



Fachbereich Wirtschaftswissenschaften  
und Informationswissenschaften

WiSe 22/23

MASTERARBEIT

# Effizienz von Simplified User Interfaces in Software-Tutorials - Eine Eyetracking-Studie

Zur Erlangung des akademischen Grades Master of Arts (M.A.)

Eingereicht am: 06. Februar 2023



<b>Vorgelegt von:</b>	<b>Pia Bsdurrek</b>
Studiengang:	MIDMM18 Informationsdesign und Medienmanagement
Erstgutachter:	Prof. Dr. phil. Michael Meng Hochschule Merseburg
Zweitgutachter:	Dipl.-Soz. Christian Heckel Hochschule Merseburg



## Abstract

*The design of pictures in learning-oriented contexts is empirically well secured by the cognitive theory of multimedia learning and others. Many principles give clues for the design of efficient pictures. With regard to action-oriented documents, such as tutorials, however, there is still a wide research gap. While screenshots and visual signaling techniques have already proven to be efficient under certain boundary conditions, Simplified User Interfaces (SUI) represent a still almost unexplored design technique for screenshots. This study aims to investigate the effect of reduced details in SUIs and to provide indications how efficient SUIs are in comparison to conventionally used types of pictures. Therefore, an eye tracking study was conducted, in which subjects (n=24) were asked to complete a software tutorial for GIMP image manipulating software. The tutorial is available in three versions which only differ on their types of pictures: SUIs, signalized and raw screenshots. The results provide indications that SUIs are tending to be efficient for the identification of relevant areas in an image. However, while working with the tutorial, signalized screenshots tend to show a more positive effect. The subjects could execute the instructions more efficiently with the help of signalized screenshots than with SUIs or raw screenshots. Accuracy of action step execution did not differ depending on screenshot type. Further data collection is planned to verify the results using a larger sample and to investigate the effects of more complex tasks.*

Die Gestaltung von Abbildungen in einem lernorientierten Kontext ist unter anderem durch die *Cognitive Theory of Multimedia Learning* empirisch gut abgesichert. Viele Prinzipien geben Anhaltspunkte für die Gestaltung effizienter Abbildungen. In Bezug auf handlungsorientierte Dokumente, wie Tutorials, bietet sich jedoch noch eine breite Forschungslücke. Während sich Screenshots und visuelle Signalisierungstechniken bereits unter gewissen Randbedingungen als effizient bewiesen haben, stellen Simplified User Interfaces (SUIs) eine noch nahezu unerforschte Designtechnik für Screenshots dar. Ziel dieser Studie ist es, den Effekt reduzierter Detailinformationen in SUIs zu untersuchen und Hinweise dafür zu liefern, wie effizient SUIs im Vergleich zu herkömmlich verwendeten Abbildungsarten sind. Hierfür wurde ein Eyetracking-Experiment durchgeführt, bei dem Proband\*innen (n=24) ein Software-Tutorial für ein Bildbearbeitungsprogramm bearbeiten. Das Tutorial liegt in drei Varianten vor, die sich lediglich hinsichtlich ihrer enthaltenen Abbildungsart unterscheiden: SUIs, signalisierte sowie unbearbeitete Screenshots. Die Ergebnisse liefern Hinweise darauf, dass SUIs sich tendenziell effizient für die Identifizierung relevanter Bereiche in einer Abbildung zeigen. Bei der Bearbeitung des Tutorials zeigen jedoch signalisierte Screenshots einen positiveren Effekt. Die Proband\*innen konnten die Handlungsanweisungen mit signalisierten Screenshots

effizienter ausführen als mit SUIs oder unbearbeiteten Screenshots. Die Genauigkeit der Ausführung der Handlungsanweisungen unterscheidet sich nicht hinsichtlich der Screenshot-Varianten. Weitere Datenerhebungen sind geplant, um die Ergebnisse anhand einer größeren Stichprobe zu verifizieren und die Auswirkungen komplexer Aufgabenstellungen zu untersuchen.

---

**Keywords / Schlagwörter:**

Simplified User Interface, Screenshots, Signalization / Signalisierung, Software-Tutorial, Eyetracking

## **Danksagung**

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Masterarbeit unterstützt und motiviert haben.

Zuerst gebührt mein Dank Prof. Dr. phil. Michael Meng und Dipl.-Soz. Christian Heckel, die meine Arbeit betreut und begutachtet haben. Für die hilfreichen Anregungen und konstruktive Kritik bei der Erstellung der Arbeit, aber auch für die Motivation in zähen Phasen nicht aufzugeben, möchte ich mich herzlich bedanken.

Ein weiterer Dank gilt allen Proband\*innen, welche sich freiwillig bereit erklärt haben, an meiner Eyetracking-Studie teilzunehmen. Ohne ihr Interesse und ihre Einsatzbereitschaft wäre die Arbeit nicht möglich gewesen.

Bei meinen Eltern möchte ich mich bedanken, da sie mir mein Studium durch ihre Unterstützung überhaupt erst möglich gemacht haben.

Außerdem danke ich Sarah für ihr stets offenes Ohr, ihre Motivation und für das Korrekturlesen meiner Masterarbeit.

Abschließend gilt mein besonderer Dank meinem Partner Benjamin für seinen starken emotionalen Rückhalt und seine Unterstützung, das Korrekturlesen und die unbändige Motivation, die er mir während der langen Zeit entgegengebracht hat.

# Inhaltsverzeichnis

Abstract.....	I
Abkürzungsverzeichnis.....	VI
<b>1. Einleitung .....</b>	<b>1</b>
1.1 Fragestellungen .....	2
1.2 Aufbau .....	3
<b>2. Theoretische Grundlagen .....</b>	<b>4</b>
2.1 Multimediale Lerntheorien und -prinzipien.....	4
2.1.1 <i>Cognitive Load Theory</i> .....	5
2.1.2 <i>Cognitive Theory of Multimedia Learning</i> .....	8
2.2 Lernförderliche Abbildungen.....	12
2.2.1 <i>Einflussfaktor: Vorwissen der Lernenden</i> .....	13
2.2.2 <i>Einflussfaktor: Relevanz der Abbildungen</i> .....	14
2.2.3 <i>Einflussfaktor: Aufmerksamkeitssteuerung</i> .....	16
2.2.4 <i>Einflussfaktor: Visuelle Komplexität der Abbildungen</i> .....	18
2.3 Abbildungen in Software-Tutorials.....	20
2.3.1 <i>Software-Tutorials</i> .....	22
2.3.2 <i>Screenshots</i> .....	25
2.3.3 <i>Visuelle Signalisierung</i> .....	28
2.3.4 <i>Simplified User Interfaces</i> .....	29
2.4 Hypothesen.....	35
<b>3. Methodisches Vorgehen .....</b>	<b>36</b>
3.1 Methode: Eyetracking .....	36
3.1.1 <i>Visuelle Wahrnehmung und Blickbewegungen</i> .....	39
3.1.2 <i>Eyetracking-Systeme und ihre Funktionsweise</i> .....	40
3.1.3 <i>Messwerte und Analysemöglichkeiten</i> .....	43
3.2 Studienüberblick .....	45
3.2.1 <i>Proband*innenauswahl und -zuordnung</i> .....	45
3.2.2 <i>Versuchsaufbau und Technik</i> .....	46
3.3 Studiendesign und Material .....	48
3.3.1 <i>Erhebung empirischer Daten</i> .....	48
3.3.2 <i>Kalibrierung / Validierung des Eyetrackers</i> .....	50
3.3.3 <i>Orientierungstest</i> .....	50
3.3.4 <i>Software-Tutorial</i> .....	52
3.3.5 <i>Abschließende Befragung</i> .....	55
3.4 Datenanalyse.....	56
3.4.1 <i>Empirie</i> .....	56
3.4.2 <i>Analyse der Eyetracking-Daten</i> .....	57
3.4.3 <i>Export und statistische Auswertung</i> .....	62
<b>4. Ergebnisse .....</b>	<b>64</b>
4.1 Persönlicher Hintergrund .....	64
4.1.1 <i>Vorerfahrungen</i> .....	64

4.2 Orientierungstest .....	66
4.2.1 Reaktionszeiten im Orientierungstest .....	66
4.2.2 Fixationsanzahl im Orientierungstest.....	67
4.3 Tutorial.....	69
4.3.1 Bearbeitungszeit des Tutorials .....	69
4.3.2 Fehlerrate im Tutorial.....	72
4.3.3 AOI-Auswertung .....	74
4.4 Bewertung des Tutorials und der Abbildungen.....	77
4.4.1 Subjektive Bewertung des Tutorials .....	77
4.4.2 Subjektive Bewertung der Abbildungen.....	79
<b>5. Diskussion.....</b>	<b>83</b>
<b>6. Fazit.....</b>	<b>91</b>
6.1 Ausblick .....	93
Anhang .....	i
Literaturverzeichnis.....	xv
Abbildungsverzeichnis .....	xxiii
Tabellenverzeichnis .....	xxvi
Selbstständigkeitserklärung.....	xxviii

## Abkürzungsverzeichnis

AOI	Area of Interest
CLT	Cognitive Load Theory (Theorie der kognitiven Belastung)
CATLM	Cognitive-affective theory of Learning with Media (Kognitiv-affektive Theorie des Lernens mit Medien)
CTML	Cognitive Theory of Multimedia Learning (Kognitive Theorie des multimedialen Lernens)
PC	Personal Computer
SUI	Simplified User Interface Zusätzlich auch Gruppenbezeichnung für Proband*innen, welche in der vorliegenden Studie ein Tutorial mit Simplified User Interfaces verwendet haben
SIG	Signalisierung Gruppenbezeichnung für Proband*innen, welche in der vorliegenden Studie ein Tutorial mit Screenshots mit Signalisierungen verwendet haben.
SCR	Screenshot Gruppenbezeichnung für Proband*innen, welche in der vorliegenden Studie ein Tutorial mit unbearbeiteten Screenshots verwendet haben.

**„Einfachheit ist die ultimative Form der Perfektion.“**

(Leonardo Da Vinci (\*1452 †1519)  
Universalgelehrter, Maler, Bildhauer, Ingenieur u.v.m.)

~ ~ ~ ~ ~

**„Weniger ist mehr.“**

(Ludwig van der Rohe (\*1886 †1969)  
Bedeutender Architekt, Vertreter des Minimalismus in der Architektur)

~ ~ ~ ~ ~

**„Vollkommenheit entsteht offensichtlich nicht dann, wenn man nichts mehr hinzuzufügen hat, sondern wenn man nichts mehr wegnehmen kann.“**

(Antoine de Saint-Exupéry (\*1900 †1944)  
Schriftsteller und Pilot)

~ ~ ~ ~ ~

**„Make simple tasks simple.“**

(Bjarne Stroustrup (\*1950)  
Informatiker, Professor und Entwickler von C++)

(Zitate aus dem Original frei übersetzt von der Autorin,  
Quelle: Interaction Design Foundation 2021)

## 1. Einleitung

Die obenstehenden Zitate bekannter Persönlichkeiten der letzten Jahrhunderte treffen im Grunde alle denselben Kern. Eine kurze und einfache Lösung eines Problems, ist die beste. Dies ist ebenso die Grundaussage des KISS-Prinzips, welches in den 1960er Jahren entstanden ist und mittlerweile Einzug in viele verschiedene Branchen gefunden hat. Das Akronym "KISS" steht dabei für „Keep it simple, stupid“ (alternativ auch bekannt als „Keep it simple stupid“, „Keep it short and simple“ und andere Varianten) (vgl. Interaction Design Foundation 2021). Im Rahmen des Software-Engineerings bedeutet das KISS-Prinzip, „[...] dass Systeme möglichst einfach und nicht komplex sein sollen.“ „KISS fordert Einfachheit. Dieses Prinzip sollte beherzigt werden, auch wenn seine Umsetzung nicht immer einfach ist.“ (vgl. Goll 2019: 40, vgl. auch Smith 1983: 3f). Damit gilt das KISS-Prinzip, nicht nur für die Softwareentwicklung, als zentrales Gestaltungsprinzip. Denn je einfacher ein System zu verstehen ist, desto besser funktioniert es und wird auch umso eher von den Benutzer\*innen angenommen. Wichtig ist dabei nur, dass die minimierte Komplexität dabei nicht die Funktionalität des Systems einschränkt (vgl. Boatman 2020, Interaction Design Foundation 2021). Doch wenn die Software oder das System an sich verständlich und ohne unnötige Komplexitäten gestaltet sein soll, sollten diese Grundsätze nicht ebenso für die

Dokumentation oder das Tutorial der entsprechenden Software gelten? Schließlich soll mit diesem Material der Umgang mit der Software gelernt werden.

Für Lernmaterialien haben sich im Laufe der Zeit viele multimediale Lerntheorien und -prinzipien entwickelt. An diesen kann man sich orientieren, um sein Material möglichst effektiv zu gestalten. Eins der bekanntesten Prinzipien ist das Multimediaprinzip von Richard. E. Mayer. Es lässt sich mit einem Satz zusammenfassen: „Menschen lernen besser von Wörtern und Bildern als von Wörtern allein.“ (übersetzt nach Mayer 2021: 117). Wenn Bilder oder Grafiken also prinzipiell zu einer Verbesserung des Lernprozesses dienen, wie müssen diese Grafiken dann gestaltet sein? Üblicherweise werden in Software-Dokumentationen und -Tutorials Screenshots verwendet (siehe Kapitel 2.3.2). Diese können unter Umständen, zum Beispiel bei einer komplexen Softwareoberfläche, sehr detailliert und unübersichtlich sein. Es kann mitunter schwerfallen, die richtigen Bereiche in einem Screenshot auf Anhieb zu identifizieren. Abhilfe kann hier die sogenannte Technik der Signalisierung bieten. Dafür werden zusätzliche grafische Elemente wie Rahmen, Pfeile oder farbliche Hervorhebungen dem Screenshot hinzugefügt, um die Aufmerksamkeit der Betrachtenden zu lenken. Wie in Kapitel 2.2.4 und 2.3.3 dargestellt, bietet dieses Prinzip unter Umständen Vorteile und führt zu einer effizienteren Nutzung.

Die oben genannten Zitate, sowie das KISS-Prinzip lassen aber jedoch erst einmal das Gegenteil vermuten. Demnach wäre eine Reduktion auf wesentliche Elemente von Vorteil und würde zu einer effizienteren Nutzung führen. An diesem Prinzip orientieren sich auch sogenannte Simplified User Interfaces (SUI). Darunter wird eine visuelle Darstellung einer Benutzeroberfläche verstanden, welche lediglich für das Verständnis notwendige Informationen zeigt. Alle irrelevanten Texte und Symbole werden entfernt oder stark reduziert dargestellt (vgl. Boatman 2020; siehe Kapitel 2.3.4. Abb. 2). In den letzten 10 Jahren hat diese Designtechnik auch Einzug in die Technische Kommunikation gefunden. Relevante Studien wurden jedoch bis zum heutigen Zeitpunkt nicht getätigt, um die Effizienz der reduzierten Detailinformationen empirisch zu belegen. In anderen Kontexten, wie beispielsweise bei dem Einsatz von Strichzeichnungen, liegen dahingegen bereits Studien vor (siehe Kapitel 2.2.4). Ob sich diese Ergebnisse auch auf Screenshots übertragen lassen können, bleibt bis dato offen.

## **1.1 Fragestellungen**

Dieser Forschungslücken zur Effizienz von Simplified User Interfaces in Software-Tutorials soll die vorliegende Arbeit entgegenwirken. Dabei soll der Reduktion der Detailinformation in SUIs ein besonderes Augenmerk geschenkt werden. Zur Untersuchung der Forschungslücke wurden folgende zwei Forschungsfragen gestellt:

F01: Wie effizient sind Simplified User Interfaces in Software-Tutorials im Vergleich zu herkömmlichen Abbildungsarten?

F02: Hat die Reduktion von Detailinformationen bei SUIs einen zusätzlichen, positiven Effekt?

Zur Untersuchung dieser Fragestellungen wurde ein Eyetracking-Experiment konzipiert, durchgeführt und ausgewertet. Dabei werden die Blickbewegungen von Proband\*innen aufgezeichnet, während sie mit einem Software-Tutorial arbeiten. Das Tutorial liegt dabei in drei Varianten vor, welche sich lediglich hinsichtlich ihrer enthaltenen Abbildungsarten voneinander unterscheiden. Das genauere Studiendesign wird in Kapitel 3.3 erläutert. Ziel der Studie ist es herauszufinden, wie Abbildungen in einem Tutorial gestaltet sein müssen, damit das Tutorial möglichst effizient bearbeitet werden kann.

## **1.2 Aufbau**

Das zweite Kapitel widmet sich zunächst dem theoretischen Hintergrund der Arbeit. Neben einer kurzen Einführung in bekannte multimediale Lerntheorien und -prinzipien folgt eine Abhandlung über die lernförderliche Gestaltung von Abbildungen. Im Anschluss wird der aktuelle Forschungsstand zum Untersuchungsmedium der vorliegenden Studie, dem Software-Tutorial, in den Fokus genommen und mit besonderem Hinblick auf die darin enthaltenen Abbildungsarten vorgestellt. Das Kapitel endet mit der Vorstellung der Hypothesen, welche mithilfe dieser Arbeit überprüft werden sollen.

Im dritten Kapitel werden die Methode und die eigentliche Studie präsentiert. Nach Hintergrundinformationen zur verwendeten Methodik des Eyetrackings wird daraufhin die Studie vorgestellt. Nach einem groben Überblick und Informationen zu Proband\*innen und verwendeter Technik wird genau auf den Aufbau und Ablauf der Studie sowie die dafür konzipierten Materialien eingegangen. Auch die spätere Analyse der erhobenen Daten wird erläutert. Kapitel vier stellt die Ergebnisse der Studie im Detail vor. Hier wird sowohl auf die erhobenen empirischen Daten zu den Proband\*innen und deren persönliche Einschätzungen, als auch auf die Eyetracking-Daten genau eingegangen. Die Diskussion der Ergebnisse erfolgt dann im nächsten Kapitel 5.

Im letzten Kapitel wird ein Fazit gezogen und die zuvor aufgestellten Fragestellungen und Hypothesen beantwortet. Die vorliegende Arbeit schließt schließlich mit einem Ausblick auf weitere Forschungsmöglichkeiten zum Thema Simplified User Interfaces in Software-Tutorials ab.

## 2. Theoretische Grundlagen

Im folgenden Kapitel werden zunächst die theoretischen Grundannahmen der Cognitive Load Theory (CLT) und der Cognitive Theory of Multimedia Learning (CTML) betrachtet, um zu verstehen, wie Lernende effektiv mit multimedialen Inhalten arbeiten. Im Anschluss wird der Fokus auf die lernförderliche Gestaltung von Abbildungen gelegt, wofür relevante Prinzipien der vorgestellten Theorien sowie weitere Forschungsergebnisse herangezogen werden. Ab Kapitel 2.3 wird der aktuelle Forschungsstand zum Untersuchungsgegenstand, dem Software-Tutorial mit Screenshots, Signalisierungen und Simplified User Interfaces genauer aufgearbeitet. Am Ende des Kapitels werden die aus den theoretischen Grundlagen resultierenden und dieser Arbeit zugrundeliegenden Hypothesen präsentiert.

### 2.1 Multimediale Lerntheorien und -prinzipien

Lernmaterialien und Software-Dokumentationen sollten lernförderlich gestaltet sein, um den Wissenserwerb bzw. das Verständnis bestmöglich zu fördern. Hierfür sind vor allem zwei theoretische Modelle bekannt, welche sich mit dem multimedialen Lernen auseinandersetzen. Sie versuchen kognitive Prozesse beim Lernen mit Hilfe von Bildern und Texten zu erklären. Zum einen ist das Modell der *Cognitive Load Theory (CLT)* von John Sweller (1999, 2005) zu nennen, zum anderen die *Cognitive Theory of Multimedia Learning (CTML)* von Richard Mayer (2012, 2021). Diese theoretischen Modelle konkurrieren nicht miteinander, vielmehr ergänzen sie sich. Zu Grunde liegt beiden Theorien, dass der Mensch einen visuellen und einen verbalen Informationsverarbeitungskanal besitzt. Beide weisen jedoch nur eine begrenzte Kapazität des Arbeitsgedächtnisses auf, weshalb die kognitive Belastung möglichst geringgehalten werden sollte. Dies kann beispielsweise durch eine Lenkung der Aufmerksamkeit auf relevante Bereiche erzeugt werden (siehe Kapitel 2.2.3) oder durch das Entfernen von irrelevanten Grafiken (siehe Kapitel 2.2.2) (vgl. Sung & Mayer 2012: 1619, vgl. Mayer 2021: 143 ff., 166 ff.).

Prozedurale Texte wie Software-Tutorials werden zwar grundsätzlich weniger zum reinen Zweck des Lernens gelesen. Vielmehr steht hier oft ein handlungsorientierter Ansatz des Lesens, um zu handeln beziehungsweise um zu lernen, wie man handelt im Vordergrund (siehe Kapitel 2.3.1, vgl. Redish 1989). Inwieweit Lerntheorien, wie die *Cognitive Load Theory (CLT)* und *Cognitive Theory of Multimedia Learning (CTML)*, auf solche handlungsorientierten Texte übertragbar sind, ist bisher nicht vollständig erforscht. Erste Studienergebnisse weisen jedoch darauf hin, dass sich beispielsweise der Einsatz von Abbildungen in prozeduralen Texten ähnlich positiv auf die kognitive Belastung auswirkt, wie bei lernorientierten Texten (bspw. van Genuchten et al. 2014). Auch bei Irrazabal et al.

(2016) zeigte sich ein Vorteil bei dem Einsatz von multimedial aufbereiteten, handlungsorientierten Instruktionstexten. Die Proband\*innen waren hier beispielsweise bei der Verwendung von Instruktionmaterial mit Abbildungen bzw. multimedialen Inhalten schneller in der Bearbeitung, als bei reinen Texten. Zudem machten sie mithilfe des multimedial aufbereiteten Materials weniger Fehler. Beide Studien bestätigen damit das *Multimediaprinzip* für handlungsorientierte Texte. Einen Hinweis auf die Übertragbarkeit des *Signalisierungsprinzips* von Richard Mayer auf handlungsorientierte Texte zeigt darüber hinaus eine Studie von Meng (2019). Diese bestätigt die *CTML* in der Hinsicht, dass Signalisierungen auch in Software-Tutorials die Benutzer\*innen bei der Identifikation relevanter Bildelemente unterstützen (vgl. Meng 2019: 396).

Aufgrund dieser ersten Forschungslage und der thematischen Ähnlichkeit zur Studie in der vorliegenden Arbeit, sollen die Lerntheorien und -prinzipien von Sweller und Mayer an dieser Stelle genauer betrachtet werden. Ziel der durchgeführten Studie ist es auch weitere Anhaltspunkte dafür zu liefern, ob die Prinzipien grundlegend auf visuelle Elemente in prozeduralen Handlungsanweisungen übertragen werden können.

### **2.1.1 Cognitive Load Theory**

Die erste der beiden Theorien ist die in den späten 1980er Jahren von John Sweller und Kolleg\*innen entwickelte *Cognitive Load Theory (CLT)*, nach dessen Grundannahme Lernen mit kognitiver Belastung verbunden ist (vgl. Sweller 1988). Das Arbeitsgedächtnis spielt dabei eine wichtige Rolle, insbesondere da seine Kapazität nur sehr beschränkt ist. Es kann lediglich eine sehr begrenzte Anzahl an Informationen verarbeitet werden, Miller ging hier von  $7 \pm 2$  Einheiten aus (vgl. Miller 1956, Kalyuga et al. 1999: 351). Aufgrund dieser Begrenzung des Arbeitsgedächtnisses wird Wissen laut der *CLT* in sogenannten Schemata gespeichert. Schemata sind kognitive Konstrukte, welche dazu dienen, eintreffende Informationen im Gedächtnis zu organisieren und sie im Langzeitgedächtnis abzuspeichern zu können. Hierdurch kann die kognitive Belastung reduziert werden. Das Langzeitgedächtnis des Menschen weist dabei nahezu unbegrenzte Kapazitäten auf (vgl. Kalyuga et al. 1999: 351). Aber nicht nur die Konstruktion von Schemata ist relevant, ebenso wichtig ist deren Automatisierung. Eine automatisierte Verarbeitung von eintreffenden Informationen durch Übungsprozesse trägt zu einer minimalen Belastung des Arbeitsgedächtnisses beim Lesen und Lernen bei. Ein grundlegendes Beispiel für die Kategorisierung von Informationen in Schemata stellen Buchstaben und Wörter dar. Der Mensch ist in der Lage, Zeichen (Buchstaben) zu erkennen und diese anhand zuvor abgespeicherter Schemata zu übersetzen. Damit können Texte gelesen und verstanden werden (vgl. Sweller et al. 1998: 255 ff.). Somit werden anhand der Schemata mehrere

gelernte Informationselemente miteinander verknüpft, so dass sie mit weniger kognitiver Belastung als ein einziges Element im Arbeitsgedächtnis verarbeitet werden kann (vgl. Kalyuga et al. 1999: 351).

Die kognitive Belastung wird in drei verschiedene Arten unterschieden. Diese Dreiteilung wurde später auch in die *Cognitive Theory of Multimedia Learning* integriert. Die erste Form der kognitiven Belastung ist die *Intrinsic Cognitive Load*, welche durch das Lernmaterial und dessen Komplexitätsgrad an sich definiert wird. Hier spielen bereits erlernte Schemata und ihre Verbindung mit neu erlerntem Wissen eine Rolle. Die intrinsische kognitive Belastung ist demnach niedrig, wenn wenige Interaktionen zwischen den einzelnen Lernelementen bestehen. Wenn also beispielsweise reine Vokabeln gelernt werden, wo die Aufgabe darin besteht, Bezeichnung der Wörter in zwei verschiedenen Sprachen auswendig zu lernen. Darüber hinaus müssen jedoch bei diesem Prozess keine Interaktionen zwischen dem bestehenden Vorwissen und dem neu erlernten Wissen ausgeführt werden. Schwierigkeiten beim Lernen dieser Wortpaare können jedoch trotzdem aufgrund der Anzahl der zu lernenden Elemente entstehen. Eine hohe intrinsische Belastung entsteht dahingegen, wenn zum Beispiel mathematische Formeln umgeformt werden sollen. Hierfür müssen mehrere Rechenprinzipien beherrscht und parallel angewendet werden, wodurch die kognitive Belastung steigt (vgl. Sweller & Chandler 1994: 189). Der *Extraneous Cognitive Load* bezeichnet die kognitive Belastung, die zur Verarbeitung des aufgenommenen Wissens benötigt wird. Der *Extraneous Cognitive Load* kann dabei durch die Gestaltung des Lernmaterials beeinflusst werden. Suboptimale Darstellungsformen, wie überflüssige und redundante Wiederholungen oder irrelevante Abbildungen, führen zu einer erhöhten Belastung. Ziel ist es also, das Design von Lernmaterialien so zu optimieren, dass die extrinsische Belastung reduziert wird (vgl. Sweller 2010: 125). Die dritte Form der kognitiven Belastung ist die *Germane Cognitive Load*. Während die intrinsische und extrinsischen Cognitive Load maßgeblich durch das Lernmaterial und nur nebensächlich durch die Eigenschaften des Lernenden beeinflusst wird, liegt bei der *Germane Cognitive Load* der Fokus auf dem Lernenden. Sie wird durch die Organisation und Integration von neuen Informationen in das bestehende Vorwissen verursacht und bezieht sich auf die im Arbeitsgedächtnis zur Verfügung stehenden Ressourcen. Mit der *Germane Cognitive Load* wird also der Aufwand des Lernenden bezeichnet, um das Lernmaterial zu verstehen (vgl. Sweller 2010: 125 ff.). Ziel der Cognitive Load Theory ist es, ein ausgeglichenes Verhältnis zwischen den drei kognitiven Belastungsformen zu finden, damit die begrenzte Kapazität des Arbeitsgedächtnisses nicht überlastet wird. Hierfür lassen sich 12 Gestaltungsempfehlungen aus der *Cognitive Load Theory* ableiten (siehe Tab.01).

GESTALTUNGSEMPFEHLUNGEN DER COGNITIVE LOAD THEORY	
Effekt der Zielfreiheit	Einsatz zielfreier oder zielunspezifischer Problemlöseaufgaben anstelle von Aufgaben mit einem vorgegebenen, spezifischen Ziel
Effekt ausgearbeiteter Lösungsbeispiele	Präsentation ausgearbeiteter Lösungsbeispiele anstelle von konventionellen Problemlöseaufgaben
Effekt der Problemvervollständigung	Verwendung von Problemen, bei denen der Lernende die teilweise vorgegebene Lösung selbst vervollständigen muss
Effekt der geteilten Aufmerksamkeit	Physikalische Integration multipler Informationsquellen
Modalitätseffekt	Gemeinsame Verwendung visueller und auditiver Informationen
Redundanzeffekt	Vermeidung von Redundanzen in multiplen Informationsquellen
Variabilitätseffekt	Erhöhung der Variabilität in unterschiedlichen Lernübungen
Effekt der Isolation interagierender Elemente	Isolation von Lernelementen hoher Aufgabenkomplexität
Elementinteraktivitätseffekt	Moderierender Einfluss der Elementinteraktivität auf die aufgeführten Gestaltungsempfehlungen
Imaginationseffekt	Imagination bereits gelernter Arbeitsschritte bei Experten
Expertise-Umkehr-Effekt	Moderierender Einfluss des Vorwissens des Lernenden auf die aufgeführten Gestaltungsempfehlungen
Effekt der abschwächenden Unterstützung	Abschwächung der Unterstützung von Lernenden bei ansteigender Expertise

Tab. 1: Gestaltungsempfehlungen der Cognitive Load Theory. (Quelle: Übernommen aus Rey 2007: 19.)

Für die vorliegende Arbeit zur Untersuchung von Abbildungsarten in Tutorials spielen dabei der *Effekt der geteilten Aufmerksamkeit* (auch *Split-Attention Effekt* genannt) sowie der *Expertise-Umkehr-Effekt* (*Expertise Reversal Effect*) eine Rolle. Auf letzteren wird in Kapitel 2.2.1 genauer eingegangen. Der *Effekt der geteilten Aufmerksamkeit* wird im folgenden Absatz näher betrachtet.

Beim Lernen von multimedialen Lernmaterialien ist es für den Lernenden oft notwendig, seine Aufmerksamkeit auf mehrere für den Lernprozess relevante und aufeinander bezogene Informationsquellen gleichzeitig legen zu müssen. Beispielsweise auf Texte und Bilder. Die Informationen aus beiden Quellen müssen nicht nur parallel aufgenommen, sondern auch mental integriert werden, wodurch der *Extraneous Cognitive Load* des Lernenden erhöht wird. Eine Vermeidung dieser Aufmerksamkeitsaufteilung sollte daher bei der Gestaltung von Lernmaterialien möglichst von vornherein das Ziel sein. Bei einem Diagramm hat es sich beispielsweise als hilfreich herausgestellt, Beschriftungen jeweils in direkter physischer Nähe zum zu beschriftenden Element zu platzieren, anstatt gesammelt neben oder unter der Grafik. So kann die Beschriftung direkt zugeordnet und irrelevante kognitive Prozesse durch eine mentale Integrierung der beiden Informationen vermieden werden (vgl. Sweller 1998: 278, vgl. Ayres & Sweller 2014: 206 ff.). Eine weitere Möglichkeit

besteht darin, die Aufmerksamkeit der Lernenden aktiv zu steuern, indem beispielsweise durch farbliche Signalisierungen eine Zuordnung zwischen Diagrammbereich und Text erleichtert wird. Dies wurde durch ein Experiment von Kalyuga und Kollegen nachgewiesen. Die Proband\*innen, welche signalisierte Diagramm-Text-Kombinationen verwendet hatten, empfanden die Aufgaben selbst zwar nur geringfügig einfacher als die Kontrollgruppe, sie zeigte jedoch signifikant bessere Ergebnisse in anschließenden Wissenstests (vgl. Kalyuga et al. 1999: 369). In einer Meta-Analyse hat Ginns (2006) 50 Studien zum Einfluss des *räumlichen und zeitlichen Kontiguitätsprinzips* auf den Lernerfolg von Anfänger\*innen analysiert. Die Ergebnisse zeigen, dass sich besonders bei komplexem Material ein positiver Effekt durch die Reduzierung der geteilten Aufmerksamkeit ergeben hat (vgl. Ginns 2006: 511).

Das *Split-Attention-Prinzip* entspricht auch dem (*räumlichen und zeitlichen*) *Kontiguitätsprinzip* der *CTML* (siehe Tab. 02), demnach Menschen besser lernen können, wenn zusammengehörige Texte und Abbildungen nah beieinanderstehen und gleichzeitig präsentiert werden (vgl. Mayer 2021: 207, 227). Auch die Teilung der Aufmerksamkeit durch unterschiedliche Medien, wie ein papierbasiertes, ausgedrucktes Software-Tutorial und der Ausführung der Aufgabe auf dem Computerbildschirm mithilfe von Tastatur und Maus, kann zu einem *Split-Attention-Effekt* führen. Dieser ist an der Stelle jedoch nur schwer zu vermeiden (vgl. Ayres & Sweller 2014: 213). In der vorliegenden Arbeit wurde versucht, der geteilten Aufmerksamkeit entgegenzuwirken, indem das Tutorial und die zu benutzende Software zumindest auf demselben Bildschirm präsentiert wurden. So konnte sich die Aufmerksamkeit der Proband\*innen vorrangig auf den Bildschirm konzentrieren.

### **2.1.2 Cognitive Theory of Multimedia Learning**

Neben der *Cognitive Load Theory* stellt die *Cognitive Theory of Multimedia Learning (CTML)* von Richard E. Mayer und Kolleg\*innen eine bekannte und empirisch oft untersuchte Theorie des multimedialen Lernens dar. Die *CTML* beschreibt, wie wir von Bildern und Wörtern lernen. Unter dem Begriff „Wörter“ wird im Folgenden sowohl das geschriebene als auch das gesprochene Wort verstanden. Die Theorie entstand vor mehr als 30 Jahren noch unter dem Namen „Model of meaningful Learning“ und wurde seitdem stetig weiterentwickelt (vgl. Mayer 2005: 44). Eine Erweiterung der *CTML* zur *Cognitive-Affective Theory of Learning with Media (CATLM)* erfolgte beispielsweise durch Moreno und Mayer Mitte der 2000er Jahre. Wesentliche Ergänzungen der *CATLM* betreffen zum einen die Unterteilung des Langzeitgedächtnisses in ein semantisches und episodisches Gedächtnis, zum anderen werden motivationale und metakognitive Faktoren in der Theorie betrachtet. Darüber hinaus werden in der Erweiterung, neben den bisherigen multimedialen

Lernmaterialien bestehend aus Wörtern und Texten, auch weitere Lernumgebungen wie virtuelle Realitäten miteinbezogen (vgl. Moreno & Mayer 2007: 313).

Die *Cognitive Theory of Multimedia Learning* basiert auf drei grundlegenden Annahmen zur Informationsverarbeitung des Menschen. Zum einen findet die Informationsverarbeitung durch zwei verschiedene Kanäle statt, einen für visuelle bzw. bildhafte Informationen, den anderen für auditive bzw. verbale Informationen (vgl. Mayer 2021: 34 f., Mayer 2005: 47). Mayer greift bei dieser Annahme auf die *Duale Kodierungstheorie* von Paivio (bspw. Clark & Paivio 1991, Paivio 1990) sowie das *Arbeitsgedächtnismodell* von Baddeley (1992) zurück. Die beiden Kanäle lassen sich wiederum hinsichtlich ihres Repräsentationsmodus und ihrer sensorischen Modalität unterscheiden. Unter Repräsentationsmodus wird dabei die Darbietungsform der Lernmaterialien verstanden, also ob es sich um Wörter oder Bilder handelt. Die sensorische Modalität spiegelt dabei die Sinnesorgane wider, durch die der Stimulus wahrgenommen und im Arbeitsgedächtnis repräsentiert wird. Die Ohren nehmen auditiv repräsentierte Stimuli wahr, die Augen sowohl Bilder als auch geschriebene Wörter. Als vorteilhaft für den Lernprozess zeigt es sich insbesondere, wenn sowohl der visuelle als auch der auditiv / verbale Kanal aktiviert wird. Der Mensch ist hierfür in der Lage, durch einen Kanal aufgenommene Informationen in den anderen Kanal zu übertragen. Beispielsweise kann aus einem auditiv aufgenommenen Satz ein mentales Bild von dessen Inhalt erzeugt werden (vgl. Mayer 2021: 35 f.).

Ein zweiter wesentlicher Aspekt der *CTML* betrifft die begrenzte Kapazität des Arbeitsgedächtnisses. Diese Annahme basiert ebenfalls auf Baddeleys Arbeitsgedächtnismodell und findet sich auch in der *Cognitive Load Theory* wieder. Demnach kann das menschliche Arbeitsgedächtnis nur eine sehr begrenzte Anzahl an visuellen oder verbalen Informationen gleichzeitig behalten und verarbeiten. Mayer bezieht sich dabei auf etwa 5-7 Einheiten, wobei diese durch Lerntechniken erweitert werden können. Um die begrenzte Kapazität des Arbeitsgedächtnisses bestmöglich auszunutzen zu können, müssen Entscheidungen über die Relevanz der aufgenommenen Informationen und sowohl deren Verbindung zueinander als auch deren Verbindung zu bereits erlangtem Vorwissen getroffen werden (vgl. Mayer 2005: 35 f.). Das allumfassende Ziel ist dabei, den Lernenden kognitiv nicht zu überlasten. Die dritte Annahme der *CTML* erläutert die aktive Informationsverarbeitung des Lernenden, bei der durch kognitive Prozesse ein kohärentes mentales Modell vom Lerngegenstand entwickelt wird. Hierfür ist es notwendig, relevante von irrelevanten Informationen zu selektieren und die relevanten Informationen strukturiert mit dem Vorwissen in Verbindung zu setzen (vgl. Mayer 2021: 37 f.). Zusammengefasst und visualisiert werden die Grundannahmen der *CTML* in folgendem Schaubild (Abb. 1):

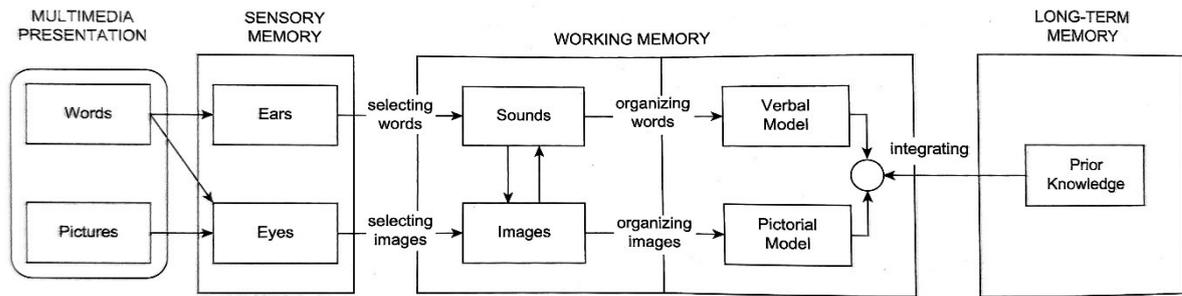


Abb. 1: Modell der Cognitive Theory of Multimedia Learning nach R. E. Mayer. (Quelle: Mayer 2021: 40.)

Deutlich wird anhand dieses Schaubildes auch die Dreiteilung der Gedächtnisspeicher in ein sensorisches, das Arbeits- und das Langzeitgedächtnis. Im sensorischen Gedächtnis werden die Wörter und Bilder aufgenommen und für sehr kurze Zeit gespeichert. Eine Auswahl an relevanten visuellen und verbalen / auditiven Informationen wird dann in das Arbeitsgedächtnis überführt und dort zwischengespeichert. Hier können dann mentale Transformationsprozesse stattfinden, indem zum Beispiel aus auditiven Informationen mentale Bilder erzeugt werden. Die Informationen werden dann durch kognitive Prozesse zur Bildung von verbalen und piktoralen mentalen Modellen weiterverarbeitet, in welche bereits bestehendes Vorwissen aus dem Langzeitgedächtnis integriert wird (vgl. Mayer 2021: 41 ff.).

Während des Lernens finden drei Arten von kognitiven Verarbeitungsprozessen statt, die die begrenzte Kapazität des Arbeitsgedächtnisses belasten. *Irrelevante Prozesse*, beispielsweise durch verwirrendes Instruktionsdesign sowie essenzielle Prozesse durch die Verarbeitung der relevanten Informationen. Hinzu kommen *generative Prozesse*, welche für ein tieferes Verständnis der Lernmaterie notwendig sind. Hierbei gehen *CTML* und *CLT* von derselben Unterteilung aus, Mayer spricht dabei jedoch von kognitiven Prozessen, während Sweller dies als kognitive Belastung bezeichnet. Ziel eines guten Instruktionsdesigns im Sinne der *CTML* ist es demnach, die extrinsische Belastung durch irrelevante Informationen zu minimieren, essenzielle mentale Prozesse bestmöglich zu unterstützen und generative Prozesse zu fördern (vgl. Mayer 2021: 51 ff.). Für diese drei Ziele wurden von Mayer und Kolleg\*innen 15 Prinzipien des multimedialen Lernens entwickelt, die in Tabelle 2 auf der nächsten Seite dargestellt werden.

<b>PRINZIPIEN DES MULTIMEDIALEN LERNENS NACH MAYER (2021) ZUR...</b>		
<b>... Entlastung der extrinsischen Belastung</b>		
Kohärenzprinzip*	Menschen lernen besser, wenn irrelevante Wörter, Abbildungen und Töne vermieden werden.	→ Irrelevantes Material weglassen
Signalisierungsprinzip*	Menschen lernen besser, wenn Hinweise hinzugefügt werden, die die wesentlichen Informationen hervorheben.	→ Wichtiges Material hervorheben
Redundanzprinzip	Menschen lernen besser von Abbildungen und Erzählungen als von Abbildungen, Erzählungen und Bildschirmtexten.	→ Gesprochenen Text nicht zusätzlich textlich abbilden
Räumliches Kontiguitätsprinzip	Menschen lernen besser, wenn zusammengehöriger Text und Abbildungen nahe zusammenstehen und nicht weit auseinander.	→ Zusammengehörigen Text und Bilder nahe zusammen abbilden
Zeitliches Kontiguitätsprinzip	Menschen lernen besser, wenn zusammengehöriger Text und Abbildungen gleichzeitig und nicht sukzessive dargestellt werden.	→ Zusammengehörige Texte und Bilder gleichzeitig präsentieren
<b>... Unterstützung essenzieller mentaler Prozesse</b>		
Segmentierungsprinzip	Menschen lernen besser, wenn eine Lerneinheit in benutzer*innen-gesteuerte Segmente unterteilt ist und nicht als kontinuierliche Einheit.	→ Lektionen in lerngerechte Einheiten aufteilen
Vorwissensprinzip	Menschen lernen besser von multimedialem Lernmaterial, wenn sie die wichtigsten Begriffe und Konzepte bereits kennen.	→ Schlüsselbegriffe und Konzepte vorab erklären
Modalitätsprinzip	Menschen lernen besser von Abbildungen und Erzählungen als von Animationen und Bildschirmtexten.	→ Geschriebene Texte durch Narration ersetzen
<b>... Förderung generativer Prozesse</b>		
Multimediaprinzip*	Menschen lernen besser von Abbildungen und Wörtern als von Wörtern allein.	→ Wörter und Bilder zusammen verwenden.
Personalisierungsprinzip	Menschen lernen besser von multimedialem Lernmaterial, wenn dieses umgangssprachlich verfasst ist und nicht in formaler Sprache.	→ Texte umgangssprachlich formulieren
Stimmprinzip	Menschen lernen besser, wenn eine freundliche, menschliche Stimme spricht anstatt einer computergenerierten Stimme.	→ Menschliche Stimmen verwenden
Bildprinzip	Menschen lernen nicht grundsätzlich besser von multimedialem Lernmaterial, wenn Sprechende zusehen sind.	→ Keine statischen Bilder von Sprechenden verwenden
Verkörperungsprinzip	Menschen lernen besser von multimedialem Lernmaterial, wenn der Lehrende ausgeprägte nonverbale Reize zeigt.	→ Lehrende sollen Gesten, Augenkontakt, Mimik und Bewegungen verwenden
Immersionsprinzip	Menschen lernen in immersiven, virtuellen 3D-Welten nicht besser als von 2D-Bildschirmpräsentationen	→ Keine virtuellen 3D-Räume verwenden
Generatives Aktivitätsprinzip	Menschen lernen besser durch praktische Anwendungen während des Lernens	→ Generative Prozesse durch Aufforderungen zu Selbsttests, zeichnen, sich vorstellen hinzufügen

Tab. 2: 15 Prinzipien des multimedialen Lernens nach Mayer (2021).

(Eigene Darstellung, übersetzt nach Mayer 2021: 399-402)

\* Für diese Arbeit relevante Prinzipien, welche im Folgenden genauer vorgestellt werden.

Relevant für die vorliegende Arbeit sind vor allem das *Multimediaprinzip*, das *Signalisierungsprinzip* sowie das *Kohärenzprinzip*. Diese Prinzipien sowie weitere relevante Aspekte zum effektiven Einsatz von Abbildungen in multimedialem Lernmaterial werden im Folgenden genauer betrachtet.

## 2.2 Lernförderliche Abbildungen

Das *Multimediaprinzip* von Mayer lautet: „People learn better from words and pictures than from words alone.“ (Mayer 2021: 117). Menschen lernen also effektiver, wenn das zur Verfügung gestellte Lernmaterial nicht nur aus reinem Text besteht, sondern zudem auch Abbildungen oder Grafiken enthält.

Zur Stützung seiner Theorie haben Mayer und Kolleg\*innen insgesamt 13 Studien durchgeführt, in denen die Effektivität von rein textbasiertem Lernmaterial im Vergleich zu multimedialem Lernmaterial (bestehend aus Abbildungen und Texten bzw. Animationen und Erzählungen) untersucht wurden. Alle Studien bestätigten dabei die Grundannahme des *Multimediaprinzips* mit einer insgesamten Effektgröße ( $d=1.35$ ) (vgl. Mayer 2021: 126). Für spezifische Studienergebnisse siehe Mayer (1989a), Mayer & Gallini (1990), Mayer et al. (1996), Mayer & Anderson (1991, 1992), Moreno & Mayer (1999), Sung & Mayer (2012) und Ponce & Mayer (2014).

In seinen theoretischen Ausführungen bezieht sich Mayer darauf, dass die Kombination aus Wörtern und Bildern effektiv dazu beitragen, sowohl ein verbales als auch ein visuelles mentales Modell vom Lerngegenstand zu entwickeln und diese beiden Modelle miteinander verknüpfen zu können. Hierdurch entsteht ein tieferes Verständnis des Lerninhaltes. Wird stattdessen reines Textmaterial verwendet, muss der Lernende selbst anhand des Gelesenen visuelle Modelle entwickeln, welches zu einer Erhöhung der kognitiven Belastung führt und die Lernleistung mindert (vgl. Mayer 2021: 122). Das Einfügen von Bildern zu Lerntexten ist jedoch nicht in jedem Falle effektiv. Sung und Mayer (2012) weisen darauf hin, dass lediglich relevante Grafiken das Lernen effizienter gestalten, dekorative oder irrelevante Grafiken nicht (vgl. Sung & Mayer 2012: 1618). Dies spiegelt sich auch im *Kohärenzprinzip* wider, welches in Kapitel 2.2.2 vorgestellt wird. Neben der Relevanz sollten die Grafiken auch eine gute Qualität haben und einen pädagogischen Wert erfüllen, indem sie beispielsweise einen klaren Bezug zum Text aufweisen (vgl. Schnotz & Bannert 2003, vgl. Schmeck et al. 2014). Außerdem stellt sich ein positiver Lerneffekt von Bildern mehr ein, wenn Bilder und dazugehörige Texte simultan präsentiert werden (siehe auch Zeitliches Kontiguitätsprinzip, Tab. 2). Hierauf weisen die Ergebnisse einer Studie von Moreno & Mayer (2002) hin.

Wie bereits dargestellt, kann sich der Einsatz von Bildern in Lernmaterialien positiv auf den Lerneffekt auswirken. Gleichzeitig haben empirische Studien jedoch auch gezeigt, dass dies nicht grundsätzlich der Fall sein muss. Im Folgenden wird daher eine für diese Arbeit relevante Auswahl an Einflussfaktoren vorgestellt, welche die Effektivität der Abbildungen beeinflussen können.

### **2.2.1 Einflussfaktor: Vorwissen der Lernenden**

Die Integrierung des Vorwissens aus dem Langzeitgedächtnis in die im Arbeitsgedächtnis neu aufgenommenen und zu verarbeitenden Informationen stellt einen wesentlichen Prozess beim multimedialen Lernen dar. Das Vorwissen ist dabei jedoch ein individuell ausgeprägter Parameter. Gleichzeitig stellt es eines der bedeutendsten Unterscheidungsmerkmale der Lernenden dar, welches sich stark auf das Lernergebnis auswirken kann. Das Vorwissen, welches Lernende mitbringen, ist dabei selten exakt identisch. Um den Einfluss des Vorwissens auf den Lerneffekt jedoch analysieren zu können, werden die Lernenden vorrangig in Anfänger\*innen und Expert\*innen kategorisiert.

Mayer und Gallini (1990) haben in einer Studie gezeigt, dass das Vorwissen der Lernenden den lernförderlichen Einsatz von Grafiken in Lernmaterialien massiv beeinflussen kann. Während bei Anfänger\*innen ein starker positiver Lerneffekt durch den Einsatz von multimedialem Lernmaterial erzeugt werden konnte, zeigten Lernende mit großem Vorwissen der entsprechenden Thematik so gut wie keinen Unterschied in ihrer Performance bei der Verwendung von reinem Textmaterial im Vergleich zu multimedialen Präsentationen (vgl. Mayer & Gallini 1990: 724). Bei drei Experimenten von Kalyuga und Kollegen zeigte sich sogar ein direkter Umkehreffekt. Während sich die zusätzliche Integrierung von Diagrammen zu einem Lerntext bei Anfänger\*innen positiv auf das Lernergebnis auswirkte, zeigte sich bei Expert\*innen ein negativer Effekt durch die zusätzlichen Abbildungen (vgl. Kalyuga et al. 1998: 1). Dieser Effekt wird in der Literatur auch als *Expertise Reversal Effect*, also *Expertise-Umkehr-Effekt*, bezeichnet (bspw. Kalyuga et al. 2003, 2010, 2014). Sowohl in die *Cognitive Load Theory* als auch in die *Cognitive Theory of Multimedia Learning* wird dieser Effekt einbezogen, um Randbedingungen für die Verwendung von Abbildungen in Lerntexten zu definieren (für CTML: vgl. Mayer 2021: 133, für CLT: vgl. Sweller 2010: 131).

Erklärt werden kann der *Expertise-Umkehr-Effekt* durch die Konstruktion und Speicherung von aufgenommenen Informationen in Schemata bzw. mentalen Modellen im Arbeitsgedächtnis der Lernenden. Wie in den bereits vorgestellten Lerntheorien wird das Wissen zur Konstruktion dieser Schemata mit dem Vorwissen des Lernenden verknüpft. Anfänger\*innen, welche sich zum ersten Mal mit der Thematik beschäftigen, fehlt dieses

Vorwissen, so dass sie mehr Unterstützung dabei benötigen, selbst Schemata zu konstruieren (vgl. Kalyuga et al. 2003: 24). Dies kann durch den Einsatz von erläuternden Grafiken gelingen, da diese den Lerninhalt visualisieren. Somit wird der Lerninhalt sowohl über den visuellen als auch über den verbalen Informationsverarbeitungskanal des Gedächtnisses aufgenommen. Dies kann bei Anfänger\*innen zu einer Reduzierung der kognitiven Belastung führen und sich somit positiv auf den Lerneffekt auswirken, da die kognitiven Kapazitäten anderweitig verwendet werden können (vgl. Kalyuga et al. 1998: 1). Erfahrene Nutzer\*innen bringen jedoch Vorwissen und damit bereits bestehende mentale Modelle bzw. Schemata mit. Zusätzliche Erklärungen durch Visualisierungen werden hier nicht benötigt, stattdessen können sie an dieser Stelle redundant wirken und die irrelevante kognitive Belastung erhöhen (vgl. Kalyuga et al. 1998: 1, vgl. Kalyuga et al. 2003: 24).

Bei der Konzeption von Lernmaterialien sollte also die Zielgruppe des Dokuments im Blick behalten und anhand dieser genau hinterfragt werden, ob und welche Abbildungen sich als effizient herausstellen.

### **2.2.2 Einflussfaktor: Relevanz der Abbildungen**

Nicht nur das Vorwissen der Lernenden allein wirkt sich jedoch auf die Effektivität von Abbildungen aus. Von großer Bedeutung ist zudem, ob ein dargestelltes Bild relevant für den Lernprozess ist oder beispielsweise nur zu Dekorationszwecken dient.

In der *CTML* geht Mayer nicht nur davon aus, dass von der Kombination aus Wörtern und Bildern besser gelernt wird als von Wörtern allein. In seinem *Kohärenzprinzip* spezifiziert er dieses Prinzip genauer, indem er sich zusätzlich auf die nötige Relevanz der Wörter und Bilder bezieht. Das *Kohärenzprinzip* besagt dabei: „People learn better when extraneous material is excluded rather than included“ (Mayer 2021: 143). Menschen lernen demnach besser, wenn zusätzliches zwar interessantes, aber für den Lernprozess irrelevantes, Material vermieden wird. Hierunter fallen sowohl irrelevante Wörter und Bilder, Symbole als auch Musik. Auf letzteres soll an dieser Stelle nicht weiter eingegangen werden, da es für die vorliegende Arbeit nicht von Relevanz ist.

Acht von Mayer und Kolleg\*innen durchgeführte Experimente haben gezeigt, dass die Reduktion von zusätzlichen verbalen Informationen, wie Messungen und Berechnungen aus Formeln, einen positiven Effekt auf den Lernerfolg hat (vgl. Mayer 2021: 155 f.). Auch ältere Studien haben diesbezüglich bereits gezeigt, dass sich das Lernen mit Zusammenfassungen positiver auf das Ergebnis von Transfertests auswirkt, als wenn die Lernenden zuvor ganze Kapitel von Lerntexten gelesen haben (vgl. Reder & Anderson 1980). In seinen Untersuchungen zu irrelevanten Bildern konzentriert sich die *CTML* zwar lediglich auf das jeweils gesamte Bild als Untersuchungsgegenstand. Im Zusammenhang

dieser Arbeit und der dazu durchgeführten Studie soll jedoch überprüft werden, ob sich dieses Prinzip auch auf lediglich irrelevante Bildteile übertragen lässt. Denn bei Simplified User Interfaces werden irrelevante Bildelemente entfernt, um die kognitive Belastung gering zu halten (vgl. Bollen & Saremba 2020a: 14). Aus diesem Grund wird das Prinzip an dieser Stelle genauer betrachtet.

Das *Kohärenzprinzip* wird in der Textforschung auch als *Seductive Detail Effect* bezeichnet (vgl. Garner et al. 1992, vgl. Schraw & Lehman 2001). Später wurden in diesem Effekt auch zwischen *Seductive Text* und *Seductive Illustrations* unterschieden (vgl. Harp & Mayer 1998). Mayer (2021) erklärt seinen Ansatz damit, dass durch irrelevantes Material die Aufmerksamkeit des Lernenden von den relevanten Kerninhalten abgelenkt wird. Dies führt zu einer unnötigen kognitiven Belastung des Arbeitsgedächtnisses. Da das Arbeitsgedächtnis jedoch nur eine begrenzte Kapazität aufweist, stehen somit weniger Ressourcen für die lernrelevanten Inhalte zur Verfügung. Außerdem beeinflussen irrelevante Inhalte den Lernenden negativ dabei, ein mentales Modell vom Lerninhalt zu erstellen (vgl. Mayer 2021: 143 ff.).

Zur Untersuchung seiner Theorie, dass irrelevante Wörter und Texte das Lernen negativ beeinflussen, haben Mayer und Kollegen neun Experimente durchgeführt. (Harp & Mayer 1997, 1998, Mayer et al. 2001, Mayer et al. 2008, Sung & Mayer 2012). Hierbei sollten die Proband\*innen mithilfe einer multimedialen Präsentation lernen, im Anschluss wurde der Lernerfolg durch einen Transfertest überprüft. Während bei einer Proband\*innengruppe irrelevantes Material eingefügt wurde, beschränkte sich das Lernmaterial der anderen Gruppe auf die relevanten Inhalte. In allen neun Experimenten konnte dabei ein starker positiver Effekt durch den Verzicht auf irrelevante Text- und Bildinhalte festgestellt werden (Median Effektgröße  $d=1.27$ ) (vgl. Mayer 2021: 150 f.). Ähnliche Ergebnisse konnten auch in neueren Studien weiterer Forschender festgestellt werden, darunter beispielsweise Wang und Adesope (2017). Ein gegenteiliges Ergebnis erzielte dahingegen die Arbeit von Schraw (1998), hier führte zusätzliches Textmaterial sogar zu einem positiven Lerneffekt. Zahlreiche weitere Studien konnten darüber hinaus zu keinem eindeutigen Ergebnis führen (vgl. bspw. Garner et al.: 1991, Wade & Adams 1990, Moreno & Mayer 2002).

Generell lassen sich drei Arten von Abbildungen in multimedialen Präsentationen unterscheiden. Zum einen sind das *instruktive, lehrreiche Grafiken*, welche das Lernen des wesentlichen Lernstoffes erleichtern und somit für das Erreichen des Lernziels relevant sind. Hinzu kommen *interessante, verlockende Grafiken*. Diese haben keinen direkten Bezug zum Lernziel, wirken auf den Lernenden jedoch interessant und aufmerksamkeitsregend. Die dritte Grafikart ist die der *dekorativen Abbildungen*. Diese

dienen lediglich zu Dekorationszwecken und haben keine Relevanz für das Lernziel (vgl. Sung & Mayer 2012: 1619). In einer Studie von Sung und Mayer (2012) wurde der Einsatz dieser drei Grafikarten untersucht. Herausgestellt hat sich dabei, dass sich der Einsatz von *instruktiven Grafiken* am signifikant positivsten auf das Lernergebnis ausgewirkt hat. Die Lernenden motiviert haben dahingegen alle drei Grafikarten gleichermaßen (vgl. Sung & Mayer 2012: 1618). Sind Grafiken irrelevant, aber trotzdem von großem Interesse für den Lernenden, kann sich dies jedoch auch negativ auf den Lernprozess auswirken (vgl. Mayer et al. 2008).

Eine weitere Randbedingung stellt die individuelle Kapazität des Arbeitsgedächtnisses dar. Sanchez und Wiley (2006) fanden beispielsweise in einem Experiment heraus, dass Proband\*innen mit einer geringeren Kapazität ihres Arbeitsgedächtnisses stärker negativ von irrelevanten Inhalten beeinflusst werden als solche mit höheren Kapazitäten (vgl. Sanchez & Wiley 2006: 344). Ähnlich dazu zeigen sich bei Rey (2014) in einem Eyetracking-Experiment Tendenzen, dass sich irrelevantes Material stärker negativ auf den Lerneffekt auswirkt, wenn die Proband\*innen eine geringe Aufmerksamkeitskontrolle aufweisen (Rey 2014: 133).

### **2.2.3 Einflussfaktor: Aufmerksamkeitssteuerung**

Zur besseren Lenkung der Aufmerksamkeit von Lernenden auf relevante Text- oder Bildstellen trägt das *Signalisierungsprinzip* der *CTML* bei. Dies lautet: „People learn better when cues are added that highlight the organization of the essential material.“ (Meyer 2021: 166).

Mayer unterscheidet in seinem *Signalisierungsprinzip* zwischen verbaler und visueller Signalisierung. Zu verbaler Signalisierung zählen beispielsweise Textauszeichnungen wie Überschriften, oder Hervorhebungen innerhalb der Typographie (z. B. durch Farben, Unterstreichungen oder Fettdruck). Einen positiven Effekt konnte er dabei in 15 von 16 Studien feststellen, welche zwischen 1983 und 2019 durchgeführt wurden. Bei rein textbasierten Studien betrug der Median der Effektgröße dabei  $d=0.75$ , während multimediale Studien  $d=0.51$  aufwiesen (vgl. Mayer 2021: 173). Unter visueller Signalisierung werden farbliche Hervorhebungen relevanter und Verblässungen unwichtiger Bildbereiche sowie Pfeile, welche auf relevante Bereiche hinweisen oder auch zeigende Gesten von Instrukteur\*innen bei Lernvideos verstanden. Beide Arten der Signalisierung wurden für das in der vorliegenden Arbeit zu untersuchende Software-Tutorial genutzt. Der visuellen Signalisierung kommt jedoch mehr Relevanz zu, da signalisierte Screenshots eine Kontrollbedingung zur Untersuchung der Effizienz von Simplified User Interfaces darstellen.

Zur Untersuchung von visueller Signalisierung führten Mayer und Kolleg\*innen 12 Studien im Zeitraum von 2001-2019 durch, von denen 11 eine positive Auswirkung von Signalisierungen auf den Lerneffekt zeigten ( $d=0.71$ ) (vgl. Mayer 2021: 168 ff.). Ebenfalls untersuchten Dönmez und Kollegen das *Signalisierungsprinzip*. Sie führten hierfür Eyetracking-Experiment durch welches zeigen konnte, dass durch Signalisierungen die Aufmerksamkeit der Lernenden effektiv auf relevante Bereiche gelenkt und kognitive Belastungen minimiert werden können. Die Proband\*innen zeigten zudem zwar einen positiveren Lerneffekt bei signalisierten Abbildungen, dieser zeigte jedoch keine signifikanten Unterschiede (vgl. Dönmez et al. 2018: 700).

Das Lernen wird durch Signalisierungen also unterstützt, da die wesentlichen Strukturen des Lernmaterials hervorgehoben, und die Aufmerksamkeit der Lernenden damit auf relevante Bereiche gelenkt wird. So kann die extrinsische Belastung der Lernenden minimiert werden. Den Lernenden wird aktiv dabei geholfen, relevante Informationen zu identifizieren und zu organisieren, wodurch ein besseres Verständnis erzielt wird (vgl. Alpizar et al. 2020: 2097, vgl. Meng 2019: 396). Ergebnisse einer von Moreno und Abercrombie durchgeführte Studie zeigen zudem, dass Proband\*innen die kognitive Belastung durch signalisiertes Lernmaterial geringer wahrnehmen (vgl. Moreno & Abercrombie 2010). Als Rahmenbedingung legt Mayer für das *Signalisierungsprinzip* die Komplexität des Lernmaterials fest. Es hat sich gezeigt, dass Lernende mehr von Signalisierungen im Lernmaterial profitieren, wenn sich das Material als komplex herausstellt. Bei simplem Material zeigt Signalisierung hingegen keine zusätzlich positiven Effekte (vgl. Mayer 2021: 402). Bestätigt werden kann diese Rahmenbedingung durch eine Studie von Jeung und Kollegen aus dem Jahr 1997, welche zu einem ähnlichen Ergebnis gelangten (Jeung et al. 1997).

Inwieweit das Vorwissen der Lernenden sich auf die Effektstärke von Signalisierungen auswirkt, ist in der aktuellen Forschungslage umstritten. Schneider et al. (2018) haben eine Metaanalyse zur Untersuchung des Signalisierungsprinzips durchgeführt. In 103 von ihnen betrachteten Studien unter über 12.000 Teilnehmer\*innen zeigt sich eine insgesamt signifikante Reduzierung der kognitiven Belastung mit mittlerer Effektgröße der Lernenden, wenn diese mit Signalisierungen versehenes Lernmaterial genutzt haben. Sowohl auf die Bearbeitungszeit als auch auf die Fixationen relevanter Bereiche wirkte sich die Signalisierung dabei positiv aus. Nicht bestätigen konnte sich bei der Metaanalyse jedoch der von Mayer aufgezeigte *Expertise-Umkehr-Effekt* im *Signalisierungsprinzip* (vgl. Schneider et al. 2018: 1). Bezüglich des Vorwissens der Lernenden zeigen Alpizar und Kollegen in einer Metaanalyse von 29 Studien jedoch das gegenteilige Ergebnis. Hier zeigt sich besonders bei Anfänger\*innen ein verstärkt positiver Effekt durch den Einsatz von

Signalisierungen (vgl. Alpizar et al. 2020: 2115). Zu einem ähnlichen Ergebnis kommt auch eine weitere Metaanalyse von Richter, Scheiter und Eitel (vgl. Richter et al. 2016: 19).

Neben dieser uneinheitlichen Forschungslage bezüglich des Vorwissens der Lernenden bestehen des Weiteren noch Forschungslücken darin, wann, wie viel und welche Art von Signalisierungen effektiv sind und welche Art von Material besser oder schlechter geeignet ist. Die Vielzahl von Signalisierungsvarianten erschweren hier die Ausgangslage (vgl. Alpizar et al. 2020, vgl. van Gog 2014).

#### **2.2.4 Einflussfaktor: Visuelle Komplexität der Abbildungen**

Welche Rolle die visuelle Komplexität einer Abbildung in einem Software-Tutorial spielt, ist bislang wenig erforscht. Die visuelle Komplexitätsforschung an sich wird jedoch schon länger vollzogen und geht auf die Gestaltpsychologie und die menschliche Wahrnehmung im Allgemeinen zurück (vgl. Ries 2020: 25). Definieren kann man visuelle Komplexität auf unterschiedliche Art und Weise. Zu den häufig verwendeten Definitionen zählen die von Snodgrass und Vanderwart: „[Visuelle Komplexität ist] der Detaillierungsgrad oder die Komplexität, die in einem Bild enthalten ist“ (übersetzt nach Snodgrass & Vanderwart 1998: 183) und die auf Rao und Lohse (1993) zurückgehende Definition von Komplexität als „Schwierigkeitsgrad bei der verbalen Beschreibung eines Bildes“ (übersetzt nach Heaps & Handel 1999: 301).

Einer der Hauptaspekte der menschlichen Wahrnehmung liegt darin, Strukturen und Anordnungen von einem Stimulus erkennen zu können. Auf diesen grundlegenden Aspekt wurden daraufhin *Gestaltgesetze* definiert (vgl. Ries 2020: 25). Zu den *Gestaltgesetzen* zählen beispielsweise die *Gesetze der Nähe*, der *Geschlossenheit*, der *Figur-Grund-Trennung*, der *Prägnanz*, der *Erfahrung* oder der *guten Fortsetzung* (siehe bspw. Koffka 1935, Alexander 2013: 25 ff.). Die *Gestaltgesetze* greifen darauf zurück, dass die menschliche Wahrnehmung zwar individuell durch die sozio-kulturellen Prägungen und Erfahrungen eines jeden Menschen beeinflusst ist, jedoch durch die Macht der Gewohnheit in vielen Aspekten auch übereinstimmt. Ein Beispiel hierfür ist die Tatsache, dass Objekte in unmittelbarer Nähe kontrastreich und groß wahrgenommen werden, wohingegen weit entfernte Objekte kontrastarm und klein erscheinen (vgl. Alexander 2013: 12 ff.).

Ein relevantes *Gestaltgesetz* bezüglich der visuellen Komplexität ist das *Gesetz der Prägnanz*, auch Einfachheits- oder Minimalprinzip genannt. Die menschliche Wahrnehmung tendiert demnach dazu, die am einfachsten erfassbare Struktur zu erkennen. Viele kognitive Prozesse sind auf Strukturierungen von Mustern und Mustererkennung in visuellen Reizen spezialisiert. Die Muster und Strukturen erleichtern dabei die Interpretation der wahrgenommenen Reize. Bei einer Vielzahl von Details oder

einem hohen Komplexitätsgrad einer Abbildung wird es schwieriger diese Muster und Strukturen zu identifizieren und die kognitive Belastung steigt (vgl. Ries 2020: 26). Zu wenige Details, an denen sich der Betrachtende orientieren kann, können sich jedoch auch negativ auswirken.

Problematisch bei dem Begriff der (visuellen) Komplexität ist, dass kein einheitliches Maß zur Messung der Komplexität vorhanden ist. Zwar wurden immer wieder Versuche vollzogen, eine Metrik zu entwickeln, wie von Geiselman et al. (1982) oder Garcia et al. (1994). Durchsetzen konnte sich jedoch keines. Dies liegt vor allem an einer fehlenden theoretischen Einbettung. Ein existierendes, valides Maß würde es Forschenden ermöglichen, Auswirkungen von reduzierten oder zusätzlichen Detailinformationen genauer zu untersuchen und zu bestimmen, wie sich dies auf die Leistungsfähigkeit auswirkt (vgl. Forsythe 2009: 161). Zudem ist zu beachten, dass verschiedene Faktoren die Wahrnehmung von Komplexität überlagern. Dazu zählt zum einen die Vertrautheit des Abbildungsinhaltes. Wenn der Betrachtende mit dem abgebildeten Objekt bereits vertraut ist, bewertet er dies meist als weniger komplex als ein ihm völlig fremdes Objekt. Außerdem spielen die Aufmerksamkeit und das Interesse, welches durch den Stimulus auf sich gezogen wird, eine Rolle. Ein verstärktes Interesse hilft dem Betrachtenden die Aufmerksamkeit zu bündeln und den Interpretationsaufwand zu minimieren (vgl. Forsythe 2009: 163).

Das Konzept der visuellen Komplexität wird unter anderem auch in der Icon-Forschung angewendet. Hier hat sich gezeigt, dass besonders sehr simple, aber auch sehr detaillierte Icons von Proband\*innen schneller erkannt werden, als solche mit einer mittleren Anzahl von Linien (vgl. Forsythe 2009: 158). Neben der Icon-Forschung gibt es auch Studien, welche die Rolle von visuellen Details in Strichzeichnung von Technischen Dokumentationen erforschen und hier zu einem ähnlichen Ergebnis kommen. Vermutungen liegen nahe, dass sich diese Ergebnisse auch auf den Detailgrad von Screenshots übertragen lassen könnten. Auch ein Screenshot dient, ähnlich wie eine Strichzeichnung, vorrangig zur Orientierung und Identifizierung von für die Aufgabe relevanten Bereichen.

Eine Studie von Alexander et al. (2016) hat beispielsweise untersucht, ob die Anzahl visueller Details von Strichzeichnungen Auswirkungen auf die Performance der Benutzer\*innen der Dokumentation hat. Hierfür wurde eine Schnellstartanleitung für einen Laserdrucker in drei Versionen mit jeweils unterschiedlich detaillierten Strichzeichnungen analysiert. Die Ergebnisse der Studie weisen darauf hin, dass eine im Detailreichtum optimierte Anleitung zu einer signifikant schnelleren Bearbeitung führt. Unabhängig von dem Detailgrad der Abbildungen zeigte sich dabei die Fehlerrate. Mit jeder Variante wurden

die Proband\*innen erfolgreich beim Bearbeitungsprozess der Anleitung unterstützt (vgl. Alexander et al. 2016: 49ff). Zu einem ähnlichen Ergebnis kam auch bereits Szlichcinski (1984). Entscheidend ist hier, dass sich weder ein sehr hoher noch ein sehr niedriger Detailgrad vorteilhafter zeigt. Zu viele Details, also eine hohe Komplexität der Abbildung, kann eher dazu führen, den Betrachtenden zu verwirren. Werden dahingegen zu viele Informationen weggelassen, kann es ebenso zu einer schlechteren Performance führen, da Objektteile und deren Ausrichtung nur schwer identifiziert werden könnten. Von Aufgabe zu Aufgabe sollte hier entschieden werden, in welchem Detailgrad die Strichzeichnung optimiert werden sollte. Denn fest steht auch, dass eine zusätzliche Optimierung mit zusätzlichem Zeitaufwand und Kosten verbunden ist. (vgl. Alexander et al. 2016: 51, vgl. Schumacher 2007).

Ältere empirische Forschungen zeigen ein heterogenes Bild von der Effizienz von visuellen Details in Strichzeichnungen. Craik und Tulving (1975) zeigen in ihrer Arbeit beispielsweise auf, dass Strichzeichnungen eher einen hohen Detailgrad aufweisen sollten. Sie argumentieren damit, dass viele Details eine intensivere Verarbeitung im Gehirn erfordern. Dies hat wiederum zur Folge, dass das Gedächtnis besser auf diese Informationen zugreifen und sie nutzen kann. Gleichzeitig gibt es jedoch auch Studienergebnisse, die besagen, dass irrelevante Details schwieriger aus dem Gedächtnis abgerufen werden können als relevante Informationen. (vgl. Alexander et al. 2016: 52). Dies belegen zum Beispiel Studien von Pedzek & Chen (1982) sowie Pedzek et al. (1988). Hieraus lässt sich, ebenso wie aus einer Studie von Dwyer (1978), ein positiver Effekt durch die Reduzierung von Detailinformationen ableiten.

Die aktuelle Studienlage bedarf also noch weiterer Forschung bezüglich eines effektiven Detailgrades von Abbildungen. Auch im Hinblick auf die visuelle Komplexität von Screenshots lassen sich bisher nur wenig Aussagen treffen. Das *Signalisierungsprinzip* nach Richard Mayer zeigt jedoch, dass das Hinzufügen von Details, um auf relevante Bereiche hinzuweisen, eine positive Wirkung auf den Lerneffekt von komplexem Material erzielen kann (siehe Kapitel 2.2.3).

Die vorliegende Studie soll dazu beitragen, Hinweise auf eine effiziente Gestaltung hinsichtlich des abgebildeten Detailgrades von Screenshots zu liefern, um die Effektivität von Software-Tutorials zu verbessern.

### **2.3 Abbildungen in Software-Tutorials**

Schon in den ersten bildlichen Darstellungen der Ägypter wurden zum Teil komplexe Arbeitsabläufe abgebildet. Diese dienten zwar noch nicht als Anleitung, jedoch sind diese

bereits so detailgetreu, dass aus ihnen bereits Arbeitsabläufe rekonstruierbar sind (vgl. Ballstaedt 1999: 74 f.). Erste Texte und Bilder, welche anleitenden Zwecken gedient haben sollen, sind hingegen bereits aus dem 11. Jahrhundert vor Christus überliefert. Mehr als 2000 Jahre später war es Leonardo da Vinci, zu Zeiten der Renaissance, der visuelle Darstellungsformen, wie die Explosionszeichnung, perfektionierte. Diese wird noch heute in vielen Technischen Dokumentationen verwendet (vgl. Piehl 2002: 57, vgl. Schwender 1999: 10). Etwa zur selben Zeit kamen auch diverse Verfahren zur Reproduktion von Bildern auf, welche gemeinsam mit dem Buchdruck zu einer leichteren Verbreitung von Gebrauchsanleitungen führten und für einen Aufstieg der modernen Technik sorgten (vgl. Ballstaedt 1999: 77).

Die erste Gebrauchsanleitung in Buchform wurde 1420, noch vor Erfindung des Buchdrucks, veröffentlicht. Neben Montage- und Handlungsanweisungen enthielt das „Feuerwerk-Handbuch“ auch bereits Warnhinweise vor Gefahrenquellen bei der Verwendung von Feuerwerken (vgl. Schwender 1999: 12). Gebrauchsanleitungen bestanden in den folgenden Jahrhunderten jedoch vorrangig aus Texten. Bilder und Grafiken hielten erst später Einzug. Bis Ende des 19. Jahrhunderts wurden diese ausschließlich in Dokumentationen der Mathematik verwendet (vgl. Eberleh 1990: 68). Einen wesentlichen Aufschwung in der Erstellung von Gebrauchsanleitungen stellte die Zeit der Industrialisierung dar, in der eine große Anzahl neuartiger Maschinen und Geräte entwickelt wurde. Während anfangs vor allem Fachpersonal, wie die Bediener\*innen von Maschinen, in die Benutzung dieser eingeführt werden mussten, wurden in den darauffolgenden Jahren auch immer mehr Anleitungen für alltägliche Haushaltsgeräte wie Telefone, Waschmaschinen oder Kühlschränke benötigt (vgl. Piehl 2002: 57 f., vgl. Schwender 1999: 12 ff.). Zum heutigen Zeitpunkt sind Gebrauchsanleitungen gar nicht mehr wegzudenken und Hersteller\*innen dazu verpflichtet diese zu jedem ihrer Produkte beizufügen.

Neben den Gebrauchsanleitungen für Maschinen und Produkte kam mit dem Aufkommen der Computertechnik ein ganz neues Feld für die Dokumentation auf. Fortan mussten nicht nur physische Prozesse abgebildet und erklärt werden, sondern auch digitale Prozesse. Noch in den 1960er Jahren wurden Computer, riesige übermäßig teure Maschinen in großen Unternehmen oder beim Militär, ausschließlich von Programmierern und Fachleuten benutzt, um komplizierte Rechenprozesse durchzuführen. Während es in den 1970er Jahren bereits erste Entwicklungen hin zu einem auch von Laien bedienbaren Computer gab, stellte sich ein Erfolg in diesem Entwicklungsprozess erst einige Jahre später ein. Mit der Einführung des IBM PC 1981 kam der erste Personal Computer (PC) erfolgreich auf den Markt. In der Folge gab es einen Computer, der nicht nur vom

Fachpersonal, sondern von Jedermann zu Hause oder im Büro genutzt werden konnte. Anfangs noch vorrangig zur Textverarbeitung und Kalkulation genutzt, entwickelte er sich schließlich auch immer mehr zum Informations- und Kommunikationsgerät (vgl. Friedemann 2000: 165, 170).

### **2.3.1 Software-Tutorials**

Die fortschreitende Entwicklung der Computertechnik machte auch ein Umdenken bei der Software-Dokumentation notwendig. Bisher waren Computerspezialisten und Programmierer, welche bereits über umfassendes Vorwissen verfügten, die Zielgruppe von Software-Dokumentationen. Daher waren umfangreiche, sachorientierte Handbücher, welche als Nachschlagewerk für spezifische Probleme genutzt werden konnten, gut für diese Zielgruppe geeignet. Durch die Entwicklung hin zum Personal Computer für Jedermann nutzten jedoch fortan auch normale Büroangestellte oder einfache Privatleute ohne Vorkenntnisse den PC. Die Ansprüche an eine Software-Dokumentation waren hier vollkommen andere. Anstatt mit einer Dokumentation spezielle Anwendungsprobleme zu lösen, war es nun gefragt, Grundfunktionen einer Software-Anwendung kennenzulernen und Grundkenntnisse zu vermitteln. Hierfür war ein vollkommen neuer, instruktiver Designansatz notwendig, der zu der Entstehung von Software-Tutorials führte (vgl. van der Meij et al. 2009: 265 f.).

Während erste Tutorials noch nach dem Standard des Dokumentendesigns konzipiert wurden, sind in Folge wissenschaftliche Forschungsergebnisse immer mehr einbezogen worden. Tutorials wurden auch unter dem Gesichtspunkt der Usability weiterentwickelt. Dem zuvor von Sticht et al. (1976, 1985) gegebenen Ansatz des Lesens, um entweder zu lernen oder zu handeln, wich nun dem von Redish (1989) beschriebenen Prinzip des Lesens, um zu lernen, wie man handelt (vgl. van der Meij et al. 2009: 267, vgl. Redish 1989: 289).

Den Ansatz des Lesens um zu lernen verfolgen vor allem Lehrbücher oder wissenschaftliche Artikel, die gelesen werden, um theoretisches Wissen zu erlangen und Schemata bilden zu können. Dies trifft in gewisser Hinsicht auch auf das Lesen von Software-Tutorials zu. In einer Studie wurde diesbezüglich gezeigt, dass sich Nutzer\*innen bei der Verwendung von Tutorials durch das gelernte Wissen Schemata der Software bilden und aufgrund derer sie in Folge agieren (vgl. Carroll und Mack 1984, Redish 1989: 291). Lesen um zu handeln beschreibt hingegen den Prozess Handlungswissen zu erlangen, welches direkt angewendet werden kann, sich aber nicht dauerhaft gemerkt werden muss. Als Beispiel können hier Installationsanleitungen genannt werden. Dieses Wissen benötigt man einmalig und muss nicht dauerhaft immer wieder abgerufen werden. Im Zweifelsfall

kann man es erneut nachschlagen (vgl. Redish 1989: 289, Sticht et al. 1985). Software-Tutorials dienen dahingegen dazu, sich mit grundlegenden Funktionen eines Programmes vertraut zu machen, relevante Aufgaben damit auszuführen und das daraus gewonnene Wissen auf andere Aufgaben mit dem Programm zu transferieren. Man lernt ein Tutorial also nicht auswendig, sondern liest es und lernt dabei, wie man mit der Software umgeht (vgl. Redish 1989: 290). Redish und Kollegen stellen dabei fest, dass Tutorials effektiv sind, wenn diese das aktive Lernen unterstützen und die Nutzer\*innen verstehen, warum sie auf diese Weise agieren sollen. Von großer Bedeutung ist zudem die Relevanz der Aufgabenstellung für ihren Arbeitskontext, der sie zum Handeln und Ausprobieren animiert (vgl. Redish 1989: 292 f.).

Auch weitere Forscher\*innen beschäftigten sich in den 1980er Jahren mit prozeduralen Handlungsanweisungen und entwickelten Methoden, wie das *Information Mapping* (vgl. Horn 1982, 1993) oder diverse Richtlinien zur Strukturierung und Konzeption von Software-Dokumentationen. Von großer Relevanz zeigte sich diesbezüglich vor allem der *Minimalismus*-Ansatz von Carroll in den 1990er Jahren, welcher von vielen Forschenden wie Hans van der Meij, David Farkas oder JoAnn Hackos weiterentwickelt wurde. *Minimalismus* beinhaltet einen aufgaben- und handlungsorientierten Ansatz zur Gestaltung von Technischen Dokumentationen, bei denen die Nutzer\*innen im Mittelpunkt stehen (vgl. Bleiel & Carroll 2021, vgl. van der Meij & Carroll 1998). Die Wahl eines handlungsorientierten Ansatzes für die Dokumentation ist hier dabei das erste Grundprinzip des *Minimalismus*. Die Dokumentation oder das Tutorial soll demnach den Lesenden zum direkten Handeln und aktiven Bearbeiten einladen. Dies sollte durch reale Aufgaben geschehen, welche sich als relevant für den Nutzenden zeigen. Dabei ist es nicht notwendig, sämtliche Funktionen oder Möglichkeiten zu erläutern. Stattdessen sollte sich kurzgefasst und auf Fachjargon verzichtet werden. Ein wichtiger Punkt in dem Konzept ist darüber hinaus die Identifikation und Eliminierung von Fehlern in der Dokumentation. Fehler sollten, wenn möglich grundsätzlich vermieden werden. Ist dies nicht möglich, sollen Hinweise auf besonders fehleranfällige Handlungsschritte vermittelt werden, welche das Erkennen und Korrigieren der Fehler erleichtert und somit aus diesen gelernt werden kann (vgl. van der Meij & Carrol 1998: 21). 2021 wurde versucht das *Minimalismus-Prinzip* in die neue Version der Norm für Nutzungsinformationen (IEC/IEEE 82079-1) aufzunehmen. Hier wird sich jedoch vorrangig auf kritische und sicherheitsbezogene Informationen und die geringstmögliche Menge anderer Informationen bezogen. Hans van der Meij bewertet diese Beschreibung in einem Interview als „It's incomplete, it's fake, and it's not well-informed.“ (Vermeulen & van der Meij 2019: ab 23:45 min).

In den 1990er Jahren wurde der instruktive Ansatz bei Software-Tutorials weiter vertieft und wissenschaftlich untersucht, so dass zahlreiche Publikationen zur Konzeption von Software-Tutorials entstanden (bspw. Farkas 1999, van der Meij 1996 / 1998). Von wissenschaftlichem Interesse waren dabei vorrangig die Fragestellung zum Effekt von deklarativen und prozeduralen Informationen, zum bereits aus dem *Minimalismus-Prinzip* hervorgehenden Ansatz zum Umgang mit Fehlern in Tutorials sowie zum Einsatz von Screenshots in Tutorials (vgl. van der Meij et al. 2009: 272). Auf den Einsatz von Abbildungen in Tutorials wird in den folgenden Unterkapiteln intensiv eingegangen.

Die Motivation der Nutzer\*innen bei der Bearbeitung eines Software-Tutorials wurde erst nach der Jahrtausendwende stärker in den Fokus gerückt. Die Forschung konzentrierte sich hier zunehmend darauf, die Erfahrung der Nutzer\*innen (User Experience) positiv unterstützen zu können und die Motivation zu fördern.

In den letzten Jahren haben, bedingt durch den technologischen Fortschritt und die Entwicklung diