The Overarching Research Question and its Empirical Answer

In the working definition that we used to grasp human morality, we have proposed the following: Moral systems have the function of regulating and identifying egoism, which in turn enables cooperation (i.e., non-zero-sum interaction) between individuals and (within/between) groups, and fosters social life between people and the evolving of human social organization. On the basis of the theoretical background we have outlined, we pursued the overarching question: Which moral system guides cooperation in different cultures? We examined this question in four empirical studies, and our analyses are based on four heterogeneous cultural entities. Comparisons between Egypt, Germany, Japan and the United States of America were at the center of the empirical examination of our research question. If we were to condense all our findings into a single sentence, we could provide the following answer: **Which moral systems guides cooperation in different cultures? Likely a universal moral system guides cooperation, which is, however, presumably culturally extended and certainly culturally edited.** A closer look at the individual countries reveals that our analyses yield culturally specific moral systems, which in turn are based on the universal foundation of the same moral domains.

In **Germany,** we find an individualizing moral system in the intuitive tendencies: fairness, property and trustworthiness are, *relatively* speaking, intuitively the most relevant moral domains. However, in dilemma scenarios, in which binding and individualizing are contrasted and deliberate cognition is required, there may be deviations from the general individualizing-oriented intuitive tendencies. Situational influences and deliberate cognition can thus mediate the impact of intuitive gut reactions. So, although binding morality plays a relatively minor role in Germany compared to the other samples, it is not unimportant per se in this cultural entity either. In all the samples studied, and thus also in the German sample, moral impartiality is a question of the focal moral domain and social relationship, at least when it comes to evaluating a (specific) deviant action with judgment (right/wrong), relevance, and the attribution of shame and guilt. In Germany, particularistic tendencies are rather observed in the context of the in-group than the family, and to some extent, conditionally cooperative interactions with strangers play a special role. On the whole, the moral system identified in the

German sample seems to come closest to what we have theoretically derived and understand as an ***individual-centered (independent) overall social orientation***. Notwithstanding the categorization presented here, the idea of an overall social orientation is more of an ideal type. The German sample corresponds most closely to this ideal type in *comparison to* the other cultural entities examined, but only *relatively* and certainly *not absolutely*. There may therefore be distinctly more pronounced individualizing moral systems with a corresponding overall social orientation. But in our studies, the German sample comes closest to what we understand by an individualizing moral system, which is why we propose this category as an answer to our research question for Germany: An **individualizing moral system** focused on the individual, that particularly emphasizes fairness, property, and trustworthiness, primarily guides cooperation in the German sample of our investigations.[155]

As for the **United States**, we note a greater empirical divergence from our pre-study ideal-typical theoretical model (see: **Chapter 1** and **2**) than is the case for Germany. In addition to intuitive tendencies towards individualizing morality, we were able to empirically identify the importance of the moral family domain. Regarding both the intuitive tendencies and parts of the deliberate preferences, as well as moral particularism, we find a weight of the family domain in the US-American sample that should not be underestimated. The latter is also specific to domains such as social relations in the US context, but if moral particularism is present, then it refers mostly to the pre-eminent position of the family. Against this background, to expect an individualizing moral system for the USA, as we predicted prior to our research, would be too simplistic. Rather, our empirical findings culminate in the conclusion that we should ideally speak of a ***familial-* and *individual-centered overall social orientation***. In addition to property, it is fairness, but above all the family domain, that characterizes the US American moral system according to our data. Consequently, we categorize the US moral system as a **mixed moral system** in *general* and as an **individualizing moral system of a familial character** in *particular*. In the USA, the moral system could be an expression of the fact that cultural developments do not necessarily follow a linear path, because, for example, a decline in impersonal trust has been observed in the US context over the past few decades (Hamamura, 2012). It will be interesting to see, also in terms of the political situation in the US in 2024, how the development of morality in the US will take shape in the coming decades and whether aspects of binding morality will perhaps regain more importance. At the very least, we can conclude that a description of the US moral system that is limited to individualizing morality

---

[155] In this context, please also note the results of data collection 2 for Germany, which are shown in the **Appendix** and give further support our conclusion.

would fall short, since significant symbols, and corresponding scripts, that emphasize the family domain are also present in the everyday situations of the cultural subjects.

The data from Japan and Egypt surprised us the most in relation to our hypotheses and the descriptions of the socio-cultural context outlined prior to the data collection, as they deviate from our expectations in a variety of ways. In fact, empirically *no* dominant ***group-centered (interdependent) overall social orientation*** can be determined for **Japan**. What we did observe in our study of intuitive moral tendencies, however, was a tight comprehensive mixed moral system. From a comparative perspective, moral conformity and moral deviance, respectively, seem to be largely independent of the respective moral domain in the Japanese context. This finding is in line with the Tightness/Looseness Theory (Gelfand et al., 2011; 2017; Roos et al., 2015) and broadens our understanding of the applicability of tightness/looseness to moral standards. What was surprising, in addition, was the empirical finding that, with one exception, moral particularism could not be identified in our data for the Japanese context; according to our data, no group occupies a special position, but moral impartiality primarily prevails in the Japanese context. Perhaps this fact in our data can also be attributed to the tight moral system, which could mean that adherence to moral principles is more important than the social affiliation of those with whom one is in potentially cooperative interaction. What could also be observed is the empirical fact that individualizing moral domains are intuitively more relevant in the Japanese context than binding moral domains. Regarding deliberate tendencies, we also found individualizing preferences in the moral dilemma analyses. Interestingly, when we compare the intuitive relevance of the binding domains in Japan with those of the two WEIRD samples in our data, we find though that they are more relevant in the Japanese context for the most part. Now, how can our findings be summarized with regard to our overarching research question? We understand Japan, against the background of our findings and our interpretation of the data, *in general* as a **tight comprehensive mixed moral system** and *in particular* as an **emerging individualizing moral system of tight and holistic orientation**. The holistic aspect (Nisbett et al., 2001) applies to the system description, since, despite the widespread individualizing dominance, we do not see any contradiction for Japan in continuing to regard binding moral domains as valuable and important. The ideal-type definition of the situation outlined prior to data collection (i.e., Model 2 to 3) is not only insufficient, but also misguiding. Our data suggests overall that the significant symbols, comprising codes and scripts, that seem to constitute everyday social situations in Japan must be defined more broadly than by just focusing on the social group.

As for our findings for **Egypt**, we are, and this should not come as a surprise after our explanations, generally inclined to be cautious, because we have doubts about the reliability of our database. All *our samples are not representative*, which must always be taken into account when interpreting our findings. But within the context of the Egyptian sample, we were able to identify more problems, which are the basis of our caution.[156] We have already discussed these problems at length, so there is no need to address them further. If we however interpret the findings which we have, we see that our ideal-typical logic of the situation derived in the theoretical Models 2 and 3 is also largely *not applicable* to Egypt, or at least does not go far enough. As was to be expected, the family domain is intuitively relevant in the Egyptian moral system. In addition, however, property and trustworthiness can also be found as intuitively relevant domains of morality. The intuitive relevance of these domains can be understood in the context of the prevailing logic of honor (Leung & Cohen, 2011; Uskul et al., 2019; 2023). What is interesting is the fact that, on the one hand, we were able to detect broad particularistic tendencies in the context of the heroism domain, which fits well with the self-assertive interdependent self-construal concept (San Martin et al., 2018), but on the other hand, we otherwise primarily find impartial tendencies. In the context of dilemma analysis, the Egyptian sample also showed primarily individualizing preferences. Against this background, our assumption of a ***group-centered (interdependent) overall social orientation*** seems, as already mentioned, to be too narrow. Accordingly, the social situations of the sample that we examined for Egypt not only contain significant symbols that make the group salient as an anchor for moral orientation, but also symbols that go beyond. Based on our data and the interpretations presented, we would categorize the moral system of our Egyptian sample *in general* as a **limited mixed moral system** and grasp it *in particular* as an **emerging individualizing moral system of a familial character**. That said, we do remain cautious about the lasting validity of the categorization of the Egyptian moral system that we propose here, and look forward to further research in Egypt, hopefully with a sample that is more representative of the Egyptian population than our own.

---

[156] We found, for example, that the Egyptian sample had the highest level of education in the comparison of the samples, which indicates a strong bias towards higher education in the Egyptian sample.

7.1.3. Four Further Key Insights

Our investigations of the human moral mind have uncovered further empirical evidence that goes beyond the classification of moral systems. To put it in a nutshell, our research was able to bring to light **four** major findings: **First,** we found strong indications that the 8 moral domains we proposed and largely borrowed from the synthesis of Moral Foundations Theory and Morality as Cooperation Theory are indeed cross-cultural components of human morality. **Fairness, trustworthiness, property, heroism, reciprocity, family, in-group, and deference are cross-cultural domains of morality; they guide cooperation between people and, as intuitively relevant elements of our evolved moral mind, form pillars of social order.** Our cross-cultural research indicates that the moral domains mentioned are human *universals*. However, following this universalism thesis, it should not be ignored that there may well be culture-specific domains of morality (see, for example; Atari et al., 2020b), and we do not assume either that we have presented an exhaustive list of all moral domains with the 8 domains suggested. For example, it will be exciting to incorporate the domain of purity (Graham et al., 2011; Atari et al., 2022b) from a cooperation-oriented moral perspective into the instruments of our theoretical perspective in the future. What our research has shown is that violations against the eight proposed moral domains lead across cultures, probably in most parts automatically, to intrapersonal processes that make individuals aware that the observed behavior is indeed subjectively relevant. In this context we assume that the effect of intuitive relevance of moral deviance in the corresponding domains is evident across cultures because human populations were fundamentally confronted with the same recurring challenges of the social organization of their survival and reproduction (Haidt & Joseph, 2007; Carlo et al., 2016; Curry, 2016; Henrich & Muthukrishna, 2021). Our perspective thus follows the canon of the cited literature and places evolutionary processes at the center of the explanation of universal human morality.

Second, the universal is only one side of the coin, the other side of which is the *culturally specific*. This duality inseparably forms a central aspect of what makes us human. Not only are humans as such a cultural species and consequently share this characteristic at the core of their being with their fellow species. Humans are also products and producers of their cultural environment (Berger & Luckmann, 2013). Even our self, the supposedly most individual element that each of us calls our own, is permeated by the social and the overarching culture. As a self we are a sociocultural shaped shaper (Markus & Kitayama, 2010, p. 421). In the course of its history, our species has come to inhabit the entire planet Earth. The production of culture was (and is) a decisive driving force behind the possibility of colonizing a wide variety of

natural environments. Culture is the biological niche of humans. As such, it encompasses the essential elements that allow survival and reproduction and reflects the realization of the interaction between humans and their environment. What is more, culture takes shape as our social environment; our psyche aligns with culture to enable us to navigate the world, which is naturally cultural for us (Henrich, 2020). Is it surprising, then, that our research not only reveals universal aspects of our human moral mind but also massive cultural differences in the calibration of the human moral mind? We think it is not surprising. **Our research has shed light on differences in moral tendencies and shown that these differences are significantly due to cultural differences. The human moral mind is thus culturally calibrated.**

Third, we derived and tested various hypotheses from a stream of theory and empirical findings from other studies. The result-driven discussions of the hypotheses can be found in the individual chapters of our four empirical investigations. We would now like to emphasize what we consistently encountered in our investigations: **individualizing morality plays a role in WEIRD cultures and beyond**. Both in the analyses that primarily addressed the intuitive, effortless, and uncontrollable tendencies of our moral mind and in the analyses that we prescribe in the area of deliberate moral cognition, a dominant individualizing morality was largely evident across cultures. Initially baffled by this result, our literature search quickly turned up findings from further studies that describe a **global trend toward individualism** and link this trend to the demands and developments of modern societies (Hamamura, 2012; Santos et al., 2017; Cai et al., 2019; Kaasa & Minkov, 2020; Minkov et al., 2021). In our view, the concept of individualizing morality largely falls under the umbrella term of what Henrich (2020) understands as the *individualism complex* and which also encompasses cultural individualism. Joseph Henrich (2020) argues in this context:

> "Focusing on one's attributes and achievements over one's roles and relationships is a key element in a psychological package that I'll clump together as the *individualism complex* (…). [This complex] is best thought of as a psychological cluster that allows people to better navigate WEIRD social worlds by calibrating their perceptions, attention, judgments, and emotions (pp. 26-27).

Strictly speaking, our research design does not allow us to make any trend statements. However, we suspect that processes of social change are at work, indicating that social modernity, and the forms of interaction that are dominant in it, particularly the protection of the individual from exploitation, require a correspondingly calibrated moral apparatus that can be understood as an adaptive concession to given societal demands. If the human moral mind is adapted to modern

society in the cultural entities examined, and in particular emphasizes individualizing morality, then various things can be derived from the dominant individualizing morality, which certainly includes fairness, trustworthiness and property, and possibly other domains as well. According to our interpretation of the data, our findings point to how our adapted psychological apparatus guides us to navigate, act and, above all, be pro-social in the face of given social demands. Furthermore, the psyche of the individual, adapted to the sociocultural context, can provide insight into what modern society consists of in terms of the demands it places on people. One of these demands seems to be to interact and, more importantly, to cooperate with relatively strangers, detached from familiar groups and the structures of conformity and deference that provide orientation. Consequently, a social organization beyond the group seems to prevail in modernity. In this context, though, it is important to emphasize that we should not assume a rigid dichotomy of "beyond the group" and "in the group"; rather, we should think of the relationship between these poles as the weights on a weighing pan: both weights are present in the pan of social modernity, but individualizing morality, that is "beyond the group", apparently weighs more heavily in the countries we have studied, at least for the time being. Our analysis clearly showed that we need appropriate comparisons with traditionally living populations, data from significantly more cultural samples, and finally a long-term perspective on societal and corresponding moral development to pursue the considerations presented in depth. The key takeaway, however, is that fairness, trustworthiness, and property represent aspects of morality across a set of four highly heterogeneous cultural entities that seemingly carry more weight than aspects of binding morality. Of course, binding morality is not unimportant, which is not what we want to say here, as illustrated by the above metaphor. Indeed, aspects of it are also (still?) relevant in WEIRD populations, as the analyses of the moral domain family in the context of the USA, e.g., revealed. But **elements of individualizing morality seem, on the whole, to be more important in modern societies across cultures, according to our findings.**

Fourth, our work raises questions in the context of binary concepts such as binding and individualizing. Following the binding and individualizing idea, a fruitful **heuristic** can be identified if, for example, a trend towards individualizing can be empirically determined across various cultural entities. As described, our investigations are limited due to their design and cannot identify such a trend. Our evidence at most provides an indication of the possibility of a cross-cultural trend towards individualizing morality. However, as such, the binary binding and individualizing heuristic also appears to be too simplistic to capture cultural reality. Cultures are diverse, and so are the individual minds that give life to the aggregated abstraction of what we understand by culture. **Sociocultural realities are made and experienced by people; they**

are multi-layered and inherently complex. What remains important to note about binary concepts — such as e.g. binding-individualizing, collectivism-individualism, particularism-impartiality, etc. — is that they are multifaceted constructs and that a mere concentration on the overarching constructs does not do justice to people's lives and their complex realities. The analyses we have presented in this work clearly underscore this point. Not only were we able to identify different intuitive moral systems across the cultural entities studied, for which the consideration of the proposed 8 moral domains was crucial. We were also able to show that deliberate moral tendencies can differ from intuitive tendencies. While this finding places our work in the canon of theory (Haidt, 2001; Tessman, 2014), we would also like to point out how complex the phenomenon of human morality is. Morality encompasses *various domains*, as well as *deliberative* and *intuitive* processes of cognition. Furthermore, our findings show that a simple binary logic of explanation does not do justice to aspects of morality such as particularism/impartiality. Moral particularism and moral impartiality are culture-dependent and moral domain-specific as well as social relation-specific.[157] Specific *aspects of morality* by themselves may possess a complexity of their own. In general, it is also important to consider whether, for example, *specific* moral actions are used to generate data and draw conclusions, or whether instruments are used to generate data on *general* moral tendencies and to derive corresponding conclusions. How much of a specific moral action is explained by a general moral tendency, and how much of a specific moral action can be generalized? These are additional questions that need to be considered in the theorization and study of human morality. So, what we want to say with the above points is the following: **Next to cultural realities, also the phenomenon of human morality is multifaceted and complex.** In addition to intuitive and deliberate processes, our morality encompasses various moral domains. Aspects of our morality can in turn comprise their own relations (see particularism/impartiality), and we need to distinguish between specific and general moral tendencies. Furthermore, morality can be researched in a multidimensional way and examined at different levels of analysis (e.g. cognition, emotion, behavior, institutions). Furthermore, human morality is situated in the field of tension between what is culturally specific and what is universal for our species. As we can see, the phenomenon of human morality is truly complex. Before making far-reaching

---

[157] Therefore, moral particularism/impartiality cannot be reduced to binding or individualizing morality per se. This observation is important in the context of our theoretical assumptions and raises questions as to whether a fundamental assignment of moral domains to either primarily promoting cooperation between individuals or cooperation between/within groups can be assumed at all. It is likely that there are only gradual and not essential differences that become prevalent in this respect, which is why a strict separation between binding and individualizing morality does not appear to us to be meaningful. The latter also became empirically evident from the fact that the higher-order MaC-DRS factors of binding and individualizing morality are indeed strongly correlated with each other and are not to be understood as orthogonal.

statements about this phenomenon, we need insights that take this complexity into account. In theory development and hypothesis generation, binary logics can play a heuristically instructive role. However, we should be aware that such heuristics are certainly too simplistic in parts and correspond more to an ideal type (Barmeyer, 2010; Weber, 2013) than to the reality of people. Corresponding deviations from the hypotheses based on binary logics should therefore be found in reality. The results of our analyses are a testament to this. It is therefore necessary to weigh up parsimony and realism. The practical consequence is that, in the context of moral research, good scientific practice requires not only a balanced theoretical framework but also a canon of methods to empirically approach the complex and multifaceted phenomenon of human morality. With our project, we have expanded the range of available measurement instruments by three additional research tools in the context of investigating human morality. We have developed a scale, a factorial survey and a series of dilemma scenarios and applied them in our cross-cultural project. All three instruments are available in four language versions and are suitable for cross-cultural psychological research.

7.1.4. Expansion of the Research Repertoire: Three New Tools for Moral Research

Our research has shown that the **Morality as Cooperation–Deviance Relevance Scale** (MaC-DRS) is a valid and reliable supplement to the scales of MFT and MaC (Graham et al., 2011; Curry et al., 2019b; Atari et al., 2022a). Our scale measures general moral deviance relevance valuations and we have designed the scale to primarily capture intuitive tendencies. MaC-DRS allows a broad coverage of moral pluralism across cultures with 8 domains, while two higher-order factors can be formed intra-culturally with binding and individualizing morality. Psychometrically, the economical 24-item short version of MaC-DRS works best, which is why we recommend this version for future research. The **dilemma scenarios** that we designed are explicitly based on what we, following Haidt (2008), understand as **binding** and **individualizing** and offer the additional possibility of working out a clear moral preference through realistic and scenically dense dilemma situations.[158] The forced choice logic leaves no room for neutrality, and the nature of the dilemma is likely to evoke and measure deliberate moral cognition. The **Moral Deviance Factorial Survey** (MDFS) allows for the variation of contextual elements of the vignettes (Auspurg & Hinz, 2015), which makes it possible to (quasi-

---

[158] In contrast to, for example, the Moral Foundations Vignettes (Clifford et al., 2015) or the Trolley Dilemma (Awad et al., 2020), we deliberately chose to create realistic rather than extreme dilemma scenarios.

)causally trace whether the different dimensions of the instrument, with their different levels, have an effect on specific and deliberate moral deviance valuations. Overall, a vignette universe with 168 vignettes is available and the design allows for variation of *gender, domain of deviance*, *reputational damage* and *social relation*(*ship*). The modular nature of the MDFS and the ability to track the variations of the dimensions and levels makes it possible to supplement research on morality with this instrument in a way that scales can hardly achieve. Like any research tool, the tools we have designed have their limitations. But when used in concert, the weaknesses of the individual instruments are partially leveled out, and together they provide the possibility for deep insights into the human moral mind. All three instruments can be considered complementary, are based on the theoretical perspective on morality presented in this work, and enable the investigation of moral tendencies from multiple perspectives.[159]

7.1.5. General Findings

Our cross-cultural investigations of the human moral mind have also revealed **general tendencies** that we would like to briefly highlight. The analyses of the vignettes of the Moral Deviance Factorial Survey (MDFS) show that deviant behavior is identified as such across seven moral domains. Across cultures, deviance is identified as rather wrong and judged accordingly. Moreover, morally deviant behavior is punished across cultures with the attribution of moral emotions. This is an indication of a cross-cultural socialization practice: as a result of perceiving morally deviant behavior, the social environment makes the moral transgressor feel the moral violation. The latter includes not only social actions such as exclusion, physical violence or punishments that lead to a loss of reputation (Henrich, 2020), but also the fact that the social environment contributes to the transgressor feeling aversive emotions (Haidt, 2003; Tangney et al., 2007; Wong & Tsai, 2007). We are supposed to feel guilty and ashamed because of moral transgressions; these emotions are attributed to the transgressor by other members of society across cultures in response to the perception of moral deviance. Comparing the results of the MaC-DRS analyses with those of the dilemma scenarios, we were also able to show that intuitive and deliberate moral tendencies do not necessarily lead to the same outcome. Corresponding effects were found across cultures. In this context, we believe it will be interesting to examine whether intuitive moral tendencies, for example, are better able to reflect

---

[159] The three instruments we developed can be viewed in full length and in all four language versions in the research plan of our third data collection, which is available online. The latest versions can be obtained from the author on request.

cultural differences than their deliberate counterpart. Our analyses support such a hypothesis, but certainly much more research is needed in this area to see clearly.

7.1.6. Moral Relevance

Moreover, our cross-cultural analyses show that moral **relevance** and moral **judgment** do not correspond in their extent when it comes to the valuation of acts of moral transgression. We argue that morally relevant actions can be considered wrong/right, but that this does not imply that they must also appear subjectively relevant to the same extent. Furthermore, we take up Curry's (2016) idea that different societies have been confronted to varying degrees with recurring social challenges that our evolved moral mind regulates, which is why moral preferences also vary from culture to culture. We argue that the concept of moral relevance is particularly suitable for investigating the hypotheses mentioned by Oliver Scott Curry. Our thoughts lead to the following: The extent of the subjective relevance of a moral violation reflects the cultural imprint of the cultural subject. The moral domains that are perceived as *more relevant* indicate a particular need for a socio-cultural entity to regulate and govern the corresponding domain, and the social order is determined to a large extent by the actions that fall within the respective domain. Thus, if people experience an intuitive and deliberate relevance when they perceive and process violations of a particular moral domain and, in the course of evaluation, attach an appropriate weight of relevance to the perceived behavior, then the domain in question also has a special place in the sociocultural history of the respective cultural entity. Relevance then is a reflection of the *intensity* with which a cultural group needs to regulate corresponding domains of morality in order to maintain the social order and to ensure the reproduction of the fundamentally interdependent social and individual organism. We assume that the concept of relevance, in contrast to the concept of judgment, is better able to reflect this intensity. However, further research is needed to determine whether the concept of relevance can make the contribution we have entrusted to it here. In our view, the field of moral relevance is still waiting for future theory development and empirical research. We have already pointed out elsewhere in this work that the concept of moral relevance has yet to prove itself alongside that of moral judgment (e.g.) in order to establish itself. The central question in this context will be whether moral relevance can contribute something to the understanding of human morality that concepts such as judgment cannot provide. It therefore remains to be seen whether and, if so, what added value can be achieved by studying moral relevance. In this work, we have presented a theoretical and partially empirical sketch of what the construct could

achieve. However, there is certainly still a lot to be done to give the idea of moral relevance the appropriate place in the context of research on human morality, be that place at the forefront or at the back. **Overall, we believe that the concept of moral relevance has the potential to serve as a field and motivation for future research.**

7.1.7. Limitations

As with all research, our investigations are not without their **limitations**. We have already pointed out and discussed various limitations in the individual chapters. Nevertheless, we would like to highlight the most important ones again below. Our cross-cultural investigations of the human moral mind are based on four cultural samples. This in itself represents a limitation, especially in the context of research into universal human tendencies. We were certainly only able to approach the latter. MaC-DRS and the other instruments are available in four language versions and, especially with regard to the Arabic and English versions, allow for future research beyond the countries we have examined. Further translations into other languages are also desirable in order to expand our understanding of human morality beyond WEIRD cultures in particular. We have examined data from Egypt, Germany, Japan, and the United States of America. Besides the fact that these are highly heterogeneous cultural entities, all four of these cultural entities are more or less modern societies. In particular, studies that focus on small-scale societies that live more traditional, non-modern lives are desirable in light of the presumed trend towards more individualism in general, but also in the context of the concepts of binding and individualizing morality. It should also be noted that our data from the cross-cultural investigation is not representative and was collected on the basis of access panels. This is a limitation of our research. Our selection of cases for the entities under investigation is well-founded and our samples are deliberately chosen, but together these research decisions also represent limitations. Our analyses are based on cross-sectional data. However, longitudinal studies are necessary to be able to identify trends in the development of moral tendencies. In particular, longitudinal studies that examine not only deliberative but also intuitive processes of moral cognition are a welcome starting point for future studies. Our survey design thus also represents a limitation. Furthermore, the analyses of the Moral Deviance Factorial Survey showed that we have by no means exhausted the full potential of our instrument, nor the full potential of our data. Here we see a further limitation of the work presented and a source of future work. Of course, the use of average marginal effects (AME) also represents a statistical limitation in some respects, because with this procedure we observe average people in the data

(Williams, 2012). With this method, we thus move away to some extent from observing real responses of people, because we are hardly, for example, average people in a variety of sociodemographic characteristics. However, the use of AMEs for the effect analysis was also justified, since we were particularly interested in measuring the effect of culture and therefore isolated sample imbalances. We were also able to take culture-specific interaction effects into account in the estimation of the variables of interest using the AMEs. Although the use of average marginal effects is a limitation, we still consider our approach to be justifiable. Overall, **the two most significant limitations** for us are the following: we did not collect behavioral data and were unable to draw any direct associations between the endorsement of different moral domains and ways of being independent and interdependent. Not only to establish the instruments we have developed, but also to give more weight to the concept of relevance and, above all, to our theoretical focus on deviance, studies that use MaC-DRS as a predictor of moral behavior are desirable and, in our view, highly promising. Cross-cultural studies of morality that combine intuitive and deliberate processes of moral cognition with behavioral data appear to us as an important goal and a worthwhile research project, in order to gain a deeper understanding of what motivates us to regulate our self, i.e., our morality. As for one of the core arguments of our theoretical framework, which is that morality is part of the self, we were unable to pursue this claim *directly* in the present thesis due to project constraints such as time pressure. In this context, we were also unable to pursue our postulated correspondence hypothesis of moral domain endorsement and ways of self-construal. We consider the latter two limitations to be the major restrictions of the present work. However, we are confident that in further and more in-depth analyses of the CIRN-3 Self-Construal Scale (Vignoles et al., 2016), we will be able to identify an instrument with which we can empirically test our postulates for the first time. This thesis was not *yet* able to provide the needed in-depth examination of the Self-Construal Scale, but the authors are already planning to undertake such an endeavor and are confident that they will be able to carry out the corresponding analyses in the near future.

7.1.8. Future Research Should Not Lose Sight of the Social Relevance of Morality

As we have seen, limitations can be starting points for future research. In addition to the already mentioned possibilities for worthwhile research projects, **the social relevance of the topic of morality** also suggests further research and practical application potential. We will only briefly touch on some aspects of the far-reaching social relevance of morality. In the context of intergroup conflict and prejudice, a fruitful contribution could be made by examining and focusing more closely on the moral orientations of the actors involved. Research could ask whether cross-group moral conformity can effectively reduce prejudice between different social groups and whether moral conformity can serve as a positive element in approaches such as the contact hypothesis. Turning to the other side of the conformity-deviance duality, we may also ask whether moral deviance committed by outgroup members is a factor in the failure of positive intergroup contact and a source of increasing prejudice (Pettigrew, 1998; Pettigrew & Tropp, 2006; Barlow et al., 2012)? Research into whether and to what extent different calibrations of our moral mind are also perceived as ethnic markers and lead to categorizations of belonging or not belonging to social groups is certainly also desirable in the context of inter- and intra-group dynamics (Kurzban et al., 2001; Hogg & Reid, 2006; Henrich & McElreath, 2007; Chudek & Henrich, 2011; Turner & Reynolds, 2012; Pietraszewski et al., 2014). We therefore believe that future analyses of the effects of moral conformity and deviance in intra- and intergroup dynamics continue to represent an important area of social psychological endeavor (Abrams, 2000; 2002; Frings et al., 2012; McKee et al., 2024), and could hold the potential to contribute to the mitigation of intergroup conflicts and prejudiced social relations. Furthermore, a migration and integration perspective on moral deviance also promises socially fruitful insights. For example, the Moral Deviance Factorial Survey presented by us could be extended to include the dimension *member of host country/member of migratory group* in order to investigate whether acts of moral deviance are evaluated differently for different social groups and whether such evaluations may possibly promote prejudices, perceived threats and discriminatory attitudes (Esser , 1999; Fiske, 2000; McLaren, 2003; Stephan et al., 2009; Dovidio et al., 2010; Tooby & Cosmides, 2010; Schreiber & Iacoboni, 2012). In our view, it will be important in general to further explore the duality of human morality, which is fed by universalism and the culturally specific, but also to communicate research on this to society as a whole. Kohlberg and Hersh wrote as early as 1977: "Whether we like it or not schooling is a moral enterprise" (p. 53). The institutional teaching of moral standards that goes beyond a

WEIRD perspective and, by incorporating our evolutionarily developed universal foundations, also takes into account the culture-specific calibration of morality and the *relative* differences in terms of what is morally relevant, can only be for the benefit of any society. We must not forget that we live in a globalized world, and given that migration has always been a part of human societies (de Haas, 2014; de Haas et al., 2020), relative cultural differences in the standards of what we consider good and bad will inevitably be brought into confrontation. It is important to learn how to avoid creating a divisive, destructive, or even anti-social space for one's own ethnocentrism (Bizumic & Duckitt, 2008; Bizumic et al., 2009), even in relation to such fundamental phenomena as morality. Communicating research findings on human morality to institutions of social learning and institutional teaching of these findings can help to recognize human diversity and engage with it appropriately. The latter applies not only to social integration and recognition of group members who may have different standards than the mainstream of a society in relative terms (Esser, 1999; 2008; Verkuyten, 2007; Honneth, 2012; Windzio, 2016), but also in general, so that we humans continue to learn to *care* about each other beyond the realm of our primary cultural editing. Human beings are diverse, and we should celebrate this diversity (Henrich, 2020). Scientific research into human morality across cultures has already achieved a great deal, but there is still much to discover. In addition to gaining knowledge, it is also important for researchers to communicate their findings, and it is up to our social institutions to pass these findings on to future generations. In our view, it is a desirable goal of moral research that, in addition to gaining knowledge, these findings themselves also make a social contribution.

7.1.9. Contributions to the Field of Moral Research and Closing Remarks

Moral deviance is at the center of our investigations of the human moral mind. With this project, we have attempted to make **three independent contributions to the field of moral research**. We see our **first contribution** in the theoretical view of morality that we have developed. Deviance and conformity are at the core of the coalescence perspective of MFT and MaC, which unfolds against a gene-culture-co-evolutionary background. We see the **second contribution** of the present project in the measurement instruments we have developed. With the Morality as Cooperation—Deviance Relevance Scale (MaC-DRS), the binding vs. individualizing dilemma scenarios and the Moral Deviance Factorial Survey (MDFS), we have tried to contribute to the toolbox in moral research. We see the **third contribution** in the four empirical investigations

of the human moral mind that we have carried out. We were able to identify 8 different moral domains across cultures. Furthermore, we were able to show the cultural calibration of the moral mind and worked out cross-cultural tendencies towards individualizing morality. We examined both intuitive moral tendencies and deliberate moral tendencies across cultures and also took a close look at moral particularism/impartiality. Furthermore, we have identified a total of four different moral systems that guide cooperation in different cultural entities. Lastly, we were able to identify open research questions and, based on our analyses, also identify pathways for future cross-cultural investigations of the human moral mind.

After what has been a challenging journey for us, the readers may hopefully forgive us for attempting a foray into the lyrical and metaphorical for the last words of this thesis: The attempt to produce new knowledge is like pouring a liquid over a seemingly solid brick statue. Three consequences can be observed. Either the liquid runs down the statue and leaves no trace that it was ever poured over the statue after the next ray of sunshine. Or the tiny mineral particles in the liquid settle in the unfilled spaces, harden the statue and make it grow. Finally, and thirdly, the liquid may also gradually wash away loose elements and eventually shatter the statue. After that, we build a new statue out of bricks and pour the flowing essence of inventiveness over it again to observe the consequences revealed to us by time. We will see what the future holds for the contribution presented in this thesis and what consequences arise from what we have learned from our investigations of the human moral mind.

Bibliography

Aden, J., Bunina, A., & Vavrik, C. (2021). Die einfaktorielle Varianzanalyse für unabhängige Stichproben und der Kruskal-Wallis-Test im psychotherapiewissenschaftlichen Kontext-Empfehlungen für Anwendung und Interpretation. *Empfehlungen für Anwendung und Interpretation. SFU Forschungsbulletin*, *9*(1), 68-86. DOI:10.15135/2021.9.1.68-86

Abrams, D. (1999). Social identity, social cognition, and the self: The flexibility and stability of self-Categorization. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 197-229). Blackwell.

Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of personality and social psychology*, *78*(5), 906-912. DOI: 10.1O37//O022-3514.78.5.9O6

Abrams, D., Marques, J., Bown, N., & Dougill, M. (2002). Anti-norm and pro-norm deviance in the bank and on the campus: Two experiments on subjective group dynamics. *Group Processes & Intergroup Relations, 5*(2), 163–182. https://doi.org/10.1177/1368430202005002922

Abrams, D., & Hogg, M. A. (2004). Metatheory: Lessons from Social Identity Research. *Personality and Social Psychology Review, 8*(2), 98–106. https://doi.org/10.1207/s15327957pspr0802_2

Aldering, H., & Böhm, R. (2020). Parochial versus universal cooperation: Introducing a novel economic game of within- and between-group interaction. *Social Psychological and Personality Science*. 11(1),36-45. https://doi.org/10.1177/1948550619841627

Altemeyer, B. (2004). Highly dominating, highly authoritarian personalities. *The Journal of social psychology*, *144*(4), 421-448. https://doi.org/10.3200/SOCP.144.4.421-448

Amodio, D. (2014). The neuroscience of prejudice and stereotyping. *Nat Rev Neurosci*. 15, 670–682. https://doi.org/10.1038/nrn3800

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, 41(5), 319-329. https://doi.org/10.1016/j.evolhumbehav.2020.07.015

Arévalo-Iglesias, G. (2024). Intergenerational cohabitation and welfare attitudes among European young adults. *Journal of European Social Policy*. DOI: 10.1177/09589287241290761

Armbruster, S. (2021). *Politisches Bewusstsein und Soziale Medien*. Springer VS.

Asao, K., & Buss, D. M. (2016). The tripartite theory of machiavellian morality: Judgment, influence, and conscience as distinct moral adaptations. In T. K. Shackelford, & R. D. Hansen (Eds.). *The Evolution of Morality* (pp. 3-25). Springer.

Atari, M., Lai, H. C., & Dehghani, M. (2020a). Sex differences in moral judgements across 67 countries. Proc. R. Soc. B.287:20201201.2020120.1http://doi.org/10.1098/rspb.2020.1201

Atari, M., Graham, J., & Dehghani, M. (2020b). Foundations of morality in Iran. *Evolution and Human Behavior*, *41*(5), 367-384. https://doi.org/10.1016/j.evolhumbehav.2020.07.014

Atari, M., Haidt, J., Graham, J., Koleva, S., Stevens, S. T., & Dehghani, M. (2022a). Morality beyond the WEIRD: How the nomological network of morality varies across cultures. PsyAr-Xiv Preprint. https://doi.org/10.31234/osf.io/q6c9r

Atari, M., Reimer, N. K., Graham, J., et al., (2022b). Pathogens are linked to human moral systems across time and space. *Current Research in Ecological and Social Psychology*, 3, 100060. https://doi.org/10.1016/j.cresp.2022.100060

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*. 6(3),128–138. DOI: 10.1027/1614-2241/a000014

Auspurg, K., Hinz, T., & Liebig, S. (2009). Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey. *Methoden, Daten, Analysen (mda)*, *3*(1), 59-96. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-126659

Auspurg, K., & Hinz, T. (2014). *Factorial survey experiments* (Vol. 175). Sage Publications.

Auspurg, K., & Hinz, T. (2015). Multifactorial experiments in surveys. In *Experimente in den Sozialwissenschaften* (pp. 294-320). Nomos Verlagsgesellschaft mbH & Co. KG. doi.org/10.5771/9783845260433

Awad, E., Dsouza, S., Shariff, A., Rahwan, I., & Bonnefon, J. F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proceedings of the National Academy of Sciences*, *117*(5), 2332-2337. https://doi.org/10.1073/pnas.1911517117

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*. 80(4), 1095-1111. https://doi.org/10.2307/1960858

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, 40(1), 8-34. https://doi.org/10.1007/s11747-011-0278-x

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364–374. https://doi.org/10.1037/0022-3514.71.2.364

Bahrami-Rad, D., Beauchamp, J., Henrich, J., & Schulz, J. (2022). Kin-based institutions and economic development. *Available at SSRN*. https://ssrn.com/abstract=4200629

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R., Harwood, J., ... & Sibley, C. G. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and social Psychology bulletin*, *38*(12), 1629-1643. DOI: 10.1177/0146167212457953

Barmeyer, C. (2010). Kulturdimensionen und Kulturstandards. In: Barmeyer, C. I., Genkova, P., & Scheffer, J. (Hg.). *Interkulturelle Kommunikation und Kulturwissenschaft. Grundbegriffe, Wissenschaftsdisziplinen, Kulturräume* (pp. 93-127). Verlag Karl Stutz.

Bar-Tal, D. (1998). Group beliefs as an expression of social identity. In S. Worchel, J. F. Morales, D. Paez, & J.-C. Deschamps (Eds.), *Social identity: International perspectives* (pp. 93-113). Thousand Oaks, CA: Sage.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*. 5(4), 323-370. https://doi.org/10.1037/1089-2680.5.4.323

Baumeister, R. F. (2010). The self. In Baumeister, R. F., & Finkel, E. J. (Eds.). (2010). *Advanced social psychology: The state of the science* (pp. 139-176). Oxford University Press.

Baumeister, R. F. (2022). *The self explained: Why and how we become who we are*. Guilford Publications.

Baumgartner, H., & Weijters, B. (2015). Response biases in cross-cultural management. In S. Ng & A. Y. Lee (Eds.), *Handbook of culture and consumer behavior* (pp. 150-180). Oxford University Press.

