Universität Bremen

Faculty 03: Mathematics/Computer Science

# Natural and Playful Interaction for 3D Digital Content Creation

Dissertation in partial fulfillment of the requirements for the
Degree of Doctor of Engineering (Dr.-Ing.)

By Thomas Fröhlich, M.Sc.
Date of doctoral viva: 5th March 2020

Reviewers:
Prof. Dr. Rainer Malaka
Prof. Dr. Frank Steinicke

## Abstract

The creation of 3D content for visualizations, games, or virtual environments is often done with tools like Blender or Maya. While these applications are very powerful, they are also hard to learn and require domain-specific knowledge in order to create appealing results. This can be challenging for beginners and non-technical users who want to create 3D content in a fast and easy way. With the rise of 3D environments and technologies such as Virtual and Augmented Reality, easy-to-learn development tools such as Swift Playgrounds for programming and Unity for the development of games and interactive applications have reduced the barrier for non-technical users. In the field of 3D content creation applications often lack accessibility for non-technical users and beginners, so their ability of expressing themselves in the form of 3D content is limited.

However, interacting with 3D content can be easy and fast-to-learn when the appropriate input technology and interaction design is used. Playful approaches to 3D interaction promise high motivation and enjoyable experiences that reduce the seriousness of the 3D creation task, providing room for experimentation and creative freedom. Further, natural user interfaces incorporate novel input technology such as motion tracking, gesture recognition, direct manipulation, or reality-based interaction. Research has shown that playful and natural interaction can benefit 3D tasks and can result in pleasurable experiences for beginners and non-technical users.

The goal and leading research question for this thesis is to explore how playful and natural interaction concepts can be used to facilitate 3D content creation for beginners and non-technical users. This thesis aims to further democratize the field of 3D content creation, making it available to a larger audience. It presents three systems that center around the concept of playful interaction and three systems incorporating the natural interaction paradigm for digital content creation. These systems were evaluated in user studies in order to answer the research question.

The results show that both playful and natural interaction can benefit content creation tasks, improving fun, captivation and attractiveness towards 3D creation tasks whereas complexity and learning effort is decreased, indicating the usefulness of playful and natural interaction for digital content creation as a starting point for 3D creation tasks.

## Zusammenfassung

Die Erstellung von 3D-Inhalten für Visualisierungen, Computerspiele oder virtuelle Umgebungen wird häufig mit Programmen wie Maya oder Blender durchgeführt. Während diese Programme sehr leistungsfähig und umfangreich sind, haben sie den Nachteil, dass Anfänger*innen und nicht-technische Nutzer*innen aufgrund der Komplexität und aufwändig zu erlernenden Bedienung nur schwer Zugang zu diesen Anwendungen finden. Mit dem Aufkommen von einfach zu nutzenden 3D-Technologien wie Virtual- und Augmented Reality wurde auch die Entwicklung entsprechender Inhalte und Umgebungen für Anfänger*innen und Einsteiger*innen vereinfacht, wie z.B. mit Swift Playgrounds für die Programmierung oder auch mit Unity für die Erstellung von virtuellen Umgebungen und Spielen. Während diese und andere Programme den Einstieg erleichtern, fehlt es im Bereich der 3D-Inhaltserstellung noch an entsprechenden Anwendungen, die die Erstellung für Anfänger*innen und nicht-technische Nutzer*innen vereinfachen und ihnen die Möglichkeit geben, sich kreativ auszudrücken.

Die Interaktion mit 3D-Inhalten kann jedoch einfach und schnell erlernt werden, wenn passende Interaktionsdesigns genutzt werden. Spielerische und natürliche Interaktion sind zwei Möglichkeiten, die eine hohe Einstiegsmotivation bieten und zusätzlich eine unterhaltsame Interaktionsform darstellen, die die Ernsthaftigkeit der 3D-Inhaltserstellung senken und somit Freiraum für das kreative Erstellen von 3D-Inhalten schaffen. Zudem bietet die natürliche Interaktion neuartige Eingabetechnologien und Interaktionsformen wie z.B. Motion Tracking, Gesteninteraktion, Direct Manipulation oder auch Reality-based Interaction, womit bereits gezeigt werden konnte, dass spielerische und natürliche Interaktion Vorteile für verschiedene 3D-Anwendungen haben und in einem positiven Nutzererlebnis münden können.

Das Ziel und die leitende Forschungsfrage für diese Arbeit ist herauszufinden, inwiefern natürliche und spielerische Interaktion gewinnbringend für die Erstellung von 3D-Inhalten im Hinblick auf Anfänger*innen und nicht-technische Nutzer*innen eingesetzt werden können. Darüber hinaus zielt diese Arbeit darauf ab, die 3D-Inhaltserstellung für Anfänger*innen und nicht-technische Nutzer*innen zugänglich zu machen. Es werden drei prototypische Systeme auf Grundlage spielerischer Interaktionsformen und drei Systeme auf Grundlage natürlicher Interaktion präsentiert. Diese Systeme wurden in Nutzerstudien auf ihren positiven Einfluss auf das Nutzererlebnis hin überprüft. Die Ergebnisse zeigen, dass beide Interaktionsformen die 3D-Inhaltserstellung vereinfachen und ein positives, vergnügliches, fesselndes und attraktives Nutzererlebnis herbeiführen können. Gleichzeitig verringern sie die Komplexität und den Lernaufwand für die untersuchten Aufgaben. Spielerische und natürliche Interaktion stellen somit einen einfachen Startpunkt für die Erstellung von 3D Inhalten dar und können für Anfänger*innen und nicht-technische Nutzer*innen ein leichter Einstieg in die Erstellung von 3D-Inhalten sein.

# Danksagung

Mein größter Dank gilt meiner Frau Johanna. Du hast mir zugehört, mich beraten, mir den Rücken frei gehalten und hast dieses Ziel so viele Jahre mitgetragen. Ohne dich hätte ich das nicht geschafft.

Ich danke meinem Doktorvater Rainer Malaka für die Freiheit in der Forschung, das allgegenwärtige Vertrauen und für das Verständnis und die Unterstützung privater und beruflicher Entscheidungen. Ich danke meinen lieben Kolleg*innen der AG Digitale Medien der Universität Bremen, insbesondere Dimi, Thomas, Susanne, Anke, Dirk, Svenja und Irmgard für die großartige Unterstützung und die tolle Zusammenarbeit.

Ein großer Dank gilt zudem den Studierenden Timo Stabbert, Jan von Oehsen, Jan Tilger, Matthias Klug, Felix Unger und Christopher Pollok, mit denen ich zusammen forschen und deren Arbeiten ich betreuen konnte.

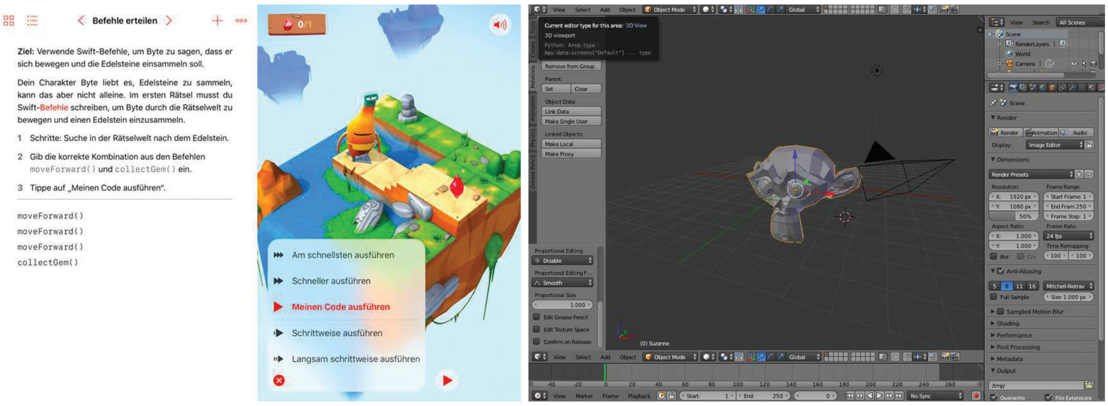
# Contents

# 1 Introduction

In 1996, Bill Gates published an essay with the title “Content is King”, referring to the growth of internet businesses and the ability of companies and individuals to create and share digital content with everyone on the internet (Gates, 2016). The phrase highlights the importance of digital content in general and in retrospective, Mr. Gates was right, considering the success of companies like Netflix and Spotify. Digital content not only describes film or music, it also refers to 3D models, animations, textures, scenes, and video games. Regarding the latter, it is nowadays usual to receive additional content in the form of extensions, where game mechanics mostly stay the same while new stories, levels, and characters are added, extending the life span of products. Quite often, publishers hand over the tools for digital content creation to the target audience in the form of (level) editors, so that the users can extend their experience themselves, sharing and copying content from online repositories, for example in games such as GTA V (Rockstar, 2013), Cities:Skylines (Colossal Order, 2015), or The Elder Scrolls V: Skyrim (Bethesda Game Studios, 2011).

While creating content for existing titles is one aspect, another important one is game development itself. When game development 20 years ago was only feasible for professionals, teams, and studios, game development has become more accessible to non-professional users. One reason for this development is the introduction of game engines such as Unity (Unity Technologies, 2005) and Unreal (Epic Games, 1998) that are free to use for individuals, providing tools that were until then only available to companies due to the high costs and the required expert knowledge. Having these powerful tools, non-expert developers are able to create games and digital content with the help of a large online community supporting with tutorials, help pages, documentation, and forum support. It can be said that game development has been democratized as it is now available to anyone with a computer and internet access, whereas in the past, these resources were only accessible for professionals.

However, when creating games or other applications, content has to be created in order for the product to have value to users. Games need characters, animations, models, and levels in order to be successful. As much as Unity and Unreal democratized game development, the creation of 3D digital content still lacks this level of accessibility. For example, creating 3D models or animations still is hard to achieve for beginners (Garwood



Befehle erteilen

**Ziel:** Verwende Swift-Befehle, um Byte zu sagen, dass er sich bewegen und die Edelsteine einsammeln soll.

Dein Charakter Byte liebt es, Edelsteine zu sammeln, kann das aber nicht alleine. Im ersten Rätsel musst du Swift-Befehle schreiben, um Byte durch die Rätselwelt zu bewegen und einen Edelstein einzusammeln.

1 Schritte: Suche in der Rätselwelt nach dem Edelstein.

2 Gib die korrekte Kombination aus den Befehlen `moveForward()` und `collectGem()` ein.

3 Tippe auf „Meinen Code ausführen“.

```
moveForward()
moveForward()
moveForward()
collectGem()
```

Figure 1.1: Left: Playful user interface of Swift Playgrounds for learning programming, Image source Apple Inc. (2020). Right: Blender user interface for 3D modeling and animation.

and Dunlop, 2014) because these users are faced with very complex tools, such as Blender (Roosendaal, 1998), or Maya (Autodesk, 1998). Blender is freely available and there is also a large online community helping beginners with creating digital content, but still, creating appealing content using these tools is a very time consuming and sometimes frustrating activity, because these applications are very complex and aimed at users having deeper knowledge of modeling and animating. When starting as a new user, wanting to create 3D content, these tools introduce a high barrier due to their complexity, task-orientation, and overall high learning effort. There are simpler tools available such as SketchUp (Trimble Inc., 2013) that make 3D modeling easier, yet these tools are still hard to learn (Cano, 2011). Further, non-technical users often fail because of too high standards and expectations regarding these tasks.

Also in game development, non-technical users have similar problems with Unity and Unreal, however, there are already tools and methods available that make the process of getting into the field easier. For example Swift Playgrounds (Apple, 2014) is an app from Apple for learning programming in a playful and easy way, see Figure 1.1. Playgrounds requires no knowledge of programming and is especially aimed at beginners. Users can learn programming in an easy and playful way without being confronted with the complex details from the beginning. Ideally, users gain self-confidence and basic knowledge of the problem area and deepen their knowledge with more classic approaches like classes, books, and self-teaching using standard software such as Xcode (Apple, 2003). Building upon their prior positive experiences, users can be more resilient regarding errors and hurdles that appear along the way.

Another problem is that novice users often don't know where to start or set goals that are too high to achieve, resulting in losing motivation and quitting after some hours or days. This "blank page syndrome", which is a form of the writer's block (Rose, 1984) makes it harder getting into the field because of lacking or too big ideas. Even if beginners make

Figure 1.2: Social hopscotch experiment showing that play is not only for children. Source: Cut (2018).

it past the first models and stay with the program, reward comes very late in the process because users are often underwhelmed by their first models and results, introducing a possible exit point. Further, when starting, beginners don't have a language (taxonomy) and basic understanding of paradigms, leading to frustration in public forums, because they don't know what question to ask and are not able to properly express themselves, again resulting in a feeling of being a newbie to the field. To sum up, novice users might be confronted with their lack of expertise in the field of 3D content in many ways, which can lead to early quitting and giving up on 3D content creation.

This thesis aims at overcoming these issues for 3D content creation, by reducing the seriousness and entry hurdles of 3D tools by introducing disruptive ways of approaching 3D interaction, with the aim of further democratizing the field of 3D content creation. It is assumed that if novice users start easily into the field, create fast and appealing results, the chances and motivation of them sticking in the field will increase.

When thinking about reducing seriousness, one can look at a fitting antonym and finds easy, friendly, funny, happy, and light, just to name a few. An activity that qualifies all of these attributes is play. Play offers a room for experimentation, exploration, and even failure. It creates the opportunity for building something new (e.g. playing with clay, in the sandbox, scribbling) and experience creative enjoyment without judgment. Play introduces the exploration of different solutions to problems, and at the same time supports the development of skills through exploratory behavior. It supports motivation and collaboration and represents a joy, fond, delight, amusing, and non-serious activity. (Garvey, 1990)

Not only is play characterized by the above presented characteristics, it also has to be noted that play is deeply grounded in the behavior of humans and can also be found in other species (Bateson and Martin, 2013). Mostly, play is researched and associated with children's play (Schwartzman, 1978). However, also adults like to engage in play and playful behavior as has been demonstrated in a social experiment where a hopscotch was drawn on a sidewalk and cameras observed the spot for one day, see Figure 1.2, (Cut, 2018). The producers counted that 129 people played the game within 10 hours while 1058 walked by without playing (ca. 12% play rate). The video demonstrates that play is not restricted by age, social status, or gender. There neither is an age limit to play nor are there limited ways to engage in play. Every human is proficient in play, because at some point in our lives, everyone enjoyed play as the above described activity (Lents, 2017). As well as play qualifies these aspects, play is simply natural for almost all lifeforms. There is no preconception of "correct" play. Play is what the player experiences as play.

As much as play is natural for us, there are also other things that are natural, making activities easy and fast to accomplish. For example, we have a basic understanding of physics, know that when we reach for an object, how to move our muscles without thinking of it. Natural behavior can also be learned, for example as in riding a bike. Starting, we have to think about all movements, accelerating, braking, keeping balance. After some practice, biking becomes natural to us, we have internalized the movements and are able to anticipate on basis of our experience what can happen in a given situation in order to act appropriately. When translating this natural understanding of interrelations to human-computer interaction (HCI) we arrive in the field on natural interaction, or natural user interfaces (NUI). NUIs build upon the notion that users should be able to "naturally" interact with computers, meaning that users can facilitate prior world-knowledge for interaction that is easy to learn and execute and available to a large audience (Wigdor and Wixon, 2011). Popular implementations of NUIs include gesture interfaces (Freeman et al., 2016; Chan et al., 2016; Walther-Franks, Herrlich, and Malaka, 2011), such as pinch and pan gestures, that some years ago found their way into consumer hardware such as iPads and mobile phones (Rendl et al., 2014). However, natural interaction still is a vivid field of research and it is still not fully understood. As of today, we can only attribute certain concepts to natural interaction, describe the interaction, and even provide guidelines implementing NUIs. However, there still is no single true definition of the concept as research and development are still ongoing.

## 1.1 Thesis Statement

This thesis investigates how to use natural and playful user interfaces for supporting 3D digital content creation for non-technical users. We focus on providing environments that are playful, creating experiences that keep users motivated in learning the deeper concepts of 3D content creation. At the same time, we aim at supporting users with building 3D content without prior knowledge, extending the democratization of the 3D content creation area. We further aim at assessing the potential of play and natural interaction for creative purposes in order to finding out how to empower non-technical users in engaging in 3D digital content creation using playful and natural interaction techniques. Our goal is not to replace professional tools, as these are very useful and needed for advanced projects and professional users.

We rather claim a need for playful and natural interfaces, because in 3D digital content creation, there is a lack of approaches and solutions that are idiosyncratic and disruptive. Playful interfaces allow for personalized experiences, creating intrinsic motivation to engage in 3D tasks by providing a safe environment (Bateson, 2014) in contrast to task oriented solutions. Task-oriented interfaces often create hurdles for beginners and non-experts. Instead the focus should be on a positive interaction and user experience. Thus, interaction design as a scientific discipline has to consider playful interaction and positive factors more extensively in order to support users with generating personalized experiences through playful interfaces because play is an important element of culture and society (Huizinga, 1949). In the interaction design literature, similar supportive claims can be found:

> *Scientific approaches to design need to be complemented by more subjective, idiosyncratic [quirky] ones. It is difficult to conceive of a task analysis for goofing around, or to think of exploration as a problem to be solved, or to determine usability requirements for systems meant to spark new perceptions. Instead, designers need to use their personal experiences as sounding boards for the systems they create.* (Gaver, 2002)
>
> *We need more objects that allow us to be playful. [...] These designs need to exist so we can make technologies ours, and our being in the world a personal affair.* (Sicart, 2014)

In line with these calls, we further motivate our objective with positive design and positive psychology, the latter concerning with eudaimonia, “the good life”, that represents a natural succession to the already existing chronology of human factors (ergonomics), usability, and user experience (Malaka, 2017). In line with the mission statement of the Delft Institute of Positive Design, in this thesis we emphasize a) the creation of possibilities as an optimistic future, supporting of human flourishing that uplifts people, enabling of meaningful activities that stimulate people, embracing rich experiences that last, and accepting responsibility: the impact of technology on people’s lives and society[1]. Based on this, we formulate our main hypothesis:

> $H_1$: The creation of 3D digital content benefits from playful and natural interaction concepts.

As we are designing and evaluating systems that center around playful and natural interaction for 3D digital content creation, we further question how inexperienced and non-technical users perceive these paradigms as beneficial in the context of use, comparing them to existing solutions. We investigate the use of sand as a natural material and playful activity for creating 3D models and virtual worlds and evaluate the system with experts from game design, computer graphics, and education. We further research the differences between natural and playful user interfaces and common approaches to 3D interaction for content creation. Thus, we postulate the following hypotheses:

> $H_2$: Interaction with gestures and natural material support fast understanding and quick learning of 3D content creation.

---

[1] https://diopd.org/about-us/mission/

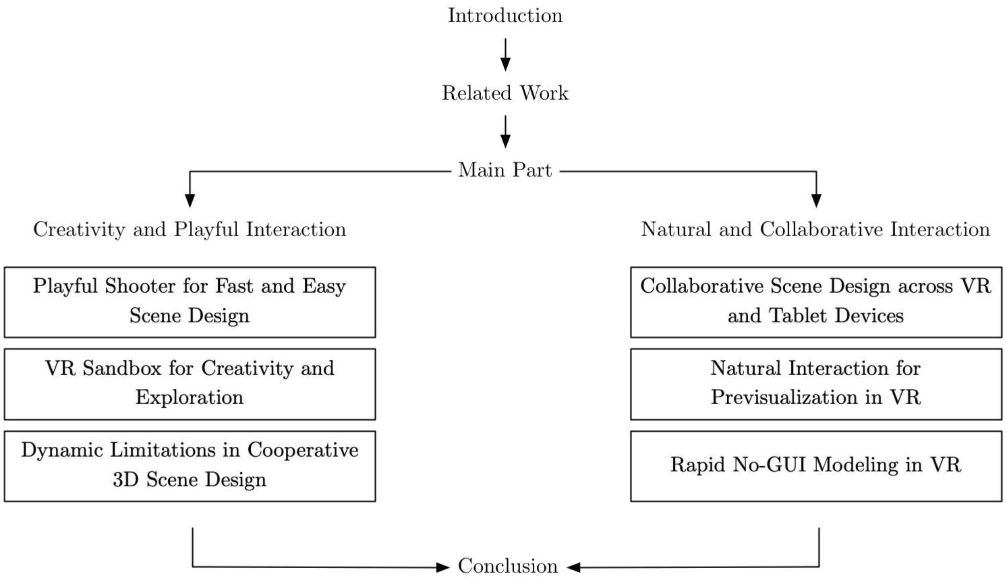


Figure 1.3: Structure of this thesis.

> $H_3$: Natural and playful interfaces are preferred by users compared to WIMP and menu interaction.

Finally, as creativity and collaboration are aspects of playful behavior, we also consider these influences in our research agenda. In lab studies, we explore how playful approaches benefit perceived user creativity, focusing on creating a positive working environment. Lastly, as collaboration can allow for increased creativity due to lateral thinking among two people, we research the effect of collaborative interfaces in two lab studies and aim to test the following hypotheses:

> $H_4$: Playful approaches to interaction design positively influence user creativity.
>
> $H_5$: Collaboration across devices and players supports a natural working style, supporting 3D content creation.

## 1.2 Outline of this work

The overall structure of this thesis is presented in Figure 1.3. In the next chapter, related work and state-of-the-art systems are presented that build the theoretical foundation of this work. Part one concerns with foundations regarding play and playful systems, while part two presents work on natural user interfaces and related concepts such as reality-based interaction.

Following, the main part of this thesis is grouped in two chapters that each include three different contributions to either playful applications (Chapter 3) and natural interaction

(Chapter 4). In Chapter 3, Creativity and Playful Interaction, we present applications that highlight playful mechanics to 3D content creation: a shooter for fast and easy scene design (Section 3.1), a VR sandbox for natural and playful modeling using real sand (Section 3.2), and a playful approach to creativity support on the basis of the game Minecraft in Section 3.3. The second chapter of the main part (Chapter 4 - Natural and Collaborative Interaction) includes a system for collaborative scene design facilitating means of natural interaction by combining VR and tablet interaction (Section 4.1), a system for natural interaction with 3D content in the area of previsualization, including a VR interface for embodied character animation (Section 4.2), and a prototype for rapid modeling in VR in Section 4.3. Finally, we conclude and discuss the findings in Chapter 5, where we summarize the main points, reflect on the work carried out, and conclude this thesis with final remarks. A short summary of the research chapters is given below:

### 1.2.1 Chapter 3: Creativity and Playful Interaction

#### *Playful Shooter for Fast and Easy Scene Design*

World and map builders for video games can be cumbersome to use because of complex user interfaces with a high learning curve and limited user experience. Often, the editor perspective doesn't match the play view (e.g. first person games), disconnecting creation and experience through constantly switching perspectives. It is further difficult to position objects in 3D space using the mouse and constantly switching between different Gizmos, often operating only in one dimension at a time. Editors also tend to be oriented on the task, lacking support for creative expression and serendipity, which are both important dimensions of design. We introduce a world builder game where the process of creating a 3D scene is play itself. Users of our tool are able to navigate directly in the 3D scene, placing objects in a playful manner by shooting them in the scene for placement and manipulation.

#### *Playful VR Sandbox for Creativity and Exploration*

Playing with sand is something most people relate to from childhood days, where this unstructured play provided room for experimentation, exploration, or cooperation. As children we develop important skills in the sandbox like proprioceptive sensing, body and space awareness, or social skills and can engage in creative behavior and open-ended play. These fond memories last a lifetime and, when given the opportunity, many adults enjoy playing with sand and rediscovering the sensations of this activity. VRBox is an interactive sandbox for playful and immersive terraforming that combines the approach of augmented sandboxes with modern Virtual Reality and mid-air gestures. The system extends existing approaches with an additional interaction and visualization layer to interact with sand and virtual objects. Furthermore, users can switch perspectives from a tabletop mode into a first person mode by teleporting directly into their own creation, allowing for an immersive experience of the landscape in full-size.

#### *Dynamic Limitations in Cooperative 3D Scene Design*

Divergent thinking is a creativity technique to create many solutions to a given problem, assuming that some of these ideas may be correct or can further lead to a correct solution. While the idea of divergent thinking creates freedom and many possibilities to choose from, there are also techniques that limit artists in a specific way in order to be more creative and generate alternative ideas. One of these is known as “creative limitation”. In contrast to divergent thinking methods, limitations work under the assumption that when the mind is forced to obey to certain rules, that at first glance limit the possibilities of action, new creative ways of overcoming the limitations arise and potentially generate creative problem solving. We expand the idea of supporting users in creating novel and creative 3D content by incorporating dynamic limitations in a 3D modeling task. For this, we modified Minecraft and conducted a user study where participants collaboratively constructed a 3D block-world. As we are interested in finding effects on self-perceived creativity and the creative self-assessment of the result, we incorporated two overall mechanisms in the game: limitations and dynamic activation.

### 1.2.2 Chapter 4: Natural and Collaborative Interaction

#### *Collaborative Scene Design across Virtual Reality and Tablet Devices*

The creation of 3D content such as virtual worlds and models is most natural in Virtual Reality because of a large variety of reasons. However, there are some aspects of VR that limit the usefulness of the approach. Specifically in collaborative design, where several users work together in achieving a common design task, VR may limit the collaboration. As an example, in VR, users are isolated from their surroundings and mostly, VR users do not share a common working place with other users. For this research, we combine the advantages of VR and tablet interaction to build a system where users can collaborate using both devices in a shared environment. We assume that depending on the devices, different roles and tasks are distributed between the users. For example, as tablets are ideal for overview, we assume that in a shared task, this is something that naturally falls in the category of a tablet use. We collect a large variety of usage-data both for the VR and tablet user. Hence, we employ an inductive research approach, observing natural behavior, generating a theory of usage based on our observations.

#### *Previsualization in Virtual Reality*

Previsualization (previs) is an essential phase in the design process of narrative media such as film, animation, and stage plays. Digital previs can involve complex technical tasks, e.g. 3D scene creation, animation and camera work, which require trained skills that are not available to all personnel involved in creative decisions for the production. In previs, creating convincing material often incorporates different tasks on various media such as 2D and 3D layout and animation. All of these media require expressive tools. The ideal previs user interface should support creative freedom and flexibility, so that users experience the previs tool as a natural extension of their own physical capabilities and can express their ideas in a seamless and intuitive way. We present natural user interfaces that speak the language of the artist and not the one of a technician by incorporating VR interaction and embodied interaction for motion capture.

***Rapid No-GUI Modeling in VR***

Beginners can feel overwhelmed by user interfaces and manifold options in 3D content creation tools. This in turn can lead to an exclusion of interested users that would like to approach the domain of 3D models, because of the time and effort those users would have to invest in becoming familiar and productive in these tools. We present a “no-GUI” system that allows for rapid prototyping of basic 3D models, focusing on natural ways of creating these models by only using gestures for system interaction and creation. We focus on rapid prototyping because we aim at supporting users that often only use these models as placeholders to be replaced in the further development of their project.

## 1.3 Publications

This thesis is based on the following publications:

- Thomas Fröhlich (2017). “Play While You Work: Productive Play for Digital Content Creation”. In: *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY ’17 Extended Abstracts.* Amsterdam, The Netherlands: ACM Press, pp. 699–702. ISBN: 978-1-4503-5111-9. DOI: `10.1145/3130859.3133228`
- Thomas Fröhlich, Jan von Oehsen, and Rainer Malaka (2017). “Productivity & Play: A First-Person Shooter for Fast and Easy Scene Design”. In: *CHI PLAY ’17 Extended Abstracts.* Amsterdam, The Netherlands: ACM Press, pp. 353–359. ISBN: 978-1-4503-5111-9. DOI: `10.1145/3130859.3131319`
- Thomas Fröhlich, Dmitry Alexandrovsky, Timo Stabbert, Tanja Döring, and Rainer Malaka (2018). “VRBox: A Virtual Reality Augmented Sandbox for Immersive Playfulness, Creativity and Exploration”. In: *The Annual Symposium on Computer-Human Interaction in Play - CHI PLAY ’18.* Melbourne, VIC, Australia: ACM Press, pp. 153–162. DOI: `10.1145/3242671.3242697`
- Timo Stabbert, Thomas Fröhlich, Dmitry Alexandrovsky, and Rainer Malaka (2017). “Extending Augmented Sandboxes with Virtual Reality Interaction”. In: *Mensch und Computer 2017 - Workshopband.* Regensburg: Gesellschaft für Informatik e.V., p. 6
- Thomas Münder, Thomas Fröhlich, and Rainer Malaka (2018). “Empowering Creative People: Virtual Reality for Previsualization”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18.* Montreal QC, Canada: ACM Press, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: `10.1145/3170427.3188612`

# 2 Related Work

This chapter presents related work and the state of the art that this thesis is based on. The main concepts center around play, playfulness and creativity, describing work in relation to the playful aspects of this work. Further, natural user interfaces will be introduced and related work in the field of human-computer-interaction (HCI) will be presented. In each section, related applications, workflows and systems that have been designed and evaluated in the field of HCI will be included, establishing a context for analysis from an HCI perspective.

## 2.1 Play, Playfulness and Creativity

### 2.1.1 Defining Play

Play has been defined from different view points in biology, psychology, and cultural contexts. The definition of the term is highly dependent on the view and context of analysis. In the following, an overview of how play is defined in the existing literature is presented. An often referenced definition is provided by Huizinga (1949) who takes a cultural perspective on the matter. Huizinga states that play has different characteristics and functions: Play is a free, voluntary, superfluous, and not ordinary activity. It is further secluded and limitless, and also a social construction. In the literature, there often is the discussion about the relatedness or non-relatedness of play and seriousness. Huizinga offers an explanation to this by saying that "*[...] the contrast between play and seriousness proves to be neither conclusive nor fixed. We can say: play is non-seriousness. But apart from the fact that this proposition tells us nothing about the positive qualities of play, it is extraordinarily easy to refute. As soon as we proceed from "play is non-seriousness" to "play is not serious", the contrast leaves us in the lurch-for some play can be very serious indeed*".

As play can be serious, he also points out that play must not be foolish by nature: *It lies outside the antithesis of wisdom and folly. The later Middle Ages tended to express the two cardinal moods of life - play and seriousness - somewhat imperfectly by opposing folie to sense, until Erasmus in his Laus Stultitiae showed the inadequacy of the contrast* (Huizinga, 1949).

Huizinga further offers an overall definition of the different characteristics of play: *[...] we might call it a free activity standing quite consciously outside "ordinary" life as being "not serious", but at the same time absorbing the player intensely and utterly. It is an activity connected with no material interest, and no profit can be gained by it. It proceeds within its own proper boundaries of time and space according to fixed rules and in an orderly manner. It promotes the formation of social groupings which tend to surround themselves with secrecy and to stress their difference from the common world by disguise or other means.*

Sicart (2014) offers a definition that in parts overlap with Huizinga's notion. He states that play is *contextual*, including the environment we play in, technologies we use and potential play companions. Here, rules can be introduced in play, further establishing and defining the play context. Moreover, play is *carnivalesque*, meaning happening between creation and destruction embedded in laughter, also *appropriative*, taking over the context in which it emerges and exists, and *disruptive* as a consequence of the appropriative character. It disrupts existing situations, taking over and "breaking the state of affairs". Further, play is *autotelic*, meaning it is being done by it's own purpose and not a fixed one and *creative*, engaging in a creative way with the world, objects, rules, contexts "through ludic interaction". Lastly, play is *personal* and the effects are individual: memories, friendships, experiences, expression or language are impacted by playing. (Sicart, 2014)

Complementing these foundational characteristics of play, Bateson and Martin (2013) summarize play by presenting the following five defining features. First, play is spontaneous, intrinsically motivated, rewarding, and autotelic. Second, play is protected from usual consequences of serious behavior, making it the "antithesis" [*sic*] of work (a dichotomy that in others have softened, e.g. Huizinga (1949)). Third, play results in novel combinations of thought, action or objects, making it a "generator of novelty". Fourth, play looks different each time and with each individual performing it, and fifth, play needs to be free of consequences, illness, repression and other hindering factors. It can only occur if fundamental preconditions have been established, see Maslov's hierarchy of needs (Maslow, 1943).

There are many more directions to the definition of the term play, very dependent on the discipline from which play is being observed and defined, for example Catherine Garvey from a children's developmental perspective (Garvey, 1990) and Bateson and Martin from a biological one (Bateson and Martin, 2013; Bateson and Nettle, 2014). A good overview of the ambiguity of the term play is presented by Sutton-Smith (2009), looking from various perspectives on the term such as evolution, development, biology, sociology, trying to sum those up with seven rhetorics, play can be looked at: play as progress, play as fate, play as power, play as identity, play as the imaginary, rhetoric of the self, and play as frivolous.

### 2.1.2 Playfulness and Playful Interaction

Having presented different views on play, in this section, playfulness will be distinguished from play. The term playfulness is often used in the context of behavior, learning, and interaction with other people or technology. Especially, playful interaction is a paradigm in the design of systems where it is assumed that playful interaction benefits the user experience. Play can be seen as an activity with specific attributes (voluntary, free, intrinsically motivated) while playfulness can be defined as a characteristic of a person (someone is playful, Barnett (1990)), state of mind (someone is in a playful mood, Youell (2008)), or a perception towards an artifact or object (this is a playful system, Kuts (2009)).

A philosophical description of playfulness is provided by Sicart (2014). The argument is that in a world full of machines that are designed to fulfill their purpose with full functional and computational power, playfulness can be a disruptive element towards computerization and "a carnivalesque attack on the seriousness of computers, on the system-driven thinking that gives maximum importance to the dictates and structures of a formal structure". Here, also the distinction of play and playfulness lies in play being the activity and playfulness an attitude towards an activity (Sicart, 2014).

Research has taken different views on the distinction between play and playfulness. Youell (2008) states that playfulness is an essential part of play. She further adds cultural aspects to the definition by adding the notion that self-experienced playfulness alone can only be experienced if an individual had a playful experience with another person and that absent playfulness in early development is a predictor for emotional, psychological, and cognitive development. Here, the distinction between play and playfulness is not strict, rather playfulness is defined as a part of play.

Work on the relationship between play and playfulness in children's development claims that play is not merely an activity. Rather, play can be looked at as an "internal predisposition to be playful", meaning that play is a characteristic of an individual, putting focus on the playful child instead of child at play (Barnett, 1990). This notion is not a novel one and is taken by many authors and goes back to Sigmund Freud "who regarded the child's play as expressive of personality patterns and internal desires" (Barnett, 1990). She builds upon research by Lieberman (1965), who defined five component dimensions of playfulness: physical spontaneity, social spontaneity, cognitive spontaneity, manifest joy, and sense of humor. These dimensions were used in questionnaire assessment for child playfulness where each item was assessed both quantitatively and qualitatively. The former represents the counted appearance of a behavior while the latter measures a certain quality of each dimension.

