RWTH Aachen University
Chair of Communication Science
Prof. Dr. phil. Martina Ziefle
Anne Kathrin Schaar, M.A.

Master's Thesis

How Does It Feel?
Presenter Experience and Evaluation While Using Canvas Presentation Tools

How Does It Feel?
Gefühle und Bewertungen des Vortragenden bei Präsentationen mit Canvas Presentation Tools

By
Philipp Wacker
Matriculation Number: 297777
Philipp.Wacker@rwth-aachen.de

Aachen, 07.09.2014

Abstract

Today, many presentations are delivered with the assistance of visual aids. The majority of presentations are authored and presented using specific software such as *Microsoft's PowerPoint*. These software are often called slideware as they employ a slide metaphor originating from technical projectors. This format has been criticized repeatedly for the linear structure and the missing relationship between slides. Canvas presentation tools are a recent development that aim to overcome these problems. Such tools allow the author to arrange the presentation content on a zoomable canvas and enable the presenter to either follow a predefined sequence of viewports or present ad-hoc without such a path. Previous studies have shown that canvas presentation tools assist the author better during the creation of a presentation and that the audience perceives to being able to follow the presentation better. The experience for the presenter, however, has not been evaluated yet.

This thesis presents a study that was conducted to measure the experience of the presenter. Participants gave two presentations—one with slideware, one with a canvas presentation tool—and reported on their feelings using self-report techniques. The results show that the experience is strongly linked to the user characteristics such as prior presentation experience or spatial ability and technological proficiency. While more technological savvy and less experienced presenters feel more pleasure during canvas presentations, the opposite is true for presenters with less technological expertise and more presentation experience. Furthermore, how much a presenter enjoys presenting affects the experience with those liking to present feeling less different between the presentation formats.

Überblick

Heutzutage werden viele Präsentationen mithilfe visueller Hilfen und speziellen Präsentationsprogrammen gehalten. Soche Programme orientieren sich in ihrem Aufbau oftmals an mechanischen Projektoren und folgen einer Folienmetapher, nach der Informationen in sequenzieller Form über mehrere Folien verteilt präsentiert werden. Dieses Format wurde wiederholt kritisiert, unter anderem dafür, dass es keine direkt erkennbare Verbindung zwischen den einzelnen Folien gibt. Ein neuerer Ansatz sind die sogenannten Canvas Presentation Tools, die diese Probleme ansprechen. Diese Programme erlauben es dem Autoren einer Präsentation, die Informationen frei auf einer Leinwand anzuordnen und mithilfe von Animationen einen Pfad durch diese Informationen zu erstellen. Während einer Präsentation hat der Vortragende die Möglichkeit diesem Pfad zu folgen, oder aber sich spontan durch die Inhalte zu bewegen. Vorangegangene Studien haben gezeigt, dass Canvas Presentation Tools den Ersteller einer Präsentation besser unterstützen als folienbasierte Programme und dass die Zuhörer das Gefühl haben, der Präsentation besser folgen zu können. Wie es für einen Vortragenden ist, ein Canvas Presentation Tool zu benutzen, wurde allerdings noch nicht untersucht.

Diese Arbeit stellt eine Studie vor, die durchgeführt wurde, um die Gefühle des Vortragenden während einer Präsentation mit Canvas Presentation Tools festzustellen. Teilnehmer der Studie haben je zwei Präsentationen gehalten, eine mit einem folienbasierten Präsentationsprogram und eine mit einem Canvas Presentation Tool. Ihre Gefühle wurden mithilfe von Selbstbewertungsfragebögen erhoben. Die Ergebnisse deuten darauf hin, dass das Empfinden während einer Präsentation stark von den Vorerfahrungen der Person — wie z.B. der Präsentationserfahrung oder der Technikerfahrung — abhängt. Während technisch erfahrene Personen und Personen, die noch nicht sehr lange präsentieren, sich besser während der Nutzung von Canvas Presentation Tools fühlen, ist das Gegenteil bei weniger technisch erfahrenden Personen und langjährigen Vortragenden der Fall. Ein weiterer Einflussfaktor ist, wie gerne der Vortragende präsentiert. Für jemanden, der gerne präsentiert, ist es egal welches Präsentationsformat genutzt wird.

Contents

List of Figures

Acknowledgements

First of all, I want to thank all participants of my study for taking part and giving these great and diverse presentations. I guess I have now seen almost every possible way the lineup of a soccer team can be presented in.

Furthermore, I thank my advisors, who assisted me with their knowledge and feedback throughout this thesis. A special thanks to Leonhard Lichtschlag, who provided valuable feedback and insights into canvas presentation tools.

Last but not least, I want to thank Mechthild, Peter, Bastian, and Michaela for their constant support while I was writing this thesis.

Chapter 1

Introduction

Before computers, presentation visuals were limited to showing physical slides or transparencies one after another using a slide or overhead projector. Even though personal computers and visual projectors do not require such limitations, several presentation software (e.g., *Microsoft*'s *PowerPoint*[1], *Apple*'s *Keynote*[2] and *Apache*'s *OpenOffice Impress*[3]) follow a slide show metaphor. These presentation software are used in a variate of different contexts like education, business meetings, and many other areas (cf. Good [2003]; Thielsch and Perabo [2012]). However, this format has been criticized multiple times as it would force a presenter to follow it's rules (e.g., Tufte [2006]). A recent development in this regard are the so called *canvas presentations*. Canvas presentations remove the necessity for a linear slide structure partially or completely and enable the author to place content freely on an infinite canvas. During the delivery, presenters can either follow a predefined path through the information, or present ad-hoc by moving manually through the content.

Previous studies regarding canvas presentations focused on the authoring [Lichtschlag et al., 2009] and the reception [Lichtschlag et al., 2012b] of canvas presentations. They found that authors create canvas presentations with more ease and follow better presentation principles. On the reception side, they showed that the audience perceived that they could follow a canvas presentation more easily. However, the person delivering a canvas presentation has not been looked at yet. How does the canvas format affect the presenter? How does the presenter feel while giving such a presentation? These are the questions that are answered in this thesis.

[1] http://office.microsoft.com/en-us/powerpoint/
[2] https://www.apple.com/mac/keynote/
[3] https://www.openoffice.org/product/impress.html

To get an understanding of the different aspects involved in the experience of a presentation, the related literature is looked at first (chapter 2). This review is split into two basic parts. One describes presentations and their different facets (chapter 2.1.1). Furthermore, slideware (chapter 2.1.2) and canvas presentations (chapter 2.1.3) are defined in more detail as well as problems and solutions of each software named. In chapter 2.2, the research on emotions and feelings is looked at. First, it is clarified what emotions and feelings are and what the difference between them is (chapter 2.2.1). Then, ways to measure emotions and feelings are explained in chapter 2.2.2. This is split up into chapters for self-report measures and physiological measurements. Finally, moods are defined and a way presented how they can be measured (chapter 2.2.3).

Chapter 3 describes the study that was conducted to answer the research questions and explore the hypotheses. In this chapter, the general decisions made regarding the structure are described first (chapter 3.1). Chapter 3.2 summarizes the creation of the materials used in the study. This includes a chapter on typical tasks and situations during presentations (chapter 3.2.1) and the description of the presentations held during the study (chapter 3.2.3). The different parts of the study are explained in chapter 3.3. In the final part of the chapter, the hypotheses to be examined are formulated (chapter 3.4).

In chapter 4, the results of the study are laid out. First, the sample of participants that took part in the study is described (chapter 4.1). Following this, the hypotheses are examined (chapter 4.2). Chapter 4.3 reports observations about the style of the presentations that participants gave in the study.

The results of the study are discussed in chapter 5 before the limitations are mentioned in chapter 6. The final chapter (chapter 7) sums everything up and gives an outlook onto possible future research topics.

To increase the readability of this thesis, the forms and instructions of the study are presented in English although the study was conducted in German. The original materials are listed in the appendix. Furthermore, an unspecified third person is referred to as “she” throughout this thesis.

Chapter 2

Literature Review

This chapter provides an overview of the research areas needed to answer the research questions. First, details are given about presentations in chapter 2.1. This includes information on the creation and delivery of presentations in general (chapter 2.1.1) as well as an overview of slideware (chapter 2.1.2) and canvas presentations (chapter 2.1.3). Chapter 2.2 reports on research about the definition of emotions and feelings (chapter 2.2.1) and how they can be measured (chapter 2.2.2). The final part of this chapter defines moods and describes a way to measure them (chapter 2.2.3).

2.1 Presentations

This chapter first explains the various parts of presentations including the role of the presenter (chapter 2.1.1) before detailing slideware and the problems associated with them (chapter 2.1.2). Finally, chapter 2.1.3 explains canvas presentations as a recent development that addresses some of these problems.

2.1.1 Presentation Basics

A presentation is defined as being

> *The action, means, or manner of presenting something to view; [...] Now also: a display or show of information, materials, etc.; a lecture (esp. one illustrated with visual displays)* [OED Online, 2014, Def. 5b].

Such presentations are held in a variety of contexts and often use slide shows as visual aids [Good, 2003, p. 1; Thielsch and Perabo, 2012, p. 116]. While this definition focuses on the *delivery* of a presentation, Lichtschlag [2008] includes three additional aspects of presentations: *research*, *authoring*, and *reuse* [p. 8f.]. Good [2003] compares the task of structuring the speech and the creation of visuals to the process of producing texts [p. 2f.]. The parts of this process cover the same aspects named by Lichtschlag. Based on the analysis of Hunter and Begoray [1990], these parts would be *generating*, *organizing*, *composing*, and *revising* [Good, 2003, p. 3].

Research According to Lichtschlag, the author has to gather knowledge about the topic of the presentation in the *research* phase—if she does not possess this knowledge already from previous presentations or because she is an expert in the field [Lichtschlag, 2008, p. 8]. This corresponds to the *generating* task for the creation of texts [Good, 2003, p. 3].

Authoring The *authoring* of a presentation is explained as the task of structuring the knowledge and also creating visual aids. While the last aspect would be a common task, it would not be imperative as talks could be held without such visuals [Lichtschlag, 2008, p. 9]. Good [2003] mentions that presenting would be the art of "[arranging] ideas in a way that is understandable and holds the audience's interest"[p. 8]. The author would have to assess the expectations and needs of different audiences, while following the "conventions of the presentation format" [p. 4]. In the context of producing texts, *organizing*—order and filter the information—and *composing*—create a usable product—match this aspect [Good, 2003, p. 3].

Delivery The focus of the definition given above lies on the *delivery* of a presentation. This aspect describes the situation during which the pieces of information are actually presented to an audience. As stated by Lichtschlag [2008], this would be a stressful activity for the speaker as it concerns a public speaking situation during which she would be observed by others and the performance probably having an influence on her career [p. 11]. Such situations would often create anxiety, which would be also influenced by situational factors such as it's novelty or prior experience [Beatty, 1988]. This stress would be increased by unforeseen situations (e.g., improvising to adapt to changed time constraints) [Good, 2003, p. 10]. The presenter is also an important factor for the quality of the talk [Thielsch and Perabo, 2012, p. 120] as she not only decides which information to use [Good, 2003, p. 2] but also the style to use to present the information, which has been shown to have an impact on the audience (e.g., Blokzijl and Andeweg [2005]).

Reuse Lichtschlag names *reuse* as an aspect of presentations since the author would often need to create multiple versions of a presentation adapted to different audiences or time restrictions [p. 12] (see also Moscovich et al. [2004]). These versions would differ in the amount of slides used as well as the sequence they are presented in. According to Good [2003], it is also common in the context of text creation to include additional information and adjust the existing structure (the *revising* aspect of text creation) [p. 3].

2.1.2 Slideware

This chapter defines slideware and mentions several points of criticism.

While the first slide shows consisted of fixed images that were shown sequentially using an overhead or slide projector, nowadays PCs and digital projectors are used during presentations [Good, 2003, p. 1]. Over the past years presentation software has been frequently used to assist the delivery of a presentation—especially in the areas of education and business [Thielsch and Perabo, 2012, p. 116]. This kind of software is often called *slideware* (e.g., [Lichtschlag, 2008, p. 1], Farkas [2009]) and with an approximate 96% market share in 2012, *Microsoft*'s *PowerPoint*[1] is by far the most often used program to deliver presentations [Thielsch and Perabo, 2012]. Other software to give slide show presentations are *Adobe*'s *Acrobat*[2] (29%), *Apple*'s *Keynote*[3] and *Apache*'s *OpenOffice Impress*[4] (each ∼10%) [Thielsch and Perabo, 2012].

Each slide of the slideware presentation has the same size and the author has to distribute the information onto different slides. The author is able to style and arrange content (e.g, text, images, or videos) freely for each slide. However, the slides themselves have to be ordered in a linear sequence. During a presentation, the presenter moves sequentially through the slides. She has the ability to advance a slide or go back as well as jump to specific slides in the sequence by using predefined hyperlinks on the slides or by using other slide switching options such as the navigator in *Apple*'s *Keynote* (cf. figure 2.1). This navigator is provided on the presenter screen. Slideware often offers such a presenter screen—a set of information and options that are shown on a secondary display during the presentation. Beside the feature to jump to a specific slide in the presentation, the information on this screen can include images of the current and upcoming slide, time information as well as notes for the presenter.

[1] http://office.microsoft.com/en-us/powerpoint/
[2] http://www.adobe.com/#acrobat
[3] https://www.apple.com/mac/keynote/
[4] https://www.openoffice.org/product/impress.html

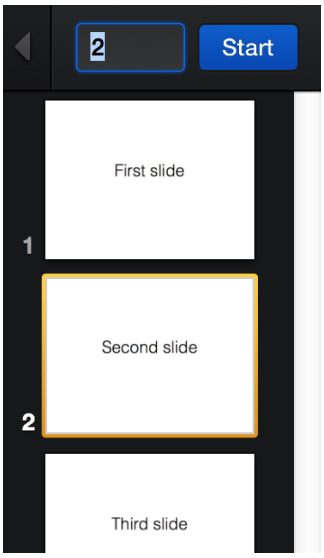


Figure 2.1: Example of the slide navigator in *Keynote*. The presenter can either type a slide number into the text field or click on a specific slide to jump to it.

The linear arrangement of slides originates from *PowerPoint* which began as an 'offline' program that was intended for the creation of transparencies and other presentation materials, and therefore follows the same linear structure as slide shows [Moscovich et al., 2004]. According to Moscovich et al. this linear structure "disregards the way that talks are typically created, edited, and presented in practice". Various authors have criticized this issue as well (e.g., [Tufte, 2006, p. 23]) and Lichtschlag [2008] calls this the problem of "*time dominance*" [p. 45]. If a presenter needs to go to a specific slide, this would present an "unaided linear search" [Good, 2003, p. 10] and would lead to a "rapid-fire clicking" through the slides [Moscovich et al., 2004]. Furthermore, the fixed linear structure of slides would present a problem for reusing the document in different presentation situations as information could not be easily complemented or left out using the same document [Lichtschlag, 2008, p. 45].

Although slide shows would be good in combining visual images with spoken word, Good [2003] criticizes that they would not be good in communicating the structure of a talk [p. 8]. In order to solve this, presenters would use outline or overview slides to present the layout of the talk or visualize connections between different topics. However, Good mentions problems with this approach as it would, for example, require additional effort to not only generate such overviews but also to adapt them once changes occur in the structure of the presentation [p. 8]. Another problem—not exclusively connected to the structure of the talk—would be linked with the complexity of some topics and structures. A great complexity would make it difficult to fit summaries on one slide, and therefore would force

the author to split them between multiple slides [p. 5, 8f.]. Lichtschlag [2008] names this problem "*content cutting*" and it would be the audiences and the presenters new task to connect the slides [p. 43f.]. This problem is especially important since studies and guidebooks suggest to keep individual slides clean and not clutter them with information (e.g., Blokzijl and Naeff [2004]), thereby reducing the possible amount per slide even further.

Another problem mentioned by Lichtschlag [2008] is the "*detail trap*" [p. 45], which is also connected to the missing overviews. Due to the limited amount of space on a computer display, it would be hard to get "global awareness" [Good, 2003, p. 7] of the presentation. Therefore, the presenter would be more likely to focus on details such as "beautify[ing] the individual slide" [Lichtschlag, 2008, p. 45].

An example of a slideware application and the problems mentioned is shown in figure 2.2.

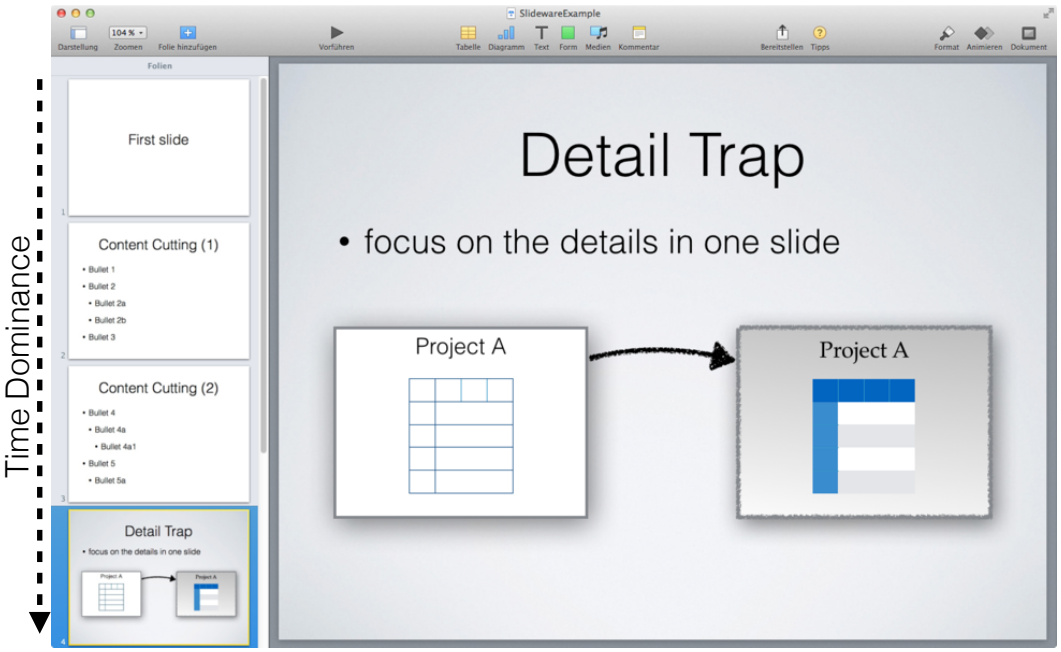


Figure 2.2: The authoring window of *Keynote* as an example of a slideware application and a visualization of the problems mentioned by Lichtschlag [2008].

2.1.3 Canvas Presentations

In order to address the problems mentioned in chapter 2.1.2, a new area of presentation software was developed. This chapter presents canvas presentation tools as a proposed solution to these problems.

Canvas presentation tools belong to the *zoomable user interfaces* (*ZUIs*) that aim at using additional options computers offer [Good, 2003, p. 1]. Some still use the slide metaphor but enable the author to define multiple paths through the slides and give the presenter the possibility to switch between paths during the delivery (e.g., Moscovich et al. [2004]), therefore removing the *time dominance*. Other software abolish the slide concept completely and allow free transformation (e.g., *Fly* by Lichtschlag [2008] or *Prezi*[5]). Here, the author is able to place information directly on a two-dimensional canvas and change the view by animated panning and zooming [Good, 2003, p. 12]. *Canvas presentations* are presentations that were created with this kind of software [Lichtschlag et al., 2012a]. For a presentation, the author can define a sequence of viewports to create one (e.g., *Prezi*) or multiple (e.g., *Fly*) paths. While delivering a presentation, the presenter can then decide whether to follow the path or present freely [Lichtschlag, 2008, p. 47].

According to Good [2003], the authoring process with a canvas presentation tool is similar to working on a whiteboard and it would exploit the human memory for spatial location [p. 12]. This is not only beneficial for the presentation context but authoring in general. *Prezi*, for example, offers a collaboration feature, enabling multiple persons to work simultaneously on a canvas [Laufer et al., 2011]. Furthermore, Good [2003] says that the inputs and corresponding animations simulate real-world actions such as moving a paper on a table, looking more closely at something, or stepping back to get an overview [p. 12]. For Lichtschlag [2008], canvas presentation tools belong to the "organic interfaces", which are interfaces shaped after the users mental model rather than trying to communicate a mental model to the user [p. 17f.].

In Good [2003]'s opinion, canvas presentations enable the user to communicate relationships among information via the positioning of elements, and to communicate hierarchy by zooming in for more details [p. 13]. Furthermore, he states that the structure would become part of the presentation through these meaningful relationships and that overviews would be already available by zooming out [p. 13]. Good also suggests that these overviews would help the presenter since she could do a visual search rather than a linear search when looking for information in reaction to questions [p. 16].

[5] https://prezi.com

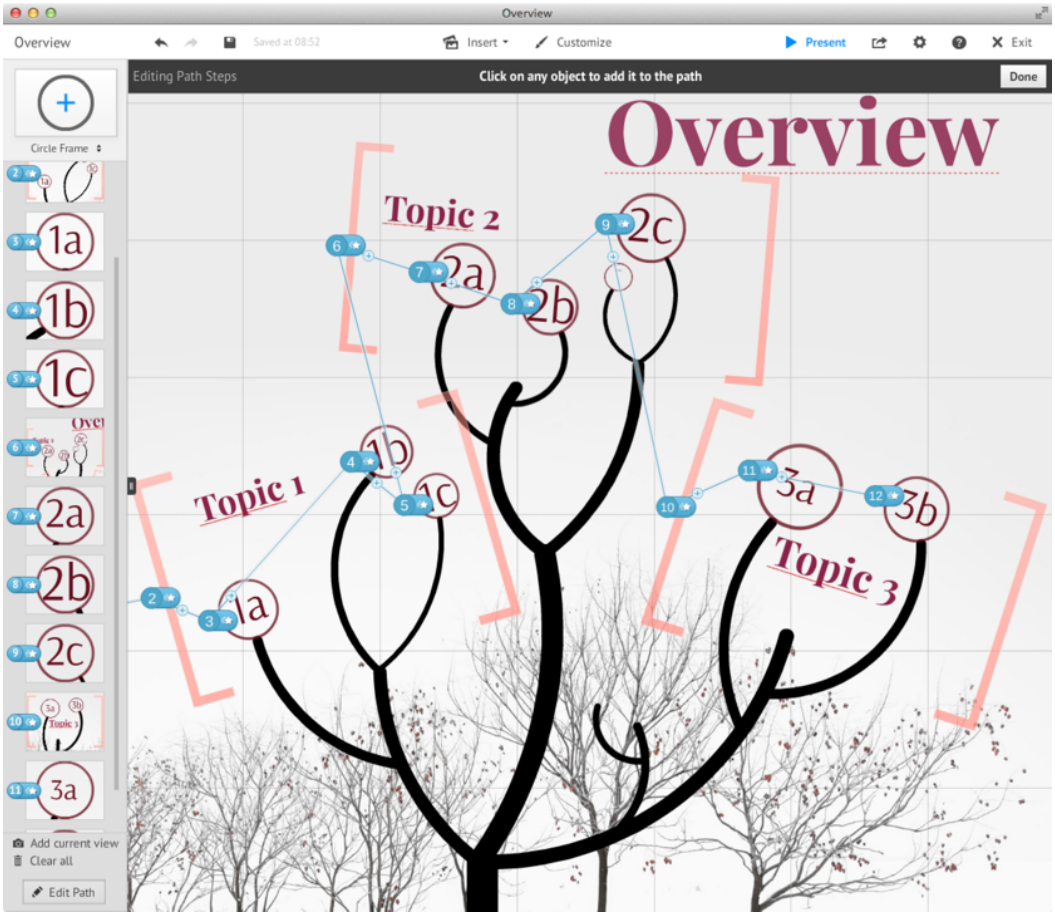


Figure 2.3: The *Prezi* authoring mode as an example of a canvas application. The blue line represents a predefined path through the content. During a presentation, the frames (summarized on the left side) would be brought into focus one after the other. The transition would be animated using panning and zooming animations.