Behr, D., Braun, M., & Dorer, B. (2016). Measurement instruments in international surveys. GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_006

Benninghaus H. (2007). *Deskriptive Statistik. Eine Einführung für Sozialwissenschaftler.* 11. Auflage. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90739-0

Beckmann, K., Glemser, A., Heckel, C. et al., (2016). *Demographische Standards*. eine gemeinsame Empfehlung des ADM, Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V., der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und des Statistischen Bundesamtes. Wiesbaden Ausgabe 2016, 6., überarb. Aufl. 2016

Berger, P. L., & Luckmann T. (2013). *Die gesellschaftliche Konstruktion der Wirklichkeit*. 25. Ed. Fischer Taschenbuch Verlag.

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural equation modeling*, *7*(4), 608-628. https://doi.org/10.1207/S15328007SEM0704_5

Bizumic, B., & Duckitt, J. (2008). "My group is not worthy of me": Narcissism and ethnocentrism. *Political Psychology*, *29*(3), 437-453. https://doi.org/10.1111/j.1467-9221.2008.00638.x

Bizumic, B., Duckitt, J., Popadic, D., Dru, V., & Krauss, S. (2009). A cross-cultural investigation into a reconceptualization of ethnocentrism. *European Journal of Social Psychology*, *39*(6), 871-899. https://doi.org/10.1002/ejsp.589

Bjørnskov, J. (2021). Civic honesty and cultures of trust. *Journal of Behavioral and Experimental Economics*, 92, (2021), 101693. https://doi.org/10.1016/j.socec.2021.101693

Blake, P., McAuliffe, K., Corbit, J. *et al.* (2015). The ontogeny of fairness in seven societies. *Nature*, 528, 258–261. https://doi.org/10.1038/nature15703

Boehnke, K., & Hadjar, A. (2015). Forschungsdesigns und statistische Verfahren. In Hurrelmann, K., Grundmann, M., & Walper, S. (Hrsg.), *Handbuch Sozialisationsforschung* (pp. 320-334). Beltz.

Boehnke, K. (2022a). Let's compare apples and oranges! A plea to demystify measurement equivalence. *American Psychologist*, 77(9), 1160. https://doi.org/10.1037/amp0001080

Bogner, K., & Landrock, U. (2016). Response Biases in Standardised Surveys. GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_016

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In Kreckel R. (Hg.), *Soziale Ungleichheiten* (Soziale Welt Sonderband 2) (pp. 183-198). Originalbeitrag, übersetzt von Reinhard Kreckel.

Bourdieu, P. (2014). *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft*. Suhrkamp taschenbuch wissenschaft. 24 Aufl.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *PNAS*, 108(2),10918–10925. https://doi.org/10.1073/pnas.1100290108

Brandt, H., & Moosbrugger, H. (2020). Planungsaspekte und konstruktionsphasen von tests und fragebogen. In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 39-66). 3. Auflage. Springer Verlag.

Brandt, H. (2020). Explorative Faktorenanalyse (EFA). In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 575-614). 3. Auflage. Springer Verlag.

Brewer, M. B., & Gardner, W. (1996). Who is this "We"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71*(1), 83–93. https://doi.org/10.1037/0022- 3514.71.1.83

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of cross-cultural psychology*, *1*(3), 185-216.

Brosnan, S. F., & De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, *425*(6955), 297-299. https://doi.org/10.1038/nature01963

Brown, G. T., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, *40*(1), 66-90. https://doi.org/10.1080/1743727X.2015.1070823

Brown, G. R., Cross, C. P., & Laland, K. (2022). Evolutionary approaches to human behavior. In J. J. Bolhuis, L-A. Giraldeau, & J. A. Hogan (Eds.). *The Behavior of animals, function, and evolution. Second Edition* (pp- 456-488). Wiley & Sons.

Brüntrup, G. (2004). Zur Kritik des Funktionalismus. In Köhler, W. R., & Mutschler, H-D. (eds.): *Ist der Geist berechenbar?* (pp. 58-76). Darmstadt: 2004.

Cai, H. J., Huang, Z. H., & Jing, Y. M. (2019). Living in a changing world: The change of culture and psychology. In D. Matsumoto & H. C. Hwang (Eds.), *Oxford handbook of culture and psychology* (pp.1070−1116), 2nd ed., Oxford University Press.

Carlo, G., Christ, C., Laible, D., & Gilseven, Z. (2016). An evolving and developing field of study: Prosocial morality from a biological, cultural and developmental perspective. In T. K. Shackelford, & R. D. Hansen (Eds.). *The Evolution of Morality* (pp. 53- 76). Springer.

Carter, N. L., & Weber, M. J. (2010). Not pollyannas: High generalized trust predicts lie detection ability. *Social Psychological and Personality Science*, 1(3), 274-279. DOI: 10.1177/1948550609360261

Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika, 30*(1), 99–121. https://doi.org/10.2333/bhmk.30.99

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of cross-cultural psychology*, *31*(2), 187-212. DOI:10.1177/0022022100031002003

Cheung, M. W. L., & Chan, W. (2002). Reducing uniform response bias with ipsative measurement in multiple-group confirmatory factor analysis. *Structural Equation Modeling*, *9*(1), 55-77. https://doi.org/10.1207/S15328007SEM0901_4

Choi, J.-K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *SCIENCE*, 318(5850), 636– 640. doi: 10.1126/science.1144237

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 151–192), 4th ed., McGraw-Hill.

Cieciuch, J., Davidov, E., Schmidt, P., & Algesheimer, R. (2019). How to obtain comparable measures for cross-national comparisons. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *71*(S1), 157-186. DOI: https://doi.org/10.1007/s11577-019-00598-7

Cole, D. P. (2003). Where have the Bedouin gone?. *Anthropological quarterly*, 235-267. DOI:10.1353/anq.2003.0021

Cooper, M. (2020). Neoliberalism's family values: welfare, human capital, and kinship. In D. Plehwe, Q. Slobodian, & P. Mirowski (Eds.). *Nine lives of neoliberalism* (pp. 95-119). Verso Books.

Cooper, D. J., Kagel, J. H. (2015). Other-Regarding Preferences: A Selective Survey of Experimental Results. In Kagel, J. H., Roth, A. E. (Hg.). *The handbook of experimental economics Vol. 2* (pp. 217-289).

Cosmides, L., & Tooby, J. (1997). *Evolutionary psychology: A primer.* Center for Evolutionary Psychology, Santa Barbara. http://cogweb.ucla.edu/ep/EP-primer.html

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Science*, 2011, 15(5), 218-240. https://doi.org/10.1016/j.tics.2011.03.003

Chudek, M. Muthukrishna, M. Henrich, J. (2016). Cultural evolution. In D. M. Buss (Eds.). *The Handbook of Evolutionary Psychology.* Vol. 2: *Integrations* (pp. 749-769). Wiley & Sons.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: A standardized stimulus database of scenarios based on moral foundations theory. *Behavior research methods*, *47*(4), 1178-1198. DOI 10.3758/s13428-014-0551-2

Cohn, A., Maréchal, A. M., Tannenbaum, D., & Zünd, C. L. (2019). Civic honesty around the globe. *Science*, 365(6448),70–73. DOI: 10.1126/science.aau8712

Cottam, M., L., Dietz-Uhler, B., Mastors E., Preston T. (2010). *Introduction to Political Psychology*. Second Edition. Psychology Press. Taylor and Francis Group. New York and Hove

Creanza, N., Kolodny, O., & Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *PNAS*, vol. 114 (30), 7782-7789. https://doi.org/10.1073/pnas.1620732114

Crocker, J., Canevello, A., & Brown, A. A. (2017). Social motivation: Costs and benefits of selfishness and otherishness*. Annu. Rev. Psychol.* 2017. 68, 299–325. https://doi.org/10.1146/annurev-psych-010416-044145

Cross S. E., Hardin E. E., & Gercek-Swing, B. (2011). The what, how, why, and where of self-construals. *Personality and Social Psychology Review*, 15(2), 142-179. https://doi.org/10.1177/1088868310373752

Curtin, C. M., Barret, H. C., Bolyanats, A., et al., (2020). Kinship intensity and the use of mental states in moral judgment across societies. *Evolution and Human Behavior*, 41, 415-429. https://doi.org/10.1016/j.evolhumbehav.2020.07.002

Curry, S. O. (2016) Morality as cooperation: A problem-centered approach. In T. K. Shackelford, & R. D. Hansen (Eds.). *The Evolution of Morality* (pp. 27-51). Springer.

Curry, S. O., Mullins, D. A., & Whitehouse, H. (2019a). Is it good to cooperate? Testing morality as cooperation in 60 societies. *Current Anthropology*. 60(1), 47-69. https://doi.org/10.1086/701478

Curry, S. O., Chesters, M. J., & Van Lissa, C. J. (2019b). Mapping morality with a compass: Testing the theory of ´morality-as-cooperation with a questionnaire. *Journal of Research in Personality*. 78, 106-124. https://doi.org/10.1016/j.jrp.2018.10.008

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual review of sociology*, *40*, 55-75. https://doi.org/10.1146/annurev-soc-071913-043137

de Haas, H. (2014). What drives human migration. *Migration: A COMPAS Anthology. Oxford: International Migration Institute.* http://compasanthology. co. uk/drives-human-migration.

de Haas, H., Castles, S., & Miller, M. J. (2020). *The age of migration: International population movements in the modern world*. 6th edition. macmillan international. RED GLOBE PRESS.

De Dreu, C. K., Fariña, A., Gross, J., & Romano, A. (2022). Prosociality as a foundation for intergroup conflict. *Current opinion in psychology*, *44*, 112-116. https://doi.org/10.1016/j.copsyc.2021.09.002

De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annu. Rev. Psychol.*, *59*, 279-300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Diekmann, A. (2013). *Spieltheorie. Einführung, Beispiele, Experimente*. Rowohlts Enzyklopädie.

Diekmann A., & Lindenberg, S. (2015). Cooperation: Sociological aspects. In J. D. Wright, (Ed.), *International Encyclopaedia of the Social & Behavioral Sciences* (pp. 862-866), 2nd Edition, Elsevier (revised version of Diekmann and Lindenberg 2002).

Dimaggio, P., & Markus, H. R. (2010). Culture and social psychology: Converging perspectives. *Social Psychology Quarterly*, 73(4), 347-352. https://doi.org/10.1177/0190272510389010

DiPrete, T. A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology*, *34*(1), 271-310. https://doi.org/10.1111/j.0081-1750.2004.00154.x

Dolata, U. (2011). Soziotechnischer Wandel als graduelle Transformation. *Berliner Journal für Soziologie*, *21*(2), 265-294. DOI 10.1007/s11609-011-0153-0

Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V., M. (2010). Prejudice, stereotyping, and discrimination: Theoretical and empirical overview. In Dovidio, John, F., Hewstone, Miles, & Glick, Peter (eds.). *The SAGE Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 3-29). London, England: Sage.

Elias, N. (1976). *Über den Prozeß der Zivilisation. Soziogenetische und psychogenetische Untersuchungen. Erster Band. Wandlungen des Verhaltens in den westlichen Oberschichten des Abendlandes.* Frankfurt: Suhrkamp taschenbuch wissenschaft 1. Aufl.

Ellemers, N., Van Der Toorn, J., Paunov, Y., & Van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*, *23*(4), 332-366. DOI: 10.1177/1088868318811759

Engelmann, H. B., Schmid, B., De Dreu, C. K. W., Chumbley, J., & Fehr, E. (2019). On the psychology and economics of antisocial personality. *PNAS*, 2019, 116(26), 12781-12786. https://doi.org/10.1073/pnas.1820133116

Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics* (2019), 953–1019. doi:10.1093/qje/qjz001

Esser, H. (1999). Die Situationslogik ethnischer Konflikte: Auch eine Anmerkung zum Beitrag „Ethnische Mobilisierung und die Logik von Identitätskämpfen "von Klaus Eder und Oliver Schmidtke (ZfS 6/98). *Zeitschrift für Soziologie*, *28*(4), 245-262.

Esser, H. (2002a). *Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln*. Campus-Verlag.

Esser, H. (2002b). Ehekrisen: Das (Re-) Framing der Ehe und der Anstieg der Scheidungsraten/Marital Crises: The (Re-) Framing of Marriage and the Increase in Divorce Rates. *Zeitschrift für Soziologie*, *31*(6), 472-496.

Esser, H. (2008). Assimilation, ethnische Schichtung oder selektive Akkulturation? Neuere Theorien der Eingliederung von Migranten und das Modell der intergenerationalen Integration. *Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderhefte*, *60*(48), 81-107.

Esser, H. (2010). Sinn, Kultur, Verstehen und das Modell der soziologischen Erklärung In Monika Wohlrab-Sahr (Ed.). *Kultursoziologie. Paradigmen – Methoden – Fragestellungen* (pp. 309-335).VS Verlag für Sozialwissenschaften.

Falk, A., & Szech, N. (2013). Morals and marktes. *SCIENCE*, 340, 707-711. DOI: 10.1126/science.1231566

Falk, A., & Szech, N. (2015). Institutions and morals: A reply. *European Journal of Political Economy*. 40, 391-394. https://doi.org/10.1016/j.ejpoleco.2015.07.003

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *TRENDS in Cognitive Sciences*, 8(4), 185-190. doi:10.1016/j.tics.2004.02.007

Fischer, R. (2004). Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *Journal of Cross-Cultural Psychology*, *35*(3), 263-282. DOI: 10.1177/002202210426412

Fischer, R., & Milfont, T. L. (2010). Standardization in psychological research. *International Journal of Psychological Research*, *3*(1), 88-96. DOI:10.21500/20112084.852

Fischer, G., & Formann, A. (2007). Methodenlehre. In U. Kastner-koller, & O. Deimann, (Hg.), *Psychologie als Wissenschaft* (pp. 59-75). facultas.wuv.

Fiske, S. T. (2000). Stereotyping, prejudice, and discrimination at the seam between the centuries: Evolution, culture, mind, and brain. *European Journal of Social Psychology, 30*(3), 299–322. https://doi.org/10.1002/(SICI)1099-0992(200005/06)30:3<299::AID-EJSP2>3.0.CO;2-F

Flinkenflogel, N., Novin, S., Huizinga, M., & Krabbendam, L. (2017). Gender moderates the influence of self construal priming on fairness considerations. *Front. Psychol.* 8:503. doi: 10.3389/fpsyg.2017.00503

Frings, D., Hurst, J., Cleveland, C., Blascovich, J., Abrams, D. (2012). Challenge, threat, and subjective group dynamics: reactions to normative and deviant group members. *Group Dyn. Theory Res. Pract.*16, 105–121. DOI: 10.1037/a0027504

Frith, C., & Frith, U. (2005). Theory of mind. *Current biology*, *15*(17), R644-R645. doi: 10.1016/j.cub.2005.08.041.

Fuchs-Heinritz, W., Klimke, D., Lautmann, R., Rammstedt, O., Staheli, U., Weischer, C., & Wienold, H. (Hrsg.) (2011). *Lexikon zur Soziologie*. 5. Auflage. Springer VS.

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American journal of epidemiology*, *173*(7), 761-767. DOI: 10.1093/aje/kwq439

Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "we" relationship: Self-construal priming mirrors cultural differences in judgement. *Psychological Science*, 10(4), 321-326. https://doi.org/10.1111/1467-9280.00162

Gavrilets, S., & Richardson, P. J. (2017) Collective action and the evolution of social norm internalization. *PNAS*, 2017, 144(23), 6068-6073. https://doi.org/10.1073/pnas.1703857114

Gazzaniga, M., Heatherton, T., & Halpern, D. (2017). *Psychologie*. 1. Aufl. Beltz.

Gäde, J. C., Schermelleh-Engel, K., & Brandt, H. (2020a). Konfirmatorische Faktorenalayse (CFA). In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 615-660). 3. Auflage. Springer Verlag.

Gäde, J. C., Schermelleh-Engel, K., & Werner, C. S. (2020b). Klassiche methoden der reliabilitätsschätzung. In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 305-334). 3. Auflage. Springer Verlag.

Gelfand, M., Raver, J. L., Nishii, L., *et al.* (2011). Differences Between Tight and Loose Cultures: A 33- Nation Study. *SCIENCE*, 332, 1100 (2011). DOI: 10.1126/science.1197754

Gelfand, M., Harrington, J. R., & Jackson, J. C. (2017). The strength of social norms across human groups. *Perspectives on Psychological Science*,12(5), 800–809. DOI: 10.1177/1745691617708631

Gintis, H. (2003). The hitchhiker´s guide to altruism: Gene-culture coevolution and the internalization of norms. *J. theor. Biol.* 2003, 220:407-418. https://doi.org/10.1006/jtbi.2003.3104

Giesselmann, M., & Windzio, M. (2013). *Regressionsmodelle zur Analyse von Paneldaten* (Vol. 1). Springer-Verlag.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029–1046. https://doi.org/10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. J Pers Soc Psychol. *Author manuscript*; available in PMC 2012 Aug 1. 1-39. Published in final edited form as: J Pers Soc Psychol. 2011 Aug; 101(2): 366–385.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, p., & Ditto, P. H., (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55-130. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham J., Meindl P., Beall, E., Johnson K. M., & Zhang, L. (2016). Cultural differences in moral judgement and behavior, across and within societies. *Current Opinion in Psychology* 2016, 8, 125-130. https://doi.org/10.1016/j.copsyc.2015.09.007

Greenwald, A. G., & Lai, C. K. (2020). Implicit Social Cognition. *Annual Review of Psychology*. 2020. 71, 419-445. https://doi.org/10.1146/annurev-psych-010419-050837

Greshoff, R. (2008). Verstehen und Erklären bei Hartmut Esser. In Greshoff, R., Kneer, G., & Schneider, W. L. (Eds.). *Verstehen und erklären. Sozial-und kulturwissenschaftliche Perspektiven* (pp. 413-443)*. München: Fink*.

Greshoff, R., Huinink, J., & Schimank, U. (2011). *Aggregation dynamics of multi-level actor constellations: Introductory paper*. Welfare Societies Working Papers, 2. Bremen: Forschungsverbund "Welfare Societies", 2011

Habermas, J. (1995). *Theorie des kommunikativen Handelns*. *Band 1: Handlungsrationalität und gesellschaftliche Rationalisierung*. Suhrkamp taschenbuch wissenschaft. Erste Auflage 1995. Surkamp Verlag.

Hadjar, A., Boehnke, K., Knafo, A., Daniel, E., Musiol, A. L., Schiefer, D., & Möllering, A. (2012). Parent-child value similarity and subjective well-being in the context of migration: An exploration. *Family Science*, *3*(1), 55-63. DOI:10.1080/19424620.2011.671502

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., *et al.,* (2022). World Values Survey: Round Seven - Country-Pooled Datafile Version 4.0. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi:10.14281/18241.18

Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism–collectivism in the United States and Japan. *Personality and social psychology review*, *16*(1), 3-24. https://doi.org/10.1177/10888683114115

Han S., & Humpreys, G. (2016). Self-construal: a cultural framework for brain function. *Curr Opin Psychol*. 2016, 8:10-14. https://doi.org/10.1016/j.copsyc.2015.09.013

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. In Adler, J. E., & Rips, L. J. (Eds.), *Reasoning. Studies of Human Inference and Its Foundations* (pp. 1024-1052). Cambridge University Press.

Haidt J. (2003). The moral emotions. In Davidson R. J., Scherer K. R., Goldsmith H. H. (Eds.), *Handbook of affective sciences* (pp. 852-870). New York, NY: Oxford University Press.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, 3(1), 65–72. https://doi.org/10.1111/j.1745-6916.2008.00063.x

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.). *The innate Mind*. Vol. 3. (pp. 367-391).

Haidt, J., Graham, J., & Joseph, C. (2009). Above and Below Left–Right: Ideological Narratives and Moral Foundations. *Psychological Inquiry*, 20, 110-119. DOI: 10.1080/10478400903028573

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological methods*, *15*(3), 234. doi: 10.1037/a0019623

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annu. Rev. Psychol*. 2017, 68,155-186. https://doi.org/10.1146/annurev-psych-010416-044201

He, J., & van de Vijver, F. (2012). Bias and equivalence in cross-cultural research. *Online readings in psychology and culture*, *2*(2), 8. DOI:10.9707/2307-0919.1111

He, J., Chi, R., & van de Vijver, F. J. (2021). People Use Scales Differently: Dealing with Survey Response Styles in Cross-Cultural Research1. *Journal of Intercultural Communication & Interactions Research*, *1*(1), 83-100. https:/doi.org/10.3726/jicir.2021.1.0006

Hemmerich, W. (2016). StatistikGuru: Rechner zur Adjustierung des α-Niveaus. Retrieved from https://statistikguru.de/rechner/adjustierung-des-alphaniveaus.html

Hemmerich, W. (2020). StatistikGuru: Bonferroni–Holm Korrektur. Retrieved from https://statistikguru.de/lexikon/bonferroni-holm-korrektur.html

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and human behavior*, *30*(4), 244-260. https://doi.org/10.1016/j.evolhumbehav.2009.03.005

Henrich, J., Boyd, R., Bowles, S., et al. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral & Brain Sciences*, 28, 795-815. doi:10.1017/S0140525X05000142

Henrich, J., & McElreath, R. (2007). Dual inheritance theory: The evolution of human cultural capacities and cultural evolution. In R. Dunbar, & L. Barrett (Eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 555-570). Oxford University Press.

Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about cultural Evolution. *Hum Nat* (2008), 19:199 – 137. https://doi.org/10.1007/s12110-008-9037-1

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). The weirdest people in the world? (*RatSWD Working Paper Series, 139*). Rat für Sozial- und Wirtschaftsdaten (RatSWD).

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2010b). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484. http://www.jstor.org/stable/40544708

Henrich, J., P. (2016). *The secret of our success. How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Henrich, J. (2020). *The weirdest people in the world: How the West became psychologically peculiar and particularly prosperous.* Penguin Books, Penguin Random House UK.

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology,* 72, 207-240. https://doi.org/10.1146/annurev-psych-081920-042106

Hewstone, M., & Cairns, E. (2001). Social psychology and intergroup conflict. In D. Chirot & M. E. P. Seligman (Eds.), *Ethnopolitical warfare: Causes, consequences, and possible solutions* (pp. 319–342). American Psychological Association. https://doi.org/10.1037/10396-020

Heyes, C. & Moore, R. (2021). Henrich, Heyes and Tomasello on the cognitive foundations of cultural evolution. In Jamshid t. J., Kendal, J., & Kendal, R. (eds), *The Oxford Handbook of Cultural Evolution* (online edn, Oxford Academic, 23 Feb. 2023), https://doi.org/10.1093/oxfordhb/9780198869252.001.0001

Hoerster, N. (2022). *Was ist Moral? Eine ganz kleine Einführung: [Was bedeutet das alles?]*. Reclam Verlag.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skita L. J. (2014). Morality in everyday life. *Science,* 345(6202), 1340-1343. DOI: 10.1126/science.1251560

Hogg, M. A., Abrams, D., Otten, S., & Hinkle, S. (2004). The social identity perspective: Intergroup relations, self-conception, and small groups. *Small group research*, *35*(3), 246-276. DOI: 10.1177/1046496404263424

Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication theory*, *16*(1), 7-30. https://doi.org/10.1111/j.1468-2885.2006.00003.x

Hogg, M. A., Adelman, J. R., & Blagg, R. D. (2010). Religion in the face of uncertainty: An uncertainty-identity theory account of religiousness. *Personality and social psychology review*, *14*(1), 72-83. DOI: 10.1177/1088868309349692

Höfling, V., & Moosbrugger, H. (2020). Standards für psychologisches testen. In Moosbrugger, H., & Kelava, A. (Hrsg.). Testtheorie und Fragebogenkonstruktion (pp. 197-216). 3. Auflage. Springer Verlag.

Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review, 65*(2), 117–127. https://doi.org/10.1037/h0042501

Honneth, A. (2012). *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte*. Suhrkamp taschenbuch wissenschaft. 7.Aufl.

Horne, C., & Cutlip, A. (2002). Sanctioning costs and norm enforcement: An experimental test. *Rationality and Society*, *14*(3), 285-307. https://doi.org/10.1177/1043463102014003002

Houghten, D., P. (2009). *Political psychology. situations, individuals, and cases*. 270 Madison Ave: Routledge.

Huddy, L., Gutting, R., Feldman S. (2015). Intergruppenvorurteile und Stereotype. In Zmerli, S., Feldman, O. (Hrsg.) *Politische Psychologie* (pp. 253-270), 2015.

Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social work and social sciences review*, *11*(1), 36-51.

Hügli, A., & Lübcke, P. (2013). *Philosophie-Lexikon. Personen und Begriffe der abendländischen Philosophie von der Antike bis zur Gegenwart*. Rowohlts enzyklopädie.

Huinink, J. (1995). *Warum noch Familie? Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft*. Campus Verlag.

Huinink, J., & Schnettler, S. (2024). Zur Evolution der sozialen Ungleichheit. In *Handbuch Sozialstrukturanalyse* (pp. 1-36). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-39759-3_58-1

Hunsberger, B., & Jackson, L. M. (2005). Religion, meaning, and prejudice. *Journal of social issues*, *61*(4), 807-826. https://doi.org/10.1111/j.1540-4560.2005.00433.x

Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the moral foundations questionnaire across 27 countries. *Assessment*, 27(2) 365–372. https://doi.org/10.1177/1073191118817916

ISSP Research Group (2018). International Social Survey Programme: Religion III - ISSP 2008. *GESIS Data Archive, Cologne. ZA4950 Data file Version 2.3.0, https://doi.org/10.4232/1.13161.*

Iyer, R., Koleva, S., Graham, J., Ditto, P., Haid, J. (2012). Understanding Libertarian Morality: The Psychological Dispositions of Self-Identified Libertarians. *PLoS ONE* 7(8): e42366. doi:10.1371/journal.pone.0042366

Jann, B. (2005). *Einführung in die Statistik*. (Mohr, A., (Ed.), 2th Ed.). Oldenbourg Wissenschaftsverlag GmbH

Jaspers, K. (1956). Das Kollektiv und der Einzelne. In *Mensch und Menschlichkeit. Eine Vortragsreihe* (pp. 65-76). Alfred Kröner Verlag Stuttgart.

Jessen, P. L., Huinink, J., Kühnen, U., & Boehnke, K. (2022). *Pilot study as part of the project: "Variances in Human Morality from a Cross-Cultural Perspective - An Inquiry on the Role of Self-Construal"*. PsychArchives. https://doi.org/10.23668/psycharchives.12183

Jessen, P. L., Huinink, J., Kühnen, U., & Boehnke, K. (2023). *Validation Study: Morality as Cooperation—Deviance Relevance Scale (MaC-DRS)*. PsychArchives. https://doi.org/10.23668/psycharchives.13059.

Jessen, P. L., Huinink, J., Kühnen, U., Murayama, A., Skitka, L., Teas, P., Miniesy, R., Rabie, D., Ahmed, S., & Boehnke, K. (2024). *Research Plan - Variances in Human Morality and Self-Construal from a Cross-Cultural Perspective - adjusted*. PsychArchives. https://doi.org/10.23668/psycharchives.14630

Jetter, M., Laudage, S., & Stadelmann, D. (2019). The intimate link between income levels and life expectancy: global evidence from 213 years. *Social Science Quarterly*, *100*(4), 1387-1403. https://doi.org/10.1111/ssqu.12638

Judd, C. M., Yzerbyt, V. Y., & Muller, D. (2014). Mediation and moderation. In H. T. Reis, & C. M. Judd (Eds.) *Handbook of research methods in social and personality psychology* (653-676), *2end Ed.*, Cambridge University Press.

Kaasa, A., & Minkov, M. (2020). Are the world's national cultures becoming more similar?. *Journal of cross-cultural psychology*, *51*(7-8), 531-550. DOI: 10.1177/0022022120933677

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2), pp. 263-292.

Kahneman, D. (2011). *Schnelles Denken, Langsames Denken*. 15. Aufl. Siedler Verlag, München.

Kant, I. (1788/2011). *Die drei Kritiken - Kritik der reinen Vernunft, Kritik der praktischen Vernunft, Kritik der Urteilskraft.* Anaconda Verlag Köln, 2001

Kaplan, T. J., Gimbel, S., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports*. 6:39589. DOI: 10.1038/srep39589

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of personality and social psychology*, *91*(1), 16. https://doi.org/10.1037/0022-3514.91.1.16

Keblusek, L., Giles, H., & Maass, A. (2017). Communication and group life: How language and symbols shape intergroup relations. *Group processes & Intergroup relations*, *20*(5), 632-643. https://doi.org/10.1177/1368430217708864

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology, 97*(2), 236–255. https://doi.org/10.1037/a0015999

Kitayama, S., & Imada, T. (2010). Implicit independence and interdependence: a cultural task analysis. In B. Mesquita, L. F. Barrett, & E. R. Smith (Eds.). *The Mind in Context* (pp. 174-200), Guilford Press.

Kitayama, S., & Salvador, C. E. (2024). Cultural psychology: Beyond east and west. *Annual Review of Psychology*, *75*. https://doi.org/10.1146/annurev-psych-021723-063333

Kirkland, K., Crimston, C. R., Jetten, J., *et al.* (2023). Moral expansiveness around the world: The role of societal factors across 36 countries. *Social Psychological and Personality Science*, 14(3) 305–318. DOI: 10.1177/19485506221101767

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., ... & Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social cognitive and affective neuroscience*, *5*(4), 378-384. https://doi.org/10.1093/scan/nsq005

Kohler, U., & Kreuter, F. (2017). *Datenanalyse mit Stata* [Data Analysis with Stata]. *Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung*. 5th edition. De Gruyter Oldenbourg.

Kohlberg, L. (1973). The claim to moral adequacy of a highest stage of moral judgment. *The journal of philosophy*, *70*(18), 630-646. https://doi.org/10.2307/2025030

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory into practice*, *16*(2), 53-59. https://www.jstor.org/stable/1475172

Konrad, K. (2020). Lautes Denken. In: Mey, G. & Mruck, K. (Hrsg.), Handbuch Qualitative Forschung in der Psychologie (S. 373-379). https://doi.org/10.1007/978-3-658-26887-9_41

Kotthoff, H. (2003). Aspekte der Höflichkeit im Vergleich der Kulturen. *Muttersprache*, 4. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-9330

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., Couper, M. (2004). Psychological research online. Report of board of scientific affairs `advisory group on the conduct of research on the internet. *American Psychologist Association*, 59(2), 105-117. doi: 10-1037/0003-066X.59.2.105

Krys, K., Vignoles, V. L., de Almeida, I., & Uchida, Y. (2022) Outside the "cultural binary": Understanding why Latin America collectivist societies foster independent selves. *Perspectives on Psychological Science*, 17(4) 1166–1187. https://doi.org/10.1177/17456916211029632

Kurzban, R., Tooby, J., Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Science*, 98(26), 15387-15392. doi: 10.1073/pnas.251541498

Kurzban R., Burton-Chellew M., N., & West S., A. (2015). The evolution of altruism in humans. *Annu. Rev. Psychol.* 66, 575–99. https://doi.org/10.1146/annurev-psych-010814-015355

Kurzban, R. & DeScioli, P. (2016) Morality. In D. M. Buss (Eds.). *The Handbook of Evolutionary Psychology.* Vol. 2: *Integrations* (pp. 770-787). Wiley & Sons.

Kühnen, U., & Hannover, B. (2003). Kultur, selbstkonzept und kognition. *Zeitschrift für Psychologie*, 211, 114(4), 212–224. doi:10.1026//0044-3409.211.4.212

Kühnen, U., & Kitayama, S. (2020). Agency-Detection. In T. K. Shackelford, V. A. Weekes-Shackelford (eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1-8), https://doi.org/10.1007/978-3-319-16999-6_3011-1

Kühnen, U., & Kitayama, S. (2024). Culture, social class and the dynamics of the self. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 1-18. https://doi.org/10.1007/s11577-024-00935-5

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses. *European Journal of Social Psychology*, 44(7), 701-710. https://doi.org/10.1002/ejsp.2023

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), 33267. https://doi.org/10.1525/collabra.33267

Lang M., Purzycki, B G., Apicella, C. L. *et al.,* (2019). Moralizing gods, impartiality and religious parochialism across 15 societies. *Proc. R. Soc. B.* 286: 20190202. https://doi.org/10.1098/rspb.2019.0202

Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication theory*, 15(2), 127-147. https://doi.org/10.1111/j.1468-2885.2005.tb00329.x

Layes, G. (2003). Kulturdimensionen. In A. Thomas, & E.u. Kinast u.a. (Hg.), *Handbuch interkulturelle Kommunikation und Kooperation. Bd. 1: Grundlagen und Praxisfelder* (pp. 60-73*).* Vandenhoeck & Ruprecht.

Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 64, 123-153. DOI 10.1007/s11577-012-0158-5

Lehman, D., R., Chiu, C., & Schaller, M. (2004). Psychology and culture. *Annu. Rev. Psychol.* 55: 689-714. https://doi.org/10.1146/annurev.psych.55.090902.141927

Lenzner, T., Schick, L., Hadler, P. et al. (2022): Zusammenhaltspanel FGZ: Welle 2 – Fragen zum Klimawandel, Antisemitismus und zur Gleichstellung der Geschlechter. Kognitiver Online-Pretest. GESIS Projektbericht. Version: 1.0. GESIS - Pretestlabor. Text. http://doi.org/10.17173/pretest115

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., ... & van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research*, 110, 102805. https://doi.org/10.1016/j.ssresearch.2022.102805

Leung, A. K.-Y., & Cohen, D. (2011). Within- and between-culture variation: Individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of Personality and Social Psychology*. Advance online publication. doi: 10.1037/a0022151

Lippke, S., Fischer, M. A., & Ratz, T. (2021). Physical activity, loneliness, and meaning of friendship in young individuals–a mixed-methods investigation prior to and during the COVID-19 pandemic with three cross-sectional studies. *Frontiers in psychology*, 12, 617267. https://doi.org/10.3389/fpsyg.2021.617267

Liu, X. S. (2011). Covariate imbalance and precision in measuring treatment effects. *Evaluation Review*, 35(6), 627-641. DOI: 10.1177/0193841X12439195

Lohmann, H. (2010). Nicht-Linearität und Nicht-Additivität in der multiplen Regression: Interaktionseffekte, Polynome und Splines. In C. Wolf, & H. Best (Eds.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (pp. 677-706). VS Verlag.

Lomazzi, V. (2018). Using Alignment Optimization to Test the Measurement Invariance of Gender Role Attitudes in 59 Countries. *Methods, data, analyses: a journal for quantitative methods and survey methodology (*mda), 12(1), 77-103. https://doi.org/10.12758/mda.2017.09

Mahoney, J. (2000). Path dependence in historical sociology. *Theory and society*, *29*(4), 507-548. https://www.jstor.org/stable/3108585

Mahoney, J. & Thelen, K. (2010). A Theory of Gradual Institutional Change. In Mahoney, J. & Thelen, K. (Eds.), *Explaining institutional change* (pp. 1-37). Cambridge University Press.

Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology*, *72*(1), 293-318. https://doi.org/10.1146/annurev-psych-072220-104358

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Markus, H., & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross-Cultural Psychology,* 29, 63–87. https://doi.org/10.1177/0022022198291004

Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, 5(4), 420-430. https://doi.org/10.1177/1745691610375557

Markus, H. R., & Schwartz, B. (2010). Does choice mean freedom and well-being?. *Journal of Consumer Research*, *37*(2), 344-355. https://doi.org/10.1086/651242

Marques, J., Abrams, D., & Serôdio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology, 81*(3), 436–447. https://doi.org/10.1037/0022-3514.81.3.436

Mauss, M. (1968). *Die Gabe, Form und Funktion des Austauschs in archaischen Gesellschaften*. Suhrkamp Verlag Frankfurt am Main. 1968.

Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. *Tutorials in quantitative methods for psychology*, *3*(2), 51-59.

McDonald, M. M., Navarrete, C. D., & Van Vugt, M. (2012). Evolution and the psychology of intergroup conflict: The male warrior hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 670-679.

McKee, P., Kim, H. E., Tang, H., Everett, J. A., Chituc, V., Gibea, T., ... & Sinnott-Armstrong, W. (2024). Does it matter who harmed whom? A cross-cultural study of moral judgments about harm by and to insiders and outsiders. *Current psychology*, *43*(9), 7997-8007. https://doi.org/10.1007/s12144-023-04986-3

McLaren, L. M. (2003). Anti-Immigrant Prejudice in Europe: Contact, Threat Perception, and Preferences for the Exclusion of Migrants. *Social Forces, 81*(3), 909–936. https://doi.org/10.1353/sof.2003.0038

Mesoudi, A. (2017). Pursuing Darwin´s curious parallel: Prospects for a science of cultural evolution. *PNAS*, 114 (30): 7853-7860. doi:10.1073/pnas.1620741114

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution?. *Proceedings of the Royal Society B*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of psychological research*, *3*(1), 111-130.

Minkov, M. (2017). Middle responding: An unobtrusive measure of national cognitive ability and personality. *Personality and Individual Differences*, *113*, 187-192. https://doi.org/10.1016/j.paid.2017.03.041

Minkov, M., Kaasa, A., & Welzel, C. (2021). Economic development and modernization in Africa homogenize national cultures. *Journal of Cross-Cultural Psychology*, *52*(8-9), 801-821. DOI: 10.1177/00220221211035495

Minkov, M., & Kaasa, A. (2022). Do dimensions of culture exist objectively? A validation of the revised Minkov-Hofstede model of culture with World Values Survey items and scores for 102 countries. *Journal of International Management*, *28*(4), 100971. https://doi.org/10.1016/j.intman.2022.100971

Miller, J. G., Bersoff, D. M., & Harwood, R. L. (1990). Perceptions of social responsibilities in India and in the United States: Moral imperatives or personal decisions? *Journal of Personality and Social Psychology, 58*(1), 33–47. https://doi.org/10.1037/0022-3514.58.1.33

Moebius, S. (2006). Die Gabe—ein neues Paradigma der Soziologie? Eine kritische Betrachtung der MAUSS-Gruppe. *Berliner Journal für Soziologie*, *16*, 355-370. DOI https://doi.org/10.1007/s11609-006-0031-3

Mooijman M., Meindl P., Oyserman, D., Monterosso J., Dehghani, M., & Doris, J. M., (2017). Resisting temptation for the good of the group: Binding moral values and the moralization of self-control. *Journal of Personality and Social Psychology*, 1-15. http://dx.doi.org/10.1037/pspp0000149

Moore, C., Detert, J. R., Trevino, L. K., Baker, V. L., & Mayer D., M (2012). Why employees do bad things: Moral Disengagement and Unethical Organizational Behavior. *Personnel Psychology*, 2012, 1-48. https://doi.org/10.1111/j.1744-6570.2011.01237.x

Moore, C. (2015). Moral Disengagement. *Current Opinion in Psychology*, 2015, 6, 199–204. DOI:10.1016/j.copsyc.2015.07.018

Moosbrugger, H., & Kelava, A. (2020). Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien"). In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 13-38). 3. Auflage. Springer Verlag.

Moosbrugger, H., Schermelleh-Engel, K., Gäde, J. C., & Kelava, A. (2020). Testtheorien im Überblick. In Moosbrugger, H., & Kelava, A. (Hrsg.). Testtheorie und Fragebogenkonstruktion (pp. 251-273). 3. Auflage. Springer Verlag.