As discussed, playfulness can both be viewed as a characteristic of a person or attitude as well as a state of mind. In the HCI literature, there are also systems that themselves are propagated to be playful. Examples are often found in the area of playful user interfaces and playful products. The premise under which playful experiences are designed is that through playful design, the user experience can be enhanced, that playfulness can

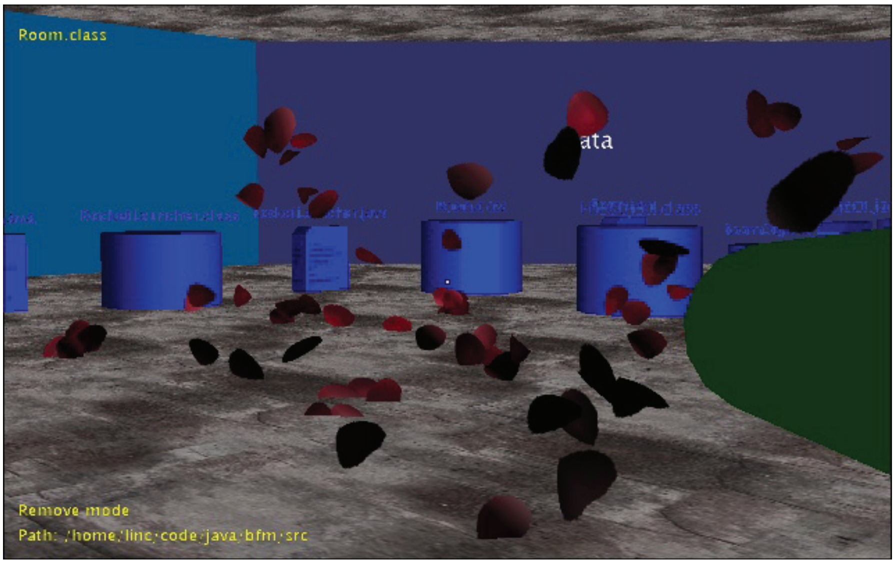


Figure 2.1: Screenshot of the Brutal File Manager. A file being destroyed by a rocket launcher. Other files are visible in the background. Image source: Johansson and Forchheimer (2005).

increase the motivation of a user towards a task or product, and that the learning process of new tasks or products can be supported (Kuts, 2009). She further provides user interface characteristics that can affect playfulness such as creative enjoyment, curiosity, exploration, fantasy, fun-in-doing, experimentation, challenge, or social interaction, to name the key aspects.

In the following, several digital playful systems will be presented, exemplifying the implementation and evaluation of the playful systems and user interfaces.

### 2.1.3 Playful User Interfaces

#### Games as Interfaces

There have been different approaches to improve the enjoyment of computer tasks by creating or modifying off-the-shelf computer games to support non-playing tasks. Chao (2001) presents a clone of the game Doom (id Software, 1993), where in contrast to the original game, processes running on a computer can be killed in the form of monsters in a 3D world. This game interface aims at supporting system administrators with process management, making this task more enjoyable and easy. Users found the software compelling and thought it was an intuitive interface. An interesting approach to file management is the “Brutal File Manager” (Johansson and Forchheimer, 2005). Users can walk in a 3D world that represents the file system of a local machine. Folders can be entered by walking into rooms that are filled with objects representing files. These can be deleted by shooting at them, see Figure 2.1. Further, a copy, duplicate, and paste functionality was implemented. Another first-person shooter metaphor was used by Christoffel and Schmitt (2002) for the exploration of libraries. The authors aimed at making the discovery of large library collections more enjoyable and modified the game Quake II (id Software, 1997). Books can be selected and information is shown by shooting at them, a second fire round displays extended information, and a third shot could even fill out the respective lending form. Brandstätter and Sommerer (2016) refer to the presented play environments, that foremost aim at a productive outcome rather than solely entertainment purpose as games for play of games for productivity (Fröhlich, Oehsen, et al., 2017).

### Tangible and Playful Interfaces

A non-fully digital implementation of playful user interfaces is represented by playful tangible systems that incorporate tangible user interfaces (TUIs). Fundamentally, Hiroshi Ishii and Ullmer (1997) presented early concepts of the tangible approach where users interact with real objects, artifacts, materials and other tangible objects in order to control a connected interactive system. Direct manipulation and immediate feedback are two of the core advantages of tangible systems, making them ideal interfaces for playful applications.

Sand or clay as interaction materials have been used in several cases in playful contexts. Piper et al. (2002) present their Illuminating Clay system for the analysis of landscapes in the landscape design process. Their system consists of a clay model that users manipulate, a laser scanner for capturing of depth information, and a projector that closes the feedback loop for visualization of different landscape analysis functions. The authors point out the simplicity and effectiveness of displaying 3D data directly onto the surface of a 3D model. Sand Garden (Reed et al., 2014) is a game where a digital landscape must be shaped, providing valleys, water, and mountains in order to make the virtual villagers happy. The system uses a projection and a second screen to display the game world. The sand itself is augmented and functions as a controller for the game. An educative use case for augmented sandboxes is presented by Sánchez et al. (2016), engaging students to experiment with topographical maps, exploring importance of water, erosion, or mountains in the evolution of a landscape. The advantages are better understanding of abstract concepts, better involvement, and more efficiently experimentation using a sand model. Inner Garden by Roo et al. (2017) is a tool to support mindful practices, which are defined as “the act of playing a deliberate and non-judgmental attention to the present moment” where “positive impact of a person’s health and subjective well-being” can be achieved.

Users create a miniature world by playing with the sand and the natural elements of the world are connected to physiological measurements like respiration, helping to stay focused on the body. This system also offers a VR mode where users can explore their creation virtually. Another interesting use case has been developed by AUDI AG (2017) where users can first form their own racing track in a non-augmented sandbox. The resulting race track is then imported into a VR driving simulator, offering direct per-

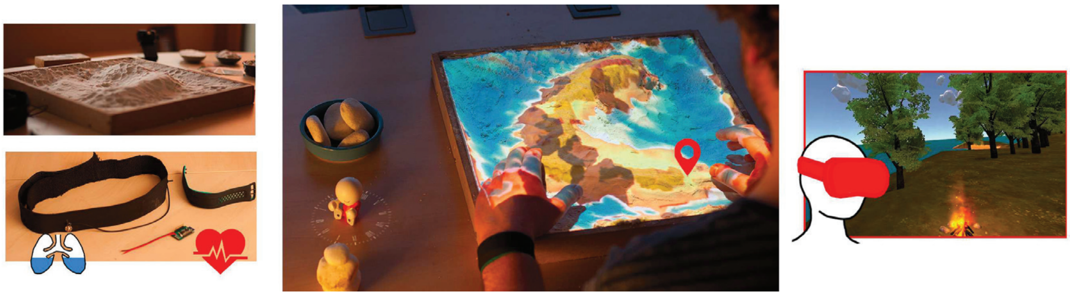


Figure 2.2: InnerGarden system for mindful practices. User's heart rate and breathing is reflected in VR, mirroring bodily state for self-reflection. Image source: Roo et al. (2017).

ception of the user's creation by driving through the digital model. However, no direct interaction with the sand in VR is supported. (Fröhlich, Alexandrovsky, et al., 2018).

### Ludic Interfaces and Interaction

The term ludic engagement has been coined by Gaver et al. (2004) in their work on the drift coffee table. With this device, users can playfully explore aerial maps of Great Britain through a tiny hole in the table center by placing objects of different weights onto the surface, see Figure 2.3. The design and prototype was driven by the question of how technology in the home can support ludic activities motivated by curiosity, exploration, and reflection in contrast to utilitarian systems that offer extrinsic motivation for engagement.

Ludic systems present an interesting approach to implementing Huizinga's concept of "Homo Ludens", humans as playful creatures (Huizinga, 1949). In contrast to playful interfaces there is no task (such as file management or scene design) that users perform in a playful manner. Ludic systems are specifically designed to enable playfulness, not just for entertainment purposes: "Such activities are not a simple matter of entertainment, or wasting time. On the contrary, they can be a mechanism for developing new values and goals, for learning new things, and for achieving new understandings." (Gaver et al., 2004). Gaver (2006) presents another ludic systems that allows looking at the skyline though a screen that shows a live view of a camera mounted on top of a 7m long mast. Through this simple system, users can experience new views of their commonly known environment, experience visual artifacts at sunset and sunrise, and even convey information in differing pleasing ways.

## 2.1.4 Playfulness and Creativity

Playfulness has been linked to creativity where a playful attitude can foster creativity and creative self-perception. This is due to openness and consequence-free creation that play and playfulness allow for. Because there are no consequences and no judgments,

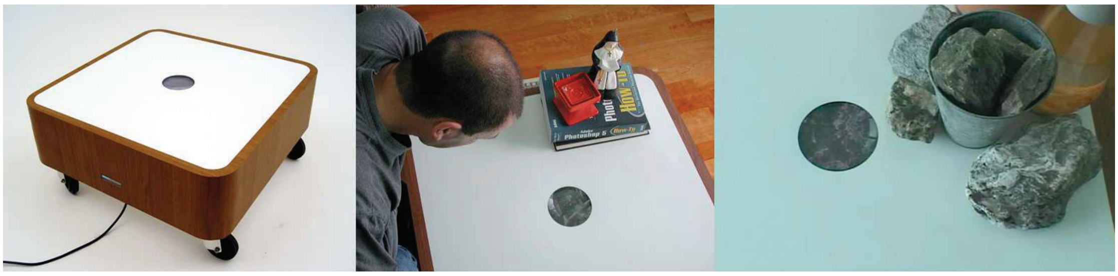


Figure 2.3: The Drift Table: a ludic interface. Users may place objects around the display, moving the aerial viewfinder around the map. Image source: Gaver et al. (2004).

play can facilitate creativity by enabling divergence in creation and through the use of existing tools, environment and context. For example, research found that the "quality of fantasy and imagination in play predicted divergent thinking over time" in children (Russ, 2003). Thus, through play, creativity can be supported. In this section, work on the boundary of creativity and play will be presented. In order to do so, first, a theoretical foundation of creativity is established, before presenting related work on systems that aim to increase user creativity.

Creativity can be observed and seen in everyday life where people are creative in finding novel solutions to problems, it can also be found in people that somehow change a way of thinking, defining a whole new field of products or research. Certainly, there are defining factors that each person adopts throughout life, including education, development, support, social background, and many more. While there are apparent differences between people, it is natural that there are differences in people's creativity.

Rhodes (1961) identified four strands, also known as the "four P's of creativity": person, process, press, and products . The term person includes "information about personality, intellect, temperament, physique, traits, habits, attitudes, self-concept, value systems, defense mechanisms, and behavior" (Rhodes, 1961). Process "applies to motivation, perception, learning, thinking, and communicating.", press "refers to the relationship between human beings and their environment.", and products implements ideas that have been turned into tangible form, such as "words, paint, clay, metal, stone, fabric, or other material." (Rhodes, 1961)

According to Kaufman and Beghetto (2009), Craft (2001), and others, creativity in people can be described using a Big-C, little-c, mini-c, and Pro-c taxonomy, see Figure 2.4. Traditionally, there is a big-c and little-c dichotomy, which distinguishes between "creative genius" and "everyday creativity". People, who are very creative but no geniuses are classified in the little-c category (Kaufman and Beghetto, 2009). Kaufman argues that the little-c category needs a further separation between people who score high on creativity tests like the Torrance test (Torrance, 1981), and people who's (creative) work is rated high among their peers in the little-c category. According to Kaufman and Beghetto (2009), there is a gap in little-c that doesn't fully account to people who gain

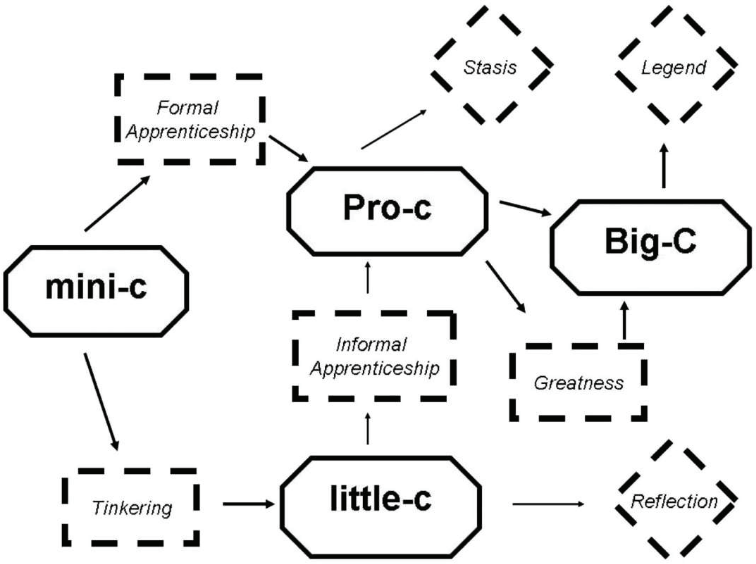


Figure 2.4: Kaufman's Four-C model (Kaufman and Beghetto, 2009)

"creative insights and interpretations involved in the learning". For this, mini-c creativity is defined as the "*novel and personally meaningful interpretation of experiences, actions, and events*" (Kaufman and Beghetto, 2009). The gain of personal knowledge and understanding is core to mini-c creativity, highlighting the importance of personal learning experiences and insight gain, that needs to be distinguished from little-c creativity as it would not attribute adequately to the differences in the creation of new knowledge.

According to Craft (2001), little-c creativity is an approach to life which is distinct from "high creativity", where people are driven to finding solutions to problems they encounter (for example losing a job, or choosing a profession) with a "can do" attitude also to life. Craft defines five criteria defining little-c, which are a) active and intentional taking of actions in the world, b) the coping with everyday challenges, both knowledge-wise and intuition-wise, c) innovation, d) "moving on", and e) problem identification and solving (Craft, 2001). Little-c and mini-c are the kind of creativity that appear in an everyday manner, where non-experts come up with novel ideas that could include drawing, writing, or building. In this thesis, little-c and mini-c are mostly targeted in the following studies and systems. Pro-c on the other hand requires expertise and training and a wider knowledge of a domain and is not targeted in this work because longitudinal studies would be required in order to investigate effects.

Two terms that are often associated with creativity are convergent and divergent thinking. These two concepts were first developed by Guildford (1967). Divergent thinking is when people generate multiple ideas and solutions to a given problem, exploring different, parallel solutions in order to find the best one. Convergent thinking on the other hand is when people formulate assumptions, models, and theories to a given problem and then

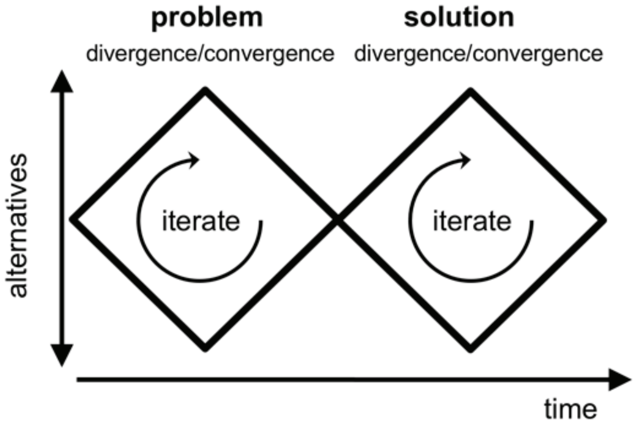


Figure 2.5: Double Diamond Framework, adapted from Design Council (2007).

coming up with one “ideal” solution to the problem. Guildford observed that creative people often employ divergent thinking in contrast to less-creative ones (Guildford, 1967). Based on Guildford’s findings, the concept of divergent thinking still holds today and it is being implemented in various creativity techniques (e.g. 6-3-5), as well as convergent thinking (e.g. six thinking hats).

Convergent and divergent thinking have been implemented in a design approach that is often related to as “double diamonds”. It represents a double diverge-converge pattern that was first introduced by Design Council (2007) and was later adapted by others, see Figure 2.5. This design method highlights the importance of generating ideas in a discovery and research phase, where many alternatives are explored. In the definition phase, the ideas converge to possible solutions in order to implement different solutions in a divergent manner. Finally, in the delivery phase, working solutions are selected for implementation in the context of use.

One could argue that the production of ideas through divergent thinking could lead to a lot of different ideas to a problem, but how useful are those ideas? Mednick (1962) picks up this issue and suggests that true creative ideas have to be useful in order to be truly creative. Mednick’s associative theory of creativity states that the probability of creating a true creative solution is predicted by the associative capabilities of individuals. In short, people who are able to connect and create connections (associations) between and across concepts are more likely to be creative (Mednick, 1962).

For this, Mednick (1962) presents three factors of achieving a creative solution: **Serendipity** (“[...] the manner of discovery to which is popularly attributed such inventions as the X ray and such discoveries as penicillin”), **Similarity** (“[...] encountered in creative writing which exploits homonymity, rhyme, and similarities in the structure and rhythm of words[...]”), and **Mediation** (“[...] the idea of relating reactive inhibition and cortical satiation[...]”) (Mednick, 1962).

One of the first researchers who discovered a link between divergent thinking and playfulness in young children was Lieberman (1965). She found that playfulness correlates with ideation fluency, spontaneous flexibility, and originality. Lieberman concludes that playful children are more likely to express divergent thinking compared to less playful individuals (Lieberman, 1965). This finding suggests that through playful activities, divergent thinking can be supported. With regard to interactive and playful systems this is interesting to note as it establishes a direct link between creativity and playfulness. More recently, Bateson (2015) confirmed this effect in a large-scale online study with 1536 participants where he found that "The link between *Acting playfully* and *Coming up with new ideas* emerged very strongly in our survey" (Bateson, 2015). Another interesting finding from this study is that in relation to Mednick's association theory, participants who rate themselves as playful and as producers of new ideas, were much more likely to come up with more associations to the tested concept than the control group.

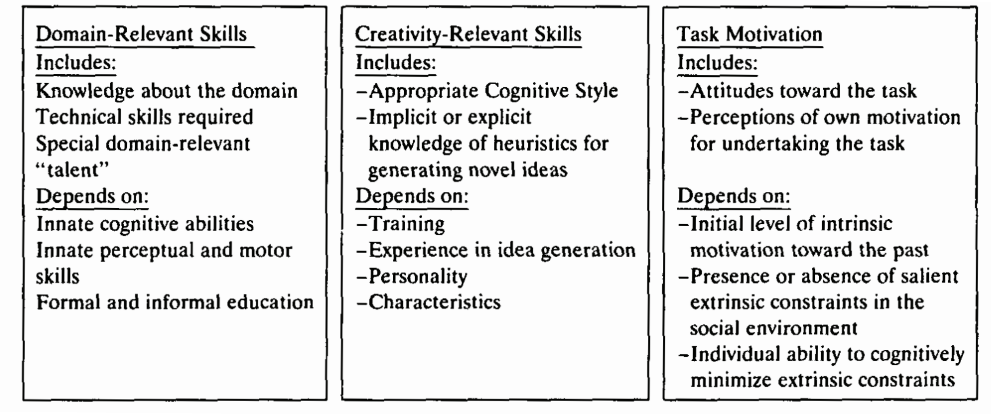
Bateson (2015) further discusses inhibiting factors of playful creativity along the work of Mihaly Csikszentmihalyi (Csikszentmihályi, 1990; Csikszentmihályi, 2013; Csikszentmihályi, 2014) and states that laziness and lack of direction can limit creative capabilities. Further, humor has proven to be a predictor of creativity, especially in combination with play as both go well together. In relation to play and creativity, Bateson argues that "By rearranging actions or thoughts, play generates novel ways of dealing with the environment, most of which lead nowhere but some of which may turn out to be useful — often at a much later date", indicating the short and long-term effects of playful activities. Further aspects include re-framing, the stepping away from known structures, towards finding new relations (Bateson, 2015).

While creativity has been heavily researched since the 1950s, the application of that knowledge in computer science was established decades later, first in the form of decision support systems (DSS). These systems provide users with information related to a certain decision making process in order to produce adequate and good decisions, but also to come up with creative responses to situations.

Elam and Mead (1990) review a three component model of creativity, with the components domain relevant skills, creativity related skills, and task motivation. In this model it is assumed that in order to develop creative responses in any domain, these three dimensions are necessary and sufficient. According to them, these components "do not represent cognitive processes, rather they represent a set of factors that control, determine, and enter into cognitive processes" (Elam and Mead, 1990).

Elam and Mead (1990) present five guidelines for the development of decision support systems that facilitate user creativity. These are

- Allow users to stop, store work session in process, then resume work later.
- Provide depth and positive tenor in its feedback.
- Make available to the user a full range of qualitative as well as quantitative decision aids.



Domain-Relevant Skills
Includes:
Knowledge about the domain
Technical skills required
Special domain-relevant "talent"
Depends on:
Innate cognitive abilities
Innate perceptual and motor skills
Formal and informal education

Creativity-Relevant Skills
Includes:
–Appropriate Cognitive Style
–Implicit or explicit knowledge of heuristics for generating novel ideas
Depends on:
–Training
–Experience in idea generation
–Personality
–Characteristics

Task Motivation
Includes:
–Attitudes toward the task
–Perceptions of own motivation for undertaking the task

Depends on:
–Initial level of intrinsic motivation toward the past
–Presence or absence of salient extrinsic constraints in the social environment
–Individual ability to cognitively minimize extrinsic constraints

Figure 2.6: Three component model of Creativity by Elam and Mead (1990).

- Technically easy to use and conceptually challenging.
- Provide an enjoyable or "fun" computing environment.

The last item relating to enjoyable or fun environments argues that users spend more time generating alternative ideas by encouraging users to examine a larger number of solutions and delay judgment. Further, the reduction of stress and external control is marked as beneficial and also the association of having "fun" being perceived as play rather than work has been shown to be linked positively to creativity. The creation of a larger number of alternatives relates to the concept of divergent thinking, as well as a delayed judgment, which is also a factor of divergent thinking. The positive attitude has been shown to benefit creative processes (Csikszentmihályi, 2013).

## 2.2 Natural User Interfaces

When humans work with machines such as computers, they need to somehow put their intent into action. For example, when buying a train ticket using a ticket machine, users can simply use their fingers to select a destination on a touch screen, navigating through the different options provided by the system, including a payment interface that may accept a variety of payment options, finally finding a printed ticket in the output compartment of the machine.

This procedure that the user interface of the ticket machine provides is already useful to many users. However, in the early days of computing, computers could only be operated by programming them to perform a certain task (e.g. punch cards). For this, "users" had to be programmers. There was no abstraction between the hardware and the software. Instructions were directly translated into machine operations. After this, operating systems were introduced that abstracted some of the functions of the hardware in order to make the programming more accessible (e.g. UNIX, DOS). It was this time, that command-line interfaces (CLIs) became popular, a first real user interface.

CLIs still required deep knowledge of the operating system, yet made many operations much easier. In the next iteration, computers were able to display more than just text, they became able to display complex graphical systems, known as graphical user interfaces (GUI), also known as WIMP (Windows, Icons, Menus, Pointers, e.g. Microsoft Windows). GUIs were the reason computers were suddenly accessible to a large number of people that could now operate computers without much knowledge about the internal processes or command lines. In the next wave of computing, devices became smaller and more powerful, resulting in tablet devices that moved away from point-and-click interfaces of stationary machines, establishing new ways of interaction: post-WIMP user interfaces such as gestural user interfaces (e.g. iPad). Using gestures it is possible to translate much of the user intend directly to the system through direct manipulation, keyboards are now only used for text input. Most of the interactions are performed through gestures (e.g. pinch and pan) (Preim and Dachselt, 2010, 463 ff.).

This brief overview of the development of user interface technology shows that with technological advancement, users require less formal knowledge about a given system with each iteration (Preim and Dachselt, 2010, 471 f.). When in early days, users had to be programmers in order to operate machines by entering very formal and correct instructions into the machines, nowadays, almost no prior knowledge and certainly no formal knowledge is needed for operation. Thus, gesture interfaces are often referred to as being natural because users can facilitate prior knowledge about the world in order to operate a tablet computer by touching interface elements, receiving immediate and useful feedback in order to progress (Preim and Dachselt, 2010, 472 f.). Still, users need to learn interfaces and the interaction with a system. However, modern user interfaces allow for easy and intuitive use of complex machines.

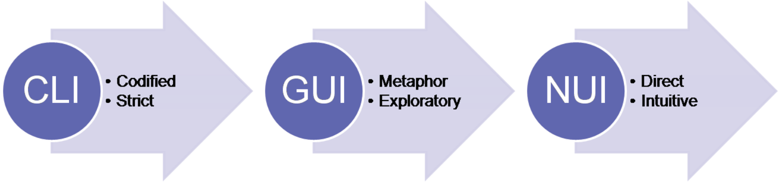


Figure 2.7: Progression of user interface technology. Source: NUI Group (2011).

Natural user interfaces take this development a step further by proposing that users can use existing (non-computer) knowledge for the interaction with machines. NUIs promise that users can rely on their world knowledge for system interaction where the computer is the one responsible for the correct interpretation of the user intent and action, and not the user (as was the case for programming, command-line interfaces, and GUIs, see Figure 2.7). This excludes the assumption that this interaction knowledge is inherited. Users still need to collect day-to-day knowledge for this kind of interaction. To sum up, the responsibility of correct system interaction is with NUIs a complete part of the computer system. Wigdor and Wixon (2011) put emphasis on the **natural user** interface instead of a **natural** user interface. The term natural user interface is a very compelling one, often used in the media and in the general public to describe interactions that "come natural" to us. Currently, there is no agreed-on general definition of the term NUI as the development and research are still ongoing. However, there are accepted assumptions about NUIs that can guide identifying interfaces that can be described as natural. Preim and Dachselt (2010) contribute the following list with different aspects of natural user interfaces :

- Publicity, familiarity
- Intuitiveness, fast access, apparent ease of use, fun and positive connotations
- Straightforward, low learning curve, little need for personal adaptation
- Easy to use (e.g. elderly, children)
- Real-world metaphors
- Familiar ways of using the own body, arms, hands, head movements
- Human-like communication with the system (speech, gaze, gestures, mimic, touch)
- High degrees of freedom (writing, sketching, speaking)
- Absence of apparent input technology (mouse, keyboard, etc.)

Figure 2.8: Videoplace system by Krueger et al. (1985) demonstrating natural gesture interaction: typing, drawing, and pinching. Source: Myron Krueger (1985).

In relation to this, there is still a need for discussion about the meaning and implications of natural user interfaces in the HCI community with workshops taking place at CHI 2018, the premier conference in HCI, dealing with the topic of "Redefining Natural User Interface" (Fu et al., 2018).

In the following, some prominent systems that represent NUIs will be presented. One of the first systems that fall into the NUI category is the "Videoplace" system by Krueger et al. (1985). Using projectors and cameras, users were able to interact with on-screen content using their own body, for example typing and drawing in mid-air. Further, they implemented a gesture set that already included the nowadays commonly known pinch gesture, see Figure 2.8.

Another groundbreaking system was developed by Bolt (1980b) at the MIT Media Lab. Users were able to interact with the system using speech and gestures in dialog-like style. For example, users were able to point at certain objects on the screen and invoke commands like "put this triangle [pointing] there" using a combination of speech and gestures. The tracking of the hand was implemented using a magnetic tracking device that was attached to the fingertip of the user, see Figure 2.9.

These systems were incredibly novel at the time and looking at them today, several interaction techniques have been adopted in consumer devices such as tablet computers, voice assistants (e.g. Alexa), and games (e.g. Microsoft Kinect, PlayStation Move). Still, 30 years later these systems present interactions that we today refer to as natural user interfaces and still, not all of the envisioned interaction methods have been introduced to the public yet, mainly because of technical limitations. In the following, some more recent systems that employ NUIs will be introduced.

For example, Mann (2007) developed a system for naturally interacting with musical instruments. Here, they took a different view on natural user interfaces by implementing physics-based interaction techniques for producing sound. With the "pneumatophone" for instance, users can produce sounds similar to a "blown bottle" using different finger positions. Other systems are the poseidophone (based on a glass-harp), or the hydraulophone (Mann, 2007).

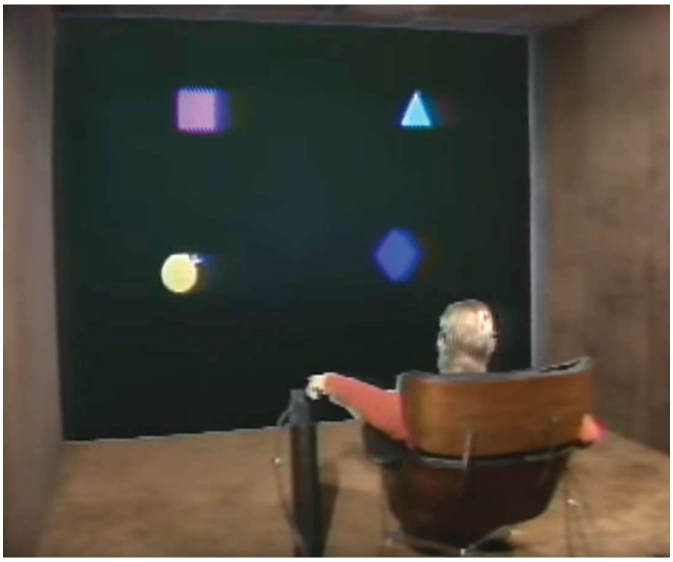


Figure 2.9: Put that there demo. Users employ speech and gesture commands to control the system. Image source: Bolt (1980a).

Innovation in HCI can also be driven by technological advancement in the consumer sector. With the introduction of the Nintendo Wii, suddenly, there was an affordable system available that allowed for bodily interactions in front of a TV screen, even including multiple actors and two-handedness. This lead to a large variety of research, implementing different systems facilitating gesture interaction in a larger space. For example, Francese et al. (2012) present a system that facilitates the Nintendo Wiimotes for map interaction in front of a large screen and found that “the more the interface is natural and involves their body in the action, the more the user are satisfied and involved in the 3D maps navigation experience.” (Francese et al., 2012).

Another example for this is the recent development in the sector of consumer VR hardware. Although VR has been an active field of research in the late 1980 and early to mid 90s, the hype wore off because the research could not be applied to consumer hardware because of the technological complexity and costs. With the introduction of the Oculus Rift, Virtual Reality research became very active again because a) sufficient quality of hardware (tracking, visualization), b) availability of development engines (e.g. Unity, Unreal), and c) low pricing. Looking at the aspects of natural interaction, VR offers many possibilities from a hardware perspective trough tracked controllers and head mounted displays, and the VR experience can be extended with other devices, providing an immersive experience. Related to this, VR incorporates intuitiveness through precise tracking, straightforwardness by using real-world metaphors (which are familiar to use), and very high degrees of freedom. It can even be argued, that the controllers represent an absence of apparent input technology as the use of them in combination with reality-based interaction metaphors becomes natural even for novice users after a short period of time.

This makes VR the ideal platform for natural user interface research and interaction in general. As an example of this, Zielke et al. (2017) present and discuss a system that allows for medical teaching in VR and AR using natural interaction metaphors. They further propose facial expression analysis, body posture, acoustic and linguistic pattern analysis, and natural language processing for increased natural interaction (Zielke et al., 2017). In the following chapter, exemplary applications of natural user interfaces in the context of previsualization will be introduced.

### 2.2.1 Natural User Interfaces for Animation and Previs

Previsualization (previs) is an essential phase in the design process of narrative media such as film, animation, and stage plays. Digital previs can involve complex technical tasks, e.g. 3D scene creation, animation and camera work, which require trained skills that are not available to all personnel involved in creative decisions for production.

Previs has been used for decades in all visual design disciplines such as film, product design, and architecture. Through the steady advancement of technology, digital previs gains more and more popularity in film, animation, and the performing arts. It is a collaborative, visual process that enables production teams to creatively explore ideas, plan technical solutions, and communicate a shared vision for efficient production (Okun and Zwerman, 2010). Although previs has many advantages for production, it is still costly, time consuming, and requires trained personnel using complex 3D tools such as Maya[1] and Blender[2]. In order to make digital previs accessible to non-technical users and to lower the costs in the pre-production phase, the EU-project "first.stage" researches and designs natural user interfaces for previs. A special interest of the first.stage project is animation. Animating digital content in a believable and interesting way can be a complex and tedious undertaking. However, animation is crucial for previs and interfaces that provide a natural access to animation are highly needed. We introduce some relevant and fundamental works in the area of performance animation that successfully implement natural user interface controls.

Implementations of full-body input as a natural interface to character animation are presented by Lee et al. (2002) by extracting user silhouettes from camera images and Chai and Hodgins (2005) who implement a more complex setup based on optical marker tracking. An affordable approach to performance animation is presented by Walther-Franks, Biermann, et al. (2012) with the “Animation Loop Station”, allowing users to create character animations layer by layer by capturing user’s movement with Kinect sensors. For system interaction, a speech interface is included so that users can fluently work on their animation without the need for a graphical interface. The “Dragimation” technique allows users to control timing in performance-based animation on 2D touch

---

[1] https://www.autodesk.com/products/maya/overview

[2] https://www.blender.org/

interfaces where they can directly interact on the characters instead of on a timeline. In a comparative study, the authors found that Dragimation performs better with regard to learnability, ease of use, mental load, and overall preference compared to timeline scrubbing and a sketch-based approach. Interview results support their findings as professionals “could well imagine benefits from using performance timing tools” in their workflow (Walther-Franks, Herrlich, Karrer, et al., 2012). This system is inspired by the work of Moscovich et al. (2005) who introduced a rigid body deformation algorithm for multi-touch character animation (Moscovich et al., 2005).

An interesting approach to augment the own character animations with rich secondary animations is the combination of performance animation with physics simulation. However, the combination of both technologies is not trivial. An approach is presented in a natural user interface that combines motion capture using the Kinect sensor and physics simulation for character animation by Liu and Zordan (2011). They provide a framework for the combination of the two technologies in order to overcome the “[...] potentially conflicting inputs from the user’s movements and physics engine.” (Liu and Zordan, 2011). Their approach and framework has been extended by Shum and Ho (2012) who present a more flexible solution to the problem of combining physical and motion capture information (Shum and Ho, 2012). Another area of complex 3D interaction is the field of digital modeling and sculpting. There is a strong need for tools that allow for natural expression in digital content creation, especially for previs, as many productions start with zero assets. This means that assets, objects, and props have to be created from scratch most of the time. For example, Herrlich, Braun, et al. (2012) investigated interface metaphors for 3D modeling and virtual sculpting, further contributing an interactive table that supports users in creating virtual 3D models on 2D interfaces (Herrlich, Braun, et al., 2012; Herrlich, Krause, et al., 2008).

Regarding virtual sculpting, a first implementation is presented by Galyean and Hughes (1991). In their system, a custom force-feedback with nine buttons is used for interaction, translating the absolute positions of the input device into a 3D mesh (Galyean and Hughes, 1991). This approach has been extended by Chen and Sun (2002) and Galoppo et al. (2007) by implementing virtual sculpting using a stylus device and a polygonal mesh instead of a voxel approach (Chen and Sun, 2002; Galoppo et al., 2007). Wesson and Wilkinson (2013) implement a more natural approach by using a Kinect sensor for deformation of a virtual mesh, while also integrating speech commands for a more fluent user experience (Wesson and Wilkinson, 2013). Natural animation can also be approached by using VR technology. For example, D. Vogel et al. (2018) designed a VR system for animation where users work with a puppeteering metaphor for character animation. They evaluated their tools with animation experts and found that it improves the speed of the workflow and fast idea implementation (D. Vogel et al., 2018).

### 2.2.2 Direct Manipulation

A well-established interaction concept that promotes intuitive control, fast learning and is easy to use is the direct manipulation concept (Shneiderman, 1984). It involves contin-

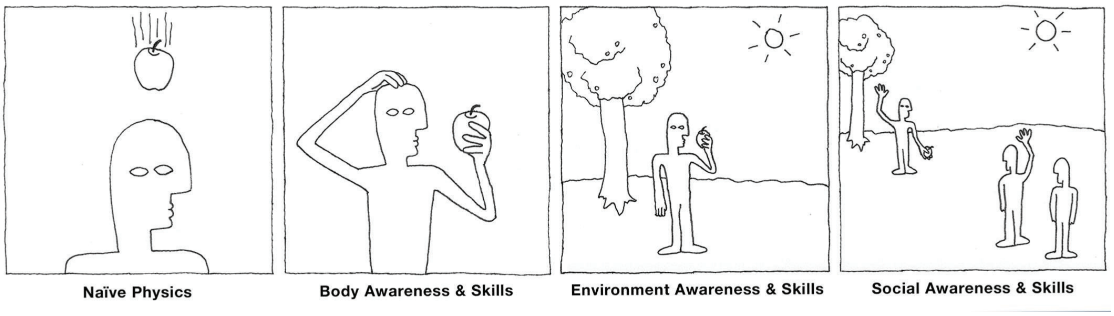


Figure 2.10: Reality-based interaction after Jacob et al. Source: Jacob et al. (2008).

uous representation of objects of interest and rapid, reversible, and incremental actions, as well as feedback. The intention of direct manipulation is to allow a user to manipulate objects presented to them, using actions that correspond loosely to manipulation of physical objects. No intermediate interface is needed to interact with an object of the domain space and actions have an immediate effect on the state of the object which is directly visible to the user. Direct manipulation systems implement continuous object representation, physical actions instead of complex syntax, rapid incremental reversible operations with immediate feedback on object impact, learnability, rapid task execution by experts, retainable operational concepts, no need for error messages, and immediate perception of goal achievement. Some limitations of the approach were identified by Frohlich (1993), which are also relevant for the application in previs software: manipulating small, distant, or attribute-rich objects under limited space, high density, or high precision; manipulating multiple objects simultaneously as a group (including group attributes) and manipulating intangible object properties.

