AFFECTIVE STATES DURING PROBLEM SOLVING

The Role of Feedback, Individual Performance, and Motivation

SANDRA BUDDE

Dissertation zur Erlangung des akademischen Grades Dr. rer. nat
Fachbereich 3– Informatik
Universität Bremen

Sandra Budde: *Affective States During Problem Solving,* The Role of Feedback, Individual Performance, and Motivation, © 2016

SUPERVISORS:
Prof. Christian Freksa, Ph.D
Prof. Dr. Christoph Hölscher

DATE OF COLLOQUIUM:
14. Sep 2015

To Justus and Jannis.

ABSTRACT

Individuals who perform a task to achieve a certain goal undergo different affective states, depending on whether they advance towards their goal or not. In this context, those individuals who are able to perform well on a task are confronted with significantly more goal-advancing situations than individuals who perform poorly. As a result, well-performing individuals can be considered to be exposed to more positive affective contexts than poorly performing individuals. However, individuals who differ with respect to performance outcome in a certain task are very likely to process available information quite differently. Along the same lines, individuals of different performance groups may evaluate an identical problem-solving condition with a different outcome. As a result, affective states relating to the same problem-solving context might also differ between individuals of different performance levels. Similarly, individuals whose goal is to achieve highly in a task might judge an identical problem-solving condition much more negatively than individuals who are less ambitious. Again, the resulting affective states are likely to differ between these two groups of individuals.

The main contribution of this thesis is a novel experimental task, namely *Luckless Mastermind*, that allows computer-aided empirical investigation of the above. In this modified version of the Mastermind problem-solving task, the game course can be predefined and is the same across individuals, independent of their actual performance. The robustness of these game courses is proven via an extensive-search algorithm. The task paradigm enables analysis of different affective states in terms of peripheral physiological response. The theoretical basis for this is elaborated upon in detail, as well as the specific task requirements for such a measurement. In the scope of this thesis, skin conductance response is utilized as the physiological signal. Based on *Luckless Mastermind*, an experimental setup has been developed and implemented to investigate three different hypotheses on affective states during problem solving.

The first hypothesis targets the objective valence of a problem-solving condition: Do individuals in general exhibit different affective states in a more goal-advancing situation compared to a goal-inhibiting one? The second hypothesis is centered around the fact that well-performing individuals are very likely to evaluate an identical problem-solving condition much differently than poorly perform-

ing individuals: Does individual performance make a difference in affective state toward a problem-solving condition of a certain valence? The third hypothesis focuses on the assumption that individuals who aim at achieving highly are likely to evaluate a certain problem-solving condition much differently than individuals who are not ambitious: Does motivation impact affective responses?

Analysis reveals that individual affective state depends on the particular valence of a problem-solving condition, i.e. whether this condition is goal advancing or goal inhibiting. Nevertheless, this effect is modulated by the general task performance of an individual. Well-performing individuals exhibit significantly different affective states in a positive compared to a negative condition. In contrast, affective states of poorly performing individuals remain unaffected by the valence of the condition. The overall motivation to achieve highly does not impact goal-related affective states during problem solving. Nevertheless, symptoms of higher alertness and arousal have been identified for highly ambitious individuals.

ZUSAMMENFASSUNG

Bei der Verrichtung mehrschrittiger und zielgerichteter Tätigkeiten zeigen Menschen eine Vielzahl von affektiven Reaktionen. Die damit zusammenhängenden affektiven Zustände werden u.a. dadurch bedingt, ob eine ausgeführte Aktion gewinnbringend war oder nicht. Die Möglichkeit innerhalb der gegebenen Bedingungen eine zielführende Aktion auszuwählen hängt jedoch stark von den eigenen Fähigkeiten ab: Menschen, die generell eine hohe Performanz bei der Erledigung einer Aufgabe haben, können dabei auf effektivere Aktionen zurückgreifen und sind dadurch auch häufiger mit einer Verbesserung der Situation hinsichtlich des Ziels konfrontiert. Hinzu kommt, dass diese High-Performer die aufgabenbezogene Information anders verarbeiten und bewerten als Menschen mit geringerer Performanz. In diesem Zusammenhang ist es wahrscheinlich, dass sich auch der individuelle affektive Zustand bei identischer Problemlösesituation in Abhängigkeit zur jeweiligen Performanz unterscheidet. Analog dazu kann auch die individuelle Motivation möglichst effizient zum Ziel zu kommen einen Einfluss auf die Bewertung der Situation und damit auf die affektive Reaktion haben. So kann die Bewertung einer ehrgeizigen Person sehr viel negativer ausfallen als die einer Person, die eine Aufgabe weniger ambitioniert verrichtet.

Der Beitrag, den diese Arbeit leistet, ist die Entwicklung der neuen computergestützten Untersuchungsmethode *Luckless Mastermind*,

die es ermöglicht die oben genannten Aspekte im Rahmen einer Problemlöseaufgabe empirisch zu untersuchen. Mit dieser modifizierten Variante des klassischen Mastermind-Problems kann der Verlauf des Lösungsprozesses im Vorhinein bestimmt werden, unabhängig von der individuellen Performanz der Testperson. Die Plausibilität bzw. Korrektheit des Spielverlaufs wird mittels einer algorithmischen Lösung, die auf erschöpfender Suche basiert, bewiesen. Der Spielaufbau ermöglicht die Analyse von affektiven Zuständen auf Basis physiologischer Reaktionen. Deren theoretische Grundlagen sind im Detail erläutert. Darauf aufbauend wird der Experimentalaufbau am Beispiel des Hautleitwertes implementiert, um die folgenden drei Fragen zu untersuchen:

1. Entscheidet die Valenz der Lösungsituation über den affektiven Zustand? D.h. hängen affektive Zustände allgemein davon ab, ob eine Testperson sich dem Ziel nähert oder nicht?

2. Unterscheiden sich diese affektiven Reaktionen je nach genereller Performanz der Testperson?

3. Zeigen ehrgeizige Personen eine andere affektive Reaktion als nicht ambitionierte, wenn sie mit der gleichen Situation konfrontiert sind?

Es zeigt sich, dass der generelle affektive Zustand von der Valenz der Problemlösesituation abhängt. Dieser Zusammenhang wird jedoch von der generellen Performanz der Person beeinflusst: Personen, die eine Aufgabe sehr performant lösen können, zeigen je nach Valenz der Situation auch unterschiedliche affektive Reaktionen. Bei Menschen mit niedriger Performanz wird dieser Unterschied nicht deutlich. Außer generellen Symptomen höherer Wachsamkeit bzw. Arousals, gibt es keine Unterschiede in der affektiven Reaktion bei ehrgeizigen Problemlösern gegenüber nicht ambitionierten.

ACKNOWLEDGMENTS

Writing this thesis would not have been possible without support of various kinds. I thank my primary supervisor, Prof. Christian Freksa, PhD, for providing enduring support during the last couple of years while allowing the greatest possible scientific freedom. This has enabled evolution of the development processes necessary to redesign and shape this thesis.

In addition, I thank my second supervisor, Prof. Dr. Christoph Hölscher. I very much appreciate his valuable feedback on the experimental setup as well as statistical analysis. In addition, discussing my work with his research group at ETH Zürich greatly helped sharpen my lines of argumentation. Financial support has been provided by SFB/ TR8.

During all the years of developing this thesis, Dr. Thomas Barkowsky has been an especially strong supporter as well as severe critic of my scientific work. Depending on which was needed to make the next necessary step, he provided the suitable kind of supervision. His professional and friendly support was of great help, especially when becoming a mother inevitably impacted thesis progress and finishing at all appeared almost impossible.

The people of CoSy provided important support of various kinds: a very good opportunity for discussions during the graduate seminar and in other meetings, as well as technical support (Alexander Mittelstorb), and organizational matters (Gracia Kranz). In particular, I would like to thank: Dr. Holger Schultheis, who provided valuable feedback for sharpening the topic of this thesis, as well as extensive advice on developing the empirical study. Without him, I would have gotten lost in weighing up one empirical approach against another; Dr. Arne Kreutzmann, for his professional support in proving the complexity of the developed algorithms as well as his friendly support in general; Julia Gantenberg, for convincing me of my ability to finish this thesis, especially when I was most in doubt; and Dr. Jan Frederik Sima for our conversations about the ups and down of interdisciplinary scientific work and how to deal with them.

Various people provided support by proofreading parts of this thesis and giving valuable feedback: a big thank you goes to Dr. Anne-Christin Kreutzmann, Dr.-Ing. Zoe Falomir Llansola, Dr. Jae Hee Lee, Dr.-Ing. Lutz Frommberger, Rasmus Wienemann, Dr. Tomi Kauppinen and Carl Schultz, Ph.D. In this context, I thank Dr. Tamara Hartke for the time and effort spent in polishing the language of this thesis, which has greatly improved its readability.

In addition, I would like to thank Carsten Fischer, Tobias Warden and Carsten Rachuy for their friendly support throughout the years.

Last but not least, the biggest thank you goes to Christian Ober-Blöbaum for his loving support over the years, as well as the time and effort he invested to create open space for me and my scientific work in our lively everyday life. Without his patience and appreciation, the times of intense developing and writing would have been much more difficult to withstand.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

APPR Affective Peripheral Physiological Response

CB Code Breaker

CM Code Maker

CPM Component Process Model

FPI-R Freiburger Persönlichkeitsinventar (revised version)

GUI Graphical User Interface

NSCR Number of Skin Conductance Responses

SCL Skin Conductance Level

SCR Skin Conductance Response

SCR AMP Skin Conductance Response Amplitude

SCR LAT Latency of Skin Conductance Response

SEC Stimulus Evaluation Check

1 INTRODUCTION

> *Human behavior flows from three main sources: desire, emotion, and knowledge.*
>
> — Plato

When individuals perform a task, they usually experience emotions that depend on how well they carry out that task. These *affective states* are highly subjective and therefore different individuals may experience the same task quite differently.

To illustrate this, let us consider a comprehensive task that needs several steps and some effort to be solved. Let us also consider two different individuals who are asked to perform that task. Individual I is highly motivated to perform well on the task. Accordingly, she is unhappy if she gets stuck in the process. She will feel challenged if she is confronted with single setbacks, but recurring setbacks will frustrate her. She will be happy if she experiences progress and will be delighted if she performs successfully. Individual II does not care much about the task and how well she performs it. Accordingly, she also does not care much about specific progress and will be happy if she solves the task without much effort. If she is confronted with difficulties and recurring setbacks, she will become bored and be annoyed by performing the task.

Depending on the particular affective states of individuals, the underlying process of task performance may proceed quite differently. Progress and success as well as setbacks and difficulties in performing a task are very likely to be modulated by individual affective states. To identify critical features that may separate good performers from poor ones, a detailed understanding of how affective states evolve while a task is performed, as well as how these states interact with task performance, is essential.

In particular, it is often assumed that good performers not only differ from poor performers with respect to pure cognitive processes and functions, but also with respect to their experienced emotions. For instance, it has been claimed that becoming a chess Grandmaster like Garry Kasparov not only depends on individual reasoning or pattern-recognition capacities but also on the motivation to win and individual emotional capacity, among others [Dai and Sternberg, 2004]. Accordingly, it is of great interest whether there are general differences in affective states that may distinguish good performers or experts from other individuals.

This affective link is crucial for more comprehensive tasks such as *problem solving*. Such a task is defined by the presence of a problem, a set of certain actions that can be performed, and particular goals that an individual wants to achieve to solve the problem (e.g. [Newell and Simon, 1972]). A problem-solving situation becomes more complex when a desired solution cannot be accomplished in a single step but requires several steps to solve the problem. At each step, certain actions are performed to finally solve the problem.

Depending on the individual's performance in solving a task, individuals may differ in selecting appropriate actions to achieve their desired solution. As a consequence, they may also encounter different conditions of progress and setbacks, which may inevitably result in individual affective states. As in the above stated example, individual characteristics such as motivation or general task performance can be assumed to interact with these affective states.

1.1 RESEARCHING AFFECTIVE STATES DURING PROBLEM SOLVING

Although the existence of progress-related affective states appears to be obvious, the underlying mechanisms have not yet been sufficiently investigated and are consequently not fully identified.

This is to some extent due to the lack of precise definitions for different affective phenomena, caused by the heterogeneity of affective sciences in which researchers from many different disciplines make their contributions (see [Frijda, 2008] for a review). Consequently, there are various competing definitions for phenomena of human emotions and applicable research instruments, each with specific requirements for a thorough analysis. As some requirements can be met more easily than others, some specific aspects of emotions are more intensively investigated.

For instance, the impact of long-lasting emotions such as *mood* is intensively studied. This is due to the wide acceptance of the view that individuals can be put into a certain mood experimentally (see [Clark, 1983] and [Mitchell and Phillips, 2007] for a review).

In contrast to this, *affective states*, which only persist for a few seconds and evolve while the individual is engaged in a certain task, are insufficiently studied. This is due, in part, to the fact that individuals are not necessarily able to describe them. In addition, asking an individual to report on affective states that accompany task execution requires either disturbing her during task performance or relying on retrospective, memorized reports of emotion after task completion. Both cases come with the risk that the provided information may be unreliable. For instance, difficulties in articulating individual affective states or personal reflections of how provided information will be perceived by the experimenter may impact self-reported affect [Picard et al., 2004].

Nevertheless, for developing a comprehensive theory of how emotions interact with problem solving, it is necessary to develop a more detailed understanding of the functions and mechanisms of these short-term affective states. In this context, it is highly important to determine how such short-term states can be experimentally manipulated to allow a precise analysis across individuals.

1.1.1 *Measuring Affective States*

To be able to analyze affective states that arise while solving a certain problem, it is necessary to reliably measure such states. These states do not occur at the end of a solving process but evolve continuously, especially when individuals monitor their progress toward a desired solution. A reliable measurement is complicated for the following two reasons. First, directly asking an individual would interrupt problem solving and might consequently interfere with the process. Second, individuals can only report on affective states that are consciously experienced, which is not necessarily the case for such short-term affective states.

Fortunately, short-term affective states are accompanied by patterns of somato-visceral responses, e.g. changes in heart activity, skin sweat distribution, breathing rhythm, and tension in certain skeletal muscles (see [Cacioppo et al., 2000] for a review). These patterns of physiological response occur in the context of other fluctuations of the physiological signals. These patterns of physiological response occur in the context of ordinary (non–task-induced) fluctuation. To analyze differences in affective physiological responses, it is thus necessary to reliably separate these responses from other signal perturbations. To do so, detailed knowledge about when to expect a specific affective state is essential. This is further complicated by the fact that affective states are elicited by internal processes whose timing is difficult to determine.

As discussed above, affective physiological responses arise any time individuals evaluate their progress towards a desired solution. Such evaluative processes can be triggered solely based on internal events managed by the individual, i.e. *intrinsically*, or triggered by external events, i.e *extrinsically*. In the former, the response is triggered solely by the individual's attempt to evaluate the outcome of planned actions before actually performing them. In this case, the time of occurrence of the resulting affective states is difficult to determine.

For the latter, an external event necessitates an evaluation of current progress with respect to a desired solution. For instance, an individual can be confronted with feedback about current progress or receive novel information about the problem that needs to be incorporated to judge the current performance and adapt behavior for proceeding with the task accordingly.

An example of a task involving feedback at each step is the Mastermind task, based on a two-player board game called *Mastermind*, which has been very popular since the 1970s. The goal of the game is to guess an arbitrary, four-digit code of colored pegs selected by one of the two players. This player, who is termed the *codemaker* (CM) [Knuth, 1977, p.1], selects the combination from six colors (multiple uses of colors is allowed) and hides it from the second player. This player, the *codebreaker* (CB) [Knuth, 1977, p.1], tries to break the code by providing a guess in terms of a four-digit code to CM. CM provides feedback for this guess using black and white pegs: the number of black pegs indicates how many digits are already of the correct color and in the correct position, while the number of white pegs indicates how many digits are of the correct color but in the wrong location. The feedback does not specify which peg corresponds to which element in the code. This cycle of guessing and receiving feedback is repeated until CB breaks the code (CM provides feedback of four black pegs) or a maximum number of steps is reached. An example Mastermind game is presented in Figure 1 on the following page.

In this game, proceeding with the next step is only possible after feedback has been received with respect to the current step. Based on the previous considerations, it is assumed that individuals who are confronted with performance-related feedback perform self-reflective evaluation, which results in an affective state. Accordingly, by measuring at which point in time feedback is provided, it is possible to determine when to expect a certain affective state as well as an affective physiological response.

1.1.2 *Analyzing Affective States*

As presented above, affective states can be analyzed by investigating their corresponding affective physiological responses. These responses are assumed to share certain features across individuals when they are experiencing the same affective state [Ekman et al., 1983]. For

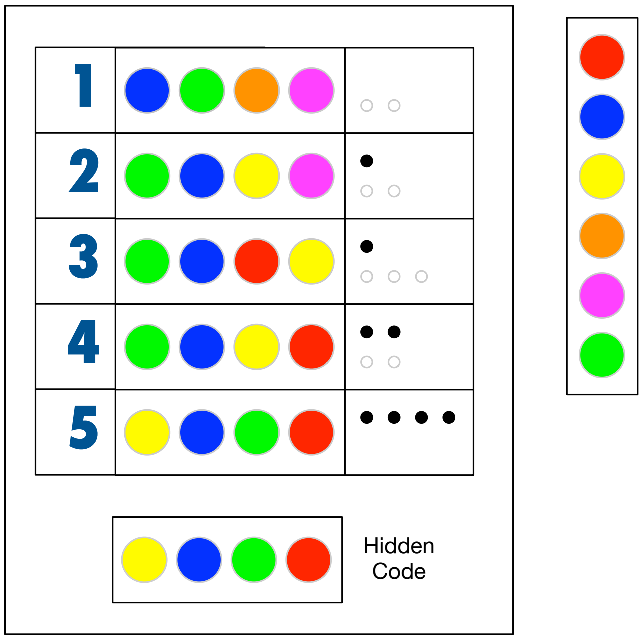

Figure 1: An illustration of a Mastermind game using colored pegs.

instance, when different individuals are confronted with a negative situation, such as general failure in solving a task, they undergo a negative affective state and exhibit a comparable affective physiological response. Along the same line, in a condition of ultimate successful performance, a positive affective state is experienced and a corresponding physiological response pattern is exhibited. In this context, the term *valence* is used to quantify how positively or negatively a person feels.

This situation may differ in the case that the valence of an affective context is not distinct but needs individual interpretation. In such a case, it depends on individual evaluative processes to judge the affective connotation of a situation. Accordingly, individuals may differ with respect to their affective states. As affective states are also subject to modulation by individual factors such as motivation, it is plausible that there is also an interaction between the two.

To test whether characteristics of the individual impact affective states during problem solving, it is necessary to provide an affective context of identical valence in such a task. Only if individuals are confronted with the very same context in a problem-solving task can affective physiological responses be compared across individuals. For this, a controlled manipulation of valence is needed.

Manipulating Affective Context

In a problem-solving task, valence of an affective context is strongly determined by the individual's progress. Negative valence is determined by little or no progress toward a solution, whereas positive valence results from increasing progress toward it.

Individual progress in such a task always depends on general task performance. Individuals who perform poorly usually experience negative valence while individuals who perform well are more likely to exhibit a positive one. Based on this, achieving identical conditions of valence for all individuals is not trivial.

In task settings that consist of only a single step, identical affective context is usually achieved by providing all individuals with identical feedback. Depending on the general task performance of an individual, this preselected feedback may not match actual progress. For instance, for an individual who performs well on a task, negative feedback mismatches her actual performance. If the individual is able to form specific expectations about individual progress before feedback is received, she will be likely to detect this mismatch and become confused. As a result, she does not experience the same negative valence as a poor performer.

To overcome this limitation, tasks are usually selected that prevent the individual from evaluating performance before feedback is received. For instance, individuals are confronted with a set of visual objects and must determine how often a certain element is contained in the set. By limiting the available time for examining the set to an interval of a few milliseconds, it is ensured that individuals cannot form a sound expectation about their own performance; actual performance is concealed from the individual. As a result, interference between expected outcome and received feedback are eliminated and identical conditions of valence can be achieved through proving identical feedback to all performing individuals.

This approach does not work for comprehensive problem-solving tasks. To successfully solve a comprehensive problem, individuals need to deliberately draw inferences on the provided information. If such information is restricted to conceal real performance from an individual, inference is not possible. Accordingly, task progress does not rely then on drawing inferences but transforms decision making into a gambling task in which other affective states are assumed to arise (cf. somatic marker hypothesis [Damasio, 2005]). It remains an open question how affective contexts can be manipulated in a comprehensive problem-solving task without violating a sound problem-solving process.

1.2 theses

In general, it is assumed that there are stable dependencies between specific progress in a task and the associated affective states an individual undergoes. Within this dissertation, the analysis of affective states during problem solving is focused on three main theses:

1. Depending on the experienced valence of task progress, an individual exhibits different affective states.
2. The specific experience of task progress may be subject to individual evaluative capabilities. As a result, individuals who perform well on a task may judge the valence of identical progress differently than individuals who perform poorly. Accordingly, depending on their general task performance, individuals exhibit different affective states.
3. These affective states also differ with respect to individual motivation to be successful, so that individuals who want to achieve higher levels of success exhibit different affective states than other individuals. Accordingly, depending on their general motivation to be successful, individuals exhibit different affective states.

To investigate these theses it is necessary i) to provide identical problem-solving conditions with a certain valence and ii) to reliably assess the exhibited affective states. Based on this, the general research problem of this dissertation is:

> How to induce and measure short-term affective states during a comprehensive problem-solving task?

1.3 approach: measuring and analyzing affective physiological responses

The approach taken in this dissertation is based on measuring and analyzing patterns of affective physiological responses that occur while individuals perform a problem-solving task.

As already presented, specific requirements have to be met before an extensive study of these affective states is possible. First, timing of the occurring affective states has to be reliably controlled. Second, affective contexts have to be purposefully manipulated to allow for comparison of affective physiological response. In this context, it is highly important to provide conditions of the same valence to all individuals.

These requirements are analyzed in detail for affective physiological responses. The original Mastermind task is introduced as a potential task setup that fits almost all of these requirements, however it lacks the possibility to manipulate the valence conditions. In addition, individuals can be lucky in this game, i.e., they identify a solution or a significant part of it by pure chance.

To remedy these problems, *Luckless Mastermind* is developed, which allows the experimenter to determine particular sequences of specific valence conditions that are valid for all individuals. This is achieved by playing Mastermind with a cheating CM who alters the solution unnoticed by CB and without violating any previous step.

To implement these changes, this dissertation introduces a formal representation of the task which allows analysis of the game with respect to general characteristics, enabling the experimenter to predetermine valence conditions. An algorithmic approach is then pursued which resolves the following two issues: First, the algorithm identifies all sequences of affective contexts that do not violate a sound problem-solving process. Second, it is utilized to determine upper thresholds for preselecting affective contexts, i.e. maximum number of affective contexts that can be manipulated depending on specific characteristics of the game. Based on this, *Luckless Mastermind* is implemented.

To test the above hypotheses of affective states during problem solving, an empirical study is designed. For this, *Luckless Mastermind* is presented on a graphical user interface (GUI) fitting specific requirements that support a reliable measurement of these response patterns. Utilizing this implementation, skin conductance responses are collected for 24 subjects. To allow a statistical analysis of affective responses, particular signal parts are separated from the complete sensor readings. The resulting signal parts are preprocessed to extract the particular features of affective responses. Based on this, these features of affective physiological responses are evaluated statistically to test the three previous theses.

1.4 CONTRIBUTIONS

The main contributions of this dissertation are:

- Identification of requirements for analyzing sequences of affective states during a comprehensive problem-solving task.

- Design of the corresponding problem-solving task, termed *Luckless Mastermind*, that

  - enables measurement of affective states in terms of physiological response and
  - provides identical feedback conditions across individuals of different performance groups without violating a sound problem-solving process.

- Identification of boundary cases for *Luckless Mastermind* games that provide maximum flexibility for preselecting feedback conditions.
- Development and implementation of an algorithmic solution for determining valid feedback sequences that are independent of an individual's actions.
- Implementation of *Luckless Mastermind* to provide preselected feedback conditions, then proceed as the common Mastermind task for any additional step.
- Implementation of GUI for playing *Luckless Mastermind* that supports measurement and analysis of patterns of affective physiological responses, i.e. skin conductance activity, of affective states to identify important features that distinguish different affective states.
- Design and execution of an empirical study that
  - utilizes skin conductance readings as measures of affective states, and
  - allows investigation of affective states with respect to
    * whether they distinguish different feedback conditions,
    * whether they differ in general across different performance groups, and
    * whether individual motivation represents a moderating factor.
- Separation of affective physiological responses from physiological signal stream and extraction of particular features of affective skin conductance responses.
- Statistical analysis of derived features to test the three theses.

1.5 OUTLINE OF THE THESIS

The remainder of this thesis is organized as follows: In Chapter 2, different phenomena of human affect and the definition of problem solving as used in this thesis are introduced. Based on this, potential interactions of different affective states and problem solving are presented. In this context, task-related short-term affective states are identified as one important remaining issue for thorough investigation.

In Chapter 3, basic challenges in measuring affective states in the context of performing a problem-solving task are presented. In addition, the state of the art in measuring and analyzing these affective phenomena is described and analyzed.

Based on previous definitions and open issues, in Chapter 4 the requirements for measuring and analyzing affective states during a comprehensive problem-solving task are elucidated. In this context, an exemplary task is introduced that fits these requirements to a large extent. Based on this, the theoretical considerations for the three previously mentioned theses are presented in more detail. In this context, characteristics of affective physiological response in skin conductance are introduced.

In Chapter 5, a modified version of the previously identified task, termed *Luckless Mastermind*, is developed to allow provision of equal feedback conditions across all possible game courses. To achieve this, formal requirements of the games, as well as specific characteristics of feedback sequences, are devised. To determine fixed game courses, i.e. preselected feedback sequences, an algorithmic approach is developed and presented.

Utilizing the previously developed task, several hypotheses about affective states and their relation to problem solving and certain feedback conditions can be investigated, as well as the potential interaction of motivation and affect. Accordingly, Chapter 6 is dedicated to the design, implementation, and execution of an empirical study to investigate the previously proposed hypotheses. This also comprises special routines for data reduction and analysis that are necessary when affective physiological responses are considered. Results are presented with respect to each previously developed hypothesis.

Finally, Chapter 7 provides a summary of this dissertation, as well as an outlook on further research which may build upon this work.

Part I

THEORETICAL BACKGROUND

2 AFFECT AND PROBLEM SOLVING: AN IMPORTANT DYAD FOR EFFICIENT COGNITIVE PROCESSING

When humans manage their everyday lives, it is always in relation to their individual goals. Such individual goals can rarely be achieved immediately, but are blocked by obstacles. To avoid these obstacles, individuals initiate problem-solving processes which allow them to perform certain actions to achieve an individual goal.

In the past, cognition and affect have been viewed as two distinct systems of human intellectual functioning that generally disrupt each other. This view has dramatically changed (see [Oatley and Jenkins, 1996] for a review) and there is now a consensus that human affect significantly modulates and interacts with successful cognitive processing (see, e.g. [Damasio, 2005]). This affective link is crucial for problem solving, as affect might impact whether individual goals can be achieved or not.

In the following section, I introduce the cognitive process of problem solving. Due to the variety of problems individuals can face when attempting to attain a goal, there are different types of human problem solving. After a short introduction to these different classes of problem solving, I focus on well-defined problem-solving tasks. Based on this, I argue that it is important to consider this class of problem solving when analyzing its interaction with affect.

Subsequently, I present in more detail the broad range of phenomena denoted by the term *affect*. I then elaborate on the ways problem solving interacts with different phenomena of human affect. In this context, I will discuss the lack of studies on achievement-related affect and its interaction with problem solving.

2.1 HUMAN PROBLEM SOLVING

When individuals encounter a situation that does not match their current individual goals, they try to perform specific actions for changing from an undesired state into a desired one. This process is termed *problem solving*.

2.1.1 *The Problem Solving Cycle*

In psychology, the process of problem solving has been described in terms of a seven-stage cycle [Bransford and Stein, 1993, Hayes, 1989, Sternberg, 1986]. During this cycle, individuals organize their knowledge and develop strategies for efficiently solving a certain problem. Although the cycle consists of stages that are depicted sequentially, it is unlikely that individuals adhere rigidly to this sequential order. Rather, steps may be skipped or repeated again until a desired solution is achieved.

1. Recognition or identification of the problem
   Identifying the problem appears to be a trivial prerequisite for solving a problem. Nevertheless, the basis of inefficient problem solving lies in an incorrect identification of the problem.

2. Definition and representation of the problem
   To be able to solve the problem at all, it must be appropriately defined and represented. In the case of an inappropriate problem representation, the solving process will be inefficient, independent of the efficiency of any subsequent solving step.

3. Develop strategies for solving the problem
   The individual forms a specific strategy to tackle the present problem, drawing on accumulated knowledge and feedback from previous attempts to solve the problem.

4. Appropriate organization of knowledge about the problem
   Before actually solving the problem, all available knowledge must be aggregated and organized in an appropriate manner. The more information an individual has with respect to the present problem, the better prepared she is for efficiently solving it.

5. Allocation of resources for solving the problem
   In order to actually solve the problem, the corresponding cognitive resources necessary to perform the task must be allocated. Depending on the priority of the task, the amount of allocated resources may vary. For high priority tasks, more resources are allocated compared to low priority tasks.

6. Monitoring of individual progress towards the goal
   For detecting ineffective actions during problem solving, current progress towards the goal must be monitored. In the case of ineffective actions, there is a re-evaluation of the approach and new strategies are formed.

7. Evaluation of the result
   After a solving process has finished, its result is checked to determine whether it matches the best possible solution. Depending on how the quality of a solution is defined, evaluating a solution differs. For instance, because the result of an arithmetic task can either be true or false, the individual solution is also measured with respect to these two categories. In the case of a more complex evaluation of a solution, such as how well received a written essay is, the corresponding evaluation by the individual is also more complex and depends on an individual selection of criteria.

There are two different classes of problems that can underly this process: *ill-defined* and *well-defined problems*.

2.1.2 *Ill-defined vs. Well-defined Problems*

Ill-defined problems [Reitman, 1964] are characterized as being unspecific with respect to particular parts of the process. For instance, in such a problem, both the initial situation and a potential goal are not clearly defined. This is also the case for the actions that can be selected to transform one situation into another [Robertson, 2001]. As an example of this kind of problem solving, let us consider the task of preparing a delicious meal. Although most adults are able to tackle this task, the result and how it has been achieved differ greatly from one individual to another. In addition, the evaluation with respect to deliciousness depends on individual taste and preferences. In this sense, all problems that involve creativity or dealing with ambiguities to solve them fall into the category of ill-defined problems.

In contrast to this, there are *well-defined problems*, which can be precisely described with respect to their initial situation, their individual goal, their expected resulting situation and a defined set of actions to be used for transforming one situation into another [Robertson, 2001]. One common example for such a task is playing a game of chess. In this game, the initial situation is defined by a certain configuration of chess pieces. The goal is to checkmate the opponent. To achieve this, there are clearly defined actions given by chess rules.

