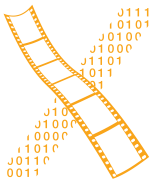


Chair for Computer Science 10 (Media Computing and Human-Computer Interaction)

RWTH AACHEN UNIVERSITY

Seeing Analog Forests and Digital Trees? — Impact of Digital Devices on the Construal Level

Master's Thesis
submitted to the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Sebastian Hueber

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof. Dr. Ulrik Schroeder

Registration date: 19. 12. 2017
Submission date: 15. 03. 2018

Eidesstattliche Versicherung

_____ _____

Name, Vorname Matrikelnummer

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Masterarbeit mit dem Titel

_____

_____

_____

selbständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

_____ _____

Ort, Datum Unterschrift

Belehrung:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

_____ _____

Ort, Datum Unterschrift

Contents

List of Figures

List of Tables

Abstract

Previous research indicated that digital devices might change the way we think. As recent trends in industry and education increasingly integrate digital devices, the possible effects of digital platforms have to be identified and understood thoroughly in order to create digital work environments from which the users benefit. While the history of comparisons between analog and digital media show that reading speed and accuracy approached each other with enhanced display quality, one cannot state the same for the abstractness of thought. It seems that the type of media itself – or its typical input interaction – evokes a higher or lower construal level, but so far no evidence to support this has been found.

In this thesis we review the history of analog and digital comparisons from the 1980s up to now. We present research done in the field of Construal Level Theory and adopt their measures to conduct our comparison between analog and digital media. We present two studies with 120 participants each in which we investigate differences in both the conceptual and perceptual construal level between analog and digital media and the effect of stylus usage.

While we could not find any significant differences between analog and digital in our studies, our results are in line with the reference results of the tests and our obtained data seems valid. We conclude that if the amount of content and the task are similar, digital devices inherently have no effect on the construal level. To the best of our knowledge, this is the first work that investigates (a) the effect of digital media on perceptual construal and (b) the effect of the interaction design of digital devices on the construal level.

Überblick

Vorangehende Forschung hat gezeigt, dass digitale Geräte einen Einfluss darauf haben könnten, wie wir denken. Da aktuelle Trends in Industrie und Lehre eine immer weitergehende Integration von digitalen Medien vorsehen, müssen die Effekte von digitalen Geräten erkannt und verstanden werden, um digitale Arbeitsbereiche zu erschaffen, die einen tatsächlichen Mehrwert für ihre Nutzer bieten. Über die letzten Jahrzehnte wurden viele Vergleiche zwischen Analog und Digital getätigt. Die Lesegeschwindigkeit und -genauigkeit auf digitalen Medien hat sich mit den Verbesserungen der Displaytechnologie stetig denen von Papier angenähert. Über die Abstraktheit der Denkens auf den jeweiligen Plattformen lässt sich jedoch keine Aussage treffen. Auf den ersten Blick lässt sich vermuten, dass die Medienart – oder die für sie typische Interaktion – ein höheres bzw. niedrigeres Construal Level hervorruft. Beweise dazu wurden jedoch bis jetzt noch nicht gesucht oder gefunden.

In dieser Arbeit untersuchen wir zunächst die Geschichte der Analog vs Digital Vergleiche von den achtziger Jahren bis ins Jetzt. Außerdem präsentieren wir Forschung im Bereich der Construal Level Theory und adaptieren ihre Messmethoden um unsere eigenen Vergleiche zwischen Analog und Digital durchzuführen. Wir präsentieren zwei Studien mit jeweils 120 Teilnehmern, in denen wir nach Unterschieden sowohl im konzeptuellen als auch im perzeptiblen Construal Level zwischen der analogen und digitalen Medien und dem Einfluss der Stylusnutzung suchen.

Auch wenn wir keine signifikanten Unterschiede zwischen Analog und Digital in unseren Studien finden konnten, entsprechen unsere Ergebnisse doch den jeweiligen Ergebnissen in der Literatur und die erhaltenen Datensätze erscheinen valide. Wir schließen daraus, dass bei ähnlicher Inhaltsmenge und Tätigkeit digital Medien keinen Effekt auf das Construal Level haben. Gemäß unseres besten Wissens ist dies die erste Arbeit, die (a) den Effekt von digitalen Medien auf das perzeptible Construal Level und (b) den Effekt des Interaktionsdesigns auf digitalen Geräten auf das Construal Level untersucht.

Acknowledgements

Thanks to Christian Cherek, my supervisor, for his valuable feedback and help acquiring participants for the first study.

Special thanks to Elisabeth Brügmann for kindly providing the test materials needed for the Picture Completion test.

I want to thank all participants for their time and willingness to help me conduct the studies. Without them this research would have not been possible.

Thanks to Prof. Dr. Jan Borchers and Prof. Dr. Ulrik Schroeder for examining this thesis.

Finally, I thank my family and friends for their support over the last months.

Conventions

Throughout this thesis we use the following conventions.

Text conventions

Definitions are set off in colored boxes.

EXCURSUS:
Excursus are detailed discussions of a particular point in a book, usually in an appendix, or digressions in a written text.

Definition:
Excursus

Source code and implementation symbols are written in typewriter-style text.

```
myClass
```

The whole thesis is written in American English. We use the plural form for the first person. Unidentified third persons are described in female form.

We use the term *analog platform* for pen and paper based tasks and workplaces. Analogously, *digital platform* refers to tasks and workplaces containing digital displays.

Chapter 1

Introduction

"The mind is [...] difficult to comprehend [...]. Most of us start by believing we already understand both human behavior and the human mind. After all, we are all human: we have all lived with ourselves all of our lives, and we like to think we understand ourselves. But the truth is, we don't."

—Don Norman [2002]

Apparently unrelated changes in our environment can have an impact on the way we behave

Did you ever notice that many people turn down the radio's volume while parking their car? Of course there is no reason how quietness should help one to see better. Often we cannot explain human behavior. So how should we explain why and how humans think?

Digital devices might affect the way we think

How human thought and decisions can be influenced by our environment has been investigated by many researchers throughout the last years. For instance, believing that a test one takes is only an exercise has a beneficial effect on the score one achieves[1]. More recently, Kaufman and Flanagan [2016] proposed that digital devices might have an impact on the way we think. In this thesis we want to investigate to which extent human construal level, i.e. the abstractness of thought, is influenced by digital devices.

[1] we will see this in Chapter 3.2

Digital devices became omnipresent

We get in contact with digital devices every day. Most of these devices have a display that conveys information to the users. Additionally, with touch screens we use displays not only to see output, but also to create input. We find displays in areas where they have not been a decade ago and many people take their smartphone with them as a daily companion. With all the increased functionality of a smartphone compared to a feature phone, people look at a display for many more hours per day than they did a few years ago. Recent trends in industry, for instance Industry 4.0 and the digitalization, put more workers in contact with computer displays and digital devices. Even in schools, modern learning approaches are taken, in which students are given tablet computers instead of books and even the blackboards are replaced with digital *smart boards*.

Digital platforms are increasingly important, e.g. in industry and education

Possible effects of digital devices on the way we think have not received much attention so far

Digital displays surround us everywhere, yet it is still unknown whether they have an impact on our performance when we use them. Over the time, display technology enhanced and nowadays reading is supposed to be comparably easy and correct on display vs on paper. The effect of digital devices on the way we think, on the other hand, has not received much attention so far.

1.1 Research Questions

Analog and digital platforms vary in presentation and interaction

One finds two main differences when comparing the traditional analog and digital platforms. Firstly, the *presentation*, i.e. paper or a display. Paper obviously has no reflections on the surface and printed text offers a very high resolution with crisp fonts. Displays, on the other hand, shine and are easier to read in low light. High-resolution displays became widely adopted just a few years ago.

Secondly, the *interaction* varies between both platforms. On paper, people use a pen to write down and mark things. On digital workplaces, computers using the desktop metaphor typically receive input with mouse and keyboard but touch screens become increasingly popular. With technological advancements, new devices were invented, e.g. tablets that can discriminate between hand and pen input, which blur

the boundaries between the conventional analog and digital interactions.

In this thesis we investigate both presentation and interaction. Hence, our research questions are as follows:

RQ1 Does the presentation of information on digital vs analog platforms have an effect on the human construal level?

RQ2 Does the input interaction have an effect on the construal level?

Does the digital platform have an inherent effect on the construal level?

In other words, this thesis investigates whether the characteristics of digital platforms have an inherent effect on the construal level. Firstly, we investigate wether people think differently when they see the same information *presented* on either a display or paper. In addition, we ask the question to which extent the pen-based interaction, that is characteristic for the analog platform, causes this effect.

1.2 Outline

Chapters 2 and 3 present the related work

In this work, we investigate the effects of digital displays on construal. In particular, we compare human performance in multiple construal tests on paper with tablet computers and the effect of input conditions, resolution and size. In Chapter 2 we present the history of the analog vs digital debate and show that the increased quality of digital displays led to more similar results between the analog and digital platforms. We then give an introduction to Construal Level Theory (CLT) in Chapter 3 and explain the connection between CLT and psychological distance. We present research that found significant differences in construal levels caused by differences in already known dimensions of psychological distance. We show why a digital platform could be another priming factor on construal level.

We present two studies with 120 participants each

In Chapter 4 we present a study with 120 participants in which we investigate the effects of tool usage in the interaction with analog and digital platforms on the conceptual construal level. We then present our main study with 120 participants in Chapter 5. Here we investigate the effect of analog and digital platforms on the perceptual construal level in a series of three experiments. While the two studies reveal no significant differences between both platforms, we emphasize the validity of our obtained data and give the implications of our results in Chapter 6.

We then conclude the thesis with a summary and point out future work in Chapter 7.

Chapter 2

History of Analog versus Digital Comparisons

The discussion how digital or analog platforms enhance working environments has been ongoing for decades, yet still there exists no final agreement to date. Different measures have been used over time to investigate the effects of analog paper and digital devices. In this chapter we take a look at the history of comparisons between analog and digital platforms and point out the shifts of results that could be found over time. We finally present the construal level as a possible differentiating characteristic of both platforms.

Results on analog and digital platforms converged with technological advancements

Noyes and Garland [2008] reviewed the history of computer- vs paper-based tasks. They observed that the results achieved on digital platforms continuously improved with technological progress. However, they state that total equivalence between the analog and digital platform seems impossible.

In the following, we take a look at the comparisons of analog and digital over the last few decades, starting with the 1980s.

Figure 2.1: The Apple Macintosh 128k was released in 1984 and had a resolution of 512×342 pixels. Image adapted from http://blog.archive.org.

2.1 The 1980s

Reading speed and proofreading accuracy were the preferred measures during the 1980s

When personal computers achieved popularity in the 1980s, small monochrome displays were common. For instance, the original Macintosh, which was released in 1984, had a 9" CRT display with a resolution of 512×342 pixels. Even with this resolution, the Macintosh's 2.5-dimensional GUI was ahead of its time as the majority of computers used during the 1980s still had full-screen textual interfaces. During this decade, rather simple measures were used when comparing analog and digital platforms, for instance reading speed or accuracy of proof reading.

Reading on a television was slower than reading in a book

Muter et al. [1982] compared reading from a book with reading on a 48cm RGB television screen. In a between-subjects design with the independent variable being the platform, 32 participants read short stories for two hours. Their evaluation showed that reading on the television was

28.5% slower than reading in a book. Additionally, they found the television to cause more eyestrain, yet this difference was not significant. Neither did they observe a significant difference in dizziness, fatigue or comprehension between the conditions.

Both the interaction and the visual presentation varied heavily between analog and digital

The authors proposed five possible reasons for the slower reading performance on the digital condition. Firstly, their participants lacked familiarity with reading from displays. Secondly, a different number of words were displayed per page: approximately 400 in a book and 120 on a video page. Thirdly, the display was only capable of displaying two thirds of the characters per line that are found on paper. Fourthly, while a book can be moved closer or farther away from the readers eyes easily while reading, the same is not true for the digital condition. Lastly, the display had a nine second delay to load a page with text.

Proofreading on paper was significantly faster than on a monochrome CRT

The aforementioned results have been further supported by Gould and Grischkowsky [1984]. Instead of a television screen, they used an IBM 3277 CRT display. These displays had a resolution of 80×42 characters, which made them to be considered as good during the time of their study. The participants' performance was measured by proof reading speed and the number of errors they made. While the participants made slightly more errors in the digital condition (CRT), Gould and Grischkowsky did not find a significant difference in errors between both conditions. However, proof-reading was about 20 - 30% faster when working on hard-copy. In terms of qualitative data, e.g. fatigue, this study could also not detect a significant effect between analog and digital.

There was no significant difference between the analog and digital platform if participants read texts on paper first

An interesting finding was done by Belmore [1985]. In her within-subject experiment, 20 participants read short text passages for one hour and had to answer multiple-choice comprehension questions. For the digital condition, she used an Apple II Plus 48K with a resolution of 40×24 characters. Overall, reading was 12% slower and participants comprehended 47% less on the computer display. However, this effect could be greatly reduced or even eliminated by presenting participants the paper condition first. As the significant difference between display and paper could

Figure 2.2: The IBM 3277 CRT was able to display 40×24 characters in green color on a dark background. Image adapted from https://vintagecomputer.ca.

only be observed if the digital condition was conducted first, Belmore concluded that reading video text may be a different mental process than reading on paper.

Light text on a dark background caused slower reading times

Cushman [1986] used microfiche and a video display terminal (VDT, see Figure 2.3) for their analog vs digital comparison. Typical during this time were displays that present light text on dark background, but the VDT setup enabled them to also present dark letters on light background. Cushman et al. were able to identify the inverse image, i.e. light text on dark, in comparison with a metal screen as a cause for slower reading. Moreover, the positive digital image, i.e. dark text on light background, was significantly more fatiguing for the eyes than the negative image or paper. The comprehension scores, on the other hand, were comparable across all conditions.

The fonts used on displays had a negative impact on proofreading accuracy

The research by Creed et al. [1987] represents another comparison between proof-reading on analog and digital platforms, but with an interesting third condition: an analog photography of the display. In a within-subjects study with 30 participants they found a significant difference in the number of errors between display and paper. Their evalu-

Figure 2.3: A microfiche reader. The flat microfiche film contains multiple images with a size of about $10\times14\text{mm}^2$ each. The device magnifies these small images and displays them on the screen. Image adapted from https://commons.wikimedia.org.

ation suggests that the major factor affecting accuracy was the display font: The performance on the photo was intermediate between the two other conditions, yet a little closer to its digital original. What is more, they found no significant difference between the time needed across the three conditions.

Single-column and two-column presentation had no effect on proofreading across conditions

In a second experiment, Creed et al. investigated whether layout had an effect on the proof-reading performance. While the results matched their previous ones, they could not find a significant difference in proof-reading performance between single-column and two-column presentation across the conditions.

In comparison with paper, reading on the digital platform was slower, more fatiguing and harder to comprehend

To conclude, reading times were slower in comparison to paper on all different display technologies used in the 1980s. Also, more fatigue and less comprehension could be observed on the digital platform. However, during this decade most study participants were not familiar with computers and display quality was relatively poor.

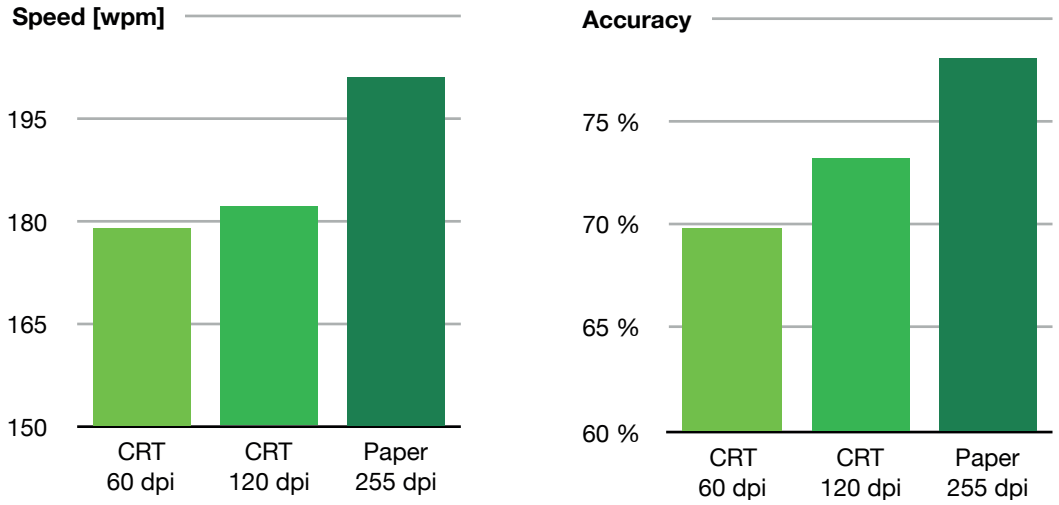


Figure 2.4: Proof-reading speed and accuracy both enhance when the resolution is increased. Adapted from Ziefle [1998]

2.2 The 1990s

CRT display technology was still predominantly used in the 1990s. However, displays were available in larger sizes and with increased resolutions.

The resolution has an effect on proofreading speed and accuracy

Ziefle [1998] compared two CRT conditions and one paper condition in a study with 20 participants. She used two monitors with 60 and 120dpi and a hard copy with a resolution of 255dpi. In terms of qualitative data, participants did not rate the text difficulty differently between conditions. As can be seen in Figure 2.4, both reading speed and accuracy increased with a higher resolution and the effect of display resolution was significant on proof-reading speed and accuracy between paper and both CRT conditions. However, there was no significant difference between the two CRT conditions. 16 of 20 participants stated their preference for reading from paper.

Reading music was not different between the analog and digital platforms

Not all studies investigated text only. Picking [1997] conducted a study in which reading music from paper was compared to reading it from a computer screen. In this study no significant differences could be found across five

different presentation forms. Neither was there a difference between analog and digital, nor did it make a difference if animations were used in the digital conditions. However, some of their participants preferred some digital conditions over the analog one. Finally, Schwartz et al. [1998] compared personality tests on analog vs digital platforms in a between-subjects study with 213 participants. Their results show significant differences in the identity status between the two conditions.

2.3 The 2000s

With further technological progress displays became larger and their resolution and brightness increased. In the 2000s significant differences between text presentation on displays and paper could no longer be found when using the same measures as mentioned before. Furthermore, for some assessment tasks, the digital versions performed significantly faster or with better scores. We will see that at least for selected tasks there is a continuing move towards similar results between digital and analog platforms.

Mayes et al. [2001] investigated conducting tests on either the analog or digital platform in a between-subjects study with 40 participants. The task was to read an article and subsequently fill out a comprehension test and a NASA-TLX. Both reading and filling out the tests were either performed on a 14″ monitor or paper. The completion time on the digital platform was significantly higher compared to the analog platform. The testing condition, on the other hand, had no significant impact on the users' performance. Additionally, they found a negative correlation between the qualitatively measured workload and the comprehension scores.

The medium had no impact on the results of a comprehension test

In order to investigate whether the observed difference between analog and digital in their previous experiment was due to an increased demand of working memory, Mayes et al. conducted a second experiment with 48 participants. While the task design was similar to the other experiment, they added a secondary task for half of the participants: the

memorization of a string of consonants. Their evaluation contains some interesting results. First and foremost, independent of the secondary task there was no significant difference between the platforms in reading speed in this experiment. Also, participants in the digital condition felt more confident regarding their performance. Only when the secondary task was introduced, participants in the digital condition remembered less information than those who worked with pen and paper.