Direct manipulation interactions can be built for 3D and 2D interfaces. However, in many cases 3D content has to be presented on a 2D medium such as a tablet or interactive surface. In such scenarios, it is not straightforward to design interaction techniques for 3D content manipulation on 2D surfaces. Thus, it is necessary to investigate how 3D content manipulation can be interacted with on 2D interfaces. For overall interaction with 3D content such as rotation, translation, and other tasks, several approaches and interaction techniques can be found in the literature (Herrlich, Walther-Franks, and Malaka, 2011; Herrlich, Walther-Franks, Schröder-Kroll, et al., 2011; Reisman et al., 2009).

### 2.2.3 Reality-based Interaction

Grounded in the post-WIMP redefinition of user interaction, Jacob et al. (2008) identified interaction styles that facilitate natural human understandings of how our world works by proposing the concept of reality-based interaction, which includes four themes: naïve physics, their own bodies, surrounding environment, and the presence of others, see Figure 2.10. Reality-based interaction makes interaction with computer systems more live-like by building upon world-knowledge, that all humans are familiar with. As this

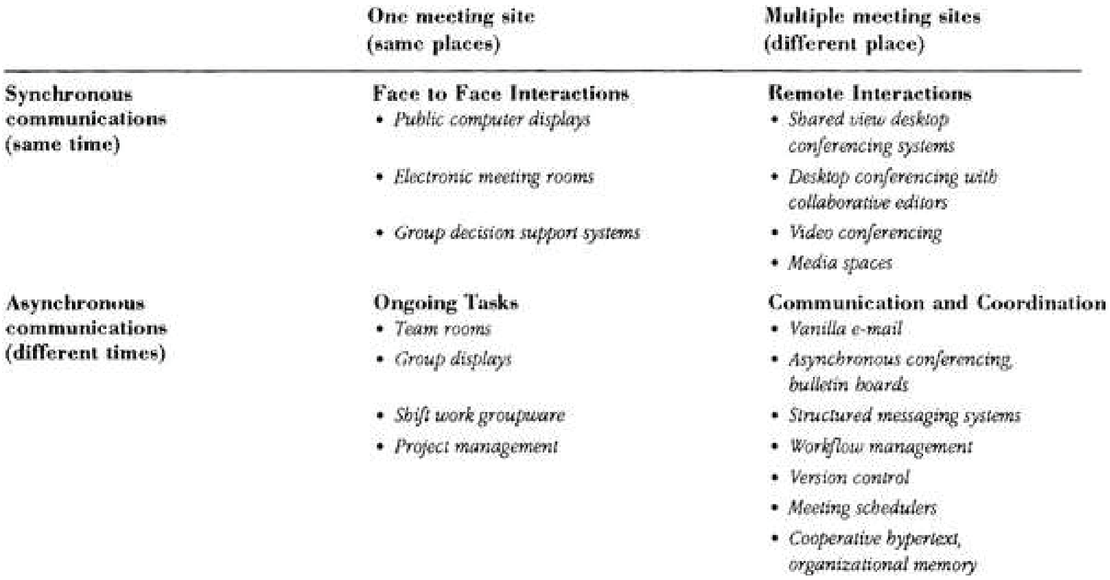
Figure 2.11: Different spaces have been created for shared interaction, reflection, personal space, and transfer. Image source: Geyer et al. (2011).

is a very promising approach to making more believable interactions that all users can relate to, the authors also state that the concept is limited in some ways and propose to give up on reality-based interaction, if some of the following qualities are more needed: "Expressive Power: i.e., users can perform a variety of tasks within the application domain, efficiency: users can perform a task rapidly, versatility: users can perform many tasks from different application domains, ergonomics: users can perform a task without physical injury or fatigue, accessibility: users with a variety of abilities can perform a task, and practicality: the system is practical to develop and produce." (Jacob et al., 2008). This list at the same time presents already known trade-offs of reality-based interaction.

In relation to this thesis, a relevant system implementing reality-based interaction for creative group work was developed by Geyer et al. (2011), see Figure 2.11. The respective use case for this systems is affinity diagramming, where multiple people work together in order to analyze a design problem or to come up with solutions to a given problem. The authors performed a study with observational techniques in order to identify interactions between people and artifacts that can be translated to reality-based interaction schemes. On basis of the results of a user study, it was found that the system was able to successfully preserve the general workflow of the method, mediated by the system.

### 2.2.4 Gesture Interfaces

One of the advantages of gestural interaction is that bi-manual interaction can be realized in the context of multi-touch interaction (Wagner et al., 2012). Free-hand or gestures-in-the-air interaction are techniques that require no physical medium in order to be executed. For this, the hands and fingers are tracked by devices such as the Microsoft Kinect or Leap Motion. Free-hand interaction is especially interesting in combination with VR, where users can work without using physical controllers. This can potentially increase the immersion and workflow of users. Free-hand interaction has been explored in virtual modeling (Kim et al., 2005), mobile augmented reality applications (Datcu and



| | **One meeting site (same places)** | **Multiple meeting sites (different place)** |
|---|---|---|
| **Synchronous communications (same time)** | **Face to Face Interactions**<br>• *Public computer displays*<br>• *Electronic meeting rooms*<br>• *Group decision support systems* | **Remote Interactions**<br>• *Shared view desktop conferencing systems*<br>• *Desktop conferencing with collaborative editors*<br>• *Video conferencing*<br>• *Media spaces* |
| **Asynchronous communications (different times)** | **Ongoing Tasks**<br>• *Team rooms*<br>• *Group displays*<br>• *Shift work groupware*<br>• *Project management* | **Communication and Coordination**<br>• *Vanilla e-mail*<br>• *Asynchronous conferencing, bulletin boards*<br>• *Structured messaging systems*<br>• *Workflow management*<br>• *Version control*<br>• *Meeting schedulers*<br>• *Cooperative hypertext, organizational memory* |

Figure 2.12: CSCW time/space continuum. Source: Baecker (2014).

Lukosch, 2013), virtual handwriting (Vikram et al., 2013), or as user interfaces for stroke rehabilitation (Khademi et al., 2014).

### 2.2.5 Cross-Device Interaction and Collaboration

In many design areas such as previs and architecture, working in small teams is common practice as different people with varying backgrounds come together in order to solve a complex design task. For example, as has been researched in the first.stage project, when planning a new theater play, a set has to be designed that fits the piece. In order to do so, set designers, directors, lighting personnel, craftsmen, and many more work collaboratively on a specific set design. Nowadays, these design tasks are supported by tools such as 3D modeling applications and computer-aided design (CAD) tools, different ways of computer-supported cooperative work (CSCW) have been researched and put into practice. In this sub-field of HCI, a large community formed in the past decades investigating ways of supporting co-located or remote group work. There are different ways of collaborating, that can be summarized in a two-dimensional space, consisting of the dimensions time and space, as introduced by Baecker (2014), see Figure 2.12.

As shown in Figure 2.12, collaboration or cooperation can happen synchronously, asynchronously, in the same place and across places. In recent years, CSCW was heavily influenced by different sub-fields in HCI, such as multimodal interaction (allowing for natural interaction through the use of different input and output modes such as speech and gesture), ubiquitous computing which is the omnipresence of computers (Scharf et al., 2013), and cross-device interaction (Scharf et al., 2013). Where multimodal interac-

Figure 2.13: The system of Wang and Lindemann. Participants use VR and a tablet device at the same time. Source: Wang and Lindeman (2014).

tion looks at different interaction modes on (often) one device, cross-device interaction considers different devices that can be used in exchange for the same task. For example, users in a 3D modeling scenario might switch from drawing on a tablet to modeling in VR inside the same environment, as presented by Kado and Hirasawa (2018): The authors present a system that combines a 3D CAD application with a VR component that was developed using the Unreal Engine 4[3], thus streamlining the design process by combining different ways of interaction. Thus, users can choose their style of interaction that is mostly suited for their task. For example, VR can be used by novices to explore 3D models and give feedback on the design, while designers work in the CAD application in order to be most expressive with the provided professional tools. Thus, cross-device interaction enables different users to work together and supports CSCW in a natural way by building upon the multimodal interaction paradigm.

Another interesting work was done by Wang and Lindeman (2014) who designed a hybrid virtual system that combines the use of VR with tablet interaction, see Figure 2.13. Here, users are immersed in VR while at the same time they can use the tablet in the same context of the task using a non-occlusive HMD. The authors argue that for some tasks tablets can be more beneficial than VR interactions and integrated both worlds into one system. Their prototype is designed for a single-user setup that uses the tablet for multi-touch interaction and an additional god-view, allowing for increased overview of the virtual world. The application for their system is a level editor that resembles basic 3D scene building capabilities. The authors found that having multiple contexts of interaction can lead to good performance beside the downside of an increased learning

---

[3] `https://www.unrealengine.com`

curve. It is pointed out that tasks are better synchronized across the different devices for better understanding of the users. Participants often made use of teleport and focus functions in order to change views and places inside the virtual environment. The tablet was mainly used for 2D tasks that require different angles and overview. VR interactions were mainly used for detailed tasks like painting, selecting and placing individual objects.

## 2.3 Summary

This chapter provided an overview of the current literature regarding playful user interfaces and systems as well as natural interaction and the application in different areas such as animation and previs. It also connected the aspects of creativity to the concept of playful interaction, building the foundation for the following research chapters. This chapter also shows that there are only a few disruptive systems that facilitate user creativity and exploration. Further, existing natural interaction systems are build for system manipulation and content interaction. There is still a lack of systems that facilitate playful and natural interaction patterns for the creation of user content such as 3D models and animations.

Therefore, in the next chapter three playful systems are presented that center around content creation using playful interaction.

# 3 Creativity and Playful Interaction

Creating 3D content can be a challenging task, especially for users without knowledge of 3D applications. We argue that non-technical users are being held back from creating 3D content because of the complexity of currently available tools and the high learning effort which can lead to motivational deficits in users. Further, novices often don't know where to start, being confronted with a "tabula rasa" situation where they find themselves in front of an empty scene, not knowing how and where to begin. What is needed are tools that support non-technical users and beginners overcoming these issues, making the access to 3D content creation more accessible.

This can be done by integrating playful interfaces to the domain, reducing the barrier for non-technical users and beginners by focusing on a positive and engaging user experience, where users are not confronted with their little to non-existing knowledge. Instead, by reducing the seriousness for these kind of tasks, users can freely experiment and create without the implicit judgment that non-familiarity might induce, possibly increasing the motivation of continuing creating 3D content.

In this chapter, three different playful and disruptive ways of digital 3D content creation are introduced. First, a system for fast and easy 3D scene generation that centers around the mechanics of a first-person shooter game. Users can shoot 3D objects into a scene, playfully engaging with 3D content interactions such as placement, scaling, and transformation by choosing different "weapons" and modes for setting a scene. Second, we introduce a playful sandbox system where users can play and interact with real sand while wearing a VR headset for increased immersion and presence. By building 3D meshes with their own hands, no prior knowledge is required and users can directly travel into their creations using VR technology. Third, using a modded version of Minecraft, we present the idea of playfully supporting creative expression by dynamically introducing creative limitations for creativity support for non-technical users.

## 3.1 Playful Shooter for Fast and Easy Scene Design

Creating user-generated 3D content for video games is mostly done with tools for map or mod making like Valve's Hammer Editor (Valve, 2017). These world or map builders can be cumbersome to use because of their complex user interfaces that are hard to learn and impose a limited user experience, as can be seen in Figure 3.1. Often, the editor perspective doesn't match the play view (e.g. first person games), disconnecting creation and experience through constantly switching perspectives. It is further difficult to position objects in 3D space using the mouse and constantly switching between different Gizmos, often operating only in one dimension at a time. Editors also tend to be oriented on the task, lacking support for creative expression and serendipity, which are both important dimensions of design.

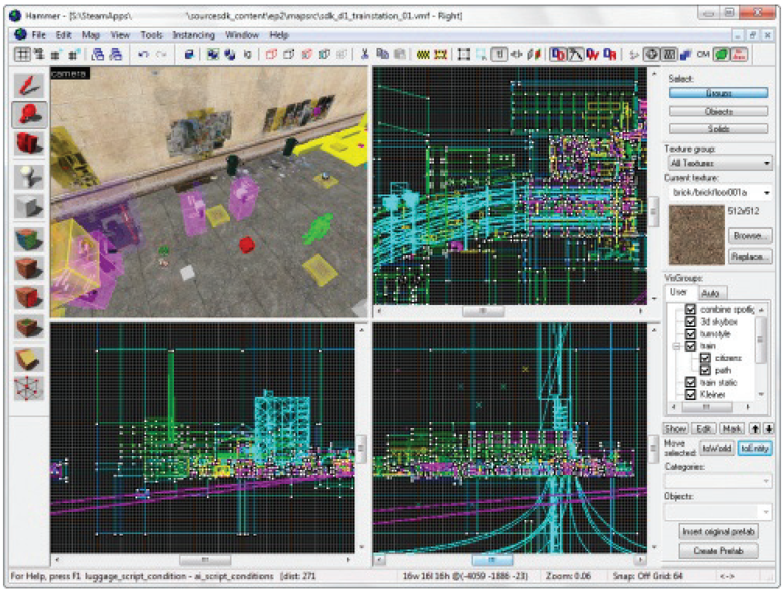


Figure 3.1: CS:GO Editor "Hammer" with a complex and hard-to-learn user interface. Image source: Valve Developer Community (2020)

In order to overcome these problems, we present a world builder game where the process of creating a 3D scene is supported by a playful game interface, see Figure 3.2. Based on the author's idea and motivation, the prototype system has been developed in the Bachelor thesis of Oehsen (2017) that the author has supervised and collaborated in. The work further has been published as Fröhlich, Oehsen, et al. (2017).

In the prototype, users are able to navigate directly in the 3D scene, placing objects in a playful manner by shooting them in the scene for placement and manipulation. We implemented four weapons: A gun for physics enabled placement, a laser gun for more precise positioning and manipulation, a sniper rifle for far distance interaction,

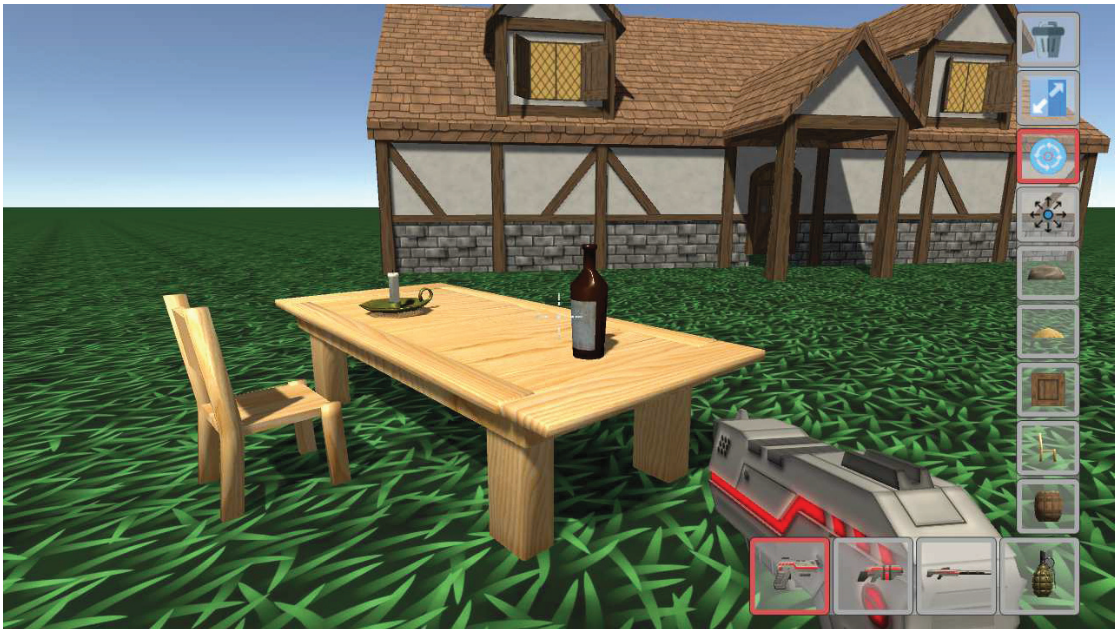


Figure 3.2: Game view with gun selected. Top right: Menu with manipulators (Delete, scale, rotate, translate, and assets), bottom right: Gun, laser rifle, sniper, grenade.

and a hand grenade for spawning multiple objects at the same time using an explosion, all allowing for playful interaction with the 3D content. For example, using our tool, users can create a small town by walking through the world, placing houses, crates, and decoration on the go, throwing a tree-grenade to spawn a forest, and using the rifle scope to place object in a detailed manner. With our approach we assume that users find the design of 3D scenes more easy and fun, feel more creative, have less difficulty positioning objects in 3D, and have a lower learning curve compared to standard editors. In this sense, we are interested in the hedonistic and pragmatic qualities of our game, as well as the overall usability of our system. Based on this, we formulate the following research questions:

**RQ1**: How creative do users feel using the game in comparison to the editor?
**RQ2**: How does the overall usability of game and editor compare?
**RQ3**: How do users rate the pragmatic and hedonistic quality of game and editor?

For the evaluation of our approach we conducted an experiment with 17 participants where we compared our game to a simplified version of the Unity editor, representing usability and workflow of standard 3D tools like Hammer. We were interested how users perform both in a free building task and in a replication task where they had to rebuild a 3D scene from a printed template. In addition to this, we conducted a short interview session with each participant, collecting feedback on the approach. Because we assume that our game is more enjoyable and creative as the editor, we were especially interested

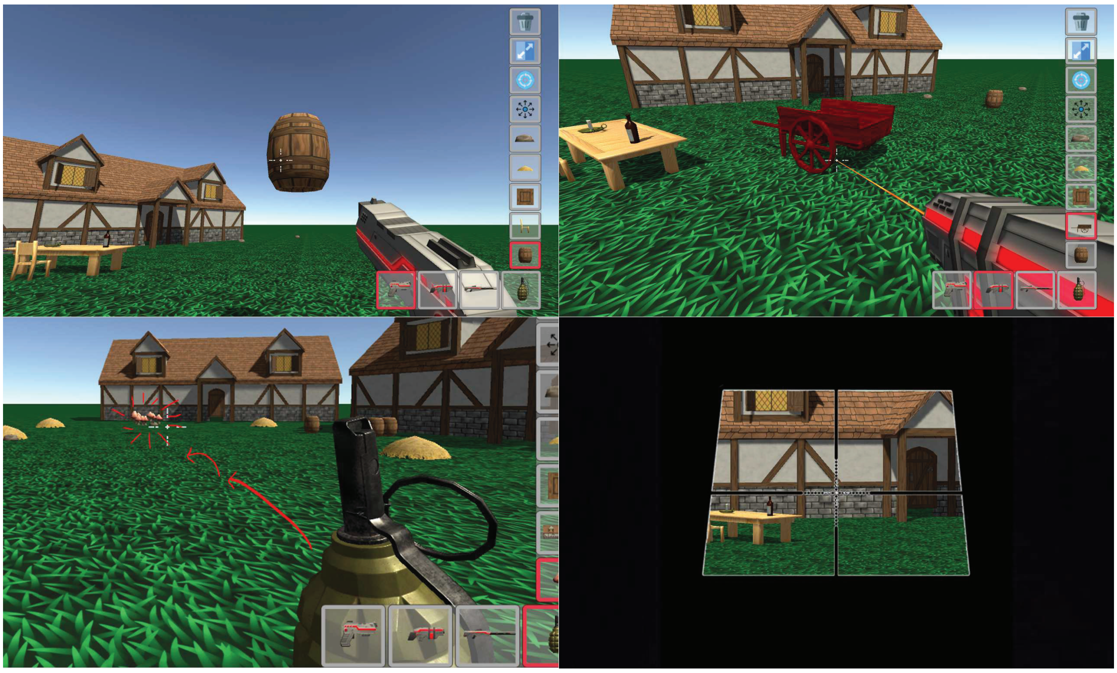


Figure 3.3: Overview of the implemented guns (clockwise): pistol, laser, sniper, grenade.

in the hedonistic and pragmatic dimensions, so we used the AttrakDiff questionnaire (Hassenzahl et al., 2003) which gives insight in both directions. For usability, we chose to use the System Usability Scale (Brooke et al., 1996).

### 3.1.1 Prototype Description

Our game uses mechanics similar to common first-person shooters. The player navigates using the WASD keys on the keyboard for movement and the mouse to look around. Objects can be placed in the scene by "shooting" them into the world using the left mouse button. Weapon loading is implemented in a Minecraft-like menu on the right side of the screen. Four different weapons for object placement and manipulation can be selected, each having individual functions. The weapons can be loaded with either 3D assets or manipulators, which are translation, rotation, and scaling. When shooting with a 3D asset loaded, the object is placed in the scene, influenced by weapon specific functions. Shooting a manipulator at an object in the scene leads to either translation, rotation, or scaling of the target object. This way, already placed objects can be repositioned, rotated, or scaled up and down.

For object placement, we implemented the four different weapons: A gun, laser gun, sniper rifle, and a hand grenade. The gun spawns objects at the muzzle, adding a forward directed force using the physics engine and when objects hit the ground, they slide a small distance before coming to halt, see Figure 3.3 top left. When aiming with

Figure 3.4: Template scene that had to be recreated in the study.

the laser gun, a red shaded preview of the selected object is drawn in the world, indicating the final placement upon shooting as depicted in Figure 3.3 top right. The sniper rifle draws a target point in the world, indicating the placement position. This weapon has a rifle scope (Figure 3.3 bottom right), allowing for precise positioning from far distances. Upon shooting, the selected object directly appears at the target position. Throwing a hand grenade into the scene spawns multiple objects in an explosion, distributing them circularly around the detonation point, see Figure 3.3 bottom left.

Manipulations to already placed scene objects can be done by shooting at a target object with one of the manipulators translation, rotation, or scaling selected. With the gun and sniper rifle, translations in X,Z space are done by applying small forces to the target object, physically moving it in the direction of shooting. Rotation is done the same way, only target objects have to be shot at the outside edges, rotating it on each shot a small bit around the Y axis. Scaling works similarly by hitting either the top or bottom half of the target object, increasing the size equally in all three dimensions. With the rail gun, objects can be directly re-positioned in X,Z dimensions by hold-shooting and moving the mouse. Scaling and rotation work the same way. Hand grenades do not have manipulation functionality. 3D objects can be selected using the mouse wheel from the right hand menu. When pressing the TAB key, another window appears and using the mouse, other objects can be added to the game menu using drag and drop.

### 3.1.2 Study

We conducted a within-subject experiment with 17 male participants (25 years on average, $SD = 5.3$), all having an undergraduate computer science background. For baseline comparison, we decided to evaluate our game against the Unity editor as a representative tool for standard world editors. In order to not over-complicate editor over game use for novices, we simplified the UI of the editor by excluding unnecessary options and highlighting task relevant elements like the asset library and providing large thumbnails of the required assets, see Figure 3.5. The experimental task was to a) freely explore either the game or editor while creating an own scene, and b) replicate a sample scene (Figure 3.4) using both tools.

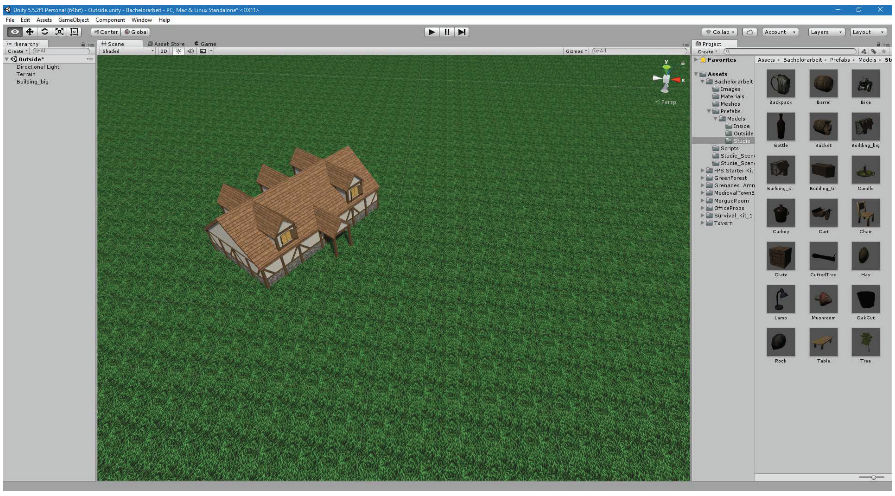


Figure 3.5: The Unity interface was simplified so that only necessary items are visible.

Our experiment has two conditions: a game and an editor condition. Every participant performed both conditions, randomly starting with either the editor or game. In each condition, after an introduction to the respective tool, participants performed a free building phase of three minutes where they could experiment to their extend. Subsequent to this, in the replication task, participants were asked to re-create a given scene within 5 minutes using the same tool. The template scene (see Figure 3.4) was shown on two print outs, each from a different perspective for better overview and were not removed during this phase. Upon reaching the five minute time limit, the task was discontinued. We did not randomize the task order (free building / replication) because of the small number of participants.

After welcoming and a general introduction, the participants were introduced to the first tool (game or editor), where the controls were explained and questions could be asked. We provided a "cheat sheet" with the controls listed for convenience. Then, the 3 minute free building phase began and after finishing, the resulting scene was saved by the experimenter. Following, using the same tool, the participants replicated the

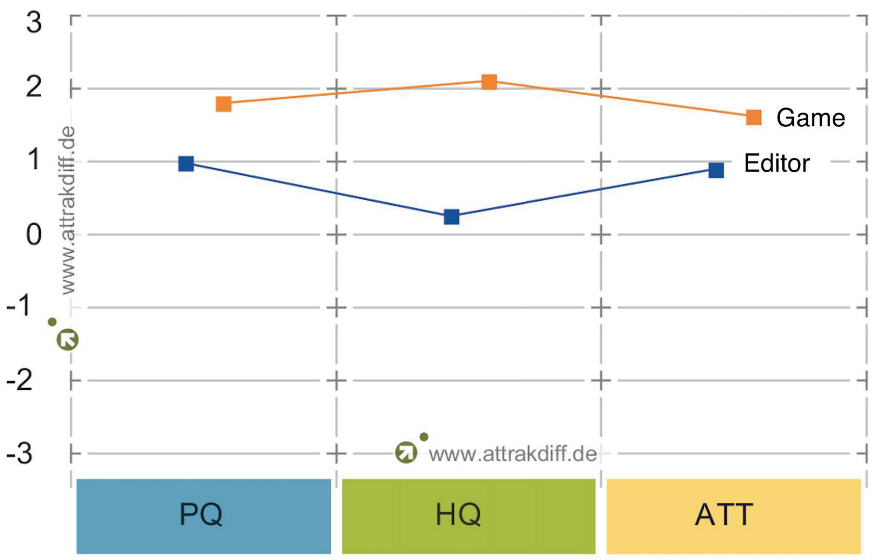


Figure 3.6: Average pragmatic (PQ), hedonistic (HQ), and attractiveness (ATT) ratings

template scene while being instructed to think-aloud. After completing the replication phase, participants filled out the AttrakDiff questionnaire and the System Usability Scale. Then, the participants performed the second condition with the same procedure, only using the alternative tool. After both conditions were finished, in the final step, a short structured interview was conducted.

### 3.1.3 Results

The System Usability Scale (SUS) yields an overall of 85 points for the game and 63 for the editor. We tested the SUS categories for significant results using an Wilcoxon-Test and found the following answers being statistically significant ($p < 0.05$). A higher score means more agreement. Q1: Would like to use the system frequently: Game 3.94 ($SD = 0.97$), Editor 3.18 ($SD = 1.13$), Q2: Found the system unnecessarily complex: Game 1.53 ($SD = 1$), Editor = 2.24 ($SD = 1.15$), Q4: Would need the support of a technical person: Game 1.47 ($SD = 0.62$), Editor 2.59 ($SD = 1.37$), Q5: Functions are well integrated: Game 4.41 ($SD = 0.87$), Editor 3.59 ($SD = 1$), Q7: Most people would learn this system very quickly: Game 4.59 ($SD = 0.8$), Editor 2.88 ($SD = 1.41$), Q8: System is very cumbersome to use: Game 1.53 ($SD = 0.62$), Editor 2.76 ($SD = 1.09$).

AttrakDiff results of the average pragmatic, hedonistic, and overall attractiveness of both the game and the editor can be seen in Figure 3.6. Figure 3.7 shows the individual ratings of the word pairs used in the questionnaire. We further applied a Wilcoxon-Test to the individual categories of the AttrakDiff questionnaire and found the following statistically significant ratings ($p < 0.05$), a higher score means better. Simple - complicated: Game 6.35 ($SD = 0.79$), Editor 4.59 ($SD = 1.62$), good - bad: Game 6.13 ($SD = 0.7$), Editor

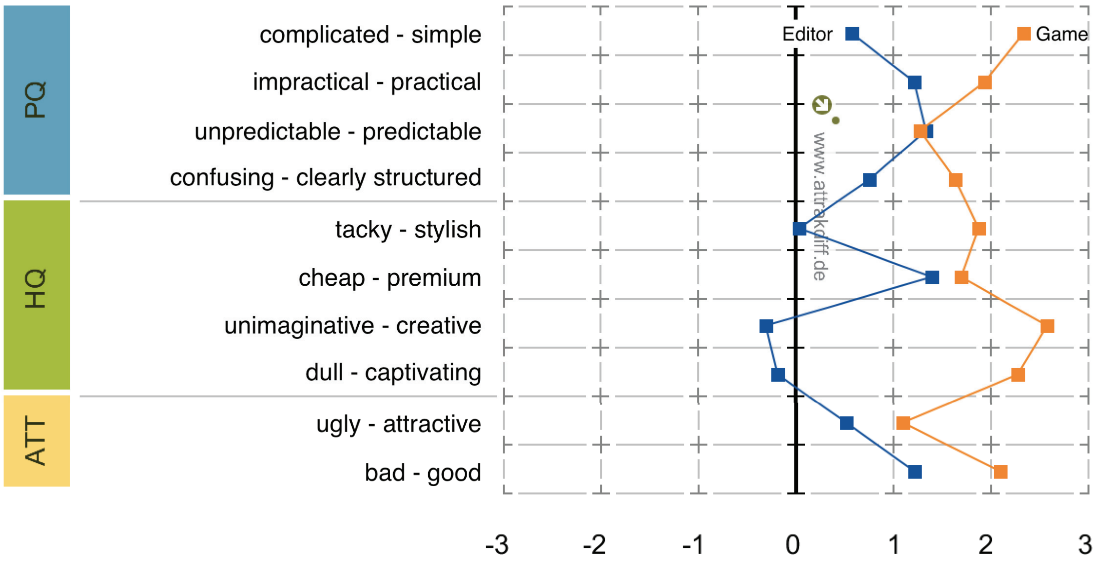


Figure 3.7: Profile of the word pairs (PQ = pragmatic quality, HQ = hedonistic quality, ATT = attractiveness)

5.24 ($SD = 1.4$), captivating - dull: Game 6.3 ($SD = 1.1$), Editor 3.8 ($SD = 1.42$), creative - unimaginative: Game 6.59 ($SD = 0.87$), Editor 3.71 ($SD = 1.89$), attractive - ugly: Game 5.88 ($SD = 1.36$), Editor 4.06 ($SD = 1.48$).

### 3.1.4 Discussion and Conclusion

For the discussion of the results, we present the main findings of our experiment. Regarding overall usability, we found statistically significant advantages of the game in comparison to the editor, where also the absolute SUS score of 85 for the game and 63 for the editor demonstrates a higher usability of game. In this regard, we found that users would like to use the game more frequently, found it less complex, would not need technical support and appreciate the learnability and ease of use.

As we were especially interested in the pragmatic and hedonistic value of our game, we included the AttrakDiff questionnaire. Here, we found that hedonistic and, interestingly, also the pragmatic quality of the game is rated higher than for the editor. This is surprising because the editor is very task oriented and pragmatic in it's design, functionality, and user experience. More data is needed to investigate this unexpected result. In relation to our research questions we can say that users felt more creative and had more fun using the game. Based on the SUS results, there is evidence that the game usability is generally higher, and the AttrakDiff results indicate that also the hedonistic and pragmatic values are also in favor of the game.

In a short interview session after the experiment we collected further verbal feedback from the participants. Half of the participants thought that they achieved a better result using the shooter, some mentioning the precision as something that could be improved. The other half was comfortable with both conditions or were in favor of the editor mode. In general, the shooter was perceived well by the majority of the participants. Positive highlights regarding the shooter were the free movement in the scene and the ego perspective that made it easier to place objects in 3D space. Additionally, the shooter was perceived as being more fun and easy to learn by most of the participants. Further, the shooter was very fast and easy to operate and placement of objects was positively affected by the built-in physics, which made interaction much more natural.

Some participants were unsure whether the shooter would be suitable for more professional use cases and also mentioned that the shooter was lacking precision. Further, tools like select, duplicate, copy and paste were mentioned to be missing. Positive about the editor was that the camera could be placed more freely. Negative comments towards the editor were positioning errors and hurdles, the unintuitive rotation of objects, and how the centered-based scaling of objects worked.

## 3.2 Playful VR Sandbox for Creativity and Exploration

Playing with sand is something most people relate to from childhood days, where this unstructured spontaneous, enjoyable, voluntary and non- goal directed activity provided room for experimentation, exploration, or cooperation (Campbell, 2004). As children we develop important skills in the sandbox like proprioceptive sensing, body and space awareness, or social skills and can engage in creative behavior and open-ended play (Jarrett et al., 2010). These memories can last a lifetime and, when given the opportunity, many adults enjoy playing with sand and rediscovering the sensations of this activity.

Augmented sandboxes offer these experiences and further add interactive digital elements to the sand, which allow for increased interaction and application possibilities. Users can play with the sand and make use of their natural ways of expression while at the same time perceiving interactive visual feedback of digital content that is projected onto the sand. Through these visuals, changes of the sand surface can be immediately experienced and reacted upon to. In the literature, there is a considerable body of research that has explored augmented sandboxes and their tangible affordances (Reed et al., 2014; Roo et al., 2017). In many application scenarios, sandboxes have been applied to offer a better understanding of spatial and geographical phenomena (H. Ishii et al., 2004; Piper et al., 2002), or to experience terrain playfully (Roo et al., 2017; Couture, 2017). Most setups use a projector mounted above the sandbox, projecting an image on the sand surface and a depth camera for tracking the shape of the sand. For example, Reed et al. (2014) use a heat map-like projection on the sand to indicate different types of landscapes.

While setups like these already offer direct interaction with sand as natural material to change the visual output, we extend existing approaches by including virtual reality (VR) for 3D visual perception and mid-air gestural interaction for system interaction. To do this, we scan the sand surface in real-time, creating a virtual representation of the sand surface in VR. Using a Vive VR headset, users can interact with the virtual sand surface and the real sand surface at the same time due to a precise mapping of the two. System interaction is performed using gesture interaction with a Leap Motion sensor. Our system offers new ways to use and explore the sand landscape, such as by placing virtual objects into the scene or by teleporting oneself into the environment.