Lichtschlag [2008] investigated the impact that this type of presentation format has during the authoring phase of a presentation. He found out that the canvas presentation tool is a better support to the author than slideware. Furthermore, presenters would arrange content less linearly while also adding overviews more frequently, therefore implementing a better presentation style [Lichtschlag et al., 2012a; Lichtschlag, 2008]. Hess [2011] looked at the delivery phase of a presentation—by examining the audience reception. He discovered that the presentation format (slideware or canvas presentation) had no effect on learning, but that audiences thought that they could follow and understand canvas presentations more easily [Hess, 2011].

Good [2003] mentions several problems associated with zoomable user interfaces such as canvas presentations. According to him, different topics could be presented in a canvas presentation with varying ease. Abstract data (e.g., mainly text-based content) would be hard to arrange spatially [p. 14]. Furthermore, he stresses that the navigation of a zoomable presentation is very important during the delivery as the audience typically sees the navigation [p. 17]. Bemtgen [2012] addressed this issue for *Fly* by providing a locking feature that freezes the content on the projector while allowing the presenter to search on a separate display.

An example of a canvas application and a predefined path through the content is visualized in figure 2.3.

The study presented in this thesis uses both a slideware and a canvas presentation tool to evaluate the differences in experience while giving a presentation with each of these tools (cf. chapter 3).

2.2 Emotions and Feelings

In this chapter research in the area of emotions is reviewed. To begin with, an existing definition of emotions is given and explained how it is connected to feelings (chapter 2.2.1). The question of how emotions and feelings can be measured is examined in chapter 2.2.2. Here, methods are described that use a person's self-reports to assess emotions (chapter 2.2.2 "Self-Reports") before connections are laid out that have been found in the area of physiological measurements (chapter 2.2.2 "Physiological Measurements"). Finally, the measurement of moods is looked at (chapter 2.2.3) since these bear a connection to emotions.

2.2.1 What Are Emotions and Feelings?

Researchers assume that emotions arise not in specific occasions alone, but would be ubiquitous with an emotional state being present at all time [Russell and Mehrabian, 1977, p. 274]. Furthermore, emotions are supposed to influence various aspects of cognition and behavior [Cacioppo and Gardner, 1999, p. 195]. However, the question what emotions and feelings are would be hard to answer. There exists no absolute definition and in fact there are numerous different attempts to formulate such a definition (cf. [Scherer, 2005, p. 696; Bradley and Lang, 2000, p. 242]).

According to Mauss and Robinson [2009], a favored model to describe *emotion* would be the *componential model*, which states that emotional reaction depends on

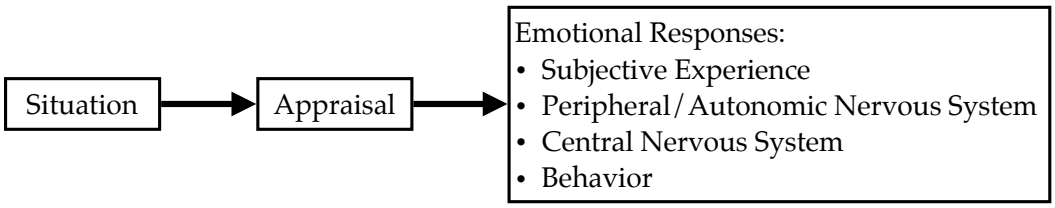


Figure 2.4: Componential model of emotional responding (after Mauss and Robinson [2009]).

the personal significance a person attributes to a specific event before triggering a response "involving subjective experience, physiology, and behavior" [p. 209f.] (cf. figure 2.4). Three components (motor expressions, bodily reactions, and subjective experience) are said to have a "long-standing status as modalities of emotion" [Scherer, 2005, p. 698]. However, different authors would vary in what components they include in the model (e.g., Scherer includes a "cognitive, information processing component" [Scherer, 2005, p. 698]).

Another point of argument is the combination of these components. Several authors propose that different components would combine to certain "basic-emotions" (e.g., Ekman [1992]), while others follow a multidimensional approach in which an emotion would be described by various combinations of dimensions (e.g., Russell and Mehrabian [1977]) [Frijda, 2008, p. 76]. Keltner and Ekman [2000] say that this is "a central question in the field of emotion" [p. 237] and some authors try to reconcile the two approaches by either distinguishing between momentary experiences (discrete emotions) and accumulated emotions (dimensional) [Keltner and Ekman, 2000, p. 239], or by proposing that a basic emotion would correspond to a combination of several dimensions [Mauss and Robinson, 2009, p. 211]. During their review on measurement methods of emotion, Mauss and Robinson [2009] come to the conclusion that the dimensional approach would be favored as more studies found connections that support a dimensional concept [p. 226f.].

The number and distinction of dimensions is not clearly defined as well. Mauss and Robinson [2009] mention *valence*, *arousal*, and *approach-avoidance* as the most common dimensions [p. 210]. The valence dimension would distinguish between states of pleasure (e.g., extreme happiness) and displeasure (e.g., extreme pain)—therefore, it is sometimes also referred to as the *pleasure* dimension (e.g., Russell and Mehrabian [1977]). The arousal dimension would range from states of low arousal (e.g., sleepy) to states of high arousal (e.g., high alert). According to Lang et al. [1993], these two dimensions explain the main variance of emotion [p. 261]. Scherer [2005] states that there are problems in naming a consistent third dimension, which would be the reason why valence and arousal are mostly used as dimensions [p. 718].

Cacioppo and Gardner [1999] name *"positivity"–"negativity"* as the broadest category in which to distinguish emotions and they assume that this originates from the physical options to *approach* an object or *withdraw* from it [p. 200]. Furthermore, the authors found two underlying influences on this dimension: a "positivity offset" and a "negativity bias". While the positivity offset describes a tendency for a mildly positive attitude towards objectively neutral stimuli—expected to motivate an organism to approach new situations instead of staying back—the negativity bias describes a tendency to react more intensely to negative stimuli—presumed to result from the higher impact and potentially lasting effects of negative results (e.g., death) [Cacioppo and Gardner, 1999, p. 205f.].

Furthermore, it is not clear how the dimensions—or the different directions—are related. For the example of "positive"–"negative", researchers disagree on whether these are inversely related—meaning that more positivity would directly relate to less negativity—or if they are independent—there would be no connection between positivity and negativity ratings [Mauss and Robinson, 2009, p. 210; Cacioppo and Gardner, 1999, p. 204].

In the componential model the term *feeling* describes the subjective experience of an emotion. Researchers found that emotional reactions can happen without subjective feelings (e.g., [Craig, 2008, p. 273; Frijda, 2008, p. 82]). However, the role of the conscious feeling and its use would remain unclear [Frijda, 2008, p. 82], with some researchers seeing them as a benefit for "emotional communication" [Craig, 2008, p. 273].

In the context of this thesis, the dimensional approach to emotions with the dimensions valence and arousal is chosen. On the one hand, these dimensions allow to classify an emotion (and therefore a feeling) according to the experienced pleasure and the intensity of the feeling. On the other hand, there are several ways to measure these dimensions readily available. The next chapter provides an overview about some of them.

2.2.2 How Can Emotions and Feelings Be Measured?

Measuring the emotion of a person is a "vexing" [Mauss and Robinson, 2009, p. 209] problem and research is done in a lot of different areas. These areas include facial expressions (e.g., Keltner and Ekman [2000]), vocal expressions (e.g., Bachorowski and Owren [2008]), and other bodily reactions (e.g., Larsen et al. [2008]). In this chapter two approaches to measure emotions and feelings are described—self-reports and physiological measurements.

According to Robinson and Clore [2002] self-reports are "the most common and potentially the best [...] way to measure a person's emotional experiences" [p. 934]. Two self-report techniques are explained: the *semantic differential* used by Osgood et al. [1957] and the *self-assessment manikin* by Bradley and Lang [1994]. Chapter 2.2.2 "Physiological Measurements" gives an overview about the different physiological measurements of emotion.

Self-Reports

According to Larsen et al. [2008], self-reports do not have a strong connection to the bodily and behavioral aspects of an emotion [p. 181]. But as feelings are defined as being the subjective experience aspect of an emotion (cf. chapter 2.2.1), Scherer [2005] declares that there would be "no access other than to ask the individual to report on the nature of the experience" [p. 712]. Mauss and Robinson [2009] come to the conclusion that self-reports would work best with the dimensional approach of emotion [p. 213].

Robinson and Clore [2002] and Clore and Robinson [2012] explored the validity of self-reports, and they describe that time would be an important factor. They report that people would access four different types of knowledge when giving an account of their emotion. As stated by Robinson and Clore [2002], these types of knowledge are (in relation to the report of feelings):

- *Experiential knowledge*: The reported feeling would be a direct interpretation of the current experience (e.g., "I am afraid.").
- *Episodic memory*: The reported feeling would be reconstructed by recalling contextual details (e.g., "I did not want to raise my arms in the roller-coaster, therefore I was afraid.").
- *Semantic memory*: The reported feeling would either be influenced by *situation-specific belief*—by what a person thinks would be felt in a situation (e.g., "People are afraid in roller-coasters, therefore I was afraid.")—or by *identity-related belief*—by beliefs about oneself (e.g., "I am easily afraid, therefore I was afraid."). [Robinson and Clore, 2002, p. 934f.]

Robinson and Clore state that these types might create different accounts of a feeling, and therefore it would be important to know which source of knowledge was used in order to report the feeling [p. 935].

They believe that the recall follows three principles: *relative accessibility*—a source of information would have to be present at the time—*dominance of the most specific source of information*—experiential knowledge would be preferred to episodic memory which would be chosen over semantic memory—and *evanescence*—which means that it would not be possible to store experiential knowledge and that episodic memory (the contextual details) would fade quickly [p. 937].

Therefore, Robinson and Clore [2002] come to the conclusion that "any delay between an experience and its report necessarily means a loss of information"[p. 935]. Furthermore, while the episodic memory would be tied to a specific event from the past, semantic memory would consist of generalized information [p. 935]. According to Clore and Robinson [2012], this might lead to reports in which the described feeling would not reflect the actual experience [p. 205]. Clore and Robinson assume that the episodic memory (the contextual details) are stored for approximately two weeks [p. 199] but that cognitive tasks would shorten this storage [p. 201]

Based on these considerations, Robinson and Clore [2002] discuss the validity of different types of self-reports. They distinguish between *online reports*—for which experiential knowledge is available—*retrospective reports*—which are deduced from episodic memory, therefore accuracy would decline over time—*prospective reports*—which would lack episodic memory as they take place in the future, thereby supposedly being influenced by beliefs—and *time-inclusive reports*—which average over a time frame and would lead to more semantic recall [p. 938f.].

Next, two types of self-report are described in greater detail—the *semantic differential* technique by Osgood et al. [1957] and the *self-assessment manikin* by Bradley and Lang [1994].

The Semantic Differential Osgood et al. [1957] used the *semantic differential* technique to measure the meaning of words and this is "among the best known research bearing on the centrality of people's net positive and negative feelings" [Cacioppo and Gardner, 1999, p. 203].

A semantic differential consists of a set of point rating scales with bipolar word pairs on either end of the scale (e.g., *good–bad*). Osgood et al. [1957] propose a seven-point rating scale with end points X and Y to which the seven rating steps correspond: extremely X—quite X—slightly X—neither X nor Y; equally X and Y—slightly Y—quite Y—extremely Y [p. 28f.] (cf. figure 2.5 *top*). A person rates a concept by indicating for each word pair where she places the concept between the extreme poles, thereby creating a profile for the concept (cf. figure 2.5 *bottom*). Osgood et al. state that a person would express both the "direction" as well as the "intensity of each judgement" using this method [p. 20]. Furthermore, it would

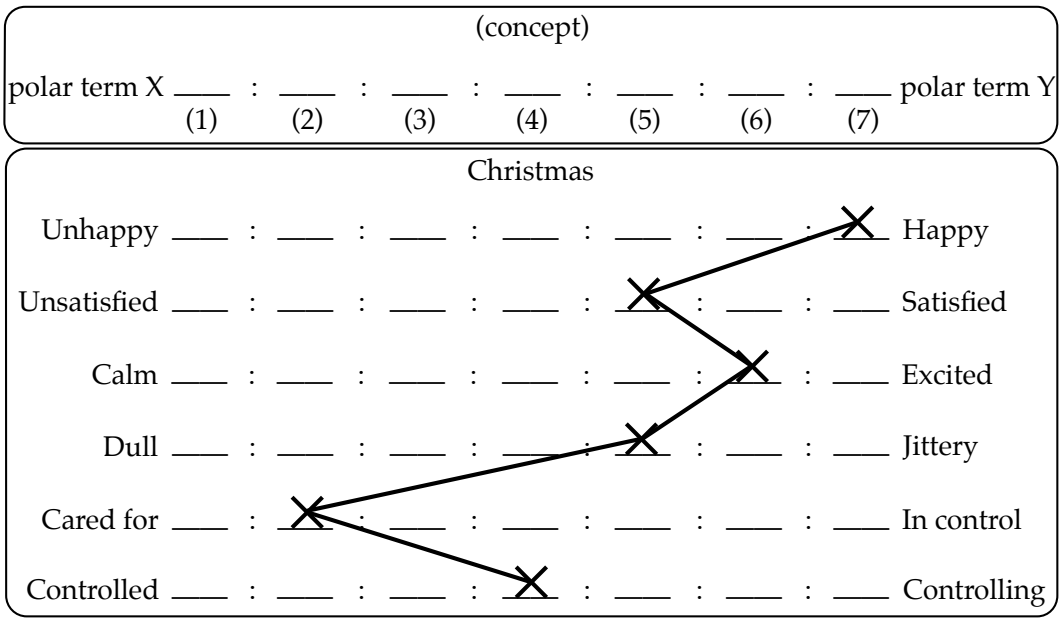


Figure 2.5: *Top*: General design of a semantic differential scale (adapted from [Osgood et al., 1957, p. 28f.]). The participant has to mark where she places the concept between the end points. The numbers correspond to extremely X(1)—quite X(2)—slightly X(3)—neither X nor Y; equally X and Y(4)—slightly Y(5)—quite Y(6)—extremely Y(7). *Bottom*: Example rating of the concept "Christmas" on six scales.

remove the "encoding fluency" by asking people to use certain words as a rating in contrast to using free formulated answers [p. 19]. The selection of terms for the end points of the scales is very important according to Osgood et al. These words would need to be "as representative as possible of all the ways in which meaningful judgments can vary, and yet small enough in size to be efficient in practice" [p. 20].

The results from Osgood et al.'s study suggest that 50% of variance in the judgments could be explained by three factors, which they named "evaluation", "activity", and "potency" (cf. [Bradley and Lang, 1994, p. 49]). Russell and Mehrabian [1977] used the semantic differential technique in the context of emotion. They conducted two studies using 18 nine-point semantic differential-type items. From their results they inferred that three dimensions underly emotional experience. These dimensions would be dimensions of "pleasure–displeasure", "arousal–nonarousal", and "dominance–submissiveness" [p. 291]. While the pleasure and arousal dimension would not share great variance, the dominance dimension covaries with the pleasure dimension [p. 291]. Nevertheless, Russell and Mehrabian concluded that dominance would be necessary to distinguish between different states of emotion (e.g., "angry from anxious")[p. 291].

The semantic differential technique has also been used to measure people's feelings. Evans [1970], for example, measured the feelings of students during math lessons and found that the instruments were sensitive enough to detect changes that occurred during these lessons [p. 105].

The Self-Assessment Manikin Bradley and Lang [1994] created a faster method to quantify the valence, arousal, and dominance dimension, which they call the *self-assessment manikin* (*SAM*). For this they use a "non verbal pictorial assessment technique"[p. 49], which would have the added benefit of not requiring the participant to speak a certain language.

For each dimension Bradley and Lang created graphic depictions that vary along the dimension. For the valence dimension, images range from a sad person who has the corners of the mouth turned down to a happy person. The arousal dimension is depicted by the change from an eyes closed, no experienced feelings person to an eyes wide open person having intense feelings. The dominance dimension is represented by a drawn person evolving from small to big (cf. figure 2.6). In their pen-and-paper version, Bradley and Lang used five images but allowed the participants to mark between the images as well, resulting in a nine-point rating scale for each dimension. They also provided a computer version that used 20 steps [p. 51f.].

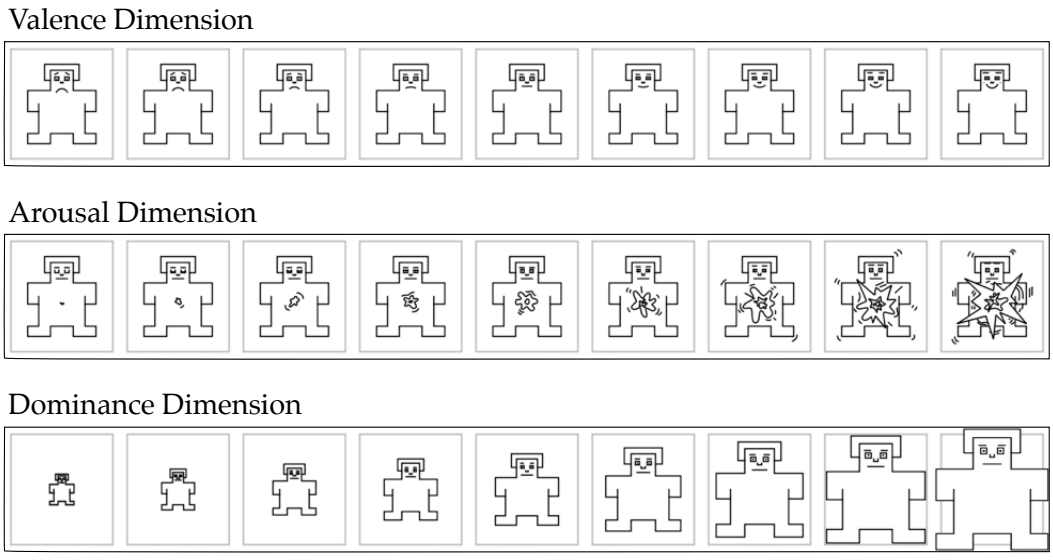


Figure 2.6: Graphic depictions of the self-assessment manikin. Participants should mark for each row, which image visualizes their emotion best (images taken from Irtel [2007]).

In order to validate their method, Bradley and Lang compared ratings from both the pen-and-paper version and the computer version to the dimensional results obtained via the semantic differential technique [p. 52f.]. Their semantic differential ratings had the same structure as the ones from Russell and Mehrabian [1977] and resulted in the same three underlying dimensions—pleasure, arousal, and dominance [Bradley and Lang, 1994, p. 53]. They found that the pleasure and arousal dimensions had high correlations with the respective SAM-scores—97% agreement for the pleasure and 94% agreement for the arousal dimension—while the dominance dimension showed less agreement (23%).

Bradley and Lang conclude that the self-assessment manikin would be a good way to assess a person's emotional reaction. Scherer [2005], however, points out that the focus on such dimensions alone would miss information about "the type of event that has produced the emotion and the appraisal process underlying the responses"[p. 718].

The study conducted in this thesis, therefore, uses both the semantic differential and the self-assessment manikin (cf. chapter 3.1). While the self-assessment manikin provides an easy measurement for the dimensions of a feeling, the semantic differential is supposed to present more details for the aspects of the feeling.

Physiological Measurements

In this chapter an overview of the research that links physiological reactions to the current emotions is presented. For this purpose, findings about the various parts of the human nervous system are reported, before a summary at the end of the chapter is provided.

Many researchers try to measure emotion by observing and recording physiological reactions (e.g., Bradley and Lang [2000]; Larsen et al. [2008]). This research focuses on different components of the nervous system: The *central nervous system* (e.g., the brain), the *autonomic nervous system* (e.g., responsible for changes in the heart rate) and the *somatic nervous system* (e.g., facial muscles) [Larsen et al., 2008, p. 181]. A lot of work in this area has been executed by Bradley and Lang [2000], who used the self-assessment manikin and physiological measurements to find correlations.

The Autonomic Nervous System According to Mauss and Robinson [2009], describing a connection between reactions of the autonomic nervous system (ANS) and emotional responses would not be easy, as it would often not be clear whether a response in the ANS shows an emotional reaction or a reaction to something else [p. 213]. The other direction would not be definite as well as "emotional stimuli do not invariably evoke reciprocal activation [...] of the autonomic nervous system" [Larsen et al., 2008, p. 189]. Therefore, Mauss and Robinson [2009] state that it would be difficult to deduct an emotion from any ANS response [p. 215].

While the connection of ANS responses to discrete emotion has not produced reliable results (e.g., Larsen et al. [2008]; Mauss and Robinson [2009]), researchers found a connection to different dimensions. Larsen et al. [2008] identified a connection between the activation of the ANS and the positive–negative dimension of emotion [p. 184], and Lang et al. [1993] also found a correlation between peak cardiac acceleration and the pleasure dimension [p. 265]. In their study 61% of the participants showed a positive correlation, with 9% being significant [Lang et al., 1993, p. 265]. Furthermore, Lang et al. report a positive relationship between skin conductance response and the arousal dimension. Here, 87% of the participants showed this relationship, with 33% being significant [p. 265].