A secondary task is more distracting on the digital platform

Noyes et al. [2004] support these findings. They conducted a similarly designed study, but with 30 participants that were "good with computers". While they did not find a significant difference in comprehension between the analog and digital platforms, they found a negative relationship between workload and comprehension. They point out that computer-based assessment tests disadvantage lower-performing individuals as they require significantly more effort. Nonetheless, the reason why participants put more effort into digital test remains unknown.

Workload and comprehension are related. Computer-based tests require more effort

Fox and Schwartz [2002] searched for differences in the answers of personality tests between the analog and digital platform. They conducted their between-subjects study with 200 males aged between 18 - 20 years. However, they did not find a significant difference between the two conditions.

The results of personality tests were similar between analog and digital

The results of Bodmann and Robinson [2004] point out advantages of working on digital devices. Their first study was a typical analog-digital comparison as within-subject study with 55 computer-literate students who were familiar with the digital testing interface that was used. While the computer interface allowed to change the entered solution multiple times, once an answer was submitted it could no longer be changed or reviewed. The digital condition only presented one question at a time. The results achieved by the participants were not significantly different between the different conditions. However, in the paper-based conditions their participants took about four minutes longer to complete the test, which is more than 10% of the allowed test time of 35 minutes.

The paper-based test took significantly more time to complete

Bodmann and Robinson conducted a second study in which they compared three different forms of digital test presentation: Firstly, an interface that closely resembles the paper interaction, in which everything is presented on one page. Hence, users can infinitely often change their answers and switch between questions. Secondly, a single question being displayed at a time, but participants are allowed to go back to a previous question. Thirdly, the digital condition from the first study, in which neither reviewing nor changing answers is allowed after submitting it. Again, there was no significant difference in the achieved scores between the three conditions. However, the last mentioned condition was 2.5 minutes faster on average compared to the other two conditions.

Not allowing users to freely switch between questions leads to a faster completion time without sacrificing the quality of answers

Mason et al. [2001] compared 30-minute knowledge tests between analog and digital platforms. In a within-subjects study with 27 participants who rated themselves confident in using unfamiliar software they could not find any significant differences in the scores between analog and digital. Likewise, their interface for the digital condition was close to the capabilities of the paper version, i.e. ability to skip questions, return to them later and change answers.

When using similar designs on both platforms, no significant difference in the scores was observable

In brief, during the 2000's the digital interaction approached the performance of the paper-based interaction. In some cases, the performance on the digital platform even exceeded the analog platform, e.g. were the participants in the study of Bodmann and Robinson [2004] faster on the digital platform. Even in the study of Mayes et al. [2001], the performance between the platforms was similar when there was no secondary task.

2.4 The 2010s

So far we have seen multiple phases of analog vs digital comparisons. In the early days reading and proof-reading were investigated and paper offered a better performance. With the increase of resolution and display quality, this effect diminished and new measures were taken. In terms of assessment and personality tests, we were able to see that

the digital platforms offer results that are similar to the analog platform. However, digital platforms require more effort to use and the exact reason could not be given so far.

In 2016, Kaufman and Flanagan [2016] took a new approach to compare analog and digital, namely whether people *think* differently if they use a digital device. If people think less abstract when they use a digital device, this might explain why filling out tests on the digital platform required more effort in the earlier mentioned studies.

Definition: *Construal Level*

CONSTRUAL LEVEL:
The construal level represents how abstract one thinks. According to Trope and Liberman [2010], high-level construals "are relatively abstract, coherent, and superordinate mental representations". Low-level construals, on the other hand, are detailed and concrete mental representations. As a consequence, people with a high construal level *see the forest rather than the trees*. In other words, the construal level of a person expands and contracts the mental horizon.
It is easy to imagine representations of actions and objects with different levels of abstraction. Abstract representations omit inconsistent or irrelevant details and are, hence, more prototypical and less ambiguous than concrete representations. Moreover, abstract representations "often convey additional information about the value of the stimulus and its relation to other stimuli" (Trope and Liberman [2010]).
For instance, one can interpret the activity "reading a paper" as "gaining knowledge". The latter abstracts away from the printed text, as the relevant aspect is the information conveyed in the text. Of course, there are multiple levels of abstractness. For actions, one usually finds the superordinate action by asking *why* the action is performed. With each iteration the superordinate action becomes more abstract.

The digital platform exposed a lower construal level

Kaufman and Flanagan [2016] presented three studies in which their users had a lower construal level when working on digital devices compared to the analog paper. They used different digital devices for different studies, includ-

Figure 2.5: The iPad 2 was released in 2011 and has a resolution of 132 ppi.

ing tablet and laptop computers. The tests chosen were the Behavior Identification Form as presented by Vallacher and Wegner [1989], a derivative of the Car Comparison Task by Fukukura et al. [2013] and a reading comprehension test. The exact reasoning for these results remains to be investigated further. The digital and analog conditions varied in multiple aspects, including size, resolution and input interaction, which means that their results might be caused by an interaction effect.

Analog and digital platforms vary in multiple aspects

Also, our brain handles information differently depending on its source. The internet — obviously accessed on digital media — is a source of information that people use every day. Thanks to the smartphone, we have quick access to any information we need from nearly everywhere. The flood of information available is overwhelming and quickly overloading the human brain, which consequently switches to less cognitively demanding mindsets. Instead of remembering the information we looked up, we remember where we found the information. Sparrow et al. [2011] call this *Google Effects*. However, these effects cannot explain the differences found by Kaufman and Flana-

The human brain prefers to remember only the location of information

gan [2016], as in total the same amount of information was presented in both conditions.

Chapter 3

Construal Level Theory

As we have seen in Chapter 2, the differences between analog and digital platforms decreased when comparing them in terms of the conventional measures. On the other hand, other research points out differences in the way we think when using digital devices. In this chapter, we present research done in the field of Construal Level Theory (CLT) as they provide us with new measures that can be used to compare analog and digital platforms on the level of abstract thought.

As already mentioned are high construal levels related to abstract representations and low construal levels related to concrete representations. CLT puts construal levels in context to the perceived distance to a represented object or action.

PSYCHOLOGICAL DISTANCE:
Trope and Liberman [2010] write that "psychological distance ist egocentric: Its reference point is the self in the here and now, and the different ways in which an object might be removed from that point — in time, in space, in social distance, and in hypotheticality — constitute different distance dimensions."
As presented by Fiedler [2007], there might be more dimensions of psychological distance. However, so far no investigation took place whether "digital vs analog platform" could be a dimension.

Definition:
Psychological Distance

Definition: *Construal Level Theory*

CONSTRUAL LEVEL THEORY:
Construal Level Theory states that the construal level and psychological distance are interrelated. The construal level of an object's representation increases together with the psychological distance of that object. A reason for this behavior is that the high-level construals less likely need to be discarded while planing a distant action.
While psychological distance and construal levels are related, they are not equal. Psychological distance, on one hand, describes *when* / *where* / *whether* / ... an event occurs. Construal levels, on the other hand, describe *what* will occur.

With an increased temporal distance our calendar becomes more vague

Let us comprehend the concept of psychological distance with a little example. Assume a friend asks you what you plan to do next Saturday. You will most likely already have *concrete* events scheduled for the day, for instance visiting your favorite restaurant at 6pm. Now assume you are asked about the same Saturday in 1 year. You will likely not have a concrete plan yet. You will give a vague — *more abstract* — answer as this day is too far away for you to already have concrete plans. Did you also notice that abstract descriptions of actions, i.e. ones with high-level construals, will bring more distant actions to your mind?

The human mind has a tendency to represent time using spatial metaphors

One more example to show how interrelated the dimensions of the psychological distance comes from Trope and Liberman [2010]: Most people will complete the sentence "a long time ago, in a ..... place" with "far away" rather than "nearby". This is due to the human mind's tendency to represent time using spatial metaphors.

Multiple methods measure the construal level as a dependent variable

In the following, we present methods to measure the construal level of a person. Additionally, we present research that measured the construal level as a dependent variable while one dimension of psychological distance was manipulated as an independent variable.

3.1 Textual Construal Tests

A well-known test to measure the construal level is the **Behavior Identification Form** (BIF) as presented by Vallacher and Wegner [1989]. The BIF was developed in the context of *Action Identification Theory*. It holds that people prefer to identify actions at higher construal levels but adopt lower-level identities for multiple reasons, e.g. the action being too difficult, unfamiliar, complex or uncomfortable.

Difficult tasks lead to lower-level construals

The test contains 25 items, each consisting out of an action and two possible identifications that differ in their abstractness.[1] The participants are asked to select exactly one identification for each action that they find to be better suitable. The BIF was designed as a pen and paper test. Vallacher and Wegner [1989] tested the BIF with a total of 1404 participants with different social and intellectual backgrounds and found similar results in their different sample groups. As it is possible to trigger a subject to switch to a lower construal level, it appears plausible to use the BIF as a dependent measure. However, it is worth noting that this test was originally not designed to be used as such.
Comparing the results of Kaufman and Flanagan [2016] with the results of Vallacher and Wegner [1989] leads to the observation that Kaufman's results in the digital condition match the results of Vallacher. However, Kaufman's results in the paper-based condition exceed the results of Vallacher.

The construal level measured by the BIF is not different across social and intellectual background

Fukukura et al. [2013] investigated the positive impact of a large psychological distance — and with it a higher construal level —- for rational decision making. In the **Car Comparison Task** participants had to compare four different fictitious car models. For each car, a summary of characteristics was presented in tabular form on one page each. The order of cars was randomized and participants were given as much time as they wanted to read an information sheet.

In the Car Comparison Task participants should identify the best alternative from a table

Fukukura et al. [2013] conducted a between-subjection study with 99 participants and three conditions. The in-

[1] See Appendix A for the complete BIF test.

The increased psychological distance significantly enhanced rational decisions

dependent variable, which defined the conditions, was the manipulation of the temporal psychological distance. In two conditions, a three minute priming activity was performed before the test. In the NEAR condition participants wrote down what they did two days ago. In the DISTANT condition, on the other hand, participants wrote down what they did today approximately 1 year ago. The last condition was the control group, which performed no priming activity. Participants in the DISTANT condition performed significantly better than the other two groups. In the DISTANT group, participants chose the best car 59% of the time. In contrast, only 34% chose the best car in the NEAR condition. The control group selected the best car only 29% of the time.

3.2 Image-based Construal Tests

GCT and SPT measure construal based on the visual structuring performance

The Kit of Factor-Referenced Cognitive Tests by Ekstrom et al. [1976] contains two tests that are used to measure the construal level as they require a strong visual structuring performance. In the **Gestalt Completion Test** (GCT), people get to see 20 fragmented images, whose components look similar to color splashes. Every image contains one object each. The task is to identify the objects presented in the images although they are not completely drawn in a time limit of four minutes. The other task is the **Snowy Pictures Task** (SPT). The goal of this task is to recognize hard-to-see objects that are concealed by strong visual grain. This tasks consists out of 24 images and has a total time limit of six minutes for all images. Examples for both tests can be seen in Figure 3.1. While both tests need similar cognitive abilities is the GCT a little harder, as it needs stronger restructuring of the stimulus set compared to the SPT.

CLT research often uses priming activities to achieve multiple conditions

Förster et al. [2004] investigated the impact of temporal psychological distance on the construal level using both the GCT and SPT. There was a total of three conditions in their studies: NEAR future, DISTANT future and control group. The manipulation of the temporal distance was achieved through a five minute task: Participants were asked to imagine their life and details about the respective day and

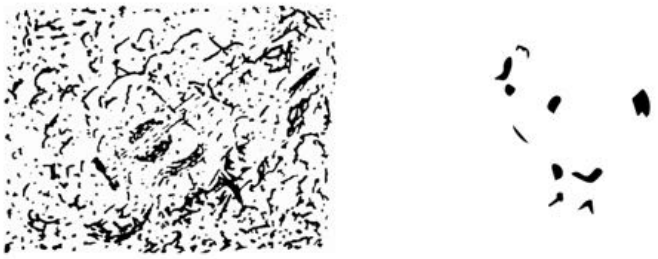


Figure 3.1: Left: A Snowy Picture depicting a guitar. Right: The representation of a chicken in the GCT. Images are adapted from Ekstrom et al. [1976].

solving the actual task either tomorrow (in the NEAR condition) or in one year (DISTANT condition). For the SPT Förster et al. [2004] used twelve images and allowed their 42 participants to fill out the test for three minutes. In the DISTANT condition, significantly more images were identified correctly compared to the NEAR condition and the control group On average, the score of the people in the DISTANT group was two images higher than in the other conditions. The GCT used by Förster et al. [2004] consisted of the first 10 items of the original GCT. Likewise, the time limit was reduced to two minutes. 45 people participated in this experiment. Similar to the results of the SPT, participants in the DISTANT condition solved more images on average. An ANOVA revealed a significant difference between the DISTANT condition and both other conditions.

SPT and GCT performance was significantly better with a larger temporal psychological distance

Other forms of manipulating the psychological distance led to significant differences in these tests, too. Smith and Trope [2006] used a computerized adaption of the GCT, in which random fragmented pictures were generated. In their study, power-priming had a significant effect on the results of the participants: Due to a higher psychological distance, the high-power condition evoked more abstract identifications.

You focus on the forest when you are in charge of the trees

Wakslak et al. [2006] used the GCT and SPT as a dependent measure of the psychological distance achieved through the manipulation of hypotheticality. In their between-subjects

Figure 3.2: A typical Navon Figure. The global letter "T" consists out of multiple local letters "S".

Hypotheticality increased the quality of the SPT and GCT answers

study 34 participants worked on both tests. Participants were told, that the SPT and GCT they received was practice material for the actual test, which would be a computerized version with 100 trials. The difference between the two conditions was which task was supposedly more likely to be conducted in the final computerized test. One group was told that with 95% likelihood the GCT was used in the computerized final test. For the other group, the the SPT was the 95% likely test. The results show, that for both tasks significantly better results were achieved when the participants believed that the task was less likely to be conducted as the final test.

3.3 Using Figures to Measure Construal

In the Navon Task subjects have to check whether two specific letters occur in a figure, either globally or locally

A different approach to measure the construal level is offered by **Navon Figures**. Navon Figures consist out of multiple equal local letters, that are arranged as one global letter as can be seen in Figure 3.2. In the most common variant of this task, people get to see a figure and have to press one of three buttons, where the first two buttons represent a letter each. If the person identifies a letter that is assigned to these buttons, this button has to be pressed. Else, the third button should be pressed. The test's score is calculated by the sum of correct button presses. Additionally, global and local letters can be analyzed separately. Navon

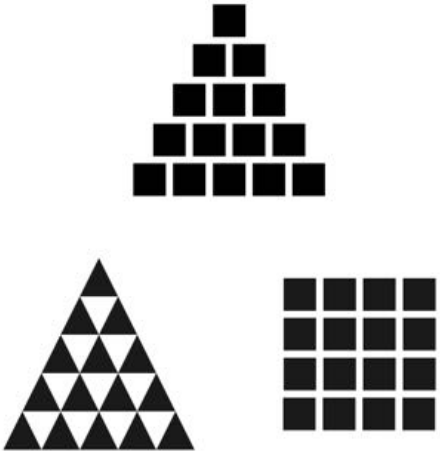


Figure 3.3: The standard figure on top is a triangle created out of squares. People with a high construal level will find the left comparison figure more similar, as they perceive the global triangular shape.

[1977] found global letters to be easier to identify than local letters. The test is language independent, as it only requires reading letters. However, the Navon Figure task is not culture independent, as found out by McKone et al. [2010].

Navon Figures are not culture independent

Liberman and Förster [2009] used Navon Figures to investigate the effect of psychological distance on the construal level with 131 participants. Their results show that an increased psychological distance, no matter in which dimension – temporal, spatial and social – increases the detection of global letters.

The detection of global letters increases with psychological distance

Kimchi and Palmer [1982] presented a concept similar to Navon Figures that is completely independent of letters. **Kimchi Palmer Figures** only consist out of squares and triangles, which either globally form a square or triangle. In this test, triples of figures as visualized in Figure 3.3 are shown to the participant. The figure on top is called the standard figure. Based on the visual impression people have from this figure, they should state whether the lower left or the lower right comparison figure looks more simi-

In the Kimchi Palmer task subjects choose between two figures where one each is more similar on the local or global level

lar to the standard figure. Each triple was presented on a white card with a size of $15 \times 13cm^2$. The test's score is the sum of correct answers in which the participant selects the comparison figure with the same global shape as the standard figure. Kimchi and Palmer [1982] found out that the number of elements, and respectively their size, has an influence on the number of answers, as multiple dense objects visually converge into a pattern.

Attention split can cause a lower construal level

Kazakova et al. [2013] investigated whether the construal level is affected by the attention split that is created when people have to pay attention to a tablet computer in their hands and a distant TV at the same time. By using Kimchi Palmer figures they were able to show that the attention split triggered a more local perceptual processing.

Images enhance immersion and cause a lower construal level

The impact of immersion on the construal level was investigated by Rim et al. [2015]. After a priming activity, which was either picture-based or textual, their participants completed a Kimchi Palmer test. People who were textually primed chose the global match, i.e. the more abstract alternative, more often. They conclude that pictures support immersion and thus reduce the psychological distance, which leads to a lower construal level.

3.4 Picture Completion Test

A low construal level is good for tasks that require attention to details

So far we only considered tests in which a high construal level leads to a better result. However, a low and more concrete construal level can also be an advantage in certain tasks, where attention to details is required. The *Picture Completion Test* is part of the Wechsler Adult Intelligence Scale, whose most recent version is WAIS-IV (Wechsler [2008]). This subtests consists out of 24 images. In each image, one important part is missing. Participants are given 20 seconds to identify the missing part. In Figure 3.4 the easiest picture in the test is depicted.

We have already mentioned the work of Wakslak et al. [2006] in which the impact of hypotheticality on construal was investigated. In addition to the SPT and GCT, they also

Figure 3.4: The first image in the Picture Completion Test in the WAIS-IV. One table leg is missing.

gave their participants the Picutre Completion Test of the Wechsler Intelligence Scale for Children (WISC). Their results show that people who where primed to have a low psychological distance, i.e. were told that this task was almost certain the one chosen in the final test, performed better. On average, participants in this condition found 2.5 more missing items in the images. This is not surprising, as a low construal level engages paying attention to details and, hence, makes it easier to find errors in images. While Wakslak et al. [2006] were able to find significant differences using this test, it is worth to note that their participants werde adults and the WISC version of this task is actually too easy for adults as the missing parts are far more obvious than in the adequate version for adults.

Performance in the Picture Completion Task increased with a lower construal level

3.5 Perception and Conception

As we have seen, different tests used to measure construal require different cognitive processes. In fact, it is possible to differentiate between construals based on the cognitive process, i.e. *perceptual* and *conceptual* construal.

Perceptual and conceptual construals relate to different cognitive processes

Definition: *Perception*

PERCEPTION:
Perception refers to the interpretation and organization of sensory information. According to Bernstein [2018] "a lot of the perceptual work that transforms sensory information into meaningful experiences happens automatically without conscious awareness." In fact, there is no knowledge required to perceive our environment. For example, babies get to know the world through perception.

Especially unknown or complex perceptions need more evaluation. For this reason *conception* refers to a higher level of information processing.

Definition: *Conception*

CONCEPTION:
The process of conception refers to us applying previously learned knowledge in order to interpret information. James [1961] describes conception as the function by which we "mark off, discriminate [...] and identify a [...] subject of discourse. Each conception is the result of our attention focusing on one specific part in a chunk of information", which obviously had to be perceived first.