More formally, a well-defined problem has a certain *problem space*, consisting of an *initial state*, a *goal state* and a potential number of *intermediate states*. One state is transformed into another by applying a *goal-directed sequence of cognitive operations* [Anderson, 1980, p. 257]. For applying such operations, a representation of the current situation has to be generated that meets three criteria [Simon, 1999]: (1) it describes the given situation, (2) it contains the set of operations that can be performed during this situation and (3) it provides a set of tests to determine whether or not a goal has been reached.

Applying operations during a given situation results in the creation of subsequent states on which other operations can be performed. Considering the application of any possible operation on any corresponding state results in a branching tree that represents all achievable situations: the *state-action space* [Robertson, 2001] which *represents an omniscient observer's view of the structure of a problem* [Kahney, 1993, p. 42]. As human mental representations are likely to be subject to noise and incompleteness and do not cover all theoretically reachable states [Newell and Simon, 1972], the problem space covers a subset of the state-action space.

Solving a well-defined problem is defined as searching the problem space by evaluating sequences of operation and their corresponding transformed states, with respect to the individual goal. Due to the combinatorial explosion of subsequent states, the evaluation of all possible sub-states exceeds human cognitive resources. Therefore, individuals rely on specific *heuristics*, i.e. search strategies, to traverse a problem tree more efficiently. Based on a proper selection of strategies, expertise in a certain problem domain can be defined as *the acquisition of knowledge that restricts the need for extensive search* [Holyak, 1995, p.271].

2.1.3 *Problem Solving in This Thesis*

The remainder of this thesis is focused on analyzing affective states while solving well-defined problems. There are two main reasons for this.

First, analyzing cognitive processes while solving ill-defined problems is difficult due to the intrinsic ambiguity of such problems [Pretz et al., 2003]. Because of this, current research on how affect modulates problem-solving behavior focuses on analyzing well-defined problems. In contrast to ill-defined problems, in well-defined problems all parts of the problem-solving process are clearly defined; the inherent structure of well-defined problems facilitates identification of potential differences in cognitive processing that are modulated by affect.

As a result, behavioral differences can be more easily analyzed with respect to differences in the problem-solving process or the underlying structure. Such well-defined tasks allow the creation of specific models of problem-solving behavior which can be compared to behavioral data of individuals [Belavkin, 2001], facilitating analysis of a problem-solving process with respect to affect-modulated differences.

Second, although everyday life appears to be dominated by ill-defined problems [Neisser, 1976], efficiently solving well-defined problems is important for academic success at school [Pretz et al., 2003]. Apart from writing an essay, taking an exam in a certain topic is most often defined by solving well-defined problems. In such a setting, the initial situation, the individual goal, and the operators to transform a certain situation into another are clearly defined. Performance is most often measured as whether a certain heuristic has been learned and reproduced or not. In this respect, the impact of individual affective states becomes highly evident when intense affective states such as anxiety are considered. In addition, achievement-related affective states, such as recovering from setbacks, can be assumed to occur [D'Mello et al., 2010].

2.2 HUMAN AFFECT – A BROAD RANGE OF PHENOMENA

Affect is used as an umbrella term for several different phenomena related to human emotion. Based on this, the term *affective state* represents the superordinate category of all different emotional conditions an individual can undergo. Examples include emotions, feelings, and moods [Clore et al., 2000,Scherer and Peper, 2001,Sloman et al., 2005].

In general, affective states are defined as temporary conditions of the human organism that reflect the representation of personal value involving multiple systems, such as cognitive and physiological systems. In this context, the term *affective* is contrasted to the term *cognitive*: whereas *cognitive* describes the binary representation of knowledge in the human mind (information can be true or false), *affective* adds a personal qualification of positivity or negativity to categories of this knowledge (see [Clore et al., 2000, pp. 29]).

In contrast to this broad definition, use of the term affect is often restricted to refer to the internal state of an individual (including mental as well as physical aspects) with respect to a specific affective phenomenon like having an emotion [Bless and Fiedler, 1995] or feeling [Diener and Oertel, 2006], or being in a certain mood [Rolls, 2000]. To avoid confusion between term usage when describing affective phenomena of differing characteristics, Picard differentiates *emotional states*, referring to the internal dynamics of having an emotion, from the superordinate concept of *affective states* in general [Picard, 2000].

In addition, the term *affect* is predominantly used to describe affective states resulting from internal evaluation of specific stimuli or parts of stimuli as having a certain affective connotation to oneself. In contrast to this internal evaluation of external stimuli, affect can also arise from an evaluation of one's own performance with respect to a certain task. As a result, individuals undergo specific affective states that are linked to their individual achievement. Based on this, it is important to distinguish affective states depending on their cause. Here, affective states that are related to a specific event or stimulus are contrasted to affective states that result from the evaluation of individual achievement with respect to a certain task. In the following section, different affective phenomena are presented while considering the distinction of stimulus-related and achievement-related affective states.

2.2.1 *Stimulus-related Affective States*

There are varying definitions for the terms denoting a specific state or process of human affect that is related to certain affective stimuli, e.g. *emotion*, *feeling*, and *mood*, and researchers within the field of affective science often use terms interchangeably. Nevertheless, a distinct definition of these different terms is important, as each of these phenomena is very likely to involve different neurophysiological mechanisms [Scherer, 2000].

Emotion

The term *emotion* is often used to describe affective states related to a specific situation or object [Clore et al., 2000]. Emotion is defined as *"an episode of coordinated changes in several, temporarily coupled, components (including cognitive processes, intraorganismic adaptation and regulation of somato-visceral systems [...]) in response to external or internal events of major significance to the organism"* [Scherer and Peper, 2001, p. 19]. Significance is often defined in relation to individual goals: emotions are elicited by evaluating an event as relevant with respect to an individual goal; the advancement of the goal results in a positive emotion, the inhibition of the goal in a negative one [Oatley, 1999]. With respect to this, situations and objects can function as instrumental reinforcers. Depending on their potential for the advancement of a goal they are categorized as rewards or punishers of a particular function(see [Rolls, 2005, p. 36]). By investigating the brain systems that are involved in the elicitation of a certain emotion, two different types of emotions can be identified: *primary* and *secondary emotions* [Damasio, 2005, p. 131]. Primary emotions are associated with bodily responses that are caused by innate response reflexes to exciting stimuli and follow the Jamesian view, which defines emotions as a pure bodily phenomenon that does not involve much mental processing [James, 1884]. Such emotions, which are most often connected to fearful stimuli, do not need higher cognitive processes to recognize the fearful stimulus but instead depend on early sensory cortices for detection and categorization, which then triggers an immediate response in certain brain areas; their most important neurophysiological areas are the limbic system circuits, the amygdala, and the anterior cingulate. The term *secondary emotions* extends the range of emotional phenomena by mechanisms to explain emotional responses resulting after voluntary and nonautonomic mental evaluation of a situation. Through experience, humans individually evaluate objects or situations and bind them to bodily responses of specific primary emotions. Therefore, secondary emotions form *"systematic connections between categories of objects and situations, on the one hand, and primary emotions, on the other"* [Damasio, 2005, p.134]. One important characteristic of emotions is their *episodic nature* [Frijda, 1986]. The emotional episode which is triggered by an external event lasts for a certain duration then fades away with decreasing intensity (see [Scherer, 2000, p.138]). It is often associated with perturbation of the autonomic nervous system [Oatley, 1999]. In sum, emotions are relatively brief episodes of coordinated changes of brain, autonomic, and behavioral changes that enable humans to respond to an internal or external event of significance (see [Davidson et al., 2003]). They are not necessarily consciously perceived by the individual.

Feeling

In general, the term *feeling* is used for the conscious and subjective experience of an emotion [Clore et al., 2000, Picard, 2000, Damasio, 2005]. Based on the underlying theory of how emotions arise, definitions slightly differ.

On the one hand, an affective state is defined as feelings if an individual is alert and aware of changes in the body state in the context of specific objects or a situation and therefore the emotion becomes consciously experienced [Damasio, 2005]. Thus, feelings are a result of continuously monitoring bodily changes. In this sense, the term is used synonymously with the term *sentiment* [Scherer and Peper, 2001]. On the other hand, feelings are also defined as consciously available internal signals providing feedback from non-conscious affective, bodily, or cognitive processes [Clore et al., 2000]. Based on this, feelings may not only arise solely from monitoring bodily changes but may also result from affective or cognitive processing.

Both definitions agree on a feeling being an emotion that is consciously perceived by an individual and which arises in conjunction with bodily changes.

In contrast to this, [Damasio, 2005] proposes the term *background feelings* to describe body states that occur in the absence of emotions. They are defined as the *"body state prevailing* between *emotions"* [Damasio, 2005, p.151]. Individuals can consciously report about these kind of feelings, although they are subtle, of low intensity, and represent the "basic" body state in the absence of emotions. Due to their longer duration and lack of specificity with respect to a certain eliciting stimulus, they fit the definition of the affective state termed *mood*.

Mood

Mood describes affective states that occur in the absence of eliciting objects [Clore et al., 2000]. They are diffuse and of low intensity, but have a relatively long duration without focusing on a specific cause, like a situation or an object. Mood can be defined as existence of individual background feelings that are uniform over a longer period of time, i.e. hours and days [Damasio, 2005]. In addition, there are usually no expressive response patterns that can be associated with a specific mood [Davidson et al., 2003].

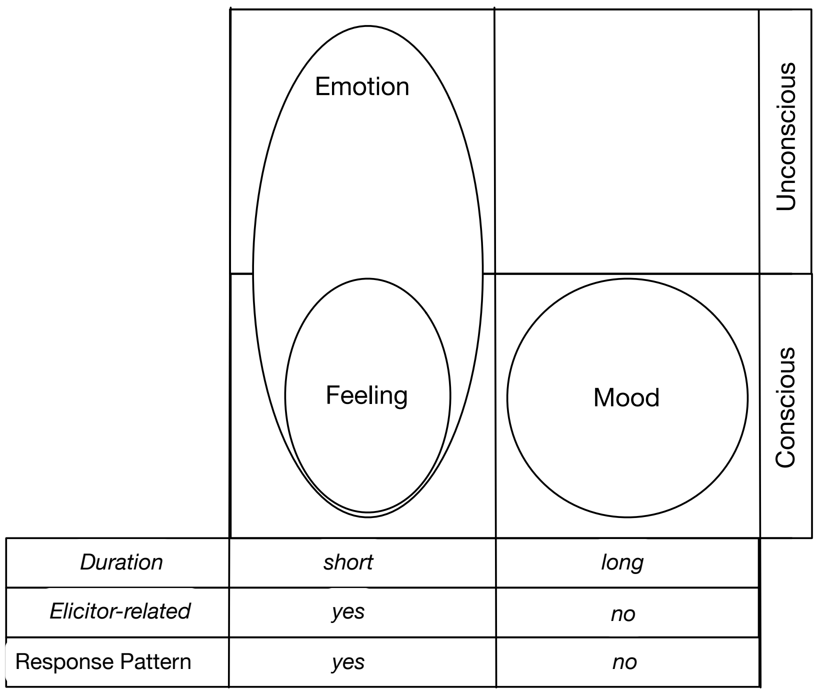

Figure 2: Categorization of stimulus-related affective states depending on their duration, whether they are consciously perceptible, their relation to a specific elicitor, and whether they become evident in a specific response pattern.

In summary, the term *emotion* describes all short-term affective states lasting for a small number of seconds which result from a specific elicitor and which become manifest in certain response patterns; they might not necessarily be consciously perceptible by the individual. *Feelings* denotes consciously perceived emotions. In contrast, *moods* are longer-lasting, i.e minutes to days, non-specific affective states that do not become evident in certain response patterns. For an illustration, see figure 2 on the previous page.

Categorizing Affective States

There are two approaches to categorizing the above affective states. First, affect can be described in terms of different dimensions and defined by quantifying it with respect to these different dimensions. Theories of affect categorization discriminate between two [Russell, 1980] or three [Mehrabian, 1980] different dimensions. The two categories common to both are *valence* and *arousal*. The measure of valence quantifies how positive or negative a person feels. Arousal measures the level of excitement of an individual. Theories considering a third category for defining affect introduce intensity *intensity* to quantify how much an affective state overwhelms an individual.

In the second approach, it is assumed that there are only a limited number of discrete and universal affective states that are stable across individuals and cultures. As a result, affect is defined using different terms for so-called basic emotions. Different authors suggests different affective states as being basic, although most authors agree on terms such as happiness, surprise, sadness, anger, fear and disgust [Ekman et al., 1972, Ortony and Turner, 1990].

Incidental vs. Integral Affect

To more specifically define a certain affective state, especially with respect to its elicitor, a distinction has been proposed between *integral* and *incidental* affective states [Blanchette and Richards, 2012]. In this classification, affective states are differentiated depending on whether their eliciting stimulus has been part of the underlying experimental task or not. For instance, in affective science, it is common to induce a specific affective state in an individual by presenting a particular stimulus which is assumed to have a certain affective connotation. Afterwards, the individual is presented with another experimental task that lacks this specific affective content and behaviour during task execution is analyzed with respect to the previously induced affective state. In this case, the underlying affective state is termed *incidental* as it results from an affective incident separate from the actual task.

Alternatively, an affective state can also be induced by the task material itself. For instance, individuals may perform a reasoning task on rather neutral material, like estimating how many people bought a cell phone in a certain time interval, or they can be confronted with a more emotional condition, such as estimating how many people died of a specific cancer in the same time interval. In the latter case, the assumed affective state results from an integral part of the task and is accordingly termed *integral affect*.

As the distinction of incidental and integral affective states is solely focused on affective content of task material, it does not account for affective states that occur while performing a specific task which is inherently non-affective. Such affective states are called *achievement emotions*.

2.2.2 *Achievement Emotions*

In general, all affective states that arise based on an individual's evaluation of an actually performed task or the respective individual's progress on a task are termed *achievement-related affective states* . Achievement in this sense is defined as *quality of activities or their outcomes as evaluated by some standard of excellence* [Pekrun et al., 2007, Heckhausen, 1991, p.15].

There are two different types of achievement-related affective states, distinguished by the focus of this internal evaluation. Affect can either arise based on i) individual progress and setbacks or ii) how one experiences performing the task itself [Pekrun, 2006]. To illustrate the latter aspect, consider an individual suffering from math anxiety who is asked to solve an arithmetic problem. In this case, it is safe to assume that the individual would experience certain affective states, primarily focused on the task itself, which are very likely to interfere with subsequent performance-related affect [Pekrun, 2006].

The definition of task-related affect is relatively new and is supported by educational psychology [Pekrun et al., 2007], based on an emerging comprehension of affect's important role in academic success. Nevertheless, earlier work on affective states resulting from success and failure [Weiner, 1985] (see [Nummenmaa and Niemi, 2004] for a review) provides additional empirical evidence for an outcome-related aspect of affect.

Achievement emotions are categorized by *valence*, *object focus*, *temporal focus* and *activation* [Pekrun et al., 2007]. As with stimulus-related affective states, *valence* discriminates achievement emotions with respect to positivity and negativity.

With respect to their *object focus*, achievement emotions are characterized as either primarily related to a specific activity or based on the outcome of an activity. In this way, *activity emotions* are distinct from *outcome emotions* [Pekrun et al., 2002, Maier, 2006].

Based on their temporal characteristics, *state achievement emotions* can be discriminated from *trait achievement emotions* [Pekrun, 2006]. Whereas the former are conceptualized as temporary occurrences in a certain situation at a specific point in time, the latter are habitual, recurring emotions that are experienced with respect to a specific achievement activity and its outcome. However, both types of emotions can occur in the same situation. To illustrate this, let us refer to the previous example of an individual suffering from math anxiety. In this specific example, it is safe to assume that the whole process of task execution is accompanied by a long-lasting negative affective state, i.e. trait achievement emotion. Nevertheless, at specific points in time, there are short-termed affective states which are related to outcome of the solving process, i.e. state achievement emotions.

Activitation categorizes achievement emotions based on whether they activate or inhibit further activities [Pekrun et al., 2007, Linnenbrink, 2007]. In this respect, the specific valence of an affective state is not decisive for its inherent activation; there are both positive and negative activating as well as deactivating achievement emotions. For instance, enjoyment, hope, and pride as well as anger, anxiety, and shame are viewed as activating affective state as they are assumed to increase motivation to actively alter a performance situation. In contrast to this, relief and relaxation as well as boredom and hopelessness are viewed as deactivating affective states.

2.3 AFFECTIVE STATES DURING PROBLEM SOLVING

As seen through the previously presented types of affect, problem solving is accompanied by several different affective states with varying temporal characteristics and focus on an eliciting object or situation. First, problem solving occurs in the presence of incidental affective states that are not task-related but result from external objects or situations. Second, there are achievement emotions that result from the task itself.

2.3.1 *Incidental Affective States*

Any problem-solving activity is framed by individual affective states that are independent of the actual task. To systematically analyze how these incidental affective states impact problem-solving behavior and performance, individuals are intentionally put into a certain affective situation and their task bebavior is compared across different affective conditions. As affective states and their variations are idiosyncratic, i.e. they are specific for an individual but not necessarily across a group of individuals, it is not safe to assume that a certain affective situation reliably elicits a certain affective state [Rottenberg et al., 2007]. Therefore, it is necessary to check whether an intended affective state has been present during task performance. To assess whether a certain affective state persisted during task performance, self-report measures are utilized to check whether an induced affective state is still present after task completion. Based on this methodology, long-term affective states, i.e. moods, are analyzed.

Analysis of short-termed affective states is hindered by the fact that any affect induction would interrupt the actual problem-solving process, as would a check for successful affect induction. In addition, by forcing an individual to report on their current affective states, the impact of these states on cognitive processes might be reduced [Keltner et al., 1993]. For these reasons, research has focused on analyzing long-term incidental affective states and their impact on problem-solving behavior.

Long-term Affect: Mood Effects

There are numerous studies on how mood effects cognitive processing. With respect to problem solving, the most important aspects are i) subjective differences in evaluative judgements, ii) modulation of individual processing style, and iii) interference with recall from memory [Schwarz and Skurnik, 2003].

First, positive mood supports a more optimistic view on individual resources and potential obstacles hindering successful performance [Brown and Mankowski, 1993,Johnson and Tversky, 1983] when compared to individuals in a negative mood. Based on this, it is assumed that individuals in a positive mood are more likely to tackle a difficult problem at all, as they evaluate themselves as more likely to reach the goal [Schwarz and Skurnik, 2003].

In addition, individuals in a negative mood show a tendency to set more ambitious goals for themselves. Together with their more pessimistic evaluation of individual potential for achieving a goal, inhibition in tackling a certain problem-solving task can be assumed [Schwarz and Skurnik, 2003].

When performing a certain problem-solving task, individuals in a negative mood are more likely to evaluate their individual solution more negatively than individuals in a positive mood [Hirt et al., 1996, Martin et al., 1993]. Based on this, individuals in a negative mood who have reached a valid but suboptimal solution are more likely to keep searching for a more optimal solution. In contrast, individuals in a positive mood stick to a suboptimal solution and are content with it. In this context, mood also impacts subjective evaluation of overall outcome. Whereas individuals in a positive mood are content with their suboptimal solution, individuals with a negative mood remain dissatisfied [Schwarz and Skurnik, 2003].

Second, mood modulates cognitive processing style in general. For instance, individuals in a positive mood are able to perform more flexible, conceptually driven, top-down processing [Clore and Huntsinger, 2007, Fiedler, 2001, Isen, 2008]. In contrast, individuals in a negative mood are more likely to perform narrowly focused and stimulus-driven bottom-up cognitive processing [Bless and Fiedler, 1995, Hertel et al., 2000, Schwarz, 1990, Schwarz, 2000]. With respect to problem solving, it significantly depends on the underlying task whether a certain mood hinders or facilitates problem solving. In the case of a task that requires more flexible and creative processing, task performance is increased for individuals in a positive mood compared to performance of individuals in a negative mood. In contrast, performance is higher for individuals in a negative mood when more detail-focused, bottom-up processing is needed to successfully perform a task. Similarly, individuals in a negative mood show higher performance in analytical reasoning tasks [Fiedler, 1988, Melton, 1995]. Nevertheless, such mood effects are ruled out if individuals are explicitly made aware of the fact that their current mood is the result of a task-independent and irrelevant source [Sinclair et al., 1994].

Third, individual mood modulates which information can be retrieved from memory. In this respect, it is more likely to retrieve material i) whose affective tone is congruent with current mood and ii) which has been memorized when having been in the same mood [Isen et al., 1978, Bower, 1981, Bower, 1991]. Based on this, performance in problem solving is improved if a task's affective tone matches current mood. For instance, if the same problem is extended by irrelevant information of different valence, individuals expect a more promising outcome in the case of a match between affective tone and individual mood [Hesse et al., 1997]. These differences in evaluation are solely based on matching affective characteristics and do not result from a mood modification due to the material as no impact of the material's affective tone on individual mood has been found.

2.3.2 *Achievement Emotions*

Performing a problem-solving task is inextricably tied to occurrences of affective states that result from interacting with the task itself. There are three different types of affective states, which differ in terms of the particular aspect of the solving process eliciting them, as well as in their temporal characteristics. First, problem solving can be accompanied by a long-term affective state representing subjective attribution of a task and its material. Second, subjective attribution of a task's material can also result in affective states with a rather limited duration. Third, as problem solving is extricably bound to achieving individual goals, any outcome evaluation in this respect results in a short-term affective state. In the following section, these three different types of affective states and their potential interactions with one another and with problem solving are presented in more detail. This framework is illustrated in figure 3.

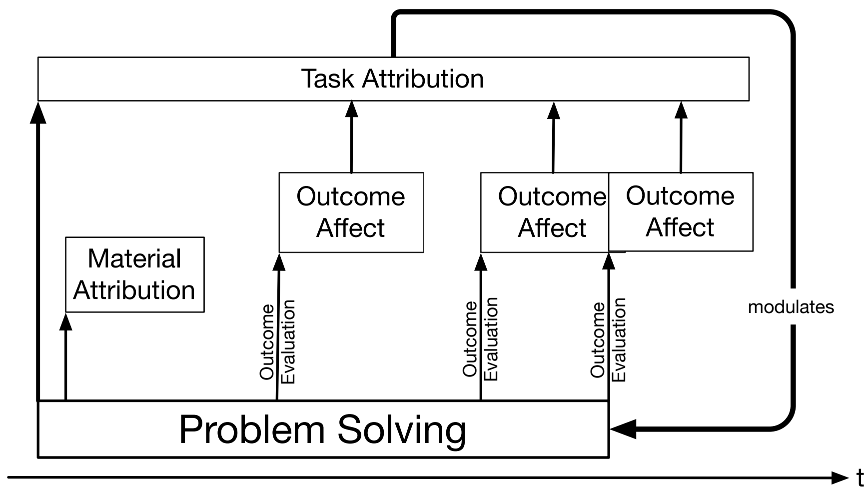

Figure 3: Illustration of task-related affective states accompanying performance of a problem-solving task that i) contains material with affective connotation at the beginning of the task and ii) consists of several step to reach a solution.

Long-Term Affect: Task Attribution

When individuals are confronted with a certain problem-solving task, a subjective attribution of the task occurs, which may result in a long-term affective state and accordingly interfere with problem-solving behavior. This state can be viewed as a task-related affective frame in which subsequent task performance takes place. Depending on individual attribution, such states vary with respect to their intensity as well as their valence. In the context of achievement emotion theory, these represent *trait-achievement emotions* [Pekrun, 2006].

Such states are modulated by domain-specific self-concepts of an individual [McCombs, 2001] and potential affective disorders. To illustrate this, let us again consider an algebraic problem-solving condition. In this context, an individual who feels confident that she likes such tasks and usually succeeds will have a more positive task attribution than an individual who is sure that she dislikes such a task and usually performs poorly.

In addition, let us consider an individual suffering from anxiety, such as math anxiety or a general text anxiety, being confronted with the same task. For such an individual, there is a negative task attribution, resulting in an intense and negative long-lasting affective state which significantly impacts subsequent problem solving [Raghunathan and Pham, 1999]. Based on this, individuals suffering from math anxiety exhibit a misjudgment of their own abilities [Richardson and Woolfolk, 1980]. In addition, they usually try to avoid anxiety-inducing situations [Ashcraft, 1995], which further interferes with successful task performance.

Fluctuations of Short-Term Affect

In addition to long-lasting mood-like affective states, problem solving is accompanied by dynamically changing short-term affective states of two different types.

On the one hand, short-term affect results from subjective attribution with respect to a certain aspect of a task's material. Such states represent *activity-focused state achievement emotions* [Pekrun, 2006]. For instance, let us consider a problem-solving task which is integrated in different cover stories of differing affective context. When individuals are confronted with such a task, there is an immediate evaluation of affective content [LeDoux, 1993, Mischel and Shoda, 1995], which results in a short-term affective state of a certain valence [Oatley et al., 2006]. As presented above, such states have been found to interact with current mood [Hesse et al., 1997]; in the case of matching valence, a more promising task outcome is expected.

On the other hand, there are short-term affective states based on individual progress with respect to individual goals. Such states represent *outcome-focused state achievement emotions* [Pekrun, 2006]. Depending on individual task performance and outcome evaluation, an individual sequence of these affective states arises. The particular dynamics of these sequences represent important information about individual problem-solving behavior and allow analysis of how these dynamics impact individual problem-solving performance [D'Mello and Graesser, 2010, McLeod, 1988]. This link is crucial, as successful problem solving does not solely depend on pure performance but also on successfully coping with its accompanying affective states [Dai and Sternberg, 2004].

IMPORTANT ROLE OF OUTCOME AFFECT Academic problem solving, in particular, is intrinsically tied to making mistakes and recovering from them [D'Mello et al., 2010]. In the case that an expected outcome does not match the actual result, individuals must deal with potential contradictions, misconceptions, and obvious contrasts [D'Mello and Graesser, 2010, Graesser et al., 2005, VanLehn et al., 2003] and are very likely to get confused. In addition, experiences of becoming frustrated may arise when individual goals are blocked by challenging obstacles [Dweck, 2002, Klein et al., 2002]. Recurrent experiences of short-term negative affective states are likely to eventually end in performance-limiting long-term affective states such as boredom and task disengagement [Csikszentmihalyi, 1990, Pekrun et al., 2010]. In extreme cases, individuals who are completely discouraged of being actively in control of their individual progress develop task anxiety [Heider, 1958, Weiner, 1985]. Similarly, persistent events of positive outcome affect lead to performance-supportive long-term flow-like affect resulting in increased task-engagement [Csikszentmihalyi, 1990].

Nevertheless, reliable findings on the interaction of individual outcome affect and problem solving-behaviour are rare, especially when focusing on specifics of individual dynamics in a single task [D'Mello et al., 2010]. This is due to methodological issues resulting from characteristics of outcome affective states that hinder their reliable measurement and analysis.

3 MEASURING SHORT-TERM AFFECTIVE STATES DURING PROBLEM SOLVING

To investigate functions and mechanisms of outcome-related affective states, it is necessary to reliably measure them. There are different techniques currently used for assessing affective states. These techniques are based on certain assumptions about the characteristics of outcome-related affective states and vary with respect to their effectiveness in a problem-solving setting. In addition, appropriate measurement techniques are limited in their applicability, as there is a lack of suitable problem-solving paradigms which would enable investigation of certain characteristics of their accompanying outcome-related affective states.

To illustrate these issues, this chapter first introduces current theory on how outcome-related affective states arise during problem solving. Next, different techniques for assessing these affective states are presented and evaluated with respect to their applicability. The remainder of this chapter focuses on the use of Affective Peripheral Physiological Responses (APPRs) as symptoms of affect: general characteristics for reliably measuring APPR are presented in detail, then the challenging issues of measuring and analyzing APPRs during a problem-solving task are illustrated.

3.1 APPRAISALS AS THE BASIS OF AFFECT

Cognitive theories of affect are based on the idea that there are evaluative processes for judging an object or situation as affective. In this respect, objects or (cognitive) events are appraised as having an affective value, i.e. as being relevant to individual goals. These evaluative processes are termed *appraisals*. They can occur consciously as well as unconsciously, and individuals are accordingly either able to report on them or not. Appraisals in this sense are viewed as mandatory antecedents of actual affective states [Frijda, 1993, Lazarus, 1991, Schutz and Davis, 2000, Smith, 1991, Scherer, 1984, Scherer, 2009, Ortony et al., 1988]. Depending on certain characteristics of these appraisal processes, it can be determined whether and which affective states are elicited by them.

Several theories about appraisal processes have emerged in parallel, each of which provides an individual set of criteria for eliciting different affective states. Nevertheless, there are specific criteria which are accepted across differing theories (see [Moors et al., 2013] for a review of the differing criteria). Theories mostly agree on two criteria for characterizing an appraisal and determining which affective state is elicited: *goal relevance* and *goal congruence*. The first represents how closely related a certain event or object is to individual goals. The latter determines whether this event or object matches or competes with individual goals. Additional criteria can be used to evaluate an event with respect to i) how certain an individual is with respect to the validity of an evaluative outcome, i.e. *certainty*, ii) the evident cause of the event, i.e. *agency*, and iii) whether one is able to actively manipulate the situation to one's own benefit, i.e. *control* vs. *coping potential*. Based on the outcomes of these different evaluations, appraisal processes may differ across individuals with respect to an identical event. As any appraisal is an antecedent of a specific affective state, different affective states are assumed to arise.

In the following section, the Component Process Model (CPM) [Scherer, 1984, Scherer, 2009] is introduced as a widely accepted theory on appraisal processes as antecedents of affective states. In the remainder of this dissertation, CPM serves as the theoretical framework for an analysis of achievement-related affective states.

3.1.1 *The CPM*

As previously presented, appraisal theories are based on the assumption that different emotions are caused by the outcome of individual evaluative processes with respect to a certain event. These evaluations occur through application of a certain set of appraisal criteria to the particular stimulus or event. Accordingly these processes are termed Stimulus Evaluation Checks (SECs). The CPM is characterized, in part, by assuming a specific set of SECs as antecedents of emotion. Figure 4 depicts the architecture of the CPM. It illustrates that SECs are mandatory for appraisal of a certain event or stimulus and occur on multiple levels of information processing. In addition, the dynamics and recursive processes of affective states are depicted as resulting from an event that is significant to the goals, needs, and values of an individual.

In [Scherer, 2009], the CPM is described as follows: An event is evaluated according to specific SECs at multiple levels of information processing. The results of these appraisal processes generally have a motivational effect, which alters the motivational level that persisted previous to the event. Based on the outcome of the appraisal process, as well as the updated motivational level, changes

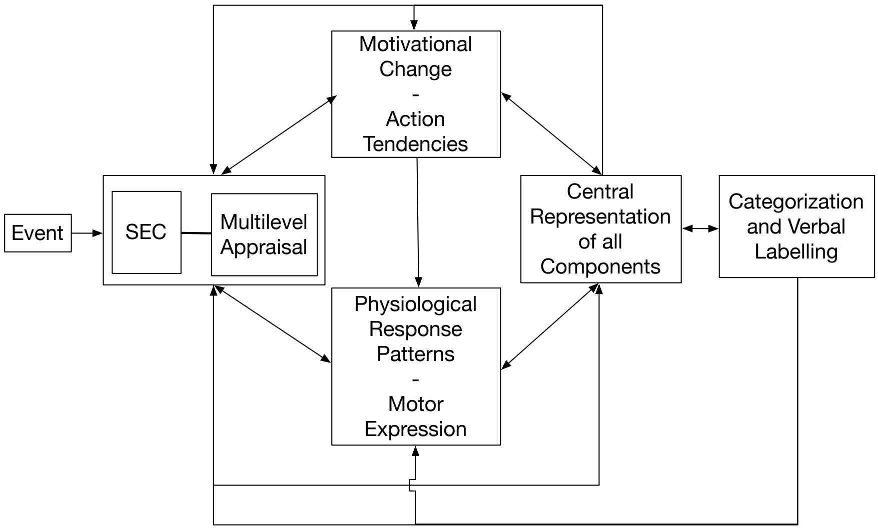

Figure 4: The architecture of CPM. Adapted from [Scherer, 2009, p.1308].

occur in both the autonomic nervous system and the somatic nervous system. These changes become evident in modifications of physiological response signals and motor expressions. All of these previous components (appraisal outcomes, motivational level, somatovisceral changes, and motor expressions) are combined into a multimodal representation, which is constantly updated when events or appraisal outcomes change. Based on this multimodal representation, parts of the affective state may become conscious and are assigned to an individual category of emotion, labeled with corresponding emotional terms, expression and metaphors.