Morgan, S. L., Winship, C. (2015). *Counterfactuals and causal inference: Methods and principles for social research*. New York, NY: Cambridge University Press. Second Edition.

Morf, C. C., & Koole, L. S. (2014). Das selbst. In K. Jonas, W. Stroebe, M. Hewstone, (Hrsg.), *Sozialpsychologie* (pp. 141-195), Springer-Lehrbuch, Springer-Verlag. doi: 10.1007/978-3-642-41091-8_5

Moses, J. W., & Knutsen, T. L. (2012). *Ways of Knowing. Competing Methodologies in Social and Political Research.* 2ed Ed. Palgrave MacMillan.

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological review*, *125*(5), 656. http://dx.doi.org/10.1037/rev0000111

Mummendey, H. D., & Bolten, H. - G. (1993). Die Impression-Management-Theorie. In D. Frey, & M. Irle, (Eds.). *Theorien der Sozialpsychologie* (pp. 57-77). (Unveränd. Nachdr. d. 1. Aufl., Vol. 3: Motivations- und Informationsverarbeitungstheorien), Huber.

Muthukrishna, M., Morgan, T. J. H., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, 37(1), 10-20. https://doi.org/10.1016/j.evolhumbehav.2015.05.004

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J. & Thue, B. (2020a). Beyond western, educated, industrial, rich, and democratic (WIRD) psychoPurlogy: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, 31(6):678-701. DOI: 10.1177/0956797620916782

Muthukrishna, M., Henrich, J., & Slingerland, E. (2020b). Psychology as a historical science. *Annu. Rev. Psychol*. 72:717–49. https://doi.org/10.1146/annurev-psych-082820-111436

Muthukrishna, M. (2021). The ties that bind us. *LSE Public Policy Review*. 2021; 2(1), 3. doi: http://doi.org/10.31389/lseppr.35

Murray, D. R., & Schaller, M. (2010). Historical Prevalence of Infectious Diseases Within 230 Geopolitical Regions: A Tool for Investigating Origins of Culture. *Journal of Cross-Cultural Psychology*, 41(1): 99–108. DOI: 10.1177/0022022109349510

Myers, D., G., Abell, J., Kolstad, A., & Sani, F. (2010). *Social psychology*. McGraw Hill.

Myers, D., G. (Ed.). (2012). *Exploring social psychology*. (6-th ed.). McGraw Hill.

Nettle, D., Gibson, M. A., Lawson, D. W., & Sear, R. (2013). Human behavioral ecology: current research and future prospects. *Behavioral Ecology* (2013), 24(5), 1031–1040. doi:10.1093/beheco/ars222.

Nietzsche, F. W. (1887/1991). *Zur Genealogie der Moral. Eine Streitschrift*. insel taschenbuch. Insel Verlag.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological review*, *108*(2), 291. https://doi.org/10.1037/0033-295X.108.2.291

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265-292). Psychology Press.

Oehlert, G. W. (2010). *A First Course in Design and Analysis of Experiments*. WH Freeman, New York, New York, USA.

Ormel, J., Lindenberg, S., Steverink, N., & Verbrugge, L. M. (1999). Subjective well-being and social production functions. *Social indicators research*, *46*, 61-90. DOI:10.1023/A:1006907811502

Ostfeld, R. S., Glass, G. E., & Keesing, F. (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in ecology & evolution*, *20*(6), 328-336. DOI:https://doi.org/10.1016/j.tree.2005.03.009

Oyserman D. (2016) What does a priming perspective reveal about culture: culture as situated cognition. *Current Opinion in Psychology*, 2016, 12, 94-99. https://doi.org/10.1016/j.copsyc.2016.10.002

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, *38*(1), 1-4. https://doi.org/10.1207/S15326985EP3801_1

Park, J., Uchida, Y., & Kitayama, S. (2016). Cultural variation in implicit independence: An extension of Kitayama et al. (2009). *International Journal of Psychology*, 51(4), 269–278, doi: 10.1002/ijop.12157.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49,* 65–85. https://doi.org/10.1146/annurev.psych.49.1.65

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of personality and social psychology*, *90*(5), 751. DOI: 10.1037/0022-3514.90.5.751

Pfattheicher, S., Nielsen, Y. A., & Thielmann, I. (2022). Prosocial behavior and altruism: A review of concepts and definitions. *Current opinion in psychology*, *44*, 124-129. https://doi.org/10.1016/j.copsyc.2021.08.021

Phillips DeZalia, R. A., & Moeschberger, S. L. (2016). Symbolic reminders of identity. In McKeown, S., Haji, R., & Ferguson, N. (eds.). *Understanding peace and conflict through social identity theory: Contemporary global perspectives* (pp. 121-134). Springer.

Pietraszewski D, Cosmides L, Tooby J. (2014). The content of our cooperation, not the color of our skin: An alliance detection system regulates categorization by coalition and race, but not sex. *PLOS ONE*. 9(2):e88534. https://doi.org/10.1371/journal.pone.0088534

Pietraszewski, D. (2016). Evolutionary Psychology. In H. Miller (Ed). *Sage Encyclopedia of Theory in Psychology* (pp. 314-318)*.* Thousand Oaks: Sage.

Pinto, I., Marques, J., Levine, J. M., Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the Black Sheep Effect? *Journal of Personality and Social Psychology*, 99, 107–119. DOI: 10.1037/a0018187

Piper, E. (2001). Preußische Tugenden im Zeitalter der totalitären Herausforderung. *Zeitschrift für Religions-und Geistesgeschichte*, *53*(1), 35-45. http://www.jstor.com/stable/23898044

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Pratto, F., Çidam, A., Stewart, A. L., Zeineddine, F. B., Aranda, M., Aiello, A., Chryssochoou, X., Cichocka, A., Cohrs, J. C., Durrheim, K., Eicher, V., Foels, R., Górska, P., Lee, I-C., Licata, L., Liu, J. H., Li, L., Meyer, I., Morselli, D., . . . Henkel, K. E. (2013). Social dominance in context and in individuals: Contextual moderation of robust effects of social dominance orientation in 15 languages and 20 countries. *Social Psychological and Personality Science, 4*(5), 587–599. https://doi.org/10.1177/1948550612473663

Prüfer, P., & Rexroth, M. (2005). Kognitive Interviews. *(GESIS-How-to, 15). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-*. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-201470

Purzycki, B. G., Henrich, J., Apicella, C., Atkinson, Q. D., Baimel, A., Cohen, E., McNamara, R. A., Willard, K., Xygalatas, D., & Norenzayan, A. (2018). The evolution of religion and morality: a synthesis of ethnographic and experimental evidence from eight societies. Religion, Brain & Behavior, 8(2), 101-132, doi: 10.1080/2153599X.2016.1267027.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental review*, *41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The transsituational influence of social norms. *Journal of personality and social psychology*, *64*(1), 104. https://doi.org/10.1037/0022-3514.64.1.104

Reicher, S., Spears, R., & Haslam, S. A. (2010). The social identity approach in social psychology. In M. S. Wetherell & C. T. Mohanty (Eds.), *Sage identities handbook* (pp. 45–62). Sage.

Reilly, B. (2013). Revisiting consanguineous marriage in the Greater Middle East: Milk, blood, and Bedouins. *American Anthropologist*, *115*(3), 374-387. https://doi.org/10.1111/aman.12023

Reynolds, K. J., Subašić, E., & Tindall, K. (2015). The problem of behaviour change: From social norms to an ingroup focus. *Social and Personality Psychology Compass*, *9*(1), 45-56. https://doi.org/10.1111/spc3.12155

Richerson, P. J., Boyd, R., & Henrich, J. (2010). Gene-culture coevolution in the age of genomics. *PNAS*, 107(2), 8985-8992. doi: 10.1073/pnas.0914631107

Ridgeway, C. L. (2006). Linking social structure and interpersonal behavior: A theoretical perspective on cultural schemas and social relations. *Social Psychology Quarterly*, 69(1), 5-16. https://doi.org/10.1177/019027250606900102

Rippl, S., & Seipel, C. (2022). *Methoden kulturvergleichender Sozialforschung: Eine Einführung*. 3. Auflage. Wiesbaden: Springer Fachmedien Wiesbaden.

Romano, A., Saral, A. S., & Wu, J. (2022). Direct and indirect reciprocity among individuals and groups. *Current Opinion in Psychology*, *43*, 254-259. https://doi.org/10.1016/j.copsyc.2021.08.003

Rosenbaum, P. R. (2005). Observational study. *Encyclopedia of statistics in behavioral science*.

Rosenbaum, P. R. (2010). Design sensitivity and efficiency in observational studies. *Journal of the American Statistical Association*, *105*(490), 692-702. https://doi.org/10.1198/jasa.2010.tm09570

Roos, P., Gelfand, M., Nau, D., & Lun, J. (2015). Societal threat and cultural variation in the strength of social norms: An evolutionary basis. *Organizational Behavior and Human Decision Processes*, *129*, 14-23. https://doi.org/10.1016/j.obhdp.2015.01.003

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing Measurement Invariance for a Second-Order Factor: A Cross-National Test of the Alienation Scale. *Methods, data, analyses: a journal for quantitative methods and survey methodology* (mda), 12(1), 47-76. https://doi.org/10.12758/mda.2017.11

Rusch, H. (2014). The evolutionary interplay of intergroup conflict and altruism in humans: a review of parochial altruism theory and prospects for its extension. *Proc. R. Soc. B*, 281: 20141539. https://doi.org/10.1098/rspb.2014.1539

Rusch, H. (2022). Heroic behavior: A review of the literature on high-stakes altruism in the wild. *Current Opinion in Psychology*, *43*, 238-243. https://doi.org/10.1016/j.copsyc.2021.07.024

Sachdeva, S., Singh, P., & Medin, D. (2011). Culture and the quest for universal principles in moral reasoning. *International journal of psychology*, *46*(3), 161-176. https://doi.org/10.1080/00207594.2011.568486

San Martin, A., Sinaceur, M., Madi, A., Tompson, S., Maddux, WW., & Kitayama, S. (2018). Self-assertive interdependence in Arab culture. *Nature Human Behaviour*, 2(11), 830–837. 10.1038/s41562-018-0435-z

Santos, H. C., Varnum, M. E., & Grossmann, I. (2017). Global increases in individualism. *Psychological science*, *28*(9), 1228-1239. DOI: 10.1177/0956797617700622

Sánchez-Villagra, M. R., & van Schaik, C. P. (2019). Evaluating the self-domestication hypothesis of human evolution. *Evolutionary Anthropology*. 2019, 28:133–143. DOI: 10.1002/evan.21777

Savalei, V., & Falk, C. F. (2014). Recovering substantive factor loadings in the presence of acquiescence bias: A comparison of three approaches. *Multivariate behavioral research*, *49*(5), 407-424. https://doi.org/10.1080/00273171.2014.931800

Schaller, M., & Neuberg, S. L. (2008). Intergroup prejudices and intergroup conflicts. In C. Crawford & D. Krebs (Eds.), *Foundations of evolutionary psychology* (pp. 401–414). Taylor & Francis Group/Lawrence Erlbaum Associates.

Schaller, M., Conway III, L. G., Peavy, K. M. (2010). Evolutionary processes. In Dovidio, J., F., Hewstone, M., Glick, P., Esses, V., M. (Hrsg.), *The Sage Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 81-96). SAGE Publications Ltd. 2010

Schermelleh-Engel, K., & Gäde, J. C. (2020). Modelbasierte methoden der reliabilitätsschätzung. In Moosbrugger, H., & Kelava, A. (Hrsg.). *Testtheorie und Fragebogenkonstruktion* (pp. 353- 368). 3. Auflage. Springer Verlag.

Schreiber, D., Iacoboni, M. (2012). Huxtables on the brain: an fMRI study of race and norm violation. *International Society of Political Psychology*. 33(3), 313-330. doi: 10.1111/j.1467-9221.2012.00879.x

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science*, *18*(5), 429-434. DOI: 10.1111/j.1467-9280.2007.01917.x

Schulz, J. F., Bahrami-Rad, D., Beauchamp, P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, 366(6466). doi: 10.1126/science.aau5141

Schulze, H. (2008). *Kleine deutsche Geschichte*. dtv. 9. Aufl.

Schwab, S., & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik*, *7*(3), 175-193. ISSN 1869-4934

Seeleib-Kaiser, M. (2014). Wohlfahrtssysteme in Europa und den USA: Annäherung des konservativen deutschen Modells an das amerikanische?. *WSI-Mitteilungen*, *67*(4), 267-276. DOI:10.5771/0342-300X-2014-4-267

Sherman, P. W., & Billing, J. (1999). Darwinian gastronomy: Why we use spices: Spices taste good because they are good for us. *BioScience*, *49*(6), 453-463. https://doi.org/10.2307/1313553

Sherratt, T. N., & Mesterton-Gibbons, M. (2015). The evolution of respect for property. *Journal of evolutionary biology*, *28*(6), 1185-1202. https://doi.org/10.1111/jeb.12648

Shilton, D., Breski, M., Dor, D., & Jablonka, E. (2020). Human social evolution: Self-domestication or self-control? *Frontiers in Psychology*, 20. doi: 10.3389/fpsyg.2020.00134

Singelis, T. M. (1994). The Measurement of Independent and Interdependent Self-construal. *PSPB*, 20(5), 580-591. https://doi.org/10.1177/0146167294205014

Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education*, *43*(5), 541-556. DOI: 10.1080/1743727X.2019.1704243

Skitka, L. J. & Conway, P. (2019). Morality. In E. J. Finkel & R. F. Baumeister (Eds.). *Advanced Social Psychology* (pp. 299-325), 2nd Ed. Oxford, U.K., Oxford University Press. https://lskitka.people.uic.edu/SkitkaConway2019.pdf

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of cross-cultural psychology*, *35*(1), 50-61. https://doi.org/10.1177/002202210326038

Smith, P. B. (2014). Sozialpsychologie und kulturelle Unterschiede. In K. Jonas, W. Stroebe, M. Hewstone, (Hrsg.), *Sozialpsychologie* (pp. 565-605), Springer-Lehrbuch, Springer-Verlag. doi: 10.1007/978-3-642- 41091-8_5

Smith, P. B., Vignoles, V. L., Becker, M., Owe, E., Easterbrook, M. J., Brown, R., ... & Harb, C. (2016). Individual and culture-level components of survey response styles: A multi-level analysis using cultural models of selfhood. *International Journal of Psychology*, *51*(6), 453-463. https://doi.org/10.1002/ijop.12293

SOEP (Ed.) (2019). SOEP-Core - 2018: Personenfragebogen. Stichproben A-L3 + N, SOEP Survey Papers, No. 608, *Deutsches Institut für Wirtschaftsforschung* (DIW), Berlin.

Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental psychology*, *56*(4), 283-294. DOI: 10.1027/1618-3169.56.4.283

StataCorp. 2013. Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.

Stephan, W. G., Ybarra, O., & Morrison, K. R. (2009). Intergroup Threat Theory. In Nelson, T. D. (Eds.) *Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 43-60). New York, NY: Psychology Press. Taylor and Francis Group.

Summerville, J. A., & Enright, E. A. (2018). The origins of infants´ fairness concerns and links to prosocial behavior. *Curr Opin Psychol*. 20, 117-121. https://doi.org/10.1016/j.copsyc.2018.01.005

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annu. Rev. Psychol.*, *58*, 345-372. https://doi.org/10.1146/annurev.psych.56.091103.070145

Talhelm, T. (2022). The rice theory of culture. *Online Readings in Psychology and Culture*, *4*(1), 7.

Tajfel, H. & Turner, J.C. (1986) The Social Identity Theory of Intergroup Behavior. In Worchel, S. & Austin, W.G., (Eds.), *Psychology of Intergroup Relation* (pp. 7-24), Hall Publishers, Chicago.

Teltemann, J., & Windzio, M. (2016). Soziologische Migrations-und Integrationsforschung. In D. B. Maehler, & H. U. Brinkmann (Eds.), *Methoden der Migrationsforschung: Ein interdisziplinärer Forschungsleitfaden* (pp. 163-190). Springer SV.

Tessman, L. (2014). *Moral failure: On the impossible demands of morality*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199396146.001.0001

Thomas, T. D., Abts, K., & Vander Weyden, P. (2014). Measurement invariance, response styles, and rural–urban measurement comparability. *Journal of Cross-Cultural Psychology*, *45*(7), 1011-1027. DOI: 10.1177/0022022114532359

Thomson R., Yuki, M., Talhelm, T., *et al.,* (2018). Relational mobility predicts social behaviors in 39 countries and is tied to historical farming and threat. *PNAS*, 115(29), 7521-7526. https://doi.org/10.1073/pnas.1713191115

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual review of psychology*, *64*, 231-255. https://doi.org/10.1146/annurev-psych-113011-143812

Tomasello, M. (2017). *Warum wir kooperieren*. Edition unseld SV. 4. Ed. Surkamp Verlag.

Tooby, J., Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. In Henrik Høgh-Olesen (Ed.), *Human morality and sociality: Evolutionary and comparative perspectives* (pp. 191–234). New York, NY: Palgrave Macmillan.

Traphagan, W., & Thompson, C. S. (2006). The practice of tradition and modernity in contemporary Japan. In C. S. Thompson & J. W. Traphagan (Eds.) *Wearing cultural styles in Japan: Concepts of tradition and modernity in practice* (pp. 2-24). State University of New York Press.

Triandis, H. C., McCusker, C., & Hui, C. H. (1990). Multimethod probes of individualism and collectivism. *Journal of Personality and Social Psychology, 59*(5), 1006–1020. https://doi.org/10.1037/0022-3514.59.5.1006

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology, 74*(1), 118–128. https://doi.org/10.1037/0022-3514.74.1.118

Turner, J. C., & Reynolds, K. J. (2012) Self-categorization theory. In P. A. M. Van Lange, A. W. Kruglanski, E: T. Higgins (red.). *Handbook of theories of social psychology* (399-417). DOI:10.4135/9781446249222.n46

Tutić, A. (2023). Thoughts on integrating evolutionary analysis into sociological action theory. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 1-26. https://doi.org/10.1007/s11577-023-00907-1

UNESCO Institute for Statistics. (2012). International standard classification of education: ISCED 2011. *Comparative Social Research*, *30*. http://dx.doi.org/10.15220/978-92-9189-123-8-en

Uskul, A. K., Cross, S. E., Günsoy, C., & Gul, P. (2019). Cultures of honor. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 793–821), 2nd ed., Guilford Press.

Uskul, A. K., Kirchner-Häusler, A., Vignoles, V. L., Rodríguez-Bailón, R., Castillo, V. A., Cross, S. E., Gezici Yalçın, M., Harb, C., Husnu, S., Ishii, K., Jin, S., Karamouna, P., Kafetsios, K., Kateri, E., Matamaros-Lima, J., Liu, D., Miniesy, R., Na, J., Özkan, Z., ... Uchida, Y. (2023). Neither Eastern nor Western: Patterns of independence and interdependence in Mediterranean societies. (in press) *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspa0000342

Utz, S. (2004). Self-construal and cooperation: Is the interdependent self more cooperative than the independent self? *Self and Identity, 3*(3), 177–190. https://doi.org/10.1080/13576500444000001

Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, *6*, 167-172. https://doi.org/10.1016/j.copsyc.2015.08.009

van de Vijver, F. J. R., & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F. J. R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). Cambridge University Press.

Van Lange, P. A. M. (2015). Generalized trust: Four lessons from genetics and culture. *Current Directions in Psychological Science,* 24, 71–76. http://dx.doi.org/10.1177/0963721414552473

Verkuyten, M. (2007). Social psychology and multiculturalism. *Social and Personality Psychology Compass*, *1*(1), 280-297. https://doi.org/10.1111/j.1751-9004.2007.00011.x

Vignoles, V. L., Owe, E., Becker, M., Smith, P. B., Easterbrook, M. J., Brown, R., . . . Bond, M. H. (2016). Beyond the 'east- west' dichotomy: Global variations in cultural models of self-hood. *Journal of Experimental Psychology: General,* 145, 966–1000. doi:10.1037/xge0000175

Völkle, M. C., & Erdfelder, E. (2010). Varianz- und Kovarianzanalyse. In C. Wolf, & H. Best (Hg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (pp. 455-494). 1. Auflage, VS Verlag.

Wasmer, M., & Baumann, H. (2019). German General Social Survey 2018: English Translation of the German" ALLBUS"-Questionnaire.

Watson, D. (1992). Correcting for acquiescent response bias in the absence of a balanced scale: An application to class consciousness. *Sociological Methods & Research*, *21*(1), 52-88. DOI:10.1177/0049124192021001003

Waytz, A., Iyer, R., Young, L., Haidt, J., & Graham, J. (2019). Ideological differences in the expanse of the moral circle. *Nature Communications*, 10(1), 1– 12. https://doi.org/10.1038/s41467-019-12227-0

Weber, M. (2013). *Wirtschaft und Gesellschaft. Max Weber-Gesamtausgabe. Band I/23: Wirtschaft und Gesellschaft. Soziologie. Unvollendet. 1919–1920* (K. Borchardt, E. Hanke, & W. Schluchter, Hrsg.). J. C. B. Mohr (Paul Siebeck).

Welkenhuysen-Gybels, J., Billiet, J., & Cambré, B. (2003). Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items. *Journal of Cross-Cultural Psychology*, *34*(6), 702-722. DOI: 10.1177/0022022103257070

Whiten, A., Ayala, F. J., Feldman, M. W., & Laland, K. N. (2017). The extension of biology through culture. *PNAS*, 114(30), 7775-7781. https://doi.org/10.1073/pnas.1707630114

Windzio, M., & Wingens, M. (2014). Soziale Beziehungen in multiethnischen Gesellschaften. Die theoretischen Grundlagen von Integration, Assimilation, Multikulturalismus und ethnischen Grenzziehungen. In E. Bicer, M.

Windzio, & M. Wingens (Eds.), *Soziale Netzwerke, Sozialkapital und ethnische Grenzziehungen im Schulkontext* (pp. 13-47). Springer SV.

Windzio, M. (2016). Children's and adolescents' peer networks and migrant integration. In Punch, S., & Vanderbeck, R. (Eds.) *Families, Intergenerationality, and Peer Group Relations* (pp. 1–20). Singapore: Springer

Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *The Stata Journal*, *12*(2), 308-331. https://doi.org/10.1177/1536867X1201200209

Wong, Y., & Tsai, J. (2007). Cultural models of shame and guilt. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 209–223). The Guilford Press

Wooldridge, J. M. (2016). *Introductory Econometrics: A Modern Approach 6th ed*. Cengage Learning.

Wrangham, R. W. (2018). Two Types of Aggression in human evolution. *PNAS*, 2018, 115(2), 245-253. https://doi.org/10.1073/pnas.1713611115

Wrangham, R. W. (2021). Targeted conspiratorial killing, human self-domestication and the evolution of groupishness. *Evolutionary human sciences*, *3*, e26. doi:10.1017/ehs.2021.20

Yang, S. (2018). *Multidimensional self-construals: Testing the model and refining measurement* (Doctoral dissertation, University of Sussex).

Żemojtel-Piotrowska, M., & Piotrowski, J. (2023). Hofstede's cultural dimensions theory. In *Encyclopedia of Sexual Psychology and Behavior* (pp. 1-4). Cham: Springer International Publishing. DOI:10.1007/978-3-031-08956-5_1124-1

Zick, A., Küpper, B., & Hövermann, A. (2011). Die Abwertung der Anderen. *Eine europäische Zustandsbeschreibung zu Intoleranz, Vorurteilen und Diskriminierung. Berlin: Friedrich-Ebert-Stiftung*.

Zuell, C., & Scholz, E. (2019). Construct equivalence of left-right scale placement in a cross-national perspective. *International Journal of Sociology*, *49*(1), 77-95. https://doi.org/10.1080/00207659.2018.1560982

Internet Sources

Allbus. (2006). Fragebogendokumentation, ALLBUS Vollversion, Studiennummer ZA4500, GESIS, Leibniz Institute for the social science. (Retrieved at 18.09.2017): https://www.gesis.org/allbus/inhalte-suche/frageboegen/. DOI=10.4232/1.10832

Altemeyer, B. (2006). *The authoritarians*. (Retrieved 2018): https://theauthoritarians.org/

Average age across countries (Retrieved at 20.03.2024):

- https://de.statista.com/statistik/daten/studie/200664/umfrage/durchschnittsalter-der-bevoelkerung-in-den-usa/
- https://de.statista.com/statistik/daten/studie/200666/umfrage/durchschnittsalter-der-bevoelkerung-in-japan/
- https://de.statista.com/statistik/daten/studie/1084430/umfrage/durchschnittsalter-der-bevoelkerung-in-deutschland/
- https://de.statista.com/statistik/daten/studie/1323636/umfrage/durchschnittsalter-der-bevoelkerung-in-aegypten/

Boehnke, K., Dragolov, G., Arant, R., & Unzicker, K. (2022b). *Gesellschaftlicher Zusammenhalt in Baden-Württemberg 2022*. (Retrieved at 14.04.2023): https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/gesellschaftlicher-zusammenhalt-in-baden-wuerttemberg-2022-langfassung

Cultural distance calculation (Retrieved 2022): http://www.culturaldistance.com/; (Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J. & Thue, B. (2020a). Beyond western, educated, industrial, rich, and democratic (WIRD) psychoPurlogy: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, 31(6):678-701. DOI: 10.1177/0956797620916782)

Definition of the word "function" (retrieved at 07.08.2023): https://www.oxfordlearnersdictionaries.com/definition/english/function_1

Definition of the word "subject" (retrieved at 20.03.2024): https://www.merriam-webster.com/dictionary/subject

Effect size calculation (Retrieved in August 2024: https://www.psychometrica.de/effect_size.html

GDP development and prognosis Egypt (Retrieved in August 2024): https://www.statista.com/statistics/377349/gross-domestic-product-gdp-in-egypt/

German basic law (Retrieved in August 2024): https://www.bundesregierung.de/breg-de/themen/75-jahre-grundgesetz/artikel-1-gg-2267756

Hofstede dimensions country comparison. (Retrieved at 28.06.2022): https://www.hofstede-insights.com/fi/product/compare-countries/

ISSP 2018 "Religion IV" - ZA No. 7570 (Retrieved at 27.03.2023): https://www.gesis.org/en/issp/modules/issp-modules-by-topic/religion/2018

Lindblom, A., Marsh, E., Motala, T. & Munyan, K. (2015) Persecution of the Rohingya Muslims. Is Genocide occurring in Myanmar's Rakhine state? A legal analysis. Allard K. Lowenstein, International Human Rights Clinic, Yale Law School, Prepared for Fortified Rights, October 2015. 1–65. available at http://www.fortifyrights.org/downloads/Yale_Persecution_of_the_Rohingya_October_2015.pdf

Life expectancy across countries (Retrieved at 20.03.2024):

- https://de.statista.com/statistik/daten/studie/749354/umfrage/lebenserwartung-in-aegypten/
- https://de.statista.com/statistik/daten/studie/273406/umfrage/entwicklung-der-lebenserwartung-bei-geburt-in-deutschland-nach-geschlecht/
- https://de.statista.com/statistik/daten/studie/18655/umfrage/lebenserwartung-in-japan/
- https://de.statista.com/statistik/daten/studie/18670/umfrage/lebenserwartung-in-den-usa/

The Inglehart-Welzel Cultural Map (Retrieved at 01.04.2023): https://www.worldvaluessurvey.org/WVSContents.jsp?CMSID=findings&CMSID=findings

The German MFQ-1 items were taken from: https://moralfoundations-org.translate.goog/questionnaires/?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=sc

The German MaC-Q items were taken from: https://tinyurl.com/3hc7p7w2

Religion in Japan (Retrieved at 11.08.2024): https://de.statista.com/statistik/daten/studie/166887/umfrage/religionen-in-japan/

Word Value Survey Questionnaire: https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp.

Appendix

Supplement Chapter 2: Case Selection Insights

Values — The Inglehart-Welzel World Cultural Map 2022

Figure 30: Inglehart-Welzel World Cultural

Source: https://www.worldvaluessurvey.org/WVSContents.jsp?CMSID=findings&CMSID=findings
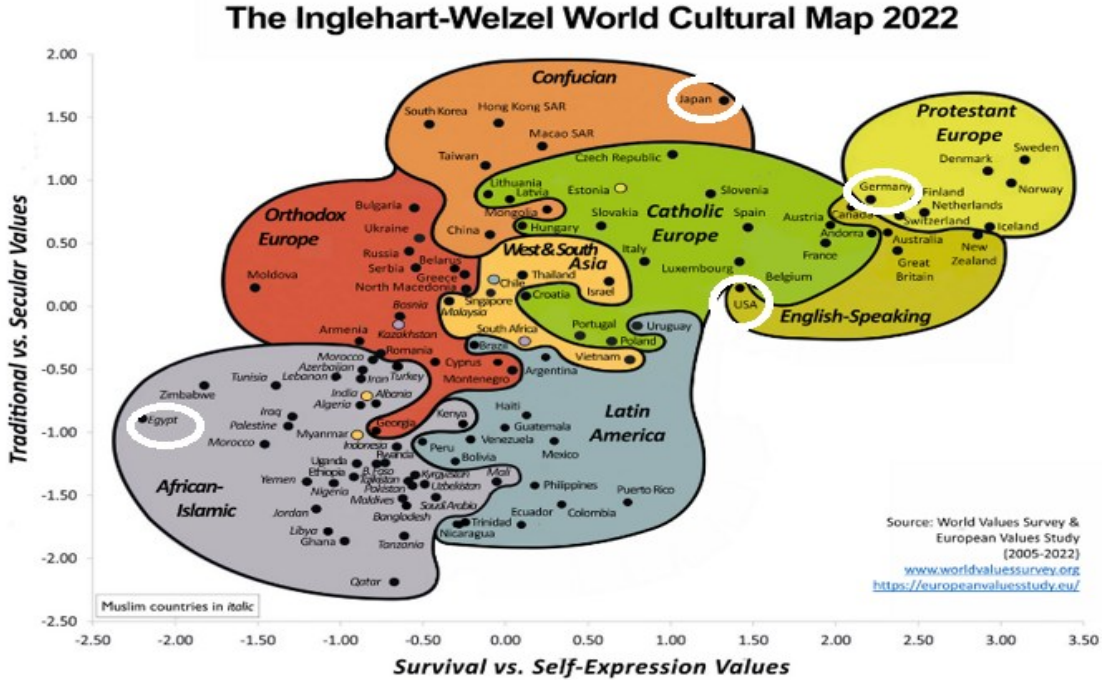
(Mean) Rankings of Values Related to Moral Domains (WVS wave 7)

The following contains rough rankings of importance for several moral domains. [160] The calculations are made with Stata and are based on WVS wave 7 data (Haerpfer et al., 2022). [161]

Family Q1 "Important in life: Family" (The lower the number, the higher the importance)

Table 62: Values across countries I

Egypt	Ø = 1.003333
Nigeria	Ø = 1.015372
Jordan	Ø = 1.024106
Lebanon	Ø = 1.0325
Vietnam	Ø = 1.04
Japan	Ø = 1.082774
Australia	Ø = 1.110372
Germany	Ø = 1.129011
Thailand	Ø = 1.129766
US	Ø = 1.131508
China	Ø = 1.14531
Brazil	Ø = 1.169694
Colombia	Ø = 1.205263

Family Q27: "One of main goals in life has been to make my parents proud" (The lower the number, the higher the importance)

Table 63: Values across countries II

Jordan	Ø = 1.137021
Egypt	Ø = 1.156616
Nigeria	Ø = 1.235628
Lebanon	Ø = 1.295833
Vietnam	Ø = 1.51
Colombia	Ø = 1.556579
Brazil	Ø = 1.630787
Thailand	Ø = 1.681943
US	Ø = 1.931555
Germany	Ø = 1.98364
China	Ø = 2.035325
Australia	Ø = 2.048549
Japan	Ø = 2.237383

---

[160] Q is the WVS abbreviation for question (item).
[161] The WVS questionnaire can be found at: https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp.

Fairness (Equal gender rights/ fairness in job) Q33 "Jobs scarce: Men should have more right to a job than women" (The lower the number, the higher the approval)

Table 64: Values across countries III

Australia	Ø = 4.165464
Germany	Ø = 4.087984
US	Ø = 3.74643
Colombia	Ø = 3.665789
Brazil	Ø = 3.504893
Thailand	Ø = 3.11134
Japan	Ø = 3.006116
China	Ø = 2.993408
Vietnam	Ø = 2.785833
Nigeria	Ø = 2.58449
Lebanon	Ø = 2.4125
Jordan	Ø = 1.746256
Egypt	Ø = 1.498331

Fairness (Equal rights/fairness immigrants & job) Q34 "Jobs scarce: Employers should give priority to (nation) people than immigrants" (The lower the number, the higher the approval)

Table 65: Values across countries IV

Germany	Ø = 3.431698
US	Ø = 2.892236
Australia	Ø = 2.839465
Brazil	Ø = 2.585648
Colombia	Ø = 2.4375
Japan	Ø = 2.336656
Nigeria	Ø = 2.313869
China	Ø = 2.228099
Thailand	Ø = 2.081155
Vietnam	Ø = 1.999167
Lebanon	Ø = 1.473333
Jordan	Ø = 1.345258
Egypt	Ø = 1.247466

Group Q61 "Trust: People you meet for the first time" (The lower the number, the higher the trust)

Table 66: Values across countries V

Australia	Ø = 2.515625
US	Ø = 2.770124
Germany	Ø = 2.805405
Vietnam	Ø = 2.8725
Thailand	Ø = 2.917172
Jordan	Ø = 2.969975
Nigeria	Ø = 3.038866
China	Ø = 3.049305
Egypt	Ø = 3.096477
Lebanon	Ø = 3.111853
Japan	Ø = 3.132541
Brazil	Ø = 3.200694
Colombia	Ø = 3.300658

Group Q62 "Trust: People of another religion" (The lower the number, the higher the trust)

Table 67: Values across countries VI

Australia	Ø = 2.173544
US	Ø = 2.225555
Germany	Ø = 2.39497
Brazil	Ø = 2.529338
Egypt	Ø = 2.535593
Jordan	Ø = 2.576068
Lebanon	Ø = 2.70529
Nigeria	Ø = 2.708367
Vietnam	Ø = 2.731667
Thailand	Ø = 2.745957
Colombia	Ø = 2.9125
China	Ø = 3.016344
Japan	Ø = 3.050847

Group Q63 "Trust: People of another nationality" (The lower the number, the higher the trust)

Table 68: Values across countries VII

Australia	Ø = 2.145915
US	Ø = 2.216543
Germany	Ø = 2.36804
Jordan	Ø = 2.705983
Vietnam	Ø = 2.846667
Thailand	Ø = 2.860999
Lebanon	Ø = 2.892707
Japan	Ø = 2.897436
Brazil	Ø = 2.907007
Nigeria	Ø = 2.933387
China	Ø = 3.072406
Colombia	Ø = 3.078947
Egypt	Ø = 3.149293

Group/Fairness Q112 "Perception of corruption in the country" (The higher the number, the higher the perception of corruption)

Table 69: Values across countries VIII

Germany	Ø = 5.583333
China	Ø = 6.492552
Australia	Ø = 6.646042
Japan	Ø = 6.876923
Thailand	Ø = 6.96601
Vietnam	Ø = 7.369167
US	Ø = 7.829003
Lebanon	Ø = 7.833333
Jordan	Ø = 8.196567
Egypt	Ø = 8.521478
Nigeria	Ø = 8.737055
Brazil	Ø = 9.454754
Colombia	Ø = 9.478289

Property Q177 "Justifiable: Claiming government benefits to which you are not entitled" (The higher the number, the higher the justifiability)

Table 70: Values across countries IX

Germany	Ø = 1.538259
Japan	Ø = 1.763117
Australia	Ø = 1.778953
Egypt	Ø = 1.789151
Nigeria	Ø = 2.003241
Thailand	Ø = 2.216362
Brazil	Ø = 2.43761
US	Ø = 2.530763
Jordan	Ø = 2.994915
China	Ø = 3.296627
Lebanon	Ø = 3.401024
Colombia	Ø = 3.777632
Vietnam	Ø = 4.524167

Property Q179 "Justifiable: Stealing property" (The higher the number, the higher the justifiability)

Table 71: Values across countries X

Egypt	Ø = 1.066109
Japan	Ø = 1.106767
Germany	Ø = 1.162304
China	Ø = 1.287884
Jordan	Ø = 1.375833
Australia	Ø = 1.401338
Brazil	Ø = 1.521086
Colombia	Ø = 1.586184
Nigeria	Ø = 1.679352
Thailand	Ø = 1.833109
US	Ø = 1.892134
Lebanon	Ø = 1.971356
Vietnam	Ø = 2.495

Values: Child Qualities Relevant in Relation to Self-Construal and Moral Domains (WVS wave 7)

Table 72: Child qualities relevant in relation to self-construal and moral domains across countries

C*	Q7**	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17
US	49.69%	55.32%	68.40%	58.87%	33.95%	69.41%	26.77%	40.32%	30.40%	29.63%	19.98%
AUS	82.29%	52.01%	43.79%	56.92%	35.36%	82.74%	21.68%	44.40%	14.45%	44.40%	18.97%
GER	83.80	69.84%	39.74%	79.74%	23.08%	84.13%	36.92%	33.51%	9.51%	5.57%	11.80%
PRC	83.98%	78.06%	71.44%	79.25%	21.94%	60.09%	40.07%	21.58%	1.32%	28.52%	5.59%
JP	84.04%	60.24%	25.06%	75.31%	40.28%	62.60%	43.68%	63.27%	4.43%	33.04%	2.73%
THAI	79.60%	43.20%	69.47%	43.20%	23.73%	52.20%	37.73%	46.27%	20.00%	34.13%	18.27%
VIE	72.42%	41.58%	51.75%	68.60%	26.17%	46.33%	32.08%	45.08%	5.00%	45.08%	55.08%
JO	86.03%	28.35%	36.41%	55.69%	16.38%	67.00%	16.71%	18.87%	77.22%	31.50%	46.88%
LB	76.00%	42.25%	44.42%	56.92%	24.33%	65.00%	34.67%	30.92%	35.33%	21.08%	9.83%
BR	72.93%	26.50%	55.45%	70.54%	15.32%	61.52%	17.99%	23.33%	36.61%	30.08%	43.19%
CO	90.07%	28.82%	24.61%	70.72%	18.22%	77.30%	23.55%	16.45%	46.64%	22.89%	48.42%
NG	89.01%	30.96%	73.48%	40.10%	14.31%	61.76%	13.90%	20.78%	72.03%	16.17%	57.15%
EG	96.25%	14.43%	60.22%	74.23%	5.84%	78.15%	21.85%	10.76%	81.73%	34.78%	55.71%

* C = Country; US refers to United States, AUS refers to Australia, GER refers to Germany, PRC refers to People´s Republic of China, JP refers to Japan, THAI refers to Thailand, VIE refers to Vietnam, JO refers to Jordan, LB refers to Lebanon, BR refers to Brazil, CO refers to Columbia, NG refers to Nigeria, EG refers to Egypt. ****** Item instruction is: "Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five!". Q + number refers to the following: Q7 Good manners; Q8 Independence; Q9 Hard work; Q10 Feeling of responsibility; Q11 Imagination; Q12 Tolerance and respect for other people; Q13 Thrift, saving money and things; Q14 Determination, perseverance; Q15 Religious faith; Q16 Not being selfish (unselfishness); Q17 Obedience.