The Somatic Nervous System A lot of research attempts to link emotional states to reactions from the somatic nervous system (SNS) (e.g., Keltner and Ekman [2000]; Bradley and Lang [2000]). According to Keltner and Ekman [2000], facial expressions show considerable connections to emotions and would work well with the discrete definition of emotions [p. 239] (also [Larsen et al., 2008, p. 187]). An interpretation of a facial expression would be done via "componential coding" [Mauss and Robinson, 2009, p. 223].

Another approach consists in measuring responses of specific muscles directly using electromyography-sensors (EMG-sensors) (e.g., Larsen et al. [2008]; Lang et al. [1993]). The results of such studies suggest that facial muscle response is connected to the valence dimension of emotion [Larsen et al., 2008, p. 189]. Lang et al. [1993] found that activity above the brow (from the *corrugator supercilii*) was inversely correlated to the pleasure ranks—81% of their participants displayed this correlation, 52% of them significant [p. 263]. Activity of a different muscle (the *zygomatic major*) exhibited a positive correlation—72% of their participants showed this behavior, 52% of them significant [p. 263].

However, Mauss and Robinson [2009] express that other effects—such as an audience—might decrease facial reactions. Therefore, the "absence of changes in facial behavior should not be equated with the absence of an emotion, and vice versa" [p. 226].

The Central Nervous System Many researchers assume that physiological reactions in the brain would lead to a better measurement of emotion [Mauss and Robinson, 2009, p. 217], and findings suggest that negative and positive emotional processing involves the left and right hemispheres differently. Negative emotions would be connected to more activity in the left hemisphere, while the opposite would be true for positive emotions [Larsen et al., 2008, p. 188]. Furthermore, frontal activity appears to be linked more to approach-withdrawal motivation than to the positive-negative dimension of emotion [Larsen et al., 2008, p. 189]. During their review, Mauss and Robinson [2009] found out that several different areas also display a connection to the approach-avoidance dimension and conclude that a combined evaluation of different areas might lead to better estimations of an emotion [p. 221].

Summary Studies and reviews of measurements of emotion came to the conclusion that there would be no "gold standard" of measuring emotional responding [Mauss and Robinson, 2009, p. 228; Scherer, 2005, p. 709]. Actually, Mauss and Robinson [2009] conclude that "correlations among multiple measures of emotion are moderate at best, small in typical studies, and inconsistent across studies" [p. 227]. While self-reports of emotion are said to be a good—and according to Scherer [2005], the only—way to assess the subjective experience of a person (her feelings), facial muscle responses show correlations with the valence, and skin conductance responses display correlations with the arousal dimension of emotion (cf. Mauss and Robinson [2009]).

As the focus of this thesis lies on the feelings of a participant, physiological measurements take a minor role. Skin conductance response as a way to assess the arousal, and EMG measurements of facial muscles for the valence dimension are considered for the conducted study (cf. chapter 3.1).

2.2.3 Measuring Moods

Below, a brief summary of moods in general is provided as well as a way to measure them.

While emotions are believed to be directed at specific events, moods are believed to be more diffuse [Larsen et al., 2008, p. 181]. Scherer [2005] describes them as long lasting underlying subjective feelings that influence a person's actions and experience [p. 705]. Furthermore, he states that moods would be less intensive than emotion [p. 702].

Similar to emotion, mood can be measured on different dimensions. According to Watson et al. [1988], *positive affect*—"the extend to which a person feels enthusiastic, active, and alert"—and *negative affect*—"dimension of subjective distress and unpleasurable engagement"—are often used to describe moods [p. 1063]. To provide a valid and reliable way to measure moods especially over periods of time, Watson et al. developed the *Positive and Negative Affect Schedule* (*PANAS*). This test consists of 20 words that describe different feelings. While 10 aim to measure the positive affect (e.g., *interested*, *strong*), the other 10 intent to measure the negative affect (e.g., *nervous*, *irritable*) [p. 1070]. For each word the participant is asked to rate the extend to which she has experienced the feeling over a specified period of time. This time can range from the current moment, over weeks, to general feelings. Watson et al. [1988] compared the results from PANAS to several other measurements of mood factors and concluded that their scales provide an efficient, reliable and valid way to assess the mood of a person [p. 1069].

The PANAS test is used in the presented study, since a study that aims to evaluate the influence of stimuli on feelings, needs to address the possibility that preexistent moods affect the reaction to the stimuli.

Chapter 3

Planning and Execution of the Study

In this chapter the study conducted to answer the question of how a presenter feels during a canvas presentation is described. Decisions regarding the structure of the study are explained in chapter 3.1, during which a broad overview about the different parts is given as well. Following this, the preparation for the experiment is outlined (chapter 3.2). This includes finding typical tasks for the presenter as well as typical situations that occur during presentations (chapter 3.2.1). Chapter 3.2.2 describes how word pairs for the semantic differential were obtained and what the final scale looks like. The structure and creation of presentations, which the participant held during the experiment, is explained in chapter 3.2.3. Further questionnaires to measure the spatial ability, demographic information, and previous knowledge of participants are presented in chapter 3.2.4. Chapter 3.3 lays out the script that was used during the study. This includes the introduction (chapter 3.3.1), during which the structure of the study was explained to the participant as well as her mood of the day measured. In chapter 3.3.2 the main part of the study is described. Here, the participant held presentations and reviewed the presentations together with the moderator. During this review, she was asked at certain times to fill out forms to report her feeling in that situation. The chapter about the conclusion of the study (chapter 3.3.3) includes details about the spatial ability test used and a questionnaire about the participant and her previous knowledge. In preparation for the evaluation, the gathered data was structured after each study. This process is explained in chapter 3.3.4. Finally, the hypotheses this study wants to explore are formulated in chapter 3.4.

3.1 General Structure

The first question that needed to be answered was whether to conduct a lab or a field study. As seen in 2.1.1 "Presentation Basics", presenting is a stressful activity. Furthermore, technological problems can happen during presentations (cf. 3.2.1 "Typical Tasks and Situations during Presentations"). Since it was planned to assess the feelings in specific situations and compare them between participants and presentation software, it had to be guaranteed that these situations occur. Therefore, it was decided to conduct the study as a lab study during which a participant had to give demo presentations with each of the two presentation formats. By choosing this format, it was possible to include the interesting tasks into the presentations as well as imitate technological problems.

One important disadvantage of this setup is the difference to a real presentation situation. If the participants do not feel as if they are giving a real presentation, it would not be possible to deduct any meaningful information regarding the feelings while giving canvas presentations. Kern et al. [1983] indicate that specific instructions lead to good role-playing results, while Higgins et al. [1979] also found out that role-playing leads to experiences similar to the real situation. Therefore, the participant was told to remember her last presentation and imagine she would be presenting in the same setting. In order to confirm the assumption that this creates feelings similar to a real presentation, a question after the study was added in which the participant was asked how the feelings during the demo presentations compared to feelings during real presentations. As this setup matches a *low-intensity* condition (cf. [Beatty and Behnke, 1991, p. 159]) it was expected that the participant experienced similar albeit muted feelings compared to real situations.

To further enhance the presentation feeling, two cameras were used to record the presentations. One of the cameras was not necessary for the evaluation but was used to give the participant the impression of being observed. The other camera, however, recorded the presentation and this recording was used to question the participant about her feelings. In order to do this, the presentation was watched together and the moderator stopped the recording at specific times. Even though the memory of feelings declines over time (cf. chapter 2.2.2 "Self-Reports"), this setup allowed to query the feelings after the presentation, making it unnecessary to interrupt the presentation. The flow of the presentation was therefore not broken and the participant was able to carry out the presentation undisturbed.

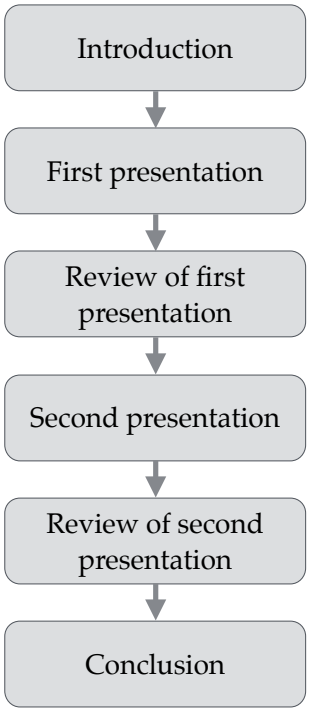


Figure 3.1: General plan of the study. Each participant gave both a canvas and a slideware presentation. The order of them was balanced.

Subsequent to the presentation, the participant was able to relive the situation and assess the feeling both by remembering and by seeing the situation on the recording—thereby reconstructing the contextual details than can help to remember the feeling (cf. chapter 2.2.2) and experience it again (cf. [Scherer, 2005, p. 700]). People are good in judging the emotions of a person they see—facial expressions for example trigger specific responses in observers [Keltner and Ekman, 2000, p. 239]—so this adds to the assessment of the situation. The sequence of the study with the various parts is visualized in figure 3.1.

A scale similar to the semantic differential was used as well as the self-assessment manikin to measure the feelings in a situation (cf. chapter 2.2.2 "The Self-Assessment Manikin"). While the self-assessment manikin was supposed to assess the broad classification of the experienced feelings, the semantic differential should provide greater detail about the different aspects of these feelings. Additionally, the skin conductance of the participant was recorded to cross-evaluate the response with the arousal rating (cf. chapter 2.2.2 "Physiological Measurements"). It was planned to measure the muscle activity of the *corrugator* as well, but the used sensor did not provide a resolution high enough to detect changes.

3.2 Preparing the Materials

This chapter describes the different tasks and problems that needed to be solved in order to create the study materials. This includes research about the typical tasks a presenter faces during presentations (chapter 3.2.1). Finding these tasks was important, since the participant should be confronted with the same situations that would happen in a real presentation. Chapter 3.2.2 lays out how the pairs for the semantic differential were obtained and how the final questionnaire looks like. The different presentations used in the study are described in chapter 3.2.3. This includes information about the topic and how the different tasks were incorporated. The final chapter (chapter 3.2.4) shows additional materials that were used to measure the participant's mood, her spatial ability, as well as the questionnaires to record her demographic information and previous experience concerning presentations.

3.2.1 Typical Tasks and Situations during Presentations

This chapter details how a list of tasks and situations that the participant should encounter during the study was created. In order to find these typical tasks and situations, the options of existing software were analyzed on one hand and further ideas were brainstormed on the other hand. This allowed to formulate a set of actions that the participant should perform during the presentation as well as to create a setup for situations that the participant should encounter.

To get a first overview over possible tasks for the presenter, the interaction options that existing presentation software offer during a presentation were looked at. Shortcut pages were checked for *PowerPoint*[1], *Keynote*[2], and *Prezi*[3]. Furthermore, the shortcuts that *Fly* provides were reviewed. Having this information, the feature set of these software was compared and the actions that are possible by all software were extracted. The remaining actions were:

- Next slide/frame
- Previous slide/frame
- Go to specific slide/frame

[1] http://office.microsoft.com/en-us/powerpoint-help/use-keyboard-shortcuts-to-deliver-your-presentation-HA102749078.aspx

[2] https://help.apple.com/keynote/mac/6.2/?lang=en#/tan951def1c9

[3] https://prezi.com/support/article/creating/keyboard-shortcuts/?lang=en

Other actions were mainly subgroups of these (e.g., *PowerPoint*'s 'go to last viewed slide') or interactions with items on the presenter screen (e.g., *Keynote*'s 'reset timer'). Since this thesis targets the feelings during canvas presentations, it was decided to put a bigger focus on the features offered in those presentations and use a slideware application that provides comparable interactions. These main features are the possibility to arrange the content spatially and move freely through the information (cf. chapter 2.1.3). A question was added to the post-study interview asking for features that the participant missed during the presentation as this setup excludes features like the presenter screen.

To find fitting scenarios for the situations that the participant should encounter, a brainstorming session with frequent presenters was conducted. This came to the conclusion that a successful presentation only consists of going forward (showing the next slide/frame). This has also been stated by Moscovich et al. [2004]. Only outer influences trigger a change in this behavior. As these influences were named:

- Questions from the audience about the content, requiring a search (linear for slideware, visual for canvas presentations (cf. Good [2003]))
- Technological difficulties (e.g., 'freezing' of a presentation)
- Time pressure

These findings were combined to create a list of situations for which the feeling of the participant should be measured. These situations were:

- Going a step forward (*Step*)
- Going back to a specific point in the presentation, therefore simulating questions from the audience. This was split up into two categories:
  - Searching a known location (*Search known*). This simulates a question for which the presenter knows where to go to display the answer.
  - Searching an unknown location (*Search unknown*). This simulates a question where the presenter has to find the answer in the presentation.
- Skipping some content of the presentation (*Skip*)
- Experiencing technological difficulties—imitated by two different problems:
  - *"Wrong input"-error*: This simulates a misinterpreted input or an unexpected reaction after an input.
  - *"Freezing"-error*: This simulates a situation during which the program appears to be unresponsive.

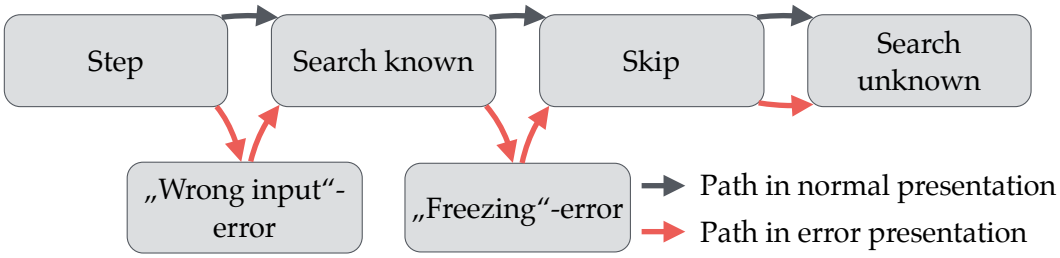


Figure 3.2: Sequence of situations in the presentations. Each participant held two presentations—one with simulated errors and one without.

The sequence of situations is visualized in figure 3.2 and chapter 3.2.3 “Description of the Presentations” elaborates on how the presentations were created to feature these situations.

3.2.2 Creating the Semantic Differential

This chapter describes how the semantic differential scale that was handed out to the participant to specify her feelings was created.

A list of categories that could be used to describe feelings during a presentation was needed to create this scale. In order to find these categories, a survey was created with *Google Forms*[4] and was sent around both at the *chair of communication science*[5] and the *chair of media computing and human-computer interaction*[6] of RWTH Aachen University. The participation was anonymous and every entry was automatically stored in a table for the evaluation. Only the timestamp of the entry was recorded beside the answers to the questions.

The survey was basically a wording assignment to list possible feelings and reactions as well as descriptions for situations during presentations. This assignment was split up into three questions such that various aspects could be targeted independently. As this was an exploratory survey, it was not tied to a specific language as the final semantic differential. Therefore, both an English and a German version were created. Both of them are attached to this thesis (cf. appendix A “Wording Survey”, figures A.1, A.2).

[4]http://www.google.com/google-d-s/createforms.html
[5]http://www.comm.rwth-aachen.de
[6]http://www.hci.rwth-aachen.de

The three questions aimed to elicit feelings or descriptions from different areas. The first question—*"When looking back to past presentations: What feelings and reactions have you experienced yourself during presentations?"*—aimed to tap into remembered feelings of the participant. With the second question—*"What feelings and reactions of a presenter can you imagine in general?"*—it was planned to gather feelings that the participant did not experience themselves but maybe heard about, seen other presenters feel, or could imagine that they could occur. Examples (*"Angry"*, *"Proud"*) were also given to trigger ideas and clarify the question. The third question—*"What other criteria to describe situations can you think of?"*—asked for more general criteria to describe situations (e.g., *"Positive – Negative"*). These three questions were used to guide the participant to think about different directions where ideas about feelings and descriptions can originate from. After sending out the survey, participants had about two weeks to answer the questions (multiple entries were encouraged) before the results were evaluated.

Overall, answers of varying length from 12 participants were received. To get to the word pairs for the semantic differential, the results were filtered in multiple steps. In the first step every individual feeling, reaction, or description was noted. Then, those naming the same concept were grouped into stacks. Using this overview, the groups that did not describe feelings, reactions, or descriptions, but influences from the environment (e.g., *"audience"*) or coping techniques (e.g., *"telling a joke"*) were eliminated. A representative from the remaining groups was then used for the further evaluation.

With these representatives, an affinity diagram (cf. [Baxter and Courage, 2005, p. 714]) was created to find an underlying structure. This resulted in a positive-negative distribution that was used to further filter the answers. Based on this overview, further post-its were eliminated from the board:

- Items that describe the talk (e.g., *"flat"*, *"fruitful"*)
- Items that describe feelings that occur before or after giving the talk (e.g., *"relief that everything is over"*)
- Items that describe a modulation of other items (e.g., *"minor panic"*)
- Items that belong to thematically similar topics (e.g., *"anger"* and *"frustration"*, *"embarrassment"* and *"shame"*). Only one was kept for each group.

Another affinity diagram was created with the remaining items that resulted in word pairs and similar topics. Using this diagram, the word pairs for the semantic differential were selected.

If the evaluation already produced two words that were opposites of each other (e.g., *"negative" – "positive"*) those were used directly. Otherwise, the words were scaled with *"very"* and/or *"not"*. The resulting word pairs were:

- very nervous – not nervous
- very surprised – not surprised
- not confident – confident
- unpleasant – pleasant
- negative – positive
- afraid – not afraid
- unsatisfied – satisfied
- sad – happy
- stressed – not stressed
- desperate – not desperate

Two more pairs were added that aim towards canvas presentation tools and were mentioned during the brainstorming sessions:

- controlled – controlling
- lost – not lost

These pairs represent a deviation from the semantic differential described in chapter 2.2.2 since the zero point is not always fixed on the middle value. For example, the pair *sad – happy* has the neutral value in the middle while *not confident – confident* has the zero value on the left side. This setup was used since it did not require an additional survey to find the semantic opposites, while providing easy to understand concepts to describe feelings. The intention here was to distract the participant as little as possible so that she could focus on describing her feelings. The positive and negative sides of the differential were not flipped for the same reason and it was expected that the moving neutral value balanced the voting tendency.

After rating all situations (cf. chapter 3.2.1) the participant was asked to rate the experience and the program overall. Two further word pairs were added for these ratings that were a result of the survey but did not exactly describe feelings of the participant. These word pairs were:

- unnatural – natural

- not familiar – familiar

The reason why these pairs were not included in the semantic differential for the situations is because they do not describe how a person feels (A person does not feel *unnatural* or *familiar*) but how a person feels *about* something (Using a program can feel *familiar*). So these two word pairs do not describe the feeling of the participant directly but the feeling about a situation or the program. Therefore, these pairs were added just in the overall semantic differential since the focus was not on how the participant feels about a situation but on how the participant feels about the used programs.

The semantic differential scale and the instructions to the user are listed in appendix B, figure B.3.

3.2.3 Description of the Presentations

This chapter describes the different presentations that the participants used during the study. This includes decisions about the used platform, software, topic, structure of the presentations, and visual appearance.

It was decided to use an *Apple iPad*[7] in the user study, which allowed a direct interaction with the presentations. Also, this lowered the differences in interacting with the different presentation tools. While the presenter needs various input commands to carry out the possible actions for a canvas presentation (cf. chapter 2.1.3), a slideware presentation is often held using a remote or pressing one button. Using an *iPad* as the control for the presentation accomplished two things: It enabled the participant to give both presentations using the same input device, which offered all necessary interactions with each presentation while hiding the assumed familiar style of slide presentations. Also it was expected that this would even out the novelty effect (cf. [Gravetter and Forzano, 2011, p. 174]) between the different presentation techniques.

[7] https://www.apple.com/ipad/

The next decision was which presentation software would be used for the user study. The options were four programs that have already been mentioned in chapter 2.1: *Prezi* and *Fly* for canvas presentations, *PowerPoint* and *Keynote* for slide presentations. The combination *Prezi* and *Keynote* was chosen for following reasons:

- *Prezi* is a commercial software while *Fly* is a research software that might not be as stable as *Prezi*. As it was planned to evaluate the feelings of the participant while using the presentation software it was decided to use the more reliable software to keep unplanned technological difficulties to a minimum.

- The decision between *Keynote* and *PowerPoint* tipped to *Keynote* because of convenience reasons. Both software offer similar features and since *Keynote* was already available, it was decided to use *Keynote*.

The topic of the presentations needed to be a subject that everyone could say something about. It should be something that people could give a talk about and remove the focus from the interaction with the presentation software. Since Germany just won the *FIFA World Cup*[8], the lineup of the German national soccer team was selected as the topic of the presentations. This topic was a good fit as the tournament was featured heavily in the media so that it could be assumed that most people had heard about the players. It was also easily possible to present this topic in a spatial way by placing players on their actual position on the pitch (cf. figure 3.3).

This setup separated the presentation in two parts. First were the players who started on the pitch initially and second the players who were on the bench at the start of the game. The sequence of the presentation was the same for both presentations formats. Starting point was the overview as shown in figure 3.3. After this, each player on the pitch was put into focus in the same order as in the broadcast on television (starting from the goalkeeper, show the defense, midfield, and striker). Some basic information as well as a photo were added for every player in the presentation. These information were their position, age and regular club (cf. figure 3.4). In addition to these information about the player, a shape close to the player was added. This shape was either a rectangle, circle, triangle or house and colored either in red, blue, gray, green or black. These shapes were included so that a participant could be asked to find a certain shape, thereby completing the *search for an unknown location* (cf. chapter 3.2.1).

[8] http://www.fifa.com/worldcup/

Figure 3.3: Overview of the presentation. Beside each player are his basic information (position, age, regular club) and a shape for the *search unknown* task.

For the condition during which technological difficulties should occur in the presentation, two situations were added that simulated such difficulties. The first simulated a misinterpreted input. In order to achieve this effect, an action was created that showed the previous player when the participant planned to go to the next player. Example: The participant moved from *Player A* to *Player B*. If she now performed the same action, instead of going to *Player C* the presentation showed *Player A* again. This simulated a wrong input or an unexpected reaction. The second situation simulated a 'freezing' of the presentation. To achieve this unresponsiveness, two 'empty' animations were added to one of the players. This resulted in a situation during which the participant had to perform an action multiple times in order to continue. Therefore, the program appeared to be unresponsive.