3.1.2 *Appraisal Criteria*

In context of CPM, four different categories for classifying SECs are proposed. These categories are *relevance*, *implications and consequences*, *coping potential*, and *norm compatibility*. The criterion "relevance" is the most important for an emotion-antecedent appraisal, as it is mandatory for further evaluations of an event. The categories comprise more detailed criteria, as follows (cf. [Scherer, 2013a, p.151]).

- Relevance consists of three sub-criteria:
  - Novelty: Is the event sudden, familiar, or unpredictable?
  - Intrinsic pleasantness: Is the event itself pleasant or unpleasant for the individual?
  - Importance for individual goals: Is the event relevant and important for individual goals or needs?

- Implications and consequences comprises five other evaluation criteria:
  - Causal attribution: Was the event caused by the individual or by somebody else? Did it happen with intent or not?
  - Outcome probability: Are the consequences of the event predictable?
  - Discrepancy from expectation: Is the event compliant with individual expectations or not?
  - Goal or need conduciveness: Are the consequences of the event positive or negative for the individual?
  - Urgency: Is an immediate response to the event necessary?
- Coping potential includes three specific criteria:
  - Control: Is the individual able to control the consequences of the event?
  - Power: Can the individual manipulate the consequences of the event?
  - Adjustment: Can the individual live with the consequences of the event?
- Norm compatibility provides two sub-criteria for evaluation:
  - Internal standards: Is the event in line with individual standards and ideals?
  - External standards: Do the events violate laws or socially accepted norms?

Valence of SECs

The outcome of SEC with respect to these criteria will result in different appraisal outcomes, which elicit different affective states. As presented earlier, such affective states are categorized with respect to their *valence*, an umbrella term for the degree of positivity or negativity. In the context of CPM, it is argued that there are different types of *valence* that depend on the associated SEC.

[Scherer, 2010] suggested that there are at least six different types of valence. More specifically, there are six different SECs with a specific built-in aspect of valence. These SECs and their associated types of valence are (cf. [Scherer, 2013a, p.154]):

- *Novelty*, associated with the valence of *familiarity*, varying along the dimensions of unknown to familiar.
- *Intrinsic pleasantness/ beauty*, associated with the valence of *pleasure*, varying along the dimensions of *pleasant to unpleasant*.

- *Goal/ need conduciveness*, associated with the valence of *contentment*, varying along the dimensions of *satisfied to disappointed*.
- *Coping potential*, associated with the valence of *power*, varying along the dimensions of *strong to weak*.
- *Compatibility to self-standards*, associated with the valence of *self-worthiness*, varying along the dimensions of *achieved to failed*.
- *Compatibility to norms/ values*, associated with the valence of *moral worthiness*, varying along the dimensions of *virtuous to wicked*.

3.1.3 *Dynamics and Interactions of Appraisals*

SECss vary with respect to the required level of information processing, as well as in their temporal dynamics. Different SECs occur at different levels of information processing, depending on how many additional semantics and inferences are needed to perform an evaluation. For instance, SECs that are based on particular stimulus characteristics can be assumed to be easily and rapidly evaluated at lower levels of processing. In contrast, SECs that require additional semantics and inferences occur at higher levels of processing. With increasing level of processing, the required processing time also increases. Four different levels are postulated (cf. [Scherer, 2013a, p.151]):

1. SECs can occur on a low *sensorimotor* level. Evaluation is based on pattern-matching mechanisms that are largely genetically determined. Criteria for evaluating a certain event consist of looking for appropriate templates that match certain characteristics of an evaluated stimulus or event. SECs for evaluating novelty and intrinsic pleasantness occur on this level of processing.
2. SECs that occur on a *schematic* level are based on learned facts that can be retrieved fairly automatically and unconsciously.
3. SECs can also occur on an *association* level, entailing various cortical association areas. Such evaluations may either happen automatically and unconsciously or involve deliberate and conscious processing.
4. SECs that occur on a *conceptual* level involve propositional knowledge, as well as individual cultural meaning systems. For such SECs, consciousness and effortful calculations in prefrontal cortical areas are mandatory.

Sequentiality of SEC Outcomes

SECs of different levels of processing are assumed to be processed in parallel but to also have *"sequential-cumulative effects on all other emotion components"* [Scherer, 2013a, p.156]. The latter results from the assumption of different processing times for SECs requiring different levels of processing. The outcome of an SEC arises much faster for evaluations that occur on a lower level of processing. As a result, such SECs yield an outcome earlier than higher level SECs. Accordingly, SECs requiring lower levels of processing immediately exhibit their effects on all other components of emotion. This also includes concurrent SECs running at higher levels of processing. Running SECs that receive effects from lower level SECs are assumed to intensify their processing accordingly.

Due to the different temporal characteristics for performing a certain evaluation, SECs with increasing levels of processing exhibit their effects in a sequential manner, considering the effect of any lower level SEC. Nevertheless, it is argued that efferent effects on other affective components are only exhibited in cases with significant justification, i.e. outcome of an SECs is relatively stabilized and important enough to justify the investment of energy. In addition, it is assumed that there is a continuous adaptation of efferent response patterns with respect to the specificity of an appraisal outcome, considering the information that results from this cumulative-sequential process.

The particular valence types of different SECs, i.e. *microvalences*, add up to a one-dimensional valence, i.e. *macrovalence* [Shuman et al., 2013]. This superordinate type of valence represents a single, homogeneous valence dimension which can serve as a driver for further cognitive processing such as decision making.

3.2 MEASURING AFFECTIVE STATES

As previously discussed, appraisals are hypothesized to have a direct, differential, cumulative impact on different components of affective states, such as conscious feelings, expressive behavior, action preparation, and physiological responses [Frijda, 1986, Lazarus, 1991, Scherer, 2009]. Consequently, affective states are usually determined by assessing a single element or a set of these components to measure the corresponding affective state.

Methods for assessing these different components differ in their requirements for a reliable measurement of affect. Based on this, some of them are more suitable for assessing affective states while a certain task is performed than others. In the following section, different measurement paradigms are introduced and evaluated with respect to their potential applicability for assessing affect during performance of a problem-solving task.

3.2.1 *Self-report*

Appraisals of task-related events are a precondition for outcome-related affective states and are assumed to become manifest in conscious feelings that individuals can report on [Frijda, 1993,Scherer, 2009]. Based on this, different appraisals and outcome-related affective states can be measured by utilizing self-report methods. These methods are based on questionnaires that support collecting information about affective states of an individual. Different questionnaires which can be distinguished with respect to two characteristics. First, available questionnaires differ with respect to their underlying theory of emotion. Second, they differ regarding the point in time for assessing a certain affective state with respect to its potential elicitor.

With respect to the first issue, the terminology of questionnaires depends on the underlying model of affect. For instance, if it is assumed that there are basic emotions whose labels are equal among different individuals, a corresponding questionnaire assesses these single and discrete labels of emotion. If affect is considered to be represented by a complex combination of different affective states, such combinations are assessed either using standardized scales, such as the Geneva Emotion Wheel (see [Scherer, 2005]; [Scherer, 2013b]), or individual scales (e.g. in [Spering et al., 2005]). To reduce issues that result from inter-individual differences in labeling affect, there are assessment tools based on pictorial representations of affect, such as the Self-Assessment Manikin [Bradley and Lang, 1994], in which affect is assessed with respect to a dimensional view of affect.

Considering the second issue, questionnaires can be utilized at different points in time with respect to the affect-eliciting event and related affective states. On the one hand, affect can be immediately assessed after its occurrence. However, when affect is assumed to arise at several points in time in a multi-step task, any affect assessment using questionnaires results in an interruption of the actual solving process. In addition, there is evidence that reporting on a current affective states reduces its impact on cognitive processing [Keltner et al., 1993]. On the other hand, affect can be assessed retrospectively, i.e. individuals are asked to report on their accompanying affective states after they have finished a certain task. This can either be solely based on individual memories or supported by video recordings of an individual while performing the task [D'Mello et al., 2010]. In both cases, these retrospective accounts are not necessarily free of memories of previous experiences [Ericsson and Simon, 1980].

3.2.2 *Expressive Behavior and Facial Expression*

As mentioned above, appraisals are assumed to become evident in expressive behavior and action preparation. To illustrate this, let us consider an individual confronted with a fearful situation. A certain part of the situation is appraised as such, and it is assumed that there is an urge to immediately get out of this situation or alter it to one's own benefit. This action tendency is accompanied by a so-called *fight-or-flight* response [Cannon, 1915]. Such a response becomes evident in body language or facial expression. In addition to such an intense affective state which appears to be important for basic survival from an evolutionary perspective, other negative as well as positive affective states are also assumed to become evident in body language [Wallbott, 1998] and facial expression. Although there is conflicting empirical evidence for existence of specific body postures as symptoms of particular affective states, specific facial expressions have been found to be associated with specific affective states and basic emotions [Ekman, 1993].

Based on this, affective states can be measured by monitoring facial expressions or body movements. This can be done in two ways. First, video recordings of the face or body can be analyzed either employing specialized algorithms for video classification or asking human experts to manually rate video recordings with respect to present affective states. Alternatively, classification of affective states can be done by analyzing the activities of certain skeletal muscles that are assumed to cause an affect-related change in facial expression or body posture.

The lack of supporting evidence for stable inter-individual body postures that distinguish different affective states might be due to restricted freedom of movement in experimental setups [Kappas, 2011]. In an experimental setting, body movement is often restricted to prevent the occurrence of movement artifacts in sensor or video recordings. This restriction of movement is likely to impact embodiment of emotion and therefore interferes with assessing affect in terms of body movements. In addition, simply being in an experimental environment may result in a unnatural restraint on behavior that impacts embodiment of affect.

Although there is supporting evidence for consistent inter-individual facial expressions with respect to basic emotions [Ekman, 1993], it is still an open issue whether these are also present in task-related achievement settings. In addition, facial expressions can be voluntarily modified to keep a potential spectator uninformed about individual affect. This is especially likely in achievement settings where individuals try to kept others in the dark about their affective states in conjunction with their individual achievement, such putting on a "poker face" when a task becomes more difficult.

3.2.3 *Physiological Responses*

Human physiological responses comprise activities of different physiological systems of the human body. Based on this, physiological responses can be distinguished with respect to the system involved. Consequently, neurophysiological responses are discriminated from responses of the peripheral physiological system. Neurophysiological responses reflect signal measures that are a symptom of direct brain activity. Peripheral physiological responses comprise bodily signals that are modulated by activity of the autonomic nervous system, such as cardiac activity, respiration, skin temperature or electrodermal response, i.e. modulation of skin sweat level.

Measuring peripheral physiological responses has certain advantages over measuring neurophysiological responses. For instance, devices for reliably measuring peripheral physiological response are less expensive than devices for measuring neurophysiological responses. In addition, the human brain is a fast-operating organ which processes masses of information in a very short time. As a result, affective processing does not take place completely separated from other processes in the brain, but is integrated in a complex response system, especially in the context of a cognitive task. This hinders a reliable distinction of affective response and other cognitive processing through *electroencephalography* (EEG) or *functional magnetic resonance imaging* (fMRI). Based on this, measurements of peripheral physiological response as symptoms of affect are widely used, although there are efforts to localize affective processes in the human brain.

Based on the assumption that appraisals become evident in expressive behavior, it appears to be straightforward to measure accompanying physiological responses such as faster breathing and a racing heart. In this case, physiological responses are viewed as side effects of behavioral changes. Nevertheless, there is significant evidence that changes in peripheral physiological responses are directly linked to appraisal processes, independent of behavioral changes. In particular, the previously discussed CPM [Scherer, 1984,Scherer, 2009] is based on the idea that a fixed sequence of appraisal evaluations triggers certain changes in activity of the autonomic nervous system, which become evident in specific physiological responses. In addition, [Smith, 1991] provide evidence that physiological activity also varies with respect to different dimensions of appraisal, such as relevance for individual goals and anticipated effort.

Consequently, patterns of APPR are widely used as a symptom of individual affective states. Due to the nature of such responses, there are certain issues that must be considered in order to reliably measure and thoroughly analyze such signal patterns. These issues are especially important when APPR is measured with respect to achievement-related affective states that accompany task performance.

3.3 MEASURING AND ANALYZING APPR

The general idea of analyzing affective states based on measurements of APPR is based on the assumption that individual appraisal processes and their related affective state become manifest in different patterns of peripheral physiological responses and that these are specific across all or most individuals. Before an such analysis can take place, corresponding patterns of APPR have to be reliably measured. As APPRs significantly differ in their temporal characteristics depending on the physiological signals involved, it is important to consider these timing issues during measurement.

3.3.1 *General Characteristics of APPR*

Basically any APPR results from an interaction between the central nervous system, the brain and the spinal cord, and the autonomic nervous system. The latter is part of the peripheral nervous system, i.e. nerves and ganglia outside of the brain and spinal cord, and acts as a control system that unconsciously manipulates visceral functions. The autonomic nervous system consists of three different subsystems: sympathetic nervous system, parasympathetic nervous system, and the enteric nervous system (see [Kreibig et al., 2010b] for a review). While the latter has the specific function of controlling the gastrointestinal system [Furness, 2008], the first two modulate the activity of several organs of the human body. The distinction of these two subsystems is based on the anatomical structure that innervates a certain peripheral part of the body from the central nervous system.

While the sympathetic nervous system is a rather fast-responding system, the parasympathetic nervous system is a more slowly responding system. Accordingly, sympathetic chronotopic control is estimated to be about 230 msec after a response is triggered by the central nervous system, while parasympathetic chronotopic control takes about 1710 msec to respond to the trigger [Berntson et al., 1993, Berntson et al., 2007].

Most organs respond to only one of the two systems; only a few, such as the iris and cardiac activity, are modulated by both of the systems. Even when certain organs are modulated by the same subsystem, different organs show a different response, depending on the triggering event and its affective connotation. Consequently, each physiological signal responds with its own latency, characteristic to a trigger. In addition, for stimulus-related short-term affective states, different latency and response characteristics have been identified for different affective states [Frijda, 1986, Lanteaume et al., 2007]. As the response time of the systems involved appears to be stable, latency variations are assumed to result from temporal characteristics of the processes involved in perception and affect processing before the related physiological response is triggered. Based on this, APPR is assumed to be affect-specific. The following section presents a widespread methodology for testing this assumption.

Evidence for Affect Specificity

Empirical evidence for affect specificity comes from experiments in which certain affective states are induced by presenting an individual with an affective stimulus and measuring the corresponding APPR. Afterwards, individuals are asked to report on the affective state they underwent after affect induction. If the individual reports having felt the intended emotion, her APPR will be compared to response patterns of other individuals who also consistently reported the induced emotion. To identify specific patterns of APPR, response patterns of a specific affective state have to be analyzed with respect to different elicitation contexts, types of stimuli, or utilized affect induction methods. With increasing numbers of experimental contexts in which APPR are found to be specific for a particular affective state, evidence increases that this APPR is specific for this particular affective state [Stemmler, 2003].

In this way, specific patterns of APPR have been identified for stimulus-related affective states and potential differences have been recognized. For instance, it has become evident that APPR differs with respect to elicitation context [Levenson, 2003]. This may either be due to methodological issues with respect to eliciting affect in the laboratory or can be viewed as evidence for context-specificity of stimulus-related patterns of APPR. In addition, there is empirical evidence that affective states and their associated somatovisceral response patterns vary for SECs of intrinsic un/pleasantness and goal conduciveness/obstructiveness [Aue and Scherer, 2011]. For instance, somatovisceral responses are significantly different when individuals are confronted with pleasant or unpleasant pictures compared to physiological profiles that result from being confronted with goal conducive or goal constructive conditions.

Nevertheless, this general methodological approach with respect to a certain stimulus facilitates analysis of patterns of APPR. As the particular point in time when an individual has been confronted with the affective stimulus is known, it is easy to determine the individual time windows during which an APPR is expected for each of the signals involved. In addition, affect is measured in an isolated setting where the only expected physiological response is an APPR. This situation is much different for patterns of task-related APPR.

3.3.2 *Characteristics of Task-related APPR*

As presented earlier, achievement emotions arise any time an individual evaluates her progress with respect to an individual goal. Such evaluative processes can be triggered solely based on internal events that are managed by the individual, i.e. *intrinsically*, or triggered by an external event, i.e *extrinsically*. For the first, evaluation is triggered by an internal cognitive event comparing current progress to individual goals, such as evaluating the outcome of planned actions before performing them. For the latter, an external event necessitates an updated evaluation of current progress with respect these goals. Examples of such events include being confronted with feedback about current progress or novel information about the problem that needs to be incorporated to judge current performance and adapt behavior for proceeding with the task accordingly.

As a result, task performance is accompanied by different APPRs at different points in time. In addition, physiological response is also modulated by cognitive processing as well as bodily movements. Patterns of APPR occur in the presence of other non-affective physiological responses, and the respective patterns are connected and interfere with each other. Consequently, before patterns of APPR can be compared with each other, they must be extracted from the collected physiological data. For this, two steps are necessary. First, particular segments of readings of physiological signals that are assumed to represent an episode of an APPR have to be separated from the remaining non-affective physiological signal readings. Second, these physiological response patterns have to be annotated with their corresponding affective value. Based on this, a certain segment of physiological data can become a specific APPR pattern.

3.4 DETERMINING APPR DURING PROBLEM SOLVING

There are two different assumptions that support a reliable measurement of APPRs during a problem-solving task. On the one hand, APPRs are assumed to arise in relation to observable actions by an individual and are extracted accordingly. On the other hand, APPRs are measured with respect to external events such as feedback.

3.4.1 *Relation to Observable Actions*

Measuring patterns of APPR in relation to an observable action that is performed by an individual is based on the assumption that action performance is followed by an evaluation of the resulting situation. This is due to the fact that an action represents an intentional manipulation of the environment; a certain action is selected with respect to its expected modification of the current situation. Accordingly, the resulting situation is evaluated as to whether or not it matches the current expectation. Depending on the outcome of this evaluation, different affective states are expected to arise.

Based on this, after a certain action has been performed there is no evaluation of individual progress towards individual goals, but rather a comparison of the expected outcome situation and the actual one. As any action is selected with respect to its expected progress toward individual goals, such an evaluation can be assumed to take place before the action is actually performed. Thus, no measurement of the current progress towards the goal is assumed after task execution. As a consequence, patterns of APPR preceding action performance are not expected to vary with respect to individual progress.

From an information-processing perspective, actions are evaluated and selected based on a search in problem space, which is represented as a branching tree. Traversing this branching tree can be done in two ways. On the one hand, subsequent problem states can only be reached one after another. As a result, only single actions can be performed. On the other hand, individuals are able to preplan and perform a certain sequence of actions, which enables them to traverse a sequence of subsequent problem states. Based on this, individuals can either perform a single action and wait for additional information before they can proceed or they can perform a sequence of actions. In both cases, there are different challenges with respect to measuring and analyzing APPRs.

Single Actions

In a well-defined problem-solving task where only single actions can be performed, preplanning and performing sequences of actions is not necessary. It is straightforward to determine the expected outcome of an action. Accordingly, it is unlikely that the actual outcome of an action mismatches the individual expectation. As previously presented, an action is evaluated with respect to individual goals before it is executed. After an action is performed, the action's outcome is evaluated as to whether it matches the individual's expectation. As a result, APPR results after action execution do not differ with respect to advantageous and disadvantageous actions.

To illustrate this let us consider a game of chess. Here, each player is allowed to perform a single action and has to wait for an opponent's action afterwards. APPR can be measured and analyzed with respect to the point in time when an individual has performed a certain action. In this task, an individual can quite accurately determine an action's outcome. As a consequence, a mismatch of actual outcome and expectation does not happen for an experienced chess player. Accordingly, APPRs that are subsequent to making a move do not differ with respect to advantageous and disadvantageous actions [Leone et al., 2012]. Nevertheless, there are difference in patterns of APPR preceding an action that discriminates advantageous from disadvantageous actions [Leone et al., 2012]. These patterns of APPR are seen as the result of evaluative processes that are targeted on the action's progress towards the individual goal (see Somatic Marker Hypothesis [Damasio, 2005] for a review). To summarize, in a well-defined problem-solving task with single steps, APPR following an action is based on evaluating the match of expectation and actual situation. APPR related to progress towards individual goals precedes action performance.

Sequences of Actions

Measuring and analyzing patterns of APPRs becomes more complex in tasks where sequences of actions can be performed. To illustrate this, let us consider the popular Tower of Hanoi task. This game consists of a stack of discs of different sizes and three rods onto which these discs can be moved. The game starts with a complete and ordered stack on one rod, smallest at the top. The task is to move this complete stack to another rod in as few moves as possible while sticking to three simple rules. First, only a single disc can be moved at a time. Second, only the uppermost disc of a stack can be moved. Third, no disc can be moved onto a smaller one.

In this game, individuals are likely to preplan certain sequences of actions and evaluate their outcome before actually performing one of them. In this case, there is no external event that allows the observer to determine both which sequence of actions is evaluated, and when this evaluation takes place. When a certain sequence of actions is selected and performed, each of its single actions is evaluated immediately after its execution. This evaluation deals with whether the specific outcome matches the previously expected outcome that has been considered during preplanning.

Accordingly, a specific APPR is only assumed in case of a mismatch between expectation and actual outcome. This can happen in two situations. First, there can be a large problem space whose action sequence to reach a goal exceeds cognitive capacities. Hence, the task can only be solved by preplanning and performing sub-sequences for moving towards a goal. In this case, the outcome of a sequence is evaluated after the last action is performed, to come up with a following sequence of actions. Second, due to limited cognitive capacities, preplanning could have been inaccurate, so a certain action's outcome does not match the individual expectation. Consequently, performance of the current sequence is interrupted, the current outcome is evaluated with respect to the goal and a novel sequence is planned.

Based on these previous issues, patterns of APPR have been analyzed in impasse situations when performing a Tower of Hanoi task [Clément and Duvallet, 2010]. In this respect, an impasse situation is present if i) there is an interruption between two subsequent actions, ii) a backward move has been performed, or iii) a rule violation has been detected, i.e. a larger disc is put onto a smaller one. Thus, in tasks where sequences of actions can be performed, APPR can only be reliably measured with respect to particular events in the problem-solving process. In addition, as cognitive capabilities determine when an individual will be confronted with an impasse situation, the occurrence of certain APPRs is modulated by the general task performance of an individual.

Role of Individual Performance

Independent of whether APPR is measured preceding action performance in a single action task or subsequent to a certain event in a task where sequences can be performed, achieving an affective context, i.e. an advantageous, a disadvantageous, or an impasse situation, highly depends on individual performance. For instance, a chess master will experience not as many disadvantageous situations than a novice than a novice confronted with the same opponent. In line with this, individuals who are able to solve a Tower of Hanoi task with five discs in an almost optimal sequence of actions, i.e. close to the minimum of 31 actions, will encounter fewer impasse situations than individuals who need more steps.

Based on this, the opportunity to measure and analyze patterns of APPR is determined by individual performance of participants: the lower individual performance is, the higher the likelihood of being confronted with disadvantageous, critical, or mismatching problem states. The opposite relation holds for advantageous and expected problem states. With increasing performance, an individual's expectations and plans are more accurate and critical events are less likely to arise. Thus, an APPR that occurs with respect to overt actions in the context of full and accurate information about the problem is solely modulated by individual performance and cannot be intentionally manipulated. This situation changes if an individual cannot acquire full information about the expected problem status. This is the case if a task provides additional information after a certain action has been performed which then has to be integrated before individual progress can be evaluated, such as explicit feedback.

3.4.2 *Relation to Provided Feedback*

When individuals are confronted with novel information about the current problem status, it results in an immediate evaluation of current progress, the outcome of previous actions, toward individual goals [D'Mello and Graesser, 2010]. Based on this, providing an individual with feedback about the current problem status triggers an immediate evaluation, which is assumed to result in an APPR. Because the point in time when feedback is presented is evident, it is possible to determine the point in time of an arising APPR. Consequently, APPR is reliably measured in relation to provided feedback.

Such a scheduled triggering of APPRs is especially supported by task paradigms that prevent an individual from evaluating actual progress without this additional information. This is the case if outcome of a performed action does not provide enough information for an extensive evaluation. This can be achieved by adding a task-external evaluation criteria that is only fed by the additional information that is provided by feedback. In such a task, individuals select a certain action based on some internal evaluation criteria, i.e. performing as well as possible. The outcome of a performed action is then evaluated with respect to the received feedback information and the other goal criteria, such as being more successful than other individuals.

There are three types of feedback. First, feedback can be a direct representation about the correctness of the current problem state, i.e. whether the current problem state is correct or not. Second, feedback can be a presentation of the correct or optimal solution of the respective problem state. Based on this, individuals must compare their individual outcome with the presented information to evaluate actual performance. Third, feedback can be a qualification about how good individual performance has been with respect to a task-external evaluation criterion, such as individual performance compared to other individuals. Again, to evaluate individual performance, provided information must be interpreted in light of individual goals and the relative state of the current problem. As timing of feedback presentation determines when an APPR arises, it is possible to manipulate temporal dynamics of APPR to some extent.

For instance, affective states have been measured in analytical reasoning tasks in which feedback has been provided with respect to the correctness of the final outcome [D'Mello et al., 2010]. In this case there are four different types of feedback: feedback can be either 1) positive or 2) negative and matches actual performance or 3) positive or 4) negative and contradicts actual performance. Based on this, affective states are assumed to be intentionally manipulated. Nevertheless, which condition can be presented at a certain point in time highly depends on performance of the individual.

To illustrate this, let us consider that all four types of feedback are supposed to be presented for the same number of tasks in the experiment. Based on this, collected patterns of APPR are to be analyzed. Nevertheless, the opportunity to receive positive feedback that matches actual performance, i.e. good performance, is decreased for an individual who is performing poorly in the task. At the same time, the opposite holds for good performers. For those individuals, there is an decreased possibility to receive negative feedback that matches actual, i.e. poor, performance.

To overcome this limitation, tasks can be selected whose underlying structure conceal actual performance. In such a task, individuals cannot accurately evaluate their individual problem state after an action has been performed but have to rely solely on presented feedback information. In this case, feedback can be presented that does not match the current performance of an individual. For instance, in a visual discrimination task in which individuals have to correctly count certain features, there can be a very limited time window to identify these features [Kreibig et al., 2010a]. In this case, individuals do not have enough time to evaluate their result in the presence of the task setup and consequently their actual performance. Individuals can only judge their task performance based on provided feedback.

The same holds true for a task that consists of several sub-trials, with feedback provided at the very end of all trials reflecting performance across all trials [Kreibig et al., 2012]. In this case, with increasing number of trials, it becomes more and more difficult to remember the individual performance in each trial. As a result, it becomes difficult to form an individual view with respect to overall task performance. Consequently, individual performance evaluation solely depends on provided feedback information. In such a feedback setting, performance evaluation becomes disconnected from single task performance.

To summarize, providing feedback allows manipulation of temporal occurrences of APPR. Intentionally manipulating the valence of feedback conditions and affective context accordingly is only possible if a task matches specific criteria, such as obfuscating real performance to the user or allowing false feedback to remain unnoticed by the individual.

3.4.3 *Analyzing Sequences of Affect*

To analyze a certain sequence of specific APPRs, it is necessary to provide a specific affective context. In addition, all resulting APPRs have to be reliably measured one after the other. Based on the previously presented issues, determining the point in time when an APPR is expected to occur is only possible for tasks in which progress evaluation requires additional information, such as feedback, to be integrated. To analyze sequences of APPRs, the task setup requires several sequential problem-solving steps and must allow manipulation of the point in time at which the individual evaluates her progress toward the general task goal. In addition, there must be controlled conditions under which APPRs occur during task performance, i.e. comparable conditions of progress across individuals.

As previously discussed, manipulating task progress is usually achieved by concealing actual progress from the individual. This approach does not fit a multi-step problem-solving task in which individuals have to deliberately draw inferences on provided information. If this information is restricted to conceal real performance from an individual, deliberate inferences are not possible anymore. Accordingly, task progress does not rely on drawing inferences, but transforms decision making into a gambling task where other affective states are assumed to arise (cf. somatic marker hypothesis [Damasio, 2005]). As there is no such sequential task that allows manipulation of affective context, researchers have developed special task paradigms to induce sequences of affective states. For instance, sequences of affective states are analyzed from repetition of a single task for which an intended affective context can be induced. In addition, affective states are elicited in a extensive problem-solving task that involves task-external feedback information at several points in time.

Repetition of Single-Step Task

To analyze the dynamics of sequences of affective states, these states are measured while individuals perform a sequence of one-step analytical reasoning tasks [D'Mello et al., 2010]. Affective context is manipulated by providing either positive or negative feedback in correspondence to actual performance or positive or negative feedback in contradiction to actual performance. As a result, individuals are assumed to undergo a certain sequence of affective states. Based on this sequence, transitions from one affective state into another have been analyzed. Nevertheless, this methodology does not allow investigation of the dynamics of affective states in a single sequential task. In addition, the possibility to provide a certain feedback pattern to an individual highly depends on an individual's general performance in the task.

Complex Task with Task-External Feedback

To analyze, for example, the impact of specific sequences of affect on problem-solving behavior, it is necessary that individuals undergo identical sequences of affective states during task performance. This is difficult to achieve, as predefined task feedback might contradict actual performance. To overcome this limitation, feedback can be employed that is not related to actual outcome of an action but relates current performance to a task-external criterion. For instance, individuals may receive feedback with respect to how well they perform in comparison to other individuals.

This approach is employed while individuals are confronted with a complex problem-solving task, e.g. the Lohhausen scenario [Dörner et al., 1983]. In this task, individuals assume the role of mayor of a small town who has to administer its social and economic situation. This is a complex task as several highly interrelated variables have to be controlled by the individual. As a result, individuals have difficulty assessing their individual performance during the task. Feedback about specific progress in comparison with other problem solvers is provided as information about current performance at several points in time during task performance [Spering et al., 2005]. Consequently, problem-solving behavior is analyzed with respect to the specific sequence of affective state. Nevertheless, this methodology does not allow analysis of how problem-solving behavior is modulated by affective states that evolve from evaluating individual progress during task performance and accompany an extensive problem-solving task. Accordingly, potential accompanying APPRs are not related to actual task performance and individual outcome evaluation but depend on classification with respect to task-external standards.