The system extends existing approaches with an additional interaction and visualization layer to interact with the sand and virtual objects. Furthermore, users can switch perspectives from a tabletop mode into a first person mode by teleporting directly into their own creation, allowing for an immersive experience of the landscape in full-size, see Figure 3.8. Based on the idea of Dmitry Alexandrovsky and the author of this thesis, a prototype system has been developed and evaluated in the Master thesis of Stabbert (2018) that the author has collaborated in and supervised. Based on this work the publications Stabbert et al. (2017) and Fröhlich, Alexandrovsky, et al. (2018) emerged.

In the context of this thesis, we evaluated the system regarding playful, natural and creative capabilities in order to assess how VR and mid-air interaction can benefit these

Figure 3.8: VRBox from first person perspective. The landscape can be shaped by interacting with real sand.

aspects. We performed the evaluation in a qualitative user study with nine experts from computer graphics, education, and game design in order to investigate playful and natural aspects of the system.

The results show a very positive attitude towards VRBox, highlighting the creative experience that the system offers due to the novel first person experience and extended interaction capabilities using virtual objects in VR. This thesis further contributes application scenarios in productive and entertainment contexts such as gaming and collaboration.

### 3.2.1 Prototype

The box has a volume of 140 × 80 × 30 cm and provides a good balance for single user as well as for collaborative scenarios. Three Kinect V1 sensors are mounted on a railing above the sand surface. The tracking volume is captured from three angles reconstructing of a full 3D point cloud of the sand model. For virtual reality, an HTC Vive[1] is used. To enhance the immersion, we use a Leap Motion[2] sensor, attached to the head-mounted display, providing information about the hand and arm positions of the users. Figure 3.9 (left) shows an overview of the incorporated components.

For providing a satisfying VR experience, the application has to run at a steady frame rate of at least 60 fps to prohibit motion sickness in VR. Hence, the visual representation of the scanned surface had to be abstracted in the form of discrete blocks with a mesh size of 98 × 59 blocks with 32 height levels resembling a Minecraft-like (Persson and

---

[1] `https://www.vive.com/`

[2] `https://www.leapmotion.com`

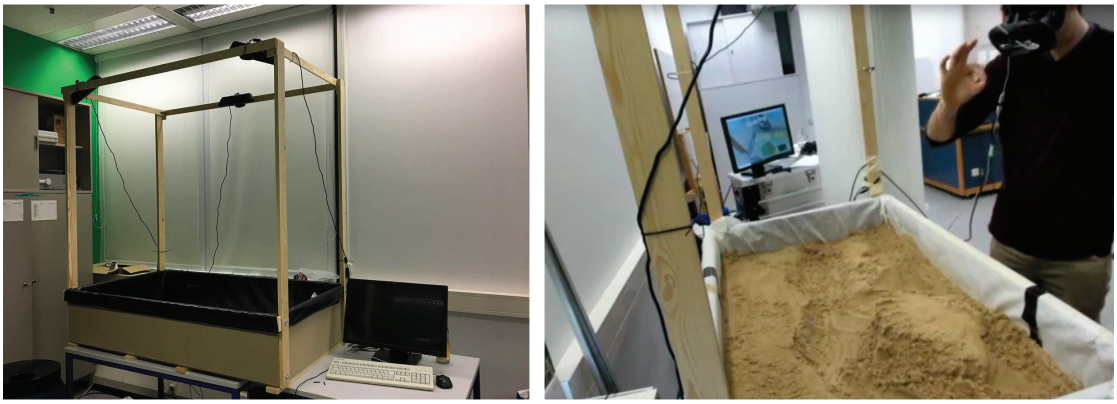


Figure 3.9: Left: The full VRBox system. Right: A user performing a pinch gesture for spawning objects.

Bergensten, 2009) look. In early tests, it was found that this resolution provides a good compromise between level of detail and latency.

**Use case**

We implemented a terraforming use case where users form a 3D world of their liking with the sand and extend their creation by adding a water level to create rivers and other waters, add different lighting moods and decoration objects like trees, flowers, and castles. In VR, users see the virtual representation of the tracked sandbox and sand surface and form a terrain with their bare hands. Furthermore, they can interact with a virtual menu using their hands. The menu items are placed to the left and right of the user in VR and are represented as bubbles that users can interact with using their hands, see Figure 3.11. The bubbles have different functions. First, a set of them can be used to decorate the sand surface with virtual objects like trees, castles, and other assets. These can be used in the sense of a pouring glass, where users grab a bubble using a pinch gesture (see Figure 3.9 right), e.g. containing a tree asset and "pour" the trees onto the sand surface, see Figure 3.10. Individual objects can also be placed in the same way. We concentrated on a small set of items that can be put onto the sand. On the left side of the user, one bubble can be used to manipulate the water level in the form of a vertical slider. Another bubble serves as a teleport item. Users can grab the teleport bubble and put it on the sand surface to mark the location where they would like to be teleported to. Then, users have to perform a "look up" gesture where they have to look up straight in order to perform the teleport. When the head reaches the highest point, the teleport is performed and when looking down, users see from a first person perspective, standing directly in the virtual world. Looking up again teleports back into tabletop mode. Further, a delete bubble and lighting bubbles (create sunset or sunrise) were implemented.

Figure 3.10: Spawning trees into the 3D world using the pinch gesture.

### 3.2.2 User Study

We conducted a qualitative user study with nine experts from computer graphics (P1, P3, P4, P5), education (P8), game design (P2, P6, P7) and HCI (P9). All participants were between 30 and 50 minutes in VR. On average, the subjects formed terrains for 20 minutes and explored the environment from first person perspective for about 10 minutes.

#### Participants

P1 works as a computer graphics (CG) researcher creating virtual testbeds for VR research projects. P2 is a 3D environment artist working at a local game studio. P3 works as a researcher developing tools for VR. P4 is a CG researcher. P5 is also a researcher with a CG background and game developer working on VR and multi user VR environments and telepresence. P6 is a game designer with a local game studio working on gameplay and game mechanics. P7 is an experienced game designer creating virtual environments and levels for a racing simulation. P8 is a teacher for mathematics and biology currently working in youth welfare services. Participant 9 is an HCI researcher in the area of educational user interfaces and a hobbyist game developer. Our participants are between 20 and 39 years of age, with eight males and one female subject.

#### Procedure

The sand was watered and stirred up, keeping sand stability constant. We smoothed the surface to provide an empty canvas for each participant. After a general introduction, participants were introduced to the system, explaining the overall functions. We demon-

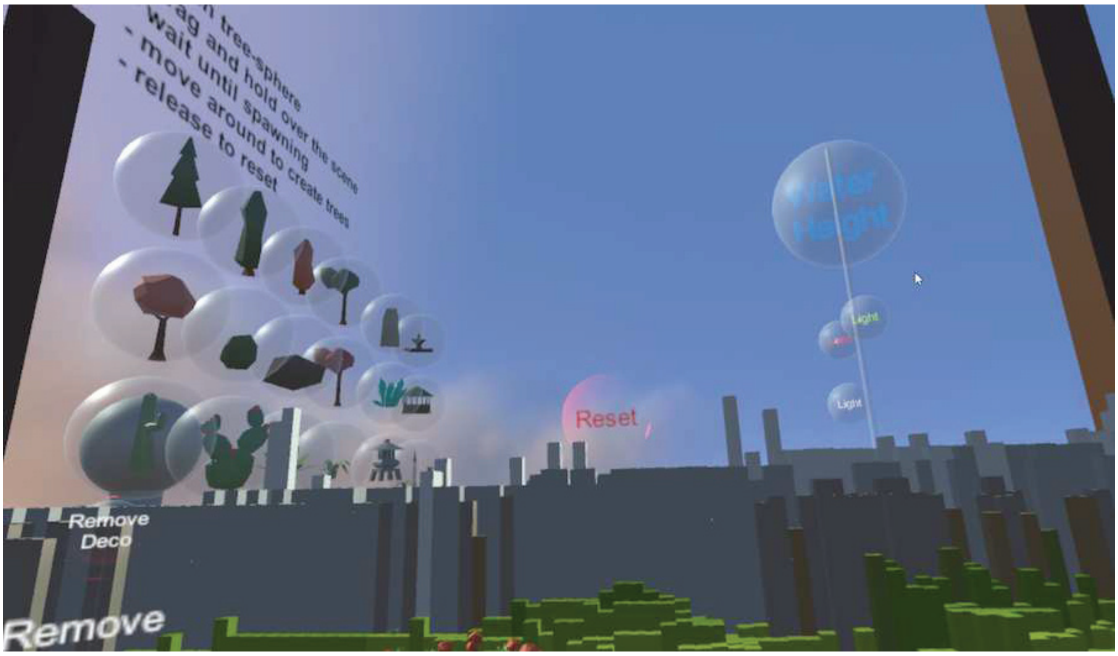


Figure 3.11: View on the VR menu. On the left are the decoration objects. On the right, the teleporter, delete, light, and water level controls.

strated that the hands should always be in front of the user so that they can be seen by the Leap sensor. After this, we started the tracking systems and again smoothed the surface. We assisted in putting on the VR headset and made sure that hand tracking worked as expected. While wearing the headset, we asked to try to form the sand surface building hills and valleys in order to get a better feeling for the system. After up to five minutes, we introduced the virtual controls next to the sandbox.

Users were now asked to create an environment of their liking, again introducing the different tools and interaction possibilities. The task was to at least create one hill or mountain, one water body and one lowland area. Further, users should add a lighting mood of their choice and place some virtual objects into the scene to extend their creation. After the scene had been created, we took various screenshots of the result from the following perspectives: exterior view in VR, exterior view of the VRBox, and inside view in VR (could be multiple). In the next step, users filled in a System Usability Scale questionnaire (Brooke et al., 1996) and subsequently performed a semi-structured interview as the last part of the study.

### Interview

For this thesis, we were mainly interested in the creative expression, playful, and natural interaction capabilities of the system. Further, we were interested in the perception of

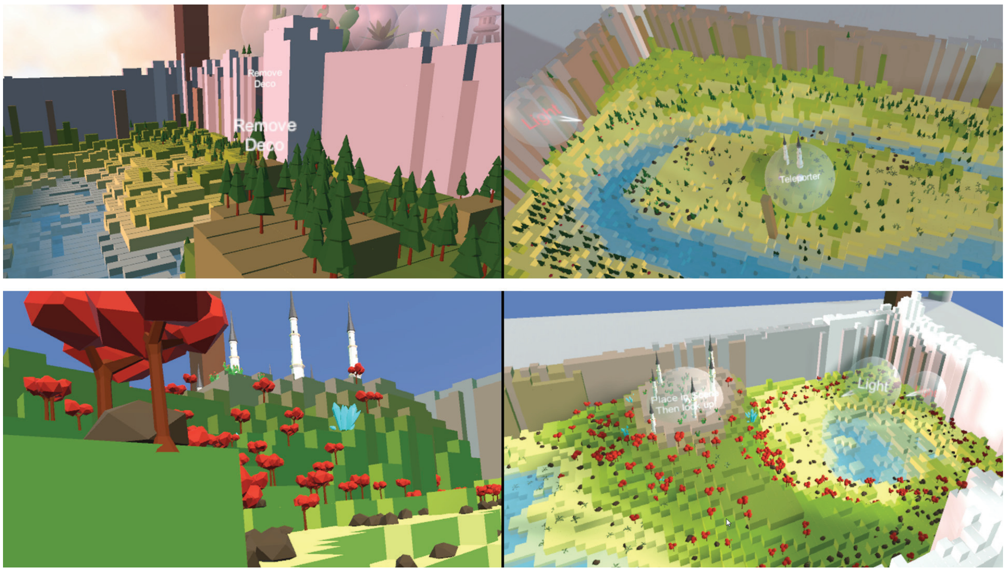


Figure 3.12: Two resulting 3D worlds (top row and bottom row) from our user study. On the left are the first-person views and on the right the tabletop views of the respective models.

sand as material, to what extend VRBox is a creative platform, and if they thought that VRBox supported creativity in a playful way.

We further asked what application scenarios users could envision for VRBox, how it could be used as a game element or core game mechanic or another application, which kinds of creative use cases users perceived as possible, and if they could think of using VRBox as an educational tool. We were interested if users could think of using VRBox in a productive context (e.g. planning tool or design tool for game worlds), how multiple users could interact with the system, and also asked for general feedback of what was especially positive or negative about interacting with our system.

### 3.2.3 Results

We first collected the answers and distributed them to the topics playfulness, creativity, and application scenarios. We further found interesting comments on the material and created another category for this. Then we summarized similar answers in order to reduce the amount of statements. In the following, we present the results along these categories.

**Playfulness and Creativity** All nine participants agreed that VRBox playfully supports creativity. P1 was ambivalent if a beautified or plain model supports imagination and thus creativity better. Simultaneous modeling in the physical and virtual medium is a creative factor for P2 and P8 because it provides power and supports playful experimentation. VRBox enables art creation by facilitating manual work without peripherals, and VR adds unlimited possibilities to extend creations (P3). The model creation is translated to something more using visuals and applying force increases the liveliness and fun (P4). Similar models could be created in 3D tools but VRBox is more interactive, fun, and playful. P5 said that a lot of things can be created and there is no pre-determination of results. Users can decide what world they want to be in as a major factor for creativity. VRBox would be useful for kids and younger people as modeling is a form of creative expression (P6). The spatial perception is supported by modeling landscape (P7).

VRBox offers more creative results than standard applications and sand interaction is something more than building a landscape block by block (P1). P2 states an increased freedom using the own hands and no peripherals, a good connection to the task without learning an interface, and the very low learning curve. In line, P8 mentions the direct access to the many parameters of the model, and the benefit of direct connection between action and result. VRBox is faster than traditional modeling and useful for non-technical people like artists compared to Blender and the like, enabling 3D content creation for everyone (P3). P4 was unsure if VRBox is more fun or if VRBox stimulates creativity. P5 was unsure if VRBox is more creative because of VR itself, and task dependency. P7 highlights creativity because of immersion and the many possibilities. P9 denied creativity because of the limitations of what can be modeled with sand.

P1 stated a disconnection between tactile and visual resolution and an indirect visual perception which makes VRBox not playful. Working with the hands made P2 feel like in childhood days, but the participant also stated that the menu interaction disconnected her from the immersion. Teleporting was playful for P4 while modeling was particularly nice for P5. Object placement was very good to bring the world to life. P6 was unsure about playful characteristics and stated, more detail would be better and more interaction could be added. Decoration and the interactivity using the own body was playful for P7 and creating an own scene was a huge attraction.

**Application Scenarios** All participants suggested either prototyping, landscape modeling, or city planning as main application scenarios. The system could also be used as an installation in museums or festivals (P6). VRBox could be integrated in the content creation pipeline in game design (P1), for idea generation, architecture visualization, and presentation in general (P4). P1 states that VRBox is not ready for production due to the low accuracy, and P6 saw it more as a toy than a tool.

For games, VRBox could be used in simulations and planning games (P5, P6), physics simulations, table tops (P5), exergames (using the weight of the sand, P1), tower defense (P1, P9), or god games (like Black & White (Peter Molyneux, 2001), P1, P5). Multi user scenarios include co-op and player vs. player games (P1, P2, P3, P4, P5, P6, P8) with synchronous playing using multiple HMDs or distributed playing (external devices). Co-op was suggested with one user modeling and the other acting on the game play (P1, P6, P8) or giving instructions from outside (P7). Here, co-presence visualization would be needed (P5, P6). City planning simulations, landscape shaping (P2, P8), game world and map creation, 3D paintings, and zen garden games have further been proposed (P1).

Furthermore, participants mentioned general application in education (P4, P6, P7), weather and climate simulation (P1, P2, P9), city transport simulation, agriculture planning (P1, P2), urban and rural infrastructure (P1), terrain modeling (P2), tech and science teaching (P2, P3), and 3D modeling (P3). Also, teaching physical and mechanical principles (P4, P8), creativity support in education (P6), art classes (P6), experiments and creativity (P2), relaxation and mind wandering (P1) were mentioned.

**Sand as Material** All participants agreed that for the use case of terraforming sand is an appropriate material, especially because it is well-known from playing with (regular) sandboxes (P1). P2 and P6 found it too dirty. P3 mentioned, that sharp edges were difficult to create with sand, and he recommended using a denser material. P6 suggested to use kinetic sand, as it brings more robustness to the forms. But he also admitted that it is costly. P8 added that clay could be added for models that are impossible to form with sand.

### 3.2.4 Discussion and Conclusion

#### Playfulness and Creativity

We found different playful aspects in our study. Working with the sand as in childhood days was mentioned several times, the teleporting feature, object placement, bringing the world to life, and that using the own body were perceived playful. For one participant the disconnection between visual and tactile made VRBox less playful. As there are mainly positive attributions for playfulness, we assess VRBox as a playful environment. Regarding creativity, all participants agreed that VRBox supports creative expression. The reasons were different: beautification, modeling of sand, endless possibilities, no predetermination, expressiveness, freedom, and low learning curve. We collected a lot of positive feedback regarding the creative capabilities and assess that with VRBox, users feel creative.

#### Visualization and Visual Abstraction of the Sand Surface

On the positive side, the block world look was generally appreciated, it was perceived as colorful, the level of abstraction was well accepted, “fits well”, was pleasant, did not interfere with the immersion and provided a good abstraction of the sand. It was further stated that a visual approximation of the details can lead to a more free expression compared to very high detail. However, participants also mentioned that a higher resolution would be better. From this we conclude that the overall block style is suitable for this kind of use case and people tend to like it, also because due to Minecraft, they are already used to it. However, more detail could be added by increasing the resolution of the blocks, keeping the overall style while increasing the detail.

### Limitations

In general, all participants showed a positive attitude towards VRBox. However, the sample size and composition of our participants needs to be considered when interpreting our findings. With only one female participant, there is a high gender bias in our study results, which needs to be considered in future work. Further, most participants are working in computer science related fields and have prior experience with VR, having a natural tendency towards the acceptance of new technology. P8, which is the only non-technical participant with an educational background was more critical about the setup and rated the system much lower compared to other participants. This is a finding which requires further investigation in future studies, because with VRBox we aim to support explorative learning, which should center around providing a positive user experience and usability for novice users in order to support the learning experience. Thus, the selection of participants should focus more on those users and also on users of the respective target groups, depending on the implementation (in our case for example students and teachers for the terraforming use case). One limitation regarding the interaction is a missing locomotion feature. Having the option to move through the virtual environment from a first-person perspective would increase the immersive and explorative capabilities of the system even more and will be integrated for future studies.

### Conclusions and Future Work

We presented the setup and implementation of VRBox, an augmented real sandbox that is used in a VR environment and combined with gesture interaction. With this prototype, we explored the integration of rich haptic interaction with natural materials into virtual reality and present a solution for tracking and presenting real physical materials in VR. We evaluated our system in a qualitative expert study with nine participants and found strong support for our assumption that VR increases the immersive and exploratory aspects of augmented sandboxes, leading to high levels of self-perceived creativity and playfulness. We further contribute application scenarios and possible use cases in gaming and 3D modeling.

We conclude that VR is a very beneficial addition to augmented sandboxes and supports immersion, participation, and control. Using free hand interaction, users can rely on their natural ways of expression without having to interact through hardware. There is a discrepancy between the tactile sensation and the visual appearance, especially because of the blocky world style. However, as our results show, users are very positive regarding this aspect and can adapt well to a little offset between what they feel and what they see. Further, the creative capabilities of our system have been highly appreciated, making it a useful tool for different kinds of tasks like modeling and gaming.

## 3.3 Dynamic Limitations in Cooperative 3D Scene Design

In any form of expression like singing, drawing, modeling, the personal creativity of the artist is one important factor for producing novel, surprising, beautiful, interesting and convincing results. Personal creativity is associated with a large variety of factors such as divergent thinking, extraversion, empathy, intrinsic motivation, intelligence and even age (Costa et al., 2015). Often, creativity is required to solve problems of any kind, for example when creating a new product, or coming up with ideas for a piece of writing. These kind of challenges are even harder to overcome if the problems are not well defined and thus hard to tackle because of the manifold ways of approaching a solution. One kind of support for these problem-solving challenges are creativity techniques. There is a large variety of different techniques that can be conducted in a group or individually. Some of the most well-known techniques are brainstorming, mind-maps, or the "Six Thinking Hats" of De Bono (2017). These techniques introduce structure to often unstructured problems, helping to generate a variety of ideas, also known as the concept of divergent thinking (Guildford, 1967). The core of divergent thinking is to create many solutions to a given problem, assuming that some of these ideas may be correct or can further lead to a correct solution. Many creative techniques center around the concept of divergent thinking, such as the 6-3-5 method (Rohrbach, 1969) or morphological analysis technique by Ritchey (1998).

While the idea of divergent thinking creates freedom and many possibilities to choose from, there are also techniques that limit artists in a specific way in order to be more creative and generate alternative ideas. One of these is known as "creative limitation". In contrast to divergent thinking methods, limitations work under the assumption that when the mind is forced to obey to certain rules, that at first glance limit the possibilities of action, new creative ways of overcoming the limitations arise and potentially generate creative problem solving. Examples of this can be found very prominently in poetry where writers limit themselves to a certain verse form or impose some sort of pattern, also known as constrained writing. In literature, this form of creative limitation is so popular that in 1960 a group of authors formed called "Oulipo", which is a short form of the french expression "L' Ouvroir de Littérature Potentielle" and roughly translates to workshop for potential literature[3]. The goal of this group is to create literature that expands the language through formally limiting themselves. A famous example is Georges Perec's novel "La Disparition" that is completely written without the letter "e" (Perec, 2001). This kind of limitation is set before the creative starts and the rules are clear throughout the artistic process.

However, there is also the possibility to introduce limitations as the creative process is ongoing. In children's play, rules are often introduced dynamically during play. For example when walking along a sidewalk, children often play a game where suddenly it is forbidden to step on the joints between the concrete slabs. This way, children have to find ways to bypass the joints and jump from slab to slab. Children make up rules during

---

[3] `https://www.oulipo.net`

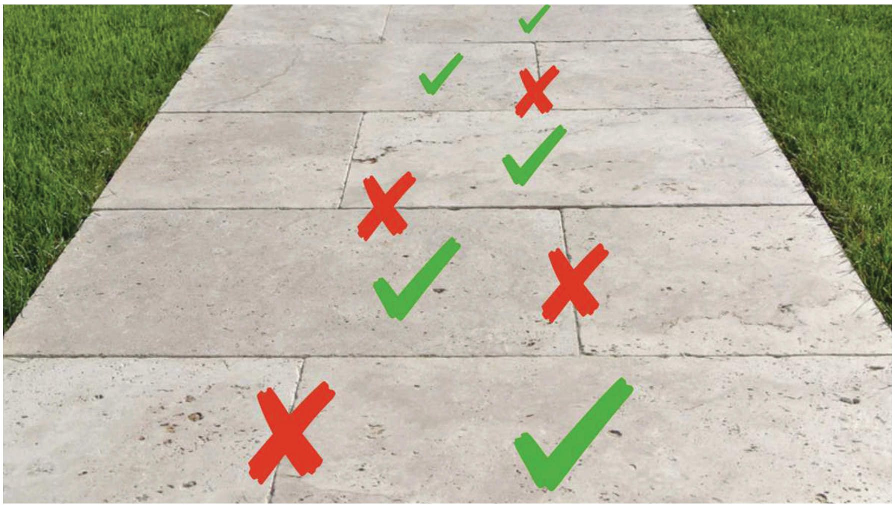


Figure 3.13: Example of a well known children's game with dynamic rules: Walking without stepping on the joints.

play all the time to make their play more engaging and fun. This kind of play naturally introduces a playful problem solving, resulting in challenge and fun. Another aspect of this kind of play is the interaction between the children. In our example, children impose dynamic rules upon the other player in order to push on the play. We believe that this cooperative factor is key to the overall game mechanic of dynamic rule introduction and would not be as engaging in a single-player scenario.

Inspired by these creativity techniques and dynamic aspects of creative problem solving, we were interested if and how creative limitations can be applied to other contexts and user groups. We concentrate on user-generated 3D content, comparable to the 3D shooter study, where we investigated how to support novice users in generating own 3D content, see Section 3.1. In this scenario, we expand the idea of supporting non-technical users in creating novel and creative 3D content by incorporating dynamic limitations to a 3D world building task in the game Minecraft (Persson and Bergensten, 2009), where we manipulate certain game mechanics (such as flying or block color) to introduce creative limitations in the game play.

Based on this idea and motivation, a prototype has been developed and evaluated in the Bachelor Thesis of Jan Tilger that the author has supervised, collaborated in, and provided the idea for (Tilger, 2018). Our main research questions are the following:

**R1** Can dynamic limitations benefit self-perceived creativity and lead to more creative results?

**R2** Are there differences based on the direction of the limitation induction (introducing/reducing)?

Figure 3.14: King's landing - The capital of the fictional world Westeros, completely built in Minecraft. Image source: Screenshot from `http://mc.westeroscraft.com/`

**R3** How do users perceive dynamically induced limitations as a technique for creativity support?

For our study, we chose Minecraft because it's the ideal playground for easy and playful 3D model generation. With Minecraft, users are able to implement their own ideas for 3D worlds in a very easy and playful way. Around this game, a large community formed in recent years, creating astonishing models. Many examples can be found on Reddit in the subreddit "Minecraft"[4]. For example, the whole "Game of Thrones" world Westeros has been modeled in Minecraft by the community WesterosCraft[5]. Other examples of creative use include computer programming using in-game blocks such as redstone. This ultimately resulted in the creation of the Minecraft Education Edition. This version of the game can be used in classrooms for collaborative education in programming, history, art, design, math and other subjects[6].

We modified Minecraft and conducted a user study where participants collaboratively constructed a 3D block-world. We were interested in finding effects of creative limitations on self-perceived creativity and the creative self-assessment of the result (i.e. how users rate their own creation). Thus, we investigated creative limitations in two different ways. First, by taking away certain game mechanics while playing the game (dynamically

---

[4] `https://www.reddit.com/r/minecraft`

[5] `https://westeroscraft.com`

[6] `https://education.minecraft.net`

introducing limitations). And second by starting with very limited set of game mechanics and reducing the limitations one after another (dynamically removing limitations).

### 3.3.1 Prototype

We chose Minecraft as our base software because its well-known for its creative mode where players are free to create 3D worlds without the need for story or game play and the game acting as a sandbox and also to circumvent unnecessary distractions like health, hunger, enemies, etc. The dynamic introduction of limitations was implemented by changing some of the game mechanics. However, in contrast to the adventure or survival mode of Minecraft, the creative mode of the game consists only of a little amount of game mechanics: adding and deleting blocks, workbench, and locomotion. This made the selection of possible limitation mechanics slightly more challenging.

#### Altering game mechanics as limitations

In the following, we present a list of our implemented mechanics and a detailed description of how these were dynamically activated.

> **Flying:** Flying is not supported in the standard game mode, but is available in the creative mode. Using the flying capability, users can fly with their avatar so that they can build more freely. Usually, in standard game mode, support structures have to be built in order to reach similar destinations. Flying works as a limitation when first having the option and then it being removed.
>
> **Mid-air:** We modified the building mechanic by including the ability to place blocks in mid-air. Normally, blocks can only be built next to other blocks or the ground (which is also made of blocks). Similar to flying, a limitation occurs when the mechanic is removed after having it used prior to the removal. This mechanic will be referenced as “mid-air”, subsequently.
>
> **Neighbor:** Only allows to build blocks next to blocks that have been placed by the other player. This mechanic presents a limitation when introduced, because without it, users are free to build wherever they want.
>
> **Variety:** Limits the block inventory. Without limitation, users can build many different block types, whereas with the limitation, only two types of blocks can be built: black and white.

In the case that limitations are dynamically introduced, the mechanics are configured in the following way: flying enabled, mid-air enabled, neighbor disabled, variety large. When starting with the reduced set, the mechanics are configured: flying disabled, mid-air disabled, neighbor enabled, variety little. The timing of the alterations is fixed. Every 130 seconds, one random mechanic is altered.

Having presented the limitation mechanics, in the next section, we outline how these limitations were implemented so that they can be introduced in a dynamic fashion and how they were used in the study.

### 3.3.2 Study Design and Procedure

We conducted a between-subject lab study with 15 groups, each consisting of 2 participants. During the user study, each team either played the introducing or reducing condition. There are two tasks that each team had to perform.

In the first task, each team had to recreate a set for the classical fairy tale "Hansel and Gretel" by the Brothers Grimm. We chose this kind of task because almost everyone is familiar with the setting and story, while at the same time leaving enough room for creative expression in the task. The specific scene that the participants had to recreate is the one where the witch is attempting to grill Hansel and Gretel in front of her house in the wood stove. Teams were randomly assigned to play either the introducing or reducing condition.

In the second task, each team performed a free building phase that lasted the same amount of time. Here, participants had the fully functional game mechanics available in order to create a scene of their own liking. This phase was introduced because we wanted to compare the self-perceived creativity rating to the provided main task in order to find out how ratings differ depending on the task type.

For counter-balancing the results regarding our limitation strategy, we decided for a between-group design where each group only played one condition and the comparative analysis was performed between the two groups.

#### Procedure

After welcoming and a short introduction, participants were presented consent forms and had to sign both this and the data use agreement. In the next step, demographic data was collected, including age, gender, occupation, motivation for Minecraft and experience with the game, gaming and creative habits, media consumption, and overall initial self-perceived creativity. In the next step, participants had the chance to become familiar with Minecraft and the game mechanics. There was no time limitation for this part, but usually this would not take more than 15 minutes per group.

Next, participants were introduced to navigate to a prepared theater stage where the replication task of the study was conducted. When both participants were ready, the task was revealed and a short introduction to the Hansel and Gretel scenario was verbally presented. After both had good understanding of what is expected, the main trial was started and both participants cooperatively performed the task. They also were informed if they belonged to the introducing or removing group. Both computer screens were recorded for post-analysis. Each time a limitation was introduced or removed, a message was displayed on each of the participant's screen. Also, the experimenter read out loud the same text, to make sure both recognized the change in the game.

After completing the first trial, both participants filled in the second part of the questionnaire. These are the questions we asked in this part and had to be answered on a 1-7 Likert scale.

- How creative do you rate your result?
- How satisfied are you with your result?
- Did the cooperative aspect of playing with a partner limit you, or helped in reaching your goal?
- How creative did you feel while playing?
- How useful do you rate the limitations for building more creative objects? (For each mechanic)
- Do you rate the limitations as useful in this scenario for building more creative objects?
- How creative do you feel after this phase?

After participants filled in the questionnaire, the second experimental phase was started. Within 12 minutes, participants were free to design whatever they wanted. Before the actual building phase started, participants had time to discuss what they wanted to build. There was no time limit for this, however, most participants only needed around 1-5 minutes for brainstorming. After completing this phase, the third and last part of the questionnaire had to be filled in, which was equivalent to the former part (see questions above). Again, video was recorded for post-analysis. A subsequent interview was performed where we were interested in the following aspects:

- Overall assessment of Minecraft for this kind of task
- Assessment of playing the modified version of Minecraft
- Especially helpful and disturbing methods (limitations) during the study
- Impact of the limitations on the gaming experience
- Perception of gaining/losing capabilities during the game (dynamic aspect)
- Overall perception of cooperative aspects of the game
- Other comments

### 3.3.3 Results

We present quantitative data that we collected with post-trial questionnaire and qualitative data from the post-trial interviews.

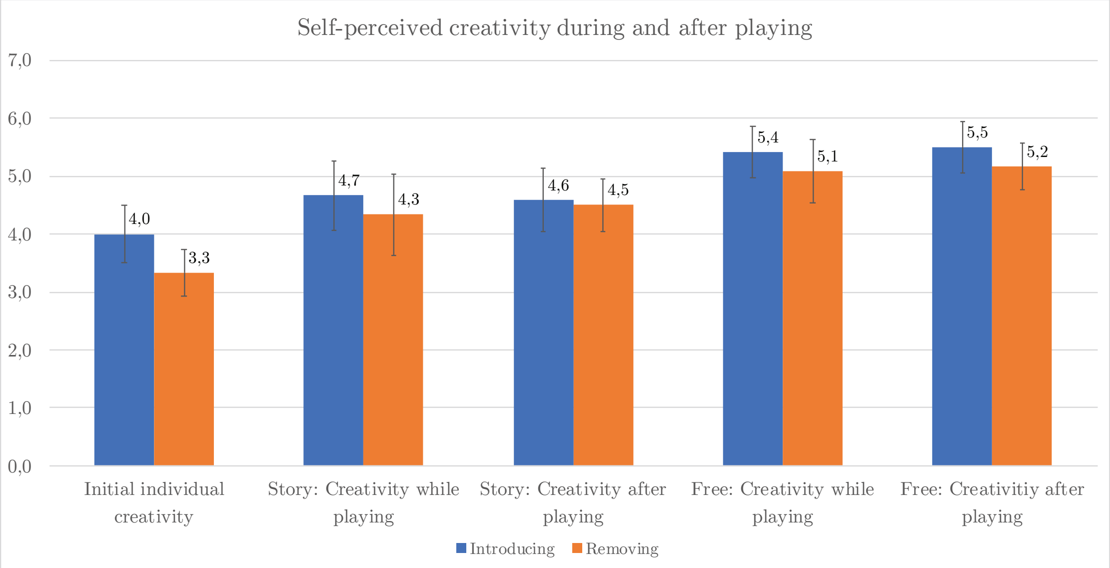


Figure 3.15: Comparison of creativity ratings: initial creativity, creativity while playing, and creativity after playing.

### Questionnaire Results

**Results "introducing" group** A paired Wilcoxon test revealed a significant effect between initial creativity ($M = 4, SD = 1.04$) and creativity after the free building phase ($M = 5.5, SD = 0.90$) with $p < 0.005$. Perceived creativity while building differed between story phase ($M = 4.67, SD = 1.3$) and free phase ($M = 5.42, SD = 0.9$) with $p < 0.05$. Perceived creativity after building differed between story phase ($M = 4.58, SD = 1.16$) and free phase ($M = 5.5, SD = 0.90$) with $p < 0.05$. (See Figure 3.15)

Satisfaction with the result differed significantly between the story phase ($M = 4.83, SD = 1.34$) and free building phase ($M = 6.08, SD = 0.99$) with $p < 0.005$. (See Figure 3.16)

**Results "removing" group** Initial creativity ($M = 3.33, SD = 0.89$) and creativity while story phase ($M = 4.33, SD = 1.5$) differed with $p < 0.05$. Also, initial creativity and creativity after story phase ($M = 4.5, SD = 0.9$) differed with $p < 0.05$. The same is true for initial creativity and creativity during free phase ($M = 5.1, SD = 1.16$) and after free phase ($M = 5.17, SD = 0.84$) with $p < 0.05$ and $p < 0.006$, respectively, see Figure 3.15.