Figure 3.4: Example of the information shown for one player. The gray rectangle is one of the targets for the *search unknown* task.

Since these situations happened in either the canvas presentation or the slide presentation, this resulted in at least four different presentations that were used during the study. The various shapes (as shown in figure 3.4) were placed differently for the presentation with errors and the presentation without errors to avoid memorization effects. The placement between the canvas presentation and the slide presentation, however, was the same. A version without the shapes was added for each program because the participant was allowed to practice with the presentation program and familiarize herself with the presentation content. Using this version, the participant was not able to memorize any placements of the shapes before giving the presentation. So overall, three presentations were created for each software (one without shapes and without errors, one with shapes and without errors, and one with shapes and with errors) resulting in six presentations.

Regarding the visual appearance of the presentations, the same layout was used for both presentation types. In their current versions *Keynote* offers a lot more options than *Prezi* to adjust the look of the content. Therefore, only the possibilities of *Prezi* were used in order to avoid different ratings due to different appearances.

To achieve this, a screenshot was taken of every step of the canvas presentation and these images were added in the slide presentation. "Moving in" was used as the animation between slides in *Keynote* since the action for going to the next slide on an *iPad* is sliding from right to left.

During the presentation, the participant saw the current position on the *iPad*. The only difference between the presentation programs was that *Keynote* showed an information bar above the current slide. This bar included information about the current slide number, if there was currently an animation running, the time, and options to draw on the screen, display presenter notes, and close the presentation.

3.2.4 Additional Materials

This chapter describes what additional materials were used both to gather information from the participant (chapter 3.2.4 "Questionnaires") but also to assist in recording and storing the data (chapter 3.2.4 "Technological Devices and Programs").

Questionnaires

In addition to the semantic differential described before (chapter 3.2.2), different questionnaires and scales were used and created to elicit data from the participant. The participant needed to sign a consent form prior to the study. This consent form included information about the purpose and sequence of the study as well as risks and benefits. Beside signing that she had these information explained to her or read and understood the consent form, the participant also had the choice to enter her E-Mail address to participate in the raffle of a 20€ Amazon Gift Card. The whole consent form is attached in appendix B, figure B.1.

As it was desired to exclude that different moods affect the feelings of the participant, the PANAS test was used to measure positive and negative affect of the participant for that day (cf. chapter 2.2.3). The translation provided by Krohne et al. [1996] was taken since the study was conducted in German. The test is listed in appendix B, figure B.2.

To measure the valence and arousal of the participant in certain situations, the self-assessment manikin described in chapter 2.2.2 "The Self-Assessment Manikin" was used. It was decided to use only the images measuring valence and arousal since this had to be filled out several times and the dominance scale did not correlate

well with results from the semantic differential (cf. Bradley and Lang [1994]). A nine-point rating scale (images taken from Irtel [2007]) was prepared together with the instruction to mark in each row which image depicts the feeling in the shown situation best. These instructions and the used images are in appendix B, figure B.4.

After each presentation and the following review, the participant was asked how it felt using the software. Additionally, she was asked five questions about the systems and presentations overall after both the presentations and the corresponding reviews were completed. These questions were:

1. *"What difference in feelings did you experience between the two programs?"*
2. *"What is your opinion about the programs and the kinds of presentations? What did you like, what did you miss?"*
3. *"Have you experienced special moments during past presentations? Moments where something worked very good or went completely wrong? How did that feel?"*
4. *"How did your presentation style change over time?"*
5. *"Comparing the presentations today with your normal presentations: how different were your feelings today?"*

While the first two questions were aimed to elicit comparisons between the programs and the general evaluation, questions three and four targeted the previous experience with presentations. The last aspect of the third question ("How did that feel?") was asked after the participant mentioned a specific situation. With the last question it was desired to verify the assumption that the participant experiences similar feelings during the study compared to real presentations (cf. chapter 3.1).

Since it was desired to see whether spatial ability influences the experience of presenting, the participant's spatial ability was measured using the Paper Folding test by [Ekstrom et al., 1976, p. 176]. This test has been used in numerous studies to assess the spatial ability (e.g., Arning and Ziefle [2009]; Hegarty et al. [2003]). It consists of images that show a paper being folded multiple times with the last image showing where a hole is made through the paper. The participant has to select the correct arrangement of holes once the paper is unfolded out of five possible solutions (cf. figure 3.5). The result is the number of items that were solved correctly (maximum of 20). The instructions for this test were translated into German and they are listed in appendix B, figure B.5.

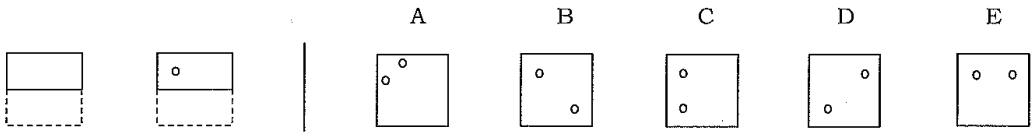


Figure 3.5: Example task of the spatial ability test by Ekstrom et al. [1976]. The left side shows how a paper is folded and where a hole is punched through the layers. The participant has to find the correct arrangement of the holes on the unfolded paper out of the options on the right.

To gather previous knowledge and demographic information, a questionnaire was created that the participant had to fill out. The information she had to give were:

- Age
- Gender
- Occupation
- *"How long ago was your last presentation?"* (one week, one month, one year, longer, never)
- *"How often do you give a presentation?"* (multiple times per week, per month, per year, more rare, never)
- *"In what context do you normally present?"*
- *"At what age did you start presenting?"*
- *"At what age did you present the most?"*
- *"How often have you used canvas presentation tools?"*(multiple times per week, per month, per year, more rare, never)
- *"How often have you used slide presentation tools?"* (multiple times per week, per month, per year, more rare, never)
- *"How much do you enjoy giving presentations?"* (five-point scale from *very much* over *neutral* to *not at all*)

It was planned to compare different groups with different attributes (e.g., 'presentation age', 'gender') against each other using these information. This questionnaire is listed in appendix B, figure B.6.

Technological Devices and Programs

Several technological devices were used to help collect the data. Beside the *iPad* that was used by the participant to give the presentation, an *Apple iPhone 5* and a *Canon MVX 35i* were used to record the presentations. Even though only the *iPhone* recording was used to review the presentation with the participant, the *Canon* was running with a tape that was overwritten after the experiment (cf. chapter 3.1). The *iPhone* was also used in addition to the notes to record the participant's responses to the qualitative questions.

To record the skin conductance (electrodermal activity), the EDA-sensor of a *BITalino board*[9] was used. The data recorded by the sensor were sent to a *MacBook Pro* where they were processed by a *Java* application. This application used *Live-Graph*[10] to display the data in real time as well as store them in a CSV file. The real time visualization allowed to verify and check the data that came from the sensor. It was also possible to mark entries during the presentation to indicate where the situations happened that were interesting for the evaluation. In order to evaluate the data from the EDA-sensor, a makro for *Microsoft Excel*[11] was written to calculate the skin conductance response with the method used by Greenwald et al. [1989].

3.3 Study Script

This chapter describes the sequence of the experiment. Every experiment started with the introduction (chapter 3.3.1) after which the participant gave her presentations. The instructions for these and the steps of the following review are described in chapter 3.3.2. Chapter 3.3.3 shows the last phase of the experiment during which the final information of the participant were collected. After the experiment, the data was entered for evaluation. The encoding for this is outlined in chapter 3.3.4.

A collection of forms that the participant had to fill out (e.g., to report her feeling for a specific situation) was prepared as well as a document to record the responses of the qualitative questions and take notes. The document for the participant included the questionnaires described above. The consent form was followed by the PANAS test and the forms to report the feeling for the specific situations and the programs overall. Behind these pages were the spatial ability test and the questionnaire about the personal information.

[9] http://www.bitalino.com
[10] http://www.live-graph.org
[11] http://office.microsoft.com/en-us/excel/

Another document was used to note for each program the technique that the participant used when completing different tasks (e.g., did she click through the pages or used the zooming feature?). Additionally, the responses to the open questions were recorded.

3.3.1 Introduction

Each experiment started with a short explanation of the purpose of the study. The participant was told that this study wanted to take a look at the feelings of a presenter while giving a talk using one of two different presentation formats. Both canvas presentations and slide presentations were explained shortly before an overview of the study and the next steps was given.

Following the basic information, the participant filled out the consent form, which was explained beforehand. After this, the participant was asked to fill out the PANAS test. She was reminded to evaluate her experience over the current day and answer spontaneously. To familiarize the participant with the scales used to evaluate the feelings, both the semantic differential scale and the self-assessment manikin were explained to her. The instruction to mark between the word pairs was made clear and she was given time to read through the pairs to prevent any uncertainties. For the self-assessment manikin, it was explained that she should mark the two images (one for each row) that depict her experienced feeling best. For the first row (the valence rating), it was described that the leftmost image shows a person who has the corners of the mouth turned down and the person is sad while the rightmost image has the corners drawn up and is feeling very good. The second row was described as changing from an eyes closed, no experienced feelings person to an eyes wide open person having intense feelings. If the participant had no further questions at this point, the study continued with the explanations and demonstrations for the presentations.

3.3.2 Presentations

To give an overview about the different programs and their features, the possible interactions were demoed for the participant. For this, the version without the shapes aside of the players was used. The demo started with the program that the user would use second so that the program she would use first would be explained right before she used it (e.g., if the participant had to use *Prezi* first and *Keynote* second, the demo would start with *Keynote* and move on to *Prezi*). The order in which the participant used the applications and which presentation had the simulated errors was balanced across participants.

The information given during the demonstration of the programs were:

- For *Keynote*:
  - The only difference between the *iPad* screen and the image on the projector was the menu bar, which was not important for the current presentation.
  - The participant was able to proceed to a new slide by tapping on the screen. Here it was also pointed out that the display on the *iPad* only updated after the animation on the projector finished, resulting in a short delay.
  - It was furthermore described how to use swiping motions to move to the next slide or return to the previous slide.
  - The last interaction described for the slide presentation tool was the slide menu. By swiping into the screen from the left border, a list of all slides in the presentation appeared. The participant was able to select any slide here and jump directly to it. By tapping anywhere on the display this slide menu disappeared.
- For *Prezi*:
  - While the starting image looked the same as in the slide presentation, it was shown that the image on the *iPad* directly mirrored what was seen on the projector. It was demonstrated that this stayed the same as the participant dragged the finger around on the screen and moved the content this way.
  - The second set of actions described were the pinch and zoom motions to move closer to the content or farther away.
  - There was a predefined sequence of frames (a path) for the presentation and pressing on the sides of the display moved through this sequence. Here, it was pointed out that there was still the possibility to move around the presentation and that tapping either of the controls continued the path from where it was left off.
  - It was also shown that double tapping a player brought him into focus and moved the position on the path here as well. This meant that the next and previous controls now started from this player instead of the last selected player.
  - After this action had happened during the pilot study and caused confusion, it was also demonstrated that holding the forward button caused a jump to the final frame of the path while holding the back button zoomed back to the overview.

After these demos, the EDA-Sensor was attached to the participant's hand. The contacts were placed on the second element of the index and middle finger of the non-dominant hand—as this hand was used to hold the *iPad*—to not impair the mobility. This placement is suggested by [Blascovich et al., 2011, p. 31f.] and allowed to use the dominant hand to navigate and gesture on the *iPad*.

Following this, the participant had as much time as wanted to familiarize herself with the controls of the application as well as review the given information and prepare for the talk. The presentation was switched to the appropriate version with shapes once the participant was confident controlling the program and was ready for the talk. Also, she was instructed to remember the last presentation she gave and how she felt in that situation. She should imagine the room to be full and that she wants to give a good presentation. After this, both the cameras and the recording application for the skin conductance data were started.

During the presentation, the moderator acted as an interested audience member that smiles and acknowledges the information given. The live visualization of the data was monitored to spot any problems. Furthermore, markings were placed in the data for the predefined situations in order to be able to evaluate the skin conductance data. These situations are visualized in figure 3.6 and were:

- *Step*:
  - 1a First presentation: The step from Mats Hummels to Jerome Boateng.
  - 1b Second presentation: The step from Jerome Boateng to Philipp Lahm.
- *Search known*:
  - 2a First presentation: After the participant described Miroslav Klose, she was asked to show Philipp Lahm again.
  - 2b Second presentation: After the participant described Thomas Müller, she was asked to go back to Manuel Neuer.
- *Search unknown*:
  - 3a First presentation: When the participant was finished, she was asked to show the player that has a gray triangle beside him.
  - 3b Second presentation: After the presentation, the participant was asked to go to the player with a blue circle.
- *Skip*:
  - 4a First presentation: As the participant reached Per Mertesacker, she was told that she should skip the last players and go to the coach because the time would run out.

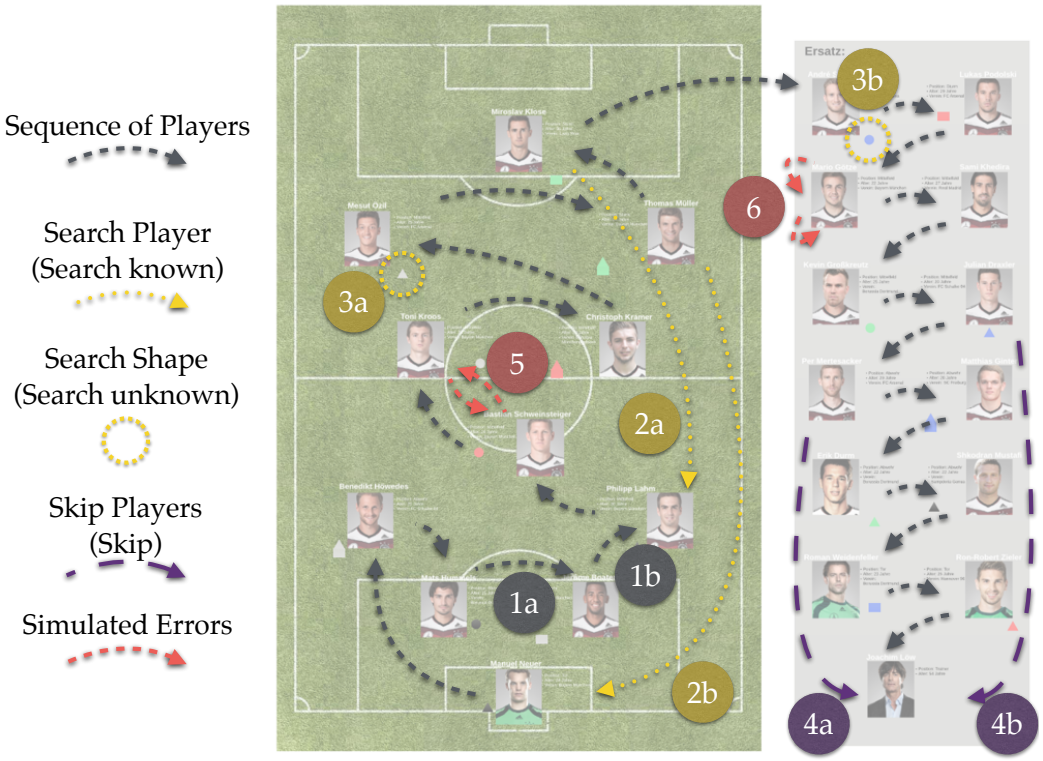


Figure 3.6: Visualization of the paths and tasks during the presentations. The numbers indicate the situations that were evaluated with the participant. The letters *a* and *b* express whether the situation was used and reviewed in the first or second presentation respectively.

4b Second presentation: The instruction was the same, but she was asked to perform it when she reached Julian Draxler (one player before Per Mertesacker)

- *"Wrong input"-error*:

  5 This error occurred when the participant tried to move from Toni Kroos to Christoph Kramer. Instead of showing Kramer, the presentation went back to Bastian Schweinsteiger.

- *"Freezing"-error*:

  6 The 'empty' animations happened when the participant tried to continue from Mario Götze.

The mark for an event was placed about a second after the question or after the error happened to give the participant time to understand the tasks or realize the problem. If the participant did not encounter the specified situations (e.g., because she held the canvas presentation without following the path), either a similar action was marked (e.g., for *Step*: going from an overview of the defense to an overview of the midfield) or she was instructed to perform a certain action (e.g., displaying a specific player and showing the following players to trigger the errors).

The video recording from the *iPhone* was transferred to the laptop and the situations were shown to the participant subsequent to the presentation. After each situation, the participant was asked to judge and remember the feeling and fill out the questionnaires. This was followed by the open ended questions for which the answers were recorded both on the *iPhone* and on paper. These steps were repeated for the second presentation as well.

3.3.3 Final Questions

After the presentations and the presentation specific questions were answered, the open ended questions regarding the comparison of the programs and the prior presentation experiences of the participant were asked (cf. chapter 3.2.4). As before, the answers were recorded on the *iPhone* as well as on paper.

The next task for the participant was the spatial ability test. Beside the written instructions, the purpose and the necessary steps were explained using the provided example and solution (cf. appendix B, figure B.5). The participant was given three minutes for each part of the test.

When the participant had completed the spatial ability test, she was presented with the final questionnaire about her personal information and previous experience with presentations in general and presentation tools in particular. After she had filled out this form, any further question that she had about the study were answered before she was thanked for the participation and guided outside.

3.3.4 Entering the Data for Evaluation

In order to collect all gathered data in one place and prepare them for the evaluation, the information of each participant was entered into an *Excel* file. Each row represented the information of one participant and the first column described which application was first used for presenting and which presentation was the error presentation. The following columns included the personal information as well

as the previous experience with presentations. For this, the questions regarding the last presentation, presentation frequency, how often canvas or slide presentation tools were used before, and how much the participant likes to present were coded using values between one and five. The numeric values (age, starting age for presentations, age when presenting the most) were stored as well as the occupation, gender and the context in which the participant normally presents. Furthermore, a *presentation age* was calculated by subtracting the starting age from the current age and a *technological expertise* variable by calculating the mean of the ratings of how often the participant used canvas or slide presentation tools respectively. To support the connection that more technological expertise results in a higher technological expertise rating, the mean of the rating was subtracted from five, resulting in a representation in which zero corresponds to low technological expertise, and four corresponds to high technological expertise. For the spatial ability, the correct items from the Paper Folding test were counted and for the PANAS test the values for the positive as well as the negative affect were calculated.

Following these information about the participant, the data for each program and each situation were entered. The skin conductance values were calculated using the *Excel* makro (cf. chapter 3.2.4) and the values from the semantic differential were coded using values from one to nine where higher values represent feelings closer to the positive site of the scale. For the search tasks and the skipping task, it was noted how the participant navigated as well (e.g., if she used the slide menu of *Keynote* or swiped through the slides). The answers to the qualitative question about a specific tool were entered behind the ratings for the presentation tool overall. Here, the notes about the presentation (e.g., if the participant did not use the predefined path in the canvas presentation) were added as well. The answers to the final questions about the presentations and the experienced feelings were noted in the last columns of the document.

3.4 Hypotheses

This chapter outlines the hypotheses that this study wants to address. One general hypothesis regards the comparison between canvas and slide presentation tools, while other hypotheses arise from the setup of the study.

The central question that this thesis wants to answer is how a presenter feels while using a canvas presentation tool. Therefore, the main hypothesis is:

(H1) Feelings in canvas presentations are rated differently than feelings in slide presentations.

Since the study design features both presentations with technological difficulties as well as presentations without technological problems, two hypotheses address this:

(H2) Presentations with technological difficulties are rated differently than presentations without technological problems.

(H3) Technological problems are the main reason why a presenter experiences a *bad* presentation.

As the skin conductance was collected, the next hypothesis is:

(H4) The response in skin conductance correlates with the arousal or valence rating.

To check that the balancing of conditions worked, the hypothesis is tested that

(H5) The order of conditions (*canvas* and *slide*, *error* and *no error*) had no effect on the ratings.

The final hypothesis deals with the problem of generalizing the results gathered in this study:

(H6) Participants experienced the same feelings during the study compared to a real presentation.

Chapter 4

Evaluation

In this chapter, the hypotheses are evaluated. In chapter 4.1 the sample of participants who took part in the study is described. Following this, each of the hypotheses is explored individually (chapter 4.2). The final part of this chapter evaluates observations that have been made regarding the delivery style of the presentations.

4.1 Sample Description and Group Composition

The study was conducted between the 21st of July and the 8th of August 2014 in two seminar rooms and a classroom. Overall, 21 participants from various professions and with different presentation proficiency took part in the study. The age ranges from 17 to 66 years with a mean of 37.09 and standard deviation of 16.02.

A *presentation age* (*PAge*) was calculated by subtracting the age at which the participant first presented from her current age. This presentation age has a mean of 18.33 years with a standard deviation of 11.73 and a range from 7 to 42 years. This presentation age is taken as an estimation of the prior *presentation experience* of the presenter.

A measure for the *technological expertise* (*TE*) was also created by calculating the mean of the frequency questions ("How often have you used canvas presentation tools/slideware?"). A zero in this distribution means that the participant has used neither canvas presentation tools nor slideware before, while a four means that the participant uses both multiple times per week. The median of this distribution is 1 and the responses range from 0.5 to 2.5.

Regarding how much the participants *like to present* (*L*), five participants reported that they like to present very much and nine participants said that they like to present. Five approach presentations with a neutral attitude, while two only present reluctantly.

As for the *spatial ability* (*SA*) scores, participants solved between 4 to 20 items of the Paper Folding test correctly. The mean is 12.76 correct solutions with a standard deviation of 4.77.

The answers from the PANAS test were evaluated by summing up the ratings to the *positive affect* (*PA*) score and the *negative affect* (*NA*) score respectively. The positive affect ratings range from 22 to 40 and have a mean of 31.14 (standard deviation: 4.95). As for the negative affect ratings, the mean is 12.57 with a standard deviation of 2.34 and the range is from 10 to 18.

A correlation of the above mentioned variables uncovers five correlations. First, there is a significant positive relationship between the age of the participant and her presentation age (r_s=.84, p<.001). Furthermore, the age correlates with the positive affect ratings as well (r_s=.5, p<.05). Third, the presentation age has a significant inverse relationship with the number of correct items in the spatial ability test (r_s=-.48, p<.05). Spatial ability is also negatively correlated with the age of the participant (r_s=-.7, p<.001). Finally, the rating of how much a participant likes to present correlates with the negative affect ratings (r_s=.44, p<.05). Since lower *like* ratings correspond to more enjoyment in presenting, this means that people who do not like to present reported higher ratings on the negative affect score.