Part II

A NOVEL METHODOLOGY

4 REQUIREMENTS FOR ANALYZING SEQUENCES OF APPRS

This chapter presents the task-specific requirements that must be met in order to analyze patterns of APPRs. First, the main conclusions that can be drawn from the theoretical background are highlighted. Second, general requirements for analysis APPR across individuals are emphasized and electrodermal activity, i.e., skin conductance, is introduced as a physiological signal applicable to the questions considered in this thesis. Based on this, hypotheses on APPR during problem solving are presented in more detail. Third, specific prerequisites for analyzing these hypotheses are presented. Fourth, the Mastermind task is presented as a suitable problem-solving scenario for measuring sequences of APPR. Potential constraints on analysis of sequences of APPR are presented and a modification of the task is described.

4.1 CONCLUSIONS FROM THEORETICAL BACKGROUND

1. Outcome-related affective states that arise during a problem-solving task result from appraisal processes, i.e., SECs.
2. SECs have efferent effects on somatovisceral responses. Accordingly, outcome-related affective states can be investigated by analyzing their associated APPRs.
3. There is need for an experimental paradigm that allows inter-individual comparison of achievement-related APPR, especially when focusing on sequences of APPR that arise during task performance.

4.2 ANALYZING APPR ACROSS INDIVIDUALS

Peripheral physiological responses can be considered highly individual and are unlikely to be identical across individuals, especially when physiological responses are analyzed in terms of absolute values. Nevertheless, these peripheral physiological responses follow specific principles that allow analysis of these patterns in terms of APPR for a certain individual. In this context, physiological signal measurements can be analyzed with respect to intra-individual differences to distinguish different affective conditions. In addition, there are certain stable dependencies between peripheral physiological responses and affective contexts, which allows an analysis of APPRs across individuals, i.e., inter-individual analysis.

These two different types of analysis are only possible in the presence of a sound theoretical conceptualization of i) the causal cognitive event, the *indicand* [Kreibig et al., 2010a, p. 107] of an APPR, ii) a (set of) measured intervals of peripheral physiological response signals, the *indicator* of an APPR [Kreibig et al., 2010a, p. 107], and iii) the relationship between the first and the second which allows it be to termed an APPR.

In this dissertation, individual and unconscious evaluative processes, i.e SECs as described in the CPM, are considered to accompany human problem solving and are assumed to be the indicand of an APPR. Electrodermal measurements of skin conductance are employed as an indicator domain. Depending on the assumed outcome of a SEC, there is a different relationship between indicand and the indicating physiological signal. In the following section, the physiological measure of skin conductance is introduced in more detail, and specific hypotheses relating indicand and indicator of a certain APPR are formulated.

4.2.1 *Skin Conductance as Physiological Measure*

The observation and analysis of electrodermal activity is one of the most frequently used in social and psychological literature and is one of the most robust measures of autonomic nervous system activity that can be measured non-invasively [Damasio, 2005, Cacioppo and Tassinary, 1990]. There are several studies in which skin conductance, one aspect of electrodermal activity, was analyzed with respect to basic emotions; a changing level of skin conductivity was found to be interrelated with happiness, fear and sadness [Cacioppo et al., 2000].

Electrodermal activity can be measured using a pair of electrodes connected to an individual's skin and a polygraph. As the individual's somatic state changes with respect to a percept or a thought, the autonomic nervous system increases the amount of fluid secreted in the skin's sweat glands. This increase is so small that it cannot be detected by the human eye or the individual's neural sensors on the skin, but it is large enough to reduce the resistance between the two electrodes [Damasio, 2005]. In addition, this increase in sweat does not directly increase the humidity of the skin's surface but imbues the skin from the inside. Therefore, the increase in conductivity can be measured approximately one second before the skin's surface humidity increases [Schandry, 1998]. Measured skin conductance responses are characterized by three different features: i) *latency*, i.e.,

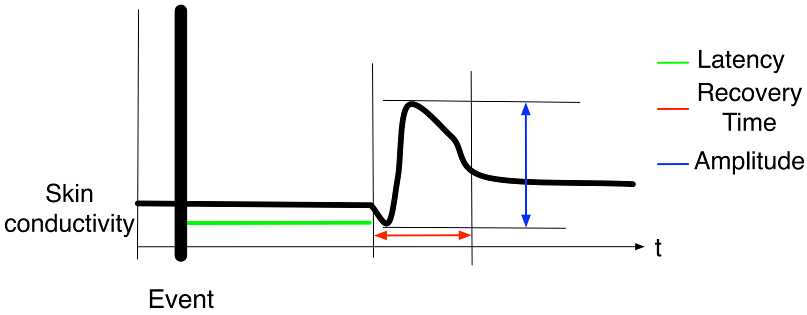

Figure 5: Response characteristics of a skin conductance response.

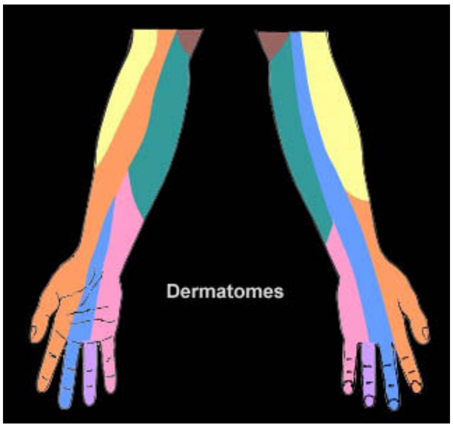

Figure 6: The different dermatomes of the human forearm highlighted in different colors. Source: `http://classes.kumc.edu/sah/resources/handkines/nerves/dermatome.htm`

time that elapses until the first signal response occurs, *recovery time*, i.e . the time needed for the signal to recover to a zero slope after a response has occurred, and *amplitude*, i.e., the maximum response. These three features are depicted in Figure 5.

Electrodermal activity is usually measured at the hand by sticking the electrodes on the skin after it has been washed with warm water (without soap, as this would result in a swelling of the epidermis and a decrease in conductivity) [Schandry, 1998]. Greasy skin can be cleaned with 70% ethanol. [Schandry, 1998] also proposes rubbing electrode gel into the skin before attaching the electrode, but this is not always necessary. Modern sensor configurations do not need this additional gel and resellers advise against the use of it. In addition, it is important to put the electrodes on parts of the hand that belong to the same dermatome, areas of the skin that area similarly innervated (see fig. 6).

Skin conductance activity consist of two different components.

1. A slowly changing *tonic* component that represents the background level of the electrodermal signal.

2. A fast-responding *phasic* component that is sensitive to novel events or stimuli and that overlays the tonic component.

Current recommendations for employing skin conductance responses in empirical investigations suggest analyzing these two components separately [Boucsein et al., 2012].

Skin conductance measures have been used to test hypotheses on goal conduciveness appraisal and their effect on skin conductance responses (see [Kreibig et al., 2010b] for a review). Individuals who are confronted with conditions of higher goal conduciveness exhibit an increase in skin conductance level. In addition, conditions of goal obstruction have been found to elicit stronger responses in skin conductance (amplitude of response) than goal conducive conditions [Van Reekum et al., 2004, Ravaja et al., 2006, Ravaja et al., 2008]. However, none of these studies separate phasic from tonic components of the response.

4.2.2 *Hypotheses*

In general, it is assumed that there are stable dependencies between specific progress in a task and the associated affective states an individual undergoes. Within this dissertation, the analysis of affective states during problem solving is focused on three main theses:

1. Depending on the experienced valence of task progress, an individual exhibits different affective states.
2. The specific experience of task progress may be subject to individual evaluative capabilities. As a result, individuals who perform well on a task may judge the valence of identical progress differently than individuals who perform poorly. Accordingly, depending on the general task performance, individuals exhibit different affective states.
3. These affective states also differ with respect to individual motivation to be successful, so that individuals who want to achieve highly exhibit different affective states than other individuals.

As previously presented, peripheral physiological responses are highly individual and are unlikely to be identical across individuals. Nevertheless, these responses have certain features that appear to have stable dependencies for individuals that share certain characteristics. For instance, when an individual is confronted with a particular condition of valence, a different APPR is exhibited as compared to being confronted with another condition of valence.

Based on the assumptions of CPM, APPR depends on whether a condition is satisfying or disappointing with respect to a specific goal. The first thesis has to be concretized with respect to the employed physiological signal, i.e., skin conductance.

There is evidence for stronger skin conductance responses in a condition of negative valence compared to a positive one. Accordingly, I hypothesize:

1. Individuals who are confronted with a condition of negative valence , i.e., goal obstruction or disappointment, exhibit a stronger skin conductance response compared to their skin conductance response in a positive condition.

Valence of individual progress in a task may be explicitly given, e.g. by task-related feedback indicating whether a previous action has been correct or not. Furthermore, the current problem status may also provide potential information about the specific task progress. In this latter case, valence of individual progress is not explicitly given but has to be perceived: Individual evaluative processes are required to judge specific progress.

Individuals who perform well on a task are likely to assess information more accurately. As a consequence, well-performing individuals who are confronted with two different conditions of progress, i.e., both goal conducive and obstructive, are expected to evaluate these two conditions as two differing ones. In contrast, such a clear distinction of progress may not necessarily occur for poorly performing individuals.

As a result, well-performing individuals are expected to exhibit different patterns of APPR for positive and negative conditions of progress. Such clearly distinguishable differences in APPRs are not expected for poorly performing individuals. Accordingly, I hypothesize:

2a. Individuals who perform well on a task exhibit skin conductance responses of varying strength, depending on valence of task progress, i.e., goal conducive vs. obstructive. Individuals who are performing poorly on the task are not expected to exhibit such clear-cut differences in strength of skin conductance response.

Individuals who perform well on a task are likely to process provided information more efficiently. Accordingly, results of progress evaluation may be available much faster for this performance group. Based on this, efferent effects on physiological components of emotion can also be expected to arise faster for good performers. Accordingly, I hypothesize:

2b. Individuals who perform well on a task exhibit a skin conductance response to current progress faster than individuals who are not performing well.

As described in Section 3.1.3 on page 33, there are four different levels of appraisal processes. These levels have efferent sequential-cumulative effects on physiological responses. In addition, there is a continuous adaptation of efferent response patterns with respect to the specificity of an appraisal outcome resulting from cumulative-sequential information.

Interaction of appraisal processes of the first and the second level, i.e., of relevance and implication assessment, have been studied [Kreibig et al., 2012]. Interaction of the second level with higher levels of appraisal processes is still to be investigated. For instance, it remains an open question whether the appraisal outcome of implication assessment interacts with appraisal processes for evaluating normative significance, i.e., whether one measures up to one's own standards.

If such an interaction is present, not only valence of task progress, i.e., goal conduciveness and obstruction, determines APPR but also whether the specific progress complies with individual self-standards. In this context, individuals who want to achieve highly might be more strongly affected by both positive and negative feedback than those individuals who are not achievement oriented. Based on this, I hypothesize:

3. Individuals who have high self-standards of achievement orientation exhibit stronger skin conductance responses for both positive and negative valence conditions than individuals who are not achievement oriented.

In the context of a problem-solving task that involves more than a single step, it is of great interest whether or not these hypotheses are also valid for more than a single step.

4.3 REQUIREMENTS FOR EVALUATING HYPOTHESES

To evaluate these hypotheses in the context of a multi-step problem-solving task, the employed task setup has to meet specific requirements for measurement, as well as analysis, of the related APPR. First, the temporal characteristics of APPR have to be considered i) to allow determination of the point in time at which an APPR is assumed to occur as well as ii) to prevent a convergence of two subsequent APPRs. Second, the task setup must allow measurement of specific indicand events with respect to the hypotheses to be evaluated, i.e., different conditions of outcome-related valence. This enables analysis of APPRs with respect to different conditions of valence for a single individual. Third, as hypotheses of APPR can only be tested across individuals if they are confronted with the very same valence conditions during task performance, it is mandatory to induce specific valence conditions that are shared by all individuals. These three different issues are elaborated upon in the next section.

4.3.1 *Temporal Characteristics of APPR*

As previously presented, the point in time at which an APPR occurs in a problem-solving scenario can only be reliably determined when an individual is forced by a task-related event to re-evaluate current problem-solving status with respect to a specific goal. For instance, in the case of a two-player game setup, this re-evaluation always results after an opponent's action with respect to one's own progress. Accordingly, resulting APPRs can be viewed as related to individual task performance and achievement.

Depending on the physiological signals of an APPR that are involved, specific characteristics of response duration have to be considered. For instance, as previously presented, skin conductance is a relatively slowly responding signal. It also requires some seconds to recover from an affective episode, i.e., return a zero slope, after an APPR has occurred. As a result, in a task setup that involves very short intervals between two different indicands, related APPR are very likely to converge. Accordingly, the two APPR interfere with each other, which complicates or at worst prevents an analysis.

Such a complication can be easily avoided by considering the maximum recovery time [1] of all physiological signals that are analyzed during task performance. In a well-suited task setup, the time interval between two of the opponent's action is large enough to allow all involved physiological signals to recover. For instance, in the context of skin conductance measures, the opponent's action should always occur in response to an action but not earlier than ten seconds after a previous action.

4.3.2 *Measuring Goal-Directed Valence*

In a well-defined problem-solving task, the overall goal status to be achieved is defined from the beginning. Accordingly, the problem status to be checked against the individual goal is the one that results after an opponent's action. The valence of a problem status can be

1 the time certain physiological signals require to return zero slope after response to an event

discriminated in terms of goal conduciveness and goal obstruction by measuring the difference between two subsequent statuses. Consequently, by monitoring the game course, it is possible to measure the particular conditions of individual progress and distinguish them with respect to their different conditions of valence.

As presented in the previous chapter, the likelihood of encountering conditions of a particular valence significantly depend on the overall task performance of an individual. For instance, a well-performing individual is very likely to be confronted with more goal-conducive conditions than a poorly-performing individual when confronted with the very same task. In terms of a task that introduces opponent actions, this is also the case when these two different individuals are confronted with the very same opponent.

4.3.3 *Inducing Specific Valence Conditions*

Nevertheless, an opponent is able to manipulate the progress of a poorly performing individual. For instance, an opponent who is able to play an optimal game may decide to make a less efficient move to support a goal-conducive condition for a poorly-performing individual. As a result, it is also easy to achieve conditions of positive valence for poorly-performing individuals. This situation is more complex when well-performing individuals have to be confronted with a negative, i.e., goal obstructive, condition.

In a task where an individual can easily obtain full information about the problem, a well-performing individual is very unlikely to experience negative conditions of task progress. In such a case, any action can be thoroughly analyzed with respect to its consequences in terms of providing advantages for an opponent before it is actually performed. A well-performing individual will only perform actions that provide maximum benefit for herself and the fewest opportunities to an opponent. As a result, the likelihood of encountering a disadvantageous, i.e., goal-obstructive, condition is significantly lower.

From an information-processing perspective, this can be described as a situation in which an individual is able to represent the complete problem space and has efficient strategies for significantly reducing it with each step of the task. Depending on the specific task setup, this situation can be modified so that a game course does not solely depend on individual performance but also involves an element of chance. In this case, well-performing individuals can encounter a disadvantageous situation by accident. There are two different task characteristics supporting this:

1. The problem space exceeds human mental capabilities. The individual has to decide which part of it should be traversed first, and by chance a solution to the problem may be in the part that is not considered. In this way, a disadvantageous situation is encountered by accident.

2. The information that is provided about the problem is only partial. As a result, specific actions have to be selected with respect to individual hypotheses about their outcome. Whether or not these hypotheses are met not only depends on general task performance but also on pure chance. Accordingly, well-performing individuals are confronted with goal-conducive task conditions by accident.

One popular example that fits both of these characteristics is the Mastermind task. In the following section, it is presented in more detail, along with the potential for modifying this task to intentionally manipulate valence conditions.

4.4 MASTERMIND TASK SETUP

Mastermind is a two-player board game that has been very popular since the 1970s. The goal of the game is to guess an arbitrary four-digit code of colored pegs that has been selected by one of the two players. This player, who is termed Code Maker (CM) [Knuth, 1977, p.1], selects the combination out of six possible colors (multiple use of colors is allowed) and hides it from the second player. This player, termed Code Breaker (CB) [Knuth, 1977, p.1], tries to break the code by providing a guess to CM in terms of a four-digit code. CM provides feedback on this guess using black and white pegs; the number of black pegs indicates how many digits are already of the correct color and in the correct position, the number of white pegs indicates how many digits are of the correct color but in the wrong location. The feedback leaves open which peg corresponds to which element in the code. This cycle of guessing and receiving feedback is repeated until CB breaks the code (CM provides feedback of four black pegs) or a maximum number of steps is reached.

Figure 7 on the following page illustrates one round of this game. In this example, CM has selected YELLOW, BLUE, GREEN, RED as the hidden code. To begin, CB provides BLUE, GREEN, ORANGE, MAGENTA as a guess. This guess contains two colors (BLUE and GREEN) that are in the hidden code but are in the wrong positions. As a result, CM provides two white pegs as feedback.

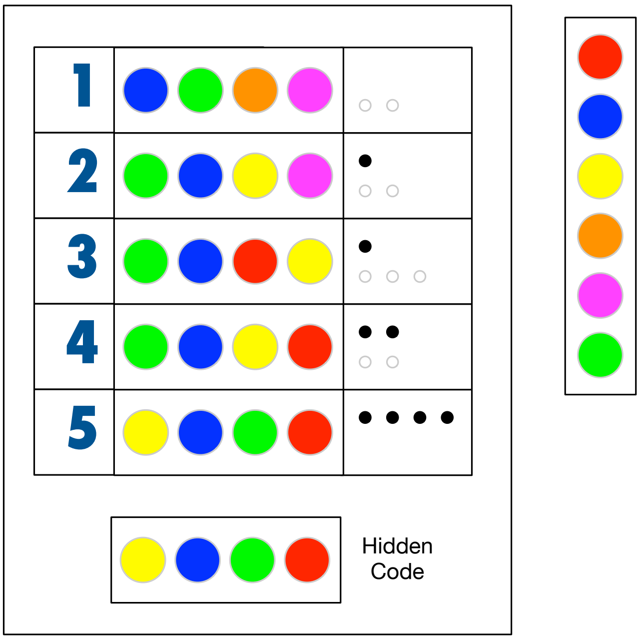

Figure 7: Illustrating a Mastermind game using colored pegs.

In the next step, CB provides GREEN, BLUE, YELLOW, MAGENTA as a guess. This code contains one color (BLUE) at the correct position and two colors (GREEN and YELLOW) that are also contained in the hidden code. Based on this, CM provides one black peg and two white pegs as feedback.

The game proceeds in this manner. In the fifth step, CB provides the guess (YELLOW, BLUE, GREEN, RED) that matches the hidden code. CM provides feedback indicating that all colored pegs are at the correct positions, four black pegs.

Based on this original task, Mastermind has been widely used for analyzing human problem-solving behavior. For this application, CM is replaced by a computer player, which sets the code randomly and generates the corresponding feedback for any game step accordingly.

4.4.1 *Dealing with a Large Problem Space*

When participants solve a Mastermind game, they have to deal with an abundance of information. At the beginning, there are very many potential solutions, which results in a high uncertainty about the hidden code. To decrease this uncertainty, participants have to consider the feedback provided for each step. Through this process, the set of potential solutions shrinks with each step until the correct solution can be selected.

For such a game, the level of difficulty depends on the number of possible solutions for a code. This number results from the number of elements that constitute the final code, the number of colors that can be selected and whether multiple uses of colors is allowed or not (see Table 1). Depending on whether multiple color use is allowed or not, there are 1296 or 360 potential solution codes for the example game that is illustrated in Figure 7 on the previous page.

Table 1: Solution set sizes of different Mastermind games with increasing levels of code length, depending on whether multiple use of color is allowed.

n digits out of c colors	Solution set size Multiple color use	Solution set size Single color use
2 out of 4	16	12
3 out of 5	125	60
4 out of 6	1296	360
5 out of 8	32768	6720

Optimal strategies

For each Mastermind game that is defined by n and c, there is a $f(n, c)$ which determines the minimum number of guesses that are needed to find an arbitrary code k. $f(n, c)$ greatly depends on the strategy that is used to solve the problem.

Many such optimal strategies have been computed for small values of n and c. As Mastermind is NP-hard [Stuckman and Zhang, 2006], a calculation for bigger numbers has not been possible yet.

For a Mastermind game with $n = 4$ and $c = 6$, Knuth put forward a deterministic strategy that allows it to be solved in five steps (i.e., four guesses are necessary to determine the code k, the last one represents k) [Knuth, 1977]. In [Koyama and Lai, 1993], a strategy was developed with $f(n, c)$= 4.34 (i.e., also five total guesses are necessary). For $n \leqslant 8$ and $c \leqslant 10$, a number of optimal winning strategies are known whose resulting $f(n, c)$ is a result of computation [Goddard, 2004, Jäger and Peczarski, 2009]. The resulting numbers are presented in Table 2 on the following page.

There is only a limited number of results that actually prove sharp bounds for larger numbers. In early findings, the asymptotic number of attempts necessary to guess a code for all numbers of n and c with $k \leqslant n^{1-\epsilon}, \epsilon \succ 0$ constant, have been established up to a constant factor in precision [Chvátal, 1983]. Recently, sharp bounds have also been proven for $c = n$; such Mastermind games can be solved in $O(n \log n)$ guesses [Doerr et al., 2012].

p \ c	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	1	3	3	4	4	5	5	6	6	7
3	1	3	4	4	5	5	6	**6**	**6**	**7**
4	1	4	4	4	5	5	**6**	**6**	**7**	**7**
5	1	4	4	5	**5**					
6	1	5	5	**5**						
7	1	5	**5**							
8	1	6								

Table 2: Computed values of minimum $f(n, c)$ for $n \leqslant 8$ and $c \leqslant 10$. Bold type represents results from [Jäger and Peczarski, 2009]. Plain type indicates computational results of [Goddard, 2004].

4.4.2 *Human Problem-Solving Strategies*

Performance in the Mastermind game is defined by the number of code-feedback steps needed for inferring the hidden code. Good performers need fewer steps than average performers; poor performers need more. As the reduction of the solution set is highly dependent on the exploitation of feedback, good performers differ from poor performers in two ways: first, in the exploitation of provided feedback [Best, 1990], and second, in the use of strategies for receiving useful feedback that allows the individual to reduce the solution set [Hussy, 1989, Laughlin et al., 1982].

In [Hussy, 1989], the Mastermind game is viewed as an example of handling an abundance of information. In the beginning, there is high uncertainty about the hidden code, resulting from the large number of potential solutions. This uncertainty is decreased by considering the feedback provided for any step. The set of solutions that are consistent with each given feedback decreases with every step until the correct solution can be selected. Following this, performance during the Mastermind task is highly dependent on the strategies used to get useful feedback that diminishes the solution set.

The authors claim that good performers differ from average and poor performers in the use of information-reduction strategies. Good performers benefit from efficient mental operations that allow consideration of a few hypotheses by simultaneously utilising all necessary information. Average and poor performers lack these efficient mental operations and must use information-reduction strategies to minimize the solution set. Potential strategies are to (i) sequentially identify the colors in a code, (ii) to sequentially infer the correct position of a colored peg, and (iii) to reduce the amount of information by only considering feedback that has recently been provided, e.g. in the last or next to last feedback step. Average performers differ from poor performers in that they utilize one of the three strategies in a successful way. Poor performers either unsuccessfully use such strategies, e.g. their differing performance is the result of incorrect inference steps due to an information overflow, or do not follow any strategy.

[Laughlin et al., 1982] found evidence in good performers for the use of two strategies whose application results in better performance in a simplified Mastermind task, the focusing and the tactical strategies. When applying the focusing strategy, subjects aggregate hypotheses about the hidden code with respect to different dimensions. These assumptions are validated by integrating them into one solving step rather than testing a single hypothesis. Subjects who use a tactical strategy provide guesses whose feedback information results in a split of the set of solutions. As a result, there are two solution sets of approximately the same size: one containing tenable solutions and the other containing untenable solutions with respect to the anticipated feedback.

The identified strategies differ in their demands on cognitive processing: whereas the focusing strategy is less demanding and is theoretically less efficient, the tactical strategy makes strong demands on the inference systems and insight into the anticipated feedback. The authors claim that subjects stick to the use of a certain strategy throughout a solving process, and that the first attempt already identifies which strategy a subject is following.

[Best, 1990] argues that subjects do not follow a certain strategy throughout the solving process but rather act strategically with respect to their success in information exploitation. Good performers differ from average or poor performers in that the information gain resulting from the anticipated feedback to their provided codes is very high. Therefore, performance in solving Mastermind tasks does not depend on the application of a certain strategy but rather on the employment of particular classes of feedback patterns that result in particular classes of hypotheses about the solution.

4.4.3 *Role of Chance*

In the above-mentioned classical version of the Mastermind game, chance plays a role with respect to identifying significant parts of the code by accident. This is particularly evident with regard to the first hypothesis given by CB. If, by chance, this hypothesis matches the code to a certain degree, a smaller number of subsequent steps

Hidden Code: 2 1 5 0 Guess: 1 5 3 4 FB: 0.2

0 1 2 3	0 1 2 5	0 1 4 2	0 1 5 2	0 2 1 3	0 2 1 5	0 2 4 1
0 2 4 3	0 2 4 5	0 2 5 1	0 2 5 3	0 3 1 2	0 3 2 1	0 3 2 5
0 3 4 2	0 3 5 2	0 4 1 2	0 4 2 1	0 4 2 3	0 4 2 5	0 4 5 2
2 0 1 3	2 0 1 5	2 0 4 1	2 0 4 3	2 0 4 5	2 0 5 1	2 0 5 3
2 1 0 3	2 1 0 5	2 1 4 0	2 1 5 0	2 3 0 1	2 3 0 5	2 3 1 0
2 3 4 0	2 3 5 0	2 4 0 1	2 4 0 3	2 4 0 5	2 4 1 0	2 4 5 0
3 0 1 2	3 0 2 1	3 0 2 5	3 0 4 2	3 0 5 2	3 1 0 2	3 1 2 0
3 2 0 1	3 2 0 5	3 2 1 0	3 2 4 0	3 2 5 0	3 4 0 2	3 4 2 0
4 0 1 2	4 0 2 1	4 0 2 3	4 0 2 5	4 0 5 2	4 1 0 2	4 1 2 0
4 2 0 1	4 2 0 3	4 2 0 5	4 2 1 0	4 2 5 0	4 3 0 2	4 3 2 0
5 0 1 2	5 0 2 1	5 0 2 3	5 0 4 2	5 1 0 2	5 1 2 0	5 2 0 1
5 2 0 3	5 2 1 0	5 2 4 0	5 3 0 2	5 3 2 0	5 4 0 2	5 4 2 0

Table 3: The 84 remaining solution codes for a guess of 1 5 3 4 in a 4-of-6, single-color-use Mastermind game with hidden code 2 1 5 0. The presented feedback pattern is 0.2.

might be necessary to identify the code. This results from the fact that incidental matches significantly reduce the number of potential candidates for the code; the amount of information to be considered is reduced. As a result, the same game can differ with respect to difficulty for participants of the same performance group.

To illustrate this, consider a Mastermind game with a four-digit code consisting of the following elements: 0, 1, 2, 3, 4, 5. Elements can only be contained once in the code. As a result, a guess like 2 5 5 0 is considered not valid during such a game, as 5 appears more than once. For such a game, there are 360 potential solution codes.

Let us consider a solution code 2 1 5 0. Depending on how closely CB's first guess matches the hidden code, the number of possible solutions in the solution set decreases. For instance, if CB provides 1 5 3 4 as a guess, two colors match the hidden code. As a result, CB receives 0.2 as feedback. Based on this, there are still 84 remaining solution codes in the solution set (see table 3). In the case that CB provides 5 1 2 4 as a guess for the first step, there is already one matching color position as well as two matching colors. As a result CB receives 1.2 as feedback. Based on this, there are 72 remaining solution codes (see table 4 on the following page). The situation is even more positive in the case that CB provides 5 1 0 2 as an initial guess. In this case, there is one matching color position as well as three matching colors. As a result, CB receives 1.3 as feedback. Based on this, there are eight potential solution codes (see table 5 on the next page).

To overcome this issue, [Funke and Hussy, 1979] suggest preselecting the first feedback that is given to the participant, i.e., that each

Hidden Code: 2 1 5 0		Guess: 5 1 2 4		FB: 1.2		
0 1 4 2	0 1 4 5	0 1 5 2	0 2 1 4	0 2 5 4	0 4 2 1	0 4 2 5
0 5 1 4	0 5 2 1	1 0 2 5	1 0 5 4	1 2 0 4	1 2 3 4	1 3 2 5
1 3 5 4	1 4 2 0	1 4 2 3	1 5 0 4	1 5 2 0	1 5 2 3	1 5 3 4
2 0 1 4	2 0 5 4	2 1 0 5	2 1 3 5	2 1 4 0	2 1 4 3	2 1 5 0
2 1 5 3	2 3 1 4	2 3 5 4	2 5 0 4	2 5 3 4	3 1 4 2	3 1 4 5
3 1 5 2	3 2 1 4	3 2 5 4	3 4 2 1	3 4 2 5	3 5 1 4	3 5 2 1
4 0 2 1	4 0 2 5	4 1 0 2	4 1 0 5	4 1 3 2	4 1 3 5	4 1 5 0
4 1 5 3	4 3 2 1	4 3 2 5	4 5 2 0	4 5 2 3	5 0 1 2	5 0 4 1
5 0 4 2	5 2 0 1	5 2 1 0	5 2 1 3	5 2 3 1	5 2 4 0	5 2 4 3
5 3 1 2	5 3 4 1	5 3 4 2	5 4 0 1	5 4 0 2	5 4 1 0	5 4 1 3
5 4 3 1	5 4 3 2					

Table 4: The 72 remaining solution codes for the guess 5 1 2 4 in a 4-of-6, single-color-use Mastermind game with hidden code 2 1 5 0. The presented feedback pattern is 1.2.

Hidden Code: 2 1 5 0		Guess: 5 1 0 2		FB: 1.3		
0 1 2 5	0 5 1 2	1 0 5 2	1 2 0 5	2 1 5 0	2 5 0 1	5 0 2 1
5 2 1 0						

Table 5: The 8 remaining solution codes for the guess 5 1 0 2 in a 4-of-6, single-color-use Mastermind game with hidden code 2 1 5 0. The presented feedback pattern is 1.3.

participant receives the same feedback response after the first guess. To achieve this, the hidden code is determined after a participant has given her first guess, based on that guess and the preselected feedback pattern. In this way, difficulty levels are uniform across participants for the first step. Nevertheless, chance matches are still possible for subsequent steps.

The central point of Hussy's approach is that it allows selection of a specific feedback pattern that is provided to every individual in the first step. Based on this idea, one can imagine a version of the game in which subsequent game steps can be adapted in terms of provided feedback and the hidden code adapted accordingly. This constitutes the starting point for developing *Luckless Mastermind*.

5 LUCKLESS MASTERMIND

In this chapter, an extension of the popular Mastermind game is introduced. The main idea of *Luckless Mastermind* is to remove the possibility of guessing significant parts of a hidden code by chance, presenting individuals with preselected patterns of feedback. Due to special characteristics of the Mastermind game, this is possible to do while maintaining sound problem solving. Thus, the same sequence of feedback can be provided to different individuals, resulting in the same situation of goal obstruction or conduciveness. This is achieved by developing a Mastermind variant in which CM is allowed to cheat and adapt a hidden solution to match a preselected feedback pattern. In this way, it is possible to provide comparable conditions across participants.