**Between group analysis** We found no significant effects between the groups.

### Interview Results

Interviews have been performed group-wise for all 15 groups, with the exception for group 4 where no interview was performed because of time constraints of the par-

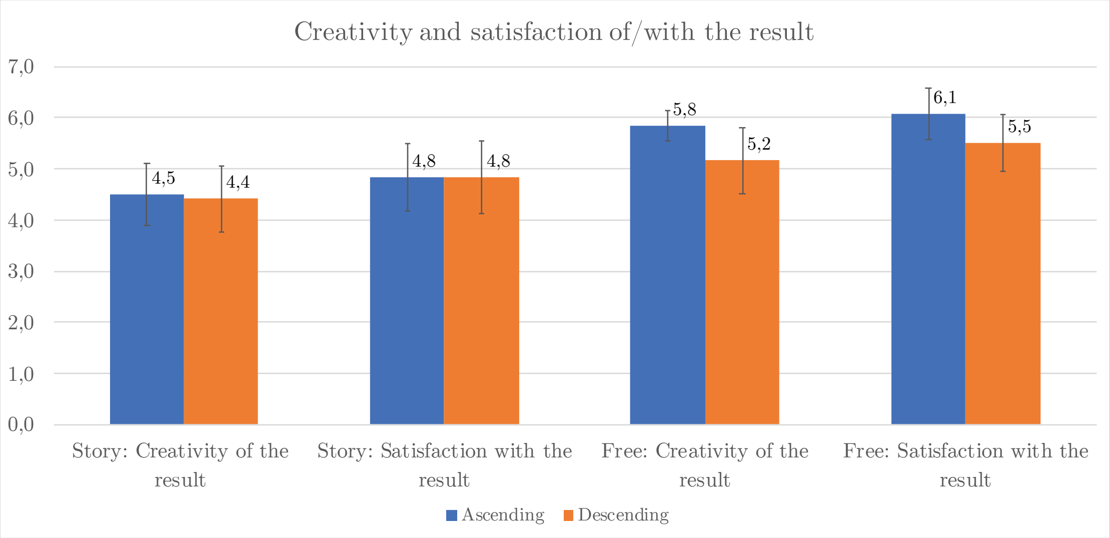


Figure 3.16: Comparison of satisfaction with the result and creativity of the result after story and free phase.

ticipants. Uneven-numbered groups belonged to the “introducing”-type, while even-numbered groups received the “removing” treatment.

**Playing the modified version** Positively mentioned were the switching of plans on the fly (3,15), the feeling of creativity (9,10), the time factor and overall challenge (10,11,13), game flow (8), and generally liked the approach (2,3,6,7,10). Neutral comments mentioned an overall positive experience after getting used to the manipulation (3,5), that starting with the reduced set was easier for increased focus (12,14), and that one is forced to think outside of the box. On the downside, it was mentioned that participants missed certain block types and building options when the set was reduced (1,13,14), that the game flow was somewhat disturbed by the dynamic changes (9). One participant mentioned that is was bad to work with a reduced set when knowing of all the possibilities of the full game (14), while 8 mentioned no improved feeling of creativity.

**Especially helpful mechanics** Flying was perceived as most helpful for all groups apart from 7, followed by placing blocks mid-air (2,3,5,6,12), direct block exchange (10,13,15), and the limited (5,12) and increased inventory (13,14).

**Especially disturbing mechanics** Most mentioned the limited inventory, specifically when only black and white blocks were available (1,5,7,8,9,11,14). Further, the block-on-block limitation was perceived as disturbing by groups 3 and 12 and the lack of flying (9,14).

**Impact of the limitations on the gaming experience** There were various comments on the overall evaluation of the gaming experience under the imposed limitations. On the positive side, participants stated that Minecraft was more fun because of the additional challenge (1,6,11,13) and due to introducing limitations in the first place (2). Some felt more excited (3,10) because of personal adaptation (3) and perceived increased fun due to time challenge (3,10,12) especially for coming up with novel creations (12). The coop aspect increased fun for groups 5,9,12,15 with coop having more impact than limitations for group 9. The gaming experience was refreshing for creating new designs (12) and for 5 and 6 no more nor less fun, however the personal adaptation was still challenging (5). Group 10 stated that limitations don't make for more beautiful creations but for creating new ideas and models. Group 15 perceived them as neutral. For them, the task played an important role and the setting on the stage helped.

On the negative side, fun was decreased when participants were not able to conduct the pre-defined plan because of the limitations (2), while others felt frustrated because of the limitations in the beginning, however with increased possibilities, this got better (7,8). Positively mentioned was being forced to solve problems creatively (7). Group 8 stated that limitations were annoying but also challenging, while 14 mentioned that limitations are limiting when the full set of mechanics is already known, and the time limit for new players was annoying (14).

**Perception of introducing/reducing mechanics during the game (dynamic aspect)** Game play became faster and more agile for groups 6 and 8 (little impact, just more effort) and reducing limitations was appreciated by group 12 (even very easy in the end). Group 14 could not execute their plan and pointed out the necessity to cooperate in order to be successful. Group 2 switched tactics back and forth and discussed how to progress under the evolving situation. Groups with "introducing" game play simply altered their goals (1, color limitation) or couldn't finish their plan (5), were forced to rethink with increasing limitations (7), found color limitation too difficult and mostly stuck with their initial plan (11). Flying was a crucial limitation for group 13, resulting in a knock out with the color limitation. Group 15 compensated for lack of flying by building ladders and adapted their building style and when the plan couldn't be fulfilled they started fooling around and abandoned their goals.

### 3.3.4 Discussion and Conclusion

We formulated the following research questions:

- **R1** Can dynamic limitations benefit self-perceived creativity and lead to more creative results?
- **R2** Are there differences based on the direction of the limitation induction (introducing/removing)?
- **R3** How do users perceive dynamically induced limitations as a technique for creativity support?

Our results demonstrate significant effects regarding the self-perceived creativity. The introducing group showed increased self-perceived creativity at the end of the study, compared to the initial creativity. Perceived creativity was also greater after the free

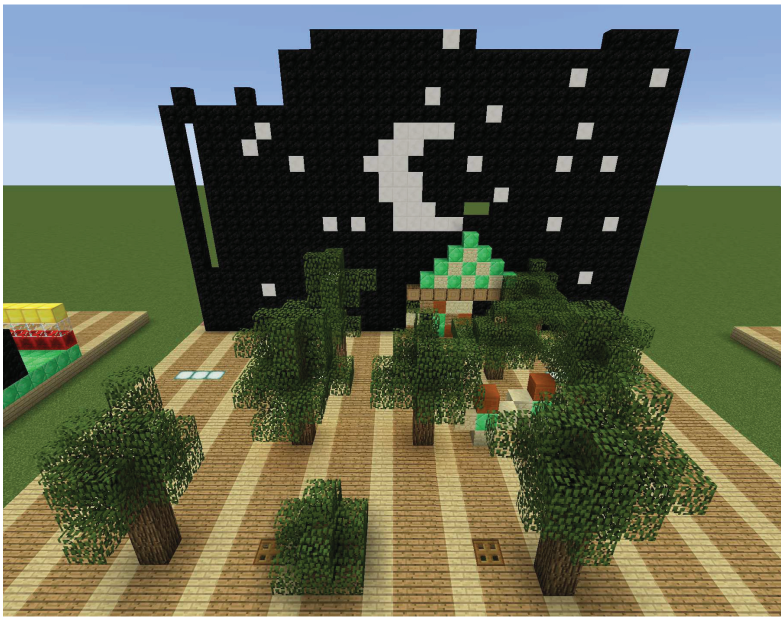


Figure 3.17: Work of one “introducing” group. After having only black/white blocks left, a night background with stars and moon was created for the Hansel and Gretel scene.

phase both while and after execution. This effect is consistent in the removing group, where participants felt more creative during and while both story and free phase. There are more significant effects regarding self-perceived creativity in the removing group compared to the introducing group. However, looking at the average initial creativity of both groups, we noticed that group “removing” rated themselves significantly less creative than group “introducing”. Putting this into context, we assume that the dominance of these effects can partly be attributed to this factor. However, as there are still significant effects in the “introducing” group that also can be found in the other group, we assess this result as a stable effect between the initial and post-self-perceived creativity measure. Considering the direction, meaning introducing and removing of limitations, we cannot clearly identify reliable indications of which of these strategies worked better. This is supported by the non-existing effects between the two groups.

It can be argued that the increase of self-perceived creativity in comparison to the initial creativity assessment is due to solely playing Minecraft, without any impact from the limitations. However, there are within-group effects in group “introducing” where self-perceived creativity during and after both free play and story phase is higher in the free condition. In this condition, no limitations were present, which could mean that limitations had a rather negative effect on self-perceived creativity. Another factor that could contribute to this is the task itself where participants were not able to express themselves as creatively as in the free condition.

Figure 3.18: Scene of a “removing” group. Starting off with full limitations, major parts of the scene were created with black/white blocks. After having colored blocks, a new oven was built.

To sum up, **R1** cannot be answered positively with full certainty, as the collected data does not clearly indicate a positive effect of limitations on self-perceived creativity. In order to answer this question, more research has to be carried out. Specifically, a study design that clearly separates limitations from a clearly defined baseline condition. Also, the cooperative aspect of our study is prone to generate noise that should be eliminated in further research. However, we collected promising and interesting results that indicate the potential of the approach. As we found in the interviews, a larger amount of participants appreciated the limitations as a fun addition to the game, pointed out that it helped to create more designs (if not better), and forced them to rethink their approach. The latter is a very promising result as lateral creativity is a key to creating more and possibly more creative results.

Regarding **R2**, we found that both conditions increased self-perceived creativity in similar ways, there seems to be no difference between the groups that indicates a clear preference for either of the strategies.

On the basis of the post-hoc interviews, we conclude that our participants were mainly positive regarding the concept of dynamic limitations (**R3**). It was stated that the game was more fun and exciting. It has to be mentioned that the coop-aspect of the game certainly played a role as well. However, specific feedback on the idea of introducing limitations reveals a positive attitude towards our idea. We highlight that some groups were not able to create what they had in mind because of missing mechanics. This can be discussed controversially, as on the one hand, this is exactly what we wanted to achieve with the idea of creating hurdles: that users have to adapt in order to progress. On the other hand, if users complain about missing mechanics when wanting to create somethings specific, it can be seen negatively as our approach is limiting users in some ways.

We conclude that there is a chance that dynamic limitations can positively influence perceived creativity, also supported by the interview results. However, our approach can be viewed as a “double-edged sword” as some users may feel restricted in completing an idea worth pursuing. A final judgment if this is good or bad depends on the situation. If the goal is to support lateral thinking, it may be beneficial, if it is longitudinal, there might be issues.

# 4 Natural and Collaborative Interaction

Natural user interfaces provide familiar, intuitive, and easy to use interaction, employing real-world metaphors such as reality-based interaction, gesture, body, speech input and in general a high degree of freedom. Because of this, NUIs are ideal candidates for the creation of 3D content such as previsualizations, animations, models, and more. In this chapter, we introduce three different systems that represent natural interaction in the context of 3D content creation.

First, we take on the question of how to combine the input methods that VR systems and tablet computers offer. Both systems have advantages based on how the hardware is designed. However, there are downsides to each of the input methods. For example, VR excludes users from the real world and there are problems with mode switches inside a work context. Tablets have rather small screens, but can be shared among multiple users and are better suited for quick content interaction, which is a little more complicated in VR (see also Chapter 2.2.3). The second system presents VR interaction for previs in the film, animation, and stage context that has been developed in the first.stage project. Further, we show how to use embodied interaction for the creation of character animations. Lastly, we introduce a rapid VR modeling application where users can prototype 3D models without the use of any graphical user interface. Here, only gesture interaction is used for model creation and system interaction.

## 4.1 Collaborative Scene Design across Virtual Reality and Tablet Devices

The creation of 3D content such as virtual worlds and models can be considered as being natural in Virtual Reality because of the following reasons. For example, VR adds immersive visualization and easy manipulation of 3D content. Through the use of head-mounted displays (HMDs) users can perceive 3D content with believable depth sensation, being able to walk around and look around virtual objects. Further, room-scale tracking allows for natural movement (in a restricted area) and object manipulation using tracked controllers in a natural and direct way. VR controllers are used as extensions of the own body, employing intuitive ways of interaction by grabbing, holding, rotating and placing objects in VR. Further, the increased immersion and understanding of scale and distances is especially useful for 3D design tasks such as scene or world design. There are already a large number of VR tools available on the market that prove the useful nature of VR for creative tasks that center around creation and building like TiltBrush[1] and GravitySketch[2].

However, in situations where multiple people work together, limitations of VR arise. In co-located design, multiple people work together in a shared environment, for example in an office. Such environments establish a social context between the users that allows for direct interaction due to the physical presence of the participants. In such settings, VR users wearing a HMD can be disconnected from the rest of a group due to physical isolation. This can result in a limited participation, less social interaction, and a decreased context awareness of bystanding users (Gugenheimer, Stemasov, et al., 2017; Gugenheimer, Mai, et al., 2019). A further disadvantage arises when other users want to actively explore the design in VR or even operate the system themselves. In these cases, users have to switch places, put on the HMD and pick up the controllers to participate, which can be a cumbersome process.

Another aspect of VR usage is system interaction such as text input or mode switching. It is often the case that VR applications work in unforeseeable ways for non-technical users, because interaction techniques can greatly differ between implementations and have to be learned on an app-basis. System interaction in VR is an ongoing research topic in HCI (Speicher et al., 2018; Surale et al., 2019) and it is still not fully known how these manipulations can best be implemented. In contrast, these aspects have mostly been solved for tablet interaction where graphical user interfaces can be controlled with touch gestures. Gesture interaction on tablets and other touch-devices nowadays is an established way of operation and similar techniques can be found across devices and operating systems (e.g. pinch and pan gestures), so that users employ these gestures almost naturally. Further, regarding co-located work, tablet computers can be shared among different users and establish a common ground for collaboration as the device can be viewed by multiple people at once and being handed over to someone else.

---

[1] `https://www.tiltbrush.com/`

[2] `https://www.gravitysketch.com/`

Figure 4.1: Users collaboratively creating a scene in VR and using a tablet.

Regarding these issues and opportunities, the aim for this research is to combine the advantages of VR and tablet interaction by building a system where co-located users can collaborate using both VR and tablet hardware in a shared environment. The aim of this research is to create a system where users can employ their natural ways of interaction in the easiest way by using common and well-known metaphors of both modalities and bring them together. Based on this motivation and idea, a prototype system has been developed and evaluated in the Bachelor thesis of Klug (2018) that the author has provided the idea for, supervised, and collaborated in.

Using the system, two users can collaborate simultaneously in VR and on a tablet device, as shown in Figure 4.1. We assume that depending on the devices, different roles and tasks are distributed between the users. For example, as tablets are suited for overview tasks, we assume that in a shared task, this is something that naturally falls in the category of a tablet use. Based on this idea, we believe that there are certain tasks that are being distributed by the users in a more or less natural way. By natural, we refer to the distribution of tasks depending on how efficient, common (depending on the device), and easy these are to perform on either of the devices.

For finding out which tasks users distribute to VR and tablet hardware, we employ an inductive research approach by collecting usage-data of the VR and tablet users. We observe user behavior, generating knowledge based on our observations. We assume that depending on the type of device, users employ different use patterns and strategies. In order to control for the device type and specific form of interaction, we implemented identical functions on both devices, considering best-practices regarding the interactions for each modality. Based on our motivation, we formulate the following research question: *How does collaboration between VR and tablet users influence task distribution and interaction in collaborative 3D scene design?*

### 4.1.1 Prototype

Our prototype is a distributed system that consists of a virtual reality system (HMD and computer) and a tablet computer. From a software perspective, both systems are integrated into a shared working environment, where the two users can work together on a shared task. One user is working with virtual reality interaction, wearing a HTC-Vive HMD and working with the Vive controllers. The other user is using an iPad Pro with a 12 inch display. Both devices offer the same interaction options so that task execution is device-independent. Moreover, we take advantage of device-specific interaction capabilities to maximize ease of use and usability for each user and modality. This means that VR interaction is tailored to match current interaction patterns like teleporting and direct manipulation. On the other hand, tablet interaction makes use of panning and pinching as these are natural ways of interaction for that kind of device. This way we intentionally prevent translation of interactions across devices. We rather make use of the device-specific ways of interaction that users are familiar with.

As both users work together in a shared environment, it should be possible to see the other user's actions directly on the own device, establishing a co-presence inside the virtual world. We implemented this co-presence by visualizing the VR user through displaying the HMD and controllers and the tablet user by an avatar. That way, both users are aware of the presence and position of each other in the virtual world.

#### Interaction

Regarding interaction, VR users use teleporting for locomotion and direct manipulation. In addition to teleporting, VR users are able to use a teleport shortcut that directly teleports the own presence next to the tablet user. This way, it is possible to examine a certain creation together from a shared perspective, a task that is common in collaborative design. Direct manipulation is supported by the Vive controllers where users can grab, rotate and translate objects and release them. In order to place new objects, a menu is provided and displayed next to the non-dominant hand controller in VR. Users are able to teleport while having an object grabbed for convenience. Assets and the whole scene have no collision and physics enabled for easy building. Manipulation of existing objects is done by grabbing, manipulating and releasing objects.

Tablet users employ gesture interaction for navigation and object manipulation. Longitudinal locomotion can be performed using a pinch-gesture as this kind of locomotion translates to "zooming" into a 2D picture, where pinching is the common gesture. Lateral movement is achieved by pan gesture with one finger. Accordingly to the pinch-gesture, this interaction is often employed for moving content on screen and a de facto standard in tablet interaction. Panning sideways orients the view to the left or right, implementing looking around in the scene. Teleporting to the VR-user is also possible for tablet users through a button press. Manipulating objects in the scene is done by double-tapping an object, which is highlighted to appear as active. Translating horizontally on the ground can be done by panning with one finger. Vertical two-finger panning allows for vertical translation and horizontal two-finger panning for rotation around the Y-axis. Tilting an object can be done using a three-finger panning gesture vertically (front and back) or horizontally (left and right). Double-tapping on an active object deactivates the object. While an object is active, tablet users cannot move around in the scene, because of the variety of manipulation options for active objects. Adding locomotion while manipulating objects was neglected because we assumed that it would confuse and burden users. However, a perspective change is possible, switching to an overview mode and back to the scene view.

As mentioned, a menu is employed in VR and on the tablet for asset browsing and selection. Both users can preview assets in the menu. Further, by selecting an asset from the library, a preview-object is spawned in the scene, allowing for in-place preview inside the real scene. In VR, assets follow the controller movement and users can explore assets from all angles by moving their hand. Selecting an asset with the other controller spawns a preview object for detailed inspection, and by grabbing the object, it can be placed in the scene. After placing the object, the menu is automatically closed. Tablet users open the menu by pressing the respective button on the GUI. It is displayed on the right side of the screen, so that it doesn't occlude the scene. Vertically swiping allows for switching between different asset panes. For easy inspection, all assets continuously rotate around the Y-axis. For creating a preview-object, tablet users can tap and hold on an asset. Releasing spawns the asset in the scene, swiping up while still tapping aborts the selection. When placing an object, the menu is closed automatically. An undo is both possible for VR and tablet users.

### Implementation

The implementation was done using Unity[3] in combination with the SteamVR-plugin[4]. Assets were used from the Poly-Toolkit from the Unity Asset Store. Additional assets were used from the Google Poly Website[5]. For networking, we used a client-server architecture on basis of Unity's High-Level networking API, extending the NetworkManager object. An interesting component of our system is the logging mechanism we implemented. As we employ an inductive approach for finding out user's preferences using either VR or tablet interaction, we collect usage data such as type of moved objects, distance of player movement, count of spawned objects, etc. We log every action that both players perform for post-analysis.

---

[3] `https://unity.com/de`

[4] `https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647`

[5] `https://poly.google.com/`

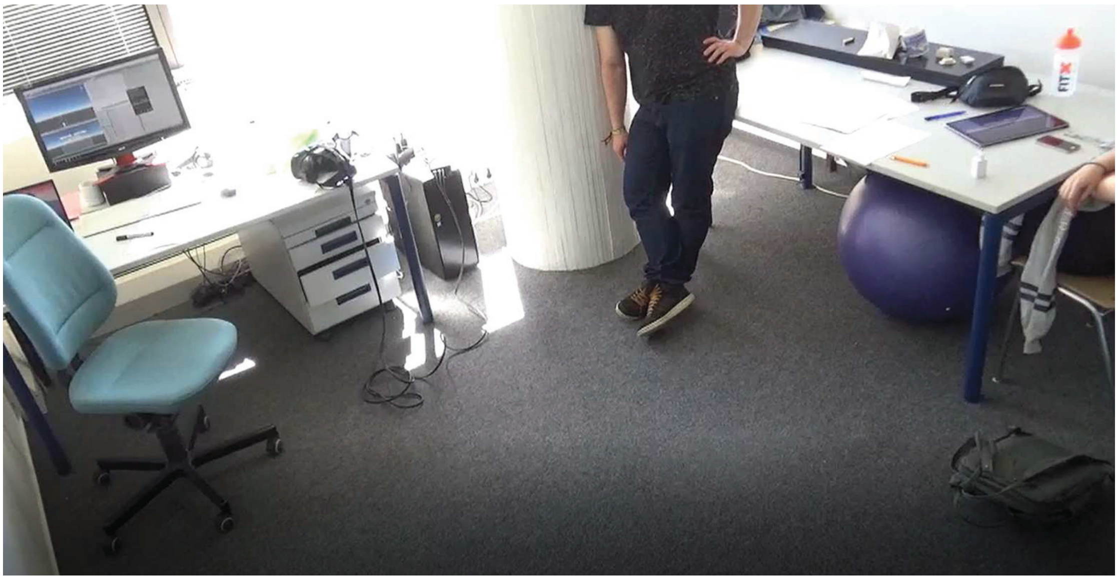


Figure 4.2: Setup of our study. Left: VR workstation with the HMD. Right: Tablet user sitting in front of the iPad Pro running the mobile version of the prototype.

### 4.1.2 Study Design and Procedure

We conducted a lab-study with 18 participants (9 teams) where we investigated the impact of the device-dependent interactions on user behavior. The overall task-setting is grounded in the tale of of Hansel and Gretel by the Grimm brothers. Specifically, users should create the scene where Hansel and Gretel are to be grilled in the wood stove near her house. We chose this setting because most people are familiar with the tale, establishing a common ground while at the same time the exact implementation of the scene is open for creative design.

#### Method

The main part consists of three different (but related) tasks. In Task 1, the witch house had to be placed and decorated according to the tale. The aim was to observe who of the participants is taking the lead. Further, we were interested in who takes the lead role in placing small objects, doing precise work and also if there is a collaborative (i.e. working together) or more cooperative (i.e. high task distribution with common goal) working style. Task 2 was about building a fence around the witches “property” and to build a front yard, including flowers, stones and an oven, which is most important for the tale. This tasks focuses on placing a lot of medium and small-sized objects in the scene. We were interested in finding differences in the overall task distribution and collaboration behavior, object-size preferences, and travel distance between the modalities. Task 3 was about creating a dense forest, with the aim to enforce placement and handling of larger objects.

Additionally to the inductive finding of natural patterns in task distribution and collaboration behavior, we were also interested in finding out how participants assess both prototypes in terms of usability and mental effort. For this, we used the SUS (Brooke et al., 1996) and NASA's raw TLX questionnaire (Hart and Staveland, 1988). In the post-interview, we asked the following three questions:

- What was particularly nice and bad about using the prototype?
- How do you rate the collaboration from your perspective?
- Would you like it better if both were using a tablet or if both were using VR?

**Procedure** The study was separated in four phases: Welcoming, introduction, main part and post-interview. In the welcoming phase, participants were informed about the aim of the study and were presented with the consent forms that they had to agree to and sign. After that, demographic data was collected: age, sex, experience with VR, tablet usage, and design of digital content. We also asked if participants already knew each other in order to better assess the collaboration behavior.

In the introduction, the concept of VR and the context of the study was introduced. Participants were randomly distributed to either condition (tablet or VR). Subsequently, participants were introduced to the prototype, and how to interact with the system. After this, there was a free phase of ca. 15 min where the team was able to test the functions of the prototype and ask questions. This phase differed in a way from the main part as here, the asset set was different compared to the main part. Figure 4.2 shows the experimental setup. On the left the workstation running the VR software, on the right a participant sitting in front of the tablet device before the beginning of the study.

In the following 30 minutes main part, the three tasks had to be completed by the team. After being introduced to the tail of Hansel and Gretel, which also included a picture of a sample setting, participants were reminded to create a harmonious scene design, the tasks were introduced one after another. After completing the main part of the study, participants filled in the SUS and raw TLX questionnaires, performed a concluding interview, which marks the end of the study.

### 4.1.3 Results

We welcomed 18 participants in 9 teams with a age-span of 21 to 29 years with an average age of 23 years. Of those were twelve participants male and six female, which distribute in five male teams, two female teams, and two mixed teams. Of the nine VR participants, five had experience in VR, and all tablet users were experienced using such devices.

We collected a wide variety of data for finding out how work and tasks are being managed and shared across users and devices. An overview and description of the variables are shown in Table 4.1.

| Variable Name | Description |
|---|---|
| Count Moved Objects | Total number of objects that were moved |
| Count Spawned Objects | Total number of newly spawned objects |
| Count Rotated Objects | Total number of rotates of all objects |
| Distance Moved Objects | Total distance of all moved objects |
| Count Menu Open | Total number of opening the menu |
| Time in menu | Total time spent having the menu open |
| Distance Player Movement | Total distance a player traveled in the scene |
| Amount Player Rotation | Accumulated degrees of player rotation |

Table 4.1: Overview of all tracked interaction data

### Numerical User Behavior

On basis of the collected data from both VR and tablet, we conducted Mann-Whitney U tests, comparing differences between the two modalities. In the following, only significant data is reported.

**Count of moved objects** This parameter describes the amount times that objects were moved for each modality. Object types are small, medium, and large. We found that in total for all object types combined, tablet users moved more objects (on average $123, SD = 68$ objects) than VR users (on average $58, SD = 10$ objects) with $p < 0.05$. Tablet users moved 1103 objects and VR users 519 objects accumulated.

Also, there is an effect for large objects in particular where tablet users moved more large objects (on average $26, SD = 18$) than VR users (on average $7, SD = 8$) with $p < 0.05$. Summing up, tablet users moved 235 large objects while VR users moved 63 large objects. Further, tablet users in total moved 316 (VR 212) small and 356 medium sized (VR 121) objects. Further details on average object move numbers are depicted in Figure 4.3.

**Number of spawned objects** This refers to the number of spawned objects per modality, divided into small, medium-sized, and large object types. We found no statistical effects for this parameter. In total, tablet players spawned 329 objects (147 small, 98 medium, 84 large), while VR users in total spawned 501 (272 small, 130 medium, 99 large) objects.

**Count of rotated objects** Tablet users rotated 796 (88,4 on average) objects in total while VR users rotated 726 (80,7 on average).

**Distance of moved objects** Describes the distance objects were moved through the scene. In general, tablet users moved all objects combined by 1742 units (519 small, 567 medium, 655 large) while VR users moved objects across 600 units (293 small, 218 medium, 89 large). Statistically, tablet users moved objects across longer distances (on average $194, SD = 140$) than VR users (on average $67, SD = 21$) with $p < 0.05$. We

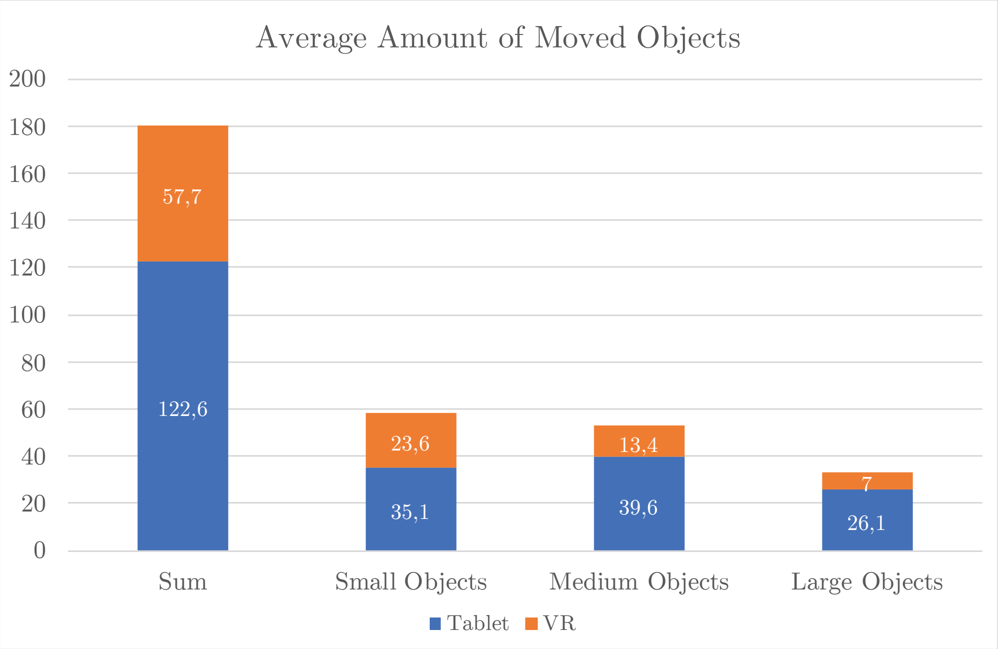


Figure 4.3: Average amount of moved objects by size and modality.

also found this to be true for large objects, where tablet users travel more (on average $73, SD = 54$) than VR users ($10, SD = 14$) with $p < 0.05$, see also Figure 4.4. There were no statistically significant effects for small and medium-sized objects.

**Menu interaction** For menu interaction, we collected data on a) count of menu openings, b) time spent in menu, and c) menu interactions such as spawn, open, close, preview). Menu openings were on average for tablet users 38 times ($SD = 18$) and VR users 57 times ($SD = 34$). Time in menu describes the time that tablet and VR users spent having the menu open. We found a close-to-alpha effect in this category. On average, VR users had the menu open for longer (on average $547s, SD = 144s$) than tablet users (on average $393s, SD = 120s$) with $p = 0.05031$. Regarding the amount of menu interactions (spawn, open, close, preview), we calculated that on average, tablet users interact menu 328 times ($SD = 154$), while VR users interact 419 times ($SD = 180$).

**Distance of player movement and rotation** Describes the distance units that users traveled. For clarification, tablet users had the pinch gesture for movement, VR users were able to teleport. There are two variables that we recorded. First, the effective movement from start of movement to end as a straight line, also known as the euclidean distance. Second, the continuous movement that is measured frame-by-frame. This includes all deviations between start and end vector. Regarding the euclidean distances, we found that on average, VR users traveled longer distances ($984, SD = 384$) than tablet users ($510, SD = 276$) with $p < 0.05$. Analyzing frame-by-frame movement, we see that on average, tablet users travel 894 units ($SD = 447$) units and VR users 1221

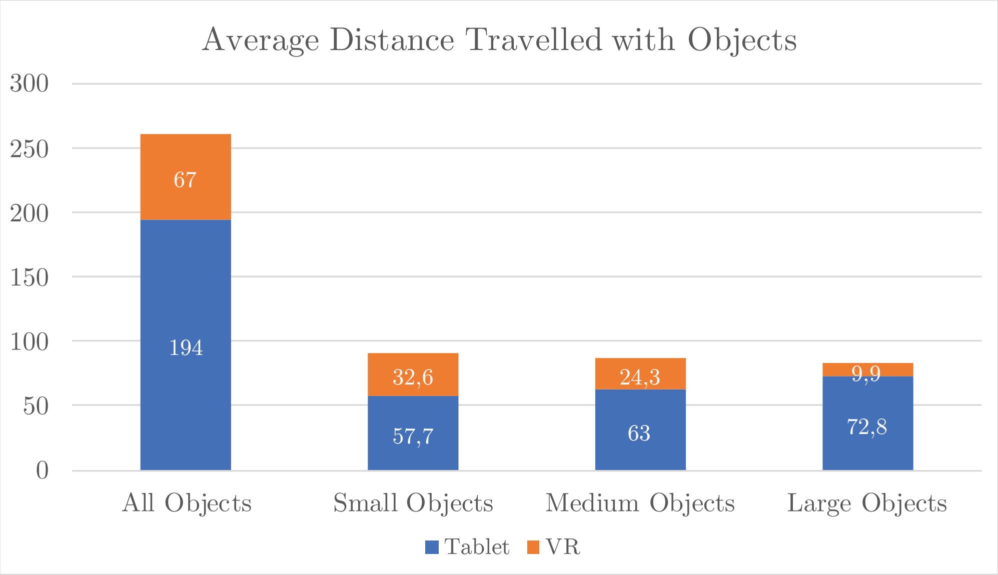


Figure 4.4: Average distances traveled with objects by object size and modality.

units ($SD = 410$), with $p > 0.05$. Regarding user rotation, we analyzed the full amount of rotation along the xyz axes for both modalities. We found differences between tablet (on average $4466°, SD = 2668$) and VR (on average $61355°, SD = 13187$). This results in a factor of 1:13,7 with $p < 0.00005$.

### Questionnaire Results

The System Usability Scale yielded no statistically significant differences, however, on average, tablet and VR users rated both prototypes on average with 74 points. The VR version scored 38.8 ($SD = 16$) points on the NASA TLX while the tablet version scored 36.3 ($SD = 16$). The full results are shown in Figure 4.5. A statistically significant effect was found for physical demand where tablet users on average reported less ($7.2, SD = 6.7$) than VR users ($47.2, SD = 26.6$) with $p < 0.005$.

### Interview Results

In this section, we present the condensed interview feedback. Re-occurring answers are annotated with the number of mentions in brackets. Other comments are individual.

**Particularly good and bad** Tablet users positively mentioned a low learning curve (2), immersion, good view from the audience's perspective, smooth operation, fast object placement at any place, good overview, and the possibility to quickly "hand over" objects to the VR user. On the downside, tablet users criticized the tilting mechanic for objects (2), lacking precision, challenging memorization of gestures, unintuitive zooming and navigation, accidental teleporting, a color glitch when moving objects, and lacking

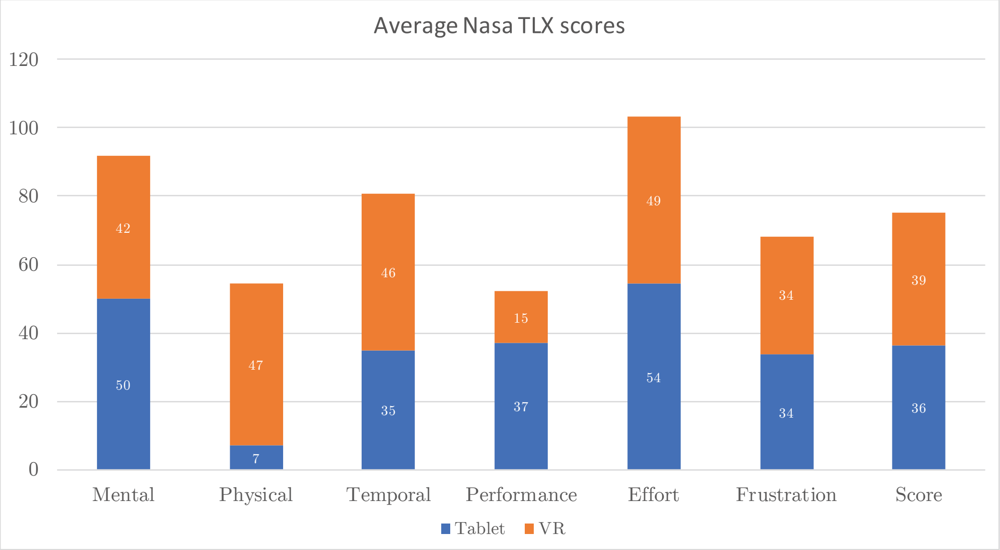


Figure 4.5: Overview of NASA TLX results for each category.

feedback on object rotation status. VR users complimented the intuitive user interface (3), fast learning, joy of use, flexibility, and overall usability. VR users disliked a hitbox glitch that made it sometimes hard to grab and move objects (6), lack of immersion, menu position, teleportation, learning of the button functions, and a missing scaling option.