Based on these correlations, it was decided to use the presentation age, the technological expertise, and how much the participant likes to present as factors for the following evaluation. Furthermore, the positive affect score is used as a covariate. Two groups were created for each factor either based on the distribution of the values, or based on semantics.

Since the histogram of the *presentation age* scores divides the participants into two groups, they were split into a low and high presentation age group (*low PAge* and *high PAge*) and the separating line is at 20 years of presentation experience. The low presentation experience group includes 16 participants and has a mean of 12.25 years of presentation experience (standard deviation: 3.45). The minimum presentation age is 7, while the maximum is 19 years. The high presentation age group has 5 members and a mean of 37.8 years (standard deviation: 4.6). Furthermore, the range is from 30 to 42 years of presentation experience.

The results of *how much the participant likes to present* were split into participants who like to present (including participants who like to present very much; *Like*)

and people who do not like to present (either neutral or reluctantly; *Don't like*). This split results in 14 participants falling into the *Like* group (9 like to present, 5 like it very much) and 7 participants in the *Don't like* group (5 gave a neutral response, 2 only present reluctantly).

The *technological expertise* was divided into a group than includes all participants that have a *TE* rating less than 1 (including 1) and a group with *TE* ratings more than 1. This value was chosen as it separates the participants into a *low TE* group—in which participants fall who have either never used any of the presentation software (*TE*=0), use one of them seldom (*TE*=0.5), use both seldom (*TE*=1), or have never used one of them and use the other multiple times per year (*TE*=1)—and a *high TE* group—the participants with the lowest score of this group (*TE*=1.5) either use one software multiple times per month and the other never or use one software seldom and the other multiple times per year. The *low TE* group includes 11 participants (2 have a *TE* score of 0.5 and 9 a score of 1) and the *high TE* group 10 participants ($6\times TE$=1.5, $2\times TE$=2.0, $2\times TE$=2.5).

It needs to be kept in mind that due to the correlation between presentation age and spatial ability as well as between like and negative affect, a finding for one of them might also originate from the other.

4.2 Exploring the Hypotheses

This chapter evaluates the data in order to find evidence for the hypotheses described in chapter 3.4. After the initial tests, the hypothesis regarding the connection between skin conductance response and arousal/valence ratings (H4) is investigated first (chapter 4.2.1). The effect that the used presentation program has on the feelings is evaluated in chapter 4.2.2. Following this are chapters that look at the impact of technological difficulties (chapter 4.2.3 and 4.2.4). Chapter 4.2.5 provides evidence that the ordering of the conditions had no effect on the ratings, and chapter 4.2.6 explores whether participants felt different in the study compared to real-world presentations.

The data gathered in the study is evaluated using an omnibus method—comparing slideware and canvas presentations regardless of the occurrence of errors, and comparing presentations with technological problems (*error*) vs. presentations without technological problems (*no error*) regardless of the program used. The ratings from the self-assessment manikin and the semantic differential have been treated as interval data, which is possible according to Scherer [2005] and Heise [1969]. Even though the data violates the normality assumption, parametric tests (e.g., MANOVA) were used for the exploration of the data. Any findings, therefore,

should be viewed with this in mind and seen only as preliminary results that need to be verified using a bigger sample.

First of all, it is verified that the overall ratings are a good representation of the presentation experience. A mean rating based on the ratings from all situations (*Step*, *Search known*, *Search unknown*, *Skip*, *"Wrong input"-error*, *"Freezing"-error*) was calculated and the correlation of this rating with the overall rating was computed. All dimensions of the overall rating have a positive correlation with their counterparts of the mean rating, and for all except *surprised–not surprised* this correlation is also significant. For *surprised–not surprised* the correlation is only marginally significant (p=.58). Based on these findings, the overall ratings are used as a representative of the feelings during a presentation for the following analysis.

The following evaluation reports the test statistic provided by Wilks' Lambda for the multivariate tests and the results provided by the Greenhouse-Geisser adjustment for the repeated-measures tests.

4.2.1 H4: The Response in Skin Conductance Correlates with the Arousal or Valence Rating

The first hypothesis addressed is *(H4) The response in skin conductance correlates with the arousal or valence rating*. During the course of the study, the EDA-sensor broke down after 11 participants. Therefore, only data for these 11 participants is available and some of them are incomplete due to malfunctions of the sensor. To answer the hypothesis, the skin conductance responses, the valence ratings, and the arousal ratings of all situations and programs combined are evaluated. This leads to a set of 95 tuples of skin conductance response, valence rating, and arousal rating.

Previous studies suggest that the skin conductance response is connected to the arousal dimension of an emotion (cf. chapter 2.2.2 "The Autonomic Nervous System"). Such a connection does not surface from the data obtained in this study. A Pearson's correlation shows neither a correlation of skin conductance response to arousal ratings (r=.026, *ns*) nor to valence ratings (r=-.1, *ns*). Therefore, hypothesis H4 is rejected. Since the skin conductance responses are only available for 11 of the 21 participants, they are not included in the following analysis.

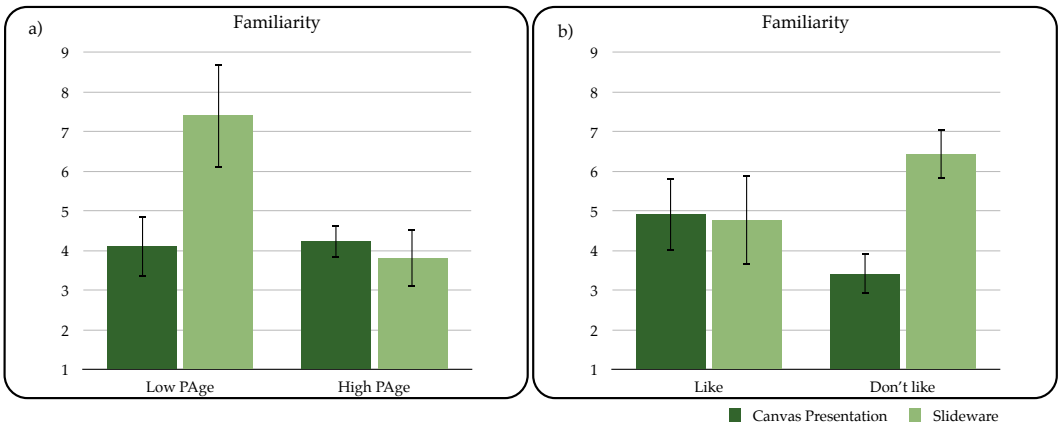


Figure 4.1: Bar charts of the significant *Program*×*PAge* (a) and *Program*×*L* (b) interactions for familiarity ratings. (a) shows that the familiarity ratings of both programs is equally low except for the slideware ratings of participants with less experience presenting. (b) shows that both programs are equally familiar for presenters who enjoy presenting, while slideware is more familiar than canvas presentations for presenters who do not like to present.

4.2.2 H1: Feelings in Canvas Presentations Are Rated Differently than Feelings in Slide Presentations

This chapter explores the hypothesis that feelings in canvas presentations differ from those in slide presentations. Beside general evaluations of valence, arousal, and semantic differential ratings, specific situations and ratings were evaluated for which a visual interpretation of the data suggested differences between the programs. These situations and ratings are the *familiarity* rating of the programs, the ratings for the *search unknown* task, and the impact of technological difficulties on each program.

Familiarity

A visual examination of the data shows differences for the familiarity ratings. A repeated-measures ANOVA with the familiarity ratings of canvas and slide presentations as dependent variables, *PAge*, *TE* and *L* as between-subjects factors, and *PA* as a covariate shows significant interactions. The first interaction is between *Program* and *PAge* ($F(1,19.45)=7.31$, $p<.05$). A graph of the means shows that participants with a high *PAge* rate both programs marginally the same (*Canvas*: 4.22, *Slide*: 3.81), while participants with a low *PAge* rate slideware clearly as more familiar (*Canvas*: 4.11, *Slide*: 7.40) (cf. figure 4.1a).

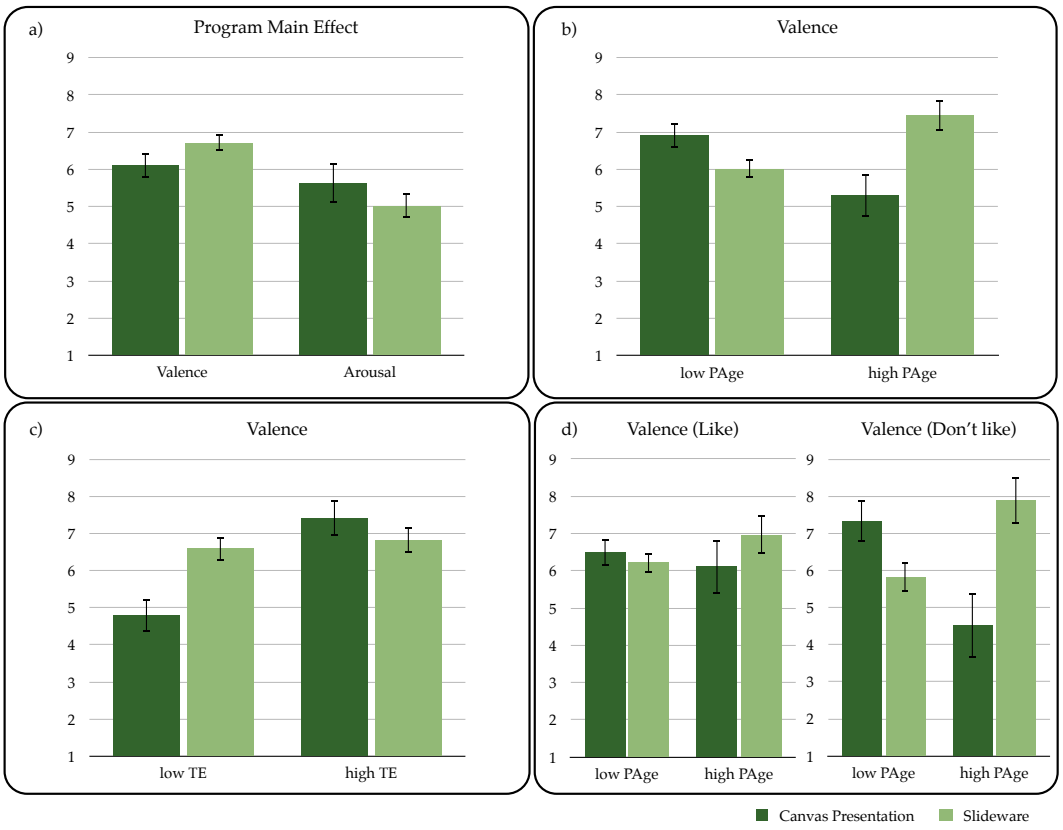


Figure 4.2: The main effect (a) indicates that participants felt more pleased during the slideware presentation and more aroused during the canvas presentation. The interactions *Program*×*PAge* (b) and *Program*×*PAge* (c) suggest that low *PAge* and high *TE* presenters felt more pleased during the canvas presentation, while the opposite appears to be true for high *PAge* and low *TE* presenters. (d) shows that presenters who like to present felt pleased regardless of *Program* and *PAge*, while not liking to present shows an increased the effect of (b).

Another interaction is between *Program* and *L* ($F(1,16.83)=6.33$, $p<.05$). The graph of the means indicates that there is no difference between the ratings of the *Like* group (*Canvas*: 4.92, *Slide*: 4.78), while participants that belong in the *Don't like* group rate the slideware as more familiar (*Canvas*: 3.42, *Slide*: 6.43) (cf. figure 4.1b).

Valence and Arousal Ratings

A repeated-measures MANCOVA was conducted with the valence and arousal ratings from the self-assessment manikin as dependent variables. *PAge*, *TE* and

L were taken as between-subjects factors and *PA* as a covariate. The results indicate a significant between-subjects effect of *TE* ($F(2,11)=8.29$, $p<.01$) and significant within-subject effects of *Program* ($F(2,11)=4.24$, $p<.05$), the interaction *Program*×*PAge* ($F(2,11)=6.47$, $p<.05$), the interaction *Program*×*TE* ($F(2,11)=4.88$, $p<.05$), and a three-way interaction *Program*×*PAge*×*L* ($F(2,11)=4.03$, $p<.05$).

Between-subjects effect of *TE* A univariate analysis shows a significant effect for the valence ratings ($F(1,12)=13.80$, $p<.01$). A comparison of the means indicates that participants with higher *TE* felt more pleasure during the presentations (*low TE*: 5.69, *high TE*: 7.14).

Main effect of *Program* A follow-up evaluation of the main effect of *Program* yielded no significant results for both valence and arousal ratings (*Valence*: $F(1,12)=1.59$, *ns*; *Arousal*: $F(1,12)=3.81$, *ns*). The graph of the means points towards a higher valence rating for slide presentations and a higher arousal rating for canvas presentations (cf. figure 4.2a).

Interaction effect *Program*×*PAge* A univariate analysis reveals a significant effect for the valence rating ($F(1,12)=11.64$, $p<.01$). The analysis of the means indicates that presenters with less presentation experience (*low PAge*) felt more pleasure during the canvas presentation (*Canvas*: 6.91, *Slide*: 6.02), while more experienced presenters rate slideware as more pleasing (*Canvas*: 5.30, *Slide*: 7.43) (cf. figure 4.2b).

Interaction effect *Program*×*TE* A significant effect for the valence rating is found by a follow-up analysis ($F(1,12)=7.45$, $p<.05$). The means indicate that presenters with less technological expertise (*low TE*) felt more pleased during the slideware presentation (*Canvas*: 4.80, *Slide*: 6.59), while more technological savvy presenters (*high TE*) rate canvas as more pleasing (*Canvas*: 7.42, *Slide*: 6.82) (cf. figure 4.2c).

Interaction effect *Program*×*PAge*×*L* The interaction effect of *Program*×*PAge*×*L* also differs significantly for the valence ratings ($F(1,12)=4.94$, $p<.05$). A comparison of the graphs shows that while it does not make a big difference for people who like to present which program they use, the situation is different for those who do not like to present. Here, the presentation experience has a big impact on the preference. Presenters with a low presentation age seem to feel more pleased using a canvas presentation tool compared to slideware, while more experienced presenters feel more pleased using slideware (cf. figure 4.2d).

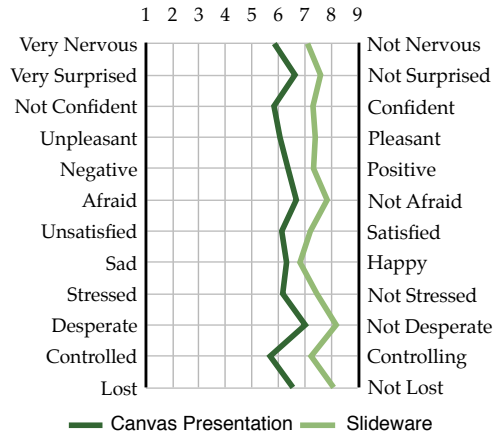


Figure 4.3: Semantic differential profile of the program ratings. Slideware shows consistently better ratings. However, no statistical significant differences are found.

Semantic Differential Ratings

A repeated-measures MANCOVA with the ratings from the semantic differential as dependent variables was conducted. *PAge*, *TE* and *L* were taken as between-subjects factors and *PA* as a covariate. Although slideware has better ratings over all dimensions, no statistically significant differences are found (cf. figure 4.3). The results indicate significant within-subject effects of the three-way interaction $Program \times PAge \times TE$ ($F(12,1)=648.21$, $p<.05$), the three-way interaction $Program \times PAge \times L$ ($F(12,1)=710.45$, $p<.05$), and a four-way interaction $Program \times PAge \times TE \times L$ ($F(12,1)=240.9$, $p=.05$).

Interaction effect *Program*×*PAge*×*TE* A univariate test of all semantic differential dimension showed that only *unsatisfied* differs significantly ($F(1,12)=6.55$, $p<.05$). A graphic comparison indicates that while presenters with a high technological expertise do not differ greatly in in their satisfaction during a presentation, presenter with low technological expertise and high presentation age seem to be more unsatisfied in the canvas condition (cf. figure 4.4).

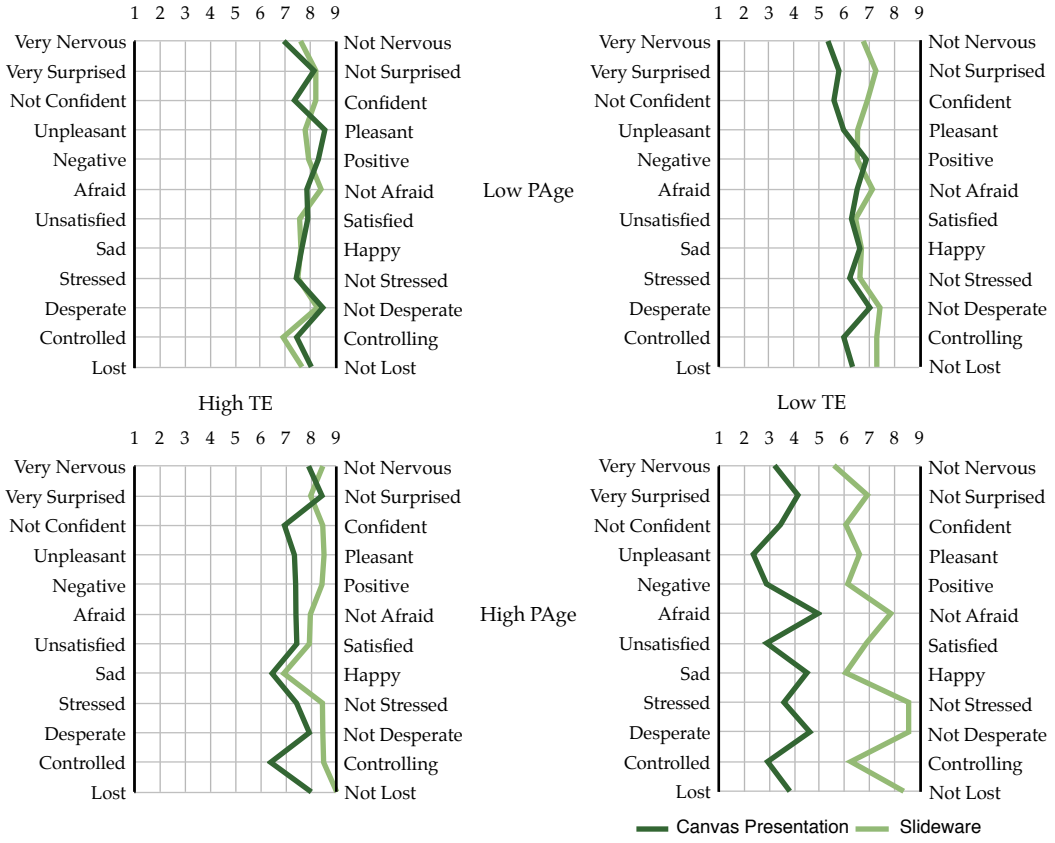


Figure 4.4: Semantic differential profiles for the *Program*×*PAge*×*TE* interaction. Only the *high PAge*, *low TE* group (bottom right) shows great differences between the programs. A statistical significant difference is found for the *unsatisfied* ratings. However, it needs to be kept in mind that this group consisted of only three people.

Interaction effect *Program*×*PAge*×*L* For this interaction effect no significant univariate effects are found. The graphs of the interaction (in appendix C, figure C.1) suggest that people who do not like to present but have a high presentation age experience more negative feelings during the canvas presentation. However, in the current study this group consists only of two persons so further studies need to be done to confirm this finding.

Interaction effect *Program*×*PAge*×*TE*×*L* This interaction is significant both for the *afraid* dimension ($F(1,12)=5.1$, $p<.05$) and the *controlled* dimension ($F(1,12)=4.97$, $p<.05$). However, this interaction is not explored in more detail due to the low number of members in each group.

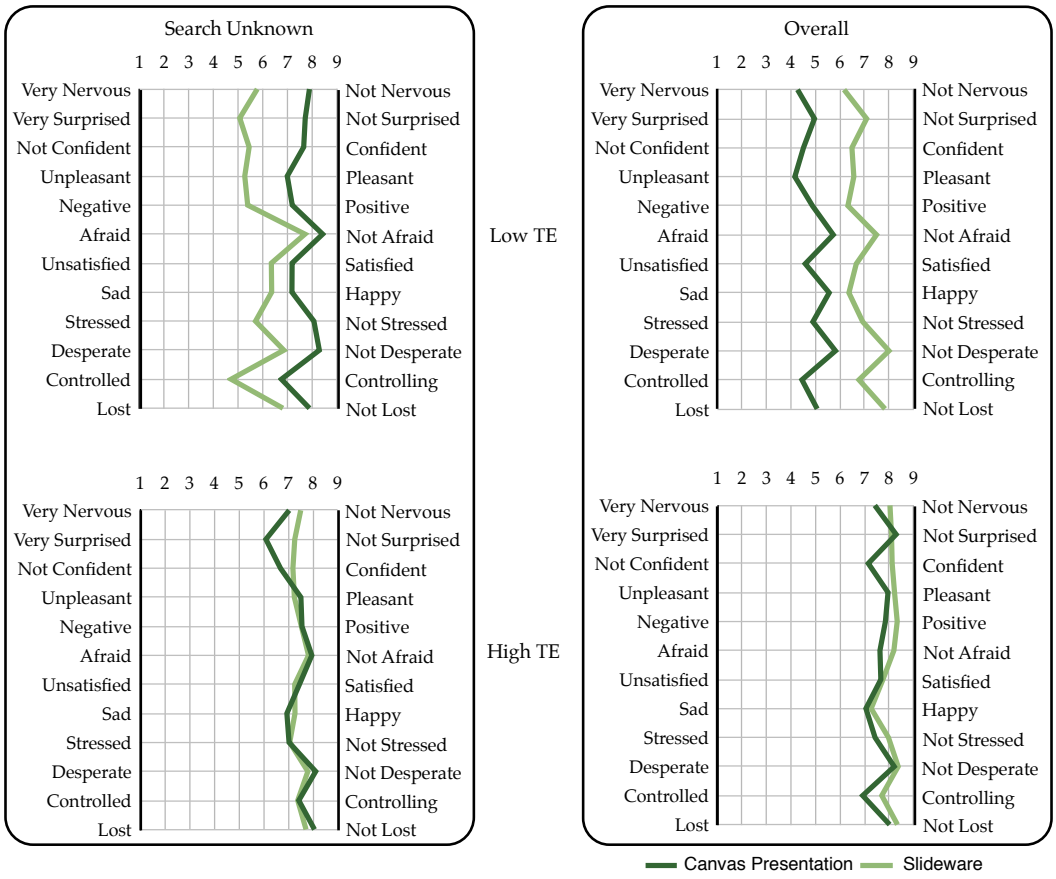


Figure 4.5: Comparison of the semantic differential ratings of the *low TE* (top) and *high TE* (bottom) groups for the *search unknown* situation (left) and the overall ratings (right). Participants from the *low TE* group gave better ratings in the *search unknown* situation compared to the overall ratings. Even though no statistically significant differences have been found for the semantic differential ratings, participants with *low TE* gave significantly different (better) valence ratings.