Cultural Distances in Comparison

Cultural distances (Muthukrishna et al., 2020A) were calculated for a set of countries. Dimensions of interest were: Overall cultural difference; difference in altruism; difference in relationship with out-groups; difference in in-group favoritism; difference in moral beliefs; difference in independence and interdependence. Items upon which the distance metric is based can be inspected under: https://michael.muthukrishna.com/cultural-distance-data/. The cultural distances and rooted trees[162] were calculated on the website provided by Muthukrishna and colleagues (2020a): http://www.culturaldistance.com/.

---

[162] Muthukrishna and colleagues (2020) note in regard to the rooted tree visualization: "This is a visualization of the Saitou et al. (1987) neighbour-joining algorithm (see also Studier and Keppler, 1988) applied to the pairwise matrix of cultural distances between selected countries on the Data Config screen. The neighbour-join algorithm attempts to pair the two most similar countries and then the two most similar pairs of countries and so on" (p. 109 — supplemental material).

Cultural Distance: Overall

Table 73: Cultural Distance: Overall

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		0.11	0.147	0.157	**0.29**	0.053	0.1	**0.255**	0.166	0.223	0.16	0.033
Brazil2005-2014	0.11		0.173	0.037	0.151	0.119	0.16	0.116	0.069	0.087	0.101	0.07
China2005-2014	0.147	**0.173**		**0.195**	0.183	0.126	0.14	0.199	**0.187**	**0.248**	0.118	0.17
Colombia2005-2014	0.157	0.037	**0.195**		0.138	0.164	0.215	0.113	0.116	0.117	0.123	0.112
Egypt2005-2014	**0.29**	**0.151**	0.183	0.138		**0.243**	**0.307**	0.03	**0.207**	0.11	**0.164**	**0.24**
Germany2005-2014	0.053	0.119	0.126	0.164	**0.243**		0.075	**0.271**	0.159	**0.254**	**0.156**	0.079
Japan2005-2014	0.1	0.16	0.14	**0.215**	**0.307**	0.075		**0.269**	**0.203**	**0.272**	**0.165**	0.118
Jordan2005-2014	**0.255**	0.116	**0.199**	0.113	0.03	**0.271**	**0.269**		0.135	0.068	0.124	**0.204**
Lebanon2005-2014	0.166	0.069	0.187	0.116	0.207	0.159	0.203	0.135		0.082	0.098	0.107
Nigeria2005-2014	**0.223**	0.087	**0.248**	0.117	0.11	**0.254**	**0.272**	0.068	0.082		0.119	0.154
Thailand2005-2014	0.16	0.101	0.118	0.123	0.164	0.156	0.165	0.124	0.098	0.119		0.139
United States2005-2014	0.033	0.07	0.17	0.112	0.24	0.079	0.118	0.204	0.107	0.154	0.139	
Viet Nam2005-2014	0.163	**0.163**	0.06	**0.196**	0.228	0.17	0.17	0.152	0.116	0.134	0.098	**0.174**

Rooted Tree

Japan2005-2014
Germany2005-2014
Australia2005-2014
United States2005-2014
Viet Nam2005-2014
China2005-2014
Thailand2005-2014
Lebanon2005-2014
Jordan2005-2014
Egypt2005-2014
Nigeria2005-2014
Colombia2005-2014
Brazil2005-2014
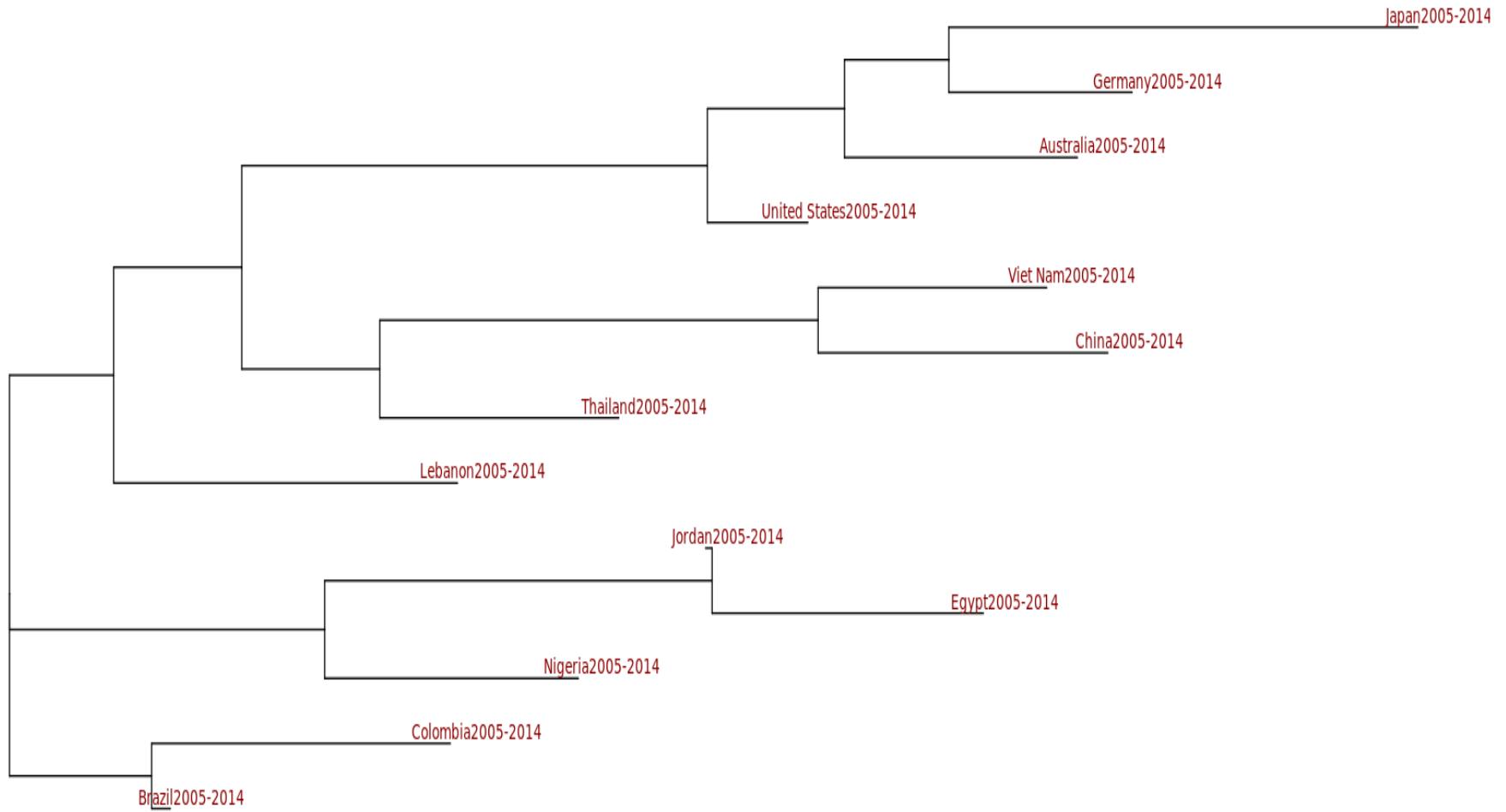
Cultural Distance: Altruism

Table 74: Cultural Distance: Altruism

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		0.104	0.135	0.101	**0.316**	0.095	0.077	**0.225**	0.173	0.162	0.144	0.03
Brazil2005-2014	0.104		0.156	0.04	0.148	0.101	0.129	0.118	0.033	0.051	0.095	0.06
China2005-2014	0.135	**0.156**		0.15	0.238	0.167	0.111	0.197	**0.188**	**0.274**	0.047	0.144
Colombia2005-2014	0.101	0.04	0.15		0.127	0.16	0.157	0.101	0.128	0.097	0.087	0.072
Egypt2005-2014	**0.316**	**0.148**	**0.238**	0.127		**0.242**	**0.3**	0.026	**0.302**	0.156	**0.149**	**0.245**
Germany2005-2014	0.095	0.101	0.167	0.16	0.242		0.093	**0.253**	0.137	0.162	**0.183**	0.074
Japan2005-2014	0.077	0.129	0.111	**0.157**	0.3	0.093		**0.229**	0.166	**0.197**	**0.146**	0.079
Jordan2005-2014	**0.225**	0.118	**0.197**	0.101	0.026	**0.253**	**0.229**		**0.212**	0.136	0.084	**0.179**
Lebanon2005-2014	0.173	0.033	0.188	**0.128**	**0.302**	0.137	0.166	0.212		0.079	0.093	0.085
Nigeria2005-2014	0.162	0.051	**0.274**	0.097	0.156	0.162	0.197	0.136	0.079		0.139	0.09
Thailand2005-2014	0.144	0.095	0.047	0.087	0.149	0.183	0.146	0.084	0.093	0.139		0.124
United States2005-2014	0.03	0.06	0.144	0.072	0.245	0.074	0.079	0.179	0.085	0.09	0.124	
Viet Nam2005-2014	**0.216**	**0.209**	0.042	**0.226**	**0.279**	**0.268**	**0.196**	0.177	0.153	**0.184**	0.045	**0.231**

Rooted Tree
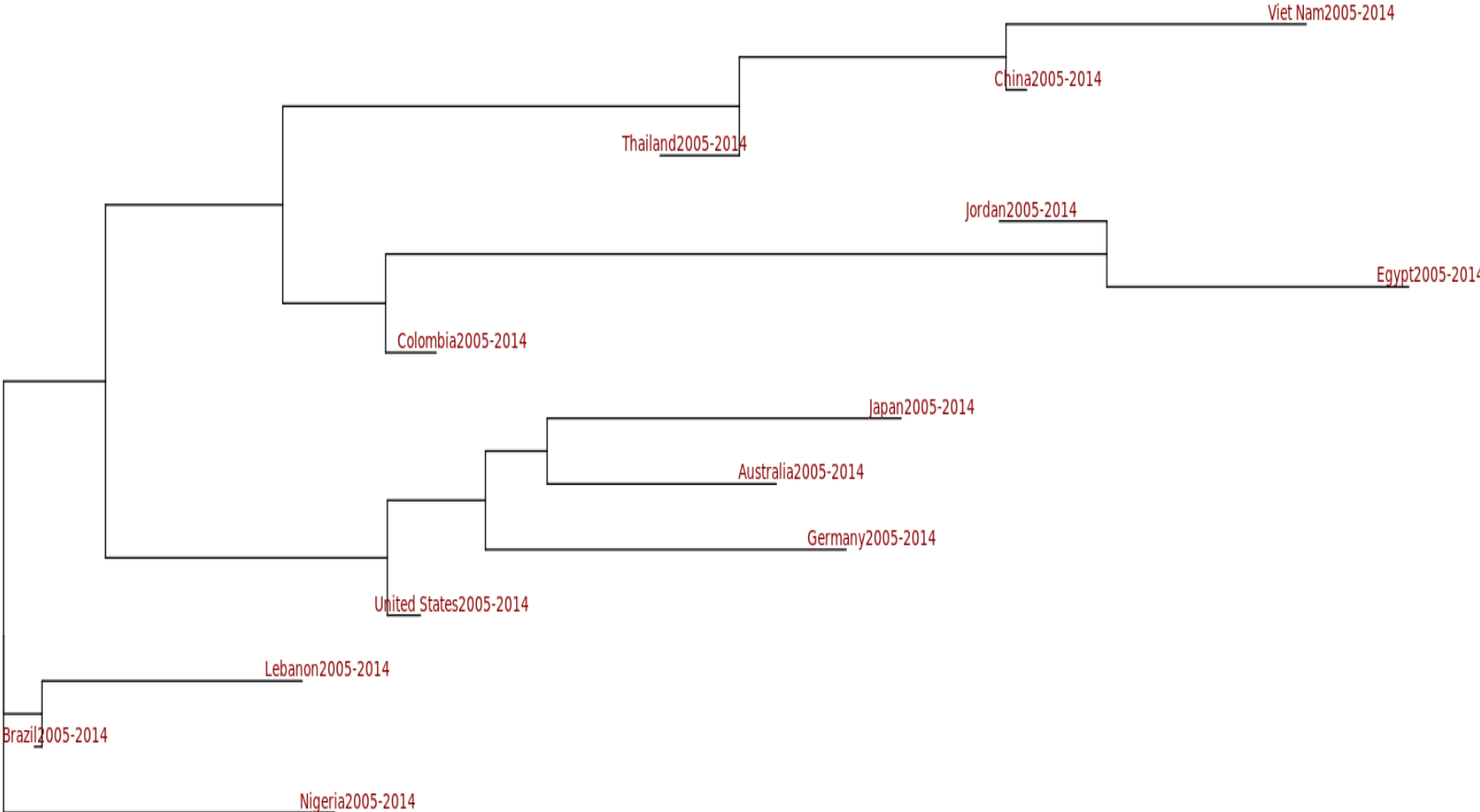
Cultural Distance: Relationship With Out-Groups

Table 75: Cultural Distance: Relationship With Out-Groups

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		**0.084**	0.077	**0.137**	**0.154**	0.033	0.061	**0.156**	**0.614**	**0.199**	**0.15**	0.034
Brazil2005-2014	0.084		0.018	0.062	0.027	0.041	0.042	0.047	0.075	0.023	0.112	0.048
China2005-2014	0.077	0.018		0.13	0.076	0.042	0.05	0.102	0.026	0.001	0.123	0.054
Colombia2005-2014	0.137	0.062	**0.13**		0.059	**0.126**	**0.1**	0.059	**0.543**	0.15	0.132	0.103
Egypt2005-2014	0.154	0.027	0.076	0.059		0.083	0.088	0.023	0.141	**0.12**	0.129	0.082
Germany2005-2014	0.033	0.041	0.042	0.126	0.083		0.02	0.104	0.189	0.052	0.092	0.008
Japan2005-2014	0.061	0.042	0.05	0.1	0.088	0.02		0.086	0.114	0.021	0.111	0.029
Jordan2005-2014	0.156	0.047	**0.102**	0.059	0.023	**0.104**	0.086		**0.205**	**0.15**	**0.196**	**0.111**
Lebanon2005-2014	**0.614**	**0.075**	0.026	0.543	**0.141**	**0.189**	**0.114**	**0.205**		0.024	0.046	**0.221**
Nigeria2005-2014	**0.199**	0.023	0.001	**0.15**	0.12	0.052	0.021	0.15	0.024		0.005	0.058
Thailand2005-2014	0.15	**0.112**	**0.123**	0.132	**0.129**	0.092	**0.111**	**0.196**	0.046	0.005		0.062
United States2005-2014	0.034	0.048	0.054	0.103	0.082	0.008	0.029	0.111	0.221	0.058	0.062	
Viet Nam2005-2014	**0.185**	0.054	0.042	**0.218**	0.113	0.097	0.093	0.117	0.002	0.008	**0.189**	**0.117**

Rooted Tree

Viet Nam2005-2014

Lebanon2005-20

Nigeria2005-2014

Thailand2005-2014

China2005-2014

Jordan2005-2014

Egypt2005-2014

Colombia2005-2014

Brazil2005-2014

Japan2005-2014

Germany2005-2014

Australia2005-2014

United States2005-2014
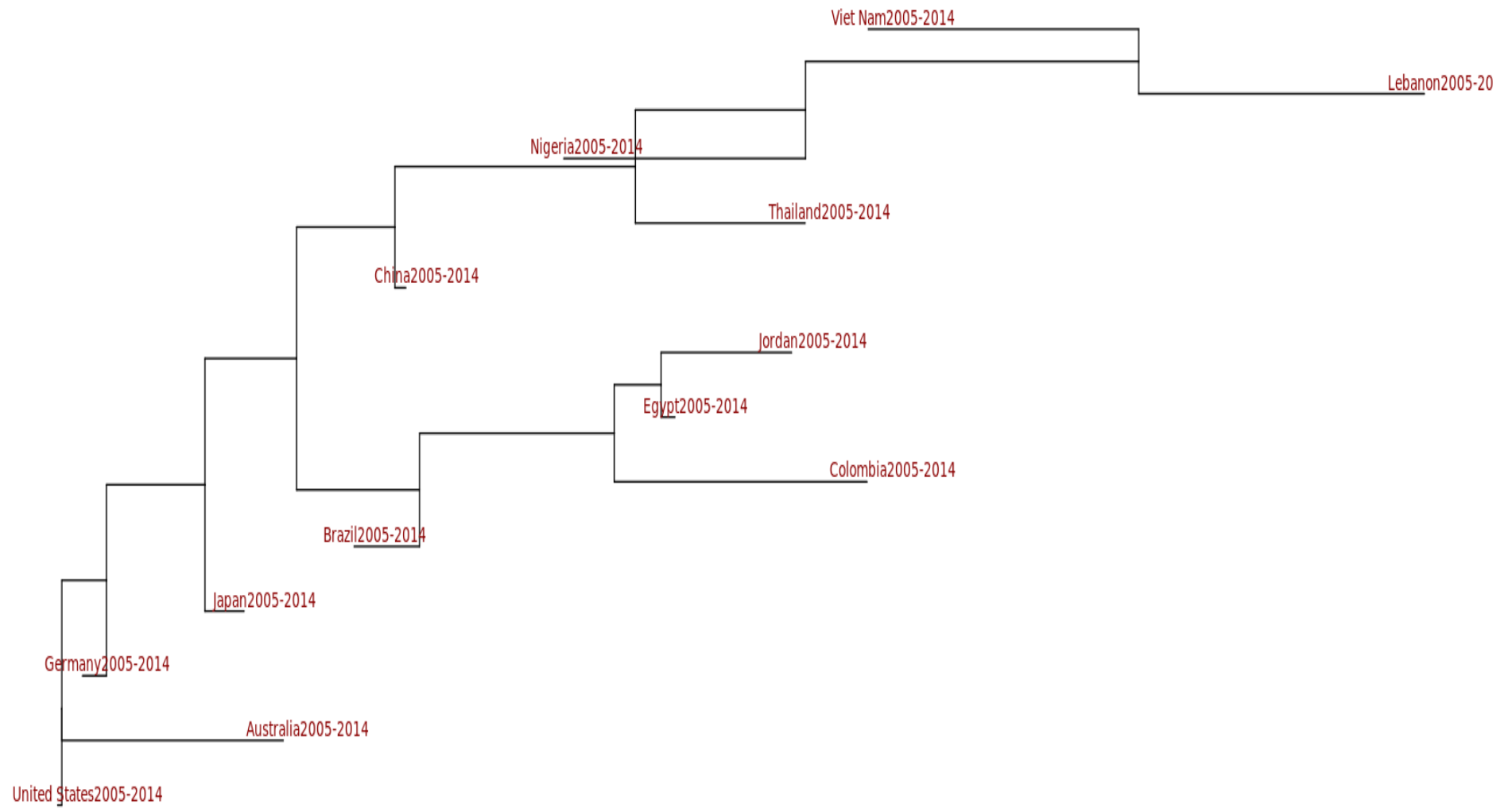
Cultural Distance: Discrimination — In-Group Favoritism

Table 76: Cultural Distance: Discrimination — In-Group Favoritism

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		0.073	**0.109**	0.071	**0.488**	0.041	0.091	**0.41**	**0.242**	**0.292**	0.084	0.006
Brazil2005-2014	0.073		0.043	0.039	0.336	0.051	0.051	0.277	0.108	0.202	0.077	0.063
China2005-2014	0.109	0.043		0.08	0.176	0.094	0.04	0.153	0.01	0.062	0.027	0.109
Colombia2005-2014	0.071	0.039	0.08		0.363	0.101	0.137	0.3	0.224	0.231	0.071	0.076
Egypt2005-2014	**0.488**	**0.336**	**0.176**	**0.363**		**0.456**	**0.348**	0.01	**0.358**	0.112	**0.216**	**0.475**
Germany2005-2014	0.041	0.051	0.094	0.101	**0.456**		0.058	**0.43**	0.183	**0.292**	**0.117**	0.034
Japan2005-2014	0.091	0.051	0.04	0.137	0.348	0.058		0.276	0.066	0.159	0.066	0.077
Jordan2005-2014	**0.41**	**0.277**	**0.153**	**0.3**	0.01	**0.43**	**0.276**		0.222	0.095	**0.166**	**0.428**
Lebanon2005-2014	0.242	0.108	0.01	0.224	0.358	0.183	0.066	0.222		0.05	0.079	0.259
Nigeria2005-2014	**0.292**	**0.202**	0.062	**0.231**	0.112	**0.292**	**0.159**	0.095	0.05		0.074	**0.324**
Thailand2005-2014	0.084	0.077	0.027	0.071	0.216	0.117	0.066	0.166	0.079	0.074		0.094
United States2005-2014	0.006	0.063	**0.109**	0.076	**0.475**	0.034	0.077	**0.428**	**0.259**	**0.324**	0.094	
Viet Nam2005-2014	0.132	0.074	0.026	0.062	0.211	0.17	0.093	0.131	0.059	0.061	0.014	0.151

Rooted Tree

United States2005-2014

Australia2005-2014

Germany2005-2014

Colombia2005-2014

Brazil2005-2014

Japan2005-2014

Jordan2005-2014

Egypt2005-2014

Nigeria2005-2014

Viet Nam2005-2014

Thailand2005-2014

China2005-2014

Lebanon2005-2014

430
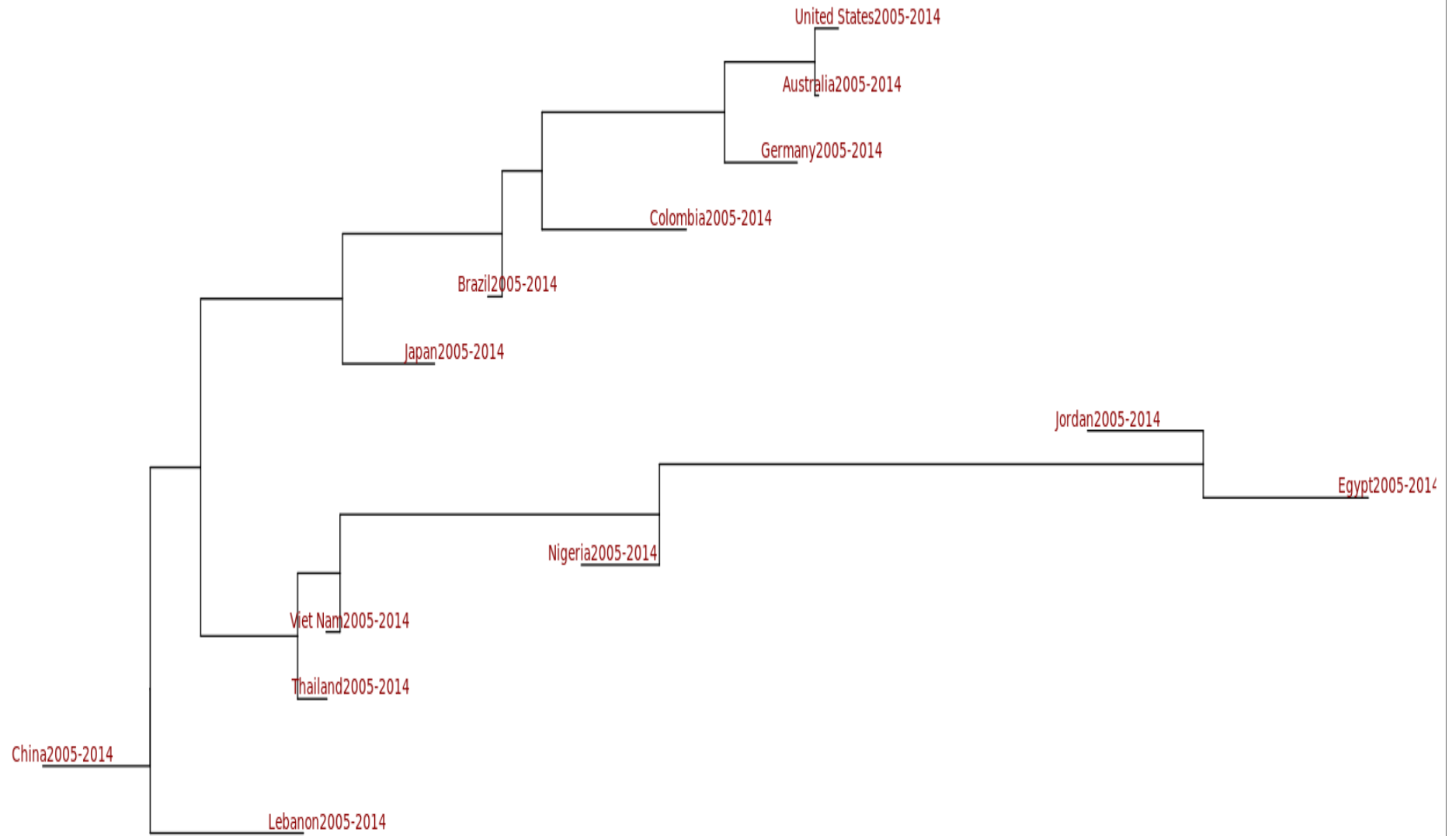
Cultural Distance: Beliefs — Morality

Table 77: Cultural Distance: Beliefs — Morality

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		**0.131**	**0.261**	**0.201**	**0.247**	0.008	0.046	**0.334**	**0.194**	**0.286**	**0.257**	0.036
Brazil2005-2014	0.131		0.099	0.034	0.067	0.102	0.163	0.116	0.06	0.122	0.055	0.071
China2005-2014	**0.261**	0.099		0.046	0.057	0.189	0.18	0.049	0.061	0.019	0.019	0.191
Colombia2005-2014	0.201	0.034	0.046		0.036	0.133	0.193	0.079	0.124	0.054	0.03	0.123
Egypt2005-2014	0.247	0.067	0.057	0.036		0.149	**0.324**	0.013	**0.194**	0.056	0.075	**0.195**
Germany2005-2014	0.008	0.102	**0.189**	**0.133**	0.149		0.05	**0.282**	0.149	**0.244**	**0.212**	0.031
Japan2005-2014	0.046	**0.163**	0.18	**0.193**	**0.324**	0.05		**0.283**	**0.217**	0.188	**0.27**	0.073
Jordan2005-2014	**0.334**	0.116	0.049	0.079	0.013	**0.282**	**0.283**		0.176	0.031	0.057	**0.24**
Lebanon2005-2014	0.194	0.06	0.061	0.124	0.194	0.149	0.217	0.176		0.1	0.038	0.116
Nigeria2005-2014	**0.286**	**0.122**	0.019	0.054	0.056	**0.244**	0.188	0.031	0.1		0.038	**0.211**
Thailand2005-2014	0.257	0.055	0.019	0.03	0.075	**0.212**	**0.27**	0.057	0.038	0.038		0.169
United States2005-2014	0.036	0.071	**0.191**	0.123	**0.195**	0.031	0.073	0.24	0.116	**0.211**	0.169	
Viet Nam2005-2014	0.155	0.076	0.007	0.05	0.091	0.124	0.129	0.1	0.028	0.022	0.087	0.106

Rooted Tree

Japan2005-2014

Australia2005-20

Germany2005-2014

United States2005-2014

Brazil2005-2014

Lebanon2005-2014

Viet Nam2005-2014

China2005-2014

Jordan2005-2014

Egypt2005-2014

Nigeria2005-2014

Colombia2005-2014

Thailand2005-2014

432
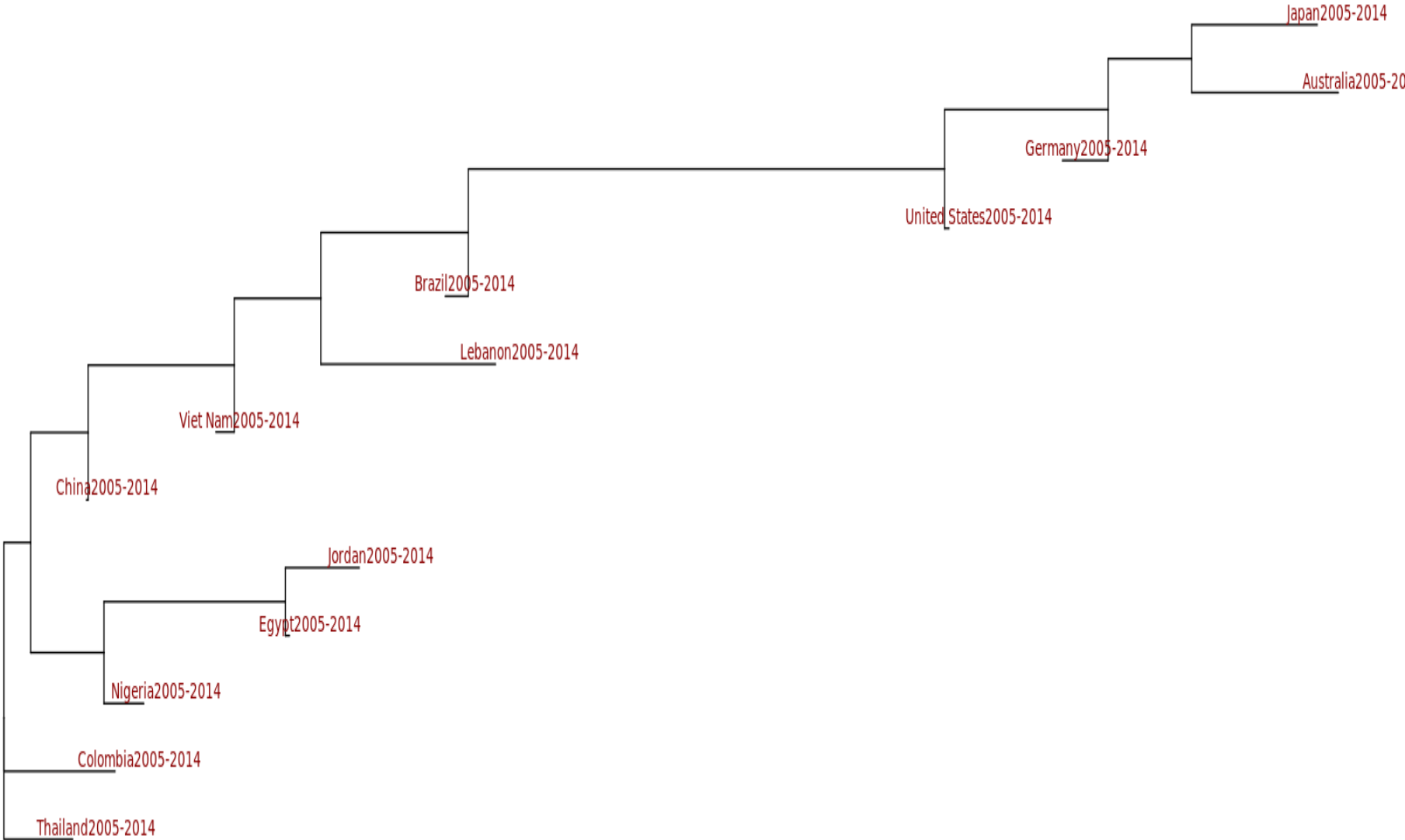
Cultural Distance: Social Relations — Independence and Autonomy

Table 78: Cultural Distance: Social Relations — Independence and Autonomy

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		0.053	0.031	0.068	0.072	0.021	0.036	0.043	0.066	0.041	0.038	0.018
Brazil2005-2014	0.053		0.054	0.008	0.038	0.107	0.118	0.024	0.013	0.024	0.032	0.031
China2005-2014	0.031	0.054		0.054	0.035	0.054	0.084	0.05	0.076	0.064	**0.07**	0.023
Colombia2005-2014	**0.068**	0.008	0.054		0.033	**0.122**	**0.149**	0.032	0.021	0.04	0.054	0.034
Egypt2005-2014	**0.072**	0.038	0.035	0.033		0.106	**0.157**	0.05	0.049	**0.072**	**0.1**	0.025
Germany2005-2014	0.021	**0.107**	0.054	**0.122**	0.106		0.034	**0.115**	**0.131**	**0.112**	**0.077**	0.041
Japan2005-2014	0.036	**0.118**	**0.084**	**0.149**	**0.157**	0.034		**0.104**	**0.151**	**0.085**	0.051	**0.086**
Jordan2005-2014	0.043	0.024	0.05	0.032	0.05	**0.115**	0.104		0.04	0.014	0.048	0.043
Lebanon2005-2014	**0.066**	0.013	**0.076**	0.021	0.049	**0.131**	**0.151**	0.04		0.025	0.041	0.034
Nigeria2005-2014	0.041	0.024	0.064	0.04	0.072	0.112	0.085	0.014	0.025		0.017	0.044
Thailand2005-2014	0.038	0.032	0.07	0.054	0.1	0.077	0.051	0.048	0.041	0.017		**0.048**
United States2005-2014	0.018	0.031	0.023	0.034	0.025	0.041	0.086	0.043	0.034	0.044	0.048	
Viet Nam2005-2014	0.044	**0.099**	**0.118**	**0.147**	**0.212**	0.073	0.02	**0.087**	**0.094**	0.051	0.023	**0.096**

Rooted Tree

Viet Nam2005-2014

Japan2005-2014

Germany2005-2014

Australia2005-2014

Thailand2005-2014

Nigeria2005-2014

Jordan2005-2014

Egypt2005-2014

China2005-2014

United States2005-2014

Lebanon2005-2014

Brazil2005-2014

Colombia2005-2014

434
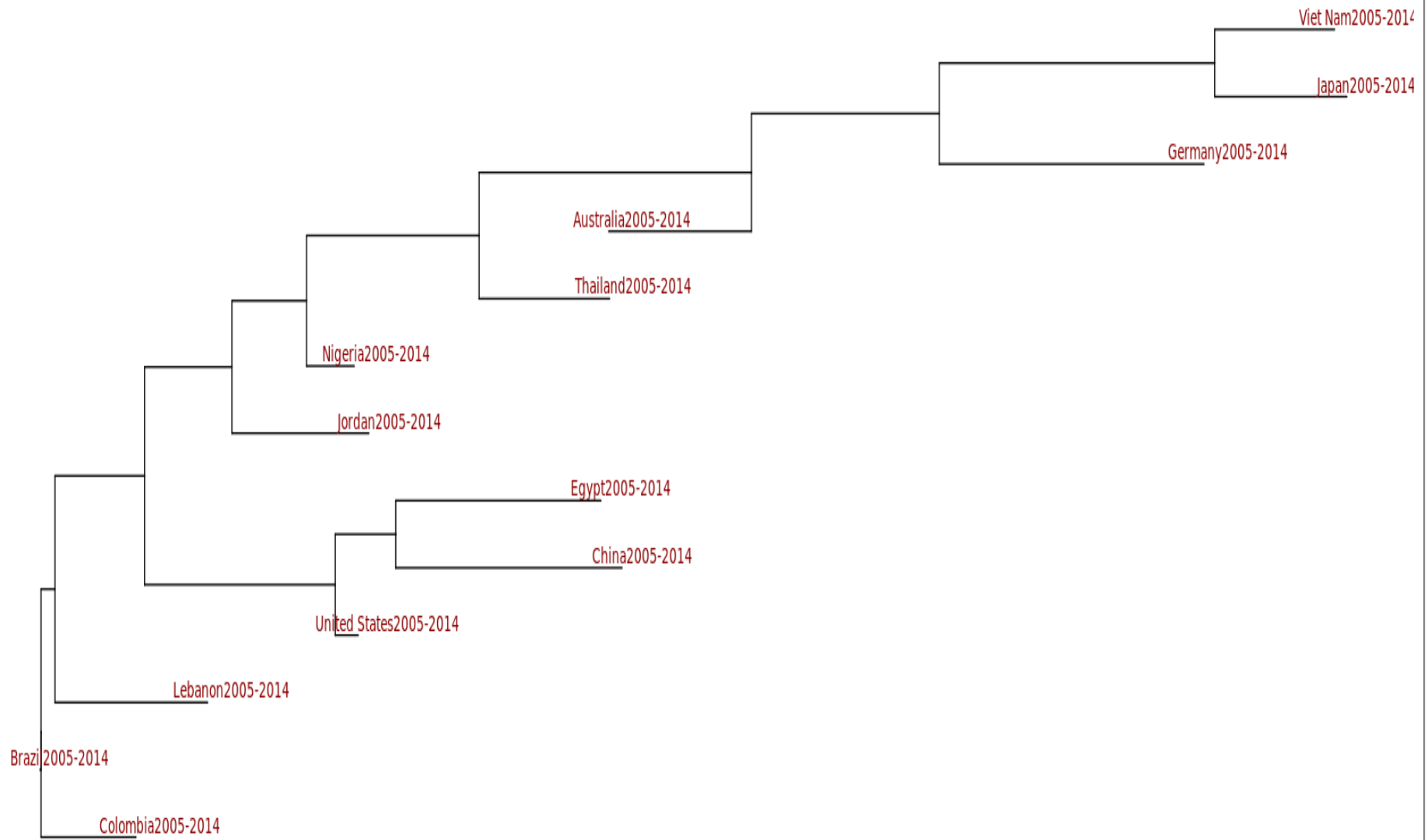
Cultural Distance: Social Relations — Interdependence and Collectivism

Table 79: Cultural Distance: Social Relations — Interdependence and Collectivism

	Australia2005-2014	Brazil2005-2014	China2005-2014	Colombia2005-2014	Egypt2005-2014	Germany2005-2014	Japan2005-2014	Jordan2005-2014	Lebanon2005-2014	Nigeria2005-2014	Thailand2005-2014	United States2005-2014
Australia2005-2014		0.078	0.093	0.088	**0.19**	0.051	0.073	**0.167**	0.106	0.108	**0.13**	0.019
Brazil2005-2014	0.078		0.088	0.032	0.1	0.082	0.082	0.096	0.055	0.06	0.105	0.075
China2005-2014	0.093	0.088		**0.112**	0.1	0.092	0.07	0.098	0.083	**0.161**	0.094	0.085
Colombia2005-2014	0.088	0.032	**0.112**		0.08	0.122	0.111	0.067	0.104	0.09	0.062	0.089
Egypt2005-2014	**0.19**	**0.1**	**0.1**	0.08		**0.179**	**0.161**	0.029	**0.195**	0.126	0.09	**0.159**
Germany2005-2014	0.051	0.082	0.092	**0.122**	**0.179**		0.062	**0.202**	0.094	**0.146**	**0.148**	0.039
Japan2005-2014	0.073	0.082	0.07	0.111	0.161	0.062		0.156	0.094	**0.158**	**0.143**	0.072
Jordan2005-2014	**0.167**	0.096	0.098	0.067	0.029	**0.202**	**0.156**		**0.142**	0.093	0.097	**0.163**
Lebanon2005-2014	0.106	0.055	0.083	0.104	**0.195**	0.094	0.094	0.142		0.08	0.099	0.076
Nigeria2005-2014	0.108	0.06	**0.161**	0.09	0.126	0.146	**0.158**	0.093	0.08		0.078	0.096
Thailand2005-2014	0.13	**0.105**	0.094	0.062	0.09	0.148	0.143	0.097	0.099	0.078		0.105
United States2005-2014	0.019	0.075	0.085	0.089	0.159	0.039	0.072	**0.163**	0.076	0.096	0.105	
Viet Nam2005-2014	**0.148**	**0.119**	0.069	**0.135**	0.087	**0.167**	0.127	0.063	**0.122**	0.101	0.056	**0.143**

Rooted Tree
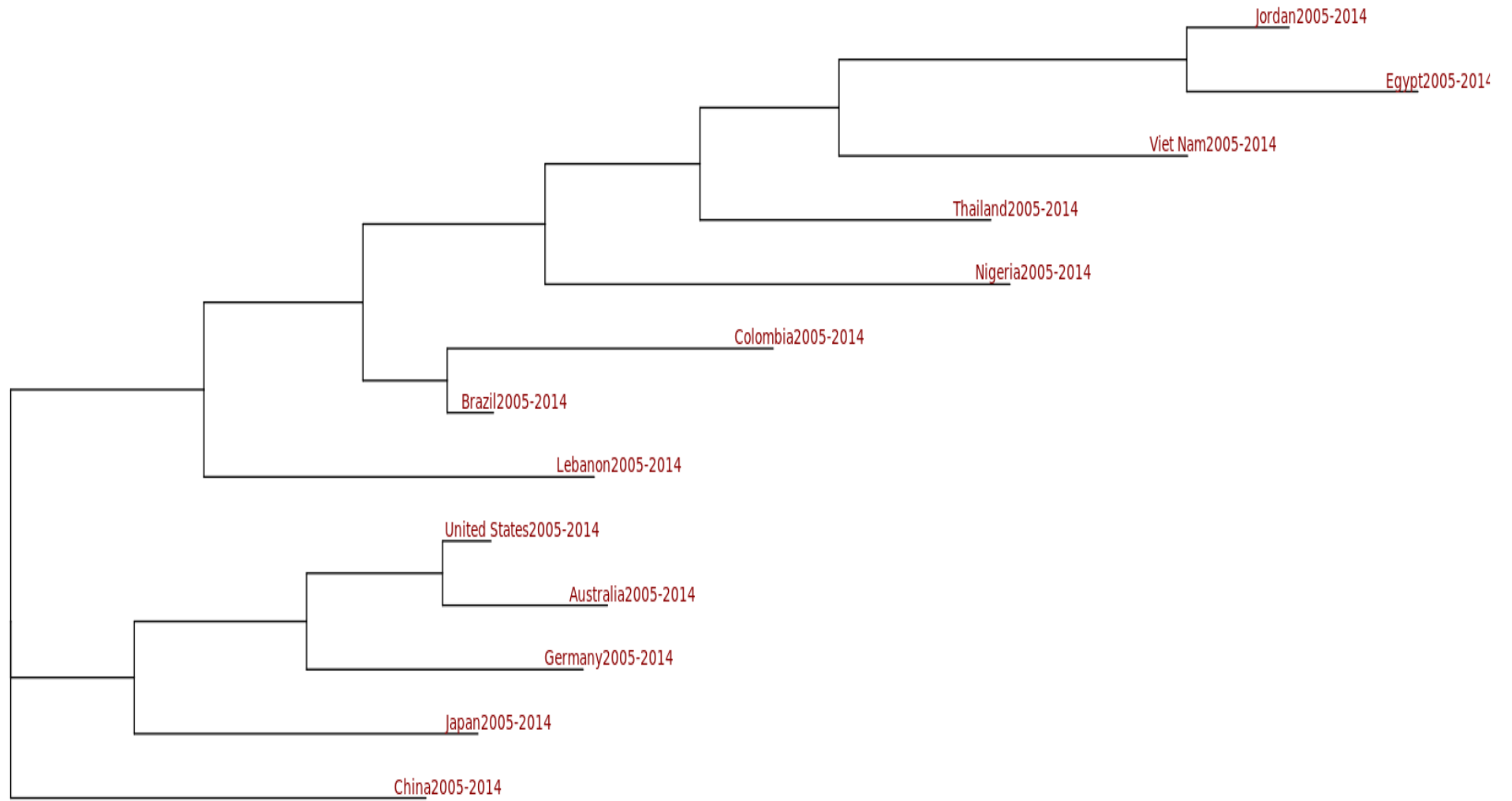
Reflections on the Age Item (Data Collection 3)

We are measuring the sociodemographic variable *age* across the four countries of our study. The countries in our study differ significantly in terms of average age and average life expectancy. A search on the statistics website Statista[163] revealed the following:

- The average life expectancy in **Egypt** is estimated to be 70.2 years at birth in 2022, while the average age in 2022 was 24.1 years.[164]

- The average life expectancy at birth in **Germany** in 2020 was 78.5 years for men and 83.4 years for women, while the average age in 2020 was 44.6 years.