Here, the basic idea of Luckless Mastermind is introduced. Specific game characteristics restrict the use of particular feedback patterns as well as the opportunity to preselect feedback for more than one step. In this context, whether codes contain a single color multiple times or not is of great importance. Accordingly, a subsequent analysis of the potential to provide preselected feedback is divided into two parts. First, games in which a single color can be used several times are analyzed with respect to their potential for providing luckless but sound problem-solving conditions. Second, these issues are presented in detail for games in which a color may only be used once in a code. Based on this, an algorithmic approach is presented that allows specification of particular feedback.

5.1 THE BASIS OF LUCKLESS MASTERMIND

The main idea of Luckless Mastermind is to return every individual who provides an arbitrary combination of guesses with the very same feedback pattern. In this respect, any provided feedback patterns must not violate a sound problem-solving process in the sense that it contradicts already received information. Accordingly, it may not represent false feedback in this respect.

The opportunity to flexibly manipulate the individual solving process is based on the large number of possible combinations at the beginning of a game. At a single step, it is not possible to provide a guess that could limit this number in such a way that only a single solution remains. Rather, limiting the number of potential codes is done by sequentially checking guesses to arrive at the single remaining one.

To illustrate this, let us consider the starting point of Luckless Mastermind, a variant of the original game. In this game variant, there are also two players, one setting the hidden code (CM) and one identifying the code by providing guesses (CB). CM's role during the game is changed from a passive to a more active one. In the classic version, CM simply sets the code and provides feedback to any received guess. In the dynamic case, CM is allowed to modify the hidden combination, e.g. in case CB provides a highly matching guess by chance during the game. Nevertheless, the modified code has to be congruent with all previous guesses and feedback patterns in that game. The opportunity to change a hidden code to another one is based on the fact that for any step in a Mastermind game, there is a modified set of remaining codes to represent the hidden one. If the current hidden code is changed to one of these codes, the problem-solving process is still sound. Depending on this novel, selected, hidden code, a specific feedback pattern is presented to CB. In this case, potential codes are selected based on a single view of the set of remaining codes. In the case of Luckless Mastermind, this view is adapted so that there is no single set of remaining codes but a partition of remaining codes depending on each available feedback pattern.

5.1.1 *Feedback-Driven Distribution of Remaining Codes*

In contrast to viewing Mastermind as a game of identifying a hidden code or adapting it appropriately during the course of the game, in Luckless Mastermind the game is viewed as the definition of sets of remaining codes based on a preselected feedback pattern, set at the beginning. Accordingly, there is no single code at the outset. Instead there is a sequence of feedback patterns that provide a set of remaining codes after each step in the game. The latter issue is highly important, as the number of remaining codes not only depends on the specific feedback pattern but also on the guess that has been provided.

To illustrate this, let us consider a small example game in which codes consist of two elements, there are three colors to choose from, and colors may be used several times in a code. Below, the corresponding element set, the feedback set and the number of possible codes are presented.

In this game, there are five different feedback patterns, one for indicating that the hidden code as been identified (**2**.$\boxed{0}$), and three for

Element Set	0, 1, 2
Feedback Set	0.0, 0.1, 1.0, 0.2, 2.0
Possible Codes	0 0; 0 1; 0 2; 1 0; 1 1; 1 2; 2 0; 2 1; 2 2;

highlighting whether parts of a given guess are correct. Nevertheless, not all of these patterns can be preselected in a game. This is due to the fact that the number and characteristics of remaining codes given a specific feedback is always determined by the associated guess.

To illustrate this, let us consider two individuals; one providing 0 0 as a guess and another providing 0 1 as a guess. Depending on each of these guesses, there is a different number of remaining codes for each of the available feedback patterns. This is illustrated below. It becomes obvious that for certain feedback patterns, there is no remaining solution possible. Consequently, such patterns must not be preselected.

Guess	0.0	0.1	1.0	0.2	2.0
0 0	1 1; 1 2; 2 1; 2 2	—	0 1; 0 2; 1 0; 2 0	—	0 0
0 1	2 2	0 0; 2 0; 1 2 1 1;	0 0 0 2; 1 1 2 1	1 0;	0 1

The opportunity to preselect certain feedback patterns depends on specific game characteristics, i.e., the number of colors in the element set, the length of possible codes, and whether or not colors may be used multiple time in a code. The latter is most important for determining the opportunity to preselect feedback in a game. Accordingly, the following considerations of feedback sequence selection are differentiated with respect to whether multiple use of color is allowed or not.

5.2 PRESELECTING FEEDBACK: MULTIPLE COLOR USE

In a Mastermind game with multiple color use, there is a much larger number of possible codes at the beginning compared to a game of the same size with single color use (cf. Table 1 on page 59). Based on this, one could assume that the possibilities for preselecting specific feedback patterns are much higher. However, the opposite is actually the case: although there are a large number of potential candidates, the possibility for providing preselected feedback is severely limited.

Feedback (FB)	Possible Hidden Codes (HC)	# HC
0.0	all codes without 0 in it	625
1.0	all codes with a single 0 in it	500
2.0	all codes with two 0's in it	150
3.0	all codes with three 0's in it	20
4.0	0 0 0 0	1

Table 6: Distribution of remaining codes covered by each feedback pattern given the guess 0 0 0 0 in a game from a set of elements 0, 1, 2, 3, 4, and 5. All feedback patterns containing white pegs cannot be used, as no element can be part of the code but at the wrong position.

This is based on the fact that any preselected feedback pattern has to be valid for any potential guess that may be provided by CB. As we have seen in the example above, there is one feedback pattern that does not result in a remaining set of potential codes: given 0 0 as a guess, there is no remaining code for feedback pattern 0.1. This characteristic is shared by all feedback patterns containing white pegs in Mastermind games of varying size with multiple color use.

To illustrate this, let us consider a 4 of 6 Mastermind game with an element set consisting of 0, 1, 2, 3, 4, and 5 and a four-element code to be provided. For such a game, there are 14 different feedback patterns (black.white): 0.0, 0.1, 1.0, 0.2, 1.1, 2.0, 0.3, 1.2, 2.1, 3.0, 0.4, 1.3, 2.2, and 4.0. In such game, the guess 0 0 0 0 has been provided. For this guess, not all available feedback patterns represent valid information with respect to the guess. The number of remaining codes within each of these possible feedback patterns is presented in Table 6. In this specific game, it is impossible that one of the code's elements is part of the code but not in the correct position. Accordingly, valid feedback patterns must not contain white pegs, but only consist of various numbers of black pegs.

5.2.1 *Determining Number of Black Pegs*

The number of black pegs to be provided in a game largely depends on the specific game characteristics. To illustrate this, let us consider another Mastermind game with a two-element code and two elements to be selected for generating codes.

Element Set	0, 1
Feedback Set	0.0, 0.1, 1.0, 0.2, 2.0
Possible Codes	0 0; 0 1; 1 0; 1 1;

For two different guesses, i.e., 0 0 and 0 1, there are a different number of possible codes:

Guess	0.0	0.1	1.0	0.2	2.0
0 0	1 1	—	0 1; 1 0	—	0 0
0 1	—	—	0 0; 1 1	1 0	0 1

It becomes evident that there is an issue with respect to the number of black pegs that can be presented to any individual. To develop a general principle, let us consider a specific feedback pattern consisting of bl black pegs. Based on this feedback pattern, bl elements of the present code have been determined. The most important aspect in this case is to balance the number of provided black pegs and the remaining colors to select elements from.

To illustrate this, let us consider two different cases in a 4 of 4 Mastermind game where individuals should receive 0.0 as feedback and the corresponding elements set is 0, 1, 2, 3.

In the first case, an individual provides 2 2 2 2 as a guess and is presented with the predetermined feedback pattern. Consequently, it is clear that a potential hidden code can only be generated by using the remaining colors, i.e., 0, 1 3.

In the second case, an individual provides 0, 1, 2, 3 and receives the same feedback. In this case, the provided feedback restricts the number of available colors to a large extent: any remaining code does not contain any of the already utilized colors. As there are no other colors to select from, no appropriate code can be generated. As a result, 0.0 is not a valid feedback pattern in this game.

In general, feedback patterns consisting of only black pegs and no white pegs indicate both the number of correct and not contained colors in a code. As individuals are not restricted in generating their particular guess, they can always present a code that contains different colors for as many of its elements as possible. In case of a feedback pattern bl.0, it is determined that bl colors of the element set are correctly contained. In addition, the remaining elements are not contained. If we consider len as length of a specific code and col as number of available colors in the element set, it holds that bl eliminates $len - bl$ potential colors.

If $bl = 0$, len number of colors are ruled out. Accordingly, the number of remaining colors is defined as $col_{remain} = col - len$. To be able to generate the remaining codes, there have to be remaining colors to select from. Accordingly, it has to hold that $\mathtt{col} - \mathtt{len} > 0$, if 0.0 should be provided as guess.

5.2.2 *Preselecting Sequences*

In a game allowing multiple uses of a color, it is not possible to preselect feedback patterns for more than one step. This is based on the fact that in such a game, the only possible feedback patterns contain information about the already correct position of elements, i.e., black pegs.

To illustrate this, let us return to the previous 4 of 6 Mastermind game and consider that CB provides 0, 1, 2, 3 as a first guess. For this initial step, there is a preselected feedback pattern that identifies a certain part of this code as already correct in color and position, i.e., it provides a number of black pegs. Due to the game characteristics, there are no white pegs involved.

As a second guess, CB provides a permutation of the first guess: 1 0 3 2. In this case, all the colors used in the previous guess have changed positions, so that any feedback pattern would violate the rule that no white pegs are allowed.

Based on this, it is not possible to preselect feedback for more than one step in a Mastermind game with multiple colors, due to the freedom of CB to provide arbitrary permutations of previous guesses. In addition, there is only a limited set of feedback patterns that can be presented in the first step.

5.3 PRESELECTING FEEDBACK: SINGLE COLOR USE

The situation is different if the available colors may not be used multiple times in a code. In this case, codes that contain a single element more than once are not valid codes and may not be used in the game. Consequently, the initial number of potential codes is smaller compared to games with multiple use of color (cf. Table 1 on page 59). However, this restriction in generating codes during the course of the game significantly increases the opportunity to provide preselected feedback.

To illustrate this, let us consider a 4 of 6 Mastermind game with an element set consisting of 0, 1, 2, 3, 4, and 5 in which colors may only be used once in a code. For such a game, there exist 14 different feedback patterns (black.white): 0.0, 0.1, 1.0, 0.2, 1.1, 2.0, 0.3, 1.2, 2.1, 3.0, 0.4, 1.3, 2.2, and 4.0.

Based on this, let us assume that an individual provides 1 2 3 4 as a first guess and receives 0.0 as feedback. In this case, none of

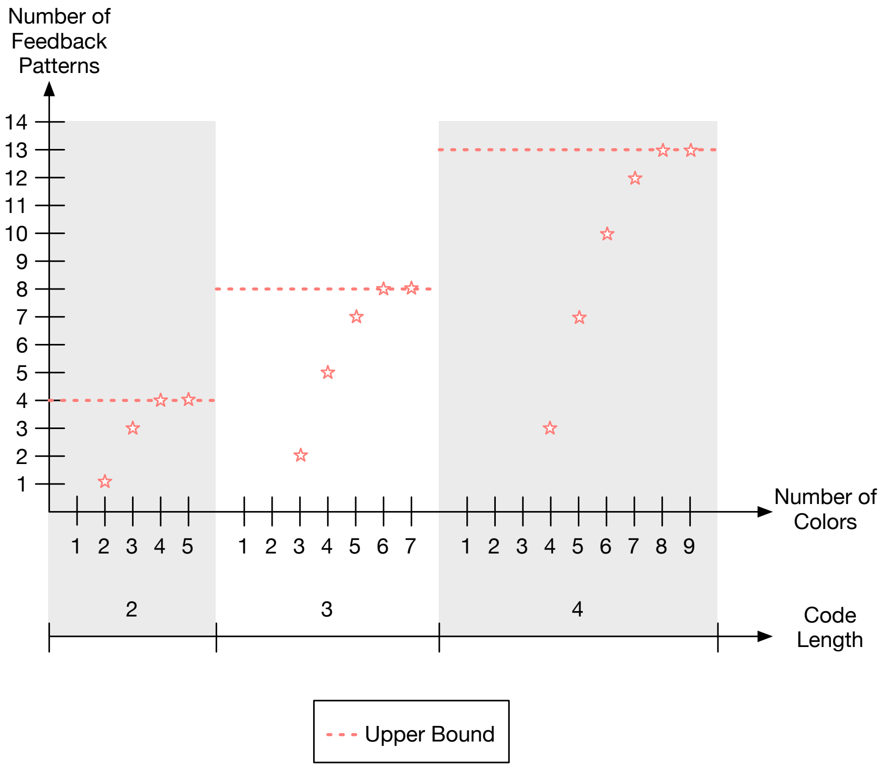

Figure 8: Number of feedback patterns that can be preselected develops with increasing values for len and col.

the elements are in the correct position (black pegs) and none of the elements are contained in the hidden code. Consequently, all colors that are part of this code are not part of the hidden solution. Based on this, there are only 0 and 5 as remaining colors. As colors may not be used more than once, it is not possible to generate a valid code.

A similar situation occurs in the case that 0.1 has been preselected as feedback. In this case, one element is part of the resulting code but the remaining three are ruled out. Accordingly, there are only three colors to select from: the one contained in the guess and the remaining two that have not yet been used. Again, this does not allow generation of a four-element code with single use of elements.

Thus it becomes evident that not all feedback patterns can be preselected in such games. More specifically, which feedback patterns can be preselected depends on the combination of elements of a single code, i.e, len, and the number of available colors in the element set, i.e., col.

In general, any feedback pattern consists of a certain number of black pegs, bl, and a certain number of white pegs, wh; each feedback pattern containing $F = (bl + wh)$ elements of a guess. Consequently, the remaining part of the code, not covered by the feedback, is not contained in a remaining code: $len - F$ elements are restricted from being used to generate codes. Consequently, this number of elements has to be removed from the available elements set. Based on this, remaining codes can only be generated based on the number of available colors: $\mathtt{col} - (\mathtt{len} - F)$. Only if there are enough elements left to set a guess' elements, remaining codes can be generated and the specific feedback pattern is valid in that game.

As a feedback pattern cannot contain more pegs than there are elements in a code, it holds that $F \leqslant \mathtt{len}$. Considering the previous issues, it also holds that:

$$\begin{aligned} & \mathtt{col} - (\mathtt{len} - F) \geqslant \mathtt{len} \quad \wedge \quad F \leqslant \mathtt{len} \\ \Leftrightarrow\ & F \geqslant 2 \cdot \mathtt{len} - \mathtt{col} \quad \wedge \quad F \leqslant \mathtt{len} \end{aligned} \tag{1}$$

Based on this, it is possible to compute upper and lower bounds for $\mathtt{len}$ of $\mathtt{col}$ for Mastermind games with single element use. Computed values for small numbers of $\mathtt{len}$ and $\mathtt{col}$ are presented in the Appendix in Table 18 on page 116. Figure 8 on the preceding page depicts the impact of different values for $\mathtt{len}$ and $\mathtt{col}$ on the number of feedback patterns that can be preselected. With increasing values for $\mathtt{col}$ and $\mathtt{len}$, there are more feedback patterns to select from for the first step.

5.3.1 *Preselecting Sequences*

In general, Mastermind tasks with single color use provide the opportunity to preselect a sequence of feedback patterns that can be presented to any individual, independent of the specific game course but not violating the basic rules of the game. However, determining such sequences is a complex task, as there is a strong interaction between an individual sequence of guesses and the preselected feedback patterns in terms of determining the set of remaining codes.

To illustrate this, let us consider three different individuals confronted with a Mastermind game as previously described: a 4 of 6 Mastermind game with an element set consisting of 0, 1, 2, 3, 4, and 5; elements may only be used once in a single code. For illustration purposes, all three individuals happen to start with the same first guess and receive the same feedback with respect to this. Subsequently, each of them continues the game quite differently.

The 1st step

In the first step of a game with single color use, any provided guess results in an equal number of remaining codes for each feedback pattern, only varying with respect to the actual permutations.

In our example, each of the three individuals provides 1 2 3 4 as a first guess. As illustrated in Table 7, the number of remaining codes differs for each feedback pattern; for some of them (0.0, 0.1, 1.0) there is no remaining code. A specific feedback pattern can only be preselected if there are still remaining codes. Based on this, let us suppose that 0.3 has been selected as the feedback pattern for the first step, as this offers the largest number of remaining hidden codes.

FB	Possible Hidden Codes (HC)	# HC
0.2	0125; 0145; 0152; 0153; 0315; 0325; 0345; 0351; 0352; 0415; 0425; 0451; 0452; 0453; 0512; 0513; 0521; 0523; 0541; 0542; 0543; 2015; 2045; 2051; 2053; 2105; 2150; 2305; 2350; 2405; 2450; 2501; 2503; 2510; 2540; 3015; 3025; 3045; 3051; 3052; 3105; 3150; 3405; 3450; 3501; 3502; 3510; 3520; 3540; 4015; 4025; 4051; 4052; 4053; 4105; 4150; 4305; 4350; 4501; 4502; 4503; 4510; 4520; 5012; 5013; 5021; 5023; 5041; 5042; 5043; 5102; 5103; 5120; 5140; 5301; 5302; 5310; 5320; 5340; 5401; 5402; 5403; 5410; 5420	84
1.1	0135; 0154; 0215; 0245; 0251; 0253; 0354; 0435; 0514; 0524; 0531; 0532; 1025; 1045; 1052; 1053; 1305; 1350; 1405; 1450; 1502; 1503; 1520; 1540; 2035; 2054; 2504; 2530; 3054; 3205; 3250; 3504; 4035; 4205; 4250; 4530; 5014; 5024; 5031; 5032; 5104; 5130; 5201; 5203; 5210; 5240; 5304; 5430	48
2.0	0235; 0254; 0534; 1035; 1054; 1205; 1250; 1504; 1530; 5034; 5204; 5230	12
0.3	0123; 0142; 0143; 0312; 0321; 0341; 0342; 0412; 0413; 0421; 0423; 2013; 2041; 2043; 2103; 2140; 2145; 2153; 2301; 2310; 2315; 2340; 2345; 2351; 2401; 2403; 2410; 2415; 2451; 2453; 2513; 2541; 2543; 3012; 3021; 3041; 3042; 3102; 3120; 3125; 3140; 3145; 3152; 3401; 3402; 3410; 3415; 3420; 3425; 3451; 3452; 3512; 3521; 3541; 3542; 4012; 4013; 4021; 4023; 4102; 4103; 4120; 4125; 4152; 4153; 4301; 4302; 4310; 4315; 4320; 4325; 4351; 4352; 4512; 4513; 4521; 4523; 5123; 5142; 5143; 5312; 5321; 5341; 5342; 5412; 5413; 5421; 5423	88
1.2	0124; 0132; 0213; 0241; 0243; 0314; 0324; 0431; 0432; 1023; 1042; 1043; 1302; 1320; 1325; 1340; 1345; 1352; 1402; 1403; 1420; 1425; 1452; 1453; 1523; 1542; 1543; 2014; 2031; 2104; 2130; 2135; 2154; 2304; 2354; 2430; 2435; 2514; 2531; 3014; 3024; 3104; 3154; 3201; 3210; 3215; 3240; 3245; 3251; 3514; 3524; 4031; 4032; 4130; 4135; 4201; 4203; 4210; 4215; 4251; 4253; 4531; 4532; 5124; 5132; 5213; 5241; 5243; 5314; 5324; 5431; 5432	72
2.1	0134; 0214; 0231; 1024; 1032; 1203; 1240; 1245; 1253; 1304; 1354; 1430; 1435; 1524; 1532; 2034; 2534; 3204; 3254; 4230; 4235; 5134; 5214; 5231	24
3.0	0234; 1034; 1204; 1230; 1235; 1254; 1534; 5234	8
0.4	2143; 2341; 2413; 3142; 3412; 3421; 4123; 4312; 4321	9
1.3	1342; 1423; 2314; 2431; 3124; 3241; 4132; 4213	8
2.2	1243; 1324; 1432; 2134; 3214; 4231	6
4.0	1234	1
For 0.0, 0.1, 1.0, there are no remaining codes.		

Table 7: Distribution of the 360 possible codes depending on each feedback pattern, given the initial guess 1 2 3 4 in a game from a set of elements 0, 1, 2, 3, 4, and 5.

The 2nd Step

For the 2nd guess, Individual 1 keeps three elements of the first guess, permutes their positions so that none of the elements matches its previous position and adds one element from the element set that has not been previously used. The resulting guess is 4315.

To be congruent with previous feedback, any preselected feedback in the 2nd step has to consist of at least two pegs. This would indicate that the color that has been dismissed in the second step is part of the potential solution code and that the newly added is not part of a potential solution code. As Individual 1 only removed a single element from the previous guess, it not possible to provide less than two pegs.

Nevertheless, for this individual, four pegs also represents a valid feedback. In that case, a wrong color has been dismissed in the second guess and the newly added color is contained in the potential solution code. In addition, there are several other possibilities for preselected feedback. In Table 8, the different remaining hidden codes, contingent upon specific feedback in the second step, are presented.

1. Guess: 1234		
FB	#HC	Possible HC
0.3	88	0123; 0142; 0143; 0312; 0321; 0341; 0342; 0412; 0413; 0421; 0423; 2013; 2041; 2043; 2103; 2140; 2145; 2153; 2301; 2310; 2315; 2340; 2345; 2351; 2401; 2403; 2410; 2415; 2451; 2453; 2513; 2541; 2543; 3012; 3021; 3041; 3042; 3102; 3120; 3125; 3140; 3145; 3152; 3401; 3402; 3410; 3415; 3420; 3425; 3451; 3452; 3512; 3521; 3541; 3542; 4012; 4013; 4021; 4023; 4102; 4103; 4120; 4125; 4152; 4153; 4301; 4302; 4310; 4315; 4320; 4325; 4351; 4352; 4512; 4513; 4521; 4523; 5123; 5142; 5143; 5312; 5321; 5341; 5342; 5412; 5413; 5421; 5423;
2. Guess: 4315		
FB	#HC	Possible HC
0.2	16	0123; 0142; 0421; 0423; 2041; 2043; 2103; 2140; 2401; 2403; 3021; 3042; 3102; 3120; 3402; 3420
1.1	12	0321; 0342; 0412; 2013; 2301; 2340; 2410; 3012; 4021; 4023; 4102; 4120
2.0	5	0312; 2310; 4012; 4302; 4320
0.3	17	0143; 2153; 2451; 2453; 2541; 2543; 3041; 3140; 3152; 3401; 3452; 3521; 3542; 5123; 5142; 5421; 5423
1.2	16	0341; 0413; 2145; 2351; 2513; 3125; 3410; 3425; 3512; 4103; 4152; 4521; 4523; 5321; 5342; 5412
2.1	7	0143; 2043; 2103; 2310; 2541; 3120; 5142
3.0	3	2145; 2340; 3140
0.4	4	4512; 4521; 5412; 5421
1.3	2	2451; 4152
2.2	4	2415; 2541; 4125; 5142
4.0	4	4315
For 0.0, 0.1, and 1.0, there is no remaining code.		

Table 8: Remaining hidden codes for Individual 1 depending on selected feedback pattern, assuming 4315 has been provided as the second guess.

Individual 2 keeps only two elements of the first guess, permutes their positions so that neither of the elements matches its position in the first guess and adds the two elements from the element set that have not been previously used. The resulting guess is 0512. In this case, Individual 2 might have selected the one element that is contained in the hidden code. Consequently, the other element of the previous guess has to receive at least one peg. In addition, as the previous guess received three feedback pegs, at least one of the two remaining elements must be contained in the hidden code. Based on this, to be congruent with previous feedback, the feedback on 2nd step must consist of at least two pegs. By the same argument, only one of the newly added elements can be correctly contained in a hidden code. As a consequence, not more than three feedback pegs are possible. In Table 9, the remaining hidden codes, depending on the feedback provided in the second step, are presented.

1. Guess: 1234		
FB	# HC	Possible Hidden Codes (HC)
0.3	88	0123; 0142; 0143; 0312; 0321; 0341; 0342; 0412; 0413; 0421; 0423; 2013; 2041; 2043; 2103; 2140; 2145; 2153; 2301; 2310; 2315; 2340; 2345; 2351; 2401; 2403; 2410; 2415; 2451; 2453; 2513; 2541; 2543; 3012; 3021; 3041; 3042; 3102; 3120; 3125; 3140; 3145; 3152; 3401; 3402; 3410; 3415; 3420; 3425; 3451; 3452; 3512; 3521; 3541; 3542; 4012; 4013; 4021; 4023; 4102; 4103; 4120; 4125; 4152; 4153; 4301; 4302; 4310; 4315; 4320; 4325; 4351; 4352; 4512; 4513; 4521; 4523; 5123; 5142; 5143; 5312; 5321; 5341; 5342; 5412; 5413; 5421; 5423;

2. Guess: 0512		
FB	#Codes	Possible Hidden Codes
0.2	22	2043; 2340; 2345; 2403; 2453; 3041; 3140; 3145; 3401; 3420; 3425; 3451; 4023; 4103; 4153; 4301; 4320; 4325; 4351; 5143; 5341; 5423
1.1	18	0143; 0341; 0423; 2543; 3042; 3402; 3410; 3415; 3452; 3541; 4013; 4302; 4310; 4315; 4352; 4523; 5342; 5413
2.0	4	0342; 0413; 3542; 4513
0.3	18	2041; 2103; 2140; 2145; 2153; 2301; 2351; 2401; 2451; 3021; 3120; 3125; 4021; 4120; 4125; 5123; 5321; 5421
1.2	16	0123; 0321; 0421; 2013; 2310; 2315; 2410; 2415; 2541; 3102; 3152; 3521; 4102; 4152; 4521; 5142
2.1	6	0142; 2513; 3012; 4012; 5312; 5412
3.0	4	0312; 0412; 3512; 4512
For 0.0, 0.1, 1.0, 0.4, 1.3, 2.2, 4.0, there is no remaining code.		

Table 9: Remaining hidden codes for Individual 2 depending on selected feedback pattern, assuming 0512 has been provided as the second guess.

Individual 3 keeps all elements of the first guess and creates a novel permutation in which none of the elements matches its previous position. Accordingly, the subsequent feedback pattern can only consist of three feedback pegs. Otherwise, elements previously reported as correct would have to be changed to incorrect or previously incorrect elements would have to become correct ones. In Table 10 on the following page, the remaining hidden codes, depending on the feedback provided in the second step, are presented.

By comparing these three different scenarios, it becomes evident that for each of them there is a different subset of possible feedback patterns. In addition, there are feedback patterns, e.g., 0.2, that repre-

Guess: 1234		
FB	#HC	Possible Hidden Codes (HC)
0.3	88	0123; 0142; 0143; 0312; 0321; 0341; 0342; 0412; 0413; 0421; 0423; 2013; 2041; 2043; 2103; 2140; 2145; 2153; 2301; 2310; 2315; 2340; 2345; 2351; 2401; 2403; 2410; 2415; 2451; 2453; 2513; 2541; 2543; 3012; 3021; 3041; 3042; 3102; 3120; 3125; 3140; 3145; 3152; 3401; 3402; 3410; 3415; 3420; 3425; 3451; 3452; 3512; 3521; 3541; 3542; 4012; 4013; 4021; 4023; 4102; 4103; 4120; 4125; 4152; 4153; 4301; 4302; 4310; 4315; 4320; 4325; 4351; 4352; 4512; 4513; 4521; 4523; 5123; 5142; 5143; 5312; 5321; 5341; 5342; 5412; 5413; 5421; 5423;

Guess: 4321		
FB	#Codes	Possible Hidden Codes
0.3	32	0142; 0143; 0412; 0413; 2013; 2043; 2103; 2140; 2145; 2153; 2403; 2410; 2415; 2453; 2513; 2543; 3012; 3042; 3102; 3140; 3145; 3152; 3402; 3410; 3415; 3452; 3512; 3542; 5142; 5143; 5412; 5413
1.2	32	0123; 0312; 0342; 0423; 2041; 2310; 2315; 2340; 2345; 2401; 2451; 2541; 3041; 3120; 3125; 3401; 3420; 3425; 3451; 3541; 4012; 4013; 4102; 4103; 4152; 4153; 4512; 4513; 5123; 5312; 5342; 5423
2.1	16	0341; 0421; 2301; 2351; 3021; 3521; 4023; 4120; 4125; 4302; 4310; 4315; 4352; 4523; 5341; 5421
3.0	8	0321; 4021; 4301; 4320; 4325; 4351; 4521; 5321;
For 0.0, 0.1, 1.0, 0.2, 1.1, 2.0, 0.4, 1.3, 2.2, 4.0, there is no remaining code.		

Table 10: Remaining hidden codes for Individual 3 depending on selected feedback pattern.

sent a possible feedback pattern in only some of these specific games (here, for Individual 1 and 2), but not in all of them. Based on this, there is only a limited number of feedback patterns (i.e., 0.3, 1.2, 2.1, 3.0) that are possible across all three situations. As a result, only these four feedback patterns represent feedback patterns that can be preselected for these three individuals.

Due to the fact that we have considered different game courses of only three different individuals that coincidentally started with the same guess, there is one main open issue. In this case, it remains unclear whether these four patterns represent viable feedback patterns for any possible game course. Individuals are flexible in selecting an arbitrary code out of the whole set of possible code combinations. Consequently, there are as many different game courses to proceed in a second step as there are valid code combinations. Accordingly, each of the four feedback patterns has to be checked to determine whether it is valid for all of these codes.

In addition, in a 4 of 6 Mastermind game, there are 10 different feedback patterns that can be preselected in the first step. Accordingly, all permutations of feedback sequences with increasing length have to be checked to determine specific sequences of feedback patterns that can be preselected.

5.3.2 *Number of Considered Game Courses*

If we assume that codes can only be provided once in a certain game, the available number of codes to be provided decreases by one with every additional game step. To illustrate this, let us consider that there are initially n potential codes to provide as guess. Subsequently, as the same guesses must not be provided twice in a game, there are $n-1$ codes to choose from. Consequently, in a third step, there are only $n-2$ possibilities to choose. Based on this, there are $n \cdot (n-1) \cdot (n-2)$ possibilities to be considered when a sequence of three feedback patterns has to be checked.

The following section presents an algorithmic solution for identifying sequences of feedback patterns that represent valid information across arbitrary game courses. To allow this in Mastermind games of varying dimensions, I first introduce formal requirements that are necessary to determine whether or not a certain feedback pattern is valid.