Conceptual and perceptual information is processed in different parts of the brain, which can even happen simultaneously. Often the boundaries between both kinds of processing are blurred, as the human brain interprets and conceptualizes most of what it perceives and is very quick at associating related knowledge to what it perceives.

A comprehensive look on the effects of digital devices on the construal level must consider perceptual and conceptual construal levels

In order to grasp a comprehensive look on construal level and how it is influenced by digital devices, it is important to consider both types of construal. Conceptual construal can be investigated by using the Behavior Identification Form or the Car Comparison Task. Instead, the Gestalt Completion Test or the Picture Completion Test can be used to measure perceptual construal.

In summary, we started our review of related work in the previous chapter with a look at the history of analog and digital comparisons. We found out that in the early days

reading times and proof-reading accuracy were better on paper. However, with the resolution having a significant effect on the performance and advancements in display technology this effect diminished. We considered assessment and personality tests and were able to point out advantages of digital testing, as digital tests were for instance faster while delivering equally good results. We presented that digital devices might have an effect on the way we think, and that this effect could be explained by psychological distance.

We then gave an introduction into Construal Level Theory in this chapter and presented various measures for perceptual and conceptual construal. Additionally, we presented research that used these measures to find significant differences across different known dimensions of psychological distance. We showed that the construal level changes if and only if the psychological distance changes. This means that if digital platforms have an impact on the psychological distance / the construal level, the presented measures from this chapter will also lead to significant differences between analog and digital platforms.

Chapter 4

Conceptual Construal Study

As we have seen, some differences between usage of analog and digital media could so far not be explained. As pointed out by Kaufman and Flanagan [2016], a difference in construal could be the reason. If digital platforms negatively effect human performance, it would be a deeply distressing result for learning and working with digital devices. Inspired by the surprising results of the impact of psychological distance on construal, we selected multiple tests presented in the previous chapter and created both digital and analog versions of them. In this chapter, we present the first study, in which we investigate the effect of digital devices on the conceptual construal of participants. In Chapter 5 we then explore how the even more fundamental perceptual construal is effected by digital devices.

We used the BIF as a dependent measure for the conceptual construal level

In the first study, we wanted to investigate the impact of digital devices on the conceptual construal, measured with the Behavior Identification Form (BIF). Kaufman and Flanagan [2016] already used this measure in a comparison between iPad 2 and paper. Furthermore, their two conditions obviously varied in multiple aspects despite being analog or digital, e.g. resolution, size, input technique. This means they might be due to an interaction effect.

This study investigated the effects of both presentation and interaction

According to our research questions, one goal of this study was to examine whether the significant difference between the analog and digital platforms persists if we use a digital device whose display is comparably large and high-resolution to paper, i.e. the output conditions are similar. The second goal was to investigate the input conditions. Often people interact with the pen in their hand while they think. So we also wanted to investigate to which extent the pen in the hand has an impact on the answers people give.

We tested two digital and one analog conditions

As a consequence, we ended up with three conditions. In the two digital conditions, named PEN & TABLET and TOUCH & TABLET, we used a 12.9" iPad Pro as its size is comparable to a DIN A4 page and its high-resolution display allows for text that looks equally crisp to paper. The PEN & PAPER condition represents a control group in which the test was executed like in its original study by Vallacher and Wegner [1989], i.e. with pen and paper.

We chose a between-subjects study design

Choosing a specific study design often comes with a trade-off, of course. For our study we decided to use a between subjects design, in which every participant completes the test only once. If we used a within-subjects design, we would have either needed to repeat the test multiple times, which obviously does not make much sense as participants are unlikely to change their selection across multiple tests once they found a suitable identification, or we would have needed to change the test into multiple different tests, which bears the danger of altering the test in such a way, that the results are not comparable. In fact we believe that both within-subjects approaches would have irritated the participants.

4.1 Hypotheses

For our first study, we phrased the following hypotheses:

H1 There is a difference in the BIF score between presentation modalities, i.e. between PEN & PAPER and PEN & TABLET.

H2 There is a difference in the BIF score between interaction modalities, i.e. between TOUCH & TABLET and the two conditions with pen-based input.

If such a differences exists, we will be able to reject the null hypothesis during the statistical testing.

4.2 Apparatus

DIN A4 paper and the 12.9" iPad Pro have a similar size

For the analog PEN & PAPER condition, we printed out the BIF on regular white DIN A4 paper. Participants selected one of the to alternative identifications by checking the according box with a pen. For the digital conditions, we used a 2015 iPad Pro with a 12.9" display. Its resolution is 2018 by 2732, i.e. a pixel density of 264ppi. With the bezels included, the size of the iPad is a few millimeters larger than DIN A4 paper in both dimensions. For PEN & TABLET, we added an Apple Pencil for input and all other input kinds except the Pencil were deactivated for this condition.

Regarding presentation, layout and size were equal on paper and iPad Pro

Using Xcode, we implemented an iOS application to conduct the study on the digital platform. Kaufman and Flanagan [2016] recommended to use the same size and layout across all conditions. By taking advantage of the PDF capabilities of Apple's `Core Graphics` base framework, we implemented a PDF renderer that supports scaling and cropping. In order to achieve absolutely equal layout and size, we used the same PDF file for the hardcopies of the BIF in the PEN & PAPER condition and in the iOS app. For the digital conditions, the PDF was scaled and repositioned so that the page was exactly at the same position when the printed out BIF was laid on top of the iPad and aligned along the bottom edge of the iPad.

The input interaction with the Apple Pencil was identical to a regular pen

The interaction with the PEN & TABLET interaction was designed to be as similar as possible to PEN & PAPER by taking care of the following three characteristics. Firstly, participants were able to freely draw with the pencil across the screen. Secondly, we left out any undo functionality. If people made a mistake, they would had to strike out an an-

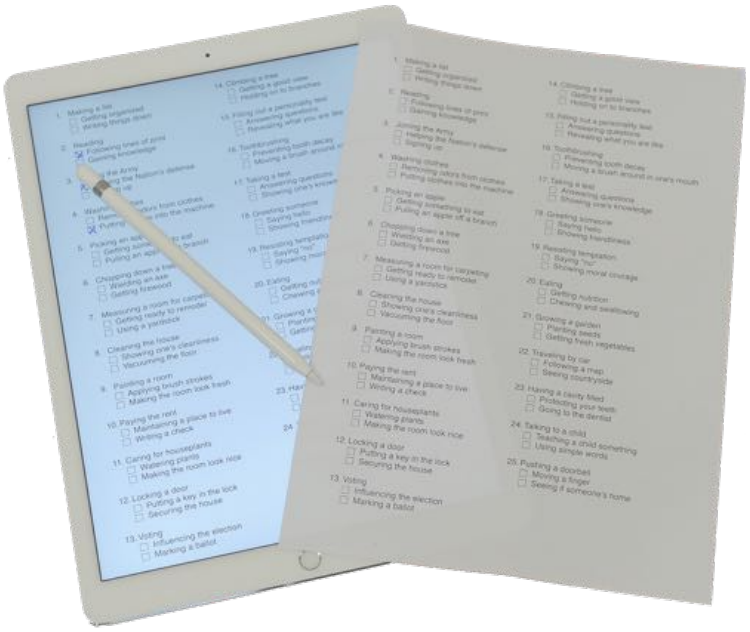


Figure 4.1: Layout and size of the BIF were identical in both analog and digital conditions.

swer and provide a new one, just like on paper. Thirdly, we implemented a palm rejection feature so that users could freely touch the display without generating input.

The palm rejection was achieved as follows: The `UIView` in which the PDF was rendered conformed to the `UIResponder` protocol and was hence able to define its own touch handling. The system then checked the type of each occurring `UITouch` event and ignored all events on the display that did not originate from the Apple Pencil.

In the TOUCH & TABLET condition a checkbox was selected by tapping

The selection of identifications in the TOUCH & TABLET condition, on the other hand, worked differently. A simple tap on the checkbox or the description next to it selected one of the identifications. Users could tap again in order to deselect it.

Both digital conditions had further details in common to achieve a higher similarity between analog and digital. Firstly, we disabled the default iOS status bar on the top

edge of the screen. Consequently, the screen was completely white except from the contents that would also be present in a printed out test. Secondly, we used an iPad with white bezels around the display, which intensified the illusion of having a homogenous white surface with black text on it. Thirdly, we disabled all hardware buttons, automatic brightness changes and dimming the screen when not used. Lastly, both the ink of the pen we provided in the PEN & PAPER condition and the virtual ink of the Apple Pencil were dark blue.

Multiple measures were taken to increase similarity of the test between all conditions

4.3 Procedure

We approached people at our lab and across the campus area and asked them whether they wanted to participate in our study. Participants were randomly assigned to one condition. The test procedure was the same across all conditions. First participants read the introduction to the test as it was originally presented by Vallacher and Wegner. Then we asked them to fill out the BIF in a maximum of ten minutes. Finally, participants filled out a short questionnaire that captured demographics and touch screen usage. If participants had vocabulary questions or were uncertain about the meaning of a specific identification, we gave them a German translation. For each identification we provided the same translation across all participants. Participants who asked more than two vocabulary questions were excluded from analysis. There was no monetary compensation for participation. However, we offered sweets.

Participants had ten minutes to fill out the BIF

4.4 Participants

We conducted this experiment with 120 participants aged between 19 and 44 (M = 24.27, SD = 3.14). Three participants were 30 or older. 47 participants were female and 72 male. One participant did not report the gender. All but one (aged 22) participant reported daily touchscreen usage. Digital pen interaction was relatively unfamiliar for most of

Our 120 participants were familiar with touch screens

our participants. They reported to rarely use a digital pen as input device (M = 0.9, SD = 1.24 on 5-point Likert scale ranging from 0 to 4).

4.5 Evaluation

Three participants in the TOUCH & TABLET condition did not answer all questions and were removed from the analysis.

	n	Mean	Median	SD
PEN & PAPER	40	14.375	15	4.139
PEN & TABLET	40	14.875	15	4.038
TOUCH & TABLET	37	14.595	15	4.408

Table 4.1: Descriptive statistics of our obtained data from the first study after removal of excluded participants.

There was no significant difference between the conditions

A Shapiro-Wilk test indicated that our results were normally distributed. In addition, Levene's test prove the homogeneity of variances. We then used a one-way ANOVA to compare the scores of all three conditions. There were no significant differences between the conditions ($F(2, 112) = .14, p = .87$).

4.6 Discussion

We could not replicate the results of Kaufman and Flanagan

The result of the ANOVA did not allow us to reject the null hypothesis. Hence, we did neither find evidence for *H1* nor *H2*. This brings us to the surprising finding that we were not able to replicate the difference between analog and digital platforms Kaufman and Flanagan [2016] found. In the following we want to compare their study with our experiment to point out possible reasons.

As already stated at the beginning of the chapter, we fathomed an interaction effect might have been the reason for their findings. On one hand, the tablet used by Kaufman

and Flanagan was a 9.7″ iPad 2, while we used a 12.9″ iPad Pro. These two devices vary in multiple aspects. First and foremost, the resolution of 264ppi of the iPad Pro is high enough to leave no discernible difference in sharpness to paper. The iPad 2's display with a 132ppi resolution, on the other hand, presents text noticeably pixelated. Second, the iPad 2 is a lot smaller than paper. Meanwhile the iPad Pro is even 3.6% larger than A4. As Kaufman and Flanagan report that they used the same font size in both analog and digital conditions, our best guess is that they might have used a two paged layout on the iPad. Third, only the iPad Pro is capable of differentiating between finger and pen. However, we did not find a significant effect between both digital conditions.

iPad 2 and iPad Pro vary in multiple aspects

Nonetheless, if the reason we obtained no significant differences between analog and digital platform was the different iPad, then our digital conditions should have performed better than the digital condition of Kaufman and Flanagan while our analog condition should have performed similar to theirs. As Figure 4.2 shows, this is however not the case. In fact, we found our results to be in line with the ones that Vallacher and Wegner [1989] obtained in a study with 1404 participants. The results from Kaufman and Flanagan for the analog condition, on the other hand, are higher than what we and previous literature were able to observe.

The construal level in the analog condition of Kaufman and Flanagan is extraordinary high

Another possible explanation for the effect could be the participants' language and culture. As we conducted the study in Germany, our participants were mostly native German speakers. The participants of Kaufman and Flanagan, however, were presumably native English speakers. Still, English proficiency at German universities usually exceeds C1 (cf. Coleman [2006]). Additionally, we excluded participants who asked more than two vocabulary questions. Culture, on the other hand, might also have an effect. The behavior "joining the army", for instance, will be accredited more importance by Americans in comparison to Germans. Still, even when leaving out this question of the BIF from our evaluation, our results remain without a significant difference.

Differences in language and culture of the participants could have an effect on the results

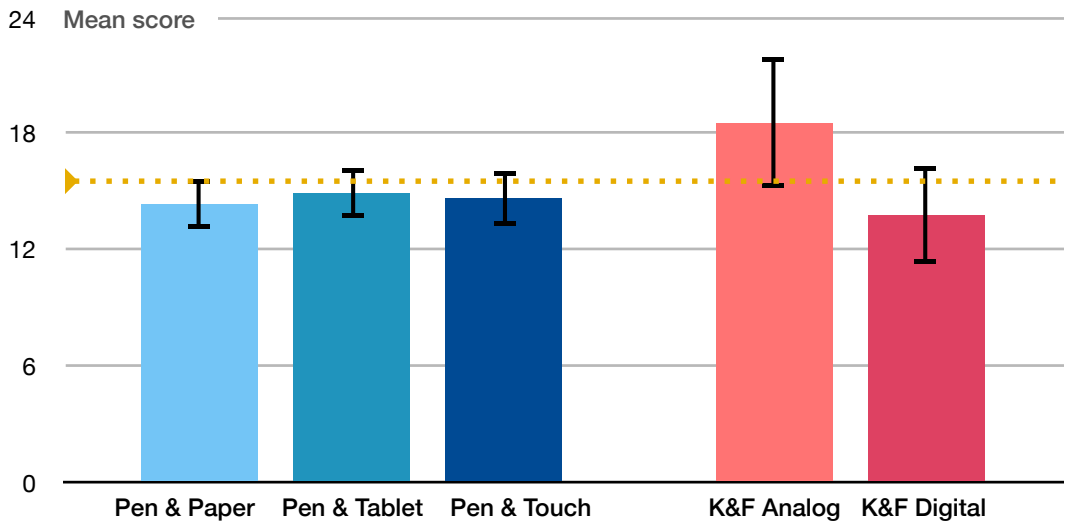


Figure 4.2: Comparison of the results we obtained in our study (blue) with the results of Kaufman and Flanagan [2016] (red). The dotted line represents the reference result from Vallacher and Wegner [1989]. Error bars indicate 95% CI.

The environment in which the study is conducted might influence the results

Finally, the differences between our results and the ones of Kaufman and Flanagan [2016] could be explained by the context and environment in which they were conducted. We conducted the experiment in the wild and Kaufman and Flanagan in the lab. Even though we avoided noisy places and made sure that our participants could fully concentrate on the task, different results in different environments are likely. As we have already shown earlier, a possible larger psychological distance can lead to a higher construal level.

The feeling of being observed should increase the BIF score

People tend to have a larger feeling of being observed in a lab environment. When taking the *Hawthorne effect* into consideration, this increased obtrusiveness will likely make participants lean towards more abstract answers, as many believe this to be the ”correct” or ”desired” behavior. For instance, multiple participants told us that they identified the difference between the two possible identifications and then kept selecting the more abstract alternative as it seemed to be the ”more intelligent” alternative for them.

Uncomfortable situations will lead to a lower BIF score

Contrarily, according to *Action Identification Theory* unfamiliar and uncomfortable situation will lead to a lower construal level. As both the participants of us and of Kaufman

and Flanagan reported daily touch screen usage, we can assume that a lower score caused by unfamiliarity with digital devices was not the reason for the differences between their and our results.

4.6.1 Effect of the Environment

The Hawthorne effect and Action Identification Theory provide contrary arguments why results of a construal test can be biased in a lab environment. Which effect dominates, however, is not clear. In order to put our results even better in the context of previous results, we analyzed the effect of the test environment by collecting an additional dataset on the analog platform. We asked 55 students who attended a lecture held in English to participate in a small study and fill out the BIF in the class. The environment of the lecture hall and the situation similar to an examination should create comparable circumstances to the related work. We applied the same procedure as in the PEN & PAPER condition.

We evaluated the BIF in a lab setting with 55 participants

The demographics of our participants for this supplemental study were similar to our previous 120 participants, namely aged between 19 and 32 (M = 22.69, SD = 2.37). 11 Participants were female. Again, all but one reported daily touch screen usage but rare usage of a digital pen as input device. Three participants were excluded from analysis as they did not answer all questions.

The new dataset obtained in the lab was normally distributed and fulfilled the requirements to compare them in a one-way ANOVA with the already existing dataset. This time our participants achieved a mean score of 13.04 in the BIF (SD = 4.21). There was, however, no significant difference between the analog tests in the lab and the already existing three conditions ($F(3, 165) = 1.747, p = .159$).

We found no significant difference between the scores obtained in the wild and in the lab

Chapter 5

Perceptual Construal Study

As we have pointed out in Chapter 3, perceptual construal happens on an even more fundamental level. The goal of this study was to identify differences in perceptual construal caused by the change of platform and interaction technique. Therefore we conducted a series of three experiments with a new group of 120 participants.

We used an iPad 2 for TOUCH & TABLET to create a larger difference to the analog condition

Remember that we did not find significant differences between the analog and digital platform in our study on the conceptual construal level. Hence, we refined one of our digital conditions to create a larger difference between analog and digital by not only using an iPad Pro but also an iPad 2. The iPad 2 already performed significantly worse than pen and paper in the conceptual construal study conducted by Kaufman and Flanagan [2016].

The other two conditions remained unchanged

Except from TOUCH & TABLET, the three conditions in this study were similar to the first study: Firstly, the PEN & PAPER with printed versions of the test material. Secondly, the PEN & TABLET condition using a 12.9″ iPad Pro and an Apple Pencil. Thirdly, the TOUCH & TABLET condition using a 9.7″ iPad 2.

In Chapter 3 we presented tests that were used to measure perceptual construal in the context of CLT. For our three ex-

Three different perceptual construal tests were selected for this study

periments, we selected one promising test from each category as a dependent measure. We made our selection based on the following three criteria. First and foremost, the tests should work language- and culture-independent. In our first study, differences between the results obtained by us and research in the US cannot be completely excluded. Consequently, we wanted our second study to minimize the error source of participants' language and culture. Second, the complete study with all tests should not take too much time to keep it interesting for the participants. Hence, the sum of all three tests should not take longer than 15 minutes. Third, we wanted to increase the validity of our results by using measures that either favor a high or a low construal level for better results. The final selection of tests was as follows:

Experiment 1 From the category of image-based construal tests, we selected the **Gestalt Completion Test** (GCT) by Ekstrom et al. [1976] (cf. Chapter 3.2). People see incomplete images in this test and need a high construal level in order to identify the objects in the images. This test was chosen over the Snowy Pictures Task as it needs less time to complete.

Experiment 2 From the category of figure-based construal tests, we rejected Navon Figures as they are known to be not culture independent. Instead, we selected **Kimchi Palmer Figures** to be used as a measure for the perceptual construal level (cf. Chapter 3.3). In this test, people compare two figures to a given figure. If they have a high construal level, they will likely select the figure with the same global arrangement.