- **Japan**: In 2022, the average life expectancy at birth in Japan is estimated to be 84.8 years, with a life expectancy of 87.8 years for women and 81.8 years for men, while the average age in 2022 was 48.7 years.

- In 2023, the average life expectancy at birth in the **US** is estimated to be around 79.7 years, with a life expectancy of around 82.2 years for women and around 77.3 years for men, while the average age in 2023 is 38.1 years.

We have tried to take this fact into account by specifying before data collection 3 how (ideally) the age distribution should be in the respective samples (in order to obtain a diverse sample in terms of age). The following parameters are decisive for us with regard to the age variable:

- **Germany and Japan**: here we are aiming for 7 "cohorts", whereby the cohorts 18-28; 29-38; 39-48; 49-58; 59-68; 69-78 should each have 106 cases and the cohort "79-85" 41 cases if possible.
- We keep the desired power-level in mind: (6 x 106 =) 636 + 41 = 677.

- **USA**: the seventh cohort/category is omitted here. So, we are aiming for the following 6 cohorts: 18-28; 29-38; 39-48; 49-58; 59-68; 69-78. Here we are aiming for 120 cases

---

[163] See: https://de.statista.com/statistik/daten/studie/749354/umfrage/lebenserwartung-in-aegypten/; https://de.statista.com/statistik/daten/studie/273406/umfrage/entwicklung-der-lebenserwartung-bei-geburt-in-deutschland-nach-geschlecht/; https://de.statista.com/statistik/daten/studie/18655/umfrage/lebenserwartung-in-japan/; https://de.statista.com/statistik/daten/studie/18670/umfrage/lebenserwartung-in-den-usa/

[164] See: https://de.statista.com/statistik/daten/studie/200664/umfrage/durchschnittsalter-der-bevoelkerung-in-den-usa/; https://de.statista.com/statistik/daten/studie/200666/umfrage/durchschnittsalter-der-bevoelkerung-in-japan/; https://de.statista.com/statistik/daten/studie/1084430/umfrage/durchschnittsalter-der-bevoelkerung-in-deutschland/; https://de.statista.com/statistik/daten/studie/1323636/umfrage/durchschnittsalter-der-bevoelkerung-in-aegypten/

for cohorts 18-28; 29-38; 39-48; 118 cases for cohorts 49-58; 59-68; and 81 cases for cohort 69-78.
- We keep the desired power-level in mind: (3 x 120 =) 360 + (2 x 118 =) 236 + (1 x 81) = 677.

- **Egypt**: Here the sixth cohort/category is dropped due to the comparatively lowest average life expectancy (among our four groups). We therefore aim for the following 5 cohorts: 18-28; 29-38; 39-48; 49-58; 59-68. We aim for 155 cases for the cohorts 18-28; 29-38; 39-48; 49-58; and 57 cases for the cohort 59-68.
- We keep the desired power-level in mind: (4 x 155 =) 620 + (1 x 57) = 677.

The specified cohorts are used as inclusion and exclusion criteria for the samples to be collected in data collection 3. However, we do not consider slight to moderate deviations from the target cohorts to be serious. Nevertheless, it should be noted that the age variable differs significantly between the countries in our study. Hence, possible effects of the age variable must therefore be statistically tested/investigated.

Supplement Chapter 3: MaC-DRS Development

Study 1 (Data Collection 1)

Table 80 presents our reliability findings for MaC-DRS from Study 1 and demonstrates all first-order dimensions come with sufficient reliability.[165]

Table 80: Study 1: Reliability of MaC-DRS Dimensions*

	Number of items	McDonald´s Omega	Cronbach´s Alpha
Family	4	0.8923 / good	0.8911 / good
Group	4	0.7951 / sufficient	0.7942 / sufficient
Deference	4	0.8251 / good	0.8204 / good
Heroism	4	0.9115 / excellent	0.9059 / excellent
Reciprocity	4	0.8327 / good	0.8334 / good
Fairness	4	0.8304 / good	0.8272 / good
Trustworthiness	4	0.8774 / good	0.8772 / good
Property	4	0.8638 / good	0.8632 / good

* Note: we report in addition to McDonald´s Omega also Cronbach´s Alpha for the latter is still more convenient in the literature (Gäde et al., 2020b); Color coding: green indicates first-order binding factors, blue indicates first-order individualizing factors and yellow those first-order factors for which we are unclear about the second-order factor assignment.

EFA Results Sample Size n = 792 (Data Collection 1)

Table 81: Varimax rotated principal factors EFA for MaC-DRS

Factor analysis/correlation				Number of obs = 792
Method: principal factors				Retained factors = 24
Rotation: orthogonal varimax				Number of params = 876
Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	4.50988	0.24336	0.1659	0.1659
Factor 2	4.26652	0.63462	0.1569	0.3228
Factor 3	3.63189	0.22431	0.1336	0.4564
Factor 4	3.40759	0.16815	0.1253	0.5817
Factor 5	3.23943	0.45617	0.1191	0.7008
Factor 6	2.78326	0.27190	0.1024	0.8032
Factor 7	2.51136	0.04275	0.0924	0.8956
Factor 8	2.46862	2.02076	0.0908	0.9863
Factor 9	0.44785	0.04115	0.0165	1.0028
Factor 10	0.40671	0.05285	0.0150	1.0178

*Note: Variance is equivalent to Eigenvalue; ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

---

[165] The questionnaires used for data collection 1, 2 and 3 can be found in the research plans available online.

Table 82: Quartimax rotated principal factors EFA for MaC-DRS

Factor analysis/correlation				Number of obs = 792
Method: principal factors				Retained factors = 24
Rotation: orthogonal quartimax				Number of params = 876

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	5.18584	0.56539	0.1907	0.1907
Factor 2	4.62044	0.87684	0.1699	0.3607
Factor 3	3.74360	0.38628	0.1377	0.4983
Factor 4	3.35732	0.22515	0.1235	0.6218
Factor 5	3.13217	0.43275	0.1152	0.7370
Factor 6	2.69942	0.56035	0.0993	0.8363
Factor 7	2.13907	0.10550	0.0787	0.9150
Factor 8	2.03357	1.60341	0.0748	0.9898
Factor 9	0.43017	0.02808	0.0158	1.0056
Factor 10	0.40208	0.06790	0.0148	1.0204

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 83: Oblique oblimin rotated principal factors EFA for MaC-DRS

Factor analysis/correlation			Number of obs = 792
Method: principal factors			Retained factors = 24
Rotation: oblique oblimin			Number of params = 876

Factor	Variance*	Proportion	Rotated factors are correlated
Factor 1**	6.94533	0.2554	
Factor 2	6.57111	0.2417	
Factor 3	6.39745	0.2353	
Factor 4	6.11618	0.2249	
Factor 5	6.06602	0.2231	
Factor 6	5.91698	0.2176	
Factor 7	5.70836	0.2099	
Factor 8	5.45759	0.2007	
Factor 9***	5.23860	0.1927	
Factor 10	4.48730	0.1650	
Factor 11	4.23350	0.1557	
Factor 12	2.47210	0.0909	
Factor 13	2.33395	0.0858	
Factor 14	2.17275	0.0799	
Factor 15	2.04674	0.0753	
Factor 16	2.02989	0.0747	
Factor 17	2.00979	0.0739	
Factor 18	1.84051	0.0677	
Factor 19	1.31377	0.0483	
Factor 20	1.03536	0.0381	
Factor 21	0.95160	0.0350	

* Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold. *** Note: only 9 factors have ≥ 3 items and factor loadings of ≥ 0.4. The 9-factor solution can be explained by the semantics of the "heroism" items (mac_31-mac_36). The heroism items load on two factors due to their semantics. This does not hinder a merging of the heroism items. Of all 9 factors, five factors have only the minimum number of (sufficient) items (i.e., 3). The remaining four factors have four items each. This explanation also applies to the other models of oblique rotation performed with the different sample sizes and is no longer mentioned in the other models.

Table 84: Reduced item-set — varimax rotated principal factors EFA*

Factor Analysis/Correlation			Number of obs = 792	
Method: Principal Factors			Retained Factors = 15	
Rotation: Orthogonal varimax			Number of params = 375	
Factor	*Variance***	*Difference*	*Proportion*	*Cumulative*
Factor 1*	3.11770	0.12762	0.1729	0.1729
Factor 2	2.99008	0.32906	0.1659	0.3388
Factor 3	2.66102	0.15763	0.1476	0.4864
Factor 4	2.50339	0.25429	0.1389	0.6253
Factor 5	2.24910	0.18639	0.1248	0.7500
Factor 6	2.06271	0.00570	0.1144	0.8644
Factor 7	2.05700	0.29189	0.1141	0.9785
Factor 8	1.76511	1.59851	0.0979	1.0764
Factor 9	0.16660	0.07849	0.0092	1.0857
Factor 10	0.08811	0.00659	0.0049	1.0906

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items. ** Note: Variance is equivalent to Eigenvalue. *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 85: Reduced item-set — oblique oblimin rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 792
Method: principal factors			Retained factors = 15
Rotation: oblique oblimin			Number of params = 375
Factor	*Variance***	*Proportion*	*Rotated factors are correlated*
Factor 1*	5.56460	0.3087	
Factor 2	5.10496	0.2832	
Factor 3	4.92408	0.2731	
Factor 4	4.90908	0.2723	
Factor 5	4.80317	0.2664	
Factor 6	4.41419	0.2449	
Factor 7	4.07522	0.2260	
Factor 8 **	3.42576	0.1900	
Factor 9	2.19386	0.1217	
Factor 10	0.59764	0.0332	

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold; **** Note: Only 8 factors fulfill the criteria of at least 3 items and factor loadings of ≥0. 30.

EFA Results Sample Size n = 628

Table 86: Varimax rotated principal factors EFA

		Number of obs = 628
Factor analysis/correlation		Number of obs = 628
Method: principal factors		Retained factors = 24
Rotation: orthogonal varimax		Number of params = 876

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	4.40234	0.12518	0.1600	0.1600
Factor 2	4.27716	0.48295	0.1554	0.3154
Factor 3	3.79422	0.40055	0.1379	0.4532
Factor 4	3.39366	0.25653	0.1233	0.5765
Factor 5	3.13713	0.27734	0.1140	0.6905
Factor 6	2.85979	0.37614	0.1039	0.7944
Factor 7	2.48364	0.03399	0.0902	0.8847
Factor 8	2.44966	1.94438	0.0890	0.9737
Factor 9	0.50528	0.07107	0.0184	0.9921
Factor 10	0.43421	0.05024	0.0158	1.0078

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level≥ 1.000 are emphasized in bold.

Table 87: Quartimax rotated principal factors EFA

		Number of obs = 628
Factor analysis/correlation		Number of obs = 628
Method: principal factors		Retained factors = 24
Rotation: orthogonal quartimax		Number of params = 876

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	5.26618	0.79213	0.1913	0.1913
Factor 2	4.47404	0.44996	0.1626	0.3539
Factor 3	4.02409	0.76219	0.1462	0.5001
Factor 4	3.26190	0.22705	0.1185	0.6187
Factor 5	3.03485	0.27111	0.1103	0.7289
Factor 6	2.76374	0.70698	0.1004	0.8293
Factor 7	2.05676	0.01702	0.0747	0.9041
Factor 8	2.03974	1.57542	0.0741	0.9782
Factor 9	0.46432	0.03481	0.0169	0.9951
Factor 10	0.42951	0.05776	0.0156	1.0107

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 88: Oblique oblimin rotated principal factors EFA

			Number of obs = 628
Factor analysis/correlation			
Method: principal factors			Retained factors = 24
Rotation: oblique oblimin			Number of params = 876

Factor	Variance*	Proportion	Rotated factors are correlated
Factor 1**	6.51816	0.2368	
Factor 2	6.44957	0.2343	
Factor 3	6.29194	0.2286	
Factor 4	6.11407	0.2222	
Factor 5	5.71966	0.2078	
Factor 6	5.56647	0.2023	
Factor 7	5.38753	0.1958	
Factor 8	5.10644	0.1855	
Factor 9	4.59767	0.1671	
Factor 10	4.31656	0.1568	
Factor 11	4.27507	0.1553	
Factor 12	3.57304	0.1298	
Factor 13	3.37417	0.1226	
Factor 14	3.20871	0.1166	
Factor 15	3.20463	0.1164	
Factor 16	2.40469	0.0874	
Factor 17	2.07955	0.0756	
Factor 18	1.72242	0.0626	
Factor 19	1.62864	0.0592	
Factor 20	1.15053	0.0418	
Factor 21	0.97615	0.0355	

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 89: Reduced item-set — varimax rotated principal factors EFA*

				Number of obs = 628
Factor analysis/correlation				
Method: principal Factors				Retained factors = 15
Rotation: orthogonal varimax				Number of params = 375

Factor	Variance**	Difference	Proportion	Cumulative
Factor 1***	3.06613	0.05768	0.1678	0.1678
Factor 2	3.00845	0.27093	0.1647	0.3325
Factor 3	2.73752	0.19706	0.1498	0.4823
Factor 4	2.54046	0.19087	0.1390	0.6213
Factor 5	2.34959	0.28147	0.1286	0.7499
Factor 6	2.06812	0.10187	0.1132	0.8631
Factor 7	1.96625	0.14425	0.1076	0.9707
Factor 8	1.82200	1.63807	0.0997	1.0705
Factor 9	0.18394	0.07868	0.0101	1.0805
Factor 10	0.10526	0.03505	0.0058	1.0805

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level≥ 1.000 are emphasized in bold.

Table 90: Reduced item-set — oblique oblimin rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 628
Method: principal factors			Retained factors = 15
Rotation: oblique oblimin			Number of params = 375

Factor	Variance**	Proportion	Rotated factors are correlated
Factor 1*	5.47503	0.2997	
Factor 2	5.31156	0.2907	
Factor 3	5.05621	0.2768	
Factor 4	4.85993	0.2660	
Factor 5	4.62496	0.2531	
Factor 6	4.25194	0.2327	
Factor 7	4.17721	0.2286	
Factor 8	3.60063	0.1971	
Factor 9	1.82460	0.0999	
Factor 10	0.85328	0.0467	

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

EFA Results Sample Size n = 574

Table 91: Varimax rotated principal factors EFA

Factor analysis/correlation				Number of obs = 574
Method: principal factors				Retained factors = 25
Rotation: orthogonal varimax				Number of params = 900

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	4.40508	0.16849	0.1608	0.1608
Factor 2	4.23660	0.59424	0.1547	0.3155
Factor 3	3.64235	0.37923	0.1330	0.4485
Factor 4	3.26312	0.06052	0.1191	0.5677
Factor 5	3.20260	0.56199	0.1169	0.6846
Factor 6	2.64062	0.10346	0.0964	0.7810
Factor 7	2.53715	0.03751	0.0926	0.8737
Factor 8	2.49965	1.91437	0.0913	0.9649
Factor 9	0.58528	0.15314	0.0214	0.9863
Factor 10	0.43214	0.02746	0.0158	1.0021

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 92: Quartimax rotated principal factors EFA

Factor analysis/correlation			Number of obs = 574	
Method: principal factors			Retained factors = 25	
Rotation: orthogonal quartimax			Number of params = 900	

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	5.11576	0.61751	0.1868	0.1868
Factor 2	4.49825	0.75888	0.1642	0.3510
Factor 3	3.73937	0.55575	0.1365	0.4876
Factor 4	3.18362	0.03972	0.1162	0.6038
Factor 5	3.14390	0.55837	0.1148	0.7186
Factor 6	2.58553	0.36485	0.0944	0.8130
Factor 7	2.22068	0.10117	0.0811	0.8941
Factor 8	2.11951	1.61657	0.0774	0.9715
Factor 9	0.50294	0.07267	0.0184	0.9899
Factor 10	0.43028	0.05230	0.0157	1.0056

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 93: Oblique oblimin rotated principal factors EFA

Factor analysis/correlation			Number of obs = 574
Method: principal factors			Retained factors = 25
Rotation: oblique oblimin			Number of params = 900

Factor	Variance*	Proportion	Rotated factors are correlated
Factor 1**	6.84652	0.2500	
Factor 2	6.30641	0.2303	
Factor 3	6.24849	0.2282	
Factor 4	6.02923	0.2201	
Factor 5	5.72137	0.2089	
Factor 6	5.69629	0.2080	
Factor 7	5.27020	0.1924	
Factor 8	4.85212	0.1772	
Factor 9	4.51795	0.1650	
Factor 10	4.40566	0.1609	
Factor 11	3.67181	0.1341	
Factor 12	3.01906	0.1102	
Factor 13	3.01507	0.1101	
Factor 14	2.44552	0.0893	
Factor 15	2.07124	0.0756	
Factor 16	1.94276	0.0709	
Factor 17	1.77297	0.0647	
Factor 18	1.50173	0.0548	
Factor 19	1.30095	0.0475	
Factor 20	1.24999	0.0456	
Factor 21	1.08728	0.0397	
Factor 22	1.06818	0.0390	
Factor 23	0.98597	0.0360	

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 94: Reduced item-set — varimax rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 574	
Method: principal Factors			Retained factors = 15	
Rotation: orthogonal varimax			Number of params = 375	
Factor	*Variance***	*Difference*	*Proportion*	*Cumulative*
Factor 1*	3.07876	0.06812	0.1705	0.1705
Factor 2	3.01064	0.32688	0.1667	0.3372
Factor 3	2.68376	0.25644	0.1486	0.4858
Factor 4	2.42732	0.25206	0.1344	0.6203
Factor 5	2.17526	0.06856	0.1205	0.7407
Factor 6	2.10670	0.06530	0.1167	0.8574
Factor 7	2.04140	0.19612	0.1130	0.9704
Factor 8	1.84528	1.0726	0.1022	1.0726
Factor 9	0.18454	0.07887	0.0102	1.0828
Factor 10	0.10567	0.02030	0.0059	1.0887

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 95: Reduced item-set — oblique oblimin rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 574
Method: principal factors			Retained factors = 15
Rotation: oblique oblimin			Number of params = 375
Factor	*Variance***	*Proportion*	*Rotated factors are correlated*
Factor 1*	5.41423	0.2998	
Factor 2	4.94234	0.2737	
Factor 3	4.68348	0.2594	
Factor 4	4.61080	0.2553	
Factor 5	4.58645	0.2540	
Factor 6	4.22177	0.2338	
Factor 7	4.06727	0.2252	
Factor 8 ****	3.03729	0.1682	
Factor 9	1.85035	0.1025	
Factor 10	1.22555	0.0679	
Factor 11	0.90129	0.0499	

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold; **** Only 8 factors have the minimal required number of manifest indicators (i.e., 3) and factor loadings of ≥0.30.

EFA Results Sample Size n = 287 (Independent EFA Sample)

Table 96: Varimax rotated principal factors EFA

Factor analysis/correlation				Number of obs = 287
Method: principal factors				Retained factors = 27
Rotation: orthogonal varimax				Number of params = 945

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	4.42476	0.59070	0.1518	0.1518
Factor 2	3.83405	0.25345	0.1315	0.2833
Factor 3	3.58060	0.08387	0.1228	0.4061
Factor 4	3.49673	0.20467	0.1199	0.5260
Factor 5	3.29205	0.06314	0.1129	0.6389
Factor 6	3.22892	0.71177	0.1107	0.7496
Factor 7	2.51715	0.35392	0.0863	0.8360
Factor 8	2.16323	1.52321	0.0742	0.9102
Factor 9	0.64002	0.02431	0.0220	0.9321
Factor 10	0.61570	0.17453	0.0211	0.9532

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 97: Quartimax rotated principal factors EFA

Factor analysis/correlation				Number of obs = 287
Method: principal factors				Retained factors = 27
Rotation: orthogonal quartimax				Number of params = 945

Factor	Variance*	Difference	Proportion	Cumulative
Factor 1**	4.53949	0.36084	0.1557	0.1557
Factor 2	4.17865	0.55089	0.1433	0.2990
Factor 3	3.62776	0.06325	0.1244	0.4234
Factor 4	3.56451	0.28329	0.1223	0.5457
Factor 5	3.28122	0.04756	0.1125	0.6582
Factor 6	3.23366	0.72503	0.1109	0.7691
Factor 7	2.50863	0.67210	0.0860	0.8552
Factor 8	1.83653	1.20475	0.0630	0.9181
Factor 9	0.63178	0.09058	0.0217	0.9398
Factor 10	0.54121	0.12336	0.0186	0.9584

*Note: Variance is equivalent to Eigenvalue. ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 98: Oblique oblimin rotated principal factors EFA

Factor analysis/correlation			Number of obs = 287	
Method: principal factors			Retained factors = 27	
Rotation: oblique oblimin			Number of params = 945	

Factor	Variance*	Proportion	Rotated factors are correlated
Factor 1**	5.72465	0.1963	
Factor 2	5.14346	0.1764	
Factor 3	4.83865	0.1660	
Factor 4	4.74063	0.1626	
Factor 5	4.64510	0.1593	
Factor 6	4.50179	0.1544	
Factor 7	4.22431	0.1449	
Factor 8	4.11916	0.1413	
Factor 9	4.02600	0.1381	
Factor 10	3.97530	0.1363	
Factor 11	3.51255	0.1205	
Factor 12	3.49393	0.1198	
Factor 13	3.35336	0.1150	
Factor 14	3.10661	0.1065	
Factor 15	2.99120	0.1026	
Factor 16	2.85948	0.0981	
Factor 17	2.83856	0.0974	
Factor 18	2.78366	0.0955	
Factor 19	2.53685	0.0870	
Factor 20	2.44344	0.0838	
Factor 21	1.90567	0.0654	
Factor 22	1.77147	0.0608	
Factor 23	1.72642	0.0592	
Factor 24	1.55331	0.0533	
Factor 25	1.50413	0.0516	
Factor 26***	1.09747	0.0376	
Factor 27	0.61318	0.0210	

*Note: Variance is equivalent to Eigenvalue; ** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold; *** Note: The principal factors EFA with oblique oblimin rotation retains 26 factors with eigenvalue ≥ 1. However, of all the factors with Eigenvalue ≥ 1 only 6 factors have the required minimal number of manifest indicators (i.e., ≥ 3 items) and factor loadings of ≥ 0.30.

Table 99: Reduced item-set — varimax rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 287	
Method: principal Factors			Retained factors = 16	
Rotation: orthogonal varimax			Number of params = 392	

Factor	Variance**	Difference	Proportion	Cumulative
Factor 1***	3.12321	0.21437	0.1667	0.1667
Factor 2	2.90884	0.27968	0.1552	0.3219
Factor 3	2.62916	0.11501	0.1403	0.4622
Factor 4	2.51415	0.15299	0.1342	0.5963
Factor 5	2.36116	0.25414	0.1260	0.7223
Factor 6	2.10702	0.02164	0.1124	0.8348
Factor 7	2.08538	0.29648	0.1113	0.9460
Factor 8	1.78889	1.57733	0.0955	1.0415
Factor 9	0.21156	0.01018	0.0113	1.0528
Factor 10	0.20138	0.00679	0.0107	1.0635

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold.

Table 100: Reduced item-set — oblique oblimin rotated principal factors EFA*

Factor analysis/correlation			Number of obs = 287
Method: principal factors			Retained factors = 16
Rotation: oblique oblimin			Number of params = 392

Factor	*Variance***	*Proportion*	*Rotated factors are correlated*
Factor 1***	4.77968	0.2551	
Factor 2	4.01998	0.2145	
Factor 3	3.97273	0.2120	
Factor 4	3.90778	0.2085	
Factor 5	3.84724	0.2053	
Factor 6	3.82618	0.2042	
Factor 7	3.47761	0.1856	
Factor 8	2.57397	0.1374	
Factor 9	2.46983	0.1318	
Factor 10	1.83397	0.0979	
Factor 11	1.42261	0.0759	
Factor 12****	1.31625	0.0702	
Factor 13	0.79225	0.0423	

* Note: This EFA is performed with the reduced set of 32 MaC-DRS items; ** Note: Variance is equivalent to Eigenvalue; *** Note: Factors with Eigenvalue above criteria level ≥ 1.000 are emphasized in bold; **** Only 8 factors have the minimal required number of manifest indicators (i.e., 3) and factor loadings of ≥ 0.30.

Study 2 (Data Collection 2)

Table 101: Questionnaire Versions and Structure (data collection 2)

Questionnaire Version	Order of instrument appearance in the questionnaire						Targeted minimum sample size
A1	**MaC-DRS**	EI *	SC **	MaC-Q	EI	Sociodemographic variables	n = 500
A2	**MaC-DRS**	EI	SC	MFQ-1	EI	Sociodemographic variables	n = 500
B1	MaC-Q	EI	SC	**MaC-DRS**	EI	Sociodemographic variables	n = 500
B2	MFQ1	EI	SC	**MaC-DRS**	EI	Sociodemographic variables	n = 500

* "EI" stands for evaluation items; ** "SC" stands for social cohesion; EI and SC are part of another study and will not be addressed here.

EFA Results MaC-DRS Long (32 Items) and Short (24 Items) Version

Table 102: EFA direct oblique oblimin rotation long MaC-DRS *

Factor analysis/correlation	Number of obs = 1162
Method: principal factors	Retained factors = 15
Rotation: oblique oblimin (Kaiser off)	Number of params = 375

Factor	Variance	Proportion	Rotated factors are correlated
Factor 1	**12.24408**	0.5006	
Factor 2	**12.16275**	0.4973	
Factor 3	**12.09738**	0.4946	
Factor 4	**10.88188**	0.4449	
Factor 5	**10.53911**	0.4309	
Factor 6	**10.02422**	0.4099	
Factor 7	**8.20845**	0.3356	
Factor 8	**8.16685**	0.3339	
Factor 9	**1.59676**	0.0653	
Factor 10	**1.48944**	0.0609	
Factor 11	**1.26142**	0.0516	
Factor 12	**1.20025**	0.0491	
Factor 13	0.84869	0.0347	
Factor 14	0.35620	0.0146	
Factor 15	0.25523	0.0104	

LR test: independent vs. saturated: chi2(496) = 4.2e+04 Prob>chi2 = 0.0000

* PFA EFA (oblique oblimin rotation) for MaC-DRS; Item-set encompasses 32 items; Factors that exceed the Eigenvalue criterion are marked in bold. Factor extraction criterion: 12 factors have an Eigenvalue (Variance) ≥ 1.000 yet only 8 factors have a sufficient number of items (≥ 3); Cut-off values (factor loadings) for items range from 0.6309 (min.) – 0.9304 (max.) and are sufficient.

Table 103: EFA direct oblique oblimin rotation short MaC-DRS *

Factor analysis/correlation	Number of obs = 1162
Method: principal factors	Retained factors = 10
Rotation: oblique oblimin (Kaiser off)	Number of params = 195

Factor	Variance	Proportion	Rotated factors are correlated
Factor 1	**9.28517**	0.5165	
Factor 2	**9.10939**	0.5067	
Factor 3	**8.75331**	0.4869	
Factor 4	**8.31803**	0.4627	
Factor 5	**7.90122**	0.4395	
Factor 6	**7.47312**	0.4157	
Factor 7	**6.89655**	0.3836	
Factor 8	**6.06862**	0.3376	
Factor 9	**1.64168**	0.0913	
Factor 10	0.27644	0.0154	

LR test: independent vs. saturated: chi2(496) = 4.2e+04 Prob>chi2 = 0.0000

* PFA EFA (oblique oblimin rotation) for MaC-DRS; Item-set encompasses 24 items; Factors that exceed the Eigenvalue criterion are marked in bold. Factor extraction criterion: 9 factors have an Eigenvalue

(Variance) ≥ 1.000 yet only 8 factors have a sufficient number of items (≥ 3); Cut-off values (factor loadings) for items range from 0.7081 (min.) – 0.9419 (max.) and are sufficient.

EFA Results MFQ-1 Relevance and Judgment Scale

Table 104: EFA direct oblique oblimin rotation MFQ-1 Relevance Scale *

Factor analysis/correlation			Number of obs = 511
Method: principal factors			Retained factors = 7
Rotation: oblique oblimin (Kaiser off)			Number of params = 84

Factor	Variance	Proportion	Rotated factors are correlated
Factor 1	**4.09440**	0.6115	
Factor 2	**3.38029**	0.5048	
Factor 3	**3.13420**	0.4681	
Factor 4	**2.50641**	0.3743	
Factor 5	**1.84471**	0.2755	
Factor 6	**1.44337**	0.2156	
Factor 7	**1.01544**	0.1516	

LR test: independent vs. saturated: chi2(496) = 4.2e+04 Prob>chi2 = 0.0000

* PFA EFA (oblique oblimin rotation) for MFQ-1 Relevance Scale; Item-set encompasses 15 items; Factors that exceed the Eigenvalue criterion are marked in bold. Factor extraction criterion: 7 factors have an Eigenvalue (Variance) ≥ 1.000 yet only 3 factors have a sufficient number of items (≥ 3); Cut-off values (factor loadings) for factors with at least 3 items range from 0.3232 (min.) – 0.8220 (max.) and are partly insufficient.

Table 105: EFA direct oblique oblimin rotation MFQ-1 Judgment Scale *

Factor analysis/correlation			Number of obs = 511
Method: principal factors			Retained factors = 7
Rotation: oblique oblimin (Kaiser off)			Number of params = 84

Factor	Variance	Proportion	Rotated factors are correlated
Factor 1	**2.42675**	0.5345	
Factor 2	**2.25726**	0.4972	
Factor 3	**1.80025**	0.3965	
Factor 4	**1.78884**	0.3940	
Factor 5	**1.66872**	0.3675	
Factor 6	**1.48758**	0.3276	
Factor 7	0.90637	0.1996	

LR test: independent vs. saturated: chi2(496) = 4.2e+04 Prob>chi2 = 0.0000

* PFA EFA (oblique oblimin rotation) for MFQ-1 Judgment Scale; Item-set encompasses 15 items; Factors that exceed the Eigenvalue criterion are marked in bold. Factor extraction criterion: 6 factors have an Eigenvalue (Variance) ≥ 1.000 yet only 3 factors have a sufficient number of items (≥ 3); Cut-off values (factor loadings) for factors with at least 3 items range from 0.2535 (min.) – 0.7221 (max.) and are partly insufficient.

EFA Results MaC-Q Relevance Scale

Table 106: EFA direct oblique oblimin rotation MaC-Q Relevance Scale*

Factor analysis/correlation	Number of obs	=	653
Method: principal factors	Retained factors =		9
Rotation: oblique oblimin (Kaiser off)	Number of params =		153

Factor	Variance	Proportion	Rotated factors are correlated
Factor 1	**7.45082**	0.4966	
Factor 2	**7.13908**	0.4758	
Factor 3	**6.34026**	0.4226	
Factor 4	**6.29554**	0.4196	
Factor 5	**6.25601**	0.4169	
Factor 6	**5.55424**	0.3702	
Factor 7	**5.14714**	0.3430	
Factor 8	**1.86360**	0.1242	
Factor 9	0.44964	0.0300	

LR test: independent vs. saturated: chi2(496) = 4.2e+04 Prob>chi2 = 0.0000

* PFA EFA (oblique oblimin rotation) for MaC-Q Relevance Scale; Item-set encompasses 21 items; Factors that exceed the Eigenvalue criterion are marked in bold. Factor extraction criterion: 8 factors have an Eigenvalue (Variance) ≥ 1.000 yet only 7 factors have a sufficient number of items (≥ 3); Cut-off values (factor loadings) for factors with at least 3 items range from 0.3887 (min.) – 0.9031 (max.) and are sufficient.

Moral Scales — Reliability Scores

Table 107: Reliability Scores MaC-DRS long version

	Number of items	*McDonald′s Omega*	*Cronbach′s Alpha*
Reliability long MaC-DRS			
Family	4	0.9378/ excellent	0.9372/ excellent
Group	4	0.9296/ excellent	0.9292/ excellent
Deference	4	0.8832/ good	0.8819/ good
Heroism	4	0.9164/ excellent	0.9119/ excellent
Reciprocity	4	0.9180/ excellent	0.9177/ excellent
Fairness	4	0.9433/ excellent	0.9425/ excellent
Trustworthiness	4	0.9617/ excellent	0.9616/ excellent
Property	4	0.9630/ excellent	0.9630/ excellent

Table 108: Reliability Scores of Moral Scales in Comparison

	Number of items	McDonald´s Omega	Cronbach´s Alpha
Reliability MFQ-1 Relevance Scale			
Purity *	3	0.5620/ poor	0.5476/ poor
Authority	3	0.6761/ questionable	0.6758/ questionable
Ingroup/Loyalty	3	. / **	0.4938/ unacceptable
Harm	3	0.7084/ acceptable	0.6995/ questionable
Fairness	3	0.7130/ acceptable	0.7092/ acceptable
Reliability MFQ-1 Judgment Scale			
Purity	3	0.6991/ questionable	0.6928/ questionable
Authority	3	0.5235/ poor	0.5157/ poor
Ingroup/Loyalty	3	0.5415/ poor	0.5378/ poor
Harm	3	0.5874/ poor	0.5871/ poor
Fairness	3	0.6121/ questionable	0.5452/ poor
Reliability MaC-Q Relevance Scale			
Family	3	0.9207/ excellent	0.9190/ excellent
Group	3	0.9433/ excellent	0.9428/ excellent
Reciprocity	3	0.8925/ good	0.8836/ good
Heroism	3	0.8885/ good	0.8839/ good
Deference	3	0.8812/ good	0.8794/ good
Fairness	3	0.8779/ good	0.8736/ good
Property	3	0.8947/ good	0.8845/ good
Reliability short MaC-DRS			
Family	3	0.9297/ excellent	0.9296/ excellent
Group	3	0.9149/ excellent	0.9139/ excellent
Deference	3	0.8717/ good	0.8707/ good
Heroism	3	0.8960/ good	0.8897/ good
Reciprocity	3	0.9056/ excellent	0.9053/ excellent
Fairness	3	0.9431/ excellent	0.9430/ excellent
Trustworthiness	3	0.9519/ excellent	0.9515/ excellent
Property	3	0.9552/ excellent	0.9549/ excellent

* Color coding: green indicates first-order binding factors, blue indicates first-order individualizing factors, and yellow indicates the domains (factors) that we consider to be between binding and individualizing. **The score could not be computed.

Intuitive Dominance of Individualizing Morality in Germany (Data Collection 2)

The German data from data collection 2 provide further evidence of an individualizing moral system. The mean values for the moral domains captured via Mac-DRS in all federal states in Germany support the statement that property, fairness and trustworthiness are more relevant than the binding domains of family, in-group and deference. The grand mean also confirms the statement just made. See the table below for the respective evidence.