5.4 FORMAL REPRESENTATION

For each Mastermind game, there is an initial solution set S_0 that contains all possible codes. The number of elements in S_0 depends on the length of the code len, the number of colors col for building a code, and whether multiple uses of a color are allowed. For each game, there is an element set $\mathrm{colors} = \{1, 2, ..., \mathrm{col}\}$. In Table 1 on page 59 sizes of solution sets for Mastermind problems of different levels of difficulty are presented.

$$|S_0| = \begin{cases} \mathrm{col}^{\mathrm{len}} & \text{for multiple color use} \\ \frac{\mathrm{col}!}{(\mathrm{col}-\mathrm{len})!} & \text{for single color use} \end{cases} \tag{2}$$

For each specific Mastermind game, there is a specific set $\mathcal{F}$ consisting of all possible feedback patterns that can be represented by black and white pegs.

$$\begin{aligned} &\mathcal{F} = \{(\mathrm{bl}, \mathrm{wh}) \mid \mathrm{bl}, \mathrm{wh} \in \{0, 1, ..., \mathrm{len}\}\} \\ &\quad \text{with} \begin{cases} \mathrm{wh} = 0 \text{ for } \mathrm{bl} = \mathrm{len} - 1 \\ \mathrm{bl} + \mathrm{wh} \leqslant \mathrm{len} \text{ otherwise} \end{cases} \end{aligned} \tag{3}$$

For a 4 of 6 Mastermind game, this set is defined as $\mathcal{F} = \{0.0,\ 0.1,\ 1.0,\ 0.2,\ 1.1,\ 2.0,\ 0.3,\ 1.2,\ 2.1,\ 3.0,\ 0.4,\ 1.3,\ 2.2,\ 4.0\}$.

Human players are not restricted in providing a specific guess. As a result, each individuals can provide any possible code as a guess. Based on this, let us consider a guess g. For multiple color use, it holds that

$$g = (g_1, ..., g_{len}) \text{ with } g_i \in colors \tag{4}$$

For single color use, it holds

$$g = (g_1, ..., g_{len}) \text{ with } g_i \in colors \wedge g_i \neq g_j \ \forall i, j \tag{5}$$

At each step, a specific guess g and a specific feedback pattern F result in a certain number of remaining codes. These codes are determined based on the previous set of remaining codes. All codes that have been part of the previous set of codes S_{prev} and for which the specific feedback pattern F is correct with respect to the specific g represent a remaining code.

Based on this, let us consider the specific feedback pattern F [Chvátal, 1983] which represents how many elements of g are already correct or contained with respect to a solution s. For this, there is a function eq which determines how many elements are correct:

$$eq(f_1, f_2) := \begin{cases} 1 & f_1 = f_2 \\ 0 & otherwise \end{cases} \tag{6}$$

Assuming that $P(g)$ represents a set containing all permutations of g, F is defined as

$$\begin{aligned} F = &(a(g,s), b(g,s)) : \\ &a(g,s) = \sum_i eq(g_i, s_i) \\ &b(g,s) = \max_{g' \in P(g)} \sum_i eq(g'_i, s_i) \end{aligned} \tag{7}$$

Based on this, $bl = a(g,s)$ represents the number of correct colors with correct positions in the code, i.e., black pegs; $wh = b(g,s) - a(g,s)$ represents the number of correct colors with the wrong position in the code, i.e., white pegs.

For any previous set of codes S_{prev}, let us consider a set of remaining codes with respect to a certain g and a specific F

$$S_{curr} = \{c \in S_{prev} | \ (a(c,g), \ b(c,g)) = F\} \tag{8}$$

If $S_{curr} \neq \emptyset$, the specific feedback is not possible with respect to a specific guess and a specific set of previously possible codes. Therefore, it is not possible across arbitrary game courses and cannot be preselected.

5.5 PRESELECTING FEEDBACK: AN ALGORITHMIC APPROACH

Based on the considerations above, it is possible to develop algorithms that check whether or not a specific sequence of feedback patterns is valid across arbitrary game courses. To illustrate this, let us first consider an algorithm that checks a single feedback pattern. Subsequently, this algorithm is extended to check a specific sequence of such patterns.

5.5.1 *Checking a Single Feedback Pattern*

The main idea of the algorithm is to check whether all sets of remaining codes that result for each possible guess g with $g \in S_0$ are non-empty with respect to a specific previous solution set S_{prev} and a certain feedback pattern F.

For this, the algorithm takes every $c \in S_{prev}$ and calculates the resulting feedback pattern F_{comp} with respect to a certain $g \in S_0$. If this F_{comp} equals the feedback pattern to be tested, c represents a remaining code with respect to F. After all $c \in S_{prev}$ have been checked in this way, there is either a non-empty set of codes for which F is valid or not. In the latter case, F can be considered as being not valid with respect to guess g and S_{prev}. As a result, the algorithm returns $\mathtt{false}$.

When this check has been done for all $g \in S_0$ and none of them returned *false*, the feedback pattern is valid independent of what an individual provides as a guess given the specific previous game situation, i.e., S_{prev}. As a result, the algorithm returns *true*. This procedure is presented in algorithm 1 on the following page.

Based on these considerations, it is possible to extend the algorithm from checking a single feedback pattern to checking a specific sequence of them.

5.5.2 *Checking a Sequence of Feedback Patterns*

The above-stated algorithm can be extended to check sequences of feedback patterns by transforming it into a recursive algorithm that sequentially checks each element of a given list of feedback patterns.

Algorithm 1 A certain feedback pattern $\mathtt{fb}$ is checked to determine whether it is valid for a previous solution set S_{prev} and a set of possible guesses S_0.

```
function IsValidFB(F, S_prev, S_0)
    for all g ∈ S_0 do
        compute S_curr = {c ∈ S_prev | F_comp(c, g) == F}
        if S_curr = ∅ then
            return false
        else
            return true
        end if
    end for
end function
```

For every j-th element of this feedback list, the algorithm has to consider $|S_0|^{j-1}$ previous sets of remaining codes. Based on this, the algorithm considers a list of feedback patterns $\mathtt{fbList}$, a set of previous solution sets S_{prev}, and the set of all possible codes S_0 as input information.

When this algorithm is initially called, $S_{prev} = S_0$. Based on this, the first element of $\mathtt{fbList}$ is selected as feedback pattern F to be checked. The algorithm then iterates over all elements of the previous set of remaining codes and all possible guesses $g \in S_0$. Based on this, it checks if there is an empty remaining solution set for any of these guesses. In this case, the current feedback is not valid. As a result, the whole sequence cannot be valid and the algorithm returns $\mathtt{false}$. In the other case, the computed solution set is stored as one of the current solution sets S_{curr}. This process is repeated for any of the previous sets of remaining codes.

If all subsets of S_{prev} have been checked and none of them resulted in a $\mathtt{false}$, the algorithm proceeds by checking whether the current F is the last element of the feedback list. If so, none of the previous checks has resulted in an empty solution set. Hence, the feedback sequence is valid across all game courses and $\mathtt{true}$ can be returned.

If there are still elements in the feedback list, these have to be checked with respect to the current sets of remaining codes S_{curr}. As a result, the algorithm is called with the tail of the current feedback list, S_{curr}, and the set of all possible guesses S_0.

In short, the algorithm performs a breadth-first search on the tree of remaining code sets. If all of these sets on one level are non-empty, it proceeds with the next level. This procedure is presented in algorithm 2 on the next page.

Based on this, the algorithm considers such sequences of feedback as valid if these represent a repetition of a single feedback pattern. This is due to the fact that the algorithm determines remaining codes

Algorithm 2 Checks sequences of feedback $\mathtt{fbList}$ based on a set of previous solution sets S_{prev} and all possible guesses S_0. Performs a breadth-first search on all solution sets of a certain level.

function IsValidFeedbackSequence(fbList, S_{prev}, S_0)
  $F \leftarrow \mathit{fbList.First}$
  for all $S \in S_{prev}$ **do**
    for all $g \in S_0$ **do** compute $S_{remain} = \{c \in S | F(c, g) == F\}$
      if $S_{remain} = \{\emptyset\}$ **then**
        return false
      else
        $S_{curr} \leftarrow S_{curr} \cup \{S_{remain}\}$
      end if
    end for
  end for
  if $|\mathit{fbList}| = 1$ **then**
    return true
  else
    IsValidFeedbackSequence($\mathit{fbList.Tail}$, S_{curr}, S_0)
  end if
end function

with respect to all possible codes for each step of the game. This implies that a previous guess is also considered a valid guess for any subsequent step. Subsequent feedback patterns have to be valid with respect to any possible guess. For a guess that equals one of a previous step, the only valid feedback pattern is the one that has been previously provided. Hence, only feedback sequences that determine a repetition of a single feedback pattern are considered valid.

Due to the fact that previous guesses are always displayed for the individual, there is no need to provide the same guess multiple times. Therefore, the algorithm has to be modified so that it takes into account the fact that a certain guess can only be provided once in a game.

5.5.3 *Algorithm for Sequence Checking to Consider Uniqueness of Guesses*

When guesses are assumed to be provided only once in an individual Mastermind game, the algorithm has to check whether every possible guess results in a non-empty set of remaining codes only with respect to codes that have not been provided previously. For this, every previous set of remaining codes has to be checked with respect to the individual sequence of guesses it resulted from.

Therefore, the algorithm is modified so that it can check elements of a feedback sequence $\mathtt{fbList}$ with respect to a set of previous guesses G_{prev}, a previous solution set and a certain guess g. To initiate testing a certain sequence of feedback patterns, the algorithm is called with an arbitrary code g, $S_{prev} = S_0$, and $G_{prev} = \emptyset$.

Before sequence checking is started, it tests whether the list of feedback has been completely traversed and if possible guesses have been considered. If this is the case, all elements of the feedback list have been checked with respect to all possible guesses. Based on that, it can be concluded that the feedback list is valid; the algorithm returns `true`.

If this is not the case, the current guess g is tested as to whether it has already been provided as a guess. If this is the case, this guess is skipped and the next possible guess $g \in S_0$ is checked with respect to the feedback list `fbList`, the previous set of remaining codes S_{prev}, and the set of previous guesses G_{prev}.

If the guess has not been provided before, the first element of the feedback list is checked. For this F the remaining solution set S_{curr} is computed. If it results in an empty set of remaining codes, the specific feedback pattern is not valid with respect to the specific guess. As a result, it is not valid across arbitrary game courses. Based on this, the overall feedback list cannot be possible across arbitrary game courses and the algorithm returns `false`.

If there is a non-empty set of remaining codes, the rest of the feedback list (`fbList.Tail`) can be checked with respect to this solution set (S_{curr}) and the set of previous guesses including the current one ($G_{prev} \leftarrow G_{prev} \cup g$), starting with the first code of the set of all possible codes S_0. Consequently, the algorithm is recursively called with these adapted parameters. If this call returns `false`, the remaining part of the feedback sequence is not valid. Consequently, the overall sequence is not valid across arbitrary game courses and the algorithm returns `false`.

If the call returns `true`, checking the rest of the feedback list was positive with respect to the specific guess g. Based on this, this remaining feedback list has to be checked with respect to the next possible guess. For this, the current guess is removed from the set of previous guesses G_{prev}, as it is not a previous guess of the next guess to be checked but rather occurs at the same step of the game.

In short, the algorithm does a depth-first search for empty solution sets. Its procedure is presented in algorithm 3 on the following page.

5.5.4 *Computed Sequences of Feedback*

Utilizing the algorithm, sequences of feedback patterns can be determined for Mastermind games of specific dimensions. Due to polynomial time, sequences are computed for small values of `len` and `col`.

Algorithm 3 Checks sequences of feedback whether they are valid across all possible game courses. Does not allow repetition of guesses.

function IsValidFBSeq($fbList$, S_{prev}, S_0, g, G_{prev})
    if $fbList == empty || g == S_0.Last$ **then**
        return $true$
    end if
    if $g \in G_{prev}$ **then**
        $g \leftarrow S_0.(IndexOf(g) + 1)$
        return IsValidFBSeq($fbList$, S_{prev}, S_0, g, G_{prev})
    else
        $fb \leftarrow fbList.First$
        compute $S_{curr} = \{c \in S_{prev} | fb(c, g) == fb\}$
        if $S_{curr} = \{\varnothing\}$ **then**
            return $false$
        else
            $G_{prev} \leftarrow G_{prev} \cup g$
            $g \leftarrow S_0.First()$
            if IsValidFBSeq($fbList.Tail$, S_{curr}, S_0, g, G_{prev}) **then**
                $G_{prev} \leftarrow G_{prev} \setminus g$
                $g \leftarrow S_0[IndexOf(g) + 1]$
                return IsValidFBSeq($fbList$, S_{curr}, S_0, g, G_{prev})
            else
                return $false$
            end if
        end if
    end if
end function

In Table 11 on the next page, the number of available sequences of preselected feedback are depicted.

For instance, in a 4 of 6 Mastermind game, there are 10 different sequences that allow preselection of feedback for two game steps. This is especially interesting as this game size represents an appropriate size for humans to deal with. These 10 sequences are (0.2, 0.2), (0.2, 1.1), (1.2, 0.3), (1.2, 1.)2, (1.2, 2.1), (2.1, 1.2), (1.1, 0.2), (1.1,1.1), (0.3,1.2), and (0.3,0.3).

		# Sequences with			
len	col	1 Step	2 Steps	3 Steps	4 Steps
1	2	1	–	–	–
1	3	1	1	–	–
1	4	1	1	1	–
1	5	1	1	1	1
2	2	1	–	–	–
2	3	3	–	–	–
2	4	4	2	–	–
2	5	4	2	–	–
2	6	4	3	–	–
2	7	4	3	–	–
2	8	4	3	1	–
2	9	4	3	1	–
2	10	4	3	1	1
2	11	4	3	1	1
2	12	4	3	1	1
3	3	2	–	–	–
3	4	5	1	–	–
3	5	7	3	–	–
3	6	8	–	–	–
3	9	8	4	–	–
3	10	8	4	–	–
3	11	8	4	–	–
3	12	8	4	1	–
4	4	3	3	–	–
4	5	7	6	–	–
4	6	10	10	–	–
4	7	12	4	1	–
4	8	13	4	1	–
4	7	13	4	1	–
5	5	4	10	–	–
5	6	9	7	1	–
5	7	13	13	–	–

Table 11: Computed numbers for quantity of sequences of feedback that can be preselected in games with small numbers for len and col.

Part III

EMPIRICAL INVESTIGATION

6 EMPIRICAL STUDY

Employing *Luckless Mastermind* allows an analysis of intra- as well as inter-individual differences in APPR. In the first case, intra-individual dynamics of APPR can be analyzed depending on different feedback conditions. In the latter case, APPR can be analyzed to show significant differences across individuals. Based on this, *Luckless Mastermind* has been employed as novel task paradigm for testing the four different hypotheses that have been presented in chapter 4.

6.1 EXPERIMENTAL DESIGN

The present study utilizes two games of *Luckless Mastermind* with preselected feedback sequences, one representing a positive feedback condition, i.e. goal conducive, the other representing a negative one, i.e goal obstructing. The game order is randomized between subjects to avoid sequence effects.

The goal obstructing feedback sequence is represented by providing the most negative feedback pattern for the first two steps of the game. This feedback is valid for all possible game courses and is represented by the 0.$\boxed{2}$. It represents the condition that a guess contains two correct colors but none of them is on its correct position.

The goal conducive feedback sequence is represented by providing 1.$\boxed{2}$ in the first step and 2.$\boxed{1}$ in the second step. In this case, an individual's guess of the first step contains two colors that are part of the hidden code and one that is already at its correct position. In the second step, there is an increase of colors with correct position. The provided guess contains two colors that are at the correct position and one that is part of the hidden code.

As preselected sequences are of limited length due to the game specifics, each Mastermind game proceeds as a common game after two steps. Consequently, the number of steps that is needed to solve the problem still depends on the individual's competence in playing Mastermind. Therefore, individual performance can be measured. Based on this, the second objective can be achieved by analyzing APPR depending on performance in the game. To achieve the third objective, the individual's *achievement orientation* can be assessed to analyze APPR with respect to this factor.

6.2 participants

24 subjects (14 women) volunteered to participate for monetary compensation; one session took 60 minutes. All of them were students at the university of Bremen. For the reason of not following the instructions, one participant has been excluded from the analysis. In addition, three participants have to be excluded from analysis due to signal measurement issues and missing data. Thus, the remaining sample of participants comprises N=20 individuals with a mean age of M = 23.55 (SD= 4.99, range: 19–43 years) consisting of n=12 female participants with a mean age of M=22.67 years (SD=1.56, range: 20–25 years) and n=8 male participants with a mean age M=24.88 (SD= 7.77, range: 19–43 years).

6.3 apparatus

Experimental instructions were presented in written form. The basic task was introduced using a common board game setup. The actual experimental task as well as its introductory advices were presented on a 15-in. Mac Book Pro with external mouse. The relaxation phase stimulus has been presented on a second 19-in. monitor attached to the laptop and positioned next to it, both at a viewing distance of approx. 50 cm.

Physiological data has been collected using a ProComp Infiniti encoder unit [Thought Technology, 2008] connected via fiber-optic cable and TT-USB adapter. For measuring cardiac activity, a pre-amplified electrocardiograph sensor has been attached to the sensor device. For measuring changes of skin conductance, a corresponding sensor has been attached that operates at 256 Hz.

The experimental setup has been implemented using Microsoft® .NET-Framework (C# and WPF) and the software development kit ProComp Infiniti capture library (TTLAPI) that is provided by the sensor manufacturer. The resulting software both controlled experimental task presentation and data collection.

6.4 procedure

The general study consisted of seven parts. First, participants were welcomed to the experiment and introduced to the upcoming experiment providing a cover story. Second, participants were asked to perform a task for assessing their skin conductance responsiveness. Third, it followed a practice phase to familiarize participants with the task. Fourth, participants underwent a relaxing phase of about five minutes duration that has been necessary to achieve a comparable baseline activation for all participants. Fifth, it followed the actual game phase during which participants played two games of Mastermind with preselected feedback and have their physiological responses measured. Sixth, there were a posttest assessment for identifying individual motivation and check whether participants had suspected the true aim of analyzing affective responses or the modification of the game courses. Seventh, participants were debriefed.

The experiment was conducted under the ethical guideline of the German Psychological Society (DGPs). To leave the participants unclear about the study's true aim, any terms that are associated to affect and emotion were not mentioned during the announcement of the experiment. Rather, the experiment's purpose was described as opportunity to collect physiological responses in an office environment that can be used as a basis for recognizing specific office activities such as web browsing, email reading, and game playing. Participants were told in advance that they will be assigned to one of these three experimental groups randomly.

In addition, participants were informed about the following criteria for exclusion from the experiment. Due to potential effects on autonomic functioning participants who i) suffer from mental-health problems like e.g. depression, mania, personality disorders, and borderline psychosis, ii) take ataractics against one or more of the previously mentioned mental impairments, iii) take any drugs like e.g cannabis, ecstasy, and LSD, iv) be on medication that impacts perception or the autonomic nervous system (e.g. beta blocker), or v) suffer from neurological impairments like disseminated sklerosis, severe craniocerebral injury, and epilepsy were not allowed to take part in the study. The same holds if people were suffering from hyperhidrosis of hands, due to potential resulting issues in measuring skin conductance. As caffein intake may also impact peripheral physiological responses, participants have been asked not to consume caffein or alike at least 2 hours before participating in the experiment.

6.4.1 *Welcome*

After a participant had arrived in the laboratory, she gave informed consent to take part in the study and had time to accustom herself in this novel environment. Afterwards, she was informed about being assigned to the experimental group whose physiological responses are collected during a game playing setup. She was informed that she will play a computerized version of the board game Mastermind with a four-element code out of six colors and single color use. In addition, she was informed that a maximum of ten guesses is allowed per game and that performance was measured solely based on the number of steps that are needed to identify a hidden code. By this, time pressure is omitted to exclude potential interaction effects with APPR. For encouraging her to do her best, she was told that monetary compensation is increased for good performers. She was asked to take a seat at a desk in front of the two presentation devices and a mouse.

6.4.2 *Test of Skin Conductance Responsiveness*

It has been shown that not all individuals are equally likely to show affective skin responses. Findings suggest that 5 to 10% up to even 25% of the population do not show skin conductance responses at all [Dawson et al., 2007]. Based on this, participants have to be assessed with respect to their individual responsiveness of skin conductance.

Any condition of pain or shock is assumed to result in a skin conductance responses, i.e. electrodermal orienting response [Boucsein, 2012]. Based on this, participants were asked to press a loud signal horn. Individuals who did not exhibit a significant skin conductance response (> 0.01 μS) within a 10 seconds responds after signal noise were assumed to be non-responders with respect to skin conductance and are excluded from analysis of physiological data.

6.4.3 *Practice Phase*

The aim of the practice phase was to ensure that any participant is familiar with solving a Mastermind problem, i.e she a) got the overall game idea and b) knew how to interpret the given information. This is achieved in a two-step process. First, participants received game playing instructions of a common Mastermind game in written form together with a board version of that game. Each participants was shown the same exemplary game course of the same Mastermind game. As the experimental setup involves playing Mastermind on a computer, participants were subsequently introduced to the corresponding Graphical User Interface (GUI).

Accustomizing to the GUI

To support a successful interaction with the GUI during the experiment, participants were introduced to its interaction design. For this, the GUI was presented and all participants were asked to interact with it a) to enter and submit a guess she wanted to provide and b) to proceed with a potential subsequent game. To provide equal conditions to all participants, there was no actual game but any provided guess is acknowledged as correct guess. Participants were informed about his fact. Time and number of trials were not restricted in this phase and participants were asked to decide when they felt comfortable with respect to the interaction. After a participant had reported on feeling familiar with the GUI, it has to be assessed whether she had correctly understood the information provided by feedback pegs.

Assessing Correct Information Exploitation

In a Mastermind task, it is important that individuals are able to correctly interpret the information that is given by a feedback pattern. Otherwise they perform badly in the task independent from their overall performance in the game. Therefore, any participant was assessed with respect to her individual ability to correctly interpret the information given by a feedback pattern.

For this, as suggested by [Funke and Hussy, 1979], each participant was asked to take the role as CM in the game. Instead of presenting guesses to the computer, she was confronted with single guesses provided by a hypothetical player and the potential hidden code. Her task was to decide for the matching feedback pattern and to provide it to a GUI. After she had provided that information, she was confronted with an answer of whether this feedback pattern has been correct or not. In the latter case, the correct feedback pattern was also presented. After this, she proceeded to the next practice phase example. This was repeated until she had provided the correct feedback three times in a row or 20 practice steps had been done. In the latter case, the participant could decide whether she wanted to have a second look at the written instructions and try for a second and last time or whether she wanted to be excluded from the experiment immediately[1].

6.4.4 *Achieving Baseline Activation*

For achieving a comparable physiological baseline activation, participants were exposed to a minimally activating baseline task. Watching a relaxing video has been identified as resulting in a more precise baseline activation than sitting quietly with eyes closed [Piferi et al., 2000]. In addition, selecting a specific stimulus for this phase allows to standardize baseline activation induction [Andreassi, 2007, p.344].

1 As all participants successfully passed the practice phase during the first trial, no participant had to be excluded for this reason

Based on this, participants watched a five minute underwater film depicting a coral reef in the South Seas that is accompanied by relaxing music. Participants were asked to sit quietly and relax. By this, it was ensured that all participants start the experiment in a comparable condition of being relaxed and minimally activated. In addition, as apparatus for measuring physiological signals have been attached before, there was also enough time to let skin conductance level restore from necessary skin preparation.

6.4.5 *Game Phase*

In this part of the experimental procedure, participants were asked to play the two games of *Luckless Mastermind*. Valence of individual progress has been experimentally manipulated by preselecting two specific feedback sequences in a 4 of 6 Mastermind game with single color use. One of the games represented being exposed to a positive progress condition, the other represents being exposed to a negative one. To rule out sequence effects, there were two experimental groups of equal size which differed in whether their group members start with the positive or the negative condition accordingly. Participants were assigned randomly to each group. During this phase, physiological responses have been measured.

Course of the Game

Each game was started by the participants, clicking a corresponding button on the GUI. Afterwards, they were presented with the GUI for playing a Mastermind game with a maximum number of ten steps, i.e. ten guesses could be provided at maximum for each game.

For the first two steps, participants provided their two respective guesses and received feedback according to the preselected sequence of feedback patterns. After these two steps, the game proceeded as a common Mastermind game. These issues were completely invisible facts facts for the participants, for whom the game implementation behaves as a common Mastermind game right from the beginning,

To achieve this, the implemented game application determined the set of remaining codes that are inline with both previous guesses and the related feedback patterns. After the game steps with fixed feedback, one of these codes was randomly selected and set as hidden code. For any other step, the game application checks any given code with respect to this hidden code and provides feedback accordingly. The game proceeded until an individual identified the hidden code or the maximum number of step, i.e. ten, was reached. In the first case, the participants were informed that they successfully identified the hidden code and that they can proceed with the next game by clicking a respective button. In the latter case, participants were informed that they did not succeed and the hidden color combination was presented. In addition, they were informed that they can proceed with the next game by clicking a respective button. After both games had been completed, participants were disconnected from the apparatus for measuring physiological signals.

6.4.6 *Posttest Assessment*

To allow an analysis of APPR with respect to the individual achievement orientation, it was necessary to assess this individual factor. Therefore, each participant was asked to complete a questionnaire of the Freiburger Persönlichkeitsinventar (revised version) (FPI-R) [Fahrenberg et al., 2010].

6.4.7 *Debriefing*

To check whether participants have become suspicious about the true aim of the study, their opinion about the study's aim had been assessed. In case a participant would reject the cover story and would mention the measurement of affective states, data would have to be excluded from analysis[2].

After this, participants were informed about the study's true aim and received their compensation. All participants received the same amount of money. It equaled the amount of money that had been announced for performing well in the experiment, independent from a participants's actual performance.

6.5 MEASURES

Electrodermal activity

In this experiment, electrodermal activity has been measured using clinically approved sensor measurement equipment, i.e. a ProComp Infiniti encoder unit [Thought Technology, 2008]. Data on electrodermal activity was collected using signals reading on changes of skin conductance at the volar part of proximal phalanges at the non-dominant hand. Two snap style silver-silver chloride pellets at a size of 8mm are attached to the middle and the ring finger of the respective hand. Skin conductance is measured at 256 Hz. Although it is rec-

2 All participants elaborated on the cover story and hypothesized about potential application aims. Therefore, no participant has been excluded due to this.

ommended to only clean the skin with water [Boucsein et al., 2012], this has not been enough in this experiment. Due to the time of the year during which the experiment has been conducted and the fair weather respectively, participants still had greasy hands due to the use of suncream and alike. Therefore, skin has been cleaned with 70% ethanol [Schandry, 1998].

Skin conductance channel was analyzed as consisting of a phasic skin conductance response to feedback presentation and a superimposed tonic level within the 10 second interval following feedback presentation. Phasic skin conductance response has been analyzed as mean level, i.e. Skin Conductance Response (SCR), in μS, sum of Skin Conductance Response Amplitude (SCR amp), in μS, and Latency of Skin Conductance Response (SCR lat), in seconds, until first significant (>0.01 μS) increase of skin conductance. SCR amp denoted the height of phasic skin conductance response, i.e. from zero slope of the signal to next peak in case the increase exceeds 0.01 μS. Tonic skin conductance level has been analyzed as Skin Conductance Level (SCL), in μS. For analyzing whether individuals exhibit a skin conductance response with respect to eliciting events at all, the number of significant SCRs within response window has been employed, i.e. Number of Skin Conductance Responses (nSCR).

Skin conductance features of tonic and physic responses have been decomposed from measured skin conductance readings. For this, the MATLAB®-based software *Ledalab* [Benedek and Kaernbach, 2010] has been employed. A continuos decomposition analysis has been performed on the particular measurement intervals. For this, the data has been down-sampled to 12 Hz. The main idea of this method is to extract the phasic driver from the overall skin conductance readings. For this, a response model of the underlying sudomotor nerve activity is employed that allows to determine the particular phasic response characteristic. The superimposed tonic level is calculated by subtracting the resulting phasic response from actual measured data. As a result of these processes, there are several different features describing phasic responses and one representing the tonic level. As features about the phasic response are calculated based on mathematical models of physiological responses, some of them are highly correlated and accordingly redundant. This is especially important if following data analysis is based on a multivariate method. In this case, only features that are not redundant should be used.

Task Performance

In these experiments, individual performance is determined based on the mean number of guesses that have been needed to identify the hidden codes of the two Mastermind games. Participants who solved a game with the last possible guess, i.e the tenth one, have to be distinguished from participants who did not solve it at all. Accordingly, unsuccessful participants received a score of eleven steps for the particular game.

Achievement Orientation

To assess the individual's *achievement orientation* FPI-R [Fahrenberg et al., 2010] is used. The test contains 138 items for which participants are asked whether the particular statement is applicable to them or not. Based on these answers, it allows to asses 10 standard dimensions of personality, i.e. life satisfaction, social orientation, achievement orientation, inhibitedness, aggressiveness, strain, somatic complaints, health concerns, and frankness, and two additional scales, i.e. extraversion and emotionality. For the scope of this experiment only the score of achievement orientation (FPI-R 3) is considered.

Individual achievement orientation, being one of the standard scales, is assessed using 12 of the 138 items of the questionnaire. Participants with high test value are considered to be achievement oriented and active, they act quickly and are ambitious and competitive in their behavior. Their self-perception represent a person of action who efficiently tackles the very important tasks. In contrast to this, participants with low test value are assumed to be marginally achievement oriented, acting with low energy and a lack of ambition in a non-competitive manner.

6.6 DATA ANALYSIS

In any population, there is a significant percentage of individuals who do not show specific skin conductance responses towards common elicitors. Accordingly, there is a preliminary analysis for identifying these non-responding individuals. Subsequently, measured skin conductance data of the remaining individuals are analyzed with respect to the four different hypotheses.

6.6.1 *Identifying Non-Responders*

Individuals who do not show a skin conductance response during the test for skin conductance responsiveness are viewed as non-responders and their data is excluded from subsequent analysis. Individuals who show a common skin conductance response a threatening stimulus exhibit at least one significant skin conductance response within the response interval (range of ten seconds) after stimulus has been presented. Such a response involves a decline of skin conductance response after a certain amplitude has been reached. Accordingly, the *Ledalab* software [Benedek and Kaernbach, 2010] detects at least one significant SCR within response window. Accordingly, individual who do not exhibit any of such a SCR are identified as non-performers and their data is excluded.

6.6.2 *Main Analyses*

All statistical analyses were performed using IBM® SPSS®. Effects of valence manipulation have been tested with respect to different features of skin conductance responses. As different features of these responses are expected to have a highly individual response characteristic with respect to a specific experimental condition, an analysis of variance (ANOVA) was performed for each of them. SCR lat measures the time that elapses until the first significant responses occurs; there is not such feature for individuals who do not exhibit a SCR during the response window. Accordingly, SCR lat can only be evaluated for 14 individuals as the remaining four do not exhibit a SCR for all of the four[3] experimental conditions.

According to the hypotheses, the experimental group was split up in two different ways: i) with respect to the general task performance and ii) with respect to the particular level of achievement orientation. Alpha level was set to .05 for each of these analyses. In case of a significant result, the corresponding analysis was followed by calculating *Cohen's d* for determining the size of the effect.

To test effects of valence manipulation in general and between good performing individuals, skin conductance features in response to feedback presentation were analyzed with a 2 (Valence) x 2 (Performance) x 2 (Game Step) mixed factorial ANOVA. For this, the sample set was split into two performance groups by performing a median split with respect to the individual average performance, i.e. average number of steps to solve a problem that were needed in the two game conditions.

Effects of valence manipulation between individual with a different achievement orientation were tested by performing a 2 (Valence) x 2 (Achievement Orientation) x 2 (Game Step) mixed factorial ANOVA. For this sample set, a quantile split was performed for comparing the upper third and the lower third of achievement oriented individuals. This was due to the fact that there was a third of participants with

3 two game sequence in two different valence conditions

the same level of achievement orientation in the exact middle of the scale of achievement orientation.