Task Specific Analysis

Based on a visual interpretation of the task data, two further scenarios were tested.

***Search unknown* task** The graphs for the *search unknown* task show a noticeable difference to other situations and the overall rating (cf. figure 4.5, figure 4.6). In order to inspect the effect of the *search unknown* task, a delta of the ratings for this task and the overall ratings was calculated. Since the overall ratings represent the mean experience over the whole presentation (cf. chapter 4.2), this rep-

resents the diversion of the *search unknown* task. The overall scores for each dimension of the semantic differential and the self-assessment manikin were substracted from the corresponding ratings of the *search unknown* task. Therefore, a positive difference marks a case in which the search task is rated better than average. A repeated-measures MANCOVA reports a significant interaction effect of *Program*×*TE* ($F(2,11)=4.5$, $p<.05$) and of *Program*×*PAge* ($F(2,11)=4.66$, $p<.05$).

The interaction effect of *Program*×*TE* is significant on the valence dimension ($F(1,12)=9.7$, $p<.01$). An analysis of the means shows, that the used program has no effect on the feelings of the *high TE* group, while it significantly affects the *low TE* group. In this group, the search task for the canvas presentation is rated with more pleasure than the overall rating. The rating for the slideware presentation, however, is rated with less pleasure compared to the overall rating.

Regarding the *Program*×*PAge* interaction, a univariate analysis shows a significant difference for the valence ratings as well ($F(1,12)=9.5$, $p<.01$). An exploration of the means shows a drastic difference for the presenters with high presentation experience. Here, presenters with a high presentation age experienced more pleasure during the *search unknown* task with the canvas presentation compared to the overall ratings. The other differences show small negative changes.

Impact of errors Another interesting effect is the impact of the simulated errors. To analyze this impact, the differences between the ratings for the *step* condition and each of the two error conditions respectively were calculated. A high difference represents a high impact of the error. The *step* ratings was chosen as the baseline since the same interaction proceeds the situation. In both cases the participant plans to go to the next slide/frame, but in the error conditions something unexpected happens.

A between-subjects MANCOVA indicates a interaction effect of *Program*×*L* for the *"Wrong input"-error* ($F(2,5)=8.47$, $p<.05$). The follow-up univariate analysis shows that both the valence and the arousal ratings are significantly different (*Valence*: $F(1,6)=9.22$, $p<.05$; *Arousal*: $F(1,6)=11.08$, $p<.05$). The means of the valence rating indicate that people who like to present are less affected by the error when they are using the canvas presentation compared to the slidware presentation (*Canvas*: 1.56, *Slide*: 2.78). The opposite is apparent for presenters who do not like to present (*Canvas*: 2.63, *Slide*: 1.72). In regard to the arousal ratings, the means indicate no differences between the programs for people who like to present (*Canvas*: 1.24, *Slide*: 1.26), while those who do not like to present reported feeling less aroused in the error condition compared to the overall ratings during the canvas presentation (*Canvas*: -2.46, *Slide*: 2.68).

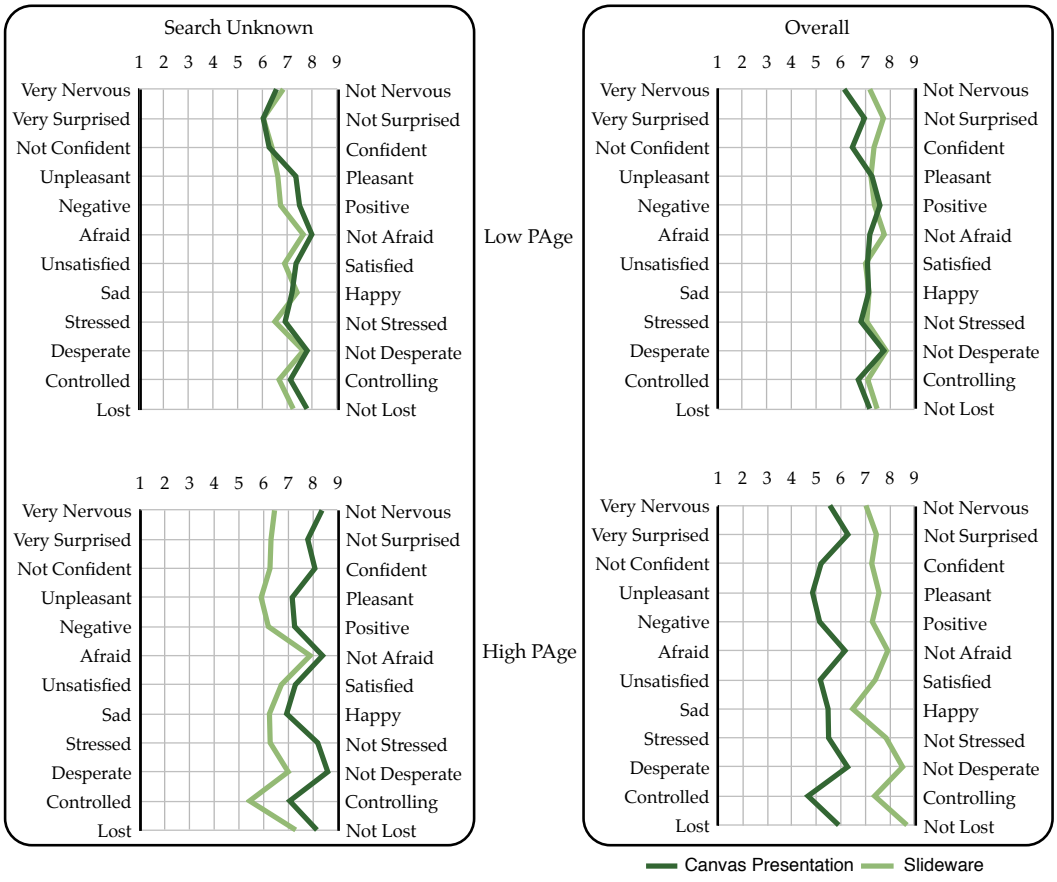


Figure 4.6: Comparison of the semantic differential ratings of the *low PAge* (top) and *high PAge* (bottom) groups for the *search unknown* situation (left) and the overall ratings (right). *High PAge* participants rated the *search unknown* situation better compared to the overall ratings. But, no statistically significant differences have been found for the semantic differential ratings. The valence ratings of participants with *high PAge*, however, are significantly better.

For the *"Freezing"-error*, a between-subjects MANCOVA shows a significant interaction effect of the used program and how much a presenter likes to present ($F(2,6)=8.43$, $p<.05$). The univariate analysis indicates a significant difference in the arousal ratings ($F(1,7)=18.06$, $p<.05$). An analysis of the means indicates that presenters who enjoy presenting reported a higher impact of the error during the canvas presentation (*Canvas*: 2.81, *Slide*: 0.66), while those who do not like to present report a higher impact on the slidware presentation (*Canvas*: -0.517, *Slide*: 3.32).

Qualitative Remarks

The first question of the final interview asked the participant to compare the experience of the two presentations. The participants gave positive answers for both of the presentation formats. For the canvas presentation tool, five users state that they liked the overview capabilities better and two more preferred the spatial arrangement. Furthermore, four participants report that the canvas presentation would have been more free than the slideware presentation. This is enhanced by the response that canvas presentations would make it easier to respond to questions (four participants). Also, participants state that presenting with the canvas presentation tool would be more fun (four participants) and that it would not be as boring as the slideware presentation (five responses). On the other hand, five participants state that slideware offered better control over the presentation and two mention the better input options. Overall, four participants felt that they were more accustomed to the slideware presentation.

Building on these findings, the hypothesis that the presentation software led to different feelings during the presentations is accepted.

4.2.3 H2: Presentations with Technological Difficulties Are Rated Differently than Presentations without Technological Problems

In order to address this hypothesis, the differences between the overall ratings of presentations with errors and the overall ratings of presentations without errors are analyzed. Two repeated-measures MANCOVAS were conducted with the valence/arousal ratings and the ratings from the semantic differential as dependent variables respectively. *PAge*, *TE*, and *L* were used as factors and *PA* as a covariate.

Valence and Arousal Ratings

The analysis of the valence and arousal ratings shows a significant between-subjects effect of *TE* ($F(2,11)=4.79$, $p<.05$) and within-subjects effects of the *Error* main effect ($F(2,11)=8.83$, $p<.01$), the interaction $Error \times PA$ ($F(2,11)=7.66$, $p<.01$), the interaction $Error \times TE$ ($F(2,11)=7.12$, $p<.05$), and the four-way interaction $Error \times PAge \times TE \times L$ ($F(2,11)=7.94$, $p<.01$).

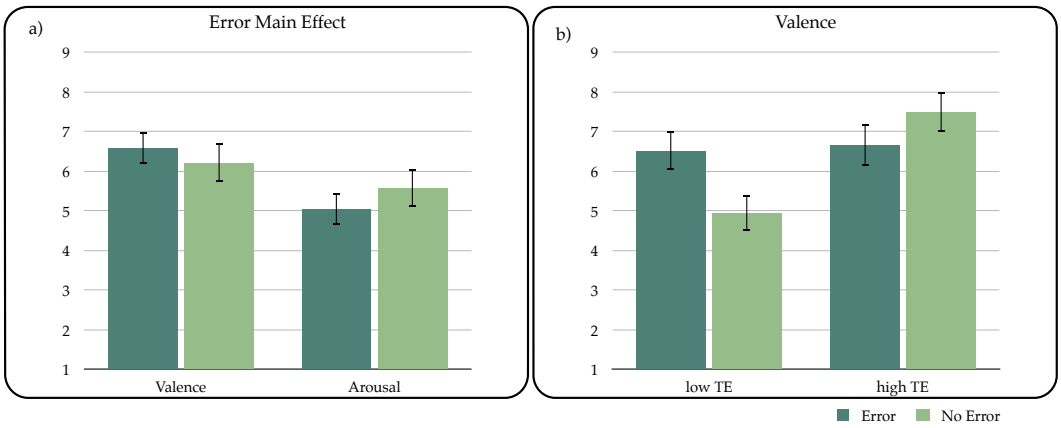


Figure 4.7: Bar charts of the significant *Error* main effect (a) and *Error*×*TE* interaction (b). The main effect is significant for the arousal ratings and the arousal ratings are higher in the *no error* presentation. The valence rating are significantly different for the *Error*×*TE* interaction with *low TE* participants being less pleased in the *no error* condition.

Between-subjects effect of *TE* A univariate analysis of the between subjects effect of *TE* indicates that the ratings are significantly different for the valence dimension ($F(1,12)=9.51$, $p<.01$). Comparing the means, it appears as if presenters with higher *TE* have more pleasure presenting regardless of errors (*Low TE*: 5.73, *High TE*: 7.08).

Main effect of *Error* Regarding the main effect of the occurrence of errors, a univariate analysis shows that the arousal ratings are significantly different ($F(1,12)=10.98$, $p<.01$). The graph and the means suggest that the arousal was higher in the presentation without errors (*Error*: 5.06, *No Error*: 5.58) (cf. figure 4.7a).

Interaction of *Error*×*PA* The interaction of *Error*×*PA* also shows a significant effect for the arousal ratings ($F(1,12)=9.85$, $p<.01$). A correlation of the *PA* scores and the arousal ratings for each condition indicate a significant positive relationship between the *PA* scores and the arousal in the *error* condition ($r_s=.493$, $p<.05$). Such a relationship is not found for the *no error* condition ($r_s=.05$, *ns*).

Interaction of *Error*×*TE* A univariate analysis of this interaction shows a significant difference for the valence ratings ($F(1,12)=4.96$, $p<.05$). Comparing the means, it becomes apparent that while there is no difference between valence ratings in the *error* condition (*Low TE*: 6.52, *High TE*: 6.67), there is a difference between the valence ratings in the *no error* condition (*Low TE*: 4.93, *High TE*: 7.49) (cf. figure 4.7b).

Interaction of *Error*×*PAge*×*TE*×*L* The four-way interaction has a significant effect on the arousal ratings ($F(1,12)=6.48$, $p<.05$). However, this interaction is not investigated in greater detail.

Based on these results, the hypothesis that the occurrence of technological problems affects the feelings during a presentation is accepted.

Semantic Differential Ratings

The semantic differential ratings showed neither significant between-subjects effects nor significant within-subjects effects.

4.2.4 H3: Technological Problems Are the Main Reason Why a Presenter Experiences a *Bad* Presentation

In order to address this hypothesis, the qualitative responses to the question about prior presentation experience are evaluated. This was the third question of the interview after both software had been used and reviewed and the participant was asked whether she has experienced any *special moments* (good or bad) in past presentations (cf. chapter 3.2.4 "Questionnaires").

In connection to the bad moments in presentations, 12 of the participants mention that they experienced technological problems in their presentations. This is by far the most often given reason for a bad presentation. Other responses indicate that the own performance can lead to bad presentations—two participants experienced situations in which they were not able to answer questions from the audience and three other participants mentioned presentations for which they were not prepared enough. Furthermore, one participant states that she experienced bad presentations during which she dropped items. Overall, three participants report that they had not experienced a bad presentation before. Based on these responses, the hypothesis that technological problems are the main reason why a presenter experiences a bad presentation is accepted.

4.2.5 H5: The Order of Conditions (*Canvas* and *Slide*, *Error* and *No Error* Had No Effect on the Ratings

To check whether the order of conditions had an effect on the ratings, the results between each condition are compared. For each comparison, a between-subjects MANOVA was conducted with the valence/arousal ratings and ratings from the semantic differential as dependent variables respectively. Depending on the comparison, the position of the software (first or second) or the position of the *error* condition (first or second) was used as a factor.

Ordering of the presentation software The test for ordering effects of the presentation software shows that there is no significant difference in the ratings for slideware and canvas presentations based on their position in the study.

- Valence and arousal ratings of the canvas presentation: $F(2,18)<1$, *ns*
- Semantic differential ratings of the canvas presentation: $F(12,8)<1$, *ns*
- Valence and arousal ratings of the slideware presentation: $F(2,18)<1$, *ns*
- Semantic differential ratings of the slideware presentation: $F(12,8)<1$, *ns*

Ordering of the error condition The comparison between ratings of the *error* condition based on the position in the study and the comparison for the *no error* condition shows that the ordering had no effect on the ratings.

- Valence and arousal ratings of the *error* condition: $F(2,18)=1.73$, *ns*
- Semantic differential ratings of the *error* condition: $F(12,8)<1$, *ns*
- Valence and arousal ratings of the *no error* condition: $F(2,18)<1$, *ns*
- Semantic differential ratings of the *no error* condition: $F(12,8)<1$, *ns*

Therefore, the hypothesis that the ordering of conditions had no effect on the ratings is accepted.

4.2.6 H6: Participants Experienced the Same Feelings during the Study Compared to a Real Presentation

The final question of the qualitative interview targeted the differences in experience between the presentations in the study and regular presentations of the participant. The answers to this question are used to address the final hypothesis—whether the participants experienced the same feelings during the presentations in the study compared to their real presentations.

All of the participants except one state that they felt similar feelings in the study presentations compared to their normal presentations. Out of these 20 participants, three clarify that the study presentations felt like ungraded, everyday presentations. Two other participants name the sensor and the cameras as a reason for the similar feelings, since they would have increased the impression of being observed. A further participant states that there would always be a basic excitement while giving a presentation, which is what she felt during the study. Finally, one explains that she would usually block out the audience during her presentations which would be the reason why she did not feel different in the study.

However, 14 participants mention that they felt less intensive feelings during the study presentations since there would have been no pressure or evaluation that pushed them to deliver a good presentation. Another reason for these less intensive feelings would be the missing audience. Further answers point to more differences between the feelings during the study presentations compared to real presentations. Ten users state that they felt different because they would usually have more time to prepare for a presentation, therefore the topic would have been known to them and they would have created the visuals themselves. Two of the participants report that the presentations in the study were harder because of this reason—one of these two is the participant who did not explicitly say that she felt similar feelings during the study presentations compared to her usual presentations.

Since the majority of participants state that they experienced similar albeit less intensive feelings during the study, the hypothesis is cautiously accepted.

4.3 Observations

This chapter reports the observations that have been made during the presentations. This includes how many participants held the presentation differently compared to the predefined sequence and what options were chosen to accomplish the tasks.

While every participant followed the sequence in the slideware presentation, seven participants gave a very free canvas presentation. Four of them did not use the path at all and the others deviated from it by regularly zooming out for overviews or changing the order in which players were presented. None of the seven participants had used canvas presentations before. Three of them fall into the *low PAge* group and three into the *high TE* group (two overlapped). Six of the seven stated that they like to present or like it very much.

In regard to the used actions to complete the different tasks, all participants except four used the advanced features for all situations. For the slideware presentation this means that the slide menu was used in favor of clicking through the slides and for the canvas presentation the free movement was used instead of tapping through the path. Out of the four who did not use this feature all the time, two used the swiping action to perform the tasks in the slideware presentation, while a third used swiping only for the skip situation. One of them later mentioned that she forgot the slide menu existed and another that she thought that the use of the slide menu was not allowed to use at all after she had been told not to use it for the regular navigation. Furthermore, two participants tapped through the path in two situations of the canvas presentation. One participant followed the path after being asked to skip to the end of the presentation, while another tapped through all players in the *search known* task.

Chapter 5

Discussion

This chapter discusses the results from the evaluation. For this, first the general findings are described before the specific comparisons are interpreted.

Overall, the results confirm the theory that canvas presentations provide more options for the presenter. A third of the participants chose to present the topic in a free manner and ignore the linear sequence at least partially. As six of the seven stated that they like to present (the seventh being neutral) and none of them had used a canvas presentation tool before, this raises the assumption that the amount of how much a person likes to present has an effect on the way presentations are held and how eager new features are tried out.

However, due to the correlation of the *like*-ratings to the *negative affect*-ratings, these factors can not be separated in following interpretations. The reason for this correlation might have arisen from the moment in time at which the PANAS test was filled out. This happened directly after a brief overview of the study was given, and therefore the participants knew that they would have to give presentations. If a person does not like to present, this might have lead to higher ratings on the items measuring the negative affect (e.g., *nervous*). Spatial ability and presentation age can also not be interpreted independently since their values correlate. As the presentation age has a high correlation with the age of the participant as well, this correlation is not unexpected—prior studies have often found a relationship between spatial ability and age (e.g., Ziefle [2008]; Ziefle et al. [2007]). Therefore, any of the following interpretations regarding the *like* ratings and *PAge* score can be attributed to the *negative affect* and *spatial ability* respectively.

In both the comparison between the different programs and the comparison between presentations with and without technological difficulties, it became apparent that participants with lower technological expertise felt significantly lower pleasure than participants with higher technological expertise. This consistent difference might be attributed to the novelty of the presentation format and the necessity to control the presentation using an *iPad*. It shows that if technology is used, the prior experience with technology affects the feelings during a presentation regardless of the presentation software.

5.1 Differences between the Programs

The results indicate that the feelings during presentations differ depending on the software used. This is especially true when considering different biographical characteristics (such as prior presentation and technological experience) of the presenter.

Familiarity Regarding the familiarity of the presentation formats, both programs are not very familiar for experienced presenters. Canvas presentations are equally unfamiliar for presenters who have presented for less than 20 years. However, they rate slideware as way more familiar. This might be attributed to the fact that slideware such as *PowerPoint* has been readily available when they started to present, while experienced presenters learned to present without slideware. Canvas presentations do not seem to benefit from this effect.

The familiarity is apparently also related to how much a person likes to present. While someone who likes to present rates both software as almost equally familiar, slideware is more familiar to those who do not like to present. This points to the conclusion that people who do not like to present might not be looking around for other programs but stick with the standard presentation tool instead. If this were the case, it would mean that the dislike of presenting does not stem from the presentation program but from the act of presenting in general.

Valence and arousal ratings The results indicate that presenters feel differently pleased and aroused during presentations depending on the used program. Overall, presenters feel more pleased and less aroused during a slideware presentation. While each dimension on its own does not differ significantly, the combination does and in a potentially stressful situation like presenting, such a difference is very important. The software used should aid the presenter and not increase the stress

level more than necessary. The qualitative responses comparing the two programs back this assumption. While the responses indicate that it would be more fun to present a canvas presentation and a slideware presentation is seen as more boring, participants also state that they felt they had better control over the slideware presentation. Generally, slideware seems to have an advantage compared to canvas presentations. However, varying presenter characteristics affect this experience differently and provide more detailed results.

In respect to the prior presentation experience, it appears as if less experienced presenters felt more pleasure using the canvas presentation, while experienced presenters felt more pleasure during the slideware presentation. These differences are interesting considering the familiarity ratings for each program. As both were equally unfamiliar for experienced presenters, the increased pleasure provides a clear advantage for slideware. On the other hand, for unexpected presenters the higher familiarity of slideware does not lead to higher pleasure. Instead the canvas presentation is rated as more pleasurable. This points to the conclusion that canvas presentations are more suitable for less experienced presenters while the opposite is the case for experienced presenters. Keeping in mind that the presentation age correlated with spatial ability, the capability of a person to visualize the spatial arrangement might also be the reason for these differences.

However, how much a person likes to present seems to have an effect on this relationship. The results show that a person who likes to present feels the same amount of pleasure regardless of presentation experience and software. On the other hand, presenters who do not like to present show the above mentioned relationship. Therefore, the conclusion that canvas presentations are better for less experienced and slidware for more experienced presenters should be limited to presenters who do not like to present. For them, this interpretation is also amplified by taking the differences in familiarity of the *Don't like* group into account. For example, presenters who do not like to present and have less experience presenting reported feeling more pleasure during the canvas presentation even though these kinds of presentations are very unfamiliar to them.

The technological expertise seems to have an effect on the canvas presentation experience as well. While the valence ratings of both high and low *TE* presenters are very similar for slideware, ratings for canvas presentations differ. Presenters with low *TE* felt less pleasure during the canvas presentation compared to the slideware presentation, whereas presenters with high *TE* experienced more pleasure during the canvas presentation. This might mean that the additional options that canvas presentations offer affect presenters with low and high *TE* differently. High *TE* presenters seem to like the increased freedom of interaction, while low *TE* presenters are negatively affected by them. Therefore, presenters should chose the presentation software to use with their technological expertise in mind.

An interesting result are the different ratings for the *search unknown* task in comparison to the overall rating. Here, especially presenters who fell either in the high *PAge* group or in the low *TE* group reported to feel more pleased in the *search unknown* situation compared to the overall experience. This points towards the interpretation that canvas presentations are helpful for search tasks even for presenters who otherwise do not feel positive during a canvas presentation. This might be attributed to the easy to reach overviews that allow a broad view on the information in the presentation. It should be more closely observed where the reasons for the less positive ratings of canvas presentations originate. If these problems could be eliminated while still retaining the freedom for search, canvas presentation tools might provide a better experience especially for the problematic user groups.

It was expected from visualizations that canvas presentation allow for better error recovery since the effects of wrong interactions would be more visible. Therefore, the impact of an error on the ratings was analyzed. This analysis did produce significant results, however, it should be kept in mind that the number of participants for each group is quite low since each participant encountered only one presentation with simulated problems. The data suggests that the impact of a technological error depends on the program used and how much a presenter likes to present. A study tailored to investigate the impact of errors might provide clearer answers.

Semantic differential ratings The semantic differential data did not produce significant results except for the satisfaction ratings in respect to the used program, presentation experience, and technological expertise. While the ratings of the high *TE* group are very similar to each other regardless of the program and the presentation age, the combination of low *TE* and high *PAge* lead to very low satisfaction ratings for the canvas presentation. Based on the graphs of the different groups (cf. figure 4.4), canvas presentations were rated worse over all dimensions. However, this data might be affected by outliers as there are only three participants who fall into this group.

Conclusion Regarding the differences between the presentation formats, the results show that the preference depends on prior experiences of the presenter. While Lichtschlag [2008] found that the free navigation helps the author to create a better presentation, the same freedom appears to affect presenters differently during a presentation. Since they are in a stressful situation, the simpler linear sequence of slideware might make the navigation easier for some presenters and it could be assumed that canvas presentations should be limited to a sequential format for the delivery as well. However, in situations such as the *search unknown* task, the free navigation is also beneficial. A solution might be a locking feature that restricts the movement to *next* and *previous* actions. A presenter could then activate the free

navigation by unlocking the features. Additionally, a simple way to get back to the presentation path could be added to enable the presenter to reach a safe place easily.

5.2 Impact of Technological Problems

This chapter discusses the effects that technological problems have on the feelings during a presentation.

The results indicate that arousal of the presenter was higher during the *no error* condition. Taking the technological expertise into account, it becomes clear that the valence rating of the *error* condition is rated similar, whereas the valence during the *no error* presentation was lower for presenters with less technological expertise than the valence rating of presenters with high technological expertise. These results are hard to interpret. The increased arousal in the *no error* condition might be explained by a *more intense positive feeling* since the valence is in both conditions roughly the same. This connection could also help to clarify the correlation of arousal and *positive affect*. In the *error* presentation, participants with more *positive affect* would therefore experience the pleasure with more intensity, while there would be no effect in the *no error* presentation. However, there is no explanation for the valence differences between the *low TE* group and the *high TE* group in the *no error* condition and why these differences are not apparent in the *error* condition.

Since overall ratings are used here to differentiate between presentation with and without technological difficulties, an alternative interpretation is that other events during the presentations affected the overall rating. An analysis of the raw valence ratings reveals two very low scores of participants from the *low TE* group for the overall rating of the *no error* condition. Based on the qualitative data, one of these likely originated from the used program and the opinion of the participant about it. The other score came from a participant who had severe problems controlling the application during the start of the presentation. Therefore, it is plausible that these two ratings were influenced by other events during the presentation. Consequently, the significant interaction effect of error condition and technological expertise might not be based on the occurrence of the errors.

As no other significant differences were found between the error conditions, this means that the occurrence of errors had no effect on the ratings. This is a contradiction of the qualitative statements in which the participant clarified that technological problems are the primary reason why a presentation is experienced as *bad*. However, the participants also noted that their feelings were not as strong as in a real presentation. Therefore, the reason why no differences between the conditions

were found might be attributed to the difference in experience between a lab study and a field study.

The qualitative responses made clear that technological problems are indeed a very important factor. Consequently, anything that reduces the risks of such problems happening is a good thing. Presentation software should take special care to avoid technological problems that originate from the software and to assist the presenter by providing easy recovery options in case something goes wrong.

5.3 Skin Conductance Response and Validity of the Study

No relationship between the skin conductance responses and the arousal or valence ratings were found in this study. This stands in contrast to previous studies of this relationship (cf. chapter 2.2.2). A number of reasons might explain this missing correlation. First, the sensor might not have been sensitive enough to track the changes or produced too much noise. Second, the formula to calculate the reaction might be incorrect. However, Greenwald et al. [1989] found correlations between skin conductance response and arousal using this formula. Another reason might be that other influences canceled experienced reactions. For example, Beatty and Behnke [1991] found that public speaking trait anxiety and heart rate were correlated under low-intensity conditions but not under high-intensity conditions. Finally, the situation in the study might not have triggered the same bodily response that would occur during a real world presentation or—as feelings can take place without bodily responses—presenting might affect only the subjective experience.

Overall, the participants reported that they felt similar to a real presentation during the study. However, since the presentation was not graded or otherwise evaluated and it was made clear that the recordings are not shown to anyone, they experienced less intensive feelings. This was expected as the setup describes a low-intensity situation according to Beatty and Behnke [1991]. Therefore, any findings from this study are likely to be more prominent in a real-world presentation. Furthermore, there are likely additional differences, and to fully understand the experience of a presenter using canvas presentations it is necessary to observe and evaluate such presentations outside of a lab.

Chapter 6

Limitations

In this chapter the limitations in regard to study design and evaluation are discussed. By design, the study was a lab study, and therefore represents a controlled situation that might not be representative of a real-world setting. Although the participants stated that they felt similar to a real presentation, the topic and visualizations were not chosen by them. For that reason, they had a shorter amount of time than normal to prepare for the talk, while also needing to get familiar with the controls of the applications. This might produce some deviation in feelings from a presentation in their usual situation. Furthermore, most of the participants had never used a canvas presentation tool before. Consequently, this represented a new situation to them which might have affected the feelings during presentations. This novelty effect was addressed by using an *iPad* for the slidware presentation as well, however, previous experience using an *iPad* for this task or *iPad* use in general have not been evaluated. The *Keynote* application on the *iPad* showed a consistent delay between the presenter input and the reaction. Several participants complained about this and as it is not linked to the presentation format a program without such a lag should be used for evaluation in further studies.

The topic selection might also influence the validity of the results. Even though the *FIFA World Cup* was featured prominently in the media in Germany, the prior knowledge about the topic might influence how comfortable a person feels while presenting it. The familiarity of the participant with soccer in general and the *FIFA World Cup* in particular was not assessed. So it is not possible to rule out that some participants felt more comfortable since they were talking about information they already knew while others had to rely on the information provided.

Regarding the creation of the presentations, there is also room for improvement. Some of the participants tried to use the slide menu for navigation all the time

instead of using the next/previous commands (tapping and sliding). However, this style of slide selection does not play the animations during slide changes and on the slide itself. Therefore, simulated errors like the *"freezing"* error would not occur. In the present study the participants were asked to close the slide menu as soon as their behavior was noted. An easy fix for this problem would be to include incremental revealing on the slide, therefore forcing the participant to close the side bar as the animations can not be triggered with an open slide menu.

A further limitation is provided by the selection of the participants. While the order of the presentation software and error conditions was balanced, the assignment of participants to these conditions was not. This resulted in uneven group compositions as differences in user attributes were not accounted for before the study. Furthermore, in the present study the spatial ability score has a correlation with the age and the presentation age of the participant. While this correlation is not surprising, it does not allow to distinguish between the different influences that these characteristics might have on feelings during the delivery of a canvas presentation. A prior screening of participants might have helped to avoid this correlation.

A revised calculation for the presentation age and the technological expertise might improve the estimation of presentation expertise and technological proficiency. For the current study, the presentation age was based solely on the difference between the first time a person presented and her current age. An inclusion of the frequency of presenting might provide a better measure for the presentation expertise. Such a frequency was recorded but many teachers stated to present only multiple times per year—although it had been clarified at the beginning of the study that this does not only include presentations assisted by presentation software. As this seemed not correct, it was decided to discard this attribute from the calculation of the presentation age, since it was not clear whether the other participants used the correct definition. Providing a clearer definition of a presentation right before the questionnaire might prevent this misunderstanding. The technological expertise calculation can also be improved. In the form used in this study, the calculation relies only on the usage of presentation software. A broader definition of technological experience and capabilities might sharpen this measure. A more exact definition for these two measures in combination with higher user numbers might allow to create more categories than only a high and low separation.

While the separation of participants into two groups for the presentation age was provided by the distribution of participants, median splits were used for technological expertise and how much a participant likes to present. This distinction is more arbitrary as it breaks up participants who reach scores close to the splitting value. In this study it was necessary for the statistical evaluation but a study with more participants might be able to analyze more distinct groups.

While the self-assessment manikin results showed significant effects for various comparisons, the semantic differential data did not. Therefore, it appears as if the goal to provide greater detail on the characteristics of the feelings was not reached. This might have several reasons. First, the dimensions used in the semantic differential maybe did not include the relevant dimensions that actually varied between the situations. Second, the amount of dimensions used might have been too high and some of them might have measured the same concept. Since a multivariate analysis was used, more dependent variables increase the difficulty to find significant effects. If many of the dependent variable measure the same dimension, this makes it less likely to find significant effects. A way to solve this would be to perform additional steps for the creation and validation of the semantic differential. For example, prior tests and factor analysis could be performed to find and filter the dimensions to include only those that measure different aspects.

The triangulation of self-report ratings and bodily response in the form of skin conductance response did not produce a connection. However, as other studies found these connections, this might have technological reasons. Furthermore, the sensor used to measure the skin conductance broke down after about half of the participants.

Another point that needs to be kept in mind is that the overall ratings were used as the description of the feelings over the presentation. However, although these ratings correlate with the mean of the ratings of each situation, the analysis in the case of *error* vs. *no error* presentations showed that other situations during the presentation effected the overall ratings. Therefore, a follow-up study might want to view the situations independently instead of comparing the overall experience.

For the statistical analysis, parametric tests like MANOVA were used even though the assumptions are not met. This increases the possibility of incorrect conclusions. Therefore, the results found in this study should be seen as tendencies and further studies need to be conducted to validate these findings.

Chapter 7

Summary and Future Work

This chapter provides a summary of the findings of this thesis and how they have been obtained (chapter 7.1). The final chapter discusses possible directions for further research and outlines the questions that need to be answered (chapter 7.2).

7.1 Summary

This thesis addressed the question of how a presenter feels during a canvas presentation. Chapter 2.1 introduced canvas presentations as an alternative to slideware. While slideware would suffer from the limitation to a linear structure and the missing global picture (chapter 2.1.2), canvas presentation would allow users to arrange and present the information freely by placing the content on an infinite canvas (chapter 2.1.3). Previous studies found that authors of presentations benefit from the canvas approach and that audiences feel as if they could follow a canvas presentation more easily. The effect that this new format has on the presenter, however, had not been evaluated yet.

In order to understand the concept of feelings, chapter 2.2 looked at the research of emotions and feelings. Emotions are seen as a multi-faceted reaction to an event and one way to classify them is to use two dimensions—valence (pleasure) and arousal. Feelings are defined as the subjective experience of an emotion and according to Scherer [2005], self-reports are the only way to record such experiences. Two self-report techniques are presented in chapter 2.2.2 "Self-Reports"—the semantic differential technique by Osgood et al. [1957] and the self-assessment manikin by Bradley and Lang [1994]. The semantic differential technique uses a set of bipolar

word pairs (e.g., *positive–negative*) for which a participants marks on a scale between the poles where she places a certain concept (e.g., life), thereby creating a profile for the concept. This technique has been used to measure feelings and has shown links to the valence and arousal dimensions of emotions. The self-assessment manikin measures these dimensions directly using images.

Chapter 2.2.2 "Physiological Measurements" summarized different approaches that use physiological reactions to assess emotions. Some of the research found correlations to the valence and arousal dimensions of an emotion. For example, skin conductance response has been correlated with arousal and facial muscle activity with valence.

The study conducted to evaluate how a presenter feels during a canvas presentation was described in chapter 3. During this study, participants gave two presentations—one with a canvas presentation tool and one with slideware. They were faced with presentation specific situations such as searching tasks and technological problems. A recording of the presentation was reviewed with the participant and she reported her feelings for each situation and overall using a customized semantic differential scale and the self-assessment manikin.

Chapter 4 explored the hypotheses (e.g., that canvas presentations are rated differently than slideware presentations), and chapter 5 discussed the results. These results suggest that biographical characteristics of participants have a great impact the feelings during a presentation in general and a canvas presentation in particular. For example technological expertise appears to affect the pleasure regardless of the used program, with presenters who have more technological expertise feeling more pleased. Based on the results, differences in technological expertise also lead to different feelings during canvas or slideware presentations. It seems as if especially presenters with less technological expertise feel less pleased using the canvas presentation tool than presenters with high technological expertise. A similar relationship has been found for the prior presentation experience. More experienced presenters appear to feel more pleasure using slideware, while less experienced presenters feel more pleased during the canvas presentation. Also, how much a person likes to present affects this relationship. Whereas presenters who do not enjoy presenting show the relationship, presenters who like to present are not affected by the presentation software used. However, as presentation experience and spatial ability are correlated in the present study, any interpretations regarding the characteristic might also be attributed to the other.

Further analysis of the search task shows that especially experienced presenters and people with less technological expertise felt more pleasure during the search task compared to the overall experience. Therefore, this suggests that the free navigation during search tasks is a very helpful feature of canvas presentations.

Qualitative feedback of the participants indicates that technological problems are the main reason why a presenter experiences a bad presentation. Therefore, presentation software should take special care to avoid technological problems and assist the presenter if they should occur. Finally, the participants stated that they experienced similar albeit muted feelings during the presentations in the study compared to real-world presentations. This suggests that the findings of this study can be expected to apply to real-world presentations as well. However, especially a presentation for which the presenter is graded or otherwise judged, might have a great impact on the experienced feelings, and therefore studies should be conducted to observe canvas presentations in real situations.

Chapter 6 discussed the limitations of the study conducted and its results. It is suggested to improve the selection of participants to be able to evaluate the effects of different characteristics in more detail. For example, spatial ability and presentation experience could not be distinguished in the sample of the current study. A more extensive screening of the participants prior to the study might allow to attribute different effects to these characteristics. Furthermore, the presentation experience and technological expertise were calculated using rudimentary information. In future studies, further information about the participants could be gathered to improve the calculations.

Another point of critique is that the semantic differential did not produce significant insights into how the participants felt during the presentations. A more thorough development and testing of the scales might help to improve this tool to provide more meaningful results.

7.2 Future Work

As the study described in this thesis was a lab study, a field study could corroborate the results. In the study the participants were given a topic to present and had to use prepared visuals. Furthermore, there was no audience except the moderator present. A study of participants who use their own visuals and use their own topic in a real-world scenario could provide further insights into how presenters feel during canvas presentations. Since the study presented in this thesis found reactions for the subjective experience but no bodily reactions, such a study could also observe this connection to see whether presenting only affects subjective feelings or bodily reactions as well.

An improved screening process might also help to gather more insight into how the spatial ability influences the feelings during a canvas presentation. In the study, it was not possible to distinguish between presentation age and spatial ability, and

therefore it is not possible to attribute findings to either of them. Since canvas presentation tools claim to be especially helpful in the way they allow to arrange content spatially, a future study could investigate this claim specifically.

The analysis showed that different user characteristics affect the experience in different ways. Future investigations should focus on specific groups to understand why they are feeling in a certain way and find solutions to improve their experience. Such groups might be people who do not enjoy presenting or people who do not present very often. Understanding such user groups might provide valuable insights into what aspects are important for their experience and how their feelings during presentations could be improved.

A more extensive study could also implement the proposed changes to the navigation options (such as the locking feature) and check whether they affect the feelings during canvas presentations. While the evaluation of this study focused on the results from the semantic differential and self-assessment manikin, qualitative responses in general could be evaluated more closely to improve the features of canvas presentation tools.

In order to better understand what the important aspects for the subjective experience during a presentation are, regression analysis could be used on the semantic differential data and the self-assessment manikin ratings. This might provide insight into which dimensions of the semantic differential have the greatest effect on the experienced pleasure and arousal.

Furthermore, it could be also investigated which situations are the most important factors for the overall experience. In order to do this, the ratings of the individual situations could be compared to the overall ratings and a regression analysis could calculate the influence that a situation has on the overall feelings.

Appendix A

Wording Survey

This appendix lists both versions of the wording survey that was used to elicit the dimensions for the semantic differential. Figure A.1 shows the German version that was sent out to participants, while figure A.2 presents the English version.

Erhebung von Gefühlen und Reaktionen während Präsentationen

Für meine Masterarbeit möchte ich die Gefühle von Vortragenden untersuchen.
Hierzu ist es zunächst wichtig festzustellen, was für Gefühle es während einer Präsentation geben kann bzw. mithilfe welcher Kriterien Situationen während einer Präsentation beschrieben werden können.

In den Textfeldern können Sie Ihren Ideen freien Lauf lassen. Sollte Ihnen zu einem späteren Zeitpunkt noch etwas einfallen, können Sie das auch gerne in einer weiteren Antwort einreichen.

Ich freue mich auf Ihre Ideen und bedanke mich recht herzlich für die Teilnahme :)

Wenn Sie auf vergangene Präsentationen zurückblicken: Was für Gefühle und Reaktionen haben Sie selbst schon einmal während einer Präsentation erlebt?

Was für Gefühle und Reaktionen können Sie sich generell für einen Vortragenden während einer Präsentation vorstellen?
Zum Beispiel kann man fröhlich sein, da das Publikum mitarbeitet, oder zornig darüber, dass es nicht mitmacht. Ebenso kann man Stolz darüber sein, etwas Gutes abgeliefert zu haben.

Nach welchen anderen Kriterien kann man Situationen beschreiben?
Neben genaueren Gefühlsreaktionen können Situationen auch allgemeiner beschrieben werden. Zum Beispiel kann man eine Situation generell als positiv oder negativ bezeichnen. Welche weiteren Beschreibungsmerkmale fallen Ihnen noch ein?

Senden

Figure A.1: German version of the wording survey to elicit the dimensions for the semantic differential.

Gathering feelings and reactions that occur during presentations

For my master's thesis I want to take a look at the feelings of presenters.
In order to do this I want to gather feelings that might occur during a presentation as well as criteria that can be used to describe situations happening while giving a presentation.

You can let your ideas run wild in the textfields below ;) If you think of something else at a later time, you are kindly invited to provide another entry.

I'm looking forward to your ideas and thank you very much for your participation.

When looking back to past presentations: What feelings and reactions have you experienced yourself during presentations?

What feelings and reactions of a presenter can you imagine in general?
For example someone can be happy that the audience participates, on the other hand they might be angry that they don't. It is also possible to be proud about delivering a great talk.

What other criteria to describe situations can you think of?
There are other, more broad descriptions of situations possible. For example a situation can be described as positive or negative in general. What other criteria come into your mind?

Senden

Figure A.2: English version of the wording survey to elicit the dimensions for the semantic differential.

Appendix B

Study Documents

This appendix includes the scales and questionnaires used in the study.

- Figure B.1 is the consent form that the participants had to sign prior to the study.
- The PANAS test to measure the mood (positive and negative affect) is shown in figure B.2.
- For each situation, the participants had to report their feeling using the semantic differential (figure B.3) and the self-assessment manikin (figure B.4).
- Figure B.5 shows the instructions for the paper folding test to measure the spatial ability.
- The final questionnaire that the participants had to fill out (for personal information and prior experience) is displayed in figure B.6.

Einverständniserklärung

Evaluierung der Gefühle während einer Canvas Präsentation

STUDIENLEITER Philipp Wacker
Media Computing Group
RWTH Aachen University
Telefon: 0151 20774223
Email: wacker@cs.rwth-aachen.de

Ziel der Studie: Das Ziel der Studie ist es, zu verstehen, wie es sich für den Vortragenden anfühlt, eine Präsentation mit einem Canvas Präsentations Programm zu geben. Die Teilnehmer werden gebeten, unter Anwendung von zwei verschiedenen Präsentationsprogrammen, einen Vortrag zu halten. Videoaufnahmen, Leitfähigkeit der Haut, Fragebögen und Interviews werden in der Analyse ausgewertet.

Ablauf: Die Teilnahme an der Studie besteht aus mehreren Phasen. Zuerst erfahren Sie Details über die verwendeten Präsentationsprogramme und haben Gelegenheit diese auszuprobieren. Mit beiden Programmen werden Sie dann einen kurzen Vortrag über die Aufstellung der deutschen Fußball Nationalmannschaft halten. Nach jedem Vortrag werden Sie das Video anschauen und gebeten zu bestimmten Stellen Auskunft über Ihre Gefühle in der Situation zu geben. Im Anschluss folgt ein Interview und ein Test für das räumliche Vorstellungsvermögen. Diese Studie sollte etwa 1,5 - 2 Stunden dauern.

Risiken/Beschwerden: Es könnte sein, dass Sie die Teilnahme an der Studie ermüdet. Sie werden mehrere Gelegenheiten haben, sich zu erholen; zusätzliche Pausen sind ebenfalls möglich. Es sind keine weiteren Risiken im Zusammenhang mit der Studie bekannt. Sollten die Aufgaben oder der Fragebogen zu anstrengend für Sie sein, können Sie die Bearbeitung sofort abbrechen.

Nutzen: Die Resultate der Studie werden genutzt, um den Einfluss von Canvas Präsentationen auf den Vortragenden besser zu verstehen.

Alternativen zur Teilnahme: Die Teilnahme an der Studie ist freiwillig. Es steht Ihnen frei, Ihre Teilnahme zurückzuziehen oder abzubrechen.

Kosten und Entschädigung: Die Teilnahme an der Studie wird Ihnen keinerlei Kosten verursachen. Während und nach ihrer Teilnahme werden für Sie Getränke und Snacks bereitstehen. Wenn Sie an der Verlosung des 20€ Amazon-Gutscheins teilnehmen wollen, geben Sie bitte unten Ihre E-Mail Adresse an.