Table 109: Raw means of MaC-DRS moral domains across Germany

Domain Relevance Across German Federal States	Fairness Mean (Std. Dev.)	Trust-worthiness Mean (Std. Dev.)	Property Mean (Std. Dev.)	Heroism Mean (Std. Dev.)	Reciprocity Mean (Std. Dev.)	Family Mean (Std. Dev.)	In-Group Mean (Std. Dev.)	Deference Mean (Std. Dev.)
Baden Württemberg (n = 272)	5.484 (1.484)	5.368 (1.642)	5.540 (1.568)	4.606 (1.437)	4.450 (1.438)	4.504 (1.555)	3.950 (1.470)	3.841 (1.499)
Bayern (n = 328)	5.459 (1.459)	5.283 (1.617)	5.429 (1.598)	4.711 (1.354)	4.518 (1.420)	4.675 (1.506)	4.102 (1.400)	3.829 (1.442)
Berlin (n = 103)	5.436 (1.436)	5.466 (1.504)	5.436 (1.562)	4.527 (1.566)	4.550 (1.505)	**4.249 (1.619)**	**3.838 (1.636)**	**3.699 (1.580)***
Brandenburg (n = 88)	5.246 (1.246)	5.325 (1.624)	5.348 (1.584)	4.556 (1.510)	4.598 (1.394)	4.481 (1.735)	3.962 (1.410)	4.147 (1.417)
Bremen (n = 50)	5.493 (1.493)	5.48 (1.427)	5.453 (1.453)	4.573 (1.519)	4.786 (1.484)	4.706 (1.621)	**4.353 (1.537)**	4.213 (1.564)
Hamburg (n = 83)	5.558 (1.558)	5.510 (1.491)	**5.686 (1.366)**	**4.871 (1.443)**	**4.799 (1.513)**	**4.891 (1.596)**	4.184 (1.471)	**4.317 (1.483)**
Hessen (n = 161)	5.521 (1.521)	5.412 (1.630)	5.583 (1.624)	4.786 (1.418)	4.753 (1.392)	4.782 (1.586)	4.171 (1.473)	4.089 (1.595)
Mecklenburg Vorpommern (n = 93)	5.440 (1.440)	5.261 (1.495)	5.605 (1.511)	4.508 (1.278)	4.512 (1.483)	4.598 (1.498)	3.989 (1.321)	4.193 (1.575)
Niedersachsen (n = 197)	5.389 (1.389)	5.311 (1.490)	5.453 (1.597)	4.700 (1.311)	4.516 (1.377)	4.629 (1.490)	3.981 (1.375)	4.201 (1.381)

Nordrhein Westfalen (n = 446)	5.462 (1.462)	5.410 (1.571)	5.546 (1.560)	4.674 (1.426)	4.566 (1.456)	4.582 (1.563)	4.033 (1.378)	4.125 (1.435)
Rheinland-Pfalz (n = 110)	5.290 (1.290)	5.166 (1.759)	5.324 (1.740)	4.778 (1.432)	4.515 (1.485)	4.557 (1.584)	3.954 (1.398)	4.109 (1.365)
Saarland (n = 56)	5.25 (1.25)	5.333 (1.455)	5.428 (1.620)	4.517 (1.550)	4.494 (1.426)	4.851 (1.600)	4.154 (1.553)	3.976 (1.439)
Sachsen (n = 110)	5.087 (1.087)	5.196 (1.553)	5.224 (1.796)	4.727 (1.468)	4.421 (1.536)	4.675 (1.502)	4.042 (1.502)	4.2 (1.666)
Sachsen-Anhalt (n = 84)	**4.936 (1.936)**	**5.011 (1.780)**	**4.956 (1.857)**	**4.214 (1.575)**	**4.380 (1.547)**	4.301 (1.706)	3.924 (1.366)	4.027 (1.506)
Schleswig-Holstein (n = 91)	**5.571 (1.571)**	**5.538 (1.226)**	5.586 (1.341)	4.824 (1.258)	4.622 (1.334)	4.666 (1.300)	4.051 (1.259)	4.179 (1.373)
Thüringen (n = 84)	5.158 (1.870)	5.174 (1.805)	5.166 (1.828)	4.527 (1.755)	4.440 (1.535)	4.432 (1.754)	3.964 (1.581)	3.972 (1.588)
N = 2,356								
Grand Mean	**5.400 (1.629)**	**5.337 (1.584)**	**5.456 (1.602)**	**4.656 (1.433)**	**4.546 (1.447)**	**4.597 (1.562)**	**4.031 (1.430)**	**4.041 (1.484)**

*Note: Numbers in blue highlight the domain specific lowest mean value across federal states; numbers in red highlight the domain specific highest mean value across federal states.

Table 110: Raw mean values of MaC-DRS higher-order factors across German federal states (data collection 2)

Moral Relevance Across German Federal States	Individualizing Mean (Std. Dev.)	Binding Mean (Std. Dev.)	General disposition of cooperation Mean (Std. Dev.)	Freq.
Baden Württemberg	5.464 (1.473)	4.098 (1.276)	4.528 (1. 276)	272
Bayern	5.390 (1.473)	4.202 (1.227)	4.614 (1. 227)	328
Berlin	5.446 (1.325)	3.928 (1.373)	4.538 (1. 373)	103
Brandenburg	5.306 (1.471)	4.196 (1.335)	4.577 (1. 335)	88
Bremen	5.475 (1.371)	4.424 (1.344)	4.68 (1. 344)	50
Hamburg	5.585 (1.328)	4.464 (1.389)	4.835 (1. 389)	83
Hessen	5.505 (1.458)	4.347 (1.260)	4.770 (1. 260)	161
Mecklenburg Vorpommern	5.436 (1.372)	4.260 (1.220)	4.510 (1. 220)	93
Niedersachsen	5.384 (1.393)	4.270 (1.231)	4.608 (1. 231)	197
Nordrhein Westfalen	5.473 (1.441)	4.247 (1.288)	4.620 (1. 288)	446
Rheinland-Pfalz	5.260 (1.612)	4.207 (1.313)	4.646 (1. 313)	110
Saarland	5.337 (1.455)	4.327 (1.387)	4.505 (1. 387)	56
Sachsen	5.169 (1.595)	4.306 (1.389)	4.574 (1. 389)	110
Sachsen-Anhalt	4.968 (1.706)	4.084 (1.396)	4.297 (1. 396)	84
Schleswig-Holstein	5.565 (1.096)	4.299 (1.079)	4.723 (1. 079)	91
Thüringen	5.166 (1.719)	4.123 (1.484)	4.484 (1. 484)	84
Total				2,356

Table 111: Grand mean MaC-DRS higher-order factors (data collection 2; German sample)

MaC-DRS Higher-Order Factors	Obs.	Mean	Std. Dev.	Min	Max
Individual-izing	2356	5.397991	1.458354	1	7
General disposition of cooperation	2356	4.601655	1.292494	1	7
Binding	2356	4.223401	1.271716	1	7c

Note: the higher-order factors are composed of the following first-order factors: fairness, trustworthiness and property (*individualizing*); reciprocity and heroism (*general disposition of cooperation*); family, in-group and deference (*binding*).

Study 3 (Data Collection 3)

EFA Insights MaC-DRS Across the Cultural Samples

We will review the data-driven EFA results for the four study groups step by step, starting with the population we are already familiar with through Study 1 and 2. Please note that the EFA results for the pan-cultural sample are not presented here, but in the main text of Chapter 3. In the data-driven approach to investigate the psychometric properties of MaC-DRS, we will focus exclusively on the EFA results obtained using oblique (oblimin) rotation and principal factor analysis.

As far as the EFA results of our third MaC-DRS study are concerned, we yield the following findings for the **GER-sample**: The **long MaC-DRS version** (32 items; KOM (overall) = 0.9693) shows 11 factors with an Eigenvalue of ≥ 1. However, the data suggests that only 8 (first-order) factors with four items each have sufficient factor loadings. The range of the respective factor loadings is from 0.424 to 0.922. Note though that only one item yields a factor loading below 0.562. Overall, the loadings are sufficient. When looking at the results, we can thus conclude that the exploratory findings from our previous studies on MaC-DRS (long version), which were tested in Germany, can be replicated. Looking next at the **MaC-DRS short version** (24 items; KOM (overall) = 0.954), the EFA yields 8 factors (of first order) with an Eigenvalue of ≥ 1 — each of these factors comprises three items with sufficient factor loadings (range: 0.653 - 0.924). The results of the German sample (Study 3; n = 751) hence replicate the results of our previous MaC-DRS (short version) studies. All in all, we found further proof of the usability of our newly developed research tool.

Taking now a look at the Japanese (**JP-sample**; n = 740) MaC-DRS results, we see the following: The EFA for the **long version** (32 items; KOM (overall) = 0.9524) anew yields 11 factors with an eigenvalue of ≥ 1. Of these factors, once again only 8 factors have a sufficient number of items (≥ 3) with an equally sufficient factor loading. It should also be noted that all factors (of first order) except for one have 4 items with good factor loadings (factor loading range: 0.609 to 0.898). The item "*Someone disregards general rules*" from the deference domain shows several cross-loadings and altogether no good loading on the factor that we would theoretically assign to the corresponding item. In the context of factor loadings, this item therefore fails the psychometric test in the JP-sample. So, if we look at the exploratory MaC-DRS (long version) results, we can see that the 8-dimensional structure also emerges from the data in the Japanese sample. However, it should be borne in mind that the item "Someone

disregards general rules" does not meet the psychometric requirements in this sample and should ideally be omitted. Turning to the **short MaC-DRS version** (24 items; KOM (overall) = 0.9343) 8 (first-order) factors emerge from the data and show an Eigenvalue of ≥ 1. All factors comprise three items each. The range of factor loadings yields sufficient values: 0.581 - 0.905. In summary, the exploratory MaC-DRS (short version) results suggest that the 8-dimensional structure can also be found in the Japanese sample. We were therefore able to replicate the previous results based on the JP-sample and the MaC-DRS items translated into Japanese.

Next, we turn to the exploratory investigation of MaC-DRS in the **US-sample** (n = 745). Inspecting first the **long MaC-DRS version** (32 items; KOM (overall) = 0.975), we find 15 factors with an eigenvalue of ≥ 1, but only 8 factors emerge from the data that have a sufficient number of items with sufficient factor loadings. If we set a minimum factor loading of 0.40, we can see that all factors, with one exception, each include four items with good loadings. The item "*Someone ignores the wishes of their own family and kin*" (family domain) only exhibits a factor loading of 0.366 in the EFA model. However, if we were to exclude this item, the factor loading range would indicate consistently solid values for an exploratory study: 0.477 - 0.898. The overall picture of the EFA results for the long version of the MaC-DRS is therefore satisfactory and we can largely replicate the results from Germany based on the US-sample and the translated items. If we now look at the **short MaC-DRS version** (24 items; KOM (overall) = 0.967) in the context of an EFA, we obtain the following results: In total 10 (first-order) factors show an Eigenvalue of ≥ 1, but only 8 factors have items with partly sufficient factor loadings. One class of items, those of the family domain, fall slightly short in the context of factor loadings compared to the others items. The family domain items have a factor loading range of 0.320 to 0.368 in our exploratory analysis. The other items, in contrast, show a factor loading range of 0.633 to 0.917 and thus good results. If we summarize the short version MaC-DRS insights, the results from Germany can therefore be replicated in the US-sample by and large, although the factor loadings of the family domain items drop somewhat and require closer examination for accuracy of fit in the subsequent confirmatory analyses.

Prior to addressing the pan-cultural analysis, we turn to the EFA MaC-DRS results for the **Egyptian sample** (n = 746). The exploratory analysis of the **long MaC-DRS version** (32 items; KOM (overall) = 0.983) yields 14 factors with an Eigenvalue of ≥ 1. This time, though, only a total of 5 factors, each comprising four items, exhibit good factor loadings (range: 0.430 to 0.946). Looking at the four items of the in-group domain, one outlier stands out (item: "*Someone is not faithful to their group*", factor loading = 0.334), while the other items have

satisfactory factor loadings (range: 0.653 to 0.827). Furthermore, the four reciprocity items do not show good factor loadings overall in the EG-sample exploratory analysis, as indicated by the range of factor loadings: 0.294 to 0.342. Finally, a rather divided picture emerges when we look at the factor loadings for the deference domain. In addition to weak factor loadings on the deference factor itself, we found several cross-loadings for two items, as the following values indicate: factor loading 0.763 (item "*Someone defies a recognized person of respect*"); factor loading 0.715 (item "*Someone does not behave according to their social position*"); factor loading 0.197 (item "*Someone disregards general rules*"); factor loading 0.132 (item "*Someone acts disrespectfully towards an authority*"). Taken together, based on the exploratory analyses conducted, we can only assume the 8-dimensional structure of the MaC-DRS for the Egyptian sample to a very limited extent. This is mainly due to two poorly loading items in the deference domain, but we also found other items that overall failed to reveal good factor loadings. We can thus, on the basis of the EG-sample, only partially replicate the MaC-DRS long version results from our previous studies. All in all, the results described for the EG-sample require special attention and further investigation in the confirmatory factor analysis that we conduct below. When inspecting the EFA for the **short MaC-DRS version** (24 items; KOM (overall) = 0.977) in the Egyptian sample, results yield in total 10 (first-order) factors with an Eigenvalue of ≥ 1. As in the long version, we also find a somewhat more complex pattern in the scale version comprising only 24 items. Three factors emerge clearly from the data. These factors are fairness, property and family that each comprise three items that show good loadings, as indicated by the factor loading range of: 0.774 to 0.917. With regard to the trustworthiness domain, the three corresponding items exhibit merely mediocre factor loading results (range: 0.329 to 0.385). In addition, the deference factor displays two items with good factor loadings (0.797 and 0.809) but also one outlier item is found with a loading of only 0.188 (deference item: "*Someone acts disrespectfully towards an authority*"). Next to deference also the in-group domain yields two items with good loading (0.672 and 0.756) and one item with a mediocre factor loading (0.399; in-group item: "*Someone is not faithful to their group*"). Furthermore, the heroism domain items show altogether poor to mediocre results in our exploratory investigation (factor loading range: 0.167 to 0.257), with one item in particular not performing well (heroism item: "*Someone does not stand up for the physical and psychological integrity of strangers out of fear*", factor loading = 0.167). Finally, for the items in the reciprocity domain, only low factor loadings can be determined, as can be seen from the corresponding range: 0.190 to 0.276. The item "*Someone does not reward a favor with something in return*" stands out in this case, as its factor loading

forms the lower end point of the corresponding range. The exploratory findings for MaC-DRS (short version) based on the Egyptian sample indicate all in all that we cannot simply assume a replication of the results of our previous studies. Rather, it is clear that problems with the 8-dimensional structure of MaC-DRS can be identified if we restrict ourselves exclusively to a data-driven approach. However, there is no question that theory-based insights are also needed and, as we have seen from our theoretical chapter, we have well-founded assumptions for the 8-dimensional MaC-DRS structure. We therefore supplement the EFA results for the Egyptian sample with the findings from the CFA to reach more clarity. See for the pan-cultural EFA and the CFA findings the main text (**Chapter 3**).

MaC-DRS Reliability Scores Across Cultures

Table 112: Reliability scores across sample: MaC-DRS short (24 item) version

MaC-DRS (short version) Factors of First-Order	Pan-cultural	GER-Sample	JP-Sample	US-Sample	EG-Sample
	(N = 2,982)	(n = 751)	(n = 740)	(n = 745)	(n = 746)
Fairness	0.9508 (0.9505)	0.9531 (0.9529)	0.9022 (0.9020)	0.9238 (0.9224)	0.9554 (0.9550)
Trustworthiness	0.9656 (0.9654)	0.9609 (0.9607)	0.8979 (0.8935)	0.9463 (0.9462)	0.9816 (0.9816)
Property	0.9735 (0.9735)	0.9630 (0.9627)	0.9330 (0.9327)	0.9620 (0.9620)	0.9826 (0.9826)
Reciprocity	0.9296 (0.9295)	0.9143 (0.9138)	0.8604 (0.8590)	0.8960 (0.8960)	0.9489 (0.9487)
Heroism	0.9448 (0.9435)	0.9030 (0.8970)	0.9302 (0.9295)	0.9282 (0.9263)	0.9579 (0.9570)
Family	0.9528 (0.9527)	0.9238 (0.9237)	0.9051 (0.9050)	0.9373 (0.9372)	0.9757 (0.9757)
In-Group	0.9289 (0.9277)	0.9157 (0.9149)	0.9297 (0.9285)	0.9025 (0.9005)	0.9330 (0.9313)
Deference	0.8565 (0.8563)	0.8708 (0.8698)	0.8217 (0.8196)	0.8511 (0.8507)	0.8595 (0.8566)

Note: The table shows McDonald's Omega and Cronbach's Alpha for each moral domain of the short MaC-DRS version (3 items per moral domain). The scores for Cronbach's Alpha are displayed in brackets.

As can be seen in ***Table 112***, the short version of MaC-DRS performs very well in terms of reliability in the four cultural samples. The lowest reliability values are consistently found for the deference domain — the lowest (McDonald's Omega) value overall is 0.8217, which is still good. Hence, the reliability values in all samples are exclusively in the range between good and excellent.

The **MaC-DRS reliability scores for the long version** are as follows (the first values correspond to McDonald's Omega followed by Cronbach's Alpha):

- **Fairness**: 0.9447 / 0.9435 (pan-cultural-sample); 0.9473 / 0.9458 (GER-sample); 0.9007 / 0.8986 (JP-sample); 0.9135 /0.9109 (US-sample); 0.9466 /0.9446 (EG-sample);

- **Trustworthiness**: 0.9762 / 0.9761 (pan-cultural-sample); 0.9698 / 0.9698 (GER-sample); 0.9300 / 0.9286 (JP-sample); 0.9628 /0.9627 (US-sample); 0.9875 /0.9875 (EG-sample);

- **Reciprocity**: 0.9459 / 0.9459 (pan-cultural-sample); 0.9254 / 0.9252 (GER-sample); 0.8843 / 0.8843 (JP-sample); 0.9212 / 0.9211 (US-sample); 0.9644 /0.9643 (EG-sample);

- **Deference**: 0.8863 / 0.8861 (pan-cultural-sample); 0.8816 / 0.8792 (GER-sample); 0.8000 / 0.7916 (JP-sample); 0.8990 /0.8989 (US-sample); 0.8988 /0.9035 (EG-sample);

- **Property**: 0.9763 / 0.9763 (pan-cultural-sample); 0.9677 / 0.9676 (GER-sample); 0.9414 / 0.9412 (JP-sample); 0.9686 /0.9686 (US-sample); 0.9814 /0.9813 (EG-sample);

- **Heroism**: 0.9595 / 0.9586 (pan-cultural-sample); 0.9286 / 0.9242 (GER-sample); 0.9442 / 0.9432 (JP-sample); 0.9482 /0.9477 (US-sample); 0.9684 /0.9685 (EG-sample);

- **Family**: 0.9496 / 0.9493 (pan-cultural-sample); 0.9042 / 0.9022 (GER-sample); 0.8964 / 0.8940 (JP-sample); 0.9420/ 0.9419 (US-sample); 0.9744 /0.9743 (EG-sample);

- **In-Group**: 0.9268 7 0.9246 (pan-cultural-sample); 0.9021 / 0.8991 (GER-sample); 0.8993 / 0.8942 (JP-sample); 0.8962 /0.8912 (US-sample); 0.94550.9448 (EG-sample).

Correlative Analysis Study 3

Table 113: MaC-DRS (short version) Pan-cultural sample: Correlation between moral domains

Factors of first-order	Fairness *	Trustworthiness	Property	Reciprocity	Heroism	Family	In-Group	Deference
Fairness **	1.0000							
Trustworthiness	0.8129	1.0000						
Property	0.7894	0.8689	1.0000					
Reciprocity	0.7118	0.8332	0.8122	1.0000				
Heroism	0.6827	0.7732	0.8220	0.7897	1.0000			
Family	0.6669	0.7670	0.8163	0.7832	0.8455	1.0000		
In-Group	0.5597	0.6658	0.6917	0.7319	0.7746	0.7997	1.0000	
Deference	0.5371	0.6376	0.6606	0.7190	0.6977	0.7060	0.7297	1.0000

Factors of first-, and second-order	Individualizing	Binding	Reciprocity	Heroism
Individualizing	1.0000			
Binding	0.7855	1.0000		
Reciprocity	0.8368	0.8193	1.0000	
Heroism	0.8090	0.8525	0.7897	1.0000

* **Note:** The table is based on the 24-item **short MaC-DRS version** and shows the correlative results for the **Pan-cultural sample** (N = 2.982). **Note:** All correlations are highly significant (0.0000) and positive. This applies to correlative analyses of first - and second-order factors. ** **Note:** Blue colored correlations mark domains of *individualizing* morality; Yellow colored correlations mark the moral domains of *reciprocity* and *heroism*; Red colored correlations mark domains of *binding* morality.

Supplement Chapter 4

Descriptive Insights *full sample* Study 3

In the following cross-cultural investigations, we focus on the findings on morality in the four cultural groups, i.e., Egypt (EG-sample), Germany (GER-sample), Japan (JP-sample), and the United States of America (US-sample). We begin with a brief introduction that highlights the descriptive insights (*full sample*) of the four groups in our cross-cultural study. These insights are then taken up again at individual points in the text. Subsequently, we will turn to the investigation and discourse of response styles. However, we will first briefly look at the characteristics of the four samples in a descriptive way.

Altogether, we were able to gather responses from a total of N = 2,982 cases in Study 3. These cases distribute as follows (**Table 114**) on the four cultural groups of this study.

Table 114: Case distribution on cultural groups in the full sample

	Freq.	Percent	Cum.
Germany	751	25.18	25.18
Japan	740	24.82	50.00
USA	745	24.98	74.98
Egypt	746	25.02	100.00
Total	2,982	100.00	

Let us now look at the *age* of the respondents in these samples. As was to be expected and as discussed elsewhere,[166] the variable *age* shows differences across the four samples that we study. The *age* distribution across the samples is as follows: GER-sample, Ø = 50.776 (median = 52); JP-sample, Ø = 50.598 (median = 50.5); US-sample, Ø = 47.114 (median = 47); EG-sample, Ø = 37.387 (median = 35).

Regarding the *gender* of the respondents, we aimed for an equal female/male distribution and built appropriate sample restrictions into the online data collection to ensure that we obtain this equal distribution. Consequently, it is hardly surprising that we can also identify an approximately equal distribution of *gender* across the sample groups, as can be seen in **Table 115** below.

---

[166] See: the reflections on the age item in the **Appendix**.

Table 115: Case distribution across gender and cultural groups (full sample)

Gender	GER-sample	JP-sample	US-sample	EG-sample	Total
Male	373	371	368	372	1,484
Female	375	367	374	374	1,490
Non-Binary	3	2	3	0	8
Total	751	740	745	746	2,982

As far as the variable *place of upbringing* is concerned, it can be seen that 32.09% of our total sample grew up in a village and 67.91% in a city. The *full sample* therefore comprises predominantly urban socialized city dwellers. For the individual samples, however, there are differences in the variable, as can be seen from the following list of respondents who grew up in a village (on the country side): *place of upbringing* GER-sample, village = 37.82%; *place of upbringing* JP-sample, village = 43.11%; *place of upbringing* US-sample, village = 30.20%; *place of upbringing* EG-sample, village = 17.29%. In particular, the EG-sample proves to be predominantly composed of people who grew up in an urban area.

Regarding the variable *place of living*, we aimed for a distribution of 70% city-dwellers and 30% village-dwellers. This distribution was achieved for three of the four cultural groups: *place of living* GER-sample, city = 69.91%; *place of living* JP-sample, city = 69.32%; *place of living* US-sample, city = 69.93%; *place of living* EG-sample, city = 89.28%. Although a general trend towards moving to a city in the course of one's life can be observed in the entire sample, it can also be seen that the EG-sample stands out here once again and is primarily an urban sample.

In the case of variable *education*, a fairly broad distribution can be demonstrated across three of the four cultural groups. However, the EG-sample again proves to be an outlier among our samples with a strong bias towards higher education. This finding can be inferred from ***Table 116***, to be found on the next page.

Table 116: Case distribution across education and cultural groups (full sample)

Education (ISCED)	GER-sample	JP-sample	US-sample	EG-sample	Total
Primary education	8	0	130	0	138
Lower secondary education	235	22	43	7	306
Upper secondary education	90	237	148	78	553
Post-secondary/non-tertiary education	137	6	59	36	238
Short cycle tertiary education	56	118	46	0	220
Bachelor´s or equivalent	88	165	167	575	995
Master´s or equivalent	113	179	59	38	389
Doctoral or equivalent	24	13	21	9	58
No response	0	0	67	12	79
Total	751	740	739	746	2,976*

* **Note:** 6 cases are not listed for they responded with ISCED category 0, i.e., "Early childhood education / no education".

The trend of the EG-sample in the *education segment* can also be underpinned by looking at the variable *years in school*. Here too, the EG-sample represents an upwardly biased outlier in the sample comparison: GER-sample, Ø = 11.350 (median = 12); JP-sample, Ø = 12.877 (median = 12); US-sample, Ø = 11.578 (median = 12); EC-sample, Ø = 14.266 (median = 15).

As far as the *net income* is concerned, a not insignificant proportion of respondents stated that they either did not receive any income on their own or did not want to answer this question (***Table 117***).

Table 117: Case distribution across net earnings filter variable and cultural groups (full sample)

Net earnings filter variable	GER-sample	JP-sample	US-sample	EG-sample	Total
Net income indicated	522	381*	356	538	1,797
No own income / No response	229	359	389	208	1,185
Total	751	740	745	746	2,982

* **Note:** 3 cases in the JP-sample and 1 case in the US indicated unrealistic high *net earnings* (above several trillions in Yen/ US Dollar) and are therefore coded as no response.

When we now look at the respective values per sample for those cases that stated their *net earnings*, we obtain the following results: GER-sample, Ø = 2434.089 (median = 2200 Euro); JP-sample, Ø = 278566.6 (median = 200000 Yen); US-sample, Ø = 16490.77 (median = 3000 US Dollar); EG-sample, Ø = 11231.21 (median = 8000 Egyptian Pound). Converted into US dollars ($) (as of 07.08.2024), our samples have the following average monthly net earnings: GER-sample ≈2658 $; JP-sample ≈1895 $; EG-sample ≈228 $. The US-sample of course remains at Ø ≈16490 $, but it should be noted here that the average is not robust against outliers. In the US-sample, we find a total of n = 79 respondents who reported a net income of over 10,000 US dollars per month. If the average net income for the US-sample is calculated without these n = 79 cases, the result for the US-sample is a net income of: Ø = 2875,785 (median = 2000 US dollars).

Regarding the variables *denomination* and *level of religiosity*, the following distributions result. The German sample is mainly Christian and includes 24.77% "Roman Catholic Church", 20.91% "Protestant Church", and 2.13% "Protestant/Evangelical free Church", 1.60% "Orthodox Church" and 0.93% of respondents that chose "Another Christian religious community". However, also a large share (i.e., 41.68%) of the sample responded with "No religion or denomination", while 3.06% decided for the response category "Can´t choose / Not specified". The remaining percent is distributed among other religious groups. Overall, the 7-point scale *level of religiosity* (higher values indicate a higher *level of religiosity*) is not very pronounced in this sample with an average of Ø = 2.926 (median = 3).[167] When asked about denomination, the Japanese sample mainly responded with "No religion or denomination" (51.22%) and also 13.78% decided not to answer the question (i.e., response category "Can´t choose / Not specified"). The largest religious group in our Japanese sample is "Buddhism": 28.65% of the respondents in the JP-sample identified themselves as Buddhists. The next largest religious group in the JP-sample is "Shinto", with 2.97%. All other response options are marginal. As far as the variable *level of religiosity* is concerned, the JP-sample, with an average value of Ø = 2.314 (median = 2), has the lowest value in the comparison of the four samples. The US-sample is also mainly Christian. In total the "Roman Catholic Church" 16.78%, "Protestant Church 17.85%", "Protestant/Evangelical free Church" 6.58%, "Orthodox Church" 1.21%, and the category "Another Christian religious community" 16.01% sum up to 58.39% of the US-Sample. Here, too, the remaining percentages are scattered among the other response options, with the largest remaining share being accounted for by "No religion or

---

[167] Note: We have coded cases that indicated "Can't choose / Not specified" with 0 in order not to lose any cases and to maintain the continuous structure of the variable.

denomination" at 25.64% and "Can´t choose / Not specified" at 7.38%. Overall, US-sample shows itself to be rather religious (variable *level of religiosity* Ø = 4.363; median = 5). In line with the dominant religion in Egypt we find in the EG-sample mainly the response option "Sunni Islam (Sunni)" (85.92%) selected. The response option "Shiite Islam (Shiite)" accounts for only 0.80% while the option "Another Islam religious community" was chosen by 4.56% of the EG-sample. Again, the remaining percentages are distributed among the other categories of the *denomination* item and represent only marginal groups with the option "Can´t choose / Not specified" holding the largest share (2.68%). Finally, the highest *level of religiosity* is found in the Egyptian sample (Ø = 5.252; median = 5) in a comparison of the four groups examined.

Taken together, the descriptive findings of the four samples reveal the following: Firstly, there is a good distribution across *age* and *education*, as well as the desired distribution for the variables *gender* and *village/city* in three of the four samples. However, the Egyptian sample is an outlier within the four samples in terms of the *education* and *village/city* variables. Consequently, we must assume that there is some degree of sample bias (He & van de Vijver, 2012), as it is difficult to compare the EG-sample with the other three groups in our study without controlling for the differences mentioned. For example, if we would not control for the *education* variable, we would have at least a partial comparison of non-comparable cases (see the following authors who discuss this issue in detail in the context of causal analysis and selection bias: Morgan & Winship, 2015). In addition, it should be noted that unobserved heterogeneity may also be associated with sampling bias. This means that unmeasured variables that are partly related to education, for example, may also contribute to further unobserved group differences and complicate the intended cultural comparison. The findings mentioned here must therefore be taken into account in the models of our analysis by including suitable covariates. However, there will still remain some uncertainty about the extent to which the GER-, JP- and US-samples are comparable with the EG-sample, which is due to the likelihood of unobserved heterogeneity. Thus, a first limitation for our further investigations already arises at this point. Furthermore, as far as *income* is concerned, the differences between groups are roughly as expected, and this is also the case in the context of the variables *level of religiosity* and *denomination*.

Although none of our samples are representative of the respective society, three out of four groups in the study show a good distribution of sample characteristics. The descriptive findings suggest however also that the results for the Egyptian sample in particular should be difficult to transfer to Egyptian society. Generalizable statements based on the Egyptian sample should therefore be avoided. The substantial results of our investigations must therefore be

evaluated against the background of the descriptive analyses and the first limitations presented. Starting from these descriptive findings, we now turn to culturally contingent ways of communicating.

Data Collection 3: Response Style Biases — Ways to Communicate *and* Nuisance in the Data (*full sample* Analyses)

So called *response styles* can be treated from two angles. One angle is emphasizing group-specific ways of responding that is irrespective of content as *nuisance* in the data. This perspective teaches us to interpret data affected by response style (bias) with caution. The reason for this is, that response styles are potentially distorting cross group (cultural) comparisons and may reveal no real group differences but display instead potentially over-/under estimated or even false non-/significant results. However, the other angle treats response styles not as a form of bias yet as a reflection of culturally pursuit *ways of communication*. Hence, response styles may be seen to hold a substantive part, i.e., a way of cultural communication, and a methodological part, i.e., nuisance in data (Smith, 2004; van de Vijver & Leung, 2011; He & van de Vijver, 2012; He et al., 2021). Both must be taken into account when approaching the world empirically and dealing with questionnaires and cross-group comparisons. As indicated before, we examine four cultural groups — i.e., Germany (GER-sample), Japan (JP-sample), United States of America (US-sample), and Egypt (EG-sample) — in our study. Before turning to the substantive analysis of moral deviance relevance in the four cultures below, we first discuss, based on empirical data, response styles as a form of cultural communication and as a source of bias in cross-cultural comparisons.

In comparing the data from the four cultural groups examined we found relevant cross-group differences, that are in parts attributable to culture specific ways of responding. Let us first discuss this by means of an introductory example. For this purpose, we use the MaC-DRS item "*Someone favors themselves in the distribution of resources*" (fairness deviance relevance item) and form histograms for this item across the four cultural groups of our study.

Figure 31: Histogram of MaC-DRS fairness deviance relevance item: Example for the use of the response format/ density within cultural groups
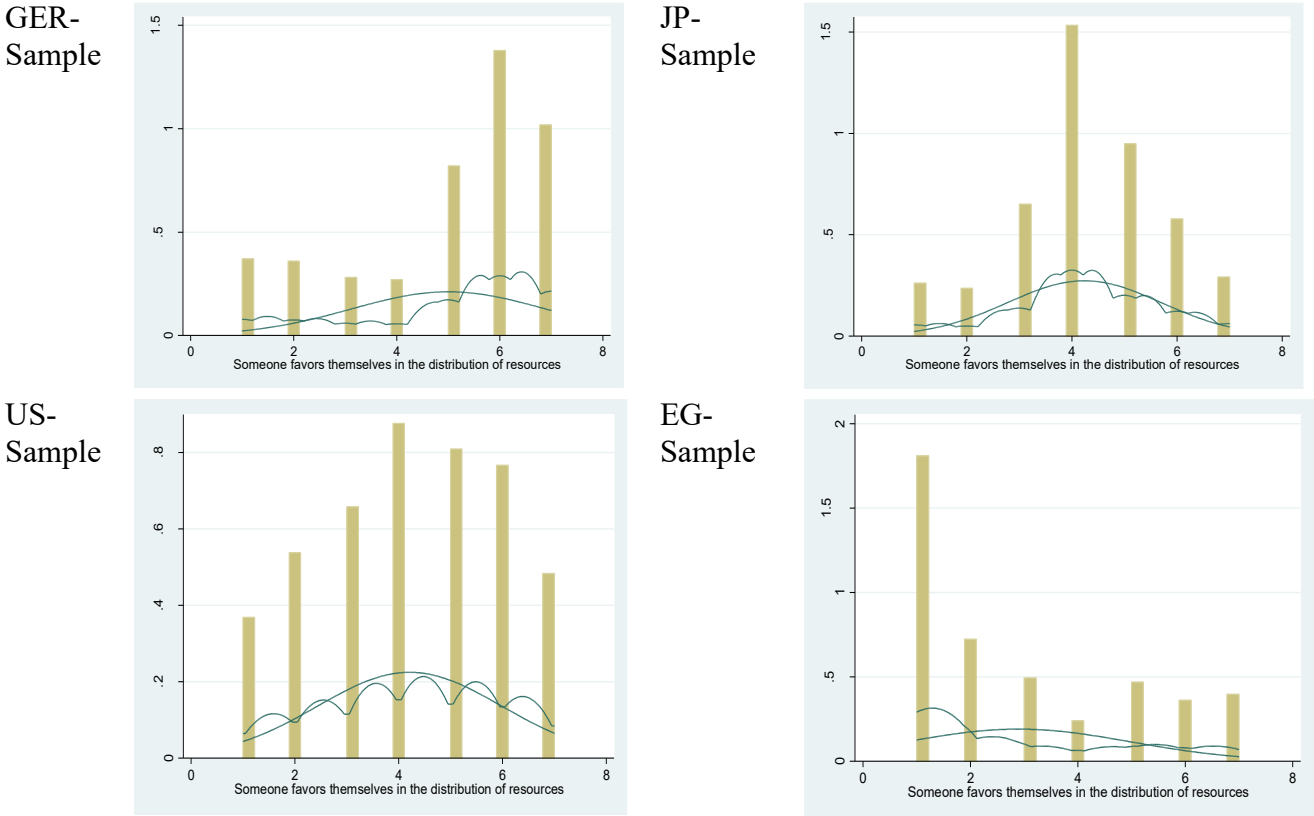
Note that MaC-DRS comprises a 7-point response format that goes from 1 "Extremely irrelevant" — having 4 "Neither irrelevant nor relevant" as a neutral response option and midpoint — to 7 "Extremely relevant". All intermediate response steps are also labeled. The example histograms demonstrate the point that we would like to make quite clearly: the GER-sample tends towards relevance as does the US-sample in general too. Both, the GER- and US-Sample also show a fairly well distributed pattern cross the response format, without clear indication by face validity to have a tendency of favoring one response category extremely more than others. In contrast, the JP-Sample markedly shows a tendency for the use of the middle-category, as evident from the high bar on the 4 ("Neither irrelevant nor relevant"). Furthermore, most cases in the EG-Sample are grouped on the extreme end-point of 1 on the response-format. The latter is indicative of an extreme response bias, and by looking at the vast size of respondents who took the 1, we have also a hunch that likely a good number of cases are potentially unwilling to answer the respective item. Taking a look at all MaC-DRS items, we further found that the exemplary pattern shown in the histograms of the GER-, US-, JP- and EG-samples indeed applies across the items of the moral scale, albeit there is also some variation. Thus, across all 32 MaC-DRS items, the GER- and US-samples show a more or less broad respectively not an extreme distribution across the response options. At least these

samples exhibit by face validity no clear favor to use one response option significantly more than others. Opposed to this, the JP-sample tends clearly towards the middle category (the 4), and the EG-sample strongly towards the endpoint, the 1 of the scale. For this reason, we have a first indication to reason that the JP-sample is likely to use a *moderacy response style* (i.e., a preference to use the middle category). The EG-sample is marked by a different pattern and displays an *extreme-response style* (i.e., a preference to use the end point(s) of a scale). If this suspicion is confirmed in the further course, it points to at least two pronounced and different types of cultural communication preferences (Smith, 2004), which may be related to the cultural logics of face (Japan) and honor (Egypt) (Leung & Cohen, 2011; Uskul et al., 2023). Our first indications are also supported by the findings of Minkov (2017), that demonstrate the prominence of a middle category responding in Japan and the reverse tendency for Egypt. The results of other studies thus also indicate that we should take the observed pattern seriously.