6.7 RESULTS

6.7.1 *Non-Responders*

Two individuals have been identified as non-responding individuals (i.e. 10% non-responders). They do not exhibit any significant SCR after pressing the horn, i.e. nSCR=0. The 18 remaining individuals exhibit at least one significant SCR after the horn event, i.e. nSCR>0. As a result, there have been 18 remaining individuals, 10 are female, mean age 22,40 years (*SD*=1.51, range: 20–24years), and 8 are male, mean age 24.88 years (*SD*=7.77, range: 19–43 years).

6.7.2 *Sample Split by Performance*

In the average, participants needed 6.06 steps (*SD*=1.59, range: 3.5–9.5) to solve the games. The median of performance equals 6.0. There are 8 individuals who needed on average less then 6 steps to solve a game; the average number of steps equals 4.63 (*SD*=0.64, range: 3.5–5.5). 10 individuals needed six and more steps to solve a game, on average 7.2 steps (*SD*=1.09, range: 6–9.5) have been required.

Sample Split by Achievement Orientation

On average, participants have an assessed level of achievement orientation of 4.5 (*SD*= 1.82, range: 1–7). The median level of achievement orientation is 5. For further analysis, there are six individuals for each of the sub-sample group. For individuals with a high level of achievement orientation, an average level of of 6.33 (*SD*=0.52, range: 6–7) has been assessed. For individuals with a low achievement orientation this average level equals 2.33 (*SD*=1.03, range: 1–4).

6.7.3 *Differences in Electrodermal Responses between Positive and Negative Valence Conditions*

To test the first hypothesis, effects of valence conditions on SCR, SCR lat, and SCL have been analyzed. Means and standard errors of analyzed features are presented in Table 12. Results of the ANOVA for each of the features of skin conductance response are presented in Table 13.

A significant effect of valence with a medium size ($p < .05$, $d = .497$) was found for SCR measure, but not for SCL. Means of SCR are significantly higher in case of a negative condition than in a positive one. SCL is also not affected by valence and there is an effect of game step sequence which is highly significant but of minor size ($p < .01$, d=.058).

Electrodermal Feature			Feedback negative	positive	mean
SCR N=18		M	**.204***	**.145***	.175
		SE	.036	.025	.028
	1.Step	M	.184	.129	.156
		SE	.032	.026	.027
	2.Step	M	.225	.161	.193
		SE	.049	.030	.036
SCR lat N=14		M	2.470	2.569	2.519
		SE	.287	.190	.190
	1.Step	M	2.701	2.717	2.709
		SE	.440	.306	.326
	2.Step	M	2.239	2.421	2.330
		SE	.280	.337	.202
SCL N=18		M	2.028	1.938	1.983
		SE	.380	.367	.367
	1.Step	M	2.074	1.982	**2.029****
		SE	.373	.389	.367
	2.Step	M	1.984	1.893	**1.938****
		SE	.366	.347	.367

Table 12: Means (M) and standard errors (SE) of measured electrodermal features considering different feedback conditions. Bold face indicates that value yielded to significant result in corresponding ANOVA (see Table 13); * indicating $p< .05$, ** indicating $p< .01$.

6.7.4 *Performance-Related Differences in Electrodermal Response towards Conditions of Positive and Negative Valence*

To test hypotheses on the interaction of valence and general task performance of the individual (H 2a & H 2b), effects of valence conditions on SCR, SCR lat, and SCL have been analyzed for the two different performance groups. Means and standard errors of analyzed features are presented in Table 14. Results of the ANOVA for each of the features skin conductance responses are presented in Table 15.

A marginally significant effect of an interaction of performance and valence has been found ($p < .063$) for SCR measure. This interaction effect is small for conditions of positive valence (d = .279), but large for conditions of negative valence (d= .920). Means of SCR of poorly

Electrodermal Feature		FB Valence	Game Step	FB Valence *Game Step
SCR N=18	SS	.063	.025	.000
	df	1	1	1
	F	4.795	1.579	.059
	p	**.044**	.227	.811
	d	**.497**	—	—
SCR lat N=14	SS	.137	2.014	.096
	df	1	1	1
	F	.105	.954	.076
	p	.751	.348	.788
	d	—	—	—
SCL N=18	SS	.140	.149	2.844E-006
	df	1	1	1
	F	1.814	13.093	.000
	p	.197	**.002**	.992
	d	—	**.058**	—

Table 13: Results of 2 x 2 x 2 mixed factorial ANOVA for each of the electrodermal features considering valence of feedback and sequence of game steps. *SS*: sum of squares. In case of a significant result, the corresponding value for Cohen's measure of sample effect size (d) is presented.

performing individuals are relatively equal across the two conditions of valence. This situation is different for well-performing individuals. When they are confronted with a positive condition, the mean of SCR is comparable to the mean SCR of poorly performing individual. When confronted with a negative condition, the mean of SCR is almost twice the mean of SCR of poorly performing individuals.

An effect of general task performance on SCR lat measure has been found that is significant (p=.012) and very large (d=1.578). For well-performing individuals, mean of SCR lat is one second smaller than for poorly performing individuals. There is no effect on SCL measures.

6.7.5 *Differences in Electrodermal Response Depending on Achievement Orientation*

To test the third hypothesis, it has been analyzed whether there is a between-individual effect or an interaction effect with respect to individuals who want to achieve highly and those who do not. Effects on SCR, SCR lat, and SCL have been analyzed. Means and standard errors of analyzed features are presented in Table 16. Results of the

Electrodermal Feature			Poor Performers			Good Performers		
			neg FB	pos FB	mean	neg FB	pos FB	mean
SCR N=18		*M*	**.135**(*)	**.130**(*)	.133	**.274**(*)	**.160**(*)	.217
		SE	.048	.034	.037	.053	.038	.041
	1.St.	*M*	.096	.094	.095	.271	.163	.217
		SE	.043	.034	.036	.048	.038	.040
	2.St.	*M*	.175	.167	.171	.276	.156	.216
		SE	.065	.040	.048	.073	.045	.053
SCR lat N=14		*M*	2.869	3.292	**3.08***	2.071	1.845	**1.96***
		SE	.406	.269	.269	.406	.269	.269
	1.St.	*M*	3.073	3.518	3.296	2.328	1.916	2.123
		SE	.622	.432	.462	.622	.432	.462
	2.St.	M	2.665	3.067	2.866	1.813	1.774	1.794
		SE	.396	.476	.286	.396	.476	.286
SCL N=18		M	1.763	1.683	1.723	2.292	2.195	2.243
		SE	.507	.474	.489	.567	.530	0.547
	1.St.	M	1.783	1.718	1.751	2.364	2.251	2.307
		SE	.497	.488	.490	.556	.546	.547
	2.St.	M	1.743	1.648	1.696	2.220	2.139	2.179
		SE	.518	.462	.489	.579	.517	.547

Table 14: Means (*M*) and standard errors (*SE*) of measured electrodermal features considering different feedback conditions and performance. Bold face indicates that value yielded to significant result in corresponding ANOVA (see Table 15); * indicating $p< .05$, (*) indicating $p < .1$

ANOVA for each of the features of skin conductance responses are presented in Table 17.

There is no effect of achievement orientation on any phasic electrodermal measure, i.e. SCR and SCR lat. A large effect ($d = .974$) has been found for SCL which is almost marginal significant ($p = .123$). Individuals who have been assessed as being achievement oriented show an almost twice as high level of SCL compared to individuals who are not achievement oriented.

6.8 DISCUSSION

The present empirical study was conducted to test the impact of i) different conditions of valence, ii) general task performance of an individual, and iii) achievement orientation on affective electrodermal response. First, it was tested whether appraisals of goal conduciveness have an effect of affective responding during an extensive sequential problem solving task. Second, it was tested whether appraisals of

Electrodermal Feature		Perf.	Perf. * FB Valence	Perf. * Game Step	Perf. * FB Valence * Game Step
SCR N=18	*SS*	.125	.053	.027	4.581E-005
	df	1	1	1	1
	F	2.274	4.002	1.709	.007
	p	.151	**.063**	.210	.933
	d	—	FB_neg: **.920** FB_pos: **.279**	—	—
SCR lat N=14	*SS*	17.641	1.475	.036	.152
	df	1	1	1	1
	F	8.738	1.132	.017	.120
	p	**.012**	.308	.898	.735
	d	**1.5779**	—	—	—
SCL N=18	*SS*	4.811	.001	.024	.004
	df	1	1	1	1
	F	.503	.017	2.106	.170
	p	.488	.899	.166	.686
	d	—	—	—	—

Table 15: Results of 2 x 2 x 2 mixed factorial ANOVA for each of the electrodermal features considering performance and potential interaction with valence of feedback and sequence of game steps. *SS*: sum of squares type III. In case of a significant or marginally significanct result, the corresponding value for Cohen's measure of sample effect size (*d*) is presented.

goal conduciveness interact with general task performance in terms of affective response. Third, it was tested whether appraisals of goal conduciveness and compatibility to self-standards have an interactive effect on affective responding.

6.8.1 *Test of Valence Hypothesis*

Tests on the hypothesis on an effect of valence in terms of goal conduciveness appraisals on affective electrodermal response showed that SCR is higher for conditions of negative valence than positive valence. These findings contradict with findings on higher SCR in case of a goal conducive condition [Pecchinenda, 1996, Kreibig et al., 2010a] than a goal obstructive condition [Kreibig et al., 2010a]. This contradiction is very likely to be due to differences in measurement. For instance, in [Pecchinenda, 1996] phasic and tonic skin conductance responses have not been split up before the analysis. In addition, measurement interval has been 5 seconds longer than in this analysis. As the measurement interval significantly exceeds 10 seconds, it is

Electrodermal Feature			Low FPI_Lei			High FPI_Lei		
			neg FB	pos FB	mean	neg FB	pos FB	mean
SCR N=12		M	.184	.124	.154	.240	.144	.192
		SE	.075	.049	.057	.075	.049	.057
	1.St.	M	.122	.097	.109	.197	.129	.163
		SE	.064	.040	.047	.064	.040	.047
	2.St.	M	.247	.150	.199	.284	.159	.222
		SE	.097	.061	.075	.097	.061	.075
SCR lat N=9		M	2.114	2.861	2.487	2.366	2.334	2.350
		SE	.322	.541	.368	.288	.484	.329
	1.St.	M	2.340	2.256	2.298	2.822	2.257	2.540
		SE	.776	.432	.488	.694	.386	.436
	2.St.	M	1.887	3.465	2.676	1.911	2.411	2.161
		SE	.383	.742	.548	.342	.663	.490
SCL N=12		M	1.266	1.327	**1.30**(**)	2.748	2.681	**2.72**(**)
		SE	.619	.575	.595	.619	.575	.595
	1.St.	M	1.313	1.376	1.345	2.768	2.793	2.781
		SE	.609	.592	.598	.609	.592	.598
	2.St.	M	1.219	1.278	1.248	2.728	2.569	2.649
		SE	.630	.561	.594	.630	.561	.594

Table 16: Means (M) and standard errors (*SE*) of measured electrodermal features considering different feedback conditions and FPI_Lei. Bold face indicates that value yielded to significant result in corresponding ANOVA (see Table 17); (**) indicating $p < .15$

likely that it also includes additional skin conductance responses that result from further cognitive processing and appraisal processes accordingly. Consequently, one measure of SCR for the whole response interval interferes with such additional electrodermal responses.

The same arguments hold for the study by [Kreibig et al., 2010a]. In this case, also tonic and phasic skin conductance responses are not split up before analysis. In addition, the measurement interval constitutes 2 minutes for which SCR is averaged. Consequently, the single measure of SCR is very likely to comprise more than a single affective electrodermal response with respect to progress towards the goal conditions.

The results of this experiment are inline with findings on stronger skin conductance response in context of goal obstructive conditions in computer games [Van Reekum et al., 2004,Ravaja et al., 2006,Ravaja et al., 2008]. In all of these studies, response windows of comparable size have been selected. In addition, in two cases also phasic and tonic skin conductance responses have been split up before analysis [Ravaja et al., 2006,Ravaja et al., 2008].

Electrodermal Feature		FPI_Lei	FPI_Lei *FB Valence	FPI_Lei *Game Step	FPI_Lei *FB Valence *Game Step
SCR N=12	*SS*	.017	.004	..003	.198
	df	1	1	1	1
	F	.225	.207	.178	.317
	p	.646	.659	.682	.591
	d	—-	—	—	—
SCR lat N=9	*SS*	.167	1.350	.1.273	.152
	df	1	1	1	1
	F	.077	1.341	.594	.120
	p	.789	.285	.466	.735
	d	—	—	—	—
SCL N=12	*SS*	24.137	.049	.004	.024
	df	1	1	1	1
	F	2.836	.888	.309	1.483
	p	**.123**	.368	.590	.251
	d	**.974**	—	—	—

Table 17: Results of 2 x 2 x 2 mixed factorial ANOVA for each of the electrodermal features considering FPI_Lei and potential interaction with valence of feedback and sequence of game steps. *SS*: sum of squares type III. In case of $p<.15$, the corresponding value sample effect size (d= Cohen's d) is presented.

The first step in this Mastermind game is based on providing an arbitrary combination of elements as guess. Accordingly, the first step could also be viewed as pure gambling. For gambling tasks, affective electrodermal response does not differ with respect to goal progress [Crone et al., 2004]. Accordingly, one could expect different response characteristics between the first (considered as result of gambling) and the second guess (considered as result of problem solving). In contrast to this, response characteristics of SCR do not differ between the first and the second step of the game. Based on this, although present empirical setup forced the participants to gamble for the first step, there is no difference in affective electrodermal response between these two steps.

6.8.2 *Test of Performance Hypothesis*

Tests of the interaction effect of performance and valence of progress towards the goal yielded to a twofold result. First, the interaction hypothesis has been confirmed for SCR. For individuals who are able to efficiently solve the problem, there is a stronger SCR in context of negative condition of progress towards a goal. In contrast, individuals who are not able to solve the same problem as efficiently appear to exhibit a quite indifferent SCR response characteristic. For such individuals, valence of a conditions does not have an impact on affective electrodermal response in terms of SCR at all.

This provides empirical evidence for the fact that progress evaluation is highly dependent on individual capabilities to judge a current status towards a certain goal. Accordingly, a differentiated SCR for conditions of positive and negative feedback is an indicator for an efficient processing of feedback information. These findings provide novel insights into matching formative, i.e. task-related feedback information to general task performance, i.e. skills needed for solving the problem [Shute, 2008].

Second, it appears that the hypothesis on faster skin conductance response due to more efficient processing has been confirmed. Individuals who have a better task performance exhibit a much faster skin conductance response than individuals who are not performing that well in general. SCR lat does not appear to be modulated by valence of a condition but purely by cognitive processing capacities. This provides evidence for the assumption that different components of electrodermal responses may represent partially independent sources of information [Dawson et al., 2007]. Achievement-related APPR appear to consist of two different components: one component that reflects pure information processing and which is stable with respect to different valence conditions, i.e. SCR lat and another component that is sensitive to the individual interpretation of valence, i.e. SCR, this interpretation being also affected by information processing capabilities.

6.8.3 *Differences in Electrodermal Response Depending on Achievement Orientation*

Tests on the interaction of appraisals of goal conduciveness and compatibility with self-standards of achievement revealed no differences in corresponding phasic electrodermal response. Both SCR as well as SCR lat remain unaffected. This suggests that appraisals of compatibility of self-standards do not become evident in differentiated phasic skin responses.

Nevertheless, previous findings emphasized the important role of individual performance for the existence of differentiated phasic electrodermal responses. Accordingly, individuals with a general low task performance are unlikely to interpret feedback information accurately. These individuals are unlikely to identify a certain condition of progress as incompatible with self-standards. The opposite can be assumed for well-performing individuals. If they are confronted with a condition of progress that violates individual self-standard, they will be able to identify such an incompatibility. As a result, it remains an open question whether such an interaction effect can be identified for the group of well-performing individuals.

Test of an effect of individual achievement orientation on tonic skin conductance response revealed a higher SCL for individuals who want to achieve highly compared to SCL of individuals who do not want to do so. SCL is viewed as symptom of general arousal [Boucsein, 2012]. Accordingly, achievement oriented individuals can be viewed to have a higher level of arousal than individuals who are not achievement oriented. Individuals who want to achieve highly are more activated than individuals who are not achievement oriented [Dawson et al., 2007]. From an information processing point of view, achievement oriented individuals may allocate more attentional resources which results in a heightened autonomic activation [Jennings, 1986].

6.9 LIMITATIONS

There are certain limitations for the present empirical study. First, the sample has a limited range of age. This is due to the fact that present study has been announced on the campus of the University of Bremen so that only students have been attracted for participation. In addition, a German-language questionnaire has been used to assess achievement orientation. As a result, only Western culture participants took part in the study. Appraisal patterns have been found to be highly similar for these cultures [Scherer, 1997]. Future research is necessary to test whether current findings can be replicates with participants of other age groups as well as language and culture groups.

Second, appraisals of compatibility with self-standards have been operationalized as dimension of achievement orientation as assessed by FPI-R. Nevertheless, there are other self-standards such as perfectionism for which an incompatibility may become more evident in APPR. In addition, the number of participants in this study have been rather small. For this reason, a more extensive analysis of the third hypothesis in terms of different performance groups has not been possible. Accordingly, a repetition of this study with an increased number or participants would allow a more comprehensive analysis of a potential interaction of individual task performance and appraisals of valence and compatibility with self-standards.

Part IV

CONCLUDING REMARKS

7 CONCLUSION & OUTLOOK

In this dissertation, I have developed a novel experimental paradigm that allows to reliably measure and analyze affective states during problem solving by utilizing affective physiological responses. By applying this experimental paradigm, a thorough analysis of affective physiological responses have been possible with respect to three different theses. This final chapter represents a recapitulation of the previous chapters and offers future prospects on subsequent lines of research. In Section 7.1, the core ideas and research, the development of the novel experimental approach, as well as its employment are summarized. In Section 7.2, the results as well as major findings are presented and critically discussed. In Section 7.3, future lines of research are presented.

7.1 SUMMARY

The aim of this dissertation is to investigate affective states that arise from performing a problem solving task by analyzing the associated physiological responses of these states. To analyze affective physiological responses across individuals, two requirements have to be met. First, temporal characteristics of affective physiological responses have to be determinable which is mandatory for separating these responses from other non-affective parts of a physiological signal. Second, differences in affective physiological responses can only be compared across individuals in case these responses are collected with respect to conditions of the same valence. These issues are elaborated on in Chapter 3.

In problem solving tasks, valence is defined by individual progress towards individual goals (cf. Chapter 2). Achieving the same conditions of valence is complicated by the fact that individual progress depends on the task performance of the individual. Identical conditions of progress are accomplished when all individual make the same progress independent from their particular task performance. This has to be achieved without the individuals being aware of it to not interfere with valence of individual progress. These issues are presented in more detail in Chapter 4.

To this end, *Luckless Mastermind*, a novel experimental task paradigm is developed that meets the two previous requirements. *Luckless*

Mastermind differs from the original Mastermind in that sense as it allows to preselect game courses that are equal for all individuals. By this, all individuals encounter the same conditions of valence.

As a prerequisite for this development, it has been proofed via an algorithm performing an extensive search that there are such game courses that are valid across all individuals, independent of particular actions. The formal requirements for this as well as the algorithmic solution to this are presented in Chapter 5. Based on this, *Luckless Mastermind* was implemented and connected to a graphical user interface to use it in an experimental setup for analyzing affective physiological responses while playing this game. In this context, temporal characteristics of affective physiological responses have been considered. For instance, to prevent two subsequent states from interfering with each other, user interaction has been optimized with respect to this. The resulting task setup is presented in Chapter 6.

Based on this task setup, three theses have been investigated which are presented in more detail in Chapter 4. Employing the previous task setup, affective physiological response were collected. To allow a statistical analysis of measured data, several steps are necessary. Due to the fact that at least 5-10% of a population do not show skin conductance responses to affective stimuli, these non-responders are identified and excluded from the analysis. Before affective physiological responses can be statistically analyzed across individuals, the corresponding features of these responses have to be extracted from collected sensor data by considering their specific temporal characteristics. For this, sensor data are separated from surrounding non-affective physiological signal readings. Based on this, tonic, i.e. slowly changing and phasic, i.e. rapidly responding component of skin conductance signal are extracted. Three features are used that allow a comprehensive analysis of the specific response patterns of different affective states (cf. Section 6.5). Different features of affective skin conductance responses are assumed to show highly individual characteristics, i.e, there is no common response characteristic that is shared by all features. As a consequence, statistical data analysis has been performed using separate ANOVAs for each of the features of affective skin responses. By this, a thorough analysis for each of these particular characteristics is possible. This yielded to the following findings for the three theses.

7.2 Discussion

This dissertation answers on the following three research questions.

1. Do affective states that arise during task performance differ in case feedback of differing valence, i.e. positive or negative, is presented?

2. Does individual performance modulate affective states? Do individuals who usually perform well on a task show different affective responses compared to poor performers in case the same feedback is presented?

3. Does the individual motivation to be successful impact those affective states in general?

The first research question analyzes whether also conditions of positive or negative feedback during task performance, i.e. before a task is finished, result in particular affective physiological responses. Such physiological response patterns have been identified for affective physiological responses that result from feedback with respect to task completion but a systematic analysis during task performance has been still outstanding.

The second research question is based on the fact that individuals of different performance groups usually differ with respect to individual information processing and exploitation. Based on this, also evaluative processes that are necessary to assess feedback with respect to current task progress are likely to be moderated by general task performance. Previous experimental setups lacked the opportunity to provide equal feedback conditions across all individuals during task performance, i.e. before task completion. Based on this, another key contribution of this dissertation is to allow such a systematic investigation of an interaction of general task performance and task-related affective states.

To answer the third research question, it is analyzed whether individual factors such as the motivation to be successful also modulate affective states that occur during task performance. Just like in the previously mentioned research question, such an analysis needs identical conditions of progress in a task so that affective states can be compared across individuals of different general task performance. Accordingly, this analysis also represents a key contribution of this dissertation.

7.2.1 *Does Valence of Feedback Matter?*

Skin conductance responses consist of two different types of responses: i) a *tonic* response component which represents a slowly changing background characteristics and ii) a rapid changing *phasic* response. Phasic skin conductance responses are sensitive to stimuli or events of different valence. Accordingly, providing different conditions of feedback can be assumed to result in such a response.

To test the first thesis, phasic skin conductance measures have been analyzed in context of task-related feedback before task completion. These measures consisted of two features: i) the average phasic driver after feedback presentation, i.e. SCR and ii) the amount of time that elapsed until the first significant phasic skin conductance response occurred after feedback presentation, i.e latency.

Analysis revealed that only SCR is sensitive to task-related feedback of varying valence. More specific, individuals exhibit a stronger skin conductance response when confronted with negative feedback compared to positive one. In contrast to this, the onset of such a skin response is not modulated by valence of feedback. Accordingly, skin conductance responses do not occur faster for feedback of differing valence.

To sum it up, the intensity of phasic skin conductance responses appeared to be modulated by valence of task-related feedback. Nevertheless, subsequent analysis revealed that this effect is not a general one but is modulated by the specific characteristics of particular subsamples, i.e. good performing individuals.

7.2.2 *Does Performance Make a Difference?*

To test the second thesis, the same skin conductance measures that have been used for testing the first thesis are analyzed with respect to two subgroups of good and poor performing individuals. In this context, good performers are defined by solving a task in as few steps as possible, poor performers need more.

Analysis of skin conductance responses to feedback with differing valence between these two performance groups yielded mixed results, also with respect to the considered skin conductance features.

First, the previous finding on the interaction of valence of feedback and SCR could only be confirmed for individuals who perform well on the task. More specific, good performers exhibited a much stronger SCR when confronted with negative feedback compared to SCR resulting from positive feedback. In contrast to this, SCR appeared to be stable across feedback conditions for poor performers; SCR did not differ with respect to positive or negative feedback. Accordingly, SCR resulting from task-related feedback significantly interacts with individual performance.

Second, analysis revealed that the latency of phasic skin conductance response generally differs with respect to an individual's general task performance. For good performers, phasic skin conductance responses to task-related feedback occur significantly earlier than for individuals who perform poorly, independent of the valence of feedback. Accordingly, latency of skin conductance response appeared to be solely affected by general task performance and not by valence of feedback.

The valence of feedback that has been used in the experimental setting does not provide a direct feedback information such as correct or incorrect. It rather allows for individual interpretation to assess its valence. Based on this, for poor performers evaluating the two different types of feedback information yielded to the same outcome. Consequently, poor performers exhibit similar affective physiological responses in both the positive and the negative case. In contrast, good performers appear to exploit provided feedback information more accurately and distinguish between positive and negative feedback. Consequently, a stronger SCR response occurs for a negative condition compared to positive one which is inline with previous findings on emotional specificity of skin conductance response [Cacioppo et al., 2000].

As a result, task-related feedback that does not have an apparent value of valence but needs interpretation does not elicit different affective states for any individual. It rather depends on individual evaluative processes to add affective tone to it. Based on this, affective physiological responses appear to be an indicator of whether feedback has been received or not.

In addition, individual task performance represents a general moderating factor. Affective states arise faster for individuals who have better skills in performing a task than other performers This is independent of the corresponding feedback and its valence respectively. Assuming that good performers are able to process information more efficiently, latency of skin conductance response appears to be solely a symptom of cognitive processing speed but not of affective processing.

7.2.3 *Does Motivation Impact Affective Response?*

To test the third thesis, SCR and latency have been used as measures of affective skin conductance response. In addition, also the tonic skin conductance response has been considered as this might reflect general characteristics with respect to motivational issues. Individual motivation to be successful in the task has been assessed using the revised version of the *Freiburger Persönlichkeitsinventar (FPI-R)* [Fahrenberg et al., 2010], a German psychometric test for assessing personality dimensions.

Analysis of tonic skin conductance revealed that individuals with higher motivation also exhibit higher tonic skin responses compared to individuals with lower motivation. Higher levels of tonic skin responses are a symptom of higher arousal and alertness [Dawson et al., 2007]. Consequently, individuals who are motivated to perform better can be viewed as being more aroused during task performance than poor performers.

Analysis of phasic skin conductance responses revealed that individual motivation does not modulate such affective physiological responses. This is inline with findings on affective responses with respect to single step tasks with feedback during which affective physiological responses have been analyzed [Kreibig et al., 2012]. Accordingly, there is no specific distinguishable characteristic that differs for individuals with a high or a low level of motivation to be successful. Based on this, the hypothesis of an interaction of valence of feedback and the individual relevance have not been confirmed.

In addition, the individual level of motivation to be successful was not consistent with general task performance. As a result, there were highly motivated individuals who performed poorly and vice versa. Accordingly, it is of question whether individual motivation to be successful interacts with affective physiological responses to feedback of differing valence only for individuals of the same performance group.

7.3 DIRECTIONS FOR FUTURE RESEARCH

The work presented in this dissertation enables future research in several directions. For instance, the interaction of different physiological signals can be analyzed to a larger extend. In addition, the developed research approach can be employed for in the domain of learning in two ways. First, the present findings on performance-related differences in affective physiological response can be exploited for improving the effectiveness of e-learning applications. Second, due to the strong link between problem solving and learning, the interaction of affect and learning can be investigated in more detail by employing the developed methodology.

7.3.1 *Analyze Other Physiological Signals*

Analysis revealed that individuals of varying motivation to be successful do not show differences in affective physiological response in terms of skin conductance. Nevertheless, previous findings on affective physiological responses provide evidence that an interaction becomes apparent when other physiological signals such as cardiac activity are considered [Kreibig et al., 2012]. Based on this, for a comprehensive analysis of appraisal processes during problem solving, the developed methodological approach can be extended by incorporating other physiological signals. Additionally, current findings on the interaction of specific physiological signals in terms of affective physiological responses with respect to varying conditions of valence can be analyzed.

7.3.2 *Improving Effectiveness of E-Learning Applications*

Based on findings on performance-modulated difference in affective physiological response, skin conductance measures can be employed as indicators for efficient feedback interpretation as well as information processing in e-learning applications.

Affective physiological responses serve as additional information whether an individual has received provided feedback information or not. Further, individuals are assumed to show specific patterns of affective physiological response when information is processed that supports or hinder in solving a current task.

In the same line, affective physiological responses can be used as indicators of current performance level. The existence of affective physiological response suggest that the individual is currently a good performer who is able to process the available information appropriately. Assuming that task performance may fluctuate over the day, it remains an open question whether these affective physiological responses are also subject to individual fluctuations. For this, more extensive studies are necessary to identify whether such responses also differ inter-individually when conducted at several times.

In this context, affective physiological responses may also serve as an indicator of individual performance improvement. Individuals who usually perform badly and do not show certain affective physiological responses may be offered additional training in a task. Based on present findings, it can be assumed that efficient training and therefore improved performance result in modified affective physiological response. As a result, the extend of training measures can be adapted until these characteristics of affective physiological responses arise.

7.3.3 *Investigating Role of Affect in Learning*

Optimal learning is facilitated if the learner is actively engaged. In case of *deep learning* [D'Mello and Graesser, 2012], individuals interpret and make use of available information to both integrate it with existing knowledge and to identify potential relationships. This process involves several key cognitive processes such as acquisition of knowledge, association of information, and causal reasoning and inference. Also, metacognitive processes such as planning, re-planning, and continuous monitoring throughout the whole process are involved in this process. [D'Mello et al., 2010].

Experimental setups that support a comprehensive study of affective states with respect to cognitive and metacognitive functions and processes that are essential for deep learning were missing so far. As a result, it has not been possible to investigate whether current assumptions on affective states in terms of stimulus-related emotion also hold for affective states that accompany learning. In addition, potential interaction effects with individual characteristics such as performance or motivation while learning have not been considered. Illuminating relations between accompanying affective states and learning offers the potential to develop novel learning environments that support an optimal exploitation of individual capabilities.

When individual perform an extensive problem solving task, the same cognitive and metacognitive processes play an important role as involved during learning. As a consequence, affective states during deep learning can be investigated by employing experimental setup involving problem solving task, i.e., *Luckless Mastermind*. In this context, individuals who perform better with respect to problem solving are also assumed to be more efficient in deep learning.

Part V

APPENDIX

8 TABLES

len	col	$F \geqslant 2 \cdot len - col)$	$F \leqslant len$	Feedback patterns	#FB
2	2	$F \geqslant 2$	$F \leqslant 2$	0.2	1
2	3	$F \geqslant 1$	$F \leqslant 2$	0.1; 1.0; 0.2	3
2	4	$F \geqslant 0$	$F \leqslant 2$	0.0; 0.1; 1.0; 0.2	4
2	5	$F \geqslant -1$	$F \leqslant 2$	0.0; 0.1; 1.0; 0.2	4
3	3	$F \geqslant 3$	$F \leqslant 3$	0.3; 1.2	2
3	4	$F \geqslant 2$	$F \leqslant 3$	0.2; 1.1; 2.0; 0.3; 1.2	5
3	5	$F \geqslant 1$	$F \leqslant 3$	0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2	7
3	6	$F \geqslant 0$	$F \leqslant 3$	0.0; 0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2	8
3	7	$F \geqslant -1$	$F \leqslant 3$	0.0; 0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2	8
4	4	$F \geqslant 4$	$F \leqslant 4$	0.4; 1.3; 2.2	3
4	5	$F \geqslant 3$	$F \leqslant 4$	0.3; 1.2; 2.1; 3.0; 0.4;1,3; 2.2	7
4	6	$F \geqslant 2$	$F \leqslant 4$	0.2; 1.1; 2.0; 0.3; 1.2; 2.1; 3.0; 0.4; 1.3, 2.2	10
4	7	$F \geqslant 1$	$F \leqslant 4$	0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2; 2.1; 3.0; 0.4; 1.3, 2.2	12
4	8	$F \geqslant 0$	$F \leqslant 4$	0.0; 0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2; 2.1; 3.0; 0.4; 1.3, 2.2	13
4	9	$F \geqslant -1$	$F \leqslant 4$	0.0; 0.1; 1.0; 0.2; 1.1; 2.0; 0.3; 1.2; 2.1; 3.0; 0.4; 1.3, 2.2	13

Table 18: Upper and lower bounds for preselected feedback pattern in a Mastermind game with single element use for for small values of len and col.