Vertraulichkeit: Alle Informationen, die während der Studienphase gesammelt werden, werden streng vertraulich behandelt. Ihre Daten werden nur durch Identifikationsnummern identifiziert. Keine Publikationen oder Berichte aus diesem Projekt werden personenbezogene Informationen über die Teilnehmer beinhalten. Wenn Sie sich bereit erklären, an dieser Studie teilzunehmen, unterschreiben Sie bitte unten.

_____ Ich habe die Hinweise auf diesem Formular gelesen und verstanden.
_____ Man hat mir die Hinweise auf dem Formular erklärt.

Name des Teilnehmers	Unterschrift des Teilnehmers	Datum
	Studienleiter	Datum

_____ Ich möchte an der Verlosung des 20€ Amazon-Gutscheins teilnehmen

E-Mail Adresse: _____

Wenn Sie Fragen zu dieser Studie haben, wenden Sie sich bitte an Philipp Wacker unter 0151 20774223, Email: wacker@cs.rwth-aachen.de

Figure B.1: Consent form informing the participant about the purpose and structure of the study.

Participant ID: _____

Dieser Fragebogen enthält eine Reihe von Wörtern, die unterschiedliche Gefühle und Empfindungen beschreiben. Lesen Sie jedes Wort und tragen dann in die Skala neben jedem Wort die Intensität ein. Sie haben die Möglichkeit, zwischen fünf Abstufungen zu wählen.

Geben Sie bitte an, wie Sie sich an diesem Tag fühlen.

	ganz wenig oder gar nicht	**ein bisschen**	**einigermassen**	**erheblich**	**äusserst**
aktiv	☐	☐	☐	☐	☐
bekümmert	☐	☐	☐	☐	☐
interessiert	☐	☐	☐	☐	☐
freudig erregt	☐	☐	☐	☐	☐
verärgert	☐	☐	☐	☐	☐
stark	☐	☐	☐	☐	☐
schuldig	☐	☐	☐	☐	☐
erschrocken	☐	☐	☐	☐	☐
feindselig	☐	☐	☐	☐	☐
angeregt	☐	☐	☐	☐	☐
stolz	☐	☐	☐	☐	☐
gereizt	☐	☐	☐	☐	☐
begeistert	☐	☐	☐	☐	☐
beschämt	☐	☐	☐	☐	☐
wach	☐	☐	☐	☐	☐
nervös	☐	☐	☐	☐	☐
entschlossen	☐	☐	☐	☐	☐
aufmerksam	☐	☐	☐	☐	☐
durcheinander	☐	☐	☐	☐	☐
ängstlich	☐	☐	☐	☐	☐

Figure B.2: Instructions and scales for the PANAS test to measure the mood of the participant.

Participant ID: ________ Situation ID: ________

1. Bitte markieren Sie zwischen den folgenden Wortpaaren, wo Sie Ihr erlebtes Gefühl in der gezeigten Situation einordnen:

Beispiel:

A										B

A			X							B

sehr nervös										nicht nervös
sehr verwundert										nicht verwundert
unsicher										sicher
unangenehm										angenehm
negativ										positiv
sehr ängstlich										nicht ängstlich
unzufrieden										zufrieden
traurig										glücklich
sehr gestresst										nicht gestresst
sehr verzweifelt										nicht verzweifelt
kontrolliert										kontrollierend
verloren										nicht verloren

Figure B.3: Final semantic differential scale that was used in the study.

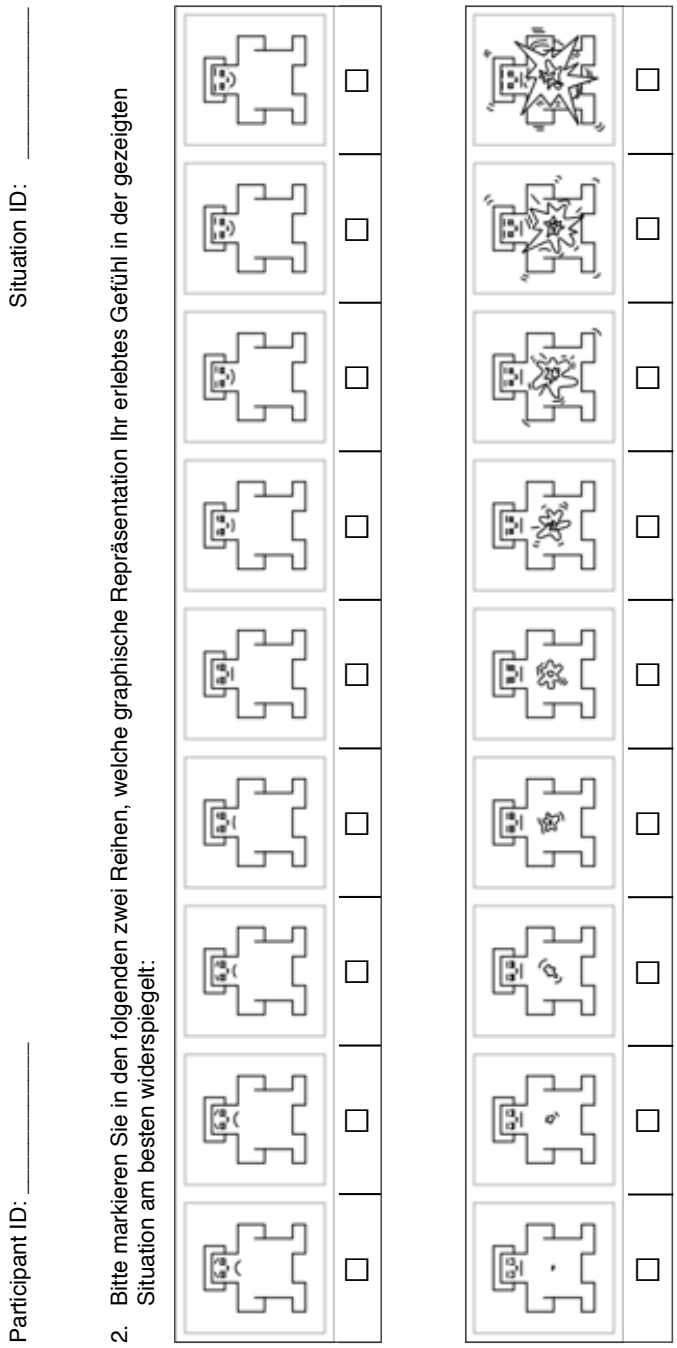
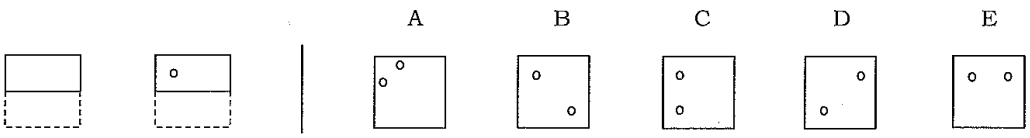


Figure B.4: Self-assessment manikin images with instructions (images taken from Irtel [2007])

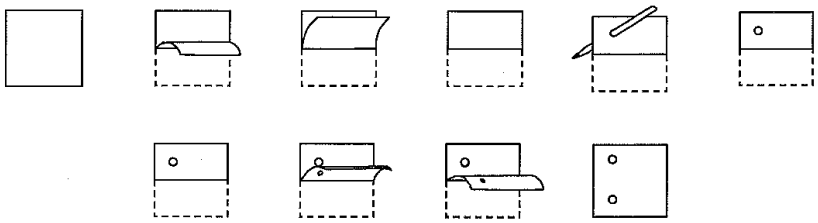
Participant ID: _____

In diesem Test sollen Sie sich das Falten und Entfalten von Papier vorstellen. Für jedes Problem in diesem Test sind einige Darstellungen auf der linken Seite der vertikalen Linie gezeichnet und andere auf der rechten Seite. Die Darstellungen auf der linken Seite zeigen, wie ein quadratisches Papier gefaltet wird und in das letzte Bild sind ein oder zwei kleine Kreise eingezeichnet. Diese zeigen an, wo das Papier durchlöchert wurde. Jedes Loch geht durch die gesamte Dicke des Papiers an dieser Stelle.
Eine der fünf Darstellungen auf der rechten Seite zeigt an, wo die Löcher sein werden, wenn das Papier komplett entfaltet wird.
Sie sollen sich entscheiden, welche Darstellung korrekt ist und sollen diese mit einem X markieren.

Versuchen Sie nun dieses Beispiel. (Hier wurde nur ein Loch durch das Papier gestoßen).



Die richtige Antwort für dieses Beispiel ist C und Sie sollten dieses mit einem X markiert haben. Die folgenden Darstellungen zeigen, wie das Papier gefaltet wurde und warum C die richtige Antwort ist.



Bei den folgenden Problemen wird auf der linken Seite jeder Faltungsschritt gezeigt und das Papier wird nicht gedreht oder bewegt bis auf die gezeigten Faltungen.
Denken Sie daran, dass die Antwort die Darstellung ist, die die Positionen der Löcher anzeigt, wenn das Papier komplett aufgefaltet wurde.

Einige der Probleme in diesem Test sind schwerer als andere. Wenn Sie ein Problem nicht lösen können, überspringen Sie dieses und machen mit dem nächsten weiter.

Sie haben drei Minuten für jeden der zwei Teile dieses Tests. Jeder Teil hat eine Seite. Wenn Sie mit Teil 1 fertig sind, STOP. Bitte fangen Sie nicht mit Teil 2 an, bevor Sie darum gebeten werden.

BLÄTTERN SIE NICHT WEITER, BEVOR SIE DARUM GEBETEN WERDEN

Figure B.5: Translation of the instructions and example for the Paper Folding test by Ekstrom et al. [1976] to measure the spatial ability.

Participant ID: _____

Zum Abschluss beantworten Sie bitte noch folgende Fragen zu Ihrer Person bzw. Ihren Vorkenntnissen:

Was ist ihre Beschäftigung (Beruf/Studium)?					
Wie alt sind Sie?					
Geschlecht	männlich ☐		weiblich ☐	keine Angabe ☐	
Wie lang ist Ihre letzte Präsentation her (ausgenommen heute)?	Eine Woche ☐	Einen Monat ☐	Ein Jahr ☐	Länger ☐	Nie ☐
Wie häufig präsentieren Sie?	Mehrmals in der Woche ☐	Mehrmals im Monat ☐	Mehrmals im Jahr ☐	Seltener ☐	Nie ☐
In welchem Kontext präsentieren Sie normalerweise?					
Mit welchem Alter haben Sie angefangen, Präsentationen zu halten?					
Mit welchem Alter haben Sie am häufigsten präsentiert?					
Wie häufig haben Sie bereits Canvas Presentation Tools wie Prezi benutzt?	Mehrmals in der Woche ☐	Mehrmals im Monat ☐	Mehrmals im Jahr ☐	Seltener ☐	Nie ☐
Wie häufig haben Sie bereits Slide Presentation Tools wie PowerPoint / Dia Projektoren benutzt?	Mehrmals in der Woche ☐	Mehrmals im Monat ☐	Mehrmals im Jahr ☐	Seltener ☐	Nie ☐
Wie gern halten Sie Präsentationen?	Sehr gern ☐	☐	Neutral ☐	☐	Sehr ungern ☐

Herzlichen Dank für Ihre Teilnahme :)

Figure B.6: Questionnaire to record demographic information and prior presentation experience.

Appendix C

Program×*PAge*×*L* Interaction Graphs

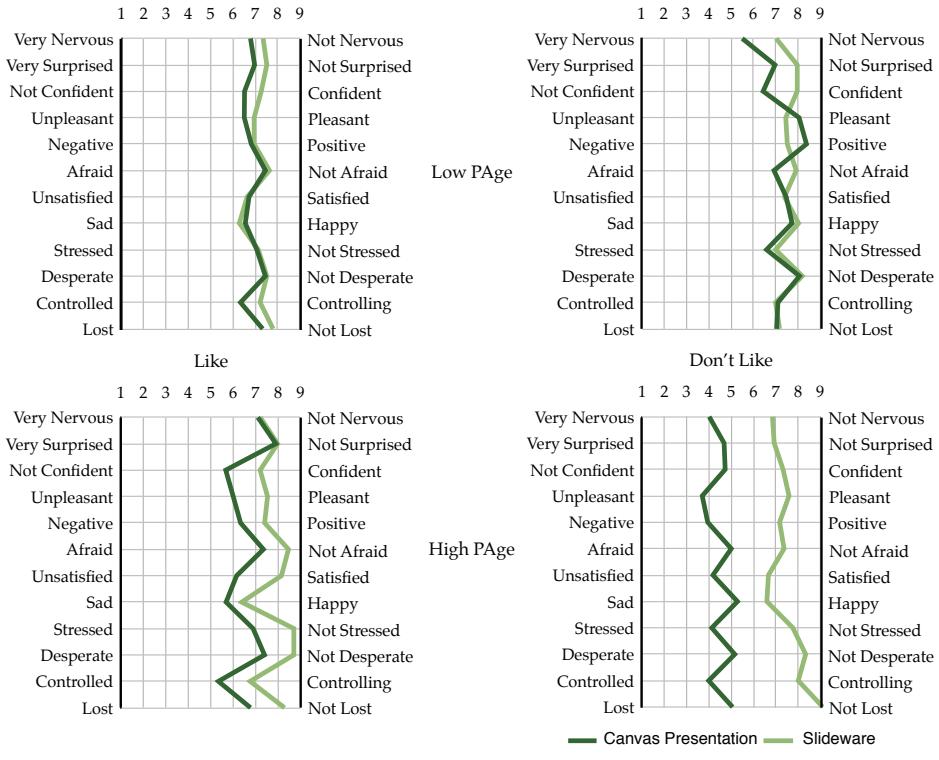


Figure C.1: *Program*×*PAge*×*L* Interaction Graphs. They suggest that people with *high PAge* but from the *Don't Like* group experience more negative feelings during the canvas presentation. However, this group consists only of two persons.

Bibliography

Arning, K. and Ziefle, M. (2009). Effects of age, cognitive, and personal factors on PDA menu navigation performance. *Behaviour & Information Technology*, 28(3):251–268.

Bachorowski, J.-A. and Owren, M. J. (2008). Vocal expressions of emotion. In Lewis, M., Haviland-Jones, J. M., and Feldman Barrett, L., editors, *Handbook of Emotions*, pages 196–210. Guilford Publications, Inc., New York, 3rd edition.

Baxter, K. and Courage, C. (2005). *Understanding your users: A practical guide to user requirements methods, tools, and techniques*. Morgan Kaufmann Publishers Inc., San Francisco.

Beatty, M. J. (1988). Situational and predispositional correlates of public speaking anxiety. *Communication Education*, 37(1):28–39.

Beatty, M. J. and Behnke, R. R. (1991). Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance. *Human Communication Research*, 18(2):147–176.

Bemtgen, C. (2012). *Fly Remote: Reciting canvas presentations with an iPad*. Bachelor's thesis, RWTH Aachen University.

Blascovich, J. J., Vanman, E., Mendes, W. B., and Dickerson, S. (2011). *Social Psychophysiology for Social and Personality Psychology*. SAGE Publications Ltd, London.

Blokzijl, W. and Andeweg, B. (2005). The effects of text slide format and presentational quality on learning in college lectures. In *Professional Communication Conference, 2005. IPCC 2005. Proceedings. International*, pages 288–299. IEEE.

Blokzijl, W. and Naeff, R. (2004). The instructor as stagehand: Dutch student responses to PowerPoint. *Business Communication Quarterly*, 67(1):70–77.

Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59.

Bradley, M. M. and Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In Lane, R. D. and Nadel, L., editors, *Cognitive neuroscience of emotion*, Series in affective science., pages 242–276. Oxford University Press, New York.

Cacioppo, J. T. and Gardner, W. L. (1999). Emotion. *Annual Review of Psychology*, 50:191–214.

Clore, G. L. and Robinson, M. D. (2012). Knowing our emotions: How do we know what we feel? In Vazire, S. and Wilson, T., editors, *Handbook of Self-Knowledge*, pages 194–209. Guilford Press, New York.

Craig, A. D. (2008). Interoception and emotion – a neuroanatomical perspective. In Lewis, M., Haviland-Jones, J. M., and Feldman Barrett, L., editors, *Handbook of Emotions*, pages 272–288. Guilford Publications, Inc., New York, 3 edition.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3):169–200.

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). *Manual for Kit of Factor-referenced Cognitive Tests*. Educational Testing Service, Princeton.

Evans, G. T. (1970). Use of the semantic differential technique to study attitudes during classroom lessons. *Interchange*, 1(4):96–106.

Farkas, D. K. (2009). Managing three mediation effects that influence PowerPoint deck authoring. *Technical Communication*, 56(1):28–38.

Frijda, N. H. (2008). The psychologists' point of view. In Lewis, M., Haviland-Jones, J. M., and Feldman Barrett, L., editors, *Handbook of Emotions*, pages 68–87. Guilford Publications, Inc., New York, 3rd edition.

Good, L. E. (2003). *Zoomable user interfaces for the authoring and delivery of slide presentations*. PhD thesis, University of Maryland.

Gravetter, F. and Forzano, L.-A. (2011). *Research Methods for the Behavioral Sciences*. Cengage Learning, Belmont.

Greenwald, M. K., Cook, E. W., and Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology*, 3(1):51–64.

Hegarty, M., Kriz, S., and Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21(4):325–360.

Heise, D. R. (1969). Some methodological issues in semantic differential research. *Psychological Bulletin*, 72(6):406–422.

Hess, T. (2011). *Expressive and Conveying Planar Presentations*. Master's thesis, RWTH Aachen University.

Higgins, R. L., Alonso, R. R., and Pendleton, M. G. (1979). The validity of role-play assessments of assertiveness. *Behavior Therapy*, 10(5):655–662.

Hunter, W. J. and Begoray, J. (1990). A framework for the activities involved in the writing process. *The writing notebook*, 8(1).

Irtel, H. (2007). PXLab: The psychological experiments laboratory. University of Mannheim. http://www.pxlab.de (accessed 05.09.2014).

Keltner, D. and Ekman, P. (2000). Facial expression of emotion. In Lewis, M. and Haviland-Jones, J. M., editors, *Handbook of Emotions*, pages 236–249. Guilford Publications, Inc., New York, 2nd edition.

Kern, J. M., Miller, C., and Eggers, J. (1983). Enhancing the validity of role-play tests: A comparison of three role-play methodologies. *Behavior Therapy*, 14(4):482–492.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., and Tausch, A. (1996). Untersuchung mit einer deutschen form der positive and negative affect schedule (PANAS). *Diagnostica*, 42:139–156.

Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3):261–273.

Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., and Cacioppo, J. T. (2008). The psychophysiology of emotion. In Lewis, M., Haviland-Jones, J. M., and Feldman Barrett, L., editors, *Handbook of emotions*, pages 180–195. Guilford Publications, Inc., New York, 3rd edition.

Laufer, L., Halacsy, P., and Somlai-Fischer, A. (2011). Prezi meeting: collaboration in a zoomable canvas based environment. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 749–752, New York. ACM.

Lichtschlag, L. (2008). *Fly: an organic authoring tool for presentations*. Diploma thesis, RWTH Aachen University.

Lichtschlag, L., Hess, T., Karrer, T., and Borchers, J. (2012a). Canvas presentations in the wild. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 537–540, New York. ACM.

Lichtschlag, L., Hess, T., Karrer, T., and Borchers, J. (2012b). Fly: Studying recall, macrostructure understanding, and user experience of canvas presentations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1307–1310, New York. ACM.

Lichtschlag, L., Karrer, T., and Borchers, J. (2009). Fly: a tool to author planar presentations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 547–556, New York. ACM.

Mauss, I. B. and Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, 23(2):209–237.

Moscovich, T., Scholz, K., Hughes, J. F., and Salesin, D. (2004). Customizable presentations. Technical report, Computer Science Department, Brown University.

OED Online (2014). "presentation, n.". Oxford University Press. http://www.oed.com/view/Entry/150685 (accessed 05.09.2014).

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The measurement of meaning*. Univer. Illinois Press, Oxford, England.

Robinson, M. D. and Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6):934–960.

Russell, J. A. and Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273–294.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4):695–729.

Thielsch, M. T. and Perabo, I. (2012). Use and evaluation of presentation software. *Technical Communication*, 59(2):112–123.

Tufte, E. R. (2006). *The Cognitive Style of PowerPoint: Pitching Out Corrupts Within*. Graphics Press, Cheshire, 2nd edition.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070.

Ziefle, M. (2008). Instruction formats and navigation aids in mobile devices. In Holzinger, A., editor, *HCI and Usability for Education and Work*, number 5298 in Lecture Notes in Computer Science, pages 339–358. Springer-Verlag, Berlin Heidelberg.

Ziefle, M., Schroeder, U., Strenk, J., and Michel, T. (2007). How younger and older adults master the usage of hyperlinks in small screen devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 307–316, New York. ACM.

Index

Typeset October 7, 2014

E r k l ä r u n g

Hiermit erkläre ich, dass ich die vorgelegte

- ☐ Hausarbeit*
- ☐ Bachelorarbeit*
- ☐ Masterarbeit*
- ☐ Magisterarbeit*
- ☐ Schriftliche Hausarbeit (Staatsarbeit)*

selbständig verfasst und - einschließlich eventuell beigefügter Abbildungen und Skizzen - keine anderen als die im Literaturverzeichnis angegebenen Quellen, Darstellungen und Hilfsmittel benutzt habe. Dies gilt in gleicher Weise für gedruckte Quellen wie für Quellen aus dem Internet.

Ich habe alle Passagen und Sätze der Arbeit, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, in jedem einzelnen Fall unter genauer Angabe der Stelle ihrer Herkunft (Quelle, Seitenangabe bzw. entsprechende Spezifizierung) deutlich als Entlehnung gekennzeichnet.

Außerdem erkläre ich, dass die vorgelegte Arbeit zuvor weder von mir noch - soweit mir bekannt ist - von einer anderen Person an dieser oder einer anderen Universität eingereicht wurde.

Mir ist bekannt, dass Zuwiderhandlungen gegen diese Erklärung eine Benotung der Arbeit mit der Note "nicht ausreichend" zur Folge haben. Ich weiß, dass Verletzungen des Urheberrechts sowie Betrugsversuche strafrechtlich verfolgt werden können und dass, wer vorsätzlich gegen eine die Täuschung betreffende Regelung verstößt, ordnungswidrig handelt. Die Ordnungswidrigkeit kann mit einer Geldbuße bis zu 50.000 Euro geahndet werden. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann außerdem eine Exmatrikulation erfolgen.

_____ _____

(Datum) (Unterschrift)

* Zutreffendes bitte ankreuzen