Once we had formed the measures for the response styles (i.e., NARS and MRS), we empirically investigated whether our assumption above could be confirmed across the cultural groups in our study. Hence, we conducted a simple one-way ANOVA (Bonferroni corrected) (Völkle & Erdfelder, 2010; Aden et al., 2021) to test for significant differences in *NARS* and *MRS* across the four cultures of our study. The comparison between the GER- and US-sample is insignificant regarding *NARS* yet all other comparisons between cultural groups show highly significant differences (Prob > F 0.000) on the *net acquiescence response style* measure. The direction of this response style is indicated by the mean on the *NARS* measure, which is apparently quite different across the groups of our study: GER-sample, *NARS* = 10.296; JP-sample, *NARS* = 1.206; US-sample, *NARS* = 9.880; EG-sample, *NARS* = -22.819. Next to a moderate positive acquiescence response style in the GER-/US-sample we find a pronounced negative effect in the EG-Sample. The low *NARS* mean value found for the JP-sample is moreover already indicative of the response style for this cultural sample. We also estimated an ANVOA (Bonferroni-corrected) for *MRS* across the four cultural groups. Again, we found highly significant effects this time indicating cross-cultural differences in *middle category response style* (Prob > F 0.000). However, this statement excludes the US American and the German sample, which show no significant differences. The findings show that the tendency to choose the middle category across the set of 60 items used for our measure is significantly higher (Prob > F 0.000) in the JP-sample compared to all other groups. Furthermore, as was already evident from the values of the *NARS* measure, the *MRS* effect estimated by ANOVA is

significantly lower (Prob > F 0.000) in the Egyptian sample compared to the other groups under investigation.[168]

In the light of our response style findings, it becomes evident that the **Japanese sample** tends to avoid potentially salient responses. This communication style is in line with interdependent self-construal (Markus & Kitayama, 1991; 2010) and fits well with a cultural logic in which it is important to mutually save face (Leung & Cohen, 2011). Both of these cultural dimensions are characteristic of Japan. In addition, our results for response style in the JP-sample are exactly in line with the theoretical considerations and evidence for the holistic system of thought prevalent in Japan. Holistic cognition is characterized, among others, by "a search for the "Middle Way" between opposing propositions" (Nisbett et al., 2001, p. 993). In the **Egyptian** honor context, it is important to note that honor is a contested attribute, that can be potentially lost (Uskul et al., 2019). So, perhaps more extreme responses in the Egyptian sample could be understood as a mode of self-assertion (San Martin et al., 2018) that functions to protect one's honor. The **GER-** and **US-samples** are found in the context of response styles between the aforementioned samples and reflect both styles, albeit not as pronounced. This may be part of communication in cultural entities that are characterized by dignity, individualism and independence in selfhood. Dignity cultures treat dignity as an inherent aspect of every person. Based on this notion, it is also important to treat each other with dignity, a logic that is akin to saving face. But that doesn't stop people from wanting to stand out and express their own opinions. The latter, in turn, is a characteristic of individualistic cultures and independence in selfhood, which for their turn are characteristic for Germany and the USA (Triandis, 2001; Kitayama et al., 2009; Cross et al., 2011; Kitayama & Salvador, 2024).

Although we believe that the results on response styles are theoretically plausible in many respects, we would like to make one further comment before we continue and refer once again to our data basis. We wonder if it is only the response styles that influence the response distribution for the MaC-DRS items discussed above? We have shown above as an example the distribution of responses to a MaC-DRS item as histograms. In this context, we continued to find clear differences between the study groups, which we could now also find in the 60 items comprising different measures of response styles (*NARS* and *MRS*). However, since the MaC-DRS response patterns of the Egyptian sample differ vastly from those of the other groups, we

---

[168] The MRS mean values across the four cultural groups in our study are as follows: GER-sample, *MRS* = 10.555; US-sample, *MRS* = 9.899; JP-sample, *MRS* = 16.391; EG-sample, *MRS* = 5.941. Note: our findings on *MRS* are also in line with the results on middle category responding found by Minkov (2017).

can well imagine that other additional and unobserved factors may also play a role.[169] We have already identified sample bias — the EG-sample differs from the other groups on several measured socio-demographic variables — and we have now also identified a clearly pronounced extreme response bias for this sample. These circumstances cast doubt on the robustness of the data collected, which form the basis for the Egyptian sample. It is therefore evident that we should be extremely cautious with the results we obtain from the analyses of the Egyptian sample. Generalizations based on our Egyptian sample are likely to be difficult to draw, and the interpretations of the data that we will present in the course of our study must be seen in the context of the shortcomings of the EG-sample and our data collection in general. Therefore, even before further analyses are introduced in the present study, there is a call for future studies based on a better database to conduct more valid analyses of Egyptian society in relation to morality.

Taken the insights gained and our notions together we overall conclude that the four groups of our study are pursuing likely different ways of communication — at least in respect to issues related to moral deviance. This is reflected in significant differences on *NARS* and *MRS* across cultural entities examined. Based on this finding, the suspicion obtained by analyzing the response distributions across MaC-DRS items receives support. Accordingly, we are of the opinion that is advisable to include *NARS* and *MRS* as covariates in respective models when analyzing the four cultural groups with regard to the substantive questions of our study.

---

[169] Our doubts are partly due to the fact that we conducted additional robustness checks of response behavior in which we designed *MRS* and *NARS* measures without the inclusion of the MaC-DRS items. The results of these analyses differ from those we have presented here, particularly for the EG-sample but also for the JP-sample. This fact suggests that there is likely not only a difference in response style between the groups, but that the EG-sample in particular gives extreme responses to the MaC-DRS items due to factors that we were unable to identify. Thus, it seems that there is not only a content-independent communication style, but also a group-specific interaction with the content of the items, i.e., with the topic of moral deviance. This also clearly underlines that particularly the EG-sample results presented in this paper should be treated with caution. In addition, cognitive interviews using the MaC-DRS items with individuals from different social classes in Egyptian society would provide further insights into the issues outlined here. Unfortunately, this was not feasible in the context of our study after data collection and remains a task for future research. Since these additional response style measures used for the sensitivity analysis are heavily dominated by the SCS-CIRN-3 items (Vignoles et al., 2016; Yang, 2018; Uskul et al., 2023) due to the exclusion of the MaC-DRS items, we also refrained from conducting in-depth analyses with these measures. More than four-fifths of the items used for these response style measures come from the self-construal scale, which is the only longer questionnaire battery included in the response style measures constructed for the sensitivity analysis. We consider this to be problematic. A detailed analysis of these measures would lead to new justification problems for these additional measures themselves and further robustness analyses would also have to follow. Overall, we thus conclude at this point solely that, in addition to differences in response style, there are likely other, unobserved differences between the cultural samples in this study that are related to the measurement of moral deviance. Also, particularly the EG-sample seems to be affected by this unobserved heterogeneity. Further research is hence clearly needed to address these issues.

Sample Adjustment

Apart from culturally different ways of communicating, the histograms shown above also suggest the possibility that the data may contain more than mere differences in response style. We have a hunch and suspect that some cases in our samples may have *poor data quality.* In the following, we will explain what we mean by poor data quality. In this regard, we will discuss two criteria for assessing poor data quality, namely *a)* consistent responses (i.e., always the same responses despite different content of items), and *b)* the completion time of the questionnaire. On this basis, we will identify cases for which poor data quality can be affirmed. To anticipate this point, we will also argue that cases characterized by poor data quality should be removed from the sample to gain more valid data for subsequent analyses based on an *adjusted sample*.

After discussing cases with potentially poor data quality, we then turn to the possibility of further reducing the number of cases in our sample by excluding *cases with multiple citizenships* from substantive cross-cultural comparisons. Here we present a socialization argument that we put forward to justify this step. Eventually, after these mainly methodological and only partially theoretical efforts, we turn to the substantive analysis and discussion of morality and culture.

Data Quality

Before proceeding to the substantive part, we would first and foremost like to take a closer look at data quality. In order to be able to evaluate the quality of the data at least partially, we tested whether a part of the respondents always, i.e., consistently and without consideration of the content, indicated one and the same response over larger item batteries. We have two longer item batteries that lend themselves to such a test, namely MaC-DRS (the scale comprises 32 items) and CIRN-SCS-3 (the scale comprises 48 items, 23 of which are positively polarized and 25 of which are negatively polarized). If we find consistent response patterns on the items of either these two scales, we regard the corresponding data to be of poor quality. Furthermore, we decide to remove cases with poor quality from the sample in order to be able to perform more valid analyses of substance. In other words: We treat cases that produce one and the same response pattern exclusively and independently of content as confounding factors in substantive analyses, and control them accordingly by removing them from the samples.

First, we inspect the 32 MaC-DRS items for consistent responses:

- n = 120 cases consistently indicated 1 (Extremely irrelevant) for all 32 MaC-DRS items.
- n = 63 cases consistently indicated 4 (Neither irrelevant nor relevant) for all 32 MaC-DRS items.
- n = 48 cases consistently indicated 7 (Extremely relevant) for all 32 MaC-DRS items.

Based on these results, we removed from the samples a total of n = 231 cases that consistently indicated a single response across all 32 MaC-DRS items.

After excluding these cases, we examined thereafter the CIRN-SCS-3 items for consistent responses:

- We found a total of n = 15 remaining cases that consistently (i.e., exclusively) gave one and the same answer to the item of the self-construal scale. On the one hand these cases responded either 5 (midpoint of the scale) to all 48 items of the self-construal scale. Or, on the other hand, these cases responded consistently either 1 (lower endpoint) or 9 (upper endpoint) to the 23 positively polarized SCS items respectively to the 25 negatively polarized SCS items.

Consequently, we classify these n = 15 cases as cases with poor data quality and remove them accordingly from the sample. After excluding these cases too, the following case distribution results for the four expressions of the variable *culture*.

Table 118: Case distribution on cultural groups in the corrected sample

	Freq.	Percent	Cum.
Germany	733	26.79	26.79
Japan	671	24.52	51.32
USA	692	25.29	76.61
Egypt	640	23.39	100.00
Total	2,736	100.00	

Based on the analysis of consistent response patterns in two longer measurement instruments of our questionnaire (32 items and 23/25 items), it can be concluded that we find cases with poor data quality in all four cultural samples. However, the samples differ significantly in terms of the number of cases with poor data quality: the German sample has (GER) = 18, the Japanese sample has (JP) = 69, the US American sample has (US) = 53 and the Egyptian sample has (EG) = 106 cases of poor data quality. These cases are characterized by the circumstance that they apparently only clicked through the online questionnaire without paying much attention to the content of the items. In order to conduct more valid analyses, we removed a total of n = 246 cases from the overall sample as described.

Questionnaire Completion Time

Let us now consider the completion time for answering all items of our questionnaire as another criterion to assess data quality. As a reminder, our survey design includes 111 items, 9 of which are text-based scenarios of moral dilemmas, that are together somewhat more reading-intensive. Furthermore, our research design includes also a factorial survey and all respondents saw 4 vignettes with 4 corresponding items (relevance, judgment, shame and guilt) per vignette. Against this background, we now use the questionnaire completion time as a further criterion to exclude cases characterized by poor data quality.

We conducted a pilot study before the actual field phase of our cross-cultural study. In addition to the technical testing of our questionnaire, the aim of this pilot study was to assess the questionnaire completion time, i.e., the average time required to fill out the entire questionnaire. The pilot study revealed ≈ 23 minutes as average questionnaire completion time.

In the following, we take a closer look at the questionnaire completion time across the four cultural groups of the present study. To do this, we first compare the full sample (N = 2982) with the sample corrected for poor data quality (consistent response patterns; N = 2736), and the four cultural groups of these samples with each other (***Table 119***). In doing so, we only look at the median (in seconds), as this is robust against outliers.

Table 119: Questionnaire Completion time across cultural groups

	GER	JP	USA	EG
Full Sample Completion Time (median)	$\approx$ 1450 sec.	$\approx$ 1138 sec.	$\approx$ 1320 sec.	$\approx$ 1270 sec.
Corrected Sample Completion Time (median)	$\approx$ 1460 sec.	$\approx$ 1163 sec.	$\approx$ 1377 sec.	$\approx$ 1300 sec.

Table 93 shows that in the *full sample*, the German cultural group spent the most time answering our items, taking approximately 24 minutes to fill out the entire questionnaire. This sample is followed by the US- and EG-sample. The JP-sample took the least time to complete the entire questionnaire, namely just under 19 minutes. Furthermore, ***Table 119*** shows that the **correction** for cases with poor data quality increases the questionnaire completion time in each sample. This fact supports that the cases excluded in the process described above are marked by poor data quality and merely clicked through the questionnaire without paying attention to the content of the items.

In the next step, we focus on the sample that has already been partially corrected. We now take a closer look at the four cultural groups in our study and inspect how many cases completed the entire questionnaire in ≤ 11 to ≤ 15 minutes, at intervals of 2 minutes (**Table 120**). Through these intervals in questionnaire completion time, we want to find out how many cases would be excluded if these minimum requirements for questionnaire completion time were to be applied. We start at 11 minutes, as the company we employed to collect data set a minimum completion time of 10 minutes for the entire questionnaire. The company did not allow a longer minimum completion time, regrettably. After all, the following comparison should lead to the exclusion of cases with a very short questionnaire completion time. Simultaneously, however, the number of cases in the cultural samples should be taken into account in order to avoid to run into problems with the statistical power due to possible case exclusions. At this point, it should be noted that a priori power analyses for our study resulted in a minimum sample size per cultural group of n = 525.[170] This sample size should also be considered against the background of an additional restriction of the data set, which is also associated with the loss of cases and is discussed further below.

Table 120: Corrected Sample: Questionnaire completion time and number of observations across cultural samples

Cultural samples	GER	JP	USA	EG	Total number of obs.
Sample n	n = 733	n = 671	n = 692	n = 640	N = 2736
Obs. completion time ≤ 11 min.	1	50	35	17	103
Obs. completion time $\leq 12{:}50$ min.*	1	115	77	47	240
Obs. completion time ≤ 13 min.	1	136	94	58	289
Obs. completion time ≤ 15 min.	43	198	145	121	507

* Note: We have added another time restriction: *12 minutes and 50 seconds*. This (minimal) time restriction for completing the questionnaire was added because we are also constraining the sample later on to include exclusively cases with solely the citizenship of the countries under investigation. In this sense, the restriction to 12 minutes and 50 seconds (minimal completion time) includes sufficiently many cases so that no problems with statistical power arise, but at the same time excludes cases that completed the questionnaire very quickly. More on this topic follows further below.

---

[170] A priori power analysis result based on Bonferroni corrected alpha-level: critical t value = 2.397, DF = 1048, n = 525 (per target group), n = 1050 (for country wise t-test comparison), N = 2100 (for four target countries) and actual power = 0.80. Note: a power of 0.80 is considered the minimum requirement for sufficient statistical power.

Considering the empirical completion time of the questionnaire in our study as another criterion of data quality, we come to the following conclusions:

- Overall, in regard to the questionnaire completion time, a look at the results presented shows that the data quality could be improved in three out of four cultural samples.
- The German sample appears again of good quality. In the GER-sample only a few cases show an unrealistically short questionnaire completion time.
- However, the same does not apply to the JP-sample. This sample has the highest number of cases that completed the entire study very quickly. For example, we would exclude n = 289 cases if we set a minimum completion time of 13 minutes. Note: this minimum completion time would still be more than 10 minutes shorter than the questionnaire completion time we found in our pilot study.
- Although to a lesser extent than in the JP-sample, there is also a not insignificant number of respondents in the EG- and US-sample who completed the questionnaire relatively quickly.
- Considering the questionnaire completion time, we should nevertheless take also a mitigating circumstance into account. The participants in our study were drawn from a pool of respondents from the company we commissioned. These respondents, on average, can be expected to have a very high level of instrument familiarity (He & van de Vijver, 2012), for some of them likely participated in other survey-based studies before. The high level of instrument familiarity is most likely also a reason for the relatively short questionnaire completion time of some cases.
- Next to a very short completion times, we found also n = 248 cases that took ≥ 50 minutes to complete the entire questionnaire. In this case, it could be assumed that the (online) completion of the questionnaire was possibly interrupted and continued at a later time. We consider the long completion times not as particularly relevant in the context of the assessment of data quality. For this reason, we will not go into detail about cases with particularly long completion times.

Taken together, our analysis of the questionnaire completion time shows that the JP- but also the EG- and US-sample contain a non-insignificant number of cases that completed the questionnaire of our study very, and in parts unrealistically quickly. This is first of all a general fact that needs to be taken into account in the context of the quality of our data. Moreover, from how we read the data the analysis suggests that as many cases as possible with an unrealistically short completion time of the questionnaire should be excluded. In this context, however, considerations of statistical power must be taken into account. As we are in the next step introducing a further important restriction of the sample, we therefore are postponing the final decision on the exclusion of cases based on the completion time of the questionnaire.

Sample Adjustment Based on a Cultural Socialization Argument: Exclusion of Cases of Multiple Citizenship

Before we apply the questionnaire-completion-time-correction and turn thereafter to the *adjusted sample*, we would like to discuss a further sample exclusion criterion. This criterion is not, as before, methodological, but exclusively theory-based and content-related. The fact that we want to conduct cross-cultural analyses means that ideally, we should only examine cases of individuals who have been socialized in the cultural mainstream of our study countries (Esser, 2008; Kitayama et al., 2009; Hadjar et al., 2012; Berger & Luckmann, 2013; Windzio & Wingens, 2014; Boehnke & Hadjar, 2015; Teltemann & Windzio, 2016; Henrich, 2020; Kühnen & Kitayama, 2024). This criterion can be measured to some extent indirectly in our data set using the citizenship variable for the respondents and the parents of the respondents. Cases with a (recent) migration background or migration background of the parents — who are also likely to be exposed to other important and influential cultural socialization influences — can be identified using the citizenship variable and filtered out of the data set. In view of this, it seems sensible to us to tailor our sample so that only cases with single German, Japanese, US-American or Egyptian citizenship are represented in the data set.

If we were to apply the exclusion criterion described, our sample would again lose n = 121 cases. The focus of our analyses is on cross-cultural similarities and differences in moral deviance relevance. We are hence explicitly interested in the effect of culture on morality. A prerequisite for our analyses is therefore that the people we study are also socialized in the respective cultures we are comparing. For this reason, we consider it appropriate to apply the exclusion criterion outlined. Accordingly, we removed n = 121 cases with non-exclusive GER-, JP-, US-, or EG-citizenship from the data set. In addition, we checked the citizenship of the respondents' parents, and after excluding the n = 121 cases only the four citizenships of our study groups can be found in these variables. Restricting the samples to be examined overall increases the probability that we will examine people who have been socialized in the cultural mainstream of the corresponding groups under investigation, which is an important aspect of comparative studies aimed at analyzing effects of culture.

The Adjusted Sample

The restriction to focus exclusively on cases of the respective cultural mainstream decreases the initial sample by the described n = 121 cases. If we now apply a strict minimal time criterion for completing the questionnaire, e.g. a minimum of 15 minutes, we quickly run into the problem of insufficient power for some of the cultural sub-samples. As it is nevertheless desirable to exclude as many cases as possible with questionnaires that were completed very quickly from substantive analyses, we have tried to find a compromise. This compromise was found in a minimum completion time of 12 minutes and 50 seconds for the entire study. Applying this restriction equates to the exclusion of n = 240 cases. After excluding cases below a minimum questionnaire completion time of 12 min. and 50 sec., the four cultural sub-samples have sufficient power and additionally some cases with poorer data quality could be removed from the data. By relying on this ***adjusted sample***, we are hence more certain to be able to examine actual cross-cultural commonalities and differences in the analyses to come.

Moral Deviance Relevance from a Cross-Cultural Perspective — MaC-DRS Findings (Full Sample Analyses)

As a result of the case exclusions described, the sample N is reduced by a total of n = 622 cases. However, not only the number of respondents in the samples changes after these reasonable case exclusions. The analyses of the values for the dependent variable of the intuitive deviance relevance of different moral domains (MaC-DRS) are also influenced by the reduction of cases in the sample. In the following, we present this change in *three* parts and explain why we reject the interpretation of the data and the resulting findings of two of these parts. When we arrive at the third part, we find ourselves at the analysis of the *adjusted sample*. The results we draw from the analysis of this sample are, in our opinion, based on the most valid data from our cross-cultural data collection. We have examined and discussed the results of the adjusted sample in detail — for corresponding insights, we refer the reader to the main text (**Chapter 4**). But to better understand the data we have collected and our sample corrections, we present the full sample analyses in the following.

We will ***first*** start with a naive cross-cultural comparison of moral deviance relevance using the *full sample*. We call this model naive for it comes without the consideration of covariates, response style differences and taking into account poor data quality of some cases. In fact, we never know the true value of our analyses (Fischer & Milfont, 2010), which is why

dataset adjustments that exclude cases without justification should be rejected in any case. However, due to the response style differences outlined above, among other things, we consider it justified to reject the results of the naive model. ***Second***, after the first part and the naive cultural comparison, we then briefly turn to further analyses with the *full sample* for the sake of completeness. These analyses will include an appropriate set of covariates, taking into account response styles as a potential bias to be controlled for, in addition to other variables and interaction terms. We argue that the model underlying the second part of analyses is good given what our questionnaire measures. Nevertheless, we suggest to reject the results obtained from this analysis. The background for our rejection is based on the cases which, in our opinion, are characterized by poor data quality. From our perspective, basing the analyses on the full, non-adjusted sample would lead to distorted results. For this reason, we will remain on the surface and refrain from in-depth examination and interpretation of the results obtained using the *full sample*. We nonetheless would like to show these results so that the reader can compare them with the analyses of the *adjusted sample*. The latter is the ***third*** and final step of our analyses and can be found in the main text.

The Naive Comparison (Full Sample)

Let us first take a relatively naive look at the MaC-DRS results. We begin by tabulating (***Table 121***) insights from an OLS regression that sequentially tests the 8 moral domains measured via MaC-DRS for cross-cultural similarities — i.e., non-significant differences in moral deviance relevance —, and differences in moral deviance relevance. The models include the 8 moral domains as consecutively tested dependent variables and *culture* as the only predictor variable. No covariates are specified for this model. Against the background of the descriptive analyses and the findings in the context of response styles, we therefore refer to this model as the naive model. As a reminder, we have shown in another chapter that there is full exact scalar measurement invariance for the first-order factors (i.e., the 8 moral domains) of MaC-DRS, which is why mean comparisons of this variable across the four cultural samples of this study are permissible.[171] The following labeling is used in ***Table 121*** to indicate (Holm-Bonferroni corrected) significance: p <0.05, ** p <0.01, *** p <0.001.[172]

---

[171] In the following analyses, we only use the 24-item short form of MaC-DRS, as both pan-cultural and country-specific CFA models have acceptable goodness of fit values and we were able to demonstrate scalar measurement invariance for this instrument.

[172] With four expressions of the variable culture, a total of 6 pairwise comparisons (i.e., GER vs. JP; GER vs. US; GER vs. EG; JP vs. US; JP vs. EG; US vs. EG) results for each test for significant differences in terms of moral deviance relevance (MaC-DRS, 8 dimensions). We are therefore in the realm of multiple testing due to the 6

Looking briefly at the results in ***Table 121***, it is evident that the EG-sample in particular differs greatly from the other three cultural groups. In our view, this is partly explainable given what we have shown above with regard to the descriptive statistics (*full sample*), the response style bias and the cases of poor data quality. Arguing from this perspective, we therefore distrust the naively estimated results. Thus, for a more valid assessment of cross-cultural similarities and differences in moral deviance relevance, we believe that the inclusion of covariates (e.g. response styles, socio-demographic variables etc.) into our estimation models, and the omission of cases with poor data is essential.

At this point, we would like to emphasize the aspect of covariates once again. We found significant differences in the sample composition between the four cultural groups, as indicated before in the section of the descriptive insights. These differences, in addition to the response style and cases with poor data quality, are likely to have affected the results of the naively estimated MaC-DRS values. In other words: without the inclusion of covariates, we are comparing cases that are actually not comparable and which, apart from the variable *culture* in which we are interested in, differ in terms of other aspects not accounted for in the naive model.

We discuss in the following again three variables as examples to make our point more prominent. Firstly, the variable *level of religiosity*. This variable asks about the subjective level of religious self-description ("Would you describe yourself as..."). The response format ranges from 1 "Extremely non-religious" to 7 "Extremely religious", with all intermediate response-options labeled. A simple (*full sample*) one-way ANOVA (Bonferroni corrected) shows highly significant differences (Prob > F 0.000) on the variable *level of religiosity* between all cultural groups in our study. These differences become also obvious from the group-specific mean values: GER-sample (n = 751) = 2. 926; JP-sample (n = 740) = 2.314; US-sample (n = 745) = 4.363; and EG-sample (n = 746) = 5.252. The respective median of the groups examined shows that the GER-sample responded most frequently with "Somewhat non-religious" and the JP-sample with "Very non-religious". The US- and EG-sample, for their part, responded most frequently with "somewhat religious". Overall, though, the Egyptian group surveyed exhibits a higher average *level of religiosity* than the US-sample.

---

pairwise comparisons. Multiple testing is associated with alpha (Type I) error cumulation, so we should use adjusted significance levels to account for this. The Bonferroni correction is criticized for being too conservative (for it increases the probability of a type II error), whereas the Holm-Bonferroni correction works with sequentially less restrictive significance levels and can therefore be considered a suitable compromise between Type I and Type II errors (Hemmerich, 2020). We apply the Holm-Bonferroni correction in the (6) pairwise comparisons of the margins obtained by the OLS covariate model, i.e., for each of the tests for significant differences with respect to the 8 moral domains. ***Table 121*** therefore displays significance levels obtained by Holm-Bonferroni correction. We calculated the Holm-Bonferroni corrected significance levels using the significance level calculator on the following website: https://statistikguru.de/rechner/adjustierung-des-alphaniveaus.html

Table 121: Naive estimation†: Full sample analysis of deviance relevance across moral domains and across cultural groups

<div align="right">N = 2,982</div>

Germany n = 751	Samples Compared (pairwise) ††	Japan n = 740	Samples Compared (pairwise)	USA n = 745	Samples Compared (pairwise)	Egypt n = 746
5.196 Fairness †††	GER vs. JP *** GER vs. USA *** GER vs. EG ***	4.490 Fairness	*JP vs. USA n.s.* JP vs. EG ***	4.482 Fairness	USA vs. EG ***	2.622 Fairness
5.066 Trustworthiness	*GER vs. JP n.s.* GER vs. USA *** GER vs. EG ***	4.951 Trustworthiness	JP vs. USA *** JP vs. EG ***	4.532 Trustworthiness	USA vs. EG ***	2.538 Trustworthiness
5.235 Property	GER vs. JP * GER vs. USA *** GER vs. EG***	5.036 Property	JP vs. USA *** JP vs. EG ***	4.670 Property	USA vs. EG ***	2.628 Property
4.340 Reciprocity	*GER vs. JP n.s.* GER vs. USA ** GER vs. EG ***	4.439 Reciprocity	JP vs. USA ** JP vs. EG ***	4.077 Reciprocity	USA vs. EG ***	2.635 Reciprocity
4.490 Heroism	*GER vs. JP n.s.* *GER vs. USA n.s.* GER vs. EG ***	4.372 Heroism	*JP vs. USA n.s.* JP vs. EG ***	4.404 Heroism	USA vs. EG ***	2.565 Heroism
4.323 Family	*GER vs. JP n.s.* *GER vs. USA n.s.* GER vs. EG ***	4.260 Family	JP vs. USA * JP vs. EG ***	4.504 Family	USA vs. EG ***	2.613 Family
3.792 In-Group	*GER vs. JP n.s.* *GER vs. USA n.s.* GER vs. EG ***	3.891 In-Group	*JP vs. USA n.s.* JP vs. EG ***	3.936 In-Group	USA vs. EG ***	2.594 In-Group
3.780 Deference	*GER vs. JP n.s.* GER vs. USA** GER vs. EG ***	3.880 Deference	*JP vs. USA n.s.* JP vs. EG ***	4.017 Deference	USA vs. EG ***	2.869 Deference

† **Note:** The naively estimated scores result from previously fit OLS model(s) that include moral deviance relevance as the dependent variable and culture as the only covariate in the model(s). On display are the scores for the cultural groups: Germany (GER), Japan (JP), United States of America (USA) and Egypt (EG); †† **Note:** To test for significant differences in moral deviance relevance, cultural groups were compared in a pairwise fashion. The asterisks indicate (Holm-Bonferroni corrected) significance: * p <0.05, ** p <0.01, *** p <0.001. Non-significant comparisons (*n.s.*) are highlighted in italics; ††† **Note:** colors indicate score ranking in country comparison (Green > Blue > Yellow > Red).

In addition, the values for the subjective level of religiosity are also partly related to different *denominations*, which for their part may offer distinct codifications of morality. As far as the variable *denomination* is concerned, we find in the GER-sample primarily Christians and cases without religious affiliation: n = 378 assign themselves to a Christian denomination and n = 313 have not indicated "No religion or denomination". As already noted, in addition to the respondents without religious affiliation (n = 379), n = 212 in the JP-sample indicated Buddhism as their denomination. In the US-sample, the majority of cases (n = 435) are distributed among Christian religious groups, and in the EG-sample, as expected, there is a predominant distribution among Islamic religious groups (n = 675). In sum, regarding religious denomination, the groups in our study markedly differ.

We attempted to achieve a comparable sample composition with regard to several socio-demographic variables in the different cultural groups of our study by setting eligibility criteria wherever possible and sensible. These criteria included, for example, *gender*, *education*, partly *age* and also the variable *place of living*, for which we aimed for approx. 30% village dwellers and approx. 70% city dwellers in each sample. However, it was not possible during data collection to achieve a reasonably similar sample composition for (e.g.) the variables *education* and *place of residence*. Regarding these variables the EG-sample differs significantly compared to the other groups. Let us briefly recall the distribution on variable *education* (ISCED-2011) by way of example. The EG-sample is dominated by highly educated respondents: 82.17% of the full EG-sample attained a Bachelor's degree (or equivalent) or above. Not only does this characteristic of the EG-sample indicate that this sample is completely non-representative for Egypt. It also stands in stark contrast to the other cultural groups in our study. The latter are much more diverse composed in terms of the variable education. Looking at only the case numbers in percentage for Bachelor's degree or above across the other groups, a very different sample composition with regard to *education* becomes evident in comparison with the EG-sample: GER-sample = 29.97%; JP-sample = 48.25%; US-sample = 33.42%. Thus, unlike the Egyptian sample, there are significantly fewer respondents with higher education in the latter three samples.

Summarizing our considerations and referring to the results of the naive estimation, we come to the following conclusion: With reference to *a)* differences in response style, *b)* differences in data quality, and *c)* differences in sample composition, we argue that the results of the naive estimation should not be interpreted. Based on our explanations, we believe that the naive estimates do not reflect reliable values. Instead, to obtain more meaningful results, it

is important to *account* for response style bias and an appropriate set of covariates, in addition to excluding cases with poor data quality, as we will discuss below.

Moral Deviance Relevance: The Covariate OLS Model (Full Sample)

In the following, we repeat the analyses for the moral deviance relevance ratings (MaC-DRS) of the 8 moral domains across the four cultural groups of our study. We base the analyses again on the *full sample*. However, this time we use an adequate covariate model and briefly inspect the margins (Williams, 2012; Wooldridge, 2016) for moral deviance relevance. Note: margins are statistical values calculated from predictions of a respective previously fit OLS-model at fixed values of some covariates, in our case this will be the variable *culture*, and the integration of the average of the non-fixed covariates of the model (Average Marginal Effect) (StataCorp. 2013). Furthermore, we briefly compare the naive models with the covariate OLS models. However, before doing so we turn to the set of *covariates* and a number of *interaction terms* that we include in each model to estimate cultural deviance relevance (margins).

Moral Deviance Relevance in the *Full Sample*: Illustration and Rejection of Results

In the following we will briefly touch on the moral deviance relevance margins obtained for the four groups of this study. The **Table 122** to be found below displays the deviance relevance margins for 8 moral domains (dependent variables) after estimating OLS regressions that each comprise the specified set of covariates and interaction terms. The analyses are based on the **full sample**. Next to the margins **Table 122** also displays whether pairwise sample comparisons of the moral deviance relevance margins show significant or non-significant differences. The following labeling is used to indicate (Holm-Bonferroni corrected) significance: p <0.05, ** p <0.01, *** p <0.001. Non-significant differences are indicated by the abbreviation *n.s.* and italics. These results refer to cross-cultural similarities. We have also color-coded the results to rank moral deviance relevance margins across the four groups (Green > Blue > Yellow > Red).

Table 122: Moral deviance relevance margins†: Full sample analysis across moral domains and across cultural groups

Germany n = 748	Samples Compared (pairwise) ††	Japan n = 738	Samples Compared (pairwise)	USA n = 736	Samples Compared (pairwise)	Egypt n = 746
4.691 Fairness †††	GER vs. JP *n.s.* GER vs. USA *** GER vs. EG ***	4.445 Fairness	*JP vs. USA n.s.* JP vs. EG *	4.173 Fairness	USA vs. EG *n.s.*	3.920 Fairness
4.464 Trustworthiness	GER vs. JP *** GER vs. USA *** GER vs. EG *	4.841 Trustworthiness	JP vs. USA *** JP vs. EG ***	4.064 Trustworthiness	*USA vs. EG n.s.*	4.114 Trustworthiness
4.612 Property	*GER vs. JP *** GER vs. USA *** GER vs. EG*	4.931 Property	JP vs. USA *** JP vs. EG ***	4.245 Property	*USA vs. EG n.s.*	4.234 Property
3.761 Reciprocity	GER vs. JP *** GER vs. USA *n.s.* *GER vs. EG n.s.*	4.353 Reciprocity	JP vs. USA *** JP vs. EG **	3.695 Reciprocity	*USA vs. EG n.s.*	3.895 Reciprocity
3.971 Heroism	GER vs. JP *** *GER vs. USA n.s.* GER vs. EG *n.s.*	4.426 Heroism	JP vs. USA *** JP vs. EG ***	4.027 Heroism	USA vs. EG *n.s.*	3.801 Heroism
3.828 Family	GER vs. JP *** GER vs. USA *** *GER vs. EG n.s.*	4.191 Family	JP vs. USA *n.s.* JP vs. EG *n.s.*	4.135 Family	USA vs. EG *n.s.*	4.130 Family
3.315 In-Group	GER vs. JP *** *GER vs. USA *** *GER vs. EG ***	3.909 In-Group	JP vs. USA *** JP vs. EG *n.s.*	3.545 In-Group	*USA vs. EG n.s.*	3.796 In-Group
3.284 Deference	GER vs. JP *** GER vs. USA** GER vs. EG ***	3.916 Deference	JP vs. USA ** *JP vs. EG n.s.*	3.576 Deference	*USA vs. EG n.s.*	3.854 Deference

† **Note:** The margins presented are statistical values calculated from predictions of the respective previously fit OLS-model at fixed values of the covariate culture and the integration of the average of the non-fixed covariates of the model (Average Marginal Effect). On display are the margins for variable "dimension of moral deviance" for the cultural groups: Germany (GER), Japan (JP), United States of America (USA) and Egypt (EG); †† **Note:** To test for significant differences in moral deviance relevance (margins), cultural groups were compared in a pairwise fashion. The asterisks indicate (Holm-Bonferroni corrected) significance: * p <0.05, ** p <0.01, *** p <0.001. Non-significant comparisons (*n.s.*) are highlighted in italics; ††† **Note:** colors indicate score ranking in country comparison (Green > Blue > Yellow > Red).

So, in a nutshell, what does *Table 122* show us? First of all, there are more cross-cultural differences than similarities: 19 out of a total of 48 possible pairwise comparisons prove to be non-significant. All other pairwise comparisons of the cultural samples display at least a significance level of * $p < 0.05$. As can be seen from the *green* coloring in *Table 122*, the **Japanese sample** consistently shows the highest deviance relevance scores across all moral domains in the sample comparisons, with the exception of the fairness deviance relevance margin. For the JP-sample, the margins indicate a general tendency towards a *mixed* moral system: Apart from the dimensions in-group and deference, all other moral domains show margins of ≥ 4.0 and thus tend towards deviance relevance. The **German sample**, on the other hand, is clearly characterized by individualizing morality in the *full sample* analysis: Only the *individualizing* domains of fairness, trustworthiness and property tend with margins of ≥ 4.0 towards moral deviance relevance. The fairness domain also appears to be the most pronounced in the German sample when comparing the cultures of our study. The **USA** also seems to be mainly characterized by an *individualizing* moral system. However, in addition to heroism, there is also the *binding* domain family in this sample which tends towards moral deviance relevance. This suggests that simply describing the moral system in the US with individualizing morality might be too narrow. Let us now turn to the **Egyptian** sample. First of all, it can be seen that the deviance relevance margins in the Egyptian sample, which were clearly lower in the naive model, now align considerably with those of the other samples. This demonstrates that the covariates play an important role in the context of the comparability of our four cultural samples and retrospectively supports their inclusion in the OLS models. Furthermore, looking at the margins for the EG-sample we can see that trustworthiness, property and the family domain tend towards deviance relevance. All other margins are below 4, the midpoint of the morality scale, and tend slightly more towards irrelevance than relevance.

Obviously, the analyses with the covariate model for the *full sample* already show moral deviance relevance margins that differ to a large extent from the results obtained on the basis of the model that we call naive. We consider these changes in the results as an improvement in validity compared to the results of the naive model, which we were able to achieve by including an appropriate set of covariates. Nevertheless, we argue that not only the naive estimation results but also the analyses of the *full sample* covariate model should be discarded as invalid. Although we have already outlined the reasons for the objection to the *full sample* above, they should be briefly repeated here for the sake of their importance.

Firstly, the *full sample* contains cases that are characterized by poor data quality. In addition, the *full sample* is not exclusively limited to cases of the respective cultural mainstream from the cultural groups studied, as indirectly indicated by multiple citizenships of several respondents. Against this background, we argue that the *full sample* should be adjusted by excluding cases in order to obtain a database that allows for more valid analyses. We propose two exclusion criteria to preclude cases: *A)* A first criterion is the exclusion of as many cases with poor data quality as possible. Cases with poor data quality are characterized by the fact that they completed the questionnaire very quickly (less than 12 minutes and 50 seconds). In addition, we consider data from cases that have demonstrably only clicked through the online questionnaire, and have provided a consistent response pattern in longer questionnaire batteries, to be of poor quality. *B)* Another criterion for adjustment is the (indirect) examination of whether the respondents were socialized in the mainstream culture of the respective cultural group. Since we are explicitly interested in cross-cultural similarities and differences, we therefore consider it important to limit the analyses to cases that have only the citizenship of the respective countries under investigation.

Hence, in the light of these objections, which we believe to be justified, we reject the results of the *full sample* covariate model. Instead, in the next step we carry out analyses based on an *adjusted sample* and then turn to the interpretation of the results that emerge from a data base that we consider more valid.

Moral Deviance Relevance in Cross-Cultural Comparison Based on Samples Tailored to Higher Education

As already mentioned elsewhere, the Egyptian sample differs from the other cultural groups in this study in several aspects of the sample composition. One of these aspects is *education*, which we would now like to take out and look at in more detail. In the following ***Table 123*** the distribution of cases on the variable *education* is displayed across cultural groups in tabular form for the *adjusted sample.*

Table 123: Educational attainment (ISCED) across groups (adjusted sample)

Highest education across countries (ISCED-2011)	Cultural entities in the sample				
	GER	JP	US	EG	Total
Primary Education	6	0	102	0	108
Lower secondary education	212	18	29	5	264
Upper secondary education	75	182	109	69	435
Post-secondary education	128	5	46	28	207
Short cycle tertiary education*	51	87	41	0	179
Bachelor´s or equivalent**	**72**	**115**	**137**	**440**	**764**
Master´s or equivalent	**101**	**128**	**49**	**29**	**307**
Doctoral or equivalent	**21**	**8**	**16**	**0**	**45**
No response (or still in school)	0	0	40	11	51***
Total	666	543	569	582	2,360

* **Note:** Short cycle tertiary education does not apply in the Egyptian education system. For this reason, no cases are displayed in this category for the EG-sample. ** **Note:** We emphasized those categories that are comprising higher education. *** **Note:** Cases with "No response" will be excluded from the following analysis.