REFERENCES

[Anderson, 1980] Anderson, J. R. (1980). *Cognitive Psychology and its Implications*. W. H. Freeman. (Cited on page 14.)

[Andreassi, 2007] Andreassi, J. L. (2007). *Psychophysiology: Human Behavior and Physiological Response*. Psychology Press, New York, NY, USA, 5th edition. (Cited on page 90.)

[Ashcraft, 1995] Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition*, 1:3–34. (Cited on page 26.)

[Aue and Scherer, 2011] Aue, T. and Scherer, K. R. (2011). Effects of intrinsic pleasantness and goal conduciveness appraisals on somatovisceral responding: Somewhat similar, but not identical. *Biological Psychology*, 86:65—73. (Cited on page 38.)

[Belavkin, 2001] Belavkin, R. V. (2001). The role of emotion in problem solving. In C.Johnson, editor, *Proceedings of the AISB'01 Symposium on Emotion, Cognition and Affective Computing*, pages 49–57. Heslington, York, England. (Cited on page 16.)

[Benedek and Kaernbach, 2010] Benedek, M. and Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190:80–91. (Cited on pages 93 and 95.)

[Berntson et al., 1993] Berntson, G. G., Cacioppo, J. T., and Quigley, K. S. (1993). Cardiac psychophysiology and autonomic space in humans: empirical perspectives and conceptual implications. *Pschychological Bulletin*, 114(2):296–322. (Cited on page 38.)

[Berntson et al., 2007] Berntson, G. G., Quigley, K. S., and Lozano, D. (2007). Cardiovascular psychophysiology. In [Cacioppo et al., 2007], pages 182–210. (Cited on page 38.)

[Best, 1990] Best, J. B. (1990). Knowledge acquisition and strategic action in "Mastermind" problems. *Memory & Cognition*, 18(1):54–64. (Cited on pages 60 and 61.)

[Blanchette and Richards, 2012] Blanchette, I. and Richards, A. (2012). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, 24(4):561–595. (Cited on page 20.)

[Bless and Fiedler, 1995] Bless, H. and Fiedler, K. (1995). Affective states and the influence of activated general knowledge. *Personality and Social Psychology Bulletin*, 21(7):766–778. (Cited on pages 16 and 24.)

[Boucsein, 2012] Boucsein, W. (2012). *Electrodermal Activity*. Springer, 2nd edition. (Cited on pages 89 and 104.)

[Boucsein et al., 2012] Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., and Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, 49:1017–1034. (Cited on pages 52 and 93.)

[Bower, 1981] Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36:129–148. (Cited on page 24.)

[Bower, 1991] Bower, G. H. (1991). Mood congruity of social judgments. In Forgas, J. P., editor, *Emotion and social judgments*, pages 31–53. Pergamon, Oxford, UK. (Cited on page 24.)

[Bradley and Lang, 1994] Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry*, 25:49–49. (Cited on page 34.)

[Bransford and Stein, 1993] Bransford, J. D. and Stein, B. S. (1993). *The IDEAL problem solver*. Freeman, New York, 2nd edition. (Cited on page 13.)

[Brown and Mankowski, 1993] Brown, J. D. and Mankowski, T. A. (1993). Self-esteem, mood, and self-evaluation: Changes in mood and the way you see you. *Journal of Personality and Social Psychology*, 64:421–430. (Cited on page 23.)

[Cacioppo et al., 2000] Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., and Ito, T. A. (2000). The psychophysiology of emotion. In Levis, M. and Haviland-Jones, J., editors, *Handbook of Emotions*, chapter 11, pages 173–191. Guilford Press, New York, NY, USA. (Cited on pages 3, 50, and 110.)

[Cacioppo and Tassinary, 1990] Cacioppo, J. T. and Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45(1):16–28. (Cited on page 50.)

[Cacioppo et al., 2007] Cacioppo, J. T., Tassinary, L. G., and Berntson, G. G. (2007). *Handbook of Psychophysiology*. Cambridge University Press, New York, NY, USA, 3rd edition. (Cited on pages 117 and 120.)

[Cannon, 1915] Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage: An account of recent researches into the function of emotional excitement*. D Appleton & Company, New York, NY, US. (Cited on page 35.)

[Chvátal, 1983] Chvátal, V. (1983). Mastermind. *Combinatorica*, 3(3-4):325–329. (Cited on pages 59 and 77.)

[Clark, 1983] Clark, D. M. (1983). On the induction of depressed mood in the laboratory: Evaluation and comparison of the velten and musical procedures. *Advances in Behaviour Research and Therapy*, 5:27–49. (Cited on page 2.)

[Clément and Duvallet, 2010] Clément, E. and Duvallet, D. (2010). Emotional-related responses to critical events in problem solving. *Current Psychology Letters*, 26(2).

[Clore and Huntsinger, 2007] Clore, G. L. and Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11(9):393–399. (Cited on page 24.)

[Clore et al., 2000] Clore, G. L., Wyer, R. S., Dienes, B., Gasper, K., Gohm, C., and Isbell, L. (2000). Affective feelings as feedback: Some cognitive consequences. In Clore, G. L. and Martin, L. L., editors, *Theories of Mood and Cognition : a User's Handbook*, pages 27–62. Lawrence Erlbaum Associates, Mahwah, New Jersey. (Cited on pages 16, 17, 18, and 19.)

[Crone et al., 2004] Crone, E. A., Somsen, R. J. M., van Beek, B., and van der Molen, M. W. (2004). Heart rate and skin conductance analysis of antecendents and consequences of decision making. *Psychophysiology*, 41:531–540. (Cited on page 102.)

[Csikszentmihalyi, 1990] Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row, New York, NY, USA. (Cited on page 27.)

[Dai and Sternberg, 2004] Dai, D. Y. and Sternberg, R. J. (2004). Beyond cognitivism: Toward an integrated understanding of intellectual functioning and development. In Dai, D. Y. and Sternberg, R. J., editors, *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development.*, pages 3–38. Lawrence Erlbaum Associates, Inc. Publishers., Mahwah, NJ. (Cited on pages 1 and 27.)

[Damasio, 2005] Damasio, A. (2005). *Descartes' Error: Emotion, Reason, and the Human Brain*. Penguin Books, New York, NY. (Cited on pages 6, 12, 17, 18, 19, 41, 45, and 50.)

[Davidson et al., 2003] Davidson, R. J., Scherer, K. R., and Goldsmith, H. H. (2003). Introduction. In *Handbook of Affective Sciences*. Oxford University Press. (Cited on pages 18 and 19.)

[Dawson et al., 2007] Dawson, M. E., Schell, A. M., and Filion, D. L. (2007). The electrodermal system. In [Cacioppo et al., 2007], pages 159–182. (Cited on pages 89, 103, 104, and 111.)

[Diener and Oertel, 2006] Diener, H. and Oertel, K. (2006). Experimental approach to affective interaction in games. In Pan, Z., Aylett, R., Diener, H., Jin, X., and Göbel, S., editors, *Technologies for E-learning and Digital Entertainment: Edutainment 2006*, volume 3942 of *Lecture Notes in Computer Science (LNCS)*, pages 507–518. Springer, Berlin, Heidelberg, New York. (Cited on page 16.)

[D'Mello and Graesser, 2010] D'Mello, S. K. and Graesser, A. (2010). Modeling cognitive-affective dynamics with hidden markov models. In Catrambone, R. and Ohlsson, S., editors, *Proceedings of the 32nd Annual Cognitive Science Society*, pages 2721–2726, Austin, TX. Cognitive Science Society. (Cited on pages 27 and 43.)

[D'Mello and Graesser, 2012] D'Mello, S. K. and Graesser, A. (2012). Dynamics of affective states during complex learning. *earning and Instruction*, (22):145–157. (Cited on page 112.)

[D'Mello et al., 2010] D'Mello, S. K., Lehman, B., and Person, N. (2010). Monitoring affect states during effortful problem solving activities. *International Journal of Artificial Intelligence in Education*, 20(4):361–389. (Cited on pages 16, 27, 34, 44, 46, and 113.)

[Doerr et al., 2012] Doerr, B., Spöhel, R., Thomas, H., and Winzen, C. (2012). Playing mastermind with many colors. In *Proc. of ACM-SIAM Symposium on Discrete Algorithms (SODA 2013)*. to appear. (Cited on page 59.)

[Dörner et al., 1983] Dörner, D., Kreuzig, H. W., Reither, F., and Stäudel, T., editors (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Huber, Bern. (Cited on page 46.)

[Dweck, 2002] Dweck, C. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In Aronson, J., editor, *Improving academic achievement: Impact of psychological factors on education*, pages 61–87. Academic Press, Orlando, FL. (Cited on page 27.)

[Ekman, 1993] Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4):384–392. (Cited on page 35.)

[Ekman et al., 1972] Ekman, P., Friesen, W. V., and Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* Pergamon Press, New York. (Cited on page 20.)

[Ekman et al., 1983] Ekman, P., Levenson, R. W., and Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616):1208–1210. (Cited on page 4.)

[Ericsson and Simon, 1980] Ericsson, K. A. and Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3):215–251. (Cited on page 35.)

[Fahrenberg et al., 2010] Fahrenberg, J., Hampel, R., and Selg, H. (2010). *FPI-R Freiburger Persönlichkeitsinventar*. 8. erweiterte Aufl. Hogrefe, Göttingen. (Cited on pages 92, 94, and 110.)

[Fiedler, 1988] Fiedler, K. (1988). Emotional mood, cognitive style, and behavior regulation. In Fiedler, K. and Forgas, J., editors, *Affect, cognition, and social behavior*, pages 100–119. Hogrefe International, Toronto. (Cited on page 24.)

[Fiedler, 2001] Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In Martin, L. L. and Clore, G. L., editors, *Theories of mood and cognition: A user's guidebook.*, pages 85–98. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US. (Cited on page 24.)

[Frijda, 1986] Frijda, N. H. (1986). *The Emotions*, chapter 3: Physiology of Emotion. Cambridge University Press. (Cited on pages 18, 33, and 38.)

[Frijda, 1993] Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition & Emotion*, 7(3–4):357–387. (Cited on pages 28 and 34.)

[Frijda, 2008] Frijda, N. H. (2008). The psychologists' point of view. In Lewis, M., Haviland-Jones, J. M., and Barett, L. F., editors, *Handbook of Emotions*, pages 68–101. Guilford Press, 3rd edition. (Cited on page 2.)

[Funke and Hussy, 1979] Funke, J. and Hussy, W. (1979). Informationsverarbeitende Strukturen und Prozesse: Analysemöglichkeiten durch Problemlöseparadigmen [information processing structures and processes: possibilities of analysis by the use of problem solving paradigms]. *Trierer Psychologische Berichte*, 6(8). (Cited on pages 62 and 90.)

[Furness, 2008] Furness, J. B. (2008). *The Enteric Nervous System*. John Wiley & Sons. (Cited on page 37.)

[Goddard, 2004] Goddard, W. (2004). Mastermind revisited. *Journal of Combinatorical Mathematics and Combinatorical Computing*, 51:215–220. (Cited on pages 59 and 60.)

[Graesser et al., 2005] Graesser, A., Lu, S., Olde, B., Cooper-Pye, E., and Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory and Cognition*, 33:1235–1247. (Cited on page 27.)

[Hayes, 1989] Hayes, J. R. (1989). *The Complete Problem Solver*. Erlbaum, Hillsdale, NJ, 2nd edition. (Cited on page 13.)

[Heckhausen, 1991] Heckhausen, H. (1991). *Motivation and action*. Springer, New York. (Cited on page 21.)

[Heider, 1958] Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons, New York, NY, USA. (Cited on page 27.)

[Hertel et al., 2000] Hertel, G., Neuhof, J., Theuer, T., and Kerr, N. (2000). Mood effect on cooperation in small groups: Does positive mood simply lead to more cooperation? *Cognition and Emotion*, 14:441–472. (Cited on page 24.)

[Hesse et al., 1997] Hesse, F. W., Kauer, G., and Spies, K. (1997). Effects of emotion-related surface similarity in analogical problem solving. *American Journal of Psychology*, 110:357–383. (Cited on pages 24 and 26.)

[Hirt et al., 1996] Hirt, E., McDonald, H. E., and Melton, R. J. (1996). Processing goals and the affect-performance link. In Martin, L. L. and Tesser, A., editors, *Striving and feeling: Interactions among goals, affect, and self-regulation*, pages 303–328. Erlbaum, Hillsdale, NJ. (Cited on page 24.)

[Holyak, 1995] Holyak, K. (1995). Problem solving. In Smith, E. E. and Osherson, D., editors, *An invitation to cognitive science: Thinking*, volume 3, pages 267–296. MIT Press, Cambridge, MA. (Cited on page 15.)

[Hussy, 1989] Hussy, W. (1989). Strategien zur Bewältigung umfänglicher problemrelevanter Informationsangebote im Altersvergleich. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21(1):24–39. (Cited on page 60.)

[Isen, 2008] Isen, A. M. (2008). Some ways in which positive affect influences decision making and problem solving. In Lewis, M., Haviland-Jones, J., and Barrett, L., editors, *Handbook of emotions*, pages 548–573. Guilford, New York, NY, 3 edition. (Cited on page 24.)

[Isen et al., 1978] Isen, A. M., Shalker, T. E., Clark, M. S., and Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology*, 36:1–12. (Cited on page 24.)

[Jäger and Peczarski, 2009] Jäger, G. and Peczarski, M. (2009). The number of pessimistic guesses in Generalized Mastermind. *Information Processing Letters*, 109:635–641. (Cited on pages 59 and 60.)

[James, 1884] James, W. (1884). What is an emotion? *Mind*, 9:188–205. (Cited on page 18.)

[Jennings, 1986] Jennings, J. R. (1986). Bodily changes during attending. In Coles, M. G. H., Donchin, E., and Porges, S. W., editors, *Psychophysiology: Systems, Processes, and Applications*, pages 268–289. Guilford Press, New York. (Cited on page 104.)

[Johnson and Tversky, 1983] Johnson, E. and Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, 45:20–31. (Cited on page 23.)

[Kahney, 1993] Kahney, H. (1993). *Problem Solving: Current issues*. Open University Press, Milton Keys, UK, 2nd edition. (Cited on page 15.)

[Kappas, 2011] Kappas, A. (2011). Emotion and regulation are one! *Emotion Review*, 3:17–25. (Cited on page 35.)

[Keltner et al., 1993] Keltner, D., Locke, K. D., and Audrain, P. C. (1993). The influence of attributions on the relevance of negative feelings to personal satisfaction. *Personality and Social Psychology Bulletin*, 19(1):21–29. (Cited on pages 23 and 34.)

[Klein et al., 2002] Klein, J., Moon, Y., and Picard, R. (2002). This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, 14(2):119–140. (Cited on page 27.)

[Knuth, 1977] Knuth, D. E. (1976-1977). The computer as master mind. *Journal of Recreational Mathematics*, 9(1):1–6. (Cited on pages 4, 57, and 59.)

[Koyama and Lai, 1993] Koyama, K. and Lai, T. W. (1993). An optimal mastermind strategy. *Journal of RecreationalMathematics*, 25:251–256. (Cited on page 59.)

[Kreibig et al., 2012] Kreibig, S. D., Gendolla, G. H., and Scherer, K. R. (2012). Goal relevance and goal conduciveness appraisals lead to differential autonomic reactivity in emotional responding to performance feedback. *Biological Psychology*, 91(3):365–375. (Cited on pages 45, 54, 111, and 112.)

[Kreibig et al., 2010a] Kreibig, S. D., Gendolla, G. H. E., and Scherer, K. R. (2010a). Psychological effects of emotional responding to goal attainment. *Biological Psychology*, 84(3):474–487. (Cited on pages 44, 50, 100, and 101.)

[Kreibig et al., 2010b] Kreibig, S. D., Schaefer, G., and Brosch, T. (2010b). Psychophysiological response patterning in emotion: Implications for affective computing. In Scherer, K. R., Baenziger, T., and Roesch, E., editors, *A Blueprint for affective computing: A sourcebook*, pages 105–130. Oxford University Press, Oxford, England. (Cited on pages 37 and 52.)

[Lanteaume et al., 2007] Lanteaume, L., Khalfa, S., Régis, J., Marquis, P., Chauvel, P., and Bartolomei, F. (2007). Emotion induction after direct intracerebral stimulations of human amygdala. *Cerebral Cortex*, 17(6):1307–1313. (Cited on page 38.)

[Laughlin et al., 1982] Laughlin, P. R., Lange, R., and Adamopoulos, J. (1982). Selection strategies for "Mastermind" problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(5):475–483. (Cited on pages 60 and 61.)

[Lazarus, 1991] Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press Inc. (Cited on pages 28 and 33.)

[LeDoux, 1993] LeDoux, J. (1993). Emotional networks in the brain. In Levis, M. and Haviland, J. M., editors, *Handbook of Emotions*, pages 109–118, New York, NY. Guilford Press. (Cited on page 26.)

[Leone et al., 2012] Leone, M. J., Petroni, A., Slezak, D. F., and Sigman, M. (2012). The tell-tale heart: heart rate fluctuations index objective and subjective events during a game of chess. *Frontiers in Human Neuroscience*, 6(273). (Cited on page 41.)

[Levenson, 2003] Levenson, R. W. (2003). Autonomic specificity and emotion. In Goldsmith, H. Hill, S. K. R. D. R. J., editor, *Handbook of Affective Sciences*, chapter 11, pages 212–224. Oxford University Press, Oxford, UK. (Cited on page 38.)

[Linnenbrink, 2007] Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In Schutz, P. A.

and Pekrun, R., editors, *Emotion in education*, pages 101–118. Academic Press, San Diego. (Cited on page 22.)

[Maier, 2006] Maier, R. P. A. J. E. M. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, 98:583–597. (Cited on page 22.)

[Martin et al., 1993] Martin, L. L., Achée, D. W. W. J. W., and Wyer, R. S. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64:317–326. (Cited on page 24.)

[McCombs, 2001] McCombs, B. L. (2001). Self-regulated learning and academic achievement: A phenomenological view. In Zimmerman, B. J. and Schunk, D. H., editors, *Self-regulated Learning and Academic Achievement: Theoretical Perspectives*, chapter 3, pages 67–123. Lawrence Erlbaum Associates, 2nd edition. (Cited on page 26.)

[McLeod, 1988] McLeod, D. B. (1988). Affective issues in mathematical problem solving: Some theoretical considerations. *Journal for Research in Mathematics Education*, 19(2):134–141. (Cited on page 27.)

[Mehrabian, 1980] Mehrabian, A. (1980). *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies.* Oelgeschlager, Gunn & Hain Publishers, Cambridge, MA. (Cited on page 20.)

[Melton, 1995] Melton, R. J. (1995). The role of positive affect in syllogism performance. *Personality and Social Psychology Bulletin*, 21:788–794. (Cited on page 24.)

[Mischel and Shoda, 1995] Mischel, W. and Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structures. *Psychological Review*, 102:244–268. (Cited on page 26.)

[Mitchell and Phillips, 2007] Mitchell, R. L. C. and Phillips, L. H. (2007). The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative mood on executive functions. *Neuropsychologia*, 45:617–629. (Cited on page 2.)

[Moors et al., 2013] Moors, A., Ellsworth, P. C., Scherer, K. R., and Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, 5(2):119–124. (Cited on page 29.)

[Neisser, 1976] Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. W. H. Freeman and Company, New York. (Cited on page 16.)

[Newell and Simon, 1972] Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Prentice Hall, Englewood Cliffs, NJ. (Cited on pages 2 and 15.)

[Nummenmaa and Niemi, 2004] Nummenmaa, L. and Niemi, P. (2004). Inducing affective states with success-failure manipulations: a meta-analysis. *Emotion*, 4(2):207–214. (Cited on page 21.)

[Oatley, 1999] Oatley, K. (1999). Emotions. In Wilson, R. A. and Keil, F. C., editors, *The MIT Encyclopedia of the Cognitive Science*, pages 273–275. MIT Press, Cambridge, MA. (Cited on pages 17 and 18.)

[Oatley and Jenkins, 1996] Oatley, K. and Jenkins, J. M. (1996). *Understanding Emotions*. Blackwell, Cambridge, MA. (Cited on page 12.)

[Oatley et al., 2006] Oatley, K., Keltner, D., and Jenkins, J. M. (2006). *Understanding Emotions*. Blackwell Publishers, Oxford, UK, 2nd edition. (Cited on page 26.)

[Ortony et al., 1988] Ortony, A., Clore, G. L., and Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press, Cambridge, U.K. (Cited on page 28.)

[Ortony and Turner, 1990] Ortony, A. and Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97:315–331. (Cited on page 20.)

[Pecchinenda, 1996] Pecchinenda, A. (1996). The affective significance of skin conductance activity during a difficult problem-solving task. *Cognition & Emotion*, 10(5):481–504. (Cited on page 100.)

[Pekrun, 2006] Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4):315–341. (Cited on pages 21, 22, and 26.)

[Pekrun et al., 2007] Pekrun, R., Frenzel, A. C., Goetz, T., and Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In Schutz, P. A. and Pekrun, R., editors, *Emotion in education*, pages 13–36. Academic Press, Amsterdam, Netherlands. (Cited on pages 21 and 22.)

[Pekrun et al., 2010] Pekrun, R., Goetz, T., Daniels, L., Stupnisky, R. H., and Raymond, P. (2010). Boredom in achievement settings:

Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3):531–549. (Cited on page 27.)

[Pekrun et al., 2002] Pekrun, R., Goetz, T., and Perry, W. T. R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist*, 37:91–106. (Cited on page 22.)

[Picard, 2000] Picard, R. W. (2000). *Affective computing*. MIT Press, Cambridge, MA. (Cited on pages 16 and 18.)

[Picard et al., 2004] Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., and Strohecker, C. (2004). Affective learning – a manifesto. *BT Technology Journal*, 22(4):253–269. (Cited on page 3.)

[Piferi et al., 2000] Piferi, R. L., Kline, K. A., Younger, J., and Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: viewing an aquatic video. *International Journal of Psychophysiology*, 37:207–217. (Cited on page 90.)

[Pretz et al., 2003] Pretz, J. E., Naples, A. J., and Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In Davidson, J. E. and Sternberg, R. J., editors, *Psychology of Problem Solving*, chapter 1, pages 3–30. Cambridge University Press, New York, NY, USA. (Cited on pages 15 and 16.)

[Raghunathan and Pham, 1999] Raghunathan, R. and Pham, M. T. (1999). All negative moods are not created equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Performance*, 79:56–77. (Cited on page 26.)

[Ravaja et al., 2006] Ravaja, N., Saari, T., Salminen, M., Laarni, J., and Kallinen, K. (2006). Phasic emotional reactions to video game events: a psychophysiological investigation. *Media Psychology*, 8(4):343—367. (Cited on pages 52 and 101.)

[Ravaja et al., 2008] Ravaja, N., Turpeinen, M., ans S. Puttonen, T. S., and Keltikangas-Järvinen, L. (2008). The psychophysiology of james bond: phasic emotional responses to violent video game events. *Emotion*, 8(1):114—120. (Cited on pages 52 and 101.)

[Reitman, 1964] Reitman, W. (1964). Heuristic decision procedures, open constraints, and the structure of ill-defined problems. In Shelley, W. and Bryan, G., editors, *Human judgments and optimality*, chapter 15, pages 282–315. Wiley. (Cited on page 14.)

[Richardson and Woolfolk, 1980] Richardson, F. C. and Woolfolk, R. T. (1980). Mathematics anxiety. In Sarason, I. G., editor, *Test anxiety: Theory, research, and application*, pages 271–290. Lawrence Erlbaum Associates, Hillsdale, NJ. (Cited on page 26.)

[Robertson, 2001] Robertson, S. I. (2001). *Problem Solving*. Psychology Press. (Cited on pages 14 and 15.)

[Rolls, 2000] Rolls, E. T. (2000). Précis of The Brain and Emotion. *Behavioural and Brain Sciences*, 23:177–234. (Cited on page 16.)

[Rolls, 2005] Rolls, E. T. (2005). *Emotion Explained*. Oxford University Press, New York, NY. (Cited on page 17.)

[Rottenberg et al., 2007] Rottenberg, J., Ray, R. D., and Gross, J. J. (2007). Emotion elicitation using films. In Coan, J. A. and Allen, J. J. B., editors, *Handbook of Emotion Elicitation and Assessment*, pages 9–28. Oxford Universtity Press. (Cited on page 23.)

[Russell, 1980] Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178. (Cited on page 20.)

[Schandry, 1998] Schandry, R. (1998). *Lehrbuch Psychophysiologie: körperliche Indikatoren psychischen Geschehens*. Beltz Psychologie Verlags Union. (Cited on pages 50, 51, and 93.)

[Scherer, 1984] Scherer, K. R. (1984). On the nature and function of emotion: a component process approach. In Scherer, K. and Ekman, P., editors, *Approaches to emotion*, pages 293–317. Erlbaum, Hillsdale, NJ. (Cited on pages 28, 29, and 36.)

[Scherer, 1997] Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. *Journal of Personality and Social Psychology*, 73:902—922. (Cited on page 104.)

[Scherer, 2000] Scherer, K. R. (2000). Psychological models of emotion. In Borod, J. C., editor, *The Neuropsychology of Emotion*, pages 137–162. Oxford University Press, New York. (Cited on pages 17 and 18.)

[Scherer, 2005] Scherer, K. R. (2005). What are emotions? and how can they be measured? *Social Science Information*, 44(4):693–727. (Cited on page 34.)

[Scherer, 2009] Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, 23(7):1307–1351. (Cited on pages xii, 28, 29, 30, 33, 34, and 36.)

[Scherer, 2010] Scherer, K. R. (2010). The component process model: A blueprint for a comprehensive computational model of emotion. In Scherer, K. R., Bänziger, T., and Roesch, E. B., editors, *Blueprint for affective computing: A sourcebook*, pages 47—70. Oxford University Press, Oxford, UK. (Cited on page 31.)

[Scherer, 2013a] Scherer, K. R. (2013a). The nature and dynamics of relevance and valence appraisals: Theoretical advances and recent evidence. *Emotion Review*, 5(2):150–162. (Cited on pages 30, 31, 32, and 33.)

[Scherer and Peper, 2001] Scherer, K. R. and Peper, M. (2001). Psychological theories of emotion and neuropsychological research. In Boller, F. and Grafman, J., editors, *Handbook of Neuropsychology*, volume 5 (Emotional Behavior and its Disorders, ed. G. Gainotti). Elsevier Academic Press, Amsterdam, Netherlands, 2nd edition. (Cited on pages 16, 17, and 18.)

[Scherer, 2013b] Scherer, K.R., S. V. F. J. . S. C. (2013b). The grid meets the wheel: Assessing emotional feeling via self-report. In Fontaine, J. R., Scherer, K. R., and Soriano, C., editors, *Components of Emotional Meaning: A sourcebook*, pages 281–298. Oxford University Press, Oxford. (Cited on page 34.)

[Schutz and Davis, 2000] Schutz, P. A. and Davis, H. (2000). Emotions and self-regulation during test taking. *Educational Psychologist*, 35:243–256. (Cited on page 28.)

[Schwarz, 1990] Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In Higgins, E. and Sorrentino, R., editors, *Handbook of motivation and cognition*, pages 527–561. Guilford Press, Guilford Press. (Cited on page 24.)

[Schwarz, 2000] Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, 14:433–440. (Cited on page 24.)

[Schwarz and Skurnik, 2003] Schwarz, N. and Skurnik, I. (2003). Feeling and thinking: Implications for problem solving. In Davidson, J. E. and Sternberg, R. J., editors, *The Psychology of Problem Solving*, chapter 9, pages 263–290. Cambridge University Press. (Cited on pages 23 and 24.)

[Shuman et al., 2013] Shuman, V., Sander, D., and Scherer, K. R. (2013). Levels of valence. *Frontiers in Psychology*, 46(1–17). (Cited on page 33.)

[Shute, 2008] Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1):153—189. (Cited on page 103.)

[Simon, 1999] Simon, H. A. (1999). Problem solving. In Wilson, R. A. and Keil, F. C., editors, *The MIT Encyclopedia of the cognitive science*, pages 674–676. Bradford Book. (Cited on page 14.)

[Sinclair et al., 1994] Sinclair, R., Mark, M. M., and Clore, G. L. (1994). Mood-related persuasion depends on misattributions. *Social Cognition*, 12:309–326. (Cited on page 24.)

[Sloman et al., 2005] Sloman, A., Chrisley, R., and Scheutz, M. (2005). The architectural basis of affective states and processes. In Fellous, J.-M. and Arbib, M. A., editors, *Who Needs Emotions?: The Brain Meets the Machine*. Oxford University Press, New York, NY. (Cited on page 16.)

[Smith, 1991] Smith, C. A. (1991). The self, appraisal, and coping. In Snyder, C. R. and Forsyth, D. R., editors, *Handbook of social and clinical psychology: The health perspective*, volume 162 of *Pergamon general psychology series*, pages 116–137. Pergamon Press, Elmsford, NY, US. (Cited on pages 28 and 37.)

[Spering et al., 2005] Spering, M., Wagener, D., and Funke, J. (2005). The role of emotions in complex problem-solving. *Cognition & Emotion*, 19(8):1252–1261. (Cited on pages 34 and 46.)

[Stemmler, 2003] Stemmler, G. (2003). Methodological considerations in the psychophysiological study of emotion. In Goldsmith, H. Hill, S. K. R. D. R. J., editor, *Handbook of Affective Sciences*, chapter 12, pages 225–255. Oxford University Press, Oxford, UK. (Cited on page 38.)

[Sternberg, 1986] Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills.* Harcourt Brace Jovanovich, San Diego, CA. (Cited on page 13.)

[Stuckman and Zhang, 2006] Stuckman, J. and Zhang, G.-Q. (2006). Mastermind is NP-complete. *INFOCOMP Journal of Computer Science*, 5:25–28. (Cited on page 59.)

[Thought Technology, 2008] Thought Technology (2008). Procomp infiniti [apparatus and software]. (Cited on pages 87 and 92.)

[Van Reekum et al., 2004] Van Reekum, C. M., Johnstone, T., Banse, R., Etter, A., Wehrle, T., and Scherer, K. R. (2004). Psychophysiological responses to appraisal dimensions in a computer game. *Cognition & Emotion*, 18:663—688. (Cited on pages 52 and 101.)

[VanLehn et al., 2003] VanLehn, K., Siler, S., Murray, C., Yamauchi, T., and Baggett, W. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, 21(3):209–249. (Cited on page 27.)

[Wallbott, 1998] Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896. (Cited on page 35.)

[Weiner, 1985] Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92:548–573. (Cited on pages 21 and 27.)