**Cooperation** Tablet users negatively commented on a lacking common agreement on what and how to build the scene (3), told that they rather performed those operations that were particularly easy for them (5) [placing large objects, rough placement]. Further negative aspects were a lacking idea of what the partner is seeing, and difficult placement of small objects. Positively mentioned were that the downsides of each modality was balanced out (2), helping each other, distribution of fine placement to VR user (3), more time for zooming in, and better overview. VR users liked the strict distribution of work (3), mutual compensation of weaknesses (2), overview function that the tablet offers, separation of fine and rough placement (4), having the big picture, and rapid placement of small objects. VR users missed a function of communicating intent and “what I’m talking about”, and commented that the different views on one scene creates the need for more communication.

**Both tablet/both VR** Tablet users stated that its of advantage having two different views for increased insight, five users declined dual tablet use, two were unambiguous of dual VR use, liked to see what others are doing, and generally highlighted the usefulness of the tablet and the bird perspective. VR users said both should have VR as modality (4) because of easier navigation and interaction and increased efficiency. Others opposed (4) and commented on the positive function of the combination because of overview, the different roles (moderator, actor) that are possible, and the advantage of tablets for large objects, and having multiple options of working on the scene was beneficial. It was further suggested that tablet users could even coordinate different VR users, distributing tasks and work places.

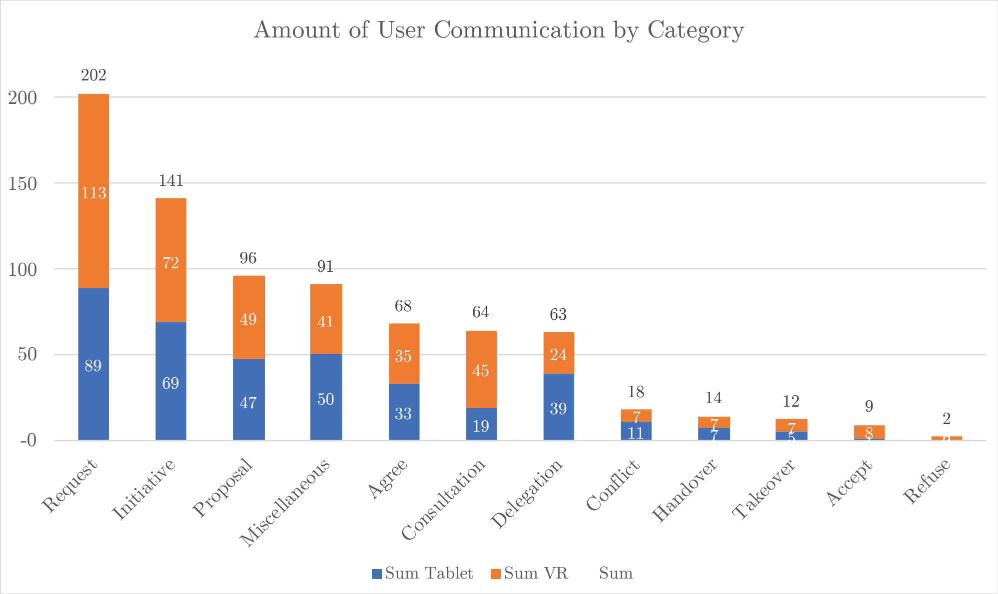


Figure 4.6: Communication that we analyzed from the videos.

### Video Observation Data

We quantified conversational data from the experiment using the coding scheme presented in Table 4.2. We counted the occurrence of any of those categories and accumulated the results based on their origin of modality. This means a phrase like “I will put the tree over there” that a participant operating the tablet expressed was counted as “Tablet” in the Initiate category. The results are presented in Figure 4.6. Out of the nine groups, only eight could be analyzed due to a failure of the memory card of the camera in the 7th group. Summing up, we collected 780 expressions, 410 of those from VR users and 370 from tablet users. We performed group-wise comparisons for statistically significant effects, but found none.

## 4.1.4 Discussion and Conclusion

One of our main observations is that users dominantly worked cooperatively instead of collaboratively. Often, after discussing the goals of the respective tasks in the user study, tablet and VR users assigned tasks to each other and started working independently. This may be due to the nature of the user study where we introduced a time frame for executing the tasks and participants felt pressured finishing the tasks in time. This can also be seen in the results of the video coding. The dominant communication pattern is request (clarifying) and initiative (communication of execution). We observed that after

| **Category** | **Description** | **Example** |
|---|---|---|
| Initiative | Communication of execution | I put that there. |
| Delegation | Delegation of a task | You put that there. |
| Handover | Handing over a task | Can you put that there? |
| Accept | Acceptance of a Handover | Yes I will put that there. |
| Refuse | Refusal of a Handover | No, I can't because I'm busy. |
| Takeover | Takeover of a task | Let me do that. |
| Request | Clarification | Do you think this looks good? |
| Proposal | Proposal of a solution | I think that has to be this way. |
| Agree | Indication of agreement | Yes I like that. |
| Conflict | Indication of a disagreement | No, I don't think so. |
| Consultation | Consultation between users | Maybe put this here? |
| Miscellaneous | Everything else | - |

Table 4.2: Overview of categories for video data analysis.

a phase of consulting on how to achieve the goals, users worked individually on their tasks, from time to time coordinating and getting feedback from the other user.

Another dominant finding is the preference of object manipulations, depending on the device type. We found that tablet users moved more and larger objects than VR users. This could be due to the nature of tablets, that provide an overview on the scene. It appears to be faster to move objects around because multiple objects are visible in a larger context compared to VR, where users experience the scene from a first-person point of view. Here, it seems to be less convenient to change perspectives more often in order to move a larger amount of objects in general. Moreover, VR users manipulated less larger objects. This might be mainly because from a first-person point of view, distances and relations are not as intuitive to understand compared to the overview on tablets. This goes in line with our finding that tablet users moved objects across longer distances. Moving an object from an overview across a scene is easier and more convenient. In the interviews, participants stated that VR users predominantly performed fine adjustments, whereas tablet users more roughly laid out objects. This is also confirmed by our finding that tablet users move longer distances with objects than VR users.

Movement over longer distances are found with VR users, which are an outcome of the teleporting, where longer distances can be overcome with only one button press. Analyzing the travel distance frame-by-frame, there was no significant difference. In our interpretation, this is the more realistic measure of user movement because not it considers movements beyond the start/end vector, including for example tablet movements where users move in circles in the scene while keeping the pinch gesture active. The fact that tablet users moved more large objects could also be influenced by our implementation. In practice, smaller objects cannot be as easily selected on tablet devices because of the smaller hit boxes, which could influence the preference in object size in

Figure 4.7: Witch house scene from the tale Hansel and Gretel created during the study.

both modalities. The same is true for VR, where some objects had too large hit boxes, leading to an accidental selection and manipulation of objects. Regarding the the SUS and NASA TLX scores, we found no differences, with the exception that physical demand was increased for VR users which is not surprising, due to the different ways of user orientation and movement between tablet and VR.

We conclude that there are device-dependent differences in the interaction style that also influence user behavior. First, tablets are primarily used for overview and rough positioning, while VR is mainly used for smaller object manipulation. We did not observe a device-dependent "lead" role because mainly, tasks were distributed and worked on in parallel. Thus, we assume that personal factors are more important in this regard. Summing up, we see potential in facilitating the advantages of different modalities for a common task. Based on our results, system design can be informed in order to increase task performance. For this, further research is needed that more deeply investigates to what extend performance indicators are affected by different modalities, such as time to completion, amount of work accomplished, etc. One example scene that has been created during our study can be seen in Figure 4.7.

## 4.2 Previsualization in Virtual Reality

This section outlines work that the author has done in the narrower context of the first.stage project that researched natural user interfaces for Previsualization (previs). The following motivation is written based on Fröhlich, Münder, et al. (2018). Previs is an essential phase in the design process of narrative media such as film, animation, and stage plays. Digital previs can involve complex technical tasks, e.g. 3D scene creation, animation and camera work, which require trained skills that are not available to all personnel involved in creative decisions for the production.

Previs has been used for decades in all visual design disciplines such as film, product design, and architecture. Through the steady advancement of technology, digital previs gains more and more popularity in film, animation, and the performing arts, see Figure 4.8. It is a collaborative, visual process that enables production teams to creatively explore ideas, plan technical solutions, and communicate a shared vision for efficient production (Okun and Zwerman, 2010). Although previs has many advantages for production, is it still costly, time consuming, and requires trained personnel using complex 3D tools e.g. Maya[6], Blender[7]. In order to make digital previs accessible to non-technical users and to lower the costs in the pre-production phase, the project “first.stage” researches and designs natural user interfaces for previs.

In order to building previs tools that speak the language of the artist instead of a technician, we have to understand how creatives work. Creative people are used to adapt their tools, workflows, and methods of expression depending on what they want to accomplish and feels the most familiar for them. For example, when creating a storyboard, drawing and sketching are preferred ways of expression in which practitioners can express themselves most efficiently, or when expressing in spatiotemporal media, video and sound editing are used to transport a vision of the project in progress. In previs, this creative process often incorporates different tasks on various media such as 2D and 3D layout and animation. All of these media require expressive tools. The ideal previs user interface should support creative freedom and flexibility, so that users experience the previs tool as a natural extension of their own physical capabilities and can express their ideas in a seamless and intuitive way.

However, no single interface can offer natural expression, precision, and task orientation at the same time, see Section 2.2.3. Further, due to the task diversity in previs, a natural user interface should consider a well-balanced workflow: moving between tasks quickly and without confusion. The first.stage user base also comprises of a broad range of users in different roles who have different workflows and stories: what is wanted for animation is not necessarily what is wanted for theater. Users should be able to choose which way of interaction they want to use for different previs tasks, selecting the one that fits their creative needs the best. In this way, we make sure that artists can be creative and

---

[6] https://www.autodesk.com/products/maya/overview

[7] https://www.blender.org/

Figure 4.8: Examples and progression of previs quality over the years. Left: Matrix Reloaded by Laurent Lavigne (ca. 2001). Middle: World War Z by Daniel D. Gregorie (Halon Entertainment, ca. 2011). Right: Avengers Infinity War (The Third Floor, ca. 2016)[9]

expressive and in turn feel natural in their work without forcing them to use an inflexible interface that doesn't account for what artists want to achieve.

Thus, a natural interaction concept for previs should map both user's needs and the previs task characteristics to meaningful and natural interaction techniques. Depending on what users want to accomplish, a range of interaction concepts should be implemented that allow for natural expression and interaction: direct manipulation via touch on 2D interfaces, spatially aware displays, tangible interaction, augmented reality, direct manipulation in 3D and via gestures, full body and embodied interaction, free-hand interaction, and speech.

### 4.2.1 The first.stage NUI Toolset

Virtual Reality (VR) emerges as a core technology for previs in the first.stage project because it provides high-precision direct interaction and manipulation for 3D content using tracked controllers and natural depth perception through stereoscopic view with head-mounted displays. VR is grounded in how humans interact with objects in the real world, making use of the human's capabilities of interaction and perception that everybody is familiar with and an expert in. This is especially prominent when looking at the HTC Vive room-scale tracking where users are able to use their body in a larger context, being able to physically walk to another location in VR instead of having to completely rely on teleport functions. Users are also able to express themselves in a very natural way using and working with their own arms, hands, and head movement. These capabilities make the use of VR very intuitive as the majority of tasks (orientation, locomotion, relocation) can be used and performed without the need to use any interface. More importantly, the visualization capabilities that recent head-mounted displays (HMD) offer, are a strong fit for natural 3D interaction. Having the possibility to experience immersive 3D worlds is one of the core assets that VR offers. No other system can provide such a realistic and

---

[9] Sources: The Matrix: `www.youtube.com/watch?v=KMMeHPGV5VE`, World War Z: `www.youtube.com/watch?v=vrakJLeKj74`, Avengers: `www.youtube.com/watch?v=30VrqKkur7w`

believable first-person experience. Because of these capabilities, in the scientific literature, VR is recently explored as a tool for Virtual Reality Exposure Therapy (VRET) (Anderson et al., 2013), demonstrating immersive visualization capabilities.

Another important aspect of interaction in VR are the controllers. While it can be argued that these offer only rudimentary control and interaction, we argue that in fact those devices offer a rich and natural experience. This is manifested in the large variety of interaction tasks that those devices can be used for. For instance, users can directly point, move, grasp, touch, and perform gestures using the controllers without having to switch modes or tools for these very common operations.

In a practical sense, there are further supporting arguments for the use of VR. First, VR is especially useful for previs as many tasks center on 3D content interaction and manipulation. For example, for animating characters, which is one of the main activities in previs, VR offers a first-person experience of the scene and editing can be done directly in place with the animation control being directly on the characters and not obfuscated through a complex user interface (Erenli and Paglia, 2019). Second, it can also be useful for rehearsals and digital production where today actors rehearse in front of a green screen with non- existent characters (P. Vogel, 2019). Using VR in the previs process at this point, actors get a sense of the scene by experiencing it themselves and in turn have a much better understanding of the context. Directors can switch into any character in the VR scene, thus discovering new perspectives and foreseeing issues and chances in the production. Third, what is most important in the end is the final result that is presented to the audience (Obernhumer and Suttheimer, 2019). With VR, every director can instantly take this perspective, being it in the theater or an animated or non-animated film. This kind of perspective change is not that easily possible in any other medium than VR. Moreover, VR and 360°technology are the future and many more productions will focus on this technology being it in production or as a medium for audience experience (Erenli and Paglia, 2019). Fourth, when planning and previsualizing exhibitions, customers and stakeholders can get a direct and immersive experience of scale, visual and technical appearance of the final design and even interact on content, depending on the detail of the previs.

However, although VR has many advantages for previs, we also believe that the technology has some limitations that are crucial in the productive context. For example, when looking at how multiple people can work together in a room on one scene, sharing ideas and brainstorming, VR is limited in this way because it excludes the VR user from the others because of the head-mounted display (HMD). Although first.stage offers multi-user access via VR, allowing distributed groups to work together, in this scenario, the most natural way of interaction is a shared medium that all users can view and work on at the same time. This problem can be solved by using tablets and mobile devices that implement AR technology.

Using touch interaction on mobile devices, users are able to employ multi-touch usage patterns that they are already familiar with, making the transition of the technology

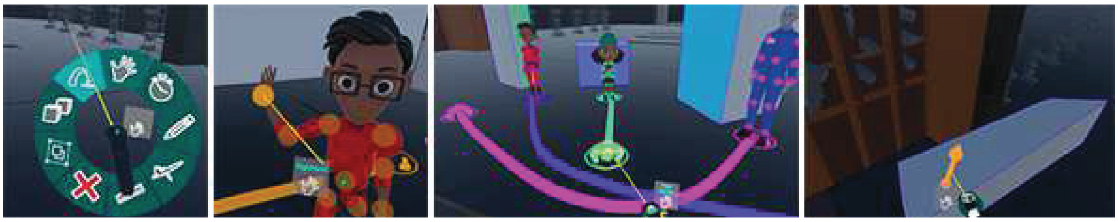


Figure 4.9: Screenshots of the first.stage VR system (from left to right): Menu, puppeteering, character path animation, modeling.

easier. Further, devices such as the iPad Pro or the Microsoft Surface offer high computational power, flexibility, and can be shared among multiple users. As these devices are also equipped with a series of sensors, these can be used to implement spatially aware capabilities that allow for 3D tracking of position and orientation. These devices are affordable and widely available, making it an ideal technology for the implementation of 2D direct manipulation interfaces, and as input devices for 3D interaction, as for example in a camera task. Further, as with VR, AR technology became recently widely available through the integration of ARKit and ARCore in iOS and Android tablets and mobile phones. Working with AR in previs is a very natural experience because users share the spatial context both in the virtual and real world, embedding advanced previs concepts directly in the real world. Directors, artists and producers can stand on a stage, in an outdoor scenery or film set and plan the production using the direct mapping that AR provides.

To sum up, in first.stage, VR is considered the core technology because of the natural interaction paradigms the technology provides for previs and the fact that it is well suited for all basic 3D previs tasks (see Figure 4.9). If needed, users can switch to touch and gesture interaction as adjacent interaction methods that complement the natural interaction capabilities of our first.stage tool set by providing collaborative features in VR, AR, and shared spaces, more expressive tools for modeling (free hand and gesture interaction) and animation (motion capture). With cross device and multi user interaction, the first.stage tools further enhance natural collaboration when working remotely and with multiple users on different devices.

Based on this, we were interested in the perception of previs users regarding the first.stage VR concept. In order to collect practitioner's opinions on this topic, we conducted an online survey to evaluate how professionals with varying technical backgrounds from the film, animation and theater domain assess the use of VR for previs.

### Survey on VR for Previs

The online survey consists of three sections: 1) general information about overall experience and professional background, 2) experience and use of previs, and 3) VR for previs. To assess the possible use of VR for previs, we expected not all participants to be fa-

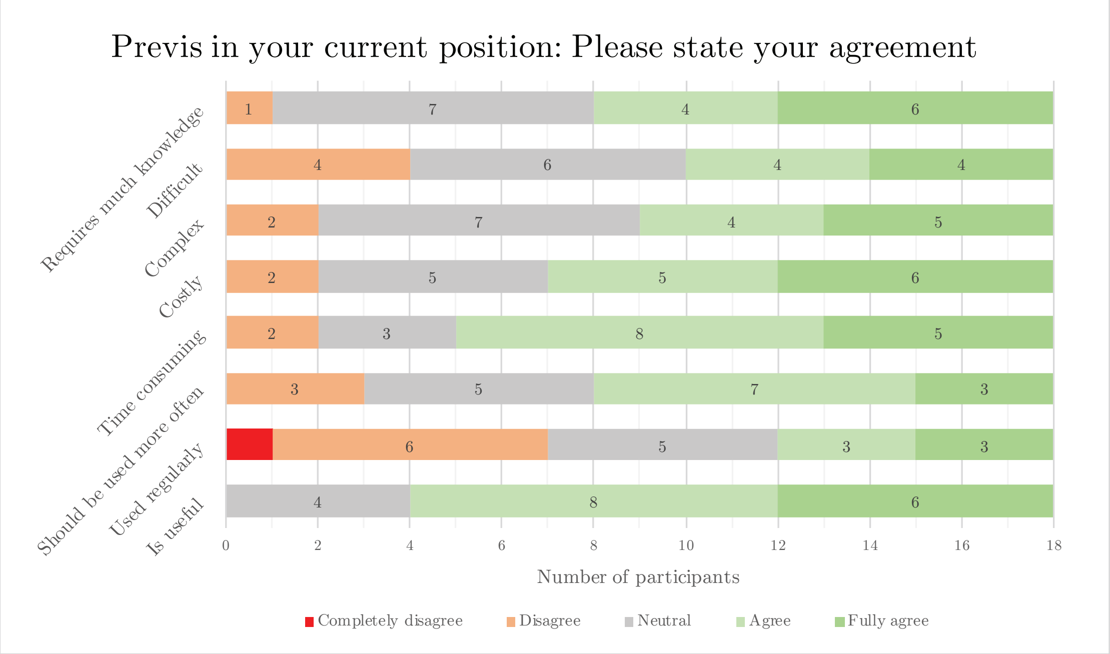


Figure 4.10: Agreement (X axis) to statements (Y axis) on the current use of previs.

miliar with VR and provided an introductory video where we showcased the application of VR for previs for the use cases set design, camera work, and animation. We made sure that participants watched the video by asking control questions about content of the video that had to be correct in order for the following answers to be included in the analysis. All participants answered the control questions correctly and thus all answers were accepted for analysis. The survey was offered in English and German. We invited approximately 400 practitioners working in film, theater, and animation via e-mail and received 18 responses. We present our findings from the survey. As the survey was distributed in German and English, we translated the German results to English and integrated both in one result set.

In the first part of the survey, general questions were asked. Out of the 18 responses, two have a theater background, eleven are from the film industry, three from animation, and two have other creative backgrounds. Most of the people worked for more than ten years in the industry (13 people), four more between six and ten years, and one participant between 3-5 years. The average age is 42 (SD = 9). Regarding the general openness to technology, eleven participants rate themselves as very open to new technology, five as open and two mediocrely open. We asked how many years of experience they have with previs and found that six participants have more than five years of previs experience, another six have no experience, and three have 1-5 years of experience with previs.

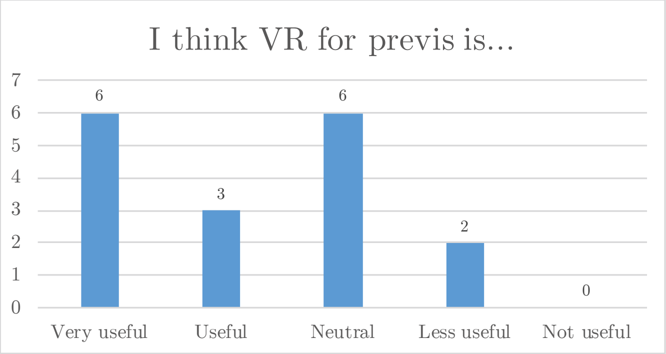


Figure 4.11: Usefulness of VR for previs tasks

In the second part, we asked participants to state their agreement to statements regarding previs in their current position. The statements and results are shown in Figure 4.10. Interestingly, previs is mostly associated with being time consuming and costly. Further, it seems that previs in general is not being used very regularly, while at the same time there is a tendency towards the wish of using it more often, also being supported by the perceived usefulness of previs. In a free text field, we further asked what overall advantages previs offers for production. Positive answers include time and cost savings, better visualization of concepts, precise planning (of complex or expensive shots), presentation and communication capabilities. Negative statements were additional effort, increased costs due to previs efforts, and reduced creativity due to detailed planning.

In the third part, an introductory video to VR was shown and the control questions were asked. After this, we asked about the general perception of VR for previs. The results are shown in Figure 4.11. There seems to be a tendency towards the perception that VR could be a useful tool for previs because six participants rate VR as "very useful", three as "useful" and six more as "neutral". In addition to this, we were interested in the reason behind this usefulness rating and provided a text field where participants could state why they rate previs high or low.

Comments on “very useful” are improved perception of space compared to desktop software, increased communication capabilities and collaboration, intuitive and fast use, better preparation for production, the feeling of “being on site”, and better identification of problems that would only appear on-site. “Useful” answers include easy operation, experience close to reality and “being there”. A “neutral” comment was that theater is being targeted at stage and not wide scenery (in the video, no theater stage was shown explicitly) and another comment questions the “pleasantness” of VR for longer use. One “less useful” response is the missing benefit of VR to the current work area, while no “not useful” comments were contributed. Furthermore, we asked which previs tasks the participants would like to do in VR and which tasks rather not. Positive responses include set design, placement of objects, characters, motion, and texturing. Further, free exploration of sets, a possibility for room or scene capture (e.g. using a 360° camera), complex cam-

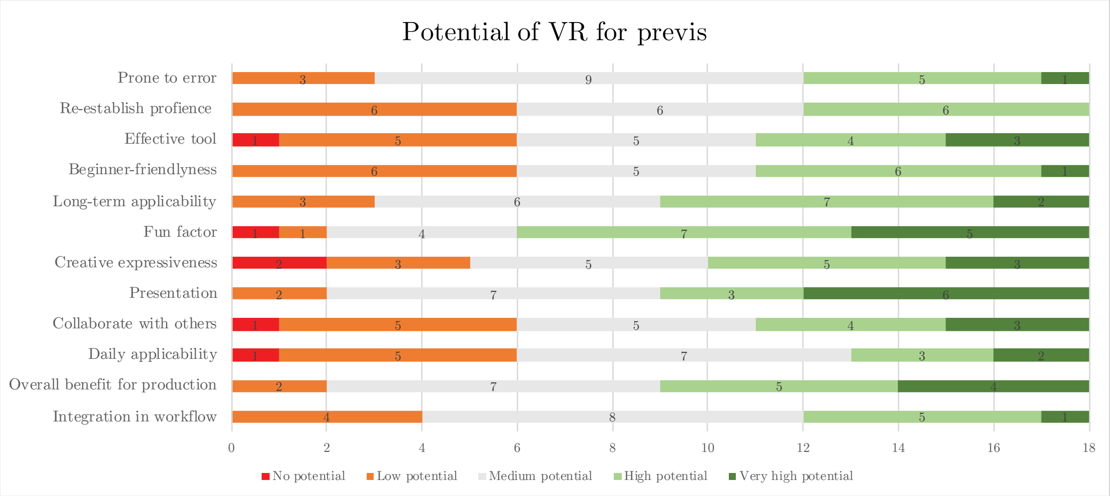


Figure 4.12: Potential of VR for previs.

era movements and placement, framing, editing and sculpting were mentioned. As not inherently VR suited tasks users listed modeling, (complex) animation, detailed editing, performance direction and also camera movement and placement. Interestingly, there are some contradicting statements because four users would like to do camera planning and two other users would rather not do the same thing. Further, animation seems to be perceived skeptically.

Having collected textual feedback on the envisioned previs tasks, we were interested in the perceived potential of virtual reality for previs in relation to several aspects of the technology. We asked the participants to rate on a scale from one (no potential) to five (very high potential) how they rate virtual reality. The results are presented in Figure 4.12, which appears to be balanced in most of the answers. It is noteworthy that although many participants point out the fun factor, there is a minority that doubts this aspect of VR, also concerning the creative expressiveness. Further, there seem to be doubts regarding the collaborative features and the daily applicability. We assume that the mostly evenly split perceptions stem from the limited presentation of the features of VR in the video. For example, collaboration, effectiveness, or expressiveness were not covered to an extend that we assume to provide enough ground for a proper judgment.

We further asked in a text field what opportunities the participants see for VR in previs and which disadvantages, drawbacks, or hurdles. Opportunities are easy access for non-computer oriented users, easy working, novelty, fast results, precise planning, problem finding, variety of application areas, immersion, understanding, and discovery of new perspectives. Participants outlined the following drawbacks: unpleasant HMD, hygienic problems, degrading fun factor, complexity of everyday use, overly virtual, limited asset database, required training for beginners, and costs for equipment.

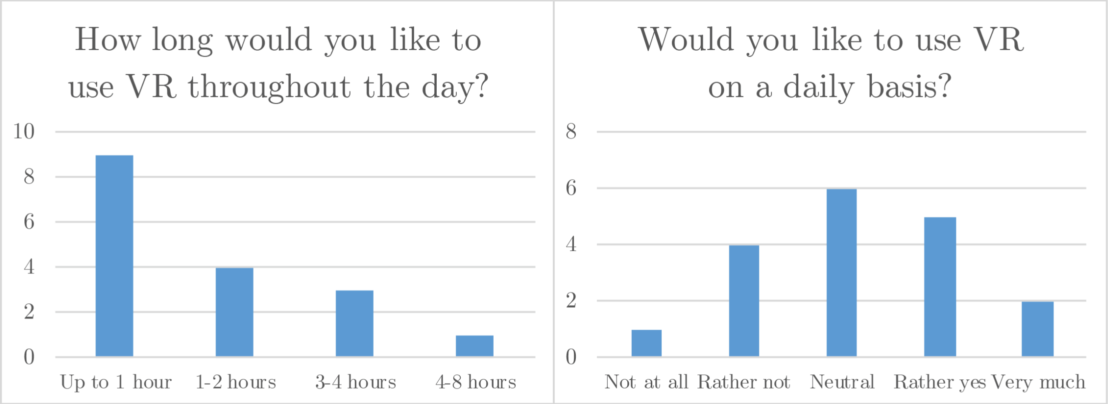


Figure 4.13: Left: VR usage throughout the day. Right: Daily use assessment of VR.

In addition to this, we asked in a text field what advantages of VR there are in comparison to current input methods like mouse and keyboard. VR is perceived as more intuitive and faster, it is easier to reach goals, immersion, natural interaction, and 3D interaction with depth. Disadvantages include isolation in VR, uncomfortable HMD use, lacking frame rate, amount and cost of equipment, learning curve, ability for deployment in working environments, and lacking precision. Further results on the daily usage of VR are presented in Figure 4.13. Most people prefer a short-time use of less than 4 hours a day, while daily use appears to be normally distributed among the answers, indicating a neutral assessment regarding this aspect.

### Discussion and Conclusion

There are many positive as well as negative attributions towards VR for previs. On the positive side, the technology is perceived as being easy, natural, and intuitive, providing immersion, overview and a creative form of expression. On the negative side, VR is mostly anticipated being hard to use all day, costly, and to cause discomfort. Further, it appears as people have not yet the perception of VR being a productivity tool, rather it being too playful and exhausting. There are also some contradicting assessments towards the usage and application of VR. On one hand, users were skeptical about performing tasks like camera handling, while others highlight the usefulness of VR for this. We assume that this originates from the lacking experience with the technology and is an artifact of the design of the online survey, that does not distinguish between experienced users and users without any VR practice.

The application of VR for previs seems to be beneficial for many aspects of use and through developing prototypes for this technology we will continue to learn how VR can contribute to productivity on a daily basis. As an example for the implementation of VR tools for previs, in the next section, a motion-capture system for recording character animations in VR is presented.

### 4.2.2 Embodied Interaction for Animation Capture

An important part of previs is character animation because only by providing believable and interesting animations, previs becomes a convincing tool for production planning. However, the creation of character animation is a task for expert animators using complex tools, requiring years of experience. Novices are usually not able to create even simple animations using keyframing, which is a de facto standard in the industry and supported by all animation tools, such as Maya or Blender. However, there is a simple solution to the problem of creation character animation: motion capturing.

With motion capturing it is possible to record live motions from humans by outfitting them with markers or tracking elements that are attached to the body. Motion capturing can be implemented in very different ways. The most common methods are optical marker tracking, electromagnetic tracking, and inertia-based tracking. Optical marker tracking is very precise and can be used in a large setup with multiple persons being recorded at the same time. A popular producer of such systems is OptiTrack[10]. The downside is that the tracking equipment is very expensive and takes a lot of time to setup and operate. Additionally, animators and technicians are needed who transfer the recorded movements onto virtual characters, reducing the usefulness for small productions, individuals and novice users. However, inertia-based tracking became affordable in the last couple of years and systems are available for as little as 2000 Dollars. Inertia-based tracking facilitates sensors that are attached to a motion-capturing suit. By wearing the suit, the sensors recognize differential movements by the measurement of the inertia of the sensors. Attaching a battery of sensors to a suit in turn leads to an affordable and good-enough precise tracking. With software, the data can be interpreted and mapped onto a virtual model, making the process of capturing and setup very easy.

However, there are still several problems with motion capture. One of the biggest issues is that individuals who work alone have problems of animating a scene completely without context. This means that when wanting to create a scene with multiple people interacting with each other, all animations of each character have to be recorded with the other movements in mind so that the complete scene makes sense after editing. In order to overcome this issue, a VR user interface for capturing embodied animations would be a promising approach. Embodied interactions resemble bodily experiences that every human is familiar with. In contrast to traditional interfaces for motion capture, this system enables users to record animations from the perspective of their own body, to slip in any other body (human or not) and perform animations from their perspective. Embodied interaction relies on the integration of interaction between humans and computers into the material and social environment. In our specific case, the recording of motion capture data via the Rokoko SmartSuit[11] allows for full body interaction and embodied performance animation though inertial sensors that can be worn in the form of a suit. Based on this idea and motivation, a prototype system has been developed and

---

[10] `https://optitrack.com/`

[11] `https://www.rokoko.com/en/products/smartsuit-pro`

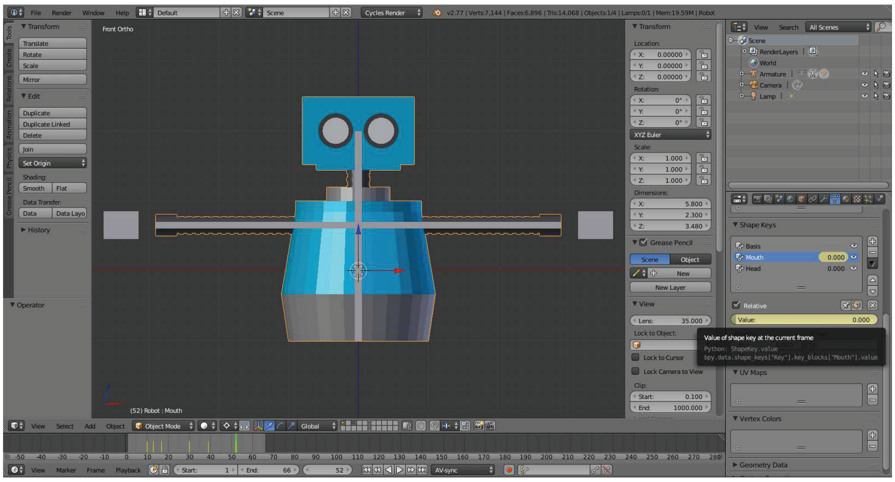


Figure 4.14: Animating a robot in Blender can be a complex undertaking for novices. Image credit: https://medium.com/@jarednielsen

evaluated in the Master thesis of Unger (2018) that the author has provided the idea for, collaborated in, and supervised in collaboration with Thomas Münder.

Using our system, it is very easy and natural to record character animations because users can use their own body as input, directly transferring their motion data to virtual characters, as can be seen in Figure 4.15. In many previs cases and especially for character animation, this makes complex and expensive keyframe animation techniques superfluous. Similar approaches to intuitive character animation using touch gestures have been provided by Walther-Franks, Herrlich, and Malaka (2011), Walther-Franks, Herrlich, Karrer, et al. (2012), and Walther-Franks, Biermann, et al. (2012).

This interaction technique is very natural because it doesn't rely on a translation between user intent and action. There are almost no graphical interfaces needed and the approach has a low learning curve regarding the animation. Using our system, one single user can record a whole crowd of people just by himself in a matter of minutes.

### Prototype and Study

In order to evaluate our system, we performed a preliminary user study with 16 participants. First, we introduce the overall setup and the user interface. Users have to put on the motion suit and place one of the HTC Vive controllers in a pocket attached to the suit to make sure the absolute movements of the user can be tracked in VR. After putting on the HMD and performing a calibration pose, users find themselves in a simple and mostly empty VR scene. When looking around in the scene, users can see that the gray standard character (see Figure 4.15) is mapped to their own movements. In VR, this creates the impression that oneself is the gray character, users are embodied.

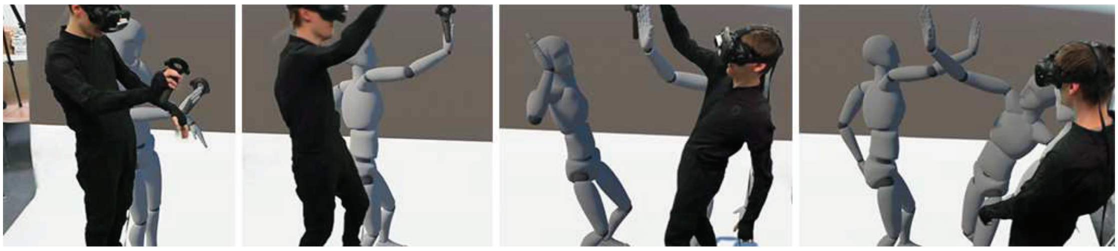
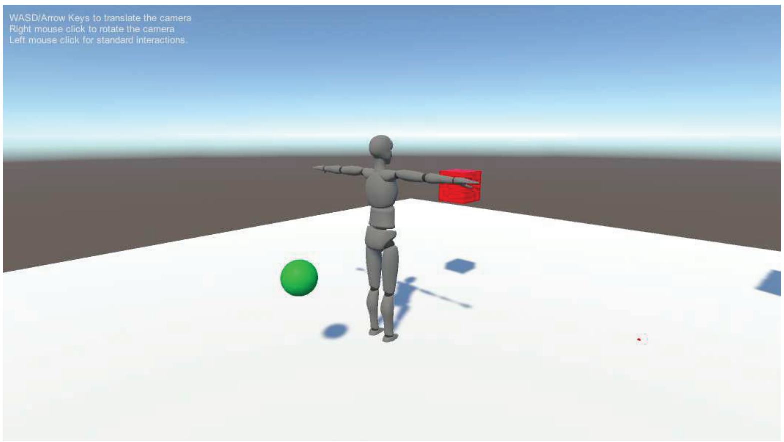


Figure 4.15: Animating a two-person "high-five" in VR (left to right). Starting the recording, animating the first character, switching places and animating the second character, and inspecting the result.