As can be seen in ***Table 123*** (highlighted in bold), people with a high level of education are strongly overrepresented in the Egyptian sample compared to respondents with lower educational attainment. We suspect that the composition of the EG-sample, especially the large number of respondents with higher levels of education, distorts the deviance relevance analyses for the EG-sample and also the transfer of the results obtained to Egyptian society. However, at this point, objections could be raised that we included the variable education (ISCED-2011) in the covariate model of our OLS estimates and thus controlled for this variable. We have indeed included education as a covariate in our estimation model. Nonetheless, we hypothesize that there is unobserved heterogeneity despite this inclusion. While this unobserved heterogeneity may be associated with higher education to some degree of correlation, we yet believe that it is not fully controlled/captured by the education covariate in our model.[173]

---

[173] Please see in this regard the discourse on hidden bias sensitivity analyses in the context of the counterfactual causal analysis framework (Rosenbaum, 2005; 2010; Morgan & Winship, 2015, pp. 429-434). The following quote

Furthermore, it is now well known that student samples in particular can be considered outliers in a variety of measurements. Also, studies on culture and morality are not unaffected by this fact (Henrich et al., 2010a). It is true, though, that the Egyptian sample, which has a strong tendency towards higher education, is not a pure student sample per se. But in this sample, more than 80% of the respondents have attended or are still attending university. Against this background, we argue below that the fact of attending institutions of higher education is associated with unobserved heterogeneity that affects the measurement of moral deviance relevance. Furthermore, we assume that this unmeasured heterogeneity leads in large parts to downward biased scores of MaC-DRS. Since in our study the Egyptian sample in particular is strongly biased towards higher education, we therefore hypothesize that in the present study it is primarily the EG-sample that is affected by this effect of unobserved heterogeneity.

To test this assumption, we repeated the MaC-DRS analyses of the 8 moral domains with a tailored sample that includes exclusively respondents with higher education. We use the *adjusted sample* as a starting point and only include cases with a university degree (minimum requirement Bachelor´s degree of equivalent) in the following analyses. Additionally, we use the same covariates (including interaction terms) in the OLS models as in the previous analysis of the *adjusted sample*. The following analysis therefore differs from the analysis already conducted only in the underlying sample data. ***Table 124*** displays the distribution of cases across the cultural groups for the analysis tailored to higher education.

Table 124: Case distribution on cultural groups: higher education sample

	Freq.	Percent	Cum.
Germany	194	17.38	17.38
Japan	251	22.49	39.87
USA	202	18.10	57.97
Egypt	469	42.03	100.00
Total	1,116	100.00	

---

from Funk and colleagues (2011), which is used in the context of doubly robust estimators, is also relevant to our argument in terms of the correlation logic referred to: "Although doubly robust property does give the analyst 2 means to achieve exchangeability, we emphasize that this method does not obviate the need to measure all confounders. Bias due to unmeasured confounders would be reduced only to the extent that these are correlated with measured characteristics that *are* included in one of the components models" (S. 763). In our case of the OLS estimation of moral deviance relevance, unobserved heterogeneity would therefore only be controlled by the education variable in the covariate model to the extent that the education variable is also correlated with the unmeasured yet influential variables. Since from a theoretical point of view it is certainly difficult to assume a perfect correlation between education and other variables, there is thus still enough room for the assumption that unobserved heterogeneity has an effect on the measurement of moral deviance relevance despite the inclusion of education as a covariate in our model, and this is precisely what we argue for.

First of all, due to the relatively low number of cases within the cultural samples, it can be stated that the *higher education sample* analyses are carried out under conditions of insufficient statistical power.[174] Consequently, the results cannot be considered reliable with regard to analysis of cross-cultural differences. Nevertheless, trends of *unchanged*, *increased*, or *decreased* moral deviance relevance margins can be derived by comparing the results obtained via the different data bases. So, to check for such trends, we accordingly compare the moral deviance relevance margins obtained from the *adjusted sample* (see: **Table 29** and **Table 31**) with those of the *sample* tailored to *higher education* (N = 1,116).

We now expect that the deviance relevance margins found in the *adjusted sample* analyses for the Egyptian cultural group (**Table 29**) are largely underestimated for Egyptian society, as the EG-sample consists predominantly of highly educated respondents. In other words: We assume that the EG-sample reflects the deviance relevance scores for Egypt solely in a distorted way and that this distortion is caused by the fact that the sample consists mainly of respondents with higher education. With this in mind, we further expect that higher education is associated with unmeasured social factors, which in turn are linked to social positioning and independence, contributing to a lower relevance of deviance in various moral domains measured by MaC-DRS. With regard to this rather general assumption, not only the deviance relevance scores of the EG-sample but also those of the other cultural groups are of interest to test our conjecture. So, given we would find lower deviance relevance margins in the four cultural groups of the *higher education sample* (N = 1,116) compared to the *adjusted sample* (N = 2.360), we see our assumption confirmed. Following this logic, and to the extent that the following analysis confirms our conjecture, we then have an empirical basis for claiming that the moral deviance relevance scores for the EG-sample in **Table 29** respectively **Table 31** (*adjusted sample*) are biased downwards and thus underestimated. Furthermore, if we indeed find consistently lower deviance relevance margins in the *sample* tailored to *higher education* (compared to the *adjusted sample*), it is necessary to at least speculate on how this effect can be explained (across cultures). As indicated before, we reason that socioeconomic positioning may play a role. All in all, we will not only conduct a sensitivity analysis in the following, but also pursue an exploratory concern, from which we may be able to derive new hypotheses for future research.

In order to investigate our conjecture empirically, we will now take a look at two tables. First, **Table 125** compares the deviance relevance margins of the *adjusted sample* (N = 2,360)

---

[174] We are referring here to the results of the a priory power analysis that we conducted before data collection of our cross-cultural study. See: **Chapter 2** for details.

and the *higher education sample* (N = 1,116). ***Table 125*** further below then shows the moral deviance relevance margins of the estimated OLS models for the sample tailored to higher education in a cultural comparison.

Although some deviance relevance margins remain relatively stable (e.g. the heroism and in-group margins in the JP-sample), ***Table 125*** demonstrates yet quite clearly that within the *higher education sample*, compared to the *adjusted sample*, the deviance relevance margins *decrease* across moral domains and all four groups. The Egyptian sample (n = 582) comprises over 80% of respondents with a degree of higher education (Bachelor's degree or higher). Furthermore, according to the findings in ***Table 125***, having attained a degree of higher education is associated with lower moral deviance relevance margins across cultures. Based on this explorative finding, we see our conjecture confirmed: the deviance relevance margins for the Egyptian sample in ***Table 29*** respectively ***Table 31 (adjusted Sample)*** are likely to be underestimated for Egyptian society. This underestimation of actual deviance relevance scores in Egyptian society may explain why, for example, deviance towards the deference domain comes with a margin of ≤ 4.0 in the *adjusted sample* analysis of the EG-sample (see: ***Table 31***).

Furthermore, when comparing the *adjusted sample* with the *higher education sample* it is apparent that deviance relevance is reduced across all moral domains and cultures studied. At this point, we can only speculatively reason about the causes of this finding. One possibility may be the insufficient statistical power of the higher education sample. However, a substantive interpretation could also be drawn taking into account the deviant actions portrayed in the MaC-DRS items and the socio-economic positioning of people with higher education. In this light we thus argue, that acquired and accumulated (economic, social and cultural) capital could buffer people from non-drastic consequences of failed cooperation.

Looking at the MaC-DRS items, it can be seen that, although the violations shown represent clear deviations from morality, they are not particularly drastic. With regard to the property domain, for example, one MaC-DRS item reads: "*Someone damages another person's property without replacing it*". For the family domain, a MaC-DRS item reads: "*Someone does not support their own family and relatives*"; for the reciprocity domain: "*Someone enjoys favors from others without reciprocating*", and finally another example (MaC-DRS deference domain): "*Someone acts disrespectfully towards an authority*". In the conceptualization of MaC-DRS, we have deliberately avoided more pointed, drastic deviance statements within the

Table 125: Moral deviance relevance margins in comparison: The adjusted sample vs. the higher education sample

Germany n = 666†	Germany n = 194††	Japan n = 543†	Japan n = 251††	USA n = 569†	USA n = 202††	Egypt n = 582†	Egypt n = 469††
4.870 Fairness	4.750 Fairness †††	4.570 Fairness	4.351 Fairness	4.291 Fairness	4.159 Fairness	4.015 Fairness	3.844 Fairness
4.601 Trustworthiness	4.526 Trustworthiness	4.993 Trustworthiness	4.526 Trustworthiness	4.107 Trustworthiness	3.646 Trustworthiness	4.239 Trustworthiness	4.043 Trustworthiness
4.772 Property	4.455 Property	5.181 Property	4.973 Property	4.336 Property	3.925 Property	4.477 Property	4.263 Property
3.906 Reciprocity	3.693 Reciprocity	4.495 Reciprocity	4.362 Reciprocity	3.709 Reciprocity	3.365 Reciprocity	3.936 Reciprocity	3.730 Reciprocity
4.115 Heroism	3.971 Heroism	4.563 Heroism	4.517 Heroism	4.072 Heroism	3.776 Heroism	4.062 Heroism	3.833 Heroism
3.961 Family	3.749 Family	4.344 Family	4.242 Family	4.239 Family	4.044 Family	4.407 Family	3.979 Family
3.412 In-Group	3.070 In-Group	4.042 In-Group	4.038 In-Group	3.603 In-Group	3.478 In-Group	4.058 In-Group	3.777 In-Group
3.374 Deference	3.015 Deference	4.043 Deference	3.845 Deference	3.603 Deference	3.456 Deference	3.960 Deference	3.644 Deference

†The scores displayed in *light grey* are the margins based on the **Adjusted Sample** (N = 2,360). These scores are shown for reasons of comparison. ††The scores displayed in *black* are the margins based on the **Higher Education Sample** (N = 1,116). [175]

---

[175] We also calculated for the lower education sample (educational attainment below Bachelor´s degree, N = 1,193) the moral deviance relevance margins (OLS covariate model) and found, apart from two exceptions (EG-sample: property = 3.944; reciprocity = 3.791), consistently higher deviance relevance margins across cultural groups. We currently have no explanation as to why the marginal value for the property and the reciprocity domain are falling in the lower education EG-subsample. Overall, however, regarding the lower education EG-sample (n = 102), the MaC-DRS analyses show moral deviance relevance (margins ≥ 4.0) for the following domains: fairness = 4.404; trustworthiness = 4.567; heroism = 4.203; family = 5.002; in-group = 4.603; deference = 4.267. These results, although based on a very small sample, largely support our argument of downward biased deviance relevance margins for the EG-sample (n = 582) of the main analysis that is resting on the adjusted sample (N = 2.360).

items for the time being, as our primary goal with MaC-DRS is to measure different moral domains across cultures as precisely as possible.[176] The moral deviations given in the items could therefore be much more drastic, which is currently not the case. To illustrate what we mean by drastic, here are two brief examples about how the moral breaches in the items could be made more extreme: "*Someone steals their credit card and robs all the money in their bank account*" (property domain), or "*Someone despises their family and actively tries to harm them*" (family domain). From our point of view, these two examples differ in their drastic from the MaC-DRS items we used to collect data. By reference to the diminished severity in the MaC-DRS items, we are aiming at a certain interpretation of the reduced deviance relevance scores in the higher education sample. We believe that it is plausible to assume that higher education can be associated with more cultural, economic and to some extent more powerful social capital (Markus & Schwartz, 2010; Bourdieu, 1983; Bourdieu, 2014; compare in particular pp. 31-56; Kühnen & Kitayama, 2024). Educational titles attained cannot be converted one-to-one into means of socioeconomic security in every society and at all times. Nonetheless, it can be justifiably assumed that educational degrees attained at least usually make a significant contribution to the social positioning of actors. With higher socioeconomic status (SES) positioning we assume furthermore higher availability of means convertible into socioeconomic security and independence (Sachdeva et al., 2011; Santos et al., 2017; Kühnen & Kitayama, 2024). In other words: non-extreme breaches of cooperation may be evaluated as less relevant if a certain degree of, above all economic and social, security protects against zero-sum consequences. From this perspective, it could be explained why non-drastic moral deviance (MaC-DRS) is classified as less relevant in the *higher education sample* compared to the *adjusted sample*.[177]

This notion is in line with findings portrayed in a literature review on social class and self-construal by Kühnen and Kitayama (2024), who state:

> "People from a lower social class may be more threat sensitive, partly because they are more likely affected by threats such as crime, violence, or poverty. (…) Accordingly, recent studies found increased neural activity in response to various external threats for people from low compared to high SES" (p. 10).

---

[176] Since in this study we were able to demonstrate the precise measurement of 8 moral domains using the MaC-DRS across 4 different cultural contexts, it would be a worthwhile challenge for the future to include more pointed, more drastic moral deviations in the items of upcoming MaC-DRS versions.

[177] Based on the interpretation presented, the question could be raised as to whether within cultural moral tightness is in part associated with socio-economic positioning.

Moral transgressions are associated with non-zero-sum consequences (e.g., failed cooperation), which can have more detrimental and thus threatening consequences for people of low socioeconomic status due to fewer overall choices, resources, and opportunities (Markus & Schwartz, 2010). Flipping this line of reasoning on its head, one encounters the argument that higher education, which tends to be associated with more economic capital and more powerful social capital, might protect against non-extreme moral transgressions. Consequently, one might expect this to be reflected in the calibration of people's moral mind, as evidenced by a lower intuitive relevance attribution for non-severe moral breaches. Precisely for the latter, we found empirical support in our study that is consistent with this reasoning.

Nonetheless, the substantive interpretation drawn here remains primarily in the realm of a speculative hypothesis. Although we have empirical findings that permit the interpretation drawn above, we need more evidence based on samples that possess (e.g.) higher sample sizes, and sufficient statistical power to safeguard this hypothesis. Apart from whether our interpretation is correct, it is nevertheless a fact shown in the data, that the *higher education sample* consistently, i.e., across moral domains and cultural groups, displays lower margins with regard to moral deviance relevance. We therefore conclude two points. Firstly, the EG-sample of our study is by no means representative of Egyptian society, as it is heavily skewed towards higher education. Our moral deviance relevance findings should therefore only be applied to Egyptian society with due caution. Secondly, based on what we have elaborated, we assume that the EG-sample findings (*adjusted Sample*; **Table 31**) underestimate the actual moral deviance relevance in Egypt´s society due to the presence of a downward bias associated with higher education.[178]

To conclude our digression, we turn in the following briefly to the comparison of moral deviance relevance across moral domains and cultural groups using the sample tailored to higher education (**Table 126**). We mention this addition due to an explorative finding indicating a tendency of cross-cultural alignment in the evaluation of moral deviance relevance between the samples, which we will discuss shortly.

Table 126 shows that the analysis of moral deviance relevance reveals non-significant differences more frequently in the *higher education sample* than in the *adjusted sample*. In the former sample, only 21 out of 48 pairwise comparisons of the cultural samples demonstrate significant differences. Conversely, with 27 non-significant pairwise comparisons of the

---

[178] Please also note that higher levels of educational attainment are associated with greater individualism (Santos et al., 2017). We therefore advise caution and restraint when attempting to generalize the EG-sample results of our study to Egyptian society, as the sample may also be biased in favor of greater individualizing morality, among other factors.

cultural samples, more than half of the comparisons are not significant. By contrast, the analyses of the *adjusted sample* (**Table 29**) show 20 non-significant pairwise comparisons. Accordingly, we find considerably more cross-cultural similarities in moral deviance relevance in the *higher education sample*. Two assumptions can be drawn from this explorative finding: *A)* due to insufficient statistical power, we find fewer significant differences in moral deviance relevance in the *higher education sample*, although in fact there are more. *B)* The finding of more non-significant pairwise comparisons across cultural groups could indicate sample *alignment* indirectly attributable to the variable (attaint degree of) educational attainment. In other words, the finding of a greater number of non-significant average marginal effects for deviance relevance across moral domains and cultural groups could indicate the possibility of a *cross-cultural alignment* that can be attributed, at least in part, to the acquisition of higher educational attainment that accompanies a socialization process in higher educational institutions (e.g., universities).[179] We cannot rule out either possibility *A)* or possibility *B)* in our study.

With these two possibilities in mind, however, we would like to point out to the importance of samples with *sufficient statistical power* (Lakens, 2014; 2022) and the *possibility of cross-cultural alignment* through socialization in institutions that have similar policies and structural foundations across cultures. Researchers interested in the effects of culture on the human mind should therefore take at best these two possibilities into account in their study designs in order to be able to reveal actual cross-cultural differences or similarities. We will now end this digression and turn back to the EG-sample in the context of our research question and the cultural difference hypotheses.

---

[179] We would like to make a brief comment on this possibility, which we refer to as *alignment hypothesis*: If there is indeed cross-cultural alignment resulting from secondary socialization in more or less the same institution (i.e., institutions of higher education such as universities), this would pose a real challenge for cross-cultural comparisons based on student samples. In this case, student samples would probably not be well suited for cross-cultural research as they would not be able to show actual cultural similarities or differences. Consequently, it would be difficult to make adequate statements of cross-cultural research, as a socialization process in more or less the same institution might have leveled the cross-cultural similarities/differences (especially the latter) by aligning intuitive tendencies. Evidence from other studies can be interpreted as to support our alignment hypothesis (Henrich, 2020, see explicitly pp. 24-34). However, if there is such a cross-cultural alignment process, partly due to institutional socialization, this does not necessarily mean that it is also reflected in the everyday life of people with higher levels of education. Cross-cultural differences in everyday life can remain, to a certain degree, unaffected by such institutional socialization. Rather, the assumption of such a process of alignment has implications for research with student samples because it might be expected that the importance of university socialization and identity is brought to the (cognitive) foreground simply by the setting of data collection (e.g. a questionnaire for data collection). In the event that our exploratory hypothesis of cross-cultural alignment proves to be true, we would have a strong indication that seems to support the argumentation of Henrich et al. (2010a): Relying solely on student samples for cross-cultural comparisons can lead to biased results. Not only can statements based on such samples not be considered generalizable, but furthermore, statements about cross-cultural differences and similarities based on student samples would have to be treated with great caution, as they potentially undermine actual cross-cultural differences.

Table 126: Higher education sample: Margins† of moral deviance relevance across moral domains and cultural groups

Germany n = 194	Samples Compared (pairwise)††	Japan n = 251	Samples Compared (pairwise)	USA n = 202	Samples Compared (pairwise)	Egypt n = 469	Egyptian Sample Compared
4.750 Fairness †††	GER vs. JP *n.s.* GER vs. USA * GER vs. EG ***	4.351 Fairness	JP vs. USA *n.s.* JP vs. EG *n.s.*	4.159 Fairness	USA vs. EG *n.s.*	3.844 Fairness	
4.526 Trustworthiness	GER vs. JP *n.s.* GER vs. USA *** GER vs. EG *n.s.*	4.526 Trustworthiness	JP vs. USA *** JP vs. EG *n.s.*	3.646 Trustworthiness	USA vs. EG *n.s.*	4.043 Trustworthiness	
4.455 Property	GER vs. JP * GER vs. USA ** GER vs. EG *n.s.*	4.973 Property	JP vs. USA *** JP vs. EG *	3.925 Property	USA vs. EG *n.s.*	4.263 Property	
3.693 Reciprocity	GER vs. JP *** GER vs. USA *n.s.* GER vs. EG *n.s.*	4.362 Reciprocity	JP vs. USA *** JP vs. EG *	3.365 Reciprocity	USA vs. EG *n.s.*	3.730 Reciprocity	
3.971 Heroism	GER vs. JP * GER vs. USA *n.s.* GER vs. EG *n.s.*	4.517 Heroism	JP vs. USA *** JP vs. EG *	3.776 Heroism	USA vs. EG *n.s.*	3.833 Heroism	
3.749 Family	GER vs. JP *n.s.* GER vs. USA *n.s.* GER vs. EG *n.s.*	4.242 Family	JP vs. USA *n.s.* JP vs. EG *n.s.*	4.044 Family	USA vs. EG *n.s.*	3.979 Family	
3.070 In-Group	GER vs. JP *** GER vs. USA * GER vs. EG ***	4.038 In-Group	JP vs. USA * JP vs. EG *n.s.*	3.478 In-Group	USA vs. EG *n.s.*	3.773 In-Group	
3.015 Deference	GER vs. JP *** GER vs. USA * GER vs. EG *	3.845 Deference	JP vs. USA *n.s.* JP vs. EG *n.s.*	3.456 Deference	USA vs. EG *n.s.*	3.644 Deference	

† **Note:** The margins presented are statistical values calculated from predictions of the respective previously fit OLS-model at fixed values of the covariate culture and the integration of the average of the non-fixed covariates of the model (Average Marginal Effect). On display are the margins for variable "dimension of moral deviance" for the cultural groups: Germany, Japan, USA and Egypt; †† **Note:** To test for significant differences in moral deviance relevance (margins), cultural groups were compared in a pairwise fashion. The asterisks indicate (Holm-Bonferroni corrected) significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Non-significant comparisons (*n.s.*) are highlighted in italics; ††† **Note:** colors indicate score ranking in country comparison (Green > Blue > Yellow > Red). Δ **Note:** we have tailored the sample to include exclusively cases with higher education. Please note also that the sample size within the particular cultural groups is insufficient in term of statistical power.

Analysis of Robustness: Comparison of the Naive Model, the Full Sample Covariate Model and the Adjusted Sample Covariate Model

The following ***Table 127*** displays the results of all three models in a comprehensive table, i.e., findings from the naïve model, the full sample covariate model and the adjusted sample covariate model are displayed. In the table we have opted for the within-sample ranking of moral deviance relevance margins.

Even though all margins change across the models and across all groups, two circumstances are particularly noticeable: firstly, the deviance relevance rank orders of the GER-, JP- and US-sample are relatively constant across all models. We strongly and, in our view, rightly advocate that only the results of the *adjusted sample* should be interpreted (row C in the ***Table 127***). Nevertheless, a certain robustness of our results can be observed across all three models for the samples mentioned. Secondly, the same robustness of the results cannot be stated for the EG-sample. This becomes evident when we look at the values of the naive model in comparison to the two covariate models and also the rank changes of the moral domains. So, what we want to show with ***Table 127*** is essentially just these two things: In the context of our research question, the MaC-DRS results for Germany, Japan and the United States of America appear to be relatively robust to our models, by and large, but the same cannot be said for the Egyptian sample. We believe it is important to keep these robustness considerations in mind when evaluating the MaC-DRS results. On the whole, we have analyzed what we believe to be the most valid database — the adjusted sample — in detail. Nevertheless, we would like to point out that our samples are not representative and show weaknesses in parts, as can be observed especially in the EG-sample. Hence, our analyses and results must be evaluated against this background.

Table 127: Moral deviance relevance (margins)† — within cultural groups ranking (highest to lowest) across samples

GER-sample A n = 751††	GER-sample B n = 748	GER-sample C n = 666	JP-Sample A n = 740	JP-Sample B n = 738	JP-Sample C n = 543	US-Sample A n = 745	US-Sample B n = 736	US-Sample C n = 569	EG-Sample A n = 746	EG-Sample B n = 746	EG-Sample C n = 582
5.235 ††† Prop.	4.691 Fair.	4.870 Fair.	5.036 Prop.	4.931 Prop.	5.181 Prop.	4.482 Fair.	4.245 Prop.	4.336 Prop.	2.869 Defer.	4.234 Prop.	4.477 Prop.
5.196 Fair.	4.612 Prop.	4.772 Prop.	4.951 Trust.	4.841 Trust.	4.993 Trust.	4.670 Prop.	4.173 Fair.	4.291 Fair.	2.635 Reci-	4.130 Fam.	4.407 Fam.
5.066 Trust.	4.464 Trust.	4.601 Trust.	4.490 Fair.	4.445 Fair.	4.570 Fair.	4.532 Trust.	4.135 Fam.	4.239 Fam.	2.628 Prop-	4.114 Trust.	4.239 Trust.
4.490 Hero.	3.971 Hero.	4.115 Hero.	4.439 Reci.	4.426 Hero.	4.563 Hero.	4.504 Family	4.064 Trust.	4.107 Trust.	2.622 Fair.	3.920 Fair.	4.062 Hero.
4.340 Reci.	3.761 Reci.	3.961 Fam.	4.372 Hero.	4.353 Reci.	4.495 Reci.	4.404 Hero.	4.027 Hero.	4.072 Hero.	2.613 Fam.	3.895 Reci.	4.058 In-Gr.
4.323 Fam.	3.828 Fam.	3.906 Reci.	4.260 Fam.	4.191 Fam.	4.344 Fam.	4.077 Reci.	3.695 Reci.	3.709 Reci.	2.594 In-Gr.	3.854 Defer.	4.015 Fair.
3.792 In-Gr.	3.315 In-Gr.	3.442 In-Gr.	3.891 In-Gr.	3.916 Defer.	4.043 Defer.	4.017 Defer.	3.576 Defer.	3.603 Def./In.	2.594 Hero.	3.801 Hero.	3.960 Defer.
3.780 Defer.	3.284 Defer.	3.374 Defer.	3.880 Defer.	3.909 In-Gr.	4.042 In-Gr.	3.936 In-Gr.	3.545 In-Gr.	3.603 In./Def.	2.538 Trust.	3.796 In-Gr.	3.936 Reci.

† **Note:** The margins presented are statistical values calculated from predictions of the respective previously fit OLS-model at fixed values of the covariate culture and the integration of the average of the non-fixed covariates of the model (Average Marginal Effect). On display are the margins for variable "dimension of moral deviance" for the cultural groups: Germany (GER-sample), Japan (JP-Sample), USA (US-sample), and Egypt (EG-sample); †† **Note:** The marker **A** is used to indicate the **naive model** (full sample); **B** is used to indicate the **full sample covariate model**; **C** is used to indicate the **adjusted sample model**. ††† **Note:** The margins are ordered by magnitude (descending order: Green > Black > Red).

Supplement Chapter 6: The Moral Deviance Factorial Survey (MDFS)

Analyzing MDFS: Regression Models for 7 Domains of Moral Deviance

Our basic OLS regression model (**Model 1**) for analyzing the Factorial Survey includes a large number of other covariates in addition to the variables of interest, *culture* and *social relationship*. The basic model, with its variables, can be found in the following *Table 128*.

Table 128: Analyzing MDFS — Basic OLS regression model (Model 1)

Model 1	
Dependent variables	Deviance Relevance; Deviance Judgment; Deviance Shame Attribution; Deviance Guilt Attribution
Variables of main interest (categorical variables)	Culture (four samples); Dimension Social Relationship (Factorial Survey; 3 expressions); Culture x Social Relationship; Moral Domain (of deviant behavior)
Independent categorical variables	Dimension Gender (factorial survey); Dimension Reputational Damage (factorial survey); Denomination; Place of Upbringing; Place of Living; Residential Mobility; Gender; Education (ISCED)
Independent continuous variables	Age; MRS; NARS; Pathogen Prevalence; Level of Religiosity;
Interaction terms	Interaction with variable culture: Age; Level of Religiosity; MRS; NARS; Pathogen Prevalence; Dimension Gender (Factorial Survey); Dimension Reputational Damage (Factorial Survey); Interaction: Level of Religiosity x Denomination

Using this model for our analyses would mean that we are using a very complex model on the one hand and that we are obviously in the area of multiple testing on the other. Consequently, the latter circumstance must be taken into account in the form of adjustments to the significance level (Holm-Bonferroni correction). The test family relating to the variable *culture* comprises a total of 51 different tests in **Model 1**. Adjusting the significance level would therefore mean that the probability of identifying effects that are not highly significant would be reduced by the model and its 51 tests of a test family. Furthermore, the model would be extremely complicated with a large number of main effects and a total of 42 interaction effects (only for the variable *culture*), which would of course also affect the interpretation of the models. Against

this background, we have decided to streamline Model 1 and to make it more frugal. We will therefore remove variables and use a second, simplified model as a starting point of analyses.

In addition to *culture* and the *moral domain of deviance*, we are primarily interested in the effect of the *social relationship* dimension (stranger; family; in-group/friend) of the Factorial Survey. As indicated before, we cannot address the entire Factorial Survey of our study with all dimensions in detail and with the necessary precision in the present paper, as this would be an excessive undertaking that would clearly go beyond the scope of this thesis. On this basis, we have decided not to examine the Factorial Survey dimensions *gender* and *reputational damage* in more detail in this study. The corresponding variables (main effects and interaction terms) are therefore not included in **Model 2** and remain as content to be investigated in future studies. Furthermore, we decided to remove the categorical education variable (ISCED) and only consider possible educational effects by means of the continuous *years in school* variable. As far as the variable *place of upbringing* is concerned, we have decided to remove this variable as well. Since we retain the variable *place of living* for Model 2, which also captures the aspect of village vs. city living, we consider the exclusion of *place of upbringing* to be justified. However, we have not yet reached the end of streamlining our initial model. Therefore, we will further simplify the second model for our analyses and eventually use a final **Model 3** adapted for the analyses of each moral domain.

As already mentioned at the beginning of this chapter, the vignettes deal with specific types of deviance from different moral domains. In contrast to the MaC-DRS analyses, the findings of the Factorial Survey cannot be compared with each other across the different moral domains nor would such a comparison be meaningful. The background to this argument is that the vignettes portray *specific*, domain-related behavior and we simply cannot assume a priori that the described behaviors are the same in their (deviance) *severity* across the domains. A within-sample deviance relevance ranking (e.g.) as in the MaC-DRS analyses is therefore neither intended nor considered expedient. Instead, we will analyze and discuss contextualized and, above all, specific deviance within 7 moral domains across cultures with regard to *relevance*, *judgment*, *shame attribution*, and *guilt attribution*.

It also follows from the above argument that the respective models for the 7 moral domains may differ from one another, as we are not aiming for a comparison across the moral domains. Therefore, based on Model 2, we will empirically derive a specific model (**Model 3**) adapted for each moral domain investigated. By empirical derivation, we mean that we remove those variables from the corresponding model that prove to be non-significant within a moral

domain for all four dependent variables. However, it is important to note that a constant model for *relevance*, *judgment*, *shame*, and *guilt* is estimated for each moral domain. On the basis of these estimation models, we then carry out our analyses.

Response Style Effects in the Moral Deviance Factorial Survey

Since the Moral Deviance Factorial Survey is a newly developed research tool, we are also interested in a methodological aspect in the context of cross-cultural measurement of moral deviance. Hence, will carry out analysis of the **response style** measures (NARS and MRS) in the context of the Factorial Survey in order to take a closer look at the methodological components to be considered in the context of applying this instrument in cross-cultural studies (Baumgartner & Weijters, 2015; He et al., 2021).

The Property Vignettes

As far as the response style measures *NARS* (*net acquiescence response style*) and *MRS* (*midpoint response style*) are concerned, we have found the following: the relevance model shows two (Holm-Bonferroni corrected) significant and positive NARS interaction effects for the *culture* variable characteristics Germany (Coeff. = .022; Std. Err. = .021; $p = 0.03$ *) and Egypt (Coeff. = .024; Std. Err. = .004; $p < 0.001$ ***). Furthermore, the shame model yields a highly significant and negative main effect of MRS: Coeff. = -.034; Std. Err. = .008; $p < 0.001$ ***, and also the guilt model shows a significant and negative MRS main effect (Coeff. = -.029; Std. Err. = .009; $p = 0.026$ *). After adjusting the significance level, no further significant response style effects could be identified. Taken together, the analyses suggest that response styles are also present in the specific property deviance models and play a culture specific role for the relevance model. Consequently, the findings suggest that the possibility of response style bias should be taken into account when analyzing the Moral Deviance Factorial Survey across cultures.

The Fairness Vignettes

The relevance model yields two (Holm-Bonferroni corrected) significant and positive NARS interaction effects with *culture*: GER-sample, Coeff. = .016; Std. Err. = .004; $p = 0.03$ *; EG-sample, Coeff. = .013; Std. Err. = .003; $p < 0.001$ ***. In the judgment model we find in addition

a significant and negative MRS main effect (Coeff. = -.003; Std. Err. = .010; p = 0.026 *), which is also present in the attribution of shame OLS model (Coeff. = -.003; Std. Err. = .001; p = 0.027 *). After applying the p-value adjustment, no further response style influences were found to be significant. All in all, the specific fairness deviance results suggest that it is advisable to take response styles into account in cross-cultural studies working with the Moral Deviance Factorial Survey.

The Trustworthiness Vignettes

As far as the specific trustworthiness OLS models in the context of the two *response style* variables are concerned, our analyses yield that only the midpoint response style (MRS) exhibits a significant impact. In the judgment (Coeff. = -.050; Std. Err. = .010; p < 0.001 ***), shame (Coeff. = -.053; Std. Err. = .010; p < 0.001 ***), as well as the guilt model (Coeff. = -.044; Std. Err. = .010; p < 0.001 ***), we found a significant (negative) main effect for MRS. No further significant response style effects were found after applying the Holm-Bonferroni correction. Consequently, neither NARS nor culture specific response style effects seems not play a prominent role in the trustworthiness deviance vignette models of our investigations.

The Heroism Vignettes

Let us turn to the *response style* measures. The OLS models yield a total of three interaction effects and two main effects. In the relevance model, there is once again a significant and positive interaction with the NARS measure and the GER-sample (Coeff. = -.014; Std. Err. = .004; p < 0.001 ***) as well as the EG-sample (Coeff. = .015; Std. Err. = .003; p < 0.001 ***). As far as the judgment and shame OLS models are concerned, the negative MRS main effect is also evident again: judgment model, MRS Coeff. = -.033; Std. Err. = .009; p = 0.035 *; shame model, MRS Coeff. = -.032; Std. Err. = .009; p = 0.03 *. Furthermore, the Holm-Bonferroni corrected p-values reveal a significant interaction effect in the judgment model for NARS, which applies exclusively to the German sample: Coeff. = .011; Std. Err. = .003; p = 0.035 *. No further significant response style effects were found. Once more, these findings suggest that response style effects should be taken into account in the context of using the Factorial Survey as a measure of valuation to moral deviance.

The Reciprocity Vignettes

Turning to the *response styles*, we find a total of three main effects attributable to MRS, and one NARS interaction effect with culture. The interaction effect is observed in the relevance OLS model for the EG-sample and is positive: Coeff. = .016; Std. Err. = .003; p < 0.001 ***. The MRS main effects, in contrast, are all negative: judgment model, MRS main effect: Coeff. = -.038; Std. Err. = .008; p < 0.001 ***; shame model, MRS main effect: Coeff. = -.034; Std. Err. = .008; p < 0.001 ***; guilt model, MRS main effect: Coeff. = -.032; Std. Err. = .008; p < 0.001 ***. No further significant NARS or MRS effects are present in the specific reciprocity deviance OLS models. Taken together, our results suggest that response styles are not meaningless in relation to the reciprocity vignettes of the Factorial Survey and should be considered when working with the respective scenarios in cross-cultural research.

The Loyalty Vignettes

Moreover, within the relevance, judgment, shame and guilt OLS models only two significant and positive NARS interaction effects are found after correcting for multiple testing. As in the models of the other moral domains, these effects apply exclusively to the German and Egyptian samples: GER-sample, Coeff. = -.017; Std. Err. = .004; p = 0.03 *; EG-sample, Coeff. = .022; Std. Err. = .004; p < 0.001 ***). No other response style effects were found to be significant. It seems to depend in part on the groups studied whether response styles exert a significant influence in the Factorial Survey data. However, since it cannot be ruled out a priori that culture-specific communication styles have an effect, the findings presented indicate once more that it is advisable to consider different response style measures in the context of the Moral Deviance Factorial Survey and cross-cultural investigations.

The Deference Vignettes

Once more we encounter a familiar picture: after the Holm-Bonferroni correction in the relevance model, we find two interaction effects for NARS and the cultural samples Germany and Egypt (GER-sample, Coeff. = -.017; Std. Err. = .004; p < 0.001 ***; EG-sample, Coeff. = .026; Std. Err. = .004; p < 0.001 ***). As far as the judgment, shame and guilt models are concerned, there are again highly significant and negative main effects for MRS: judgment model, MRS main effect: Coeff. = -.050; Std. Err. = .009; p < 0.001 ***; shame model, MRS main effect: Coeff. = -.042; Std. Err. = .008; p < 0.001 ***; guilt model, MRS main effect:

Coeff. = -.042; Std. Err. = .008; p > 0.000 ***. After applying the significance level adjustment, no further effects are below the threshold of p < 0.05. All in all, the deference vignette OLS models also show a similar effect of response styles as we have already observed in a large number of the other Factorial Survey models. Conclusively, response styles matter in cross-cultural investigations using data obtained with the Moral Deviance Factorial Survey.

Discussion: MDFS and Response Styles

The **NARS and MRS findings** show clearly that when using **MDFS** for cross-cultural studies, it is important to control for response styles. This statement is largely supported by the results across the 28 OLS models with the four dependent variables. As with MaC-DRS, the response behavior in the context of the specific deviance relevance, judgment, shame and guilt variables is not free of culture specific ways of communicating. Furthermore, the data across the models seem to indicate a pattern: the *net acquiescence response style* variable (NARS) that we designed is, with one exception, solely significant in the context of moral deviance relevance.[180] In addition, positive interaction effects found for NARS only apply for the EG- and GER-sample. The other study groups therefore show no significant effect for NARS after the Holm-Bonferroni correction and also remain unaffected by the influence of this response style in the Factorial Survey analyses. Considering MRS, we only find negative main effects. These are primarily found in the shame-OLS models, but the judgment and guilt models are also not unaffected by the midpoint response style (MRS). As a main effect, though, the results for MRS apply across groups. How the NARS and MRS pattern is to be evaluated, whether it originates from the samples on which our analyses are based — we have already mentioned at the outset that the analyses for the EG-sample cannot be assumed to be robust in the Moral Deviance Factorial Survey either —, is influenced by the MaC-DRS (relevance) items, which account for a large proportion of the NARS and MRS measures, or point to other patterns that have yet to be conclusively clarified (note that the NARS interaction effects were found exclusively for the GER- and EG-sample), needs to be addressed in future studies. What our study is able to show, however, is evidence suggesting that it is advisable to control for response styles in cross-cultural studies using the Moral Deviance Factorial Survey (van de Vijver & Leung, 2011; He & van de Vijver, 2012; Baumgartner & Weijters, 2015; He et al., 2021).

---

[180] We also found a significant and positive interaction effect between NARS and culture (GER-sample) in one of the OLS models for the dependent variable judgment.

STATUTORY DECLARATION

I hereby declare that this dissertation is my own work and that no unauthorized aids were used.

I declare that all work, quotations and ideas from others are indicated. I allow this dissertation to be checked by appropriate software as part of an investigation into allegations of plagiarism.

No part of this doctoral thesis has been accepted or is currently being submitted for another degree or qualification at this university or elsewhere.

Pay Laurin Jessen   Bremen 02.12.2024