Experiment 3 For increased validity, we wanted to add a test in which a lower construal level is helpful to achieve a better result, too. In the **Picture Completion Test** people see an image in which they should identify the missing part within 20 seconds (cf. Chapter 3.4). This task requires paying attention to details and, thus, a low construal level. Other than Wakslak et al. [2006], we used the current version of the test that is designed to be used with adults.

As in our previous study, we kept a between-subjects design in order to be able to use the tests as they were designed. Time limits and restrictions were adapted from the original task definition of each test.

5.1 Hypotheses

For our second study, we phrased the following hypotheses:

H1 There is a difference in the perceptual construal level (as measured by the three tests) between presentation modalities, i.e. between PEN & PAPER and PEN & TABLET.

H2 There is a difference in the perceptual construal level (as measured by the three tests) between interaction modalities, i.e. between TOUCH & TABLET and the two conditions with pen-based input.

5.2 Apparatus

Two different digital devices were used in addition to the analog materials

For the PEN & TABLET condition the same 2015 iPad Pro with a 12.9″ display from the first study was used. Its resolution was 2018 by 2732, i.e. a pixel density of 264ppi. The Apple Pencil was used in experiments 1 and 3, as these require written input from the participant. The iPad 2 used in the TOUCH & TABLET condition had a 9.7″ display with a resolution 1024 by 768 and a pixel density of 132 ppi. All materials used in the analog PEN & PAPER condition were printed out on regular white DIN A4 paper. Figure 5.4 shows how the Kimchi Palmer task looked like on all three conditions.

Again we took measures to assure same sizes across all conditions. Due to the smaller size of the iPad 2, however, we had to make specific adaptions for each experiment. These

changes are pointed out in the relevant section of each experiment.

5.2.1 The Test Application

The iOS application had to target different devices and operating systems

In this study we used two different devices for the digital platform. Hence, we had to implement an application compatible with the different operating systems and sizes of the iPad 2 and iPad Pro. We did this by using Xcode and the Objective-C language.

All tasks were defined in an XML-based file

There did not exist a final decision on which tests to use when we started writing the application. In addition, the materials for some tests were not available. As a consequence the application had to be flexible in terms of task design. This requirement was realized by introducing the XML-based property list `Tests.plist`, which defines the set of available tests, provides all strings, contains the file names of the used images and even contains the solutions of the test to enable automated checking. For tests with multiple elements on one page, the file even contains layout information. By introducing this textual task definition it became easy to change strings and images, add or remove questions or even completely new tasks.

A custom input widget adapted to the condition with a suitable input modality

Moreover, we knew we would conduct tests with text-based input, where the input interaction will be different. To deal with text-based input in a consistent way throughout all conditions, we introduced the `ConstrualInputField` widget. This widget provides a consistent UI and a single access point to the data entered into it, no matter whether typed in with a keyboard or written with the pencil. In fact, it also supports handwritten input with a finger, a condition that we rejected over time.

Crossing out text in the PEN & TABLET condition cleared the text field

Additionally we implemented a feature to clear the field's content by crossing out the text. When used with the Apple Pencil, the input field incorporates the same palm rejection we implemented for the first study.

Data was automatically logged into a CSV file

All data was captured and logged into a CSV file which could then be exported via AirDrop or other services right

from the application. All handwritten input was aggregated into a single image for later evaluation.

We took the same efforts to minimize distraction as in the first study. On one hand, we disabled the default iOS status bar on the top edge of the screen. On the other hand, we disabled all hardware buttons, automatic brightness changes of the screen by using the *Guided Access* feature of iOS.

Further application details can be found in the relevant chapter for each experiment. Images of the test screens can be found in Appendix B.

5.3 Procedure

We approached people at our lab and across the campus area and asked them whether they wanted to participate in our study. Participants were randomly assigned to one condition. The test procedure was the same across all conditions. However, the order in which the experiments were conducted was counterbalanced by using a Latin-square. We told people that the reason for this study was that humans might see things differently depending on their presentation. The detailed procedure for each experiment is described below in the according section. After all experiments were completed, participants filled out a short questionnaire that captured demographics and touch screen usage.

The order was counterbalanced using a Latin-square

Especially important for the digital conditions was that we conducted our study inside and sat participants in a way that possible reflections on the screen were kept at a minimum. As we wanted the test to be as language independent as possible, we allowed English and German answers and participants were allowed to switch the language of their answers inside a test if they had issues expressing their answer in a foreign language. For increased fairness between analog and digital, we tried to supervise the tasks equally on both platforms. Participants received no monetary compensation for participation. However, we offered sweets.

Participants could answer in English or German

5.4 Participants

While our 120 participants use touch screens daily, digital pen interaction was quite unfamiliar to them

We conducted this experiment with 120 participants aged between 18 and 31 (M = 22.66, SD = 2.78). 35 participants were female and 85 male. All participants reported daily touchscreen usage. Digital pen interaction was again quite unfamiliar for most of our participants. They reported to rarely use a digital pen as input device (M = 1.28, SD = 1.16 on 5-point Likert scale ranging from 0 to 4).

5.5 Experiment 1

The GCT was selected for the first experiment

In the first experiment we used the Gestalt Completion Test (GCT) as defined by Ekstrom et al. [1976]. The GCT measures the construal with 20 fragmented images that contain one object each. Participants have to identify the incompletely drawn image. A high number of correct answers is used as an indicator of a high construal level. One advantage of this test is, that in contrast to tests like the BIF, it is impossible for participants to detect a pattern between two alternative answers. Instead, every image depicts a different object and the answer has to be given freely and not selected from a set of predefined answers.

The arrangement of figures was the same as in the original test by Ekstrom et al. [1976]; we used four pages, two with six and two with four figures.

Structure and size of figures were equal to Ekstrom et al. and equal across all conditions

We made sure that the fragmented figures had the same size across all conditions. Obviously, due to the decreased screen space on the iPad 2, we had to reduce the space between the images in the TOUCH & TABLET condition in order to fit them on one screen. We slightly changed the design of our questionnaire pages in comparison to Ekstrom et al. [1976] in order to create to visually more cohesive layout that clearly puts the images and relevant text boxes in context. In our layout, both the image and the text box are wrapped in one box. Each box has a total size of $4.8 \times 5.7 cm^2$ The light gray text box has a height of $1.8cm$. Text boxes are represented by a filled rectangle rather than a line to

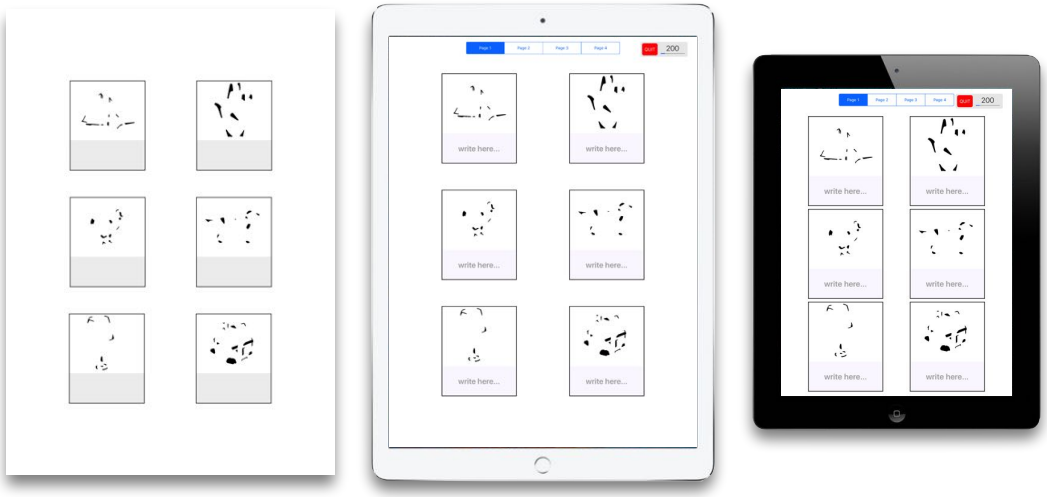


Figure 5.1: These layouts were used in each condition. Objects are sized relatively correct to each other. Paper (left) and iPad Pro (middle) share the same margins and size. On the iPad 2 margins were reduced in order to fit all 6 images on 1 page.

make the drawable area in the PEN & TABLET condition clear to the participant. On the iPad 2, text entry was realized through a software-keyboard. We disabled all word-prediction and auto-correction features in order to minimize bias of the participant towards one answer. The software keyboard was designed to move the page's contents up when it became visible and move reposition the content again when it was dismissed. In order to allow a typical interaction for the digital platform, we enabled to tab between the text boxes with either the return or tab button. In this case the page was automatically scrolled so that the next box was visible at the top of the screen

Participants could step through the textboxes on the software keyboard in the PEN & TABLET condition

In the PEN & PAPER condition the questionnaire was printed one-sided to mitigate gleaming of underneath images. We stapled the pages in the top left corner.

5.5.1 Procedure

We started with an oral introduction to the task. The introductory page contained the introduction of the task and an example image. For this page, we adapted the task de-

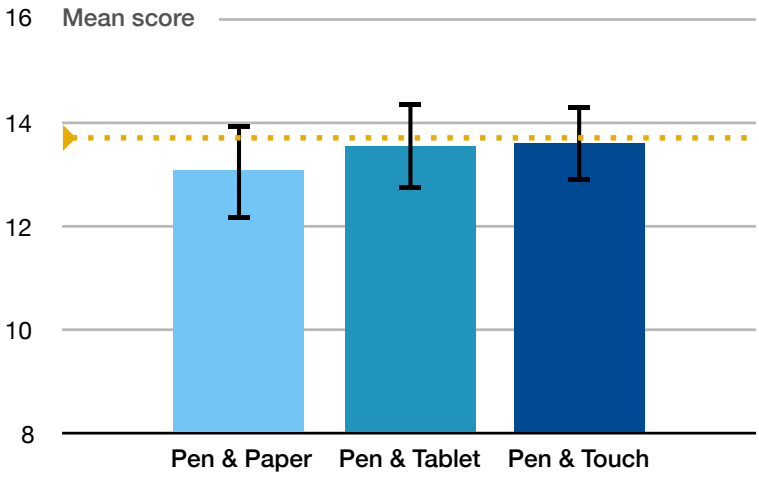


Figure 5.2: Mean of the scores achieved in the three conditions. The dotted line represents the average result of 215 college students as reported by Ekstrom et al. [1976]. Error bars indicate 95% CI.

scription from Ekstrom et al. [1976]. Furthermore, on the digital platform we added an explanation of the test's user interface and a large continue button to proceed.

Participants could freely switch between the pages during the time limit of four minutes

There was a time limit of four minutes. During this time, participants were allowed to freely switch between pages to fill in answers in all conditions. The digital platforms presented the remaining time in seconds and a progress bar in the upper right corner. For the analog platform, we informed users about their remaining time after the first two minutes and 30 seconds before the time limit.

5.5.2 Evaluation

We gave one point for a correct answer and zero points else. Hence, a total of 20 points was achievable in this test. Sometimes participants gave answers that did not match the sample solution but that were not wrong either. For instance, identifying the *chicken* as a *bird*. To assert a fair evaluation across all conditions a list of alternative correct and wrong answers was created.

Figure 5.3: These two images were the hardest for our participants. One can spot a cart on the left and a folding chair on the right. Images adapted from Ekstrom et al. [1976].

Two participants in the PEN & PAPER condition had tremendous problems with the task and were only able to identify four objects. We excluded them from the evaluation.

A Shapiro-Wilk test showed that the obtained data was normally distributed. Moreover, we checked the data for homogeneity of variances. A one-way ANOVA showed no significant differences between the conditions ($F(2, 115) = .009, p = .991$).

There was no significant difference between the conditions

	n	Mean	Median	SD
PEN & PAPER	38	13.50	14	2.203
PEN & TABLET	40	13.55	14	2.698
TOUCH & TABLET	40	13.58	14	2.417

Table 5.1: Descriptive statistics for the results of Experiment 1

An interesting observation from the analysis of our data is that participants in the PEN & PAPER were more creative in finding concrete identifications. For instance, the *bird* was identified as *dove*, *eagle*, *swallow* and *pigeon*. This observation matches the slightly lower average score we observed in this condition. The images of the *cart* and *folding chair*, that are depicted in Figure 5.3, were the hardest for our participants. Only four participants identified the chair correctly. None identified the cart.

Participants in the PEN & PAPER condition provided more concrete identifications

Figure 5.4: There were three conditions in the study on perceptual construal. From left to right: The analog paper, the iPad Pro (supporting pen input) and the smaller, low resolution iPad 2.

5.6 Experiment 2

The second experiment used Kimchi Palmer figures as dependent measures for the construal level

The second test we selected were Kimchi Palmer figures. Like Kimchi and Palmer [1982], we selected triangles and squares for our figures and created vector graphics for all of them. Kimchi and Palmer [1982] used white cards with a size of $15\times13cm^2$ to show to their participants. As we, on the other hand, also wanted to investigate digital platforms, whose displays have a fixed size, we decided to use full-sized paper instead. We did, however, make sure that our triples of figures fit inside $15\times13cm^2$ and their size was the same across all conditions. The test consists out of 16 triples of figures, four triples in each of the four size classes.

5.6.1 Procedure

The participants were told that there were no right or wrong answers

As before, we started the test with an oral introduction. We told our participants that this test was to capture their personal preference to which extent these figures look more similar. We also told them, that there is no right or wrong answer and that they should give their spontaneous answer for all images. The introductory page again summarized

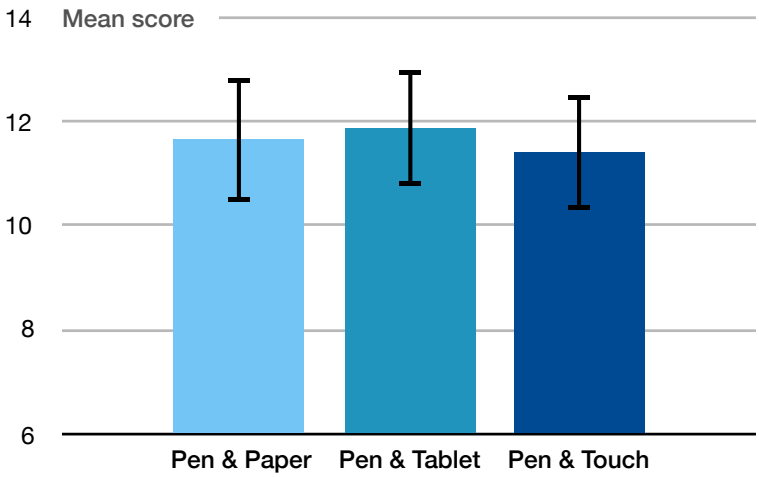


Figure 5.5: Mean of the scores achieved in the three conditions. Error bars indicate 95% CI.

the information and offered an example image with no selection indicated in order to prevent bias.

The interaction between analog and digital varies slightly in this task. In the digital conditions, participants selected a comparison figure by tapping on it. On the analog platform, we presented them one page with a figure triple each and participants either said ”left” or ”right” or pointed at their favored selection.

The order of figure triples was fixed

We maintained the same order of figures across all participants. The order of figures was selected such that a figure occurring in a triple did not occur in the next triple. There was no explicit time limit mentioned to the participants. Still, we told them that the test will likely not take longer than two minutes.

5.6.2 Evaluation

We evaluated the test by the number of answers in which participants selected the comparison figure that had the same global configuration as the standard figure. The obtained data was not normally distributed. In fact, many

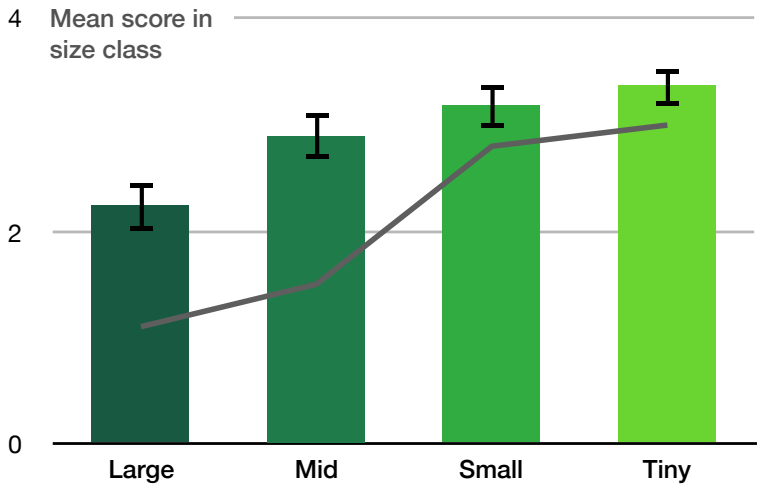


Figure 5.6: Kimchi and Palmer [1982] found a tendency to match the global shape more likely when more objects are present in the figures (gray line). Our results support this finding. See Appendix C for the definition of the size classes. Error bars indicate 95% CI.

The distribution of answers was skewed towards the maximal score

participants identified all images correctly and the distribution of scores was skewed towards the maximum. Therefore we used a Kruskal-Wallis H test for statistical analysis. The requirements for the test were met, as we observed a similar variability across the different conditions and identical skewness (towards higher scores.

There was no significant difference between the conditions

The Kruskal-Wallis H test showed no significant difference in the scores between our three conditions, $\chi^2 = 0.625, p = 0.714$. The mean rank score was 61.73 for PEN & PAPER, 56.90 for TOUCH & TABLET and 62.88 for PEN & TABLET.

	n	Mean	Median	SD
PEN & PAPER	40	11.68	12	3.812
PEN & TABLET	40	11.87	12	3.546
TOUCH & TABLET	40	11.43	12	3.551

Table 5.2: Descriptive statistics for the results of Experiment 2

5.7 Experiment 3

The final test we selected for this study was the Picture Completion Test (PCT) by Wechsler [2008]. As part of WAIS-IV, this test was originally designed and is used to measure the IQ of a person. We made high-resolution scans of the images in the stimulus book of the WAIS-IV and color-corrected them in order to achieve digital versions of the images that are as close to the original as possible.

Our digitalized test material of the PCT was as similar to the stimulus book as possible

The test consists out of 24 images in which each is missing an essential part of the image. The difficulty increases with every image. Hence, the order of images is fixed. In the original test definition, the test gets cancelled after a person did not identify the missing part in four consecutive images. For increased comparability, however, in our adaption all participants had to deal with all images. According to the test's manual, an essential evaluation criterion is, that the missing object and its location have to become clear from the participant's answer. Hence, only pointing or only providing the name of the object is not sufficient.

5.7.1 Procedure

First, an oral description of the task and an example image were used as introduction. Additionally, the introduction page on the digital platform explained the test's user interface. Conforming to the instructions in the test manual, we asked the question "what is missing in this image?" during the first few images.

Participants were explicitly asked multiple times which part is missing

The test has a time limit of 20 seconds for each picture. As we wanted to enforce a typical input interaction design of the relevant platform and condition, we slightly varied the task design as follows. The PEN & PAPER condition is conducted the same way the test was designed. We presented the participant the image and asked the question "What is missing in this image?". Participants then had to provide the object and its position verbally and/or by pointing at it. If a participant had not found the missing part within

The analog condition was conducted as defined in the test's manual

Figure 5.7: The size of the pictures varied between the conditions. However, even in the smaller image of the TOUCH & TABLET all errors were large enough to see. The depicted cow image is the tenth of the Picture Completion Test by Wechsler [2008].

In the digital conditions the missing part was visually indicated by tapping

The name of the missing part was recorded textually on the digital platform

20 seconds, we would ask for an immediate answer or else continue to the next image. In the digital conditions participants were able to select and describe the missing object and skipping to the next image on their own. In order to highlight the missing part of an image, participants were able to place a round red indicator by tapping. Placing an indicator was only possible during the first 20 seconds of seeing an image. In the PEN & TABLET condition, participants would write the objects name into a textbox underneath the image with the Apple Pencil. In the TOUCH & TABLET condition, on the other hand, they would again use an on-screen keyboard. As writing takes more time than verbally naming the missing object, the text entry field remained active even after 20 seconds.

There was no feedback on the correctness of answers

The second change we applied to the task is that we did not provide feedback if a participant identified a wrong item as the important missing part. This was done in order to create equal prerequisites on both the analog and digital platform.

The participants had to inspect all 24 images

As we did not abort the test after four failed attempts, we anticipated that participants will towards the end of the test face a series of pictures in which they will not identify the missing part. Hence, we told them beforehand that we only expect them to identify half of the images correctly in order to mitigate frustration.

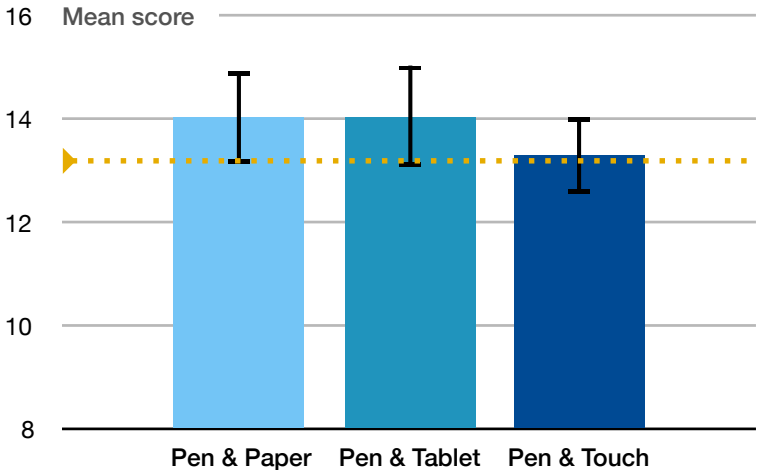


Figure 5.8: Mean of the scores achieved in the three conditions. Error bars indicate 95% CI.

5.7.2 Evaluation

Please remember that the WAIS-IV is an IQ test. Inherently, its tasks are designed to achieve normally distributed results where the average participant identifies about half of the errors correctly. Some tests of the WAIS-IV, including the Picture Completion test, are known to lead to lower results with increased age. However, all of our participants belong to the same young age group. wherefore one could expect a slightly higher performance, i.e. 13 correct answers on average. As we decided to not abort the test after four consecutive wrong answers it makes sense that the results we obtained were even a little better.

PCT performance decreases with age

	n	Mean	Median	SD
PEN & PAPER	40	13.05	14	2.926
PEN & TABLET	40	13.55	14	2.698
TOUCH & TABLET	40	13.28	13	2.417

Table 5.3: Descriptive statistics for the results of Experiment 3

There was no significant difference between the conditions

The assumptions of a one-way ANOVA were fulfilled. Therefore, it was chosen for statistical analysis. A one-way ANOVA showed no significant differences between the conditions ($F(2, 117) = .997, p = .372$).

5.8 Discussion

None of the three established construal tests showed significant differences

In summary, we conducted three different experiments in this study that all used different measures for perceptual construal. None of the statistical tests showed significant differences between the conditions and thereby did not allow us to reject the null hypothesis. Hence, we did neither find evidence for *H1* nor *H2*. As presented in Chapter 3, however, these tests have been used to point out significant differences in construals in the context of CLT. In the following we provide possible reasons for our results.

Sometimes our tests were harder than the ones used in the related work

One first reason for different results between us and the related work could be a different difficulty. For instance, in Experiment 1 we gave our participants the complete GCT test. Förster et al. [2004] only used the first half of the test, which contains more simple image than the second half. In Experiment 3 we used a completely different set of images than Wakslak et al. [2006]. They used a Picture Completion Test designed for children with their adult participants. We, on the other hand, used the appropriate most recent version of the test designed for adults. However, if the reason why we did not find any significant differences between the analog and digital platform was a too demanding difficulty, this would mean that significant differences likely only exist for simple tasks that do not require deep thinking. In this case the relevance of the findings would decrease, as workplaces and education are areas that especially incorporate complex thinking.

If the construal level on digital media was only lower at simple tasks, this difference becomes less important

Results obtained from the PCT with different image sizes will likely be different

The evaluation of Experiment 3 bears another complication, as we had to use images of different sizes between the iPad 2 and the other two conditions. With a smaller picture size, the missing details obviously become harder to see and thus can lead to a lower score. On the other hand, a construal level that is too high will also lead to a lower

score. Consequently there might have been an effect that was cancelled by an interaction effect.

An effect of the environment on our results is unlikely

We conducted this study just like our first study in the wild. Again, it is possible that one obtains different results when conducting the study in the lab even though we made sure to only use quiet places in which one could read from the digital screens without reflections. Also, we already showed in our first study that a significant difference in results between lab and in the wild is not likely. In theory the lab should create higher construal level scores as participants feel uncomfortable and observed. As our results already match the ones found in the literature, it is unlikely that we would have obtained even higher results by conducting the study in a lab setting.

Effects of language and culture can be excluded as no test was text based

The other main alteration we could identify in our first study was the difference of language and culture. This time we mitigated this problem with the selection of tests. While the instructions were still written in English, the vocabulary needed should have not been a problem for a university student. The translation of information is an additional cognitional step, of course. However, the tasks did not contain any words and we allowed our participants to answer in German. Consequently, the way they thought will have likely not been altered by the English language.

Our conditions were designed to be very similar. Finding no difference is plausible

We designed our analog and digital conditions with utmost care to make them as similar as possible. As a consequence, it is plausible, that we did not find a significant difference, as the amount of information was the same across all conditions. Moreover, finding no significant difference between the different input techniques is also plausible, as the decision which answer one wants to give is already made before the time of input.

Apparently the digital platform does not inherently trigger a lower construal level

Finally, the reason why we did not find any significances between the analog and digital platform could be that our participants simply did not have a different construal level when working on either the analog or digital platform. This could be because the digital platform does not trigger a lower-level mindset and the tasks we compared were nearly identical on both platforms.

Chapter 6

Implications

We conducted multiple construal tests and could not find any significant differences between analog and digital platforms, neither on the conceptual nor the perceptual level. In this chapter we want to provide further background to the validity of our data and its implications for interaction designers.

6.1 Interpretation of Null-Results

While one cannot prove a null hypothesis, one can still consider the validity of the data

Of course it is impossible to prove a null hypothesis by finding that no difference exists. Just because two things are equal in one aspect, they can still be different in another. Nonetheless one can assess the results of an experiment by many characteristics. First and foremost, one can consider the validity and sensitivity of the used measures. Second, one can rate the statistical quality of obtained data in terms of distribution and variance. Third, one should compare the obtained results with the related work. In the following we present three reasons why we believe in the validity of our findings.

Sensitivity and Validity of Measures.
We introduced a variety of measures for the construal level in Chapter 3 and presented related work that used these

All tests used in this thesis have been used to expose significant differences between conditions

tests. Please recall that all three tests we used in the second study are established in CLT research and have been used multiple times to expose significant differences in the construal level of participants. In addition, the BIF was not only used in the context of CLT but also by Kaufman and Flanagan [2016] who were able to spot differences between analog and digital without any further manipulation of the psychological distance.

The digital platform itself is apparently no primer for a lower psychological distance

CLT research uses various priming activities for their conditions in order to manipulate the psychological distance of their participants. As psychological distance and the construal level are interrelated, CLT research is that way able to find significant differences between their conditions. Our research, on the other hand, evaluated whether the digital platform itself has an effect on the construal level. Hence, if the digital platform has an effect on the psychological distance and the construal level, there should not exist many reasons why we did not observe a significant difference between our conditions.

Statistical Quality.

The distribution of scores was normally distributed in three of four tests

For the BIF, the GCT and the PCT the obtained data was always normally distributed. Despite the PCT being designed to deliver normally distributed results when tested with a population that is large enough, normal distributions are typically assumed when measuring human performance. Furthermore, our data had relatively low standard deviations across all tests. In fact, when looking back to the results of the BIF (Figure 4.2), we found smaller SDs and CIs than Kaufman and Flanagan [2016].

The distribution of scores in the Kimchi Palmer task was similar in all conditions

The Kimchi Palmer task turned out to be the test in which our participants got the most answers correct. As a consequence, many participants provided the correct answer to all 16 figures and our overall obtained data for this test was skewed towards the higher scores. However, it is worth noting that the variability and skewness we observed from the histograms was relatively identical across all conditions.

Comparison with Related Work.

All used tests were designed to be conducted on paper

All tests we used were originally designed to be conducted on paper. If needed, we made small changes to the test only and our digital versions were as close to the analog counterparts as possible. Hence, at least the results we obtained for the analog platform should be comparable to the reference results of the tests. In fact, as we could not find any significant differences between the conditions in any test, are our results in line with the reference results provided in the literature for all conditions. We have provided comparisons with reference results for both study 1 and all three experiments of study 2 in the relevant discussion section.

Our results match the reference results not only on the analog, but also on the digital platform

Overall, our results can be seen as a strong indication that analog and digital platforms inherently have no effect on the construal level of their user or if an effect exists, it is so small that it could not be captured with a sample size of 120 participants.

On the digital platform our results are *not* in contradiction to the ones of Kaufman and Flanagan

Please note that the motivation for this work arose from the results of Kaufman and Flanagan [2016]. However, we have seen in Figure 4.2 that their digital condition performed equal to our results or the ones found in the literature. Rather one could say that they obtained extraordinary good results on the analog platform, which provides further evidence why our obtained results are correct.

In reality there exist multiple differences between the same task on analog and digital

Of course one has to admit that in real-life use cases, work performed on digital and analog platforms differs in many ways. For instance the task of reading and annotating a book, is completely different on a tablet in comparison to a printed book. While this might be a reason why others found differences between analog and digital platforms for certain tasks, we leave the investigation of this possibility for future work. In contrast, our studies were deliberately designed to be as similar as possible on all conditions in order to eliminate extraneous variables other than the digital or analog platform and their typical input technique.

Presentation and interaction were similar in all conditions

Lastly, it is fair to say that it was foreseeable that we did not find an effect between analog and digital platforms. In Chapter 1.1 we mentioned two main differences between analog and digital: the *presentation* and the *interaction*. Re-

garding the presentation one has to say that the information content was always the same across all platforms and conditions. Finding no effect regarding the interaction makes sense, too, as the decision which answer one wants to give is made before one creates input.

6.2 Conclusions

First of all one can probably say that we found a relieving result for digital media and the whole industry. Neither digital media nor new input techniques had an inherent negative effect on the construal level of their user in our studies. While more work needs to be done to further understand the impact of digital media on the construal level, these conclusions summarize what one can learn from this thesis in an accessible format.

Impact of digital media

CONCLUSION 1:
Designers should not be afraid to integrate the digital platform into work or learning environments. The results of our study indicate that there apparently is no inherent effect of digital media that makes people think less abstract.
However, it remains important to consider the content size, the amount of information presented at once and the structure of the task.

Layout, content size and task structure remain important varying aspects between analog and digital media

The first conclusion does not mean that designers have a license for arbitrarily enforcing digital media into education and work. In our two studies layout, content size and amount of information per page were identical in all conditions. As we have seen in Chapter 2, the amount of information presented at once has an impact on our understanding of contents. Especially from the related work of the 1980s and 1990s it becomes clear that text that is hard to read has a negative impact on text comprehension and reading speed. Moreover, information overload will evoke Google Effects. While future work will analyze content size and task structure, the related work already provides a clear prediction of their effects on the construal level.

Figure 6.1: Do we see analog forests and digital trees? According to our studies this is not the case. If you read this thesis in a printed form and did not spot the missing tree trunk within a few seconds you will hopefully agree. Image taken from Wechsler [2008].

Designers of digital productivity and learning applications who want to minimize possible effects of digital media on the construal can use the well-known appearance of content on the analog platform as a cue for their designs.

Even without the existence of an inherent effect of digital media, interfaces could influence the construal level

Please remember that our results only indicate that digital media have no inherent effect on the construal level. This, however, *does not* imply that construal effects on digital platforms do not exist at all. For instance, Sparrow et al. [2011] and Kaufman and Flanagan [2016] conducted studies in which people work differently on digital media. When comparing their results to ours, one has to assume that their findings are likely caused by the used interfaces. Also, you cannot call a high or a low construal level better than the other. As we presented before, either a low or a high construal level can be advantageous for certain tasks. Hence, the impact of interfaces on the construal level should not be disregarded.

Indeed, we want to emphasize how important construal levels are for the field of Human-Computer Interaction (HCI). In our domain we try to create good interfaces so

that we empower people with digital tools. So far designers did not consider the possibility that these tools could change the way we conceptualize and think. The fact that this work is, to the best of our knowledge, the first that investigated the impact of digital devices on the perceptual construal level shows that construal levels did not receive enough attention from the research community so far.

Relevance for HCI

CONCLUSION 2:
Construal levels should be considered when evaluating the goodness of an interface. Hence, they need more attention in the field of Human-Computer Interaction.

Digital interfaces that behave similar to the analog platform will likely avoid effects on the construal level

It goes without saying that real world tasks are clearly distinct from the tasks in our experiments. Still, some tasks need paying attention to details and others focusing on the global picture. Depending on what the user should achieve while working on a digital platform, the used interface can support the user by either triggering a higher or lower construal level. When evaluating interfaces with users one should also analyze the results in terms of abstractness of thought. In case the users expose an undesired different construal level compared to the analog platform one should try a different interface design. According to our results it might be a good approach to come up with an alternative design that is more similar to the analog platform.

Chapter 7

Summary and Future Work

In this chapter we want to conclude the thesis. We first give a summary and point out the contributions of our findings to the analog vs digital discussion. Subsequently we point out the limitations of our results and how future work can extend on our findings.

7.1 Summary and Contributions

We investigated the effects of digital devices on the construal level

This thesis investigated whether digital platforms have an inherent effect on the human construal level, i.e. how abstract or concrete a person thinks. First and foremost, we examined the presentation of information on digital vs analog platforms. As alternative explanation for possible effects between analog and digital media, we also considered the effect of the interaction technique. In order to obtain a comprehensive look at the construal levels in information processing we took a look at both conceptual and perceptual construal levels.

We approached this topic with a look at the history of analog and digital comparisons structured by the last four decades. During the 1980s displays were small and had

The history of analog and digital comparison shows that reading and comprehension performance on digital devices enhanced with better display technology

a low resolution. Together with the fairly unexperienced users it led to no surprise that reading speed, text comprehension and proofreading performance was significantly worse on the digital platform. The availability of displays with higher resolutions in the 1990s enabled Ziefle [1998] to identify the resolution to have an important effect on proofreading speed and accuracy. However, the resolution of the printed paper was still lower than low-quality displays have nowadays. Starting with the 2000s the effects between analog and digital diminished and in some cases people even performed better when using digital media.

Kaufman and Flanagan proposed that digital devices have an effect on the cognition

The construal level is a measure for the abstractness of thought and interrelated to psychological distance

One important motivation of this work were the findings of Kaufman and Flanagan [2016], according to which people have a lower conceptual construal level when using digital media. As the conventional measures we presented in the history section are not considering cognitive effects, we searched for new measures in the field of Construal Level Theory (CLT). According to CLT the construal level is interrelated to the psychological distance. Hence, typical CLT research uses a priming activity in order to manipulate the psychological distance of a person and then measures the construal level by using any of the presented tests. Based on our research questions all conditions in our studies were designed to be as similar as possible. Thus we were able to reduce the number of variables to a minimum and increase the internal validity of our findings.

Study 1 analyzed the conceptual construal level and could not find a significant difference between the analog and digital platform

The first study presented the effects of digital media on the conceptual construal level by using the Behavior Identification Form (BIF) as a dependent measure. Task design, content size and information content were identical across all conditions independent of the media type. There were no significant differences in the mean scores between the conditions. We discussed possible reasons why our results did not match the ones observed by Kaufman and Flanagan [2016]. As we made sure with an additional dataset that the environment can be excluded, the only reason left was the native language of our users.

Therefore, the second study was focused on the perceptual construal level and included three image-based construal tests that are all established in CLT research. Even though

we increased the differences between the conditions by including the iPad 2 as a low-quality digital platform, we were again not able to find significant differences between the mean test scores. We discussed our results but could not identify any confounding variables that influenced them.

Study 2 showed no differences in the perceptual construal level, too

Chapter 6 wrapped up the findings of both studies and emphasizes their validity. As all four established construal tests did not expose any significant differences between the analog and digital platform, we conclude that apparently the digital platform has no inherent effect on the construal level.

All used tests are established in CLT

7.2 Limitations and Future Work

For the studies presented in this thesis the differences between analog and digital were kept at a minimum in order to find out whether digital media themselves have an inherent effect on the way people think. While it is of course not possible to prove a null hypothesis, the quality of data and similarity to the reference results of the tests provide a strong indication for the validity of our findings. Nonetheless, further work needs to be done in order to identify causes why people could change to a less abstract mindset while using digital media.

As the majority of our participants were students, the similar age, culture and social backgrounds increased the internal validity and, hence, enabled us to identify that digital media are not an issue with computer literate users. On the other hand, this selection also led to the limitation that we did not consider all age groups. For instance, users that are less familiar with technology might perform differently between analog and digital media.

One might obtain different results with people that are not computer literate

While Vallacher and Wegner [1989] were not able to find differences between social groups, nonetheless some form of anxiety when using digital devices that often occurs with older people could lead to different psychological distances and hence a change in the construal level.

Repeating the study with a different age group will enhance the validity

People could have a smaller psychological distance to an own digital device

Likewise, usage of an own digital device could have an impact on the construal level, as the personal experiences each user has with his or her tablet could evoke a smaller psychological distance and, thus, a lower construal level on digital media.

Further construal tests will likely confirm our findings

There exist more tests one can use to identify differences or similarities between both conceptual and perceptual construal level between the two media types. Our results indicate that these tests will again expose no significant differences if the task design and information content are similar across all conditions. However, this presumption also needs further investigation.

Real world tasks are different on both platforms and offer an interesting direction for future work

Ultimately, when we think outside the box, construal tests are probably not enough. Real world tasks are more complicated than tests and often include multiple processes. For example, solving a geometric problem, creating a technical drawing or reading an article with many unknown words will be executed completely differently on both platforms due to the many differences like visualization capabilities and richness of input. Consequently one needs to compare real-world tasks and how they are performed on analog or digital differently.

Changes of the construal level on digital platforms are probably due to information overload and size

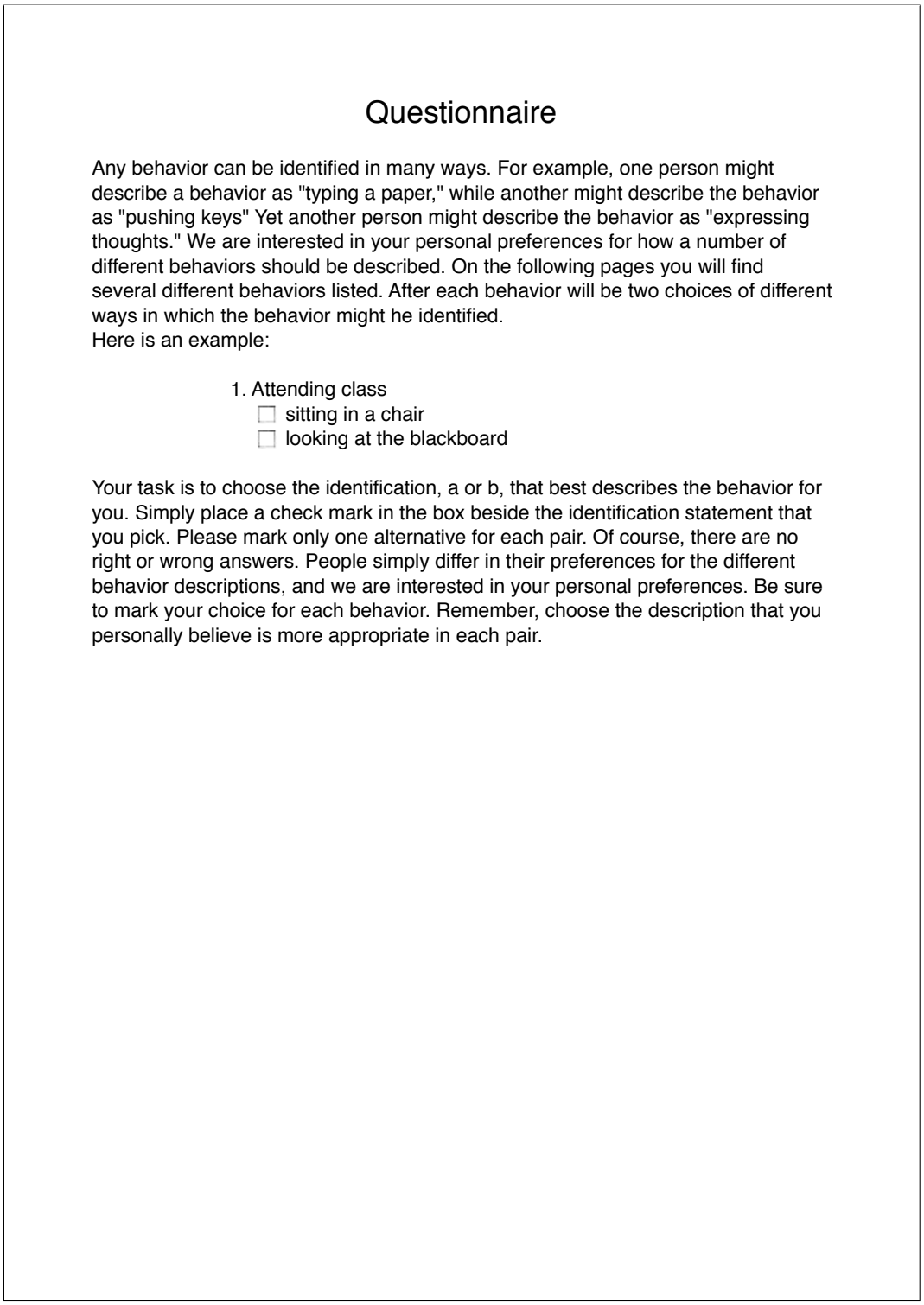
We believe that this thesis is further evidence that the differences between analog and digital platforms are probably due to information overload and different amounts of content in combination with different sizes. All theses aspects are already known to have an impact on comprehension and, hence, an effect on the construal level could also be possible.

Ultimately, the impact of digital devices on the construal level remains an important topic that requires more attention in the field of Human-Computer Interaction. This thesis provides a relieving finding as we could not observe a change in the abstractness of human mindsets across analog and digital media.

Appendix A

Material used in the Conceptual Construal Study

The following questionnaire containing the Behavior Identification Form was used in all three conditions of the first study.



Questionnaire

Any behavior can be identified in many ways. For example, one person might describe a behavior as "typing a paper," while another might describe the behavior as "pushing keys" Yet another person might describe the behavior as "expressing thoughts." We are interested in your personal preferences for how a number of different behaviors should be described. On the following pages you will find several different behaviors listed. After each behavior will be two choices of different ways in which the behavior might he identified.
Here is an example:

1. Attending class
   - ☐ sitting in a chair
   - ☐ looking at the blackboard

Your task is to choose the identification, a or b, that best describes the behavior for you. Simply place a check mark in the box beside the identification statement that you pick. Please mark only one alternative for each pair. Of course, there are no right or wrong answers. People simply differ in their preferences for the different behavior descriptions, and we are interested in your personal preferences. Be sure to mark your choice for each behavior. Remember, choose the description that you personally believe is more appropriate in each pair.

Figure A.1: Behavior Identification Form — title page

1. Making a list
   - ☐ Getting organized
   - ☐ Writing things down
2. Reading
   - ☐ Following lines of print
   - ☐ Gaining knowledge
3. Joining the Army
   - ☐ Helping the Nation's defense
   - ☐ Signing up
4. Washing clothes
   - ☐ Removing odors from clothes
   - ☐ Putting clothes into the machine
5. Picking an apple
   - ☐ Getting something to eat
   - ☐ Pulling an apple off a branch
6. Chopping down a tree
   - ☐ Wielding an axe
   - ☐ Getting firewood
7. Measuring a room for carpeting
   - ☐ Getting ready to remodel
   - ☐ Using a yardstick
8. Cleaning the house
   - ☐ Showing one's cleanliness
   - ☐ Vacuuming the floor
9. Painting a room
   - ☐ Applying brush strokes
   - ☐ Making the room look fresh
10. Paying the rent
    - ☐ Maintaining a place to live
    - ☐ Writing a check
11. Caring for houseplants
    - ☐ Watering plants
    - ☐ Making the room look nice
12. Locking a door
    - ☐ Putting a key in the lock
    - ☐ Securing the house
13. Voting
    - ☐ Influencing the election
    - ☐ Marking a ballot
14. Climbing a tree
    - ☐ Getting a good view
    - ☐ Holding on to branches
15. Filling out a personality test
    - ☐ Answering questions
    - ☐ Revealing what you are like
16. Toothbrushing
    - ☐ Preventing tooth decay
    - ☐ Moving a brush around in one's mouth
17. Taking a test
    - ☐ Answering questions
    - ☐ Showing one's knowledge
18. Greeting someone
    - ☐ Saying hello
    - ☐ Showing friendliness
19. Resisting temptation
    - ☐ Saying "no"
    - ☐ Showing moral courage
20. Eating
    - ☐ Getting nutrition
    - ☐ Chewing and swallowing
21. Growing a garden
    - ☐ Planting seeds
    - ☐ Getting fresh vegetables
22. Traveling by car
    - ☐ Following a map
    - ☐ Seeing countryside
23. Having a cavity filled
    - ☐ Protecting your teeth
    - ☐ Going to the dentist
24. Talking to a child
    - ☐ Teaching a child something
    - ☐ Using simple words
25. Pushing a doorbell
    - ☐ Moving a finger
    - ☐ Seeing if someone's home

Figure A.2: Behavior Identification Form — identifications

Appendix B

Material used in the Conceptual Perceptual Study

This appendix contains exemplary pages for each test. Please note that the sizes of the material in this thesis are scaled to the size of the pages.

For each test the conditions are presented in the following order: PEN & PAPER, PEN & TABLET, TOUCH & TABLET.

Gestalt Completion Test

4 min

This is a test of your ability to see a whole picture even though it is not completely drawn. You are to use your imagination to fill in the missing parts. Look at each incomplete picture and try to see what it is. On the line under each picture, write a word or two to describe it.

Take a look at the sample picture. It depicts a flag.

Work as rapidly as you can without sacrificing accuracy. You will have four minutes to finish all four pages of this task.

Figure B.1: Gestalt Completion Test — PEN & PAPER condition — Introduction

Figure B.2: Gestalt Completion Test — PEN & PAPER condition — Page 1

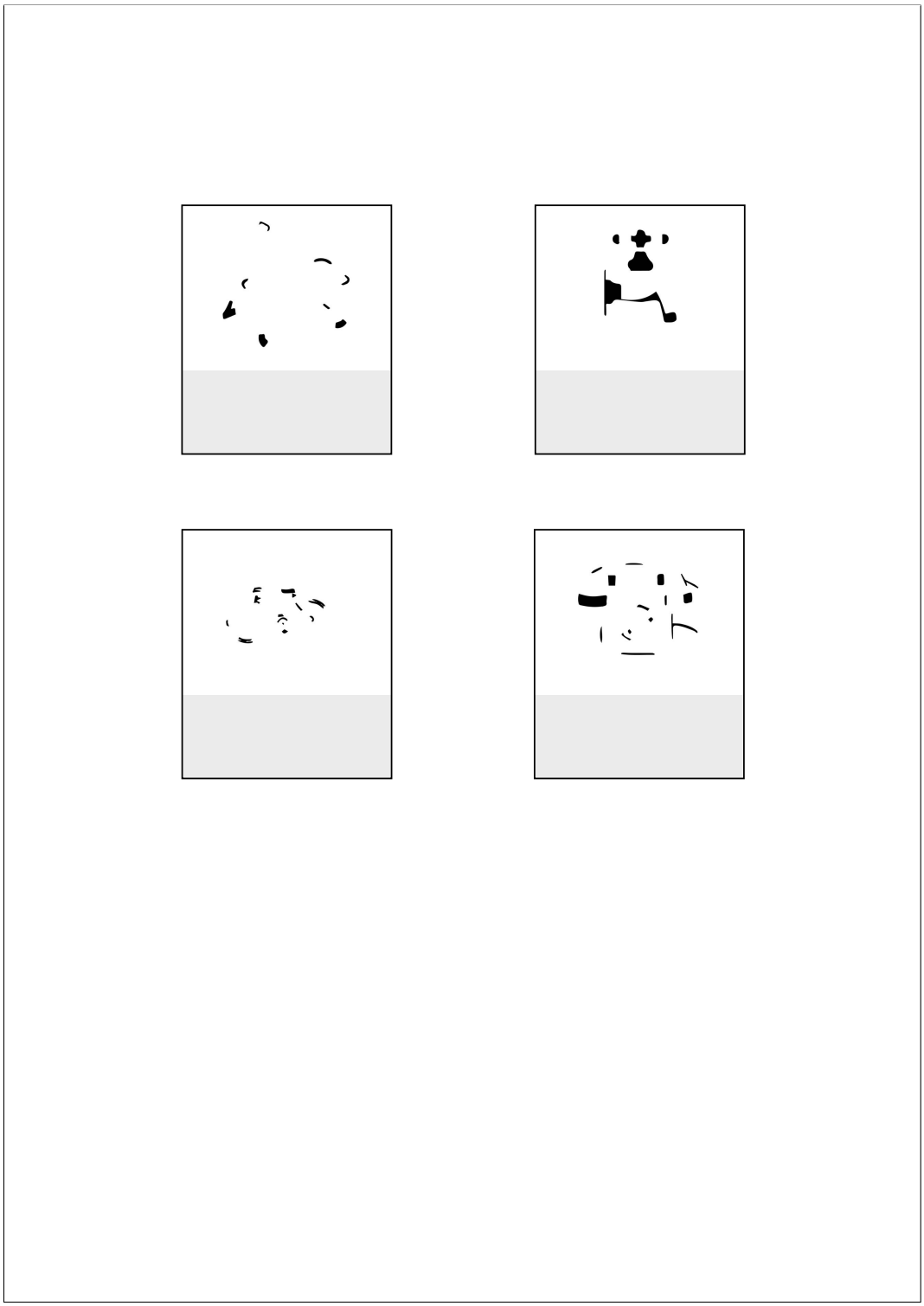


Figure B.3: Gestalt Completion Test — PEN & PAPER condition — Page 2

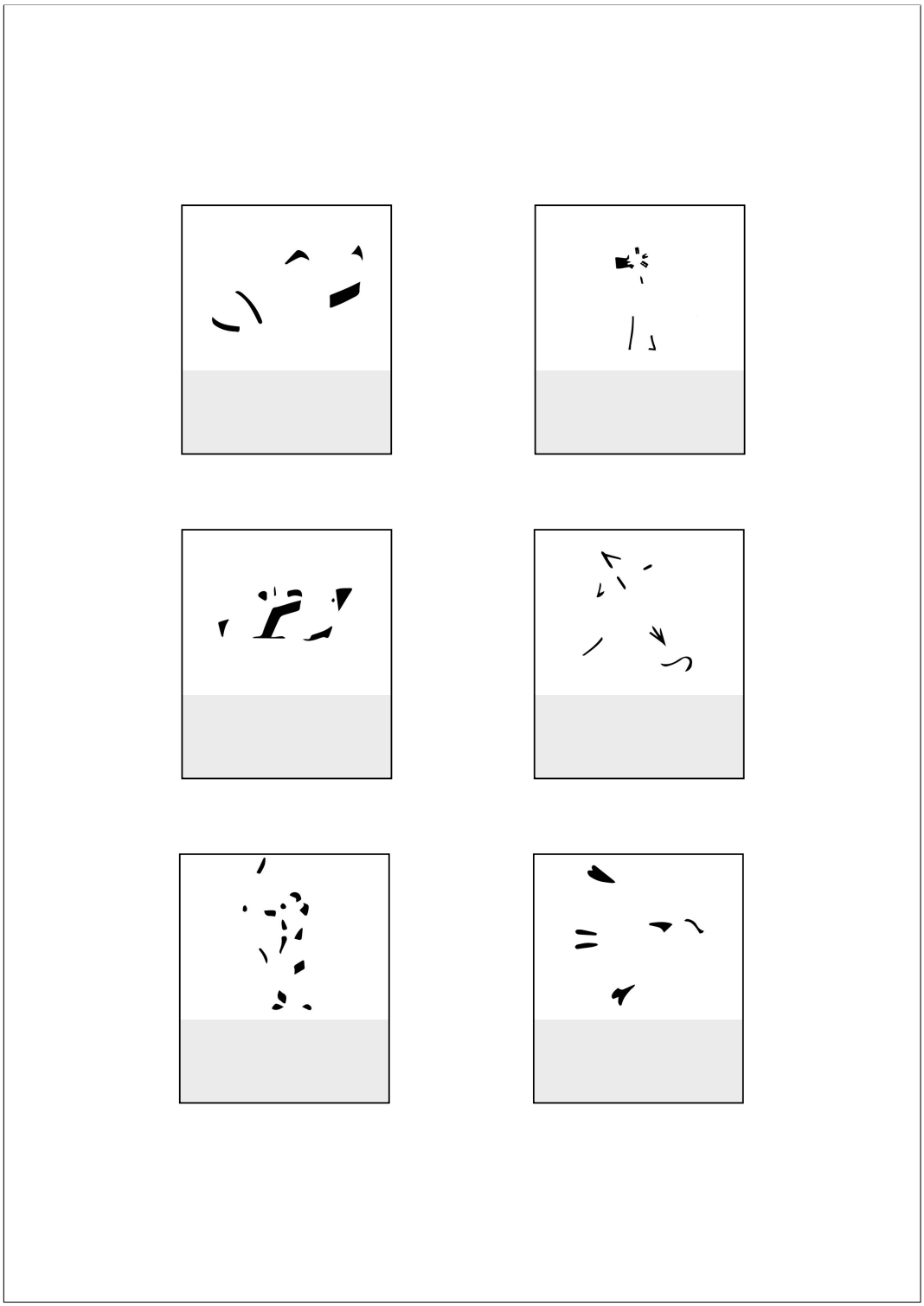


Figure B.4: Gestalt Completion Test — PEN & PAPER condition — Page 3

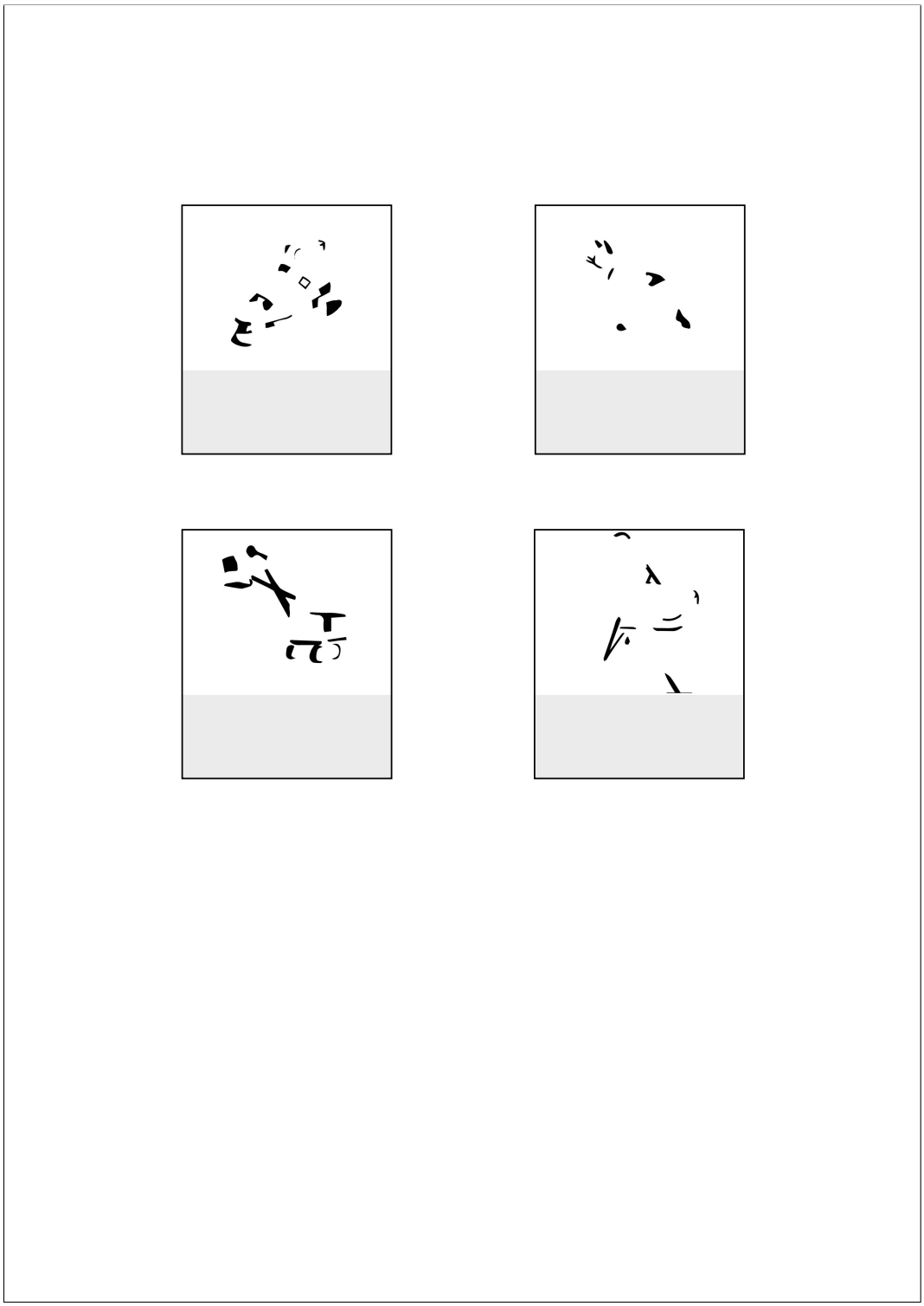


Figure B.5: Gestalt Completion Test — PEN & PAPER condition — Page 4

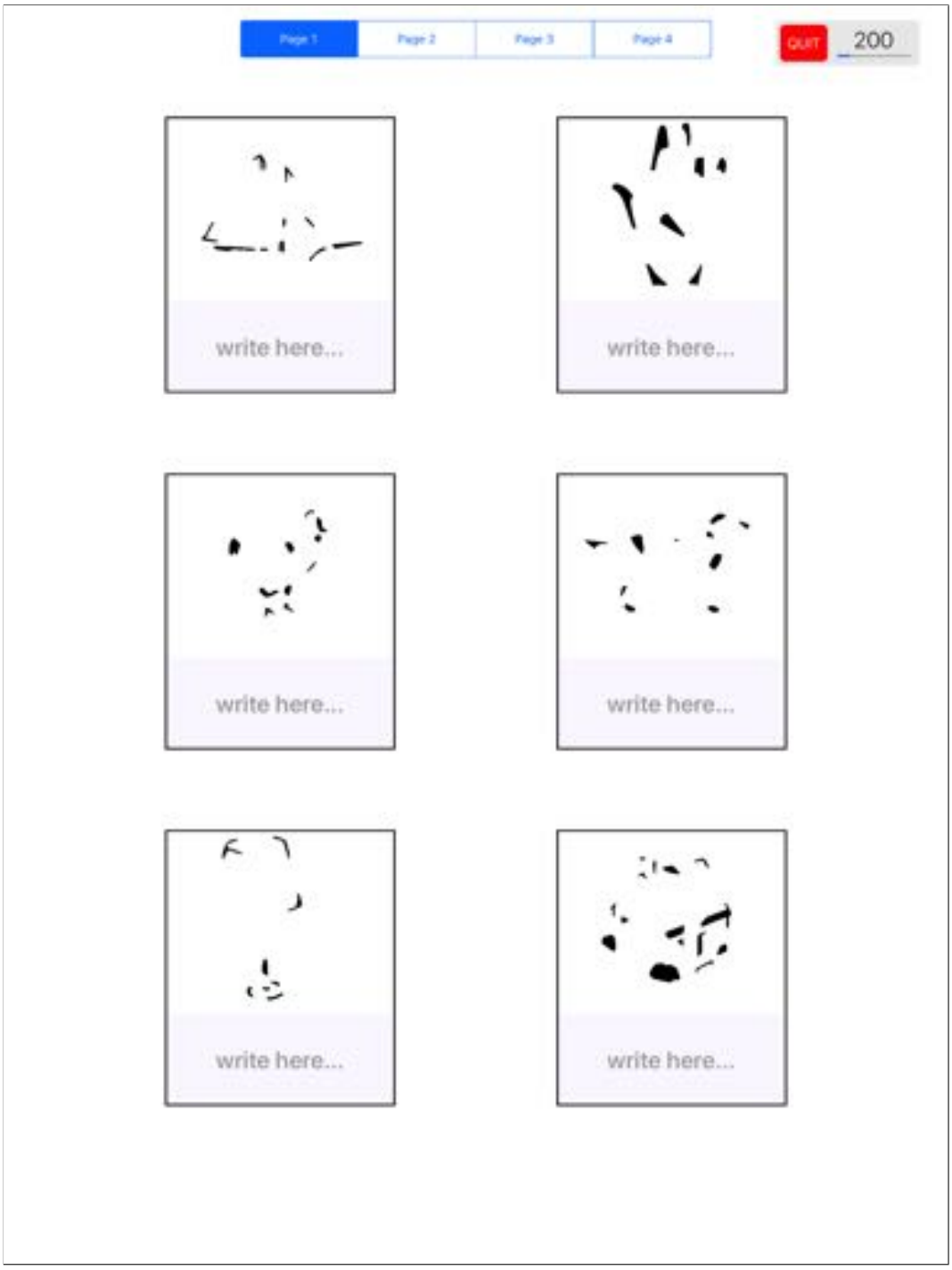


Figure B.6: Gestalt Completion Test — PEN & TABLET condition — Page 1

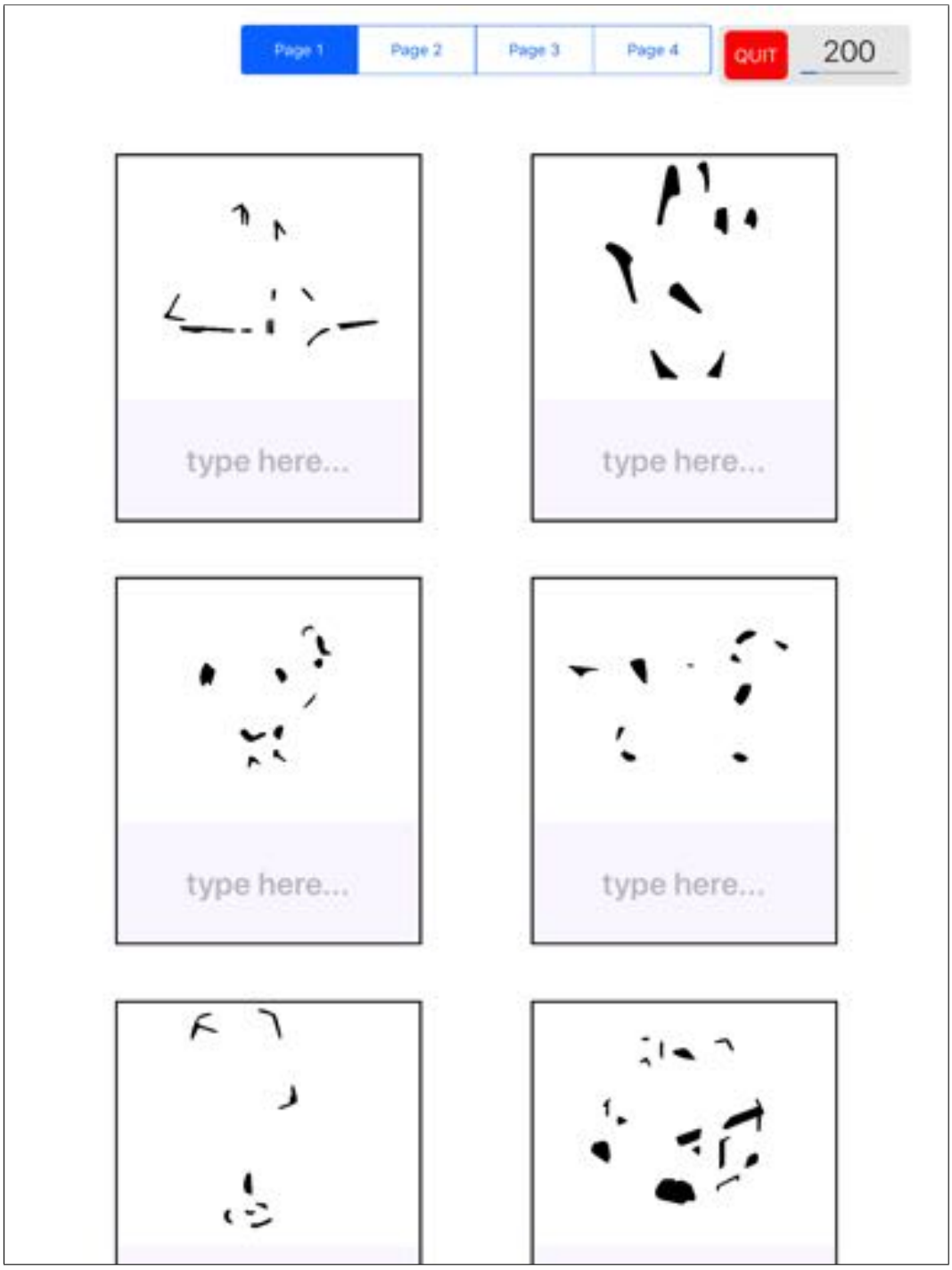


Figure B.7: Gestalt Completion Test — TOUCH & TABLET condition — Page 1

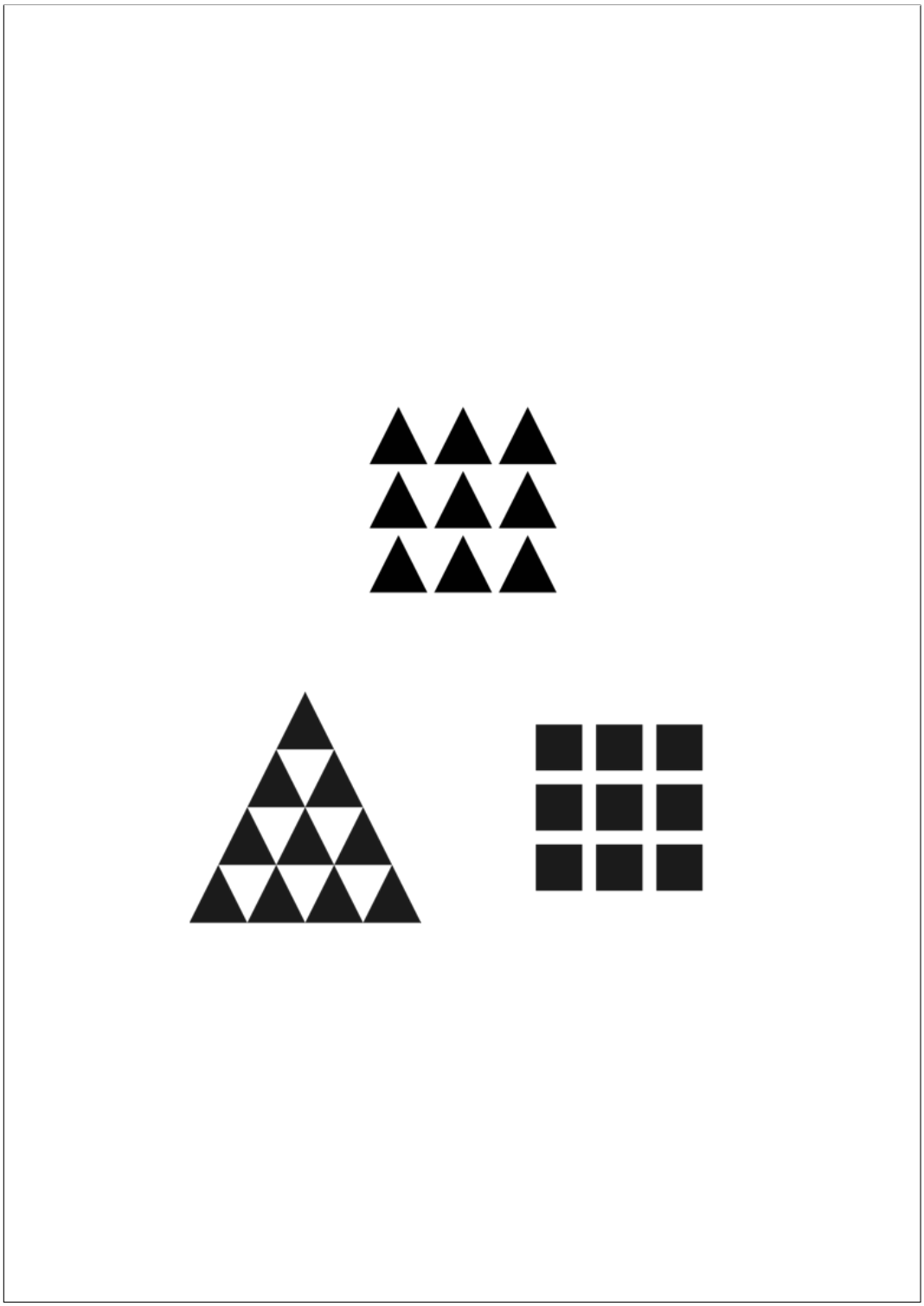


Figure B.8: Kimchi Palmer figures — PEN & PAPER condition

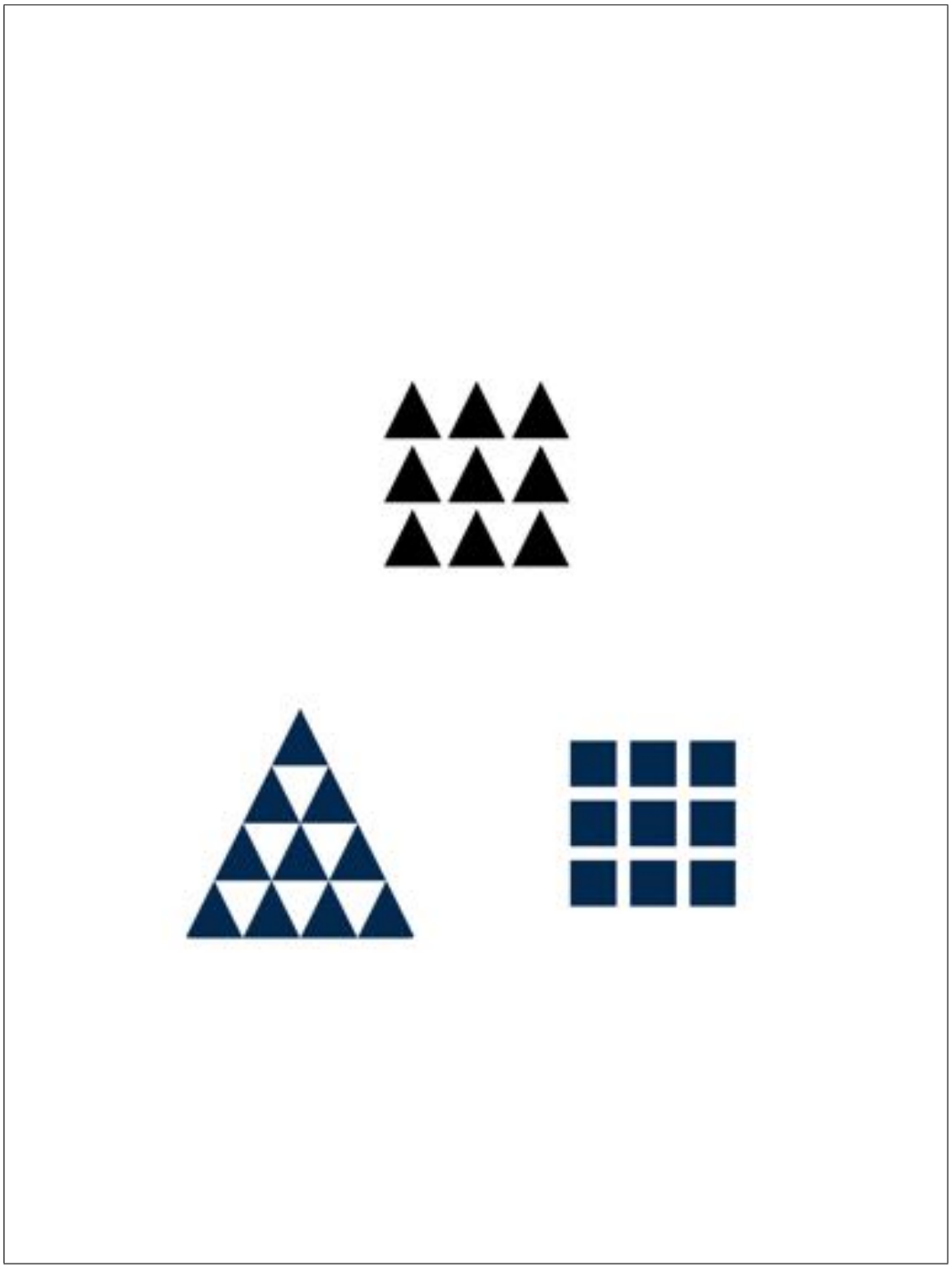


Figure B.9: Kimchi Palmer figures — PEN & TABLET condition

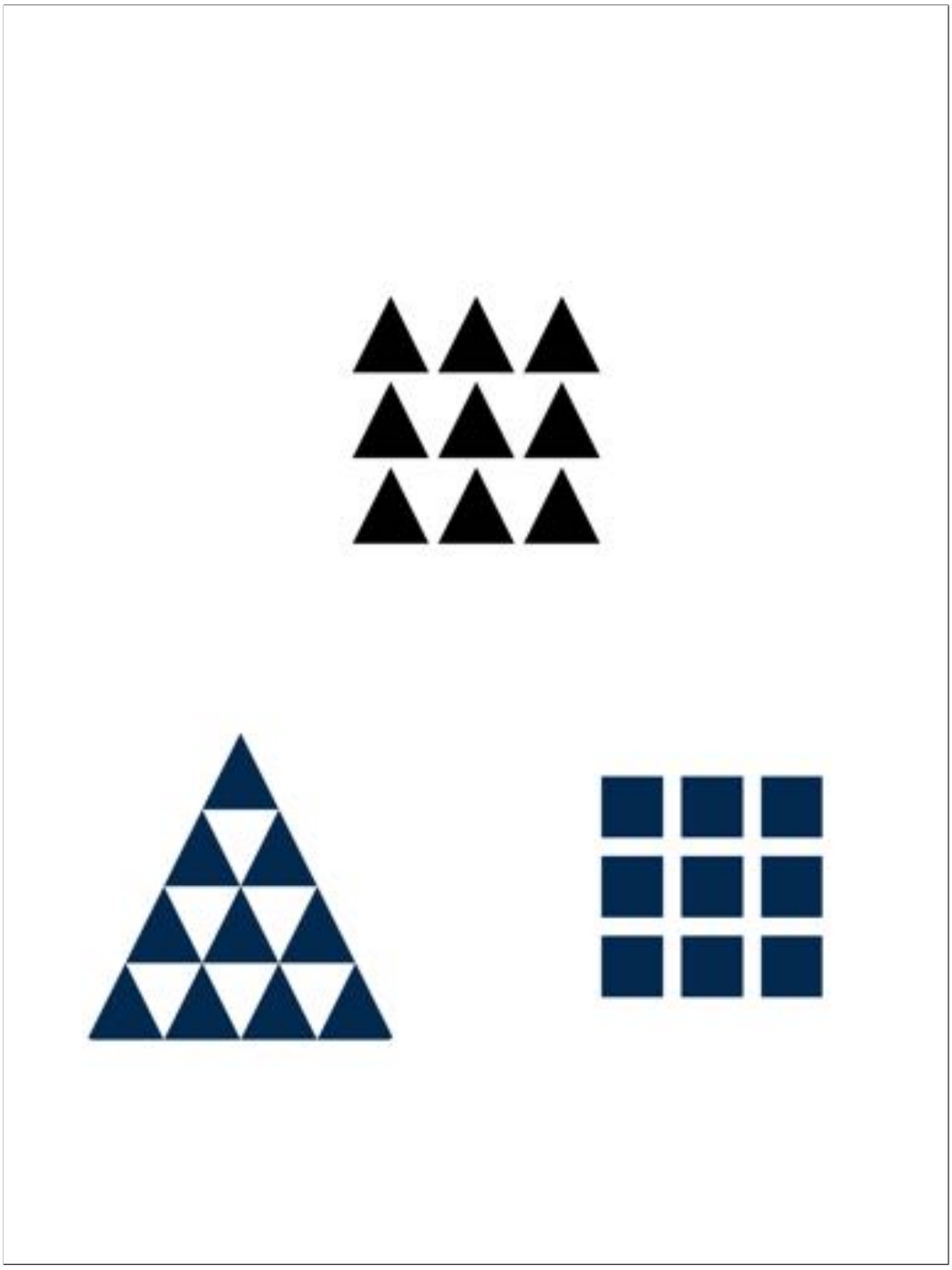


Figure B.10: Kimchi Palmer figures — TOUCH & TABLET condition

Figure B.11: Picture Completion Test — PEN & PAPER condition

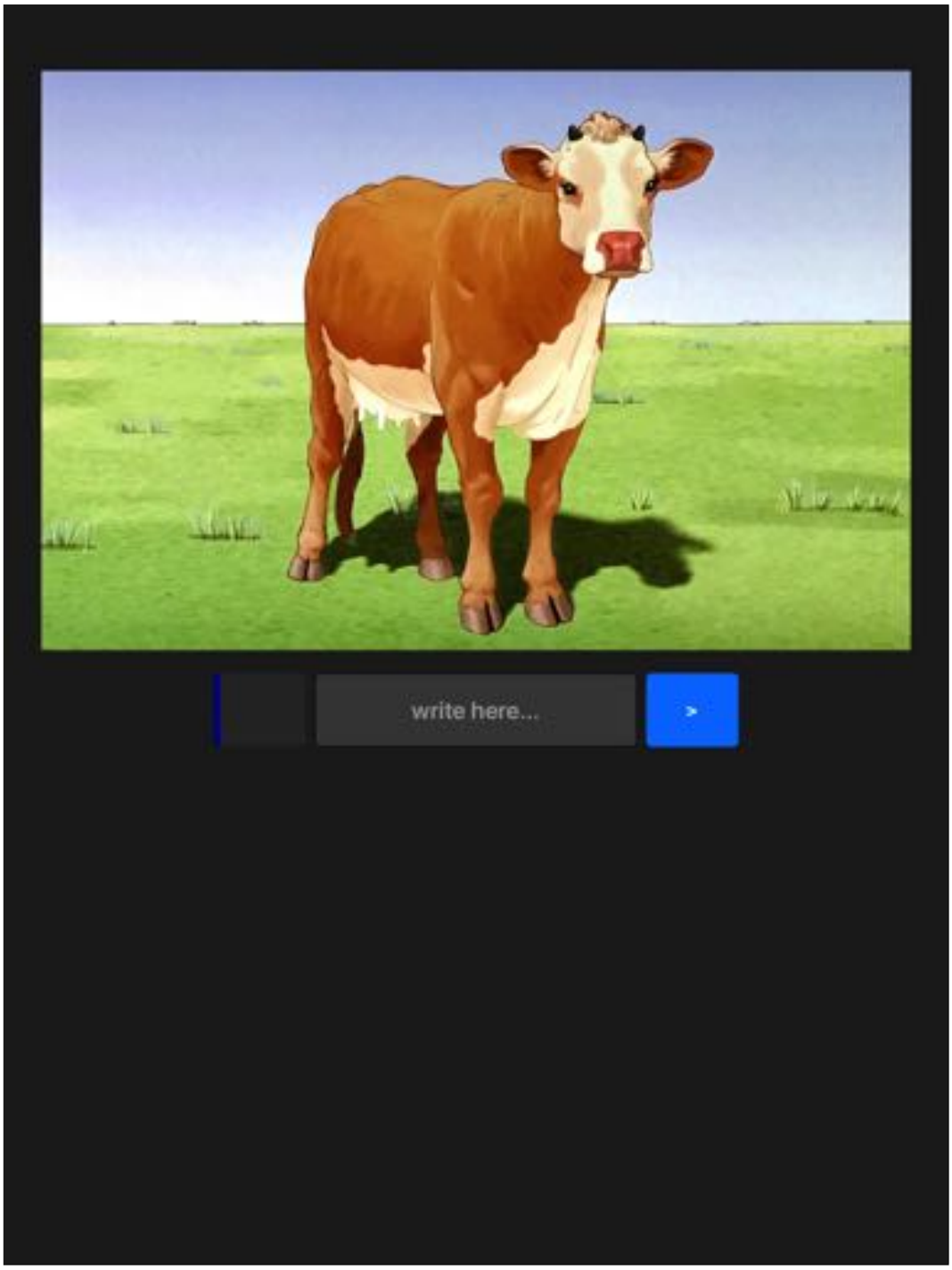


Figure B.12: Picture Completion Test — PEN & TABLET condition

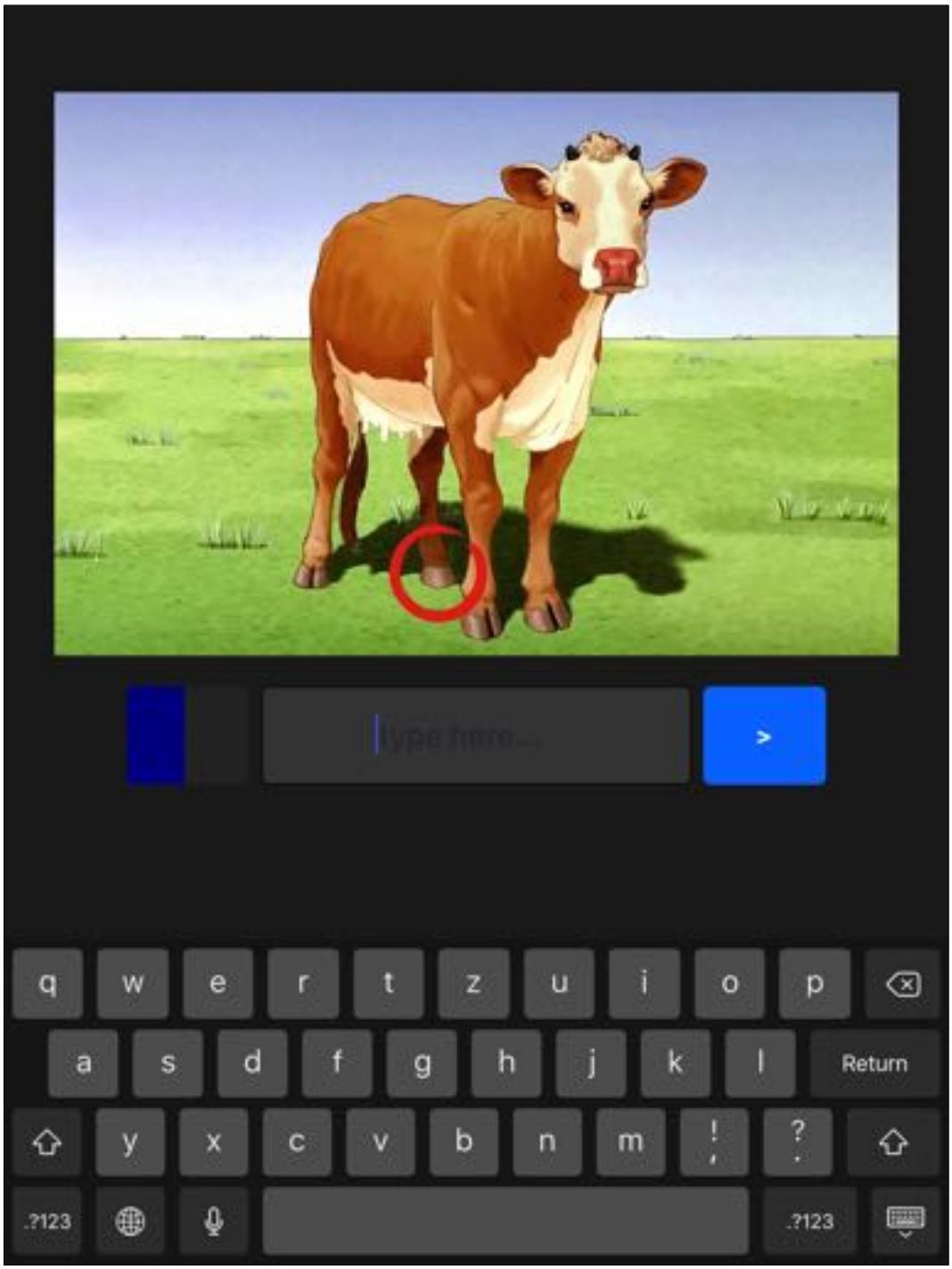


Figure B.13: Picture Completion Test — TOUCH & TABLET condition. The indicator that appeared at the position of a participant's tap location is placed on the foot

Appendix C

Kimchi Palmer Size Classes

Figure C.1 shows the different size classes used in the Kimchi Palmer task.

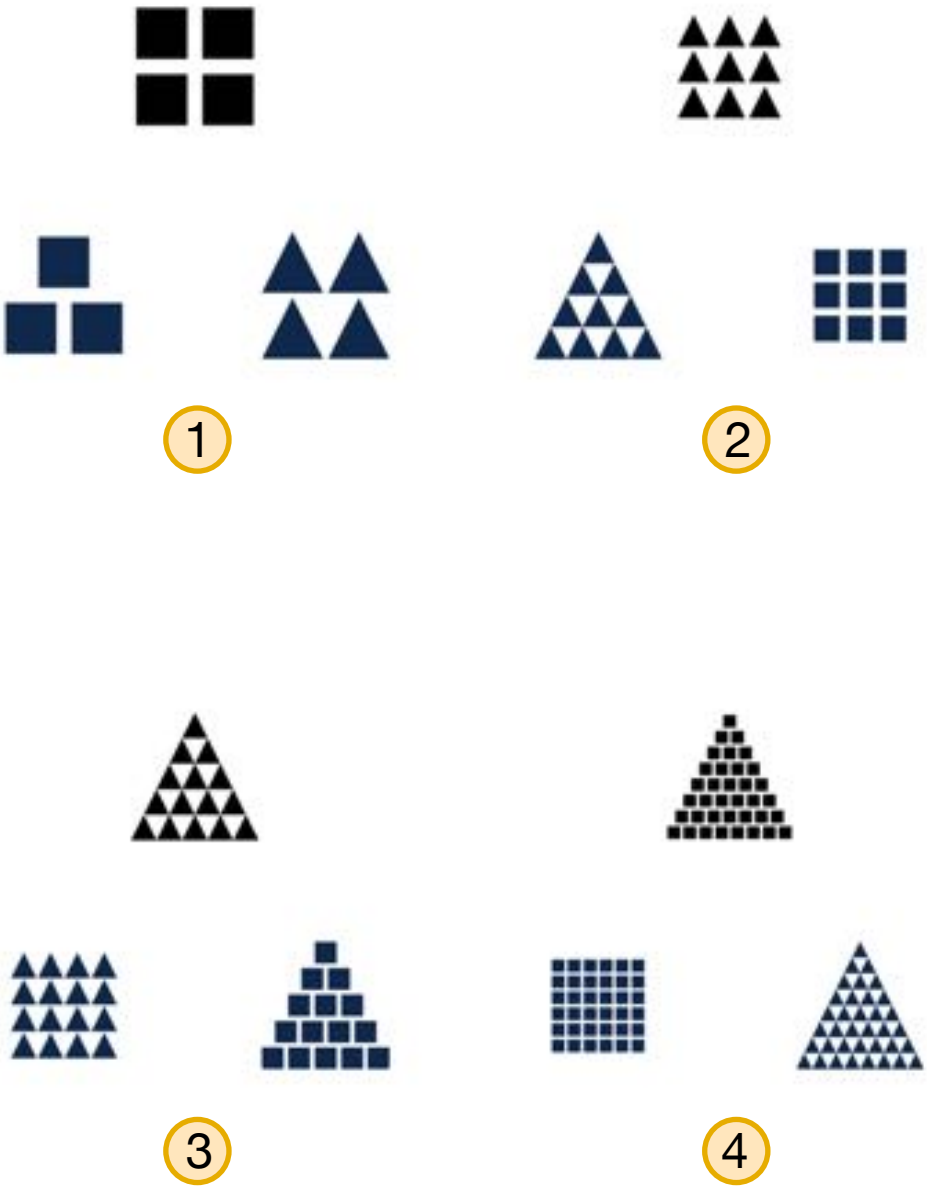


Figure C.1: Sizes classes used in the Kimchi Palmer task.
1: *Large*, global squares consist out of four objects and triangles out of three.
2: *Mid*, global squares consist out of nine objects and triangles out of 10.
3: *Small*, global squares consist out of 16 objects and triangles out of 15.
4: *Tiny*. global squares consist out of 36 objects and triangles out of 28.

Bibliography

Susan M. Belmore. Reading computer-presented text. *Bulletin of the Psychonomic Society*, 23(1):12–14, Jan 1985. URL https://doi.org/10.3758/BF03329765.

Douglas Bernstein. *Essentials of Psychology*. Cengage Learning, 2018.

Shawn M. Bodmann and Daniel H. Robinson. Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, 31(1):51–60, 2004. URL http://dx.doi.org/10.2190/GRQQ-YT0F-7LKB-F033.

James A. Coleman. English-medium teaching in european higher education. *Language Teaching*, 39(1):1–14, 2006. doi: 10.1017/S026144480600320X.

Anthony Creed, Ian Dennis, and Stephen Newstead. Proofreading on vdus. *Behaviour & Information Technology*, 6 (1):3–13, 1987. URL http://dx.doi.org/10.1080/01449298708901814.

William H. Cushman. Reading from microfiche, a vdt, and the printed page: Subjective fatigue and performance. *Human Factors*, 28(1):63–73, 1986. URL http://dx.doi.org/10.1177/001872088602800107.

Ruth B Ekstrom, John W French, Harry H Harman, and Diran Dermen. Manual for kit of factor-referenced cognitive tests. *Princeton, NJ: Educational testing service*, 1976.

Klaus Fiedler. Construal level theory as an integrative framework for behavioral decision-making research and consumer psychology. *Journal of Consumer Psychology*, 17

(2):101 – 106, 2007. URL https://doi.org/10.1016/S1057-7408(07)70015-3.

Jens Förster, Ronald S Friedman, and Nira Liberman. Temporal construal effects on abstract and concrete thinking: consequences for insight and creative cognition. *Journal of personality and social psychology*, 87(2):177, 2004.

Shaul Fox and David Schwartz. Social desirability and controllability in computerized and paper-and-pencil personality questionnaires. *Computers in Human Behavior*, 18(4):389 – 410, 2002. URL http://dx.doi.org/10.1016/S0747-5632(01)00057-7.

Jun Fukukura, Melissa J Ferguson, and Kentaro Fujita. Psychological distance can improve decision making under information overload via gist memory. *Journal of Experimental Psychology: General*, 142(3):658, 2013.

John D. Gould and Nancy Grischkowsky. Doing the same work with hard copy and with cathode-ray tube (crt) computer terminals. *Human Factors*, 26(3):323–337, 1984. URL http://dx.doi.org/10.1177/001872088402600308.

http://blog.archive.org. Photo of Apple Macintosh. URL http://blog.archive.org/2017/04/16/early-macintosh-emulation-comes-to-the-archive/.

https://commons.wikimedia.org. Bundesarchiv, Bild 183-1987-0925-014 / Grubitzsch (geb. Raphael), Waltraud / CC-BY-SA 3.0, via Wikimedia Commons. URL https://upload.wikimedia.org/wikipedia/commons/9/98/Bundesarchiv_Bild_183-1987-0925-014%2C_Leipzig%2C_Deutsche_B%C3%BCcherei%2C_Mikrofiche-Leseger%C3%A4t.jpg.

https://vintagecomputer.ca. Photo of IBM 3277 Terminal. URL https://vintagecomputer.ca/ibm-3277-terminal/.

William James. Psychology: The Briefer Course, 1961.

Geoff Kaufman and Mary Flanagan. High-low split: Divergent cognitive construal levels triggered by digital and non-digital platforms. In *Proceedings of the 2016 CHI*

Conference on Human Factors in Computing Systems, CHI '16, pages 2773–2777, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3362-7. URL http://doi.acm.org/10.1145/2858036.2858550.

Snezhanka Kazakova, Veroline Cauberghe, Mario Pandelaere, and Patrick De Pelsmacker. Can't see the forest for the trees? media multitasking leads to local perceptual processing and concrete construals in subsequent tasks. *ACR European Advances*, 2013.

Ruth Kimchi and Stephen E Palmer. Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 8(4):521, 1982.

Nira Liberman and Jens Förster. The effect of psychological distance on perceptual level of construal. *Cognitive Science*, 33(7):1330–1341, 10 2009. URL https://doi.org/10.1111/j.1551-6709.2009.01061.x.

B. Jean Mason, Marc Patry, and Daniel J. Bernstein. An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research*, 24(1):29–39, 2001. URL http://dx.doi.org/10.2190/9EPM-B14R-XQWT-WVNL.

Daniel K Mayes, Valerie K Sims, and Jefferson M Koonce. Comprehension and workload differences for vdt and paper-based reading. *International Journal of Industrial Ergonomics*, 28(6):367 – 378, 2001. URL http://dx.doi.org/10.1016/S0169-8141(01)00043-9.

Elinor McKone, Anne Aimola Davies, Dinusha Fernando, Rachel Aalders, Hildie Leung, Tushara Wickramariyaratne, and Michael J Platow. Asia has the global advantage: Race and visual attention. *Vision research*, 50 (16):1540–1549, 2010.

Paul Muter, Susane A. Latrémouille, William C. Treurniet, and Paul Beam. Extended reading of continuous text on television screens. *Human Factors*, 24(5):501–508, 1982. URL http://dx.doi.org/10.1177/001872088202400501.

David Navon. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3): 353 – 383, 1977. URL https://doi.org/10.1016/0010-0285(77)90012-3.

Donald A. Norman. *The Design of Everyday Things*. Basic Books, Inc., 2002.

Jan Noyes, Kate Garland, and Liz Robbins. Paper-based versus computer-based assessment: is workload another test mode effect? *British Journal of Educational Technology*, 35(1):111–113, 1 2004. URL https://doi.org/10.1111/j.1467-8535.2004.00373.x.

Jan M. Noyes and Kate J. Garland. Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, 51 (9):1352–1375, 2008. URL http://dx.doi.org/10.1080/00140130802170387.

Richard Picking. Reading music from screens vs paper. *Behaviour & Information Technology*, 16(2):72–78, 1997. URL http://dx.doi.org/10.1080/014492997119914.

SoYon Rim, Elinor Amit, Kentaro Fujita, Yaacov Trope, Georg Halbeisen, and Daniel Algom. How words transcend and pictures immerse: On the association between medium and level of construal. *Social Psychological and Personality Science*, 6(2):123–130, 2015. URL https://doi.org/10.1177/1948550614548728.

Seth J Schwartz, Ronald L Mullis, and Richard M Dunham. Effects of authoritative structure in the measurement of identity formation: individual computer-managed versus group paper-and-pencil testing. *Computers in Human Behavior*, 14(2):239 – 248, 1998. URL http://dx.doi.org/10.1016/S0747-5632(98)00004-1.

Pamela K Smith and Yaacov Trope. You focus on the forest when you're in charge of the trees: power priming and abstract information processing. *Journal of personality and social psychology*, 90(4):578, 2006.

Betsy Sparrow, Jenny Liu, and Daniel M. Wegner. Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043): 776–778, 2011. doi: 10.1126/science.1207745. URL

http://science.sciencemag.org/content/333/6043/776.

Yaacov Trope and Nira Liberman. Construal-level theory of psychological distance. *Psychological review*, 117(2):440, 2010.

Robin R Vallacher and Daniel M Wegner. Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social psychology*, 57(4):660, 1989.

Cheryl J Wakslak, Yaacov Trope, Nira Liberman, and Rotem Alony. Seeing the forest when entry is unlikely: probability and the mental representation of events. *Journal of Experimental Psychology: General*, 135(4):641, 2006.

David Wechsler. Wechsler Adult Intelligence Scale – Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*, 22:498, 2008.

Martina Ziefle. Effects of display resolution on visual performance. *Human Factors*, 40(4):554–568, 1998. URL http://dx.doi.org/10.1518/001872098779649355.

Index

Typeset March 13, 2018