The functions of the system include recording animations and relocating already animated puppets in the scene. Upon pressing the touchpad on the VR controller, the body movements are being recorded as long as the recording is running. When pressing again, the recording is stopped and the animated character starts an infinite playback loop of the animation. Users can grab the character and relocate it inside the scene. Now users are able to create a new recording in the same scene again by pressing the record button. A new character is spawned in place of the user and the recording cycle starts again. This ultimately leads to a scene that is filled with numerous characters, each looping though the recorded animations. With this, it is possible to record character animations of a full room of people just with one user, see the "high five" example depicted in Figure 4.15. For the study, we implemented a baseline version where users could operate the system in a similar way, however without wearing the HMD, operating the system using mouse and keyboard, watching and interacting with the computer screen instead of VR.

In a preliminary study, we evaluated the system with 16 users and collected qualitative feedback as well as quantitative data in the form of SUS and NASA TLX ratings. The participants had to perform four tasks, two basic tasks (BT 1/2) and two advanced tasks (AT 1/2). Basic tasks were to follow the steady movement of two objects inside the scene by replicating the movement (green ball moves circular, red box moves vertically) using the embodied character (see Figure 4.16). Advanced tasks were to record a "high five" scene and a handshake between two characters.

The procedure of the study was as follows. After welcoming, a general introduction and signing the consent forms, participants had to put on the mocap suit. Then, they were introduced to either the VR or desktop condition and performed the randomly assigned two first tasks. Subsequently, the SUS and TLX questionnaires had to be filled in. In the second part, users were introduced to the remaining condition and performed the other two tasks, again filling in the respective questionnaire forms. A final interview was conduced in order to get qualitative feedback on the system.



WASD/Arrow Keys to translate the camera
Right mouse click to rotate the camera
Left mouse click for standard interactions.

Figure 4.16: Study setup. Users had to perform two basic and two advanced tasks. Basic tasks included interaction with the two objects, a green sphere and a red box. Image source: Felix Unger.

### Results and Discussion

Of the 16 participants, 13 were male and 4 female. Four participants were regular VR users, eight use VR sometimes and four never use VR at all. The age span was between 21 and 29 years. The overall SUS score for the VR condition resulted in 84 points, while the desktop version reached 55 points. The average TLX score for the VR condition was 2,96, for the desktop version 4,32. A ranked-sum test resulted in a significant difference in the SUS ratings with $p < 0.005$ as well as in the TLX differences with $p < 0.05$.

The interviews revealed that 15 participants were in favor of the VR condition. 11 participants stated that in VR, it was easier to locate the character movements in the scene. Six participants stated that the VR tool felt more natural. Regarding technical issues, 7 users stated an offset between the own movements and the movements of the model, which is due to the relative tracking approach using inertia-based sensors.

Based on our initial evaluations and user feedback, we are confident, that VR can be a promising way of interaction and creating 3D previs content such as models, animations, and animated scenes. We found that most users quickly become familiar with the VR user interface and interaction styles. Even older users were able to interact with the VR prototypes after a short introduction. However, it must be said that we also recognized the need for other interfaces when it comes to integrating the tools in a larger working context. According to the notions of Jacob et al. (2008), we agree that reality-based interaction is limited in a way that certain task types, at the moment, better can be operated on other devices and interfaces, such as tablet computers or even desktop computers. However, we are confident, that with further development, some of the issues we found could be overcome, creating a more integrated, expressive, and easy to use user experience.

## 4.3 Rapid No-GUI Modeling in VR

As already stated in this thesis, (see Scene Shooter in Section 3.1, and VRBox in Section 3.2) the creation of 3D content, especially the aspect of modeling in 3D can be a challenging task for non-technical users. Available applications such as Blender or Maya are complex tools that are hard to learn for beginners. As much as these applications are ideal for their purpose, that is the creation of complex models including shaders, lighting, animation, etc., beginners can feel overwhelmed by the user interface and the manifold options. This in turn can lead to an exclusion of interested users that would like to approach the domain of 3D models, because of the time and effort those users would have to invest in becoming familiar and productive within these tools.

In this work, we are looking into how to make the creation of 3D models and prototypes more natural and approachable for novice users in order to create a pleasurable experience that motivates diving deeper into the world of 3D modeling. We are not aiming at creating detailed and complex models, the focus is on creating 3D models that represent a rough version of what users have in mind through offering rapid prototyping with natural ways of interaction. From an application point of view, in later stages, the prototyped versions that can be created with our approach can serve as a template for refinement and detailed modeling. This way, novice users can work with rough versions of 3D models that they can use along their further workflow, being able to replace those models at later stages with more detailed versions. For example, indie game developers often have the problem that they have no sufficient 3D models for their games. What often is being done is to browse free asset databases for fitting models. This process can take long time, finding out what fits the needs of the developers best. In the worst case, these users don’t find suitable models, ending up using primitive objects representing their ideas, such as cubes and spheres. These models often fit the purpose of proceeding with the software development, the problem is that these are not representative for early testers, users, funding bodies, and more.

We present a system that allows for rapid prototyping of basic 3D models, focusing on natural ways of creating these models by only using gestures for all interactions. We focus on rapid prototyping because we aim at supporting users that often only use these models as placeholders to be replaced in the further development of their project. These placeholders however often better represent the vision of the creator than arbitrary models from external sources of very primitive ones such as spheres and cubes. We further focus on no-GUI interaction because often, when novice users are confronted with unfamiliar graphical user interfaces (even in VR), they can get distracted from their original goal of creating something, ending up trying different modes and tools, ultimately becoming frustrated because they are unable to create a model that looks sufficient in a short time with relatively moderate effort. Based on this motivation and idea, a prototype application has been developed and evaluated in the Bachelor Thesis of Pollok (2018) that the author has provided the idea for, collaborated in, and supervised.

Figure 4.17: Screenshot of the game Black & White: Player performing a spiral gesture on the game world for creating an effect in the world.

Further, we assume that no-GUI interaction can increase the quality of the user workflow. This means that after getting used to using gestures for all systems interactions, users become more productive in creating models because they are not being visually distracted by menus, and effects, ending up spending more time creating instead of interacting for finding out what tool to use of what parameter to configure. At this point, we would like to highlight that our tool is aimed at quickly creating a rough version of the envisioned model so that the production of the overall project (that the model creation takes place in, such as game development) can continue without spending too much time in creating the perfect model, because these can be exchanged later in the process.

For this prototype we were inspired by the gesture interaction that the game “Black & White” by Peter Molyneux (2001) offers. In this game, players are in the role of a god to a sim world. The player himself is however not directly interacting with the world, players control a god-like creature (e.g. a huge cow-like goddess) that lives in a sim world. Players control and the teach the creature proper ways of interacting with the sim population by drawing gestures on the world surface such as spirals, rectangles, or circles, each representing an individual instruction, see Figure 4.17. For example for casing a spell, rewarding the creature for good behavior or punishing it for bad behavior. This interaction method presents a very playful and fun way of selecting and executing different instructions in the game. In our view, this mechanic can be a very beneficial interaction mechanic because it is a fun and easy way of selecting different modes and does not obstruct the user experience when being internalized.

On this basis we formulate the following research question: *How does no-GUI interaction influence user experience in comparison to menu interactions in VR?*

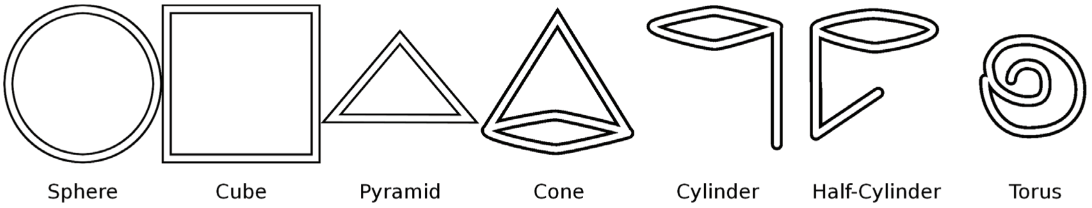


Figure 4.18: Overview of gestures for spawning the respective primitive objects.

For testing our assumptions and answering the research question, we developed a prototype application including no-GUI interaction using gestures, a baseline condition with menu interaction for comparison and conducted a user study with nine participants.

### 4.3.1 Prototype

Following our rapid prototyping strategy, we looked into what current methods there are for the fast creation of models in the real world and decided to implement a method that is similar to the popular building-blocks. Building blocks represent a fun and quick way of prototyping 3D models for children and adults alike (see also literature on Lego as a prototyping method (Mueller et al., 2014; Hofmann et al., 2016)). For modeling, users combine a set of given primitives to more complex models. These primitives are: sphere, cube, pyramid, cone, cylinder, half-cylinder, and torus. The resulting gesture set is depicted in Figure 4.18.

We implemented the following functions for prototyping: selecting primitives, spawning, relocating, and deleting primitives. A problem that always arises in systems with gesture recognition is identifying a purposeful gesture movement (beginning and end). We chose to implement a "starting" gesture that clearly marks the beginning of a gesture attempt. The starting gesture is performed with both hands, the left hand is spread out with the back of the hand facing the user, while with the right hand the gesture-drawing for spawning a primitive is performed (see Figure 4.18) using a pinch gesture, see Figure 4.19. An exemplary of the gesture variant is shown in Figure 4.20.

For development, we use Unity[12] including the SteamVR plugin for VR[13], and the Leap Motion device for hand tracking including their Interaction Engine version 1.1.0[14], the Hover-UI-Kit for the on-screen menu in the baseline condition[15] alongside with a modified version of the PDollar Point-Cloud Gesture Recognizer for identifying the different gestures [16].

---

[12] `https://unity.com/de`

[13] `https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647`

[14] `https://www.leapmotion.com/`

[15] `https://github.com/aestheticinteractive/Hover-UI-Kit`

[16] `https://assetstore.unity.com/packages/tools/input-management/pdollar-point-cloud-gesture-recognizer-21660`

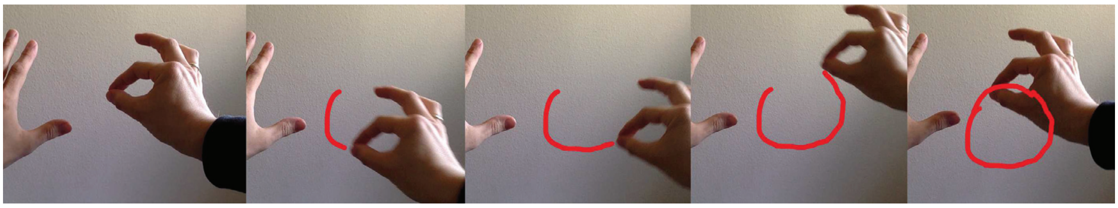


Figure 4.19: Demonstration of drawing a sphere gesture in mid-air using the starting gesture (left hand spread out facing the user, right hand performing the gesture).

### 4.3.2 Study

Our within-design study consists of two parts: using the gesture variant and using a menu variant as a baseline condition. The starting condition has been counter-balanced and there were two tasks for each condition: replication of a given model (see Figure 4.22) and a creative task where users could design a model of their liking (in this order in both conditions). There was no time limit for all tasks as we were mainly interested in the user experience and not task performance and also because most of the participants were not familiar using VR and especially not using mid-air gestures.

We used three different questionnaires in our study. The AttrakDiff for assessing pragmatic and hedonistic quality, the NASA TLX questionnaire for mental workload, and a self-designed questionnaire measuring distraction, concentration, and joy of use on a 7-point Likert scale.

The questions of our own questionnaire were the following:

- The selection of a new primitive distracted me from my original task
- I was able to constantly focus on the task
- The selection of an object was complicated
- I had fun working on the tasks

The procedure was as follows. After welcoming and a general introduction to the experiment, participants filled in the consent form and subsequently provided demographic data including age, experience with VR, and the Leap Motion device. Depending on the randomly chosen starting condition, participants were introduced to VR and general use. After that, the functions of either the gesture or menu variant were explained and participants were told to experiment. Only after both experimenter and participant were confident that all aspects of the system were understood, the first task was performed, directly followed by the second task. After completing both tasks, the three questionnaires had to be filled in. In the next step, participants executed the remaining condition, again filling in the same questionnaires after completing.

Beside the three questionnaires, we counted the following parameters. Accidentally selected primitives and the amount of how often primitives were deleted. Additionally,

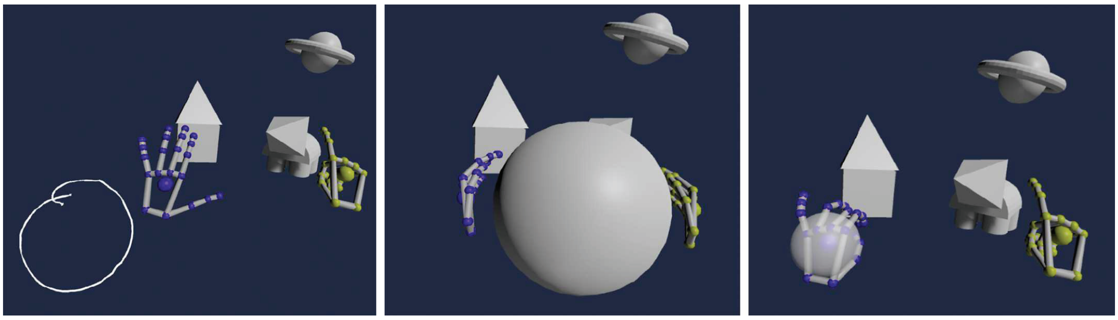


Figure 4.20: First-person view of the gesture-variant. Left: User performing a sphere gesture (sphere appearing outside of focus for better view on scene. Middle: Spawned sphere that can be scaled while pinching for size adjustment. Right: Scaled sphere ready to be placed in scene.

we counted how often the experimenter was consulted in case of questions and participants being stuck, for the creative tasks which and how many primitives were used, and on overall, how much time participants needed for all tasks. Post-hoc interviews were audio-recorded for later analysis.

### 4.3.3 Results

#### Errors and Time

We found significant differences in the error rates between gesture and menu in the replication task. Mean error rate for gesture is higher with 10.6 ($SD = 5.6$) and for menu: 1.1 ($SD = 1.1$) with $p < 0.005$. Further, deletions in the replication task were on average higher for menu 2.7 ($SD = 1.3$) and gesture 1.6 ($SD = 1.8$) with $p < 0.05$. Number of errors was also significant in the creative task for menu ($0.1, SD = 0.3$) and gesture ($7.8, SD = 10.6$) with $p < 0.05$.

Time spent on the replication task was not significant. Average time spent in the gesture condition: $498s$ ($SD = 105s$) and in menu $425s$ ($SD = 137s$). Average time spent in creative mode in the gesture condition: $338s$ ($SD = 191s$) and in menu $305s$ ($SD = 151s$). We did not include time spent in the creative task as there was no common task for all participants and thus, time spent highly differed, depending on the self-chosen goal.

#### NASA TLX

Results show significant differences in four categories and in the final score of the questionnaire. Mental demand: Menu 43 ($SD = 24$), Gesture 61 ($SD = 22$) with $p < 0.05$. Physical demand: Menu 31 ($SD = 12,7$), Gesture 49 ($SD = 16,9$) with $p < 0.05$. Temporal demand: Menu 42 ($SD = 26,8$), Gesture 67 ($SD = 26$) with $p < 0.05$. TLX score: Menu 45 ($SD = 11,3$), Gesture 59 ($SD = 14,5$) with $p < 0.005$.

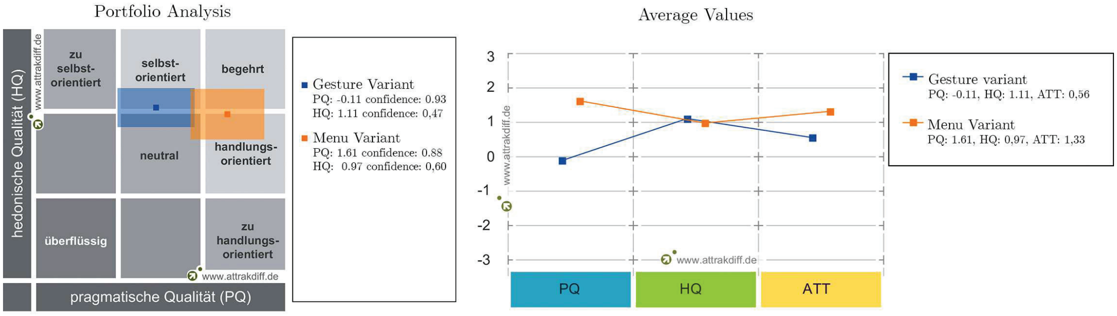


Figure 4.21: Portfolio diagram (left) and average values (right) for pragmatic quality (PQ), hedonistic quality (HQ), and attractiveness (ATT).

### Own Questionnaire

Regarding our own questionnaire, we found that participants perceived the gestural variant $(4.7, SD = 2)$ more complicated that the menu version $(2, SD = 1.9)$ and $p < 0.05$. We found no other significant effects.

### AttrakDiff

The results of the AttrakDiff questionnaire can be found in Figure 4.21. Regarding hedonistic quality, the gesture variant scored 1.11 (0.97 for the menu variant). Pragmatic quality for the gesture is -0.11 (1.61 menu). Attractiveness for gesture is 0.56 (1.33 for menu).

Further, in Figure 4.23, the distribution of the AttrakDiff word pairs is presented. A detailed analysis of differences in the AttrakDiff word pairs revealed that the dimension confusing/clear significantly differs for menu $(5.9, SD = 0.6)$ and gesture $(3.3, SD = 1.3)$ with $p < 0.05$ (higher = better).

### Interviews

We asked each participant the following five questions with the answers in a condensed form below:

**1) Which variant did you like better**
Out of nine participants, eight stated that they like the menu variant better.

**2) What is it that you like better than the other variant**
Participants rated the menu variant as being faster, more efficient, and more reliable. The selection of primitives was easier, and in contrast to the gesture variant, primitives were spawned reliably, and if errors happened, these could easily be corrected using the menu. The one participant that liked the gesture variant better, stated, that joy of use was her main criteria. Three other participants also stated, that the gesture variant was

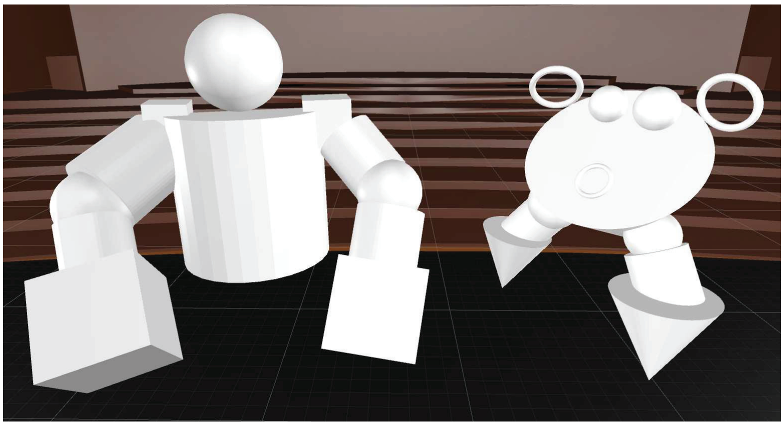


Figure 4.22: The two objects that were used as a template in the replication tasks.

more fun to use, one stated that drawing gestures was playful, two felt more creative drawing gestures, having the feeling that even the selection of the primitives is a creative act in itself.

In general it was mentioned that the tracking of the Leap Motion device was insufficient at many times. Especially when placing newly spawned objects, frustration arised. One participants stated that he would have rated the gesture control better, if there were less problems with the hand tracking. Another participant stated that especially in the menu variant, frustration came up because the menu worked fine, while the positioning of the newly spawned objects was lacking detail, which he did not rate as negative in the gesture variant, because here, both functions did not work very well.

**3) How well did you remember the gestures**
Two participants had problems remembering the gestures, and could only memorize simple gestures. Four participants needed some time but got better at recognizing the gestures also with the help of the gesture hints that were shown next to the left hand. The other three stated that they had no issues remembering the gestures. Two of those further stated that the visual hints, apart from the other issues with the hand tracking, helped them better memorize the options than the wordings in the menu.

**4) What was missing in the gesture variant**
Tracking was one of the biggest issues concerning selection and also in general use. Two participants stated that they would not change anything because it's already intuitive. Two participants thought that the gestures were not distinctive enough and provided suggestions for improvement: placing a preview object next to the gesture hint; introducing a color distinction for better differentiation of gestures and primitives; while drawing the gesture, the currently drawn path should be shown, instead of showing when finished with the gesture.

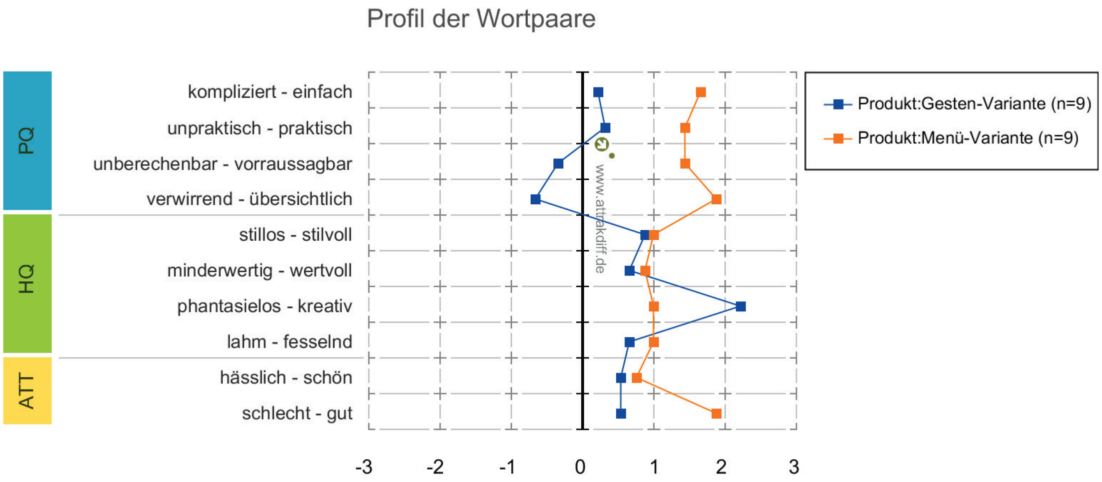


Figure 4.23: Analysis of word pairs.

**5) How do you rate your own performance using both variants?**
Three participants again stated the imprecise placement of objects as the main issue, and further told that the variant did not influence their performance. Regarding the reconstruction tasks, three said that with the menu variant it was easier or faster to select the correct object. One participant had problems identifying the correct items due to language issues. Regarding the creative tasks, a majority told that they preferred the gesture variant as this was perceived as supporting their creativity or more fun. One participant mentioned that the errors in the gesture recognition was not a problem in creative tasks as he could use the wrongly selected primitives just as well. The menu variant was perceived as more efficient than the gesture variant, at the same time also as being less creative. In one case, the menu variant even hindered the creative act.

### 4.3.4 Discussion and Conclusion

Based on the results of the three questionnaires, we conclude that the menu variant outranked the gesture variant in almost all statistically significant categories. To sum up, error rate for gesture was higher in replication and creative task, as well as number of deletions. Regarding mental workload, gestures were more demanding mentally, physically, temporally. Also, the final TLX score was significantly higher in the gesture variant.

In relation to pragmatic and hedonistic quality, we found that the menu variant was rated better in all categories apart from creative/dull. Pragmatic quality is also higher for menu, as well as attractiveness. Hedonistic quality is slightly higher for gesture. The gesture variant was further rated as being more complicated as we found from our own questionnaire.

Answering our research question “How does no-GUI interaction influence user experience in comparison to menu interactions in VR?” we have to say that our gesture variant quite negatively influenced the user experience in general and mental workload.

However, looking at the interview statements and results in favor of the gesture variant, we certainly can say that the implementation of our system was a major factor for these results. The menu variant was perceived as being more efficient, faster and more reliable, which is no surprise considering the high error rates for the gesture variant being a result of the insufficient hand tracking that negatively influenced the gesture condition. Most of the participants had little problems remembering the gestures, although there is some room for improvement regarding in-situ help for remembering. As the gesture variant more heavily depends on proper hand tracking than the menu variant, there was a categorical advantage towards the menu variant.

We conclude that if the tracking issues could be overcome, the gesture variant could be an interesting approach to our task because many participants positively rated the fun factor and creative means of the approach. Until these issues are resolved and further evaluated, the menu variant is the more feasible interaction technique for our task because it is more robust, yields a better user experience, results in less errors, and requires less mental workload.

# 5 Conclusion

This thesis is set out to explore natural and playful user interfaces for 3D digital content creation. We started out by presenting three playful approaches to 3D content creation. In Section 3.1, we investigated how game interfaces could be used for creating 3D worlds where different tools are implemented in the form of weapons, e.g. for spawning individual objects (gun) or multiple objects at the same time (grenade). Following, we presented VRBox, a system that combines natural sand interaction with playful elements for immersive 3D modeling (Section 3.2). Users can teleport themselves directly into the 3D world, exploring different view points in a very easy and intuitive way, and use their own hands in combination with sand, creating complex worlds intuitively. In Section 3.3, we focused on creativity support in playful environments by presenting a system that dynamically introduces limitations in Minecraft game play in order to facilitate creative problem solving.

Regarding natural styles of interaction, in Chapter 4, we investigated how tasks and roles of users in a collaborative tablet/VR system for creating 3D scenes naturally distribute depending on the modality, see Section 4.1. The previsualization of 3D content for film, animation, and theater is part of Section 4.2, where we introduced previs in VR and further presented a system for natural and embodied animation using motion capturing in VR. Lastly, in Section 4.3, we looked into how to create a rapid 3D scene prototyping solution that works without the use of a graphical user interface for fluid and natural interaction in VR.

In the following sections, the main hypothesis are revisited, summarizing the main insights and results of the research chapters.

## 5.1 Contributions

### Benefits of Playful and Natural Interaction for 3D Content Creation

Our first and main hypothesis is $H_1$: *The creation of 3D digital content benefits from playful and natural interaction concepts.* In our studies we found that the overall benefit of playful and natural interfaces center around a) less complexity, b) quick learnability, c) captivation and attractiveness, d) fun, and e) creativeness. Playful and natural interfaces can support and/or enhance user experience regarding to these aspects. Detailed descriptions on these positive results can be found in the research chapters. In our use-cases, the selection and implementation of the natural and playful interaction mechanics was appropriate and worked well in general. However, we also collected various feedback on limitations. These can be grouped as follows: a) limited complexity of editing, b) transfer to other tools, c) seriousness in other contexts, and d) task dependence.

By design, all of these limitations are due to or can be attributed to the user research approach that isolates tasks and user behavior, investigating and isolating effects. Looking beyond the research stage and thinking of applying these concepts in the real world more clearly highlights the importance of these limitations and user concerns. All our prototypes support a certain task and most of the interaction focuses on achieving the experimental task. Users quickly noticed missing features throughout the studies that could be overcome by more specifically designing the tools to support a large variety of functions in order to increase the applicability. Transfer to other tools can be done by supporting common file-formats so that creations can be used further along the production workflow. The aspect of seriousness is an interesting and sociological one. It can be assumed that in a professional context, the presented systems might cause inconveniences for users. Task dependence has to be considered as playful interactions often focus on one specific mechanic and it might be difficult to transfer or extend the functions without compromising the positive effects that they bring.

Summing up, introducing playful mechanics to 3D content creation can result in an overall positive user experience and low learning curve. However, one has to bear in mind that playful approaches might limit functional aspects without wanting to compromise positive characteristics of playful and natural interaction. One way of overcoming this issue can be collaborative interfaces or interfaces that can be exchanged depending on the context of use. For example, in our collaborative tablet/VR system, users mentioned that the advantages of each device could be used depending on what task was to be performed. For rough positioning and overview, tablet users were dominantly performing these tasks. VR users in contrast worked more on positioning smaller objects while using the first-person experience that VR offers to the task.

### Learning of Natural and Playful Interfaces for 3D Content Creation

Our second hypothesis is $H_2$: *Interaction with gestures and natural material support fast understanding and quick learning of 3D content creation.* One of the most important aspects is the learning curve that classic 3D interfaces have and that we aim to overcome with our approaches to playful and natural interaction for 3D content creation. We found a spectrum of results regarding this hypothesis, ranging from very intuitive and easy to learn approaches to interfaces that clearly show the limitations of natural interaction patterns. Another aspect is the technology itself, where users already have basic knowledge of interaction patterns that they can directly apply without much learning required (tablet interaction, first-person games, and VR).

Regarding the spectrum of learnability, we found that the shooter approach was very easy to understand and learn and users consistently mentioned the fast and easy aspects of this interaction style. The same is true for our Minecraft-study, where users interacted in a similar way. With VRBox, we also found that users can very quickly create appealing results with very little training. This is similar to our work on tablet and VR interaction, where users were already familiar with the basic interaction patterns. VR often is mentioned to be easy to understand and fun to use and by itself creates a very good opportunity for 3D content interaction because it enables a variety of standard interaction styles that are natural and easy to use. On the other hand, when completely removing visual information and fully relying on gestures, we find that users have a hard time learning and remembering. We are certain that if users develop a "motor memory" of specific interaction styles (e.g. pinch gestures), interactions can be very fast and easy to retrieve. In general, in our studies, users commented that the interfaces were less complex, and to be learned quickly. However, with tasks that are not pervasive in our in daily life, this is hard (if not impossible) to do. This makes learning hard and tedious, creating errors and ultimately leading to user frustration.

To conclude, there is a clear advantage of natural and playful styles of interaction for 3D content creation when it comes to learning a 3D task such as content creation. Interfaces have to be carefully designed so that the interaction centers around very familiar patterns of use that resemble already acquired world-knowledge. For this, it is best to investigate interactions that are very common and associated with fun experiences, such as VR including reality-based interaction, natural materials, and even desktop shooter-games. Another important aspect is the "fit" of interaction style and task. The better interfaces and interaction metaphors resemble the task context, the better the chances of providing a fast-to-learn interface.

### User Acceptance of Natural and Playful Interfaces in Contrast to WIMP and Menu Interaction

Our next hypothesis states: $H_3$: *Natural and playful interfaces are preferred by users compared to WIMP and menu interaction.* When having a baseline condition that resembles state-of-the-art interaction (shooter, animation, gesture), participants rated the playful and natural approach higher than the classic interaction design. Thus, we can conclude that user's acceptance of our interaction designs is higher compared to standard interfaces. However, it has to be considered that our designs cover certain parts of the whole functionality of a given program. Further, in our studies, users only worked with the system for a limited period of time, making a direct comparison difficult. It can be assumed that the usefulness of playful and natural user interfaces depends on the holistic capabilities of the system and the implementation. For example, when wanting to implement a whole world-editor using shooter mechanics, it will be difficult to design all tasks (lighting, texturing, logic) only using shooter mechanics. Further it can be assumed that the performance in terms of efficiency might be limited by playful mechanics where a task-based editor would have a higher performance in the long run. Something similar can be observed with VR as a representative natural user interface. For example, many VR tools for creative expression (e.g. Oculus Medium[1]) can support user experience and creative expression. However, long-term use with physical exertion may be problematic for professional and extended use. Further, sometimes it's hard to design interactions that bypass technological limitations or characteristics, making it a requirement to focus on other devices for specific use cases. One example is the overview/detail working style that naturally appeared in our tablet/VR study, where it's much more natural to use tablet devices to get an overview of the scene instead of doing the same thing in VR. Thus, the implementation of natural and playful mechanics can be considered as a mode that allows for fluid, fun, and engaging work. Yet, there are times where task-based approaches that are incorporated in tools such as Maya and Blender outperform natural and playful interfaces, especially in a professional context of use, where functions cannot built using playful mechanics, possibly hindering execution and in result lowering task performance.

### Creativity Support and Perceived Creativity

The fourth hypothesis is $H_4$: *Playful approaches to interaction design positively influence user creativity.* There are different aspects to creativity in the context of this work. First, users commented on the creativity of the approach. For example with the shooter, users said the game interface itself is a creative approach to 3D content creation. Second, we learned that users felt creative while working on a given task, as for example with VRBox. This feeling can be induced due to the playful approach that supports serendipitous creations due to the ever-changing sand structure: When building a mountain, a valley appears and other areas are changed without intention and purpose. This might lead to users feeling creative. Third, creativity of the result. As in our Minecraft study, users perceived their own results as creative as well as for VRBox. The perception of the creativity of the result is unarguably subjective. However, having or creating a positive attitude towards the own results is a desirable state. Fourth, creativity of the process that is the result of the inner feeling of self-perceived creativity under influence of the interaction mechanics, and goals. Throughout our studies we collected positive feedback on the creative capabilities of our systems. We learned that when users are not concerned with figuring out how to operate a system, they are free to immerse themselves in the task, which can lead to a flow-like experience where serendipitous behavior can occur in order to find new ways of creating or solving a task. Understanding the interrelations between these factors is beyond this work. However, having users immersed, forgetting about how to interact, can ultimately create an environment for creative expression.

### Collaboration for Supporting Workflow

Our final hypothesis is $H_5$: *Collaboration across devices and players supports a natural working style, supporting 3D content creation.* We observed that collaboration might

---

[1] `https://www.oculus.com/medium/`

benefit user's workflow, creativity and performance. In our Minecraft-study, users collaboratively used limitations as means for improving creative problem solving in 3D scene generation while both were working with the same user interface. In contrast, in the tablet/VR study, users cooperatively created a scene using different devices and interaction modalities. The main difference we observed in user behavior is the cooperation and collaboration between the users. In the tablet/VR study, users worked mainly in parallel, distributing tasks to each user (and device), from time to time consulting on the progress and discussing decisions. Here, users mostly spent their time working individually on different tasks and parts of the scene. We assume that this mainly was the case because the groups were assembled randomly. In contrast, in the Minecraft study, groups were pre-defined by the participants who applied in teams to the experiment. Thus, these participants knew each other and had similar working styles. Looking at the differences, there is a significant difference in team-play. Minecraft users more closely collaborated while tablet/VR users working more independently. Thus, the results have to be interpreted with this knowledge in mind. However, for both studies it's true that the benefit of working together was clearly highlighted by the majority of users. Not only is it more fun to team up with a partner, also creative engagement is supported by having another "pair of eyes", additional input and feedback and the ability to brainstorm.

We can conclude that the collaborative nature of these two studies showed that with and without prior knowledge of the participants, joy of use and motivation was increased in both conditions. If participants were not satisfied with their collaborative experience, this was due to a lack of proper communication and general unfamiliarity with the team partner, highlighting the role of knowing each other.

## 5.2 Reflection and Future Directions

Having presented different implementations of 3D content creation tools, the question of how to use these results and systems arises. When first looking at the prototypes, they can be seen to be tailored to one specific use case with limited task support and missing transferability to other tools and workflows. While this is certainly the case, nevertheless, our results can be used as indicative work that highlight the usefulness of implementing playful and natural ways of interaction to 3D content creation where they can spark joy of use and represent a low entry to 3D modeling and interaction. There are two ways in which to transfer these systems to productive use.

One is to integrate a playful mode in existing tools. For example, when looking at world builders or scene editing tools, there could be a playful mode that can be enabled in order to getting started developing a scene. This way, users have an easy and noncommittal way of starting out a line of work where they have no plan yet, but only a rough idea or feeling. Playful mechanics can support users with achieving higher goals and getting on with their work and ideas by playfully implementing rough versions of what they have in mind. At a later stage, these designs can then be used in a classical manner where they can be extended in a task-centered way. This way of implementing playful technology decreases the social pressure of “playing” at work or playing in a productive context. It is rather to be seen to experimenting, making discoveries, and being creative.

Another option would be to extend the presented approaches with more functionality, extending them to full-fledged tools that are more capable. There might be difficulties along this way, though. For example, VRBox is limited in a way that users can’t reproduce sand structures. When wanting to continue working on models, there is currently no way of recreating a surface. While there might be solutions to this problem (3D printing and shaping using motors, or holographic templates), these are not straight-forward ones.

Concluding, we found that playful and natural user interfaces offer a great chance to engage users in a positive way in otherwise complex tasks such as 3D content creation. Our prototypes can be used as inspiration for implementing playful mechanics in tools that have more functionality. When the purpose of the tool (e.g. world building for 3D shooters) and the playful mechanic match, there is a high chance of users accepting the mode as a creative starting point. This can lead to a lowering of entry hurdles that are often present in 3D applications. With natural interaction, users can facilitate interaction patterns that they already internalized. As in our motion capture example, creating animations using body tracking in combination with VR is very easy and fun for novice users, allowing them to create convincing content for their creations, such as video games or animated films.

# Bibliography

Anderson, Page L, Matthew Price, Shannan M Edwards, Mayowa A Obasaju, Stefan K Schmertz, Elana Zimand, and Martha R Calamaras (2013). “Virtual Reality Exposure Therapy for Social Anxiety Disorder: A Randomized Controlled Trial”. In: *Journal of Consulting and Clinical Psychology* 81.5, p. 751.

Apple (2003). *Xcode.* URL: `https://developer.apple.com/xcode/` (visited on 09/26/2019).

— (2014). *Swift Playgrounds.*

Apple Inc. (2020). *Swift Playgrounds.* [Online; accessed March 12, 2020]. URL: `https://www.apple.com/de/swift/playgrounds/images/overview/run_my_code_large.jpg`.

AUDI AG (2017). *Enter Sandbox.* URL: `https://www.thedrum.com/news/2018/02/15/audi-q5-enter-sandbox-real-time-vr-test-drive`.

Autodesk (1998). *Maya.* URL: `https://www.autodesk.de/products/maya/overview` (visited on 09/26/2019).

Baecker, Ronald M. (June 2014). *Readings in Human-Computer Interaction: Toward the Year 2000.* Elsevier. ISBN: 978-0-08-051574-8.

Barnett, Lynn A. (1990). “Playfulness: Definition, Design, and Measurement”. In: *Play & Culture* 3.4, pp. 319–336.

Bateson, Patrick (2014). “Play, Playfulness, Creativity and Innovation”. In: *Animal Behavior and Cognition* 2.2, p. 99. ISSN: 2372-4323. DOI: `10.12966/abc.05.02.2014`.

— (2015). “Playfulness and Creativity”. In: *Current Biology* 25.1, pp. 12–16. ISSN: 0960-9822. DOI: `10.1016/j.cub.2014.09.009`.

Bateson, Patrick and Paul Martin (2013). *Play, Playfulness, Creativity and Innovation.* Cambridge University Press. ISBN: 978-1-107-01513-5.

Bateson, Patrick and Daniel Nettle (2014). “Playfulness, Ideas, and Creativity: A Survey”. In: *Creativity Research Journal* 26.2, pp. 219–222. DOI: `https://doi.org/10.1080/10400419.2014.901091`.

Bethesda Game Studios (2011). *The Elder Scrolls V: Skyrim.*

Bolt, Richard A. (1980a). *Put That There (Original).* URL: `https://www.youtube.com/watch?v=RyBEUyEtxQo` (visited on 08/09/2019).

— (1980b). “Put-that-there: Voice and gesture at the graphics interface”. In: *ACM SIGGRAPH Computer Graphics* 14.3, pp. 262–270. ISSN: 0097-8930. DOI: `10.1145/800250.807503`.

Brandstätter, Ulrich and Christa Sommerer (2016). "Productive Gaming". In: *Entertainment Computing - ICEC 2016*. Springer, Cham, pp. 260–265. DOI: `10.1007/978-3-319-46100-7_27`.

Brooke, John et al. (1996). "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194, pp. 4–7.

Campbell, Marilyn (2004). "Value of Sandplay as a Therapeutic Tool for School Guidance Counsellors". In: *Australian Journal of Guidance and Counselling* 14. DOI: `10.1017/S103729110000251X`.

Cano, Juan Luis Chulilla (2011). "The Cambrian Explosion of Popular 3D Printing". In: *International Journal of Artificial Intelligence and Interactive Multimedia* 1.4, pp. 30–32.

Chai, Jinxiang and Jessica K. Hodgins (2005). "Performance Animation from Low-dimensional Control Signals". In: *ACM SIGGRAPH 2005 Papers*. SIGGRAPH '05. New York, NY, USA: ACM, pp. 686–696. DOI: `10.1145/1186822.1073248`. (Visited on 02/10/2019).

Chan, Edwin, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer (2016). "User Elicitation on Single-hand Microgestures". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA: ACM Press, pp. 3403–3414. ISBN: 978-1-4503-3362-7. DOI: `10.1145/2858036.2858589`.

Chao, Dennis (2001). "Doom as an Interface for Process Management". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '01*. New York, New York, USA: ACM Press, pp. 152–157. ISBN: 1-58113-327-8. DOI: `10.1145/365024.365078`.

Chen, Hui and Hanqiu Sun (2002). "Real-time Haptic Sculpting in Virtual Volume Space". In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '02. New York, NY, USA: ACM, pp. 81–88. ISBN: 978-1-58113-530-5. DOI: `10.1145/585740.585755`.

Christoffel, Michael and Bethina Schmitt (2002). "Accessing Libraries as Easy as a Game". In: *Visual Interfaces to Digital Libraries*. Springer, Berlin, Heidelberg, pp. 25–38.

Colossal Order (2015). *Cities: Skylines*. URL: `https://www.paradoxplaza.com/cities-skylines/CSCS00GSK-MASTER.html`.

Costa, Silvia da, Darío Páez, Flor Sánchez, Maite Garaigordobil, and Sonia Gondim (2015). "Personal Factors of Creativity: A Second Order Meta-Analysis". In: *Revista de Psicología del Trabajo y de las Organizaciones* 31.3, pp. 165–173. ISSN: 15765962. DOI: `10.1016/j.rpto.2015.06.002`.

Couture, Joel (2017). *Alt.CTRL.GDC Showcase: Sand Garden*. URL: `https://www.gamasutra.com/view/news/290047/AltCTRLGDC_Showcase_Sand_Garden.php` (visited on 09/26/2019).

Craft, Anna (2001). "'Little c Creativity'". In: *Creativity in Education*, pp. 45–61.

Csikszentmihályi, Mihály (1990). "The Domain of Creativity". In: *Theories of Creativity*. Sage Focus Editions, Vol. 115. Thousand Oaks, CA, US: Sage Publications, Inc, pp. 190–212.

Csikszentmihályi, Mihály (2013). *Creativity: Flow and the Psychology of Discovery and Invention.* Harper Perennial Modern Classics. ISBN: 978-0-06-228325-2.

— (2014). “Toward a Psychology of Optimal Experience”. In: *Flow and the Foundations of Positive Psychology.* Dordrecht: Springer Netherlands, pp. 209–226.

Cut (Nov. 2018). *The Hopscotch Experiment | Dirty Data | Cut.* URL: `https://www.youtube.com/watch?v=7P-PNa0bxQg` (visited on 04/29/2019).

Datcu, Dragos and Stephan Lukosch (2013). “Free-hands Interaction in Augmented Reality”. In: *Proceedings of the 1st Symposium on Spatial User Interaction.* SUI ’13. New York, NY, USA: ACM, pp. 33–40. ISBN: 978-1-4503-2141-9. DOI: `10.1145/2491367.2491370`.

De Bono, Edward (2017). *Six Thinking Hats.* Penguin UK. ISBN: 978-0-241-25753-1.

Design Council (2007). *Eleven lessons: managing design in eleven global brands. A study of the design process.* en. Tech. rep. Design Council. URL: `https://www.designcouncil.org.uk/resources/report/11-lessons-managing-design-global-brands` (visited on 04/10/2019).

Elam, Joyce J. and Melissa Mead (1990). “Can Software Influence Creativity?” In: *Information Systems Research* 1.1, pp. 1–21. ISSN: 10477047. DOI: `10.1287/isre.1.1.1`.

Epic Games (1998). *Unreal Engine.* URL: `https://www.unrealengine.com/`.

Erenli, Kai and Francesco Paglia (2019). *D1.1 Requirements Definition for Animation Productions.* Technical Report D1.1. URL: `http://first-stage.eu/deliverables.html` (visited on 09/21/2019).

Francese, Rita, Ignazio Passero, and Genoveffa Tortora (2012). “Wiimote and Kinect: Gestural User Interfaces Add a Natural Third Dimension to HCI”. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces.* Capri Island, Italy: ACM, pp. 116–123.

Freeman, Euan, Stephen Brewster, and Vuokko Lantz (2016). “Do That, There: An Interaction Technique for Addressing In-Air Gesture Systems”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16.* New York, New York, USA: ACM Press, pp. 2319–2331. ISBN: 978-1-4503-3362-7. DOI: `10.1145/2858036.2858308`.

Frohlich, David M. (1993). “The History and Future of Direct Manipulation”. In: *Behaviour & Information Technology* 12.6, pp. 315–329. ISSN: 0144-929X. DOI: `10.1080/01449299308924396`.

Fröhlich, Thomas (2017). “Play While You Work: Productive Play for Digital Content Creation”. In: *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY ’17 Extended Abstracts.* Amsterdam, The Netherlands: ACM Press, pp. 699–702. ISBN: 978-1-4503-5111-9. DOI: `10.1145/3130859.3133228`.

Fröhlich, Thomas, Dmitry Alexandrovsky, Timo Stabbert, Tanja Döring, and Rainer Malaka (2018). “VRBox: A Virtual Reality Augmented Sandbox for Immersive Playfulness, Creativity and Exploration”. In: *The Annual Symposium on Computer-Human Interaction in Play - CHI PLAY ’18.* Melbourne, VIC, Australia: ACM Press, pp. 153–162. DOI: `10.1145/3242671.3242697`.

Fröhlich, Thomas, Thomas Münder, Dirk Wenig, Tanja Döring, and Rainer Malaka (2018). *D1.6 NUI Concept and Interaction Guideline for Previs Software*. Technical Report D1.6. University of Bremen.

Fröhlich, Thomas, Jan von Oehsen, and Rainer Malaka (2017). "Productivity & Play: A First-Person Shooter for Fast and Easy Scene Design". In: *CHI PLAY '17 Extended Abstracts*. Amsterdam, The Netherlands: ACM Press, pp. 353–359. ISBN: 978-1-4503-5111-9. DOI: `10.1145/3130859.3131319`.

Fu, Limin Paul, James Landay, Michael Nebeling, Yingqing Xu, and Chen Zhao (2018). "Redefining Natural User Interface". In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. ACM. ISBN: 978-1-4503-5621-3. DOI: `10.1145/3170427.3190649`.

Galoppo, Nico, Serhat Tekin, Miguel A. Otaduy, Markus Gross, and Ming C. Lin (2007). "Interactive Haptic Rendering of High-Resolution Deformable Objects". en. In: *Virtual Reality*. Ed. by Randall Shumaker. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 215–223. ISBN: 978-3-540-73335-5.

Galyean, Tinsley A. and John F. Hughes (1991). "Sculpting: An Interactive Volumetric Modeling Technique". In: *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: ACM, pp. 267–274. ISBN: 978-0-89791-436-9. DOI: `10.1145/122718.122747`.

Garvey, Catherine (1990). *Play*. Harvard University Press. ISBN: 978-0-674-67365-6.

Garwood, Russell and Jason Dunlop (2014). "The Walking Dead: Blender as a Tool for Paleontologists with a Case Study on Extinct Arachnids". In: *Journal of Paleontology* 88.4, pp. 735–746.

Gates, Bill (2016). *Content is King*. URL: `https://medium.com/@HeathEvans/content-is-king-essay-by-bill-gates-1996-df74552f80d9` (visited on 09/26/2019).

Gaver, William (2002). "Designing for homo ludens". In: *I3 Magazine* 12.June, pp. 2–6.

— (2006). "The video window: my life with a ludic system". In: *Personal and Ubiquitous Computing* 10.2-3, pp. 60–65. DOI: `10.1007/s00779-005-0002-2`.

Gaver, William, John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars, and Brendan Walker (2004). "The drift table: designing for ludic engagement". In: *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04*. New York, New York, USA: ACM Press, p. 885. ISBN: 1-58113-703-6. DOI: `10.1145/985921.985947`.

Geyer, Florian, Ulrike Pfeil, Anita Höchtl, Jochen Budzinski, and Harald Reiterer (2011). "Designing Reality-based Interfaces for Creative Group Work". In: *Proceedings of the 8th ACM Conference on Creativity and Cognition*. New York, NY, USA: ACM, pp. 165–174. ISBN: 978-1-4503-0820-5. DOI: `10.1145/2069618.2069647`.

Gugenheimer, Jan, Christian Mai, Mark McGill, Julie Williamson, Frank Steinicke, and Ken Perlin (2019). "Challenges Using Head-Mounted Displays in Shared and Social Spaces". en. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19*. Glasgow, Scotland Uk: ACM Press, pp. 1–8. ISBN: 978-1-4503-5971-9. DOI: `10.1145/3290607.3299028`.

Gugenheimer, Jan, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio (2017). “ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17*. Denver, Colorado, USA: ACM Press, pp. 4021–4033. ISBN: 978-1-4503-4655-9. DOI: `10.1145/3025453.3025683`.

Guildford, J.P. (1967). *The Nature of Human Intelligence*. New York: McGraw-Hill.

Hart, Sandra G. and Lowell E. Staveland (Jan. 1988). “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research”. In: *Advances in Psychology*. Ed. by Peter A. Hancock and Najmedin Meshkati. Vol. 52. Human Mental Workload. North-Holland, pp. 139–183.

Hassenzahl, Marc, Michael Burmester, and Franz Koller (2003). “AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität.” In: *Mensch & Computer*. Vol. 2003, pp. 187–196.

Herrlich, Marc, Andreas Braun, and Rainer Malaka (2012). “Towards Bimanual Control for Virtual Sculpting”. en. In: *Mensch & Computer 2012: interaktiv informiert – allgegenwärtig und allumfassend!?* Oldenbourg Verlag. ISBN: 978-3-486-71879-9. URL: `http://dl.gi.de/handle/20.500.12116/7798` (visited on 02/10/2019).

Herrlich, Marc, Markus Krause, Lasse Schwarten, Jens Teichert, and Benjamin Walther-Franks (2008). “Multitouch Interface Metaphors for 3D Modeling”. In: *Proceedings of the IEEE Tabletops and Interactive Surfaces*. Amsterdam, The Netherlands: IEEE.

Herrlich, Marc, Benjamin Walther-Franks, and Rainer Malaka (2011). “Integrated Rotation and Translation for 3D Manipulation on Multi-Touch Interactive Surfaces”. In: *Smart Graphics*. Berlin, Heidelberg: Springer, pp. 146–154. ISBN: 978-3-642-22571-0.

Herrlich, Marc, Benjamin Walther-Franks, Roland Schröder-Kroll, Jan Holthusen, and Rainer Malaka (2011). “Proxy-Based Selection for Occluded and Dynamic Objects”. In: *Smart Graphics*. Berlin, Heidelberg: Springer, pp. 142–145. URL: `http://link.springer.com/10.1007/978-3-642-22571-0_15`.

Hofmann, Megan, Jeffrey Harris, Scott E. Hudson, and Jennifer Mankoff (2016). “Helping Hands: Requirements for a Prototyping Methodology for Upper-limb Prosthetics Users”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. New York, NY, USA: ACM, pp. 1769–1780. ISBN: 978-1-4503-3362-7. DOI: `10.1145/2858036.2858340`.

Huizinga, Johan (1949). *Homo Ludens - A Study of the Play-Element in Culture*. London: Routledge & Kegan Paul.

id Software (Dec. 1993). *Doom*. Richardson, Texas, USA.

— (Nov. 1997). *Quake II*. Santa Monica, Kalifornien, USA.

Ishii, H., C. Ratti, B. Piper, Y. Wang, A. Biderman, and E. Ben-Joseph (Oct. 2004). “Bringing Clay and Sand into Digital Design — Continuous Tangible User Interfaces”. English. In: *BT Technology Journal* 22.4, pp. 287–299. ISSN: 1358-3948, 1573-1995. DOI: `10.1023/B:BTTJ.0000047607.16164.16`.

Ishii, Hiroshi and Brygg Ullmer (1997). “Tangible bits: towards seamless interfaces between people, bits and atoms”. en. In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’97*. Atlanta, Georgia, United States:

ACM Press, pp. 234–241. ISBN: 978-0-89791-802-2. DOI: `10.1145/258549.258715`. URL: `http://portal.acm.org/citation.cfm?doid=258549.258715` (visited on 04/10/2019).

Jacob, Robert J K, Audrey Girouard, Leanne M Hirshfield, Michael S Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum (2008). “Reality-Based Interaction: A Framework for Post-WIMP Interfaces”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 201–210.

Jarrett, Olga, Stacey French-Lee, Nermin Bulunuz, and Mizrap Bulunuz (2010). “Play in the Sandpit: A University and a Child-Care Center Collaborate in Facilitated-Action Research”. In: *American Journal of Play* 3.2, pp. 221–237. ISSN: 1938-0399. URL: `https://eric.ed.gov/?id=EJ1070242`.

Johansson, Henrik and Daniel Forchheimer (2005). *The Brutal File Manager*. URL: `http://www.forchheimer.se/bfm/`.

Kado, Keita and Gakuhito Hirasawa (2018). “Two-way Cooperation of Architectural 3D Cad and Game Engine”. In: *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. VRCAI ’18. cited. New York, NY, USA: ACM, 22:1–22:4. ISBN: 978-1-4503-6087-6. DOI: `10.1145/3284398.3284420`. URL: `http://doi.acm.org/10.1145/3284398.3284420` (visited on 03/23/2019).

Kaufman, James C and Ronald A Beghetto (2009). “Beyond Big and Little: The Four C Model of Creativity”. In: *Review of General Psychology* 13.1, pp. 1–12. DOI: `10.1037/a0013688`. URL: `https://www.researchgate.net/profile/James_Kaufman/publication/228345133_Beyond_Big_and_Little_The_Four_C_Model_of_Creativity/links/0fcfd50cb7f1261553000000.pdf`.

Khademi, Maryam, Hossein Mousavi Hondori, Alison McKenzie, Lucy Dodakian, Cristina Videira Lopes, Steven C. Cramer, Maryam Khademi, Hossein Mousavi Hondori, Alison McKenzie, Lucy Dodakian, Cristina Videira Lopes, and Steven C. Cramer (2014). “Free-hand interaction with leap motion controller for stroke rehabilitation”. In: *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA ’14*. New York, New York, USA: ACM Press, pp. 1663–1668. ISBN: 978-1-4503-2474-8. DOI: `10.1145/2559206.2581203`. URL: `http://dl.acm.org/citation.cfm?doid=2559206.2581203`.

Kim, Hyosun, Georgia Albuquerque, Sven Havemann, and Dieter W Fellner (2005). “Tangible 3D: Hand Gesture Interaction for Immersive 3D Modeling”. In: The Eurographics Association, pp. 191–199. ISBN: 978-3-905674-06-4. DOI: `10.2312/EGVE/IPT_EGVE2005/191-199`.

Klug, Matthias (Feb. 2018). “Kollaboration zwischen Virtual Reality und Tablet beim Multi-User-Szenendesign”. Bachelor Thesis. Bremen: University of Bremen.

Krueger, Myron W., Thomas Gionfriddo, and Katrin Hinrichsen (1985). “VIDEOPLACE - An Artificial Reality”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’85. New York, NY, USA: ACM, pp. 35–40. ISBN: 978-0-89791-149-8. DOI: `10.1145/317456.317463`.

Kuts, Ekaterina (2009). “Playful User Interfaces: Literature Review and Model for Analysis”. In: *Proceedings of DiGRA 2009*. URL: `http://www.digra.org/wp-content/uploads/digital-library/09287.38077.pdf%20http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.190.3615`.

Lee, Jehee, Jinxiang Chai, Paul S. A. Reitsma, Jessica K. Hodgins, and Nancy S. Pollard (2002). “Interactive Control of Avatars Animated with Human Motion Data”. In: *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’02. New York, NY, USA: ACM, pp. 491–500. ISBN: 978-1-58113-521-3. DOI: `10.1145/566570.566607`.

Lents, Nathan H. (Jan. 2017). *Why Play Is Important*. URL: `https://www.psychologytoday.com/us/blog/beastly-behavior/201705/why-play-is-important` (visited on 07/09/2019).

Lieberman, J. Nina (Dec. 1965). “Playfulness and Divergent Thinking: An Investigation of their Relationship at the Kindergarten Level”. In: *The Journal of Genetic Psychology* 107.2, pp. 219–224. ISSN: 0022-1325. DOI: `10.1080/00221325.1965.10533661`. URL: `https://doi.org/10.1080/00221325.1965.10533661` (visited on 04/10/2019).

Liu, C. Karen and Victor B. Zordan (2011). “Natural User Interface for Physics-based Character Animation”. In: *Proceedings of the 4th International Conference on Motion in Games*. MIG’11. event-place: Edinburgh, UK. Berlin, Heidelberg: Springer-Verlag, pp. 1–14. ISBN: 978-3-642-25089-7. DOI: `10.1007/978-3-642-25090-3_1`. URL: `http://dx.doi.org/10.1007/978-3-642-25090-3_1` (visited on 02/10/2019).

Malaka, Rainer (Nov. 2017). *Keynote: Same Same but Different - Entertainment Computing und Mensch-Technik-Interaktion*. Keynote. Regensburg.

Mann, Steve (2007). “Natural Interfaces for Musical Expression: Physiphones and a Physics-based Organology”. In: *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*. NIME ’07. event-place: New York, New York. New York, NY, USA: ACM, pp. 118–123. DOI: `10.1145/1279740.1279761`. URL: `http://doi.acm.org/10.1145/1279740.1279761` (visited on 04/11/2019).

Maslow, A. H. (1943). “A theory of human motivation”. In: *Psychological Review* 50.4, pp. 370–396. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: `10.1037/h0054346`.

Mednick, Sarnoff (1962). “The associative basis of the creative process”. In: *Psychological Review* 69.3, pp. 220–232. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: `10.1037/h0048850`.

Moscovich, Tomer, Takeo Igarashi, Jun Rekimoto, Kentaro Fukuchi, and John F. Hughes (2005). “A Multi-finger Interface for Performance Animation of Deformable Drawings”. In: *Proc. of User Interface Software and Technology (UIST ’05*. ACM Press.

Mueller, Stefanie, Tobias Mohr, Kerstin Guenther, Johannes Frohnhofen, and Patrick Baudisch (2014). “faBrickation: Fast 3D Printing of Functional Objects by Integrating Construction Kit Building Blocks”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. event-place: Toronto, Ontario, Canada. New York, NY, USA: ACM, pp. 3827–3834. ISBN: 978-1-4503-2473-1. DOI: `10.1145/2556288.2557005`. URL: `http://doi.acm.org/10.1145/2556288.2557005` (visited on 04/14/2019).

Münder, Thomas, Thomas Fröhlich, and Rainer Malaka (2018). “Empowering Creative People: Virtual Reality for Previsualization”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*. Montreal QC, Canada: ACM Press, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: `10.1145/3170427.3188612`.

Myron Krueger (1985). *Videoplace, Myron Krueger (1985)*. URL: `https://www.youtube.com/watch?v=YyIWzapmOLI` (visited on 08/09/2019).

NUI Group (Mar. 2011). *Natural User Interface*. URL: `http://wiki.nuigroup.com/Natural_User_Interface` (visited on 08/09/2019).

Obernhumer, Kathrin and Frank Suttheimer (2019). *D1.3 Requirements Definition for Stage Productions*. Technical Report D1.3. URL: `http://first-stage.eu/deliverables.html` (visited on 09/21/2019).

Oehsen, Jan von (Aug. 2017). “Intuitives 3D-Leveldesign für First-Person Spiele”. Bachelor Thesis. Bremen: University of Bremen.

Okun, Jeffrey A and Susan Zwerman (2010). *The VES handbook of visual effects: industry standard VFX practices and procedures*. Taylor & Francis.

Perec, Georges (Jan. 2001). *La Disparition*. Paris: Gallimard. ISBN: 978-2-07-071523-7.

Persson, Markus and Jens Bergensten (Nov. 2009). *Minecraft*. Published: Game [Misc].

Peter Molyneux (2001). *Black & White*. Guildford, UK.

Piper, Ben, Carlo Ratti, and Hiroshi Ishii (2002). “Illuminating Clay: A 3-D Tangible Interface for Landscape Analysis”. In: ACM Press, p. 355. ISBN: 978-1-58113-453-7. DOI: `10.1145/503376.503439`.

Pollok, Christopher (Oct. 2018). “Vergleich von Menü- und Gesteninteraktion für 3D-Szenenerstellung in Virtual Reality”. Bachelor Thesis. Bremen: University of Bremen.

Preim, Bernhard and Raimund Dachselt (2010). *Interaktive Systeme Band 2*. ISBN: 978-3-642-05401-3. DOI: `10.1007/978-3-642-05402-0`. URL: `http://link.springer.com/10.1007/978-3-642-05402-0`.

Reed, SE, O Kreylos, S Hsi, LH Kellogg, G Schladow, MB Yikilmaz, H Segale, J Silverman, S Yalowitz, and E Sato (2014). “Shaping Watersheds Exhibit: An Interactive, Augmented Reality Sandbox for Advancing Earth Science Education”. In: *AGU Fall Meeting Abstracts*. Vol. 1, p. 01.

Reisman, Jason L., Philip L. Davidson, and Jefferson Y. Han (2009). “A Screen-space Formulation for 2D and 3D Direct Manipulation”. In: *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*. UIST ’09. event-place: Victoria, BC, Canada. New York, NY, USA: ACM, pp. 69–78. ISBN: 978-1-60558-745-5. DOI: `10.1145/1622176.1622190`. URL: `http://doi.acm.org/10.1145/1622176.1622190` (visited on 02/10/2019).

Rendl, Christian, Patrick Greindl, Kathrin Probst, Martin Behrens, and Michael Haller (2014). “Presstures: exploring pressure-sensitive multi-touch gestures on trackpads”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*. New York, New York, USA: ACM Press, pp. 431–434. ISBN: 978-1-4503-2473-1. DOI: `10.1145/2556288.2557146`. URL: `http://dl.acm.org/citation.cfm?doid=2556288.2557146`.

Rhodes, Mel (1961). "An Analysis of Creativity". In: *The Phi Delta Kappan* 42.7, pp. 305–310. ISSN: 0031-7217. URL: `https://www.jstor.org/stable/20342603` (visited on 04/10/2019).

Ritchey, Tom (1998). "General Morphological Analysis". In: *16th Euro Conference on Operational Analysis.*

Rockstar (2013). *Grand Theft Auto V.*

Rohrbach, Bernd (1969). "Kreativ nach Regeln". In: *Absatzwirtschaft* 12.19, pp. 73–76.

Roo, Joan Sol, Renaud Gervais, Jeremy Frey, and Martin Hachet (2017). "Inner Garden: Connecting Inner States to a Mixed Reality Sandbox for Mindfulness". English. In: ACM Press, pp. 1459–1470. ISBN: 978-1-4503-4655-9. DOI: `10.1145/3025453.3025743`.

Roosendaal, Ton (1998). *Blender.*

Rose, Mike (1984). *Writer's Block: The Cognitive Dimension.* SIU Press.

Russ, Sandra W. (July 2003). "Play and Creativity: Developmental issues". In: *Scandinavian Journal of Educational Research* 47.3, pp. 291–303. DOI: `10.1080/00313830308594`. URL: `http://www.tandfonline.com/doi/abs/10.1080/00313830308594`.

Sánchez, Sergio Álvarez, Laura Delgado Martín, Miguel Ángel Gimeno-González, Teresa Martín-Garcia, Fernando Almaraz-Menéndez, and Camilo Ruiz (2016). "Augmented Reality Sandbox: A Platform for Educative Experiences". In: *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality.* ACM, pp. 599–602.

Scharf, Florian, Christian Wolters, Michael Herczeg, and Jörg Cassens (2013). "Cross-Device Interaction Definition, Taxonomy and Applications". In: cited. IARIA, pp. 35–41. ISBN: 978-1-61208-309-4.

Schwartzman, Helen (1978). *Transformations: The Anthropology of Children's Play.* Plenum Press, New York.

Shneiderman, Ben (1984). "The Future of Interactive Systems and the Emergence of Direct Manipulation". In: *Proc. Of the NYU Symposium on User Interfaces on Human Factors and Interactive Computer Systems.* event-place: New York, USA. Norwood, NJ, USA: Ablex Publishing Corp., pp. 1–28. ISBN: 978-0-89391-182-9. URL: `http://dl.acm.org/citation.cfm?id=2092.2093` (visited on 02/10/2019).

Shum, Hubert and Edmond S.L. Ho (2012). "Real-time Physical Modelling of Character Movements with Microsoft Kinect". In: *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology.* VRST '12. event-place: Toronto, Ontario, Canada. New York, NY, USA: ACM, pp. 17–24. ISBN: 978-1-4503-1469-5. DOI: `10.1145/2407336.2407340`. URL: `http://doi.acm.org/10.1145/2407336.2407340` (visited on 02/10/2019).

Sicart, Miguel (2014). *Play Matters.* Playful Thinking. The MIT Press. ISBN: 978-0-262-02792-2. URL: `http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=826633&site=ehost-live&ebv=EB&ppid=pp_Cover`.

Speicher, Marco, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger (2018). "Selection-based Text Entry in Virtual Reality". en. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18.* Montreal QC, Canada:

ACM Press, pp. 1–13. ISBN: 978-1-4503-5620-6. DOI: `10.1145/3173574.3174221`. URL: `http://dl.acm.org/citation.cfm?doid=3173574.3174221` (visited on 09/21/2019).

Stabbert, Timo (May 2018). “Spielerische Interaktion mit einer Sandkiste in VR”. MA thesis. Bremen: University of Bremen.

Stabbert, Timo, Thomas Fröhlich, Dmitry Alexandrovsky, and Rainer Malaka (2017). “Extending Augmented Sandboxes with Virtual Reality Interaction”. In: *Mensch und Computer 2017 - Workshopband*. Regensburg: Gesellschaft für Informatik e.V., p. 6.

Surale, Hemant Bhaskar, Fabrice Matulic, and Daniel Vogel (2019). “Experimental Analysis of Barehand Mid-air Mode-Switching Techniques in Virtual Reality”. en. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI ’19*. Glasgow, Scotland Uk: ACM Press, pp. 1–14. ISBN: 978-1-4503-5970-2. DOI: `10.1145/3290605.3300426`. URL: `http://dl.acm.org/citation.cfm?doid=3290605.3300426` (visited on 09/21/2019).

Sutton-Smith, Brian (June 2009). *The Ambiguity of Play*. en. Google-Books-ID: AgA8q0TCKeIC. Harvard University Press. ISBN: 978-0-674-04418-0.

Tilger, Jan (Sept. 2018). “Steigerung der Kreativität durch kreative Einschränkungen in kollaborativen Multi-User-Szenarien”. Bachelor Thesis. Bremen: University of Bremen.

Torrance, E. Paul (1981). “Predicting the Creativity of Elementary School Children (1958-80) —and the Teacher Who "Made a Difference"”. In: *Gifted Child Quarterly* 25.2, pp. 55–62. DOI: `10.1177/001698628102500203`. URL: `https://doi.org/10.1177/001698628102500203`.

Trimble Inc. (2013). *SketchUp*. URL: `https://www.sketchup.com`.

Unger, Felix (May 2018). “Performance Animation in VR”. MA thesis. Bremen: University of Bremen.

Unity Technologies (2005). *Unity*.

Valve (2017). *Valve Hammer Editor*. URL: `https://developer.valvesoftware.com/wiki/Valve_Hammer_Editor`.

Valve Developer Community (2020). *Hammer Screenshot (User: Barracuda)*. [Online; accessed March 17, 2020]. URL: `https://developer.valvesoftware.com/wiki/File:Hammer_screen.jpg`.

Vikram, Sharad, Lei Li, and Stuart J. Russell (2013). “Writing and sketching in the air, recognizing and controlling on the fly”. In: *CHI Extended Abstracts*. DOI: `10.1145/2468356.2468567`.

Vogel, Daniel, Paul Lubos, and Frank Steinicke (2018). “AnimationVR - Interactive Controller-based Animating in Virtual Reality”. In: *Proceedings of the 1st Workshop on Animation in Virtual and Augmented Environments*. IEEE. URL: `http://basilic.informatik.uni-hamburg.de/Publications/2018/VLS18b`.

Vogel, Peter (2019). *D1.2 Requirements Definition for Film Productions*. Technical Report D1.2. URL: `http://first-stage.eu/deliverables.html` (visited on 09/21/2019).

Wagner, Julie, Stéphane Huot, and Wendy E. Mackay (2012). "BiTouch and BiPad: designing bimanual interaction for hand-held tablets". In: *CHI*. DOI: `10.1145/2207676.2208391`.

Walther-Franks, Benjamin, Florian Biermann, Nikolaas Steenbergen, and Rainer Malaka (2012). "The Animation Loop Station: Near Real-Time Animation Production". In: Springer Berlin Heidelberg, pp. 469–472. DOI: `10.1007/978-3-642-33542-6_55`. URL: `http://link.springer.com/10.1007/978-3-642-33542-6_55`.

Walther-Franks, Benjamin, Marc Herrlich, Thorsten Karrer, Moritz Wittenhagen, Roland Sch Oder-Kroll, Rainer Malaka, and Jan Borchers (2012). "DRAGIMATION: Direct Manipulation Keyframe Timing for Performance-based Animation". In:

Walther-Franks, Benjamin, Marc Herrlich, and Rainer Malaka (2011). "A Multi-touch System for 3D Modelling and Animation". In: Springer Berlin Heidelberg, pp. 48–59. DOI: `10.1007/978-3-642-22571-0_5`. URL: `http://link.springer.com/10.1007/978-3-642-22571-0_5`.

Wang, Jia and Robert Lindeman (2014). "Coordinated 3D Interaction in Tablet- and HMD-based Hybrid Virtual Environments". In: *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*. SUI '14. event-place: Honolulu, Hawaii, USA. New York, NY, USA: ACM, pp. 70–79. ISBN: 978-1-4503-2820-3. DOI: `10.1145/2659766.2659777`. URL: `http://doi.acm.org/10.1145/2659766.2659777` (visited on 03/14/2019).

Wesson, Bradley and Brett Wilkinson (2013). "Evaluating Organic 3D Sculpting Using Natural User Interfaces with the Kinect". In: *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. OzCHI '13. event-place: Adelaide, Australia. New York, NY, USA: ACM, pp. 163–166. ISBN: 978-1-4503-2525-7. DOI: `10.1145/2541016.2541084`. URL: `http://doi.acm.org/10.1145/2541016.2541084` (visited on 02/10/2019).

Wigdor, Daniel and Dennis Wixon (2011). *Brave NUI World*. ISBN: 978-0-12-382231-4.

Youell, Biddy (2008). "The importance of play and playfulness". In: 10.2, pp. 121–129. DOI: `10.1080/13642530802076193`.

Zielke, Marjorie A, Djakhangir Zakhidov, Gary Hardee, Leonard Evans, Sean Lenox, Nick Orr, Dylan Fino, and Gautham Mathialagan (2017). "Developing Virtual Patients with VR/AR for a Natural User Interface in Medical Teaching". In: Serious Games and Applications for Health (SeGAH), 2017 IEEE 5th International Conference on. IEEE, pp. 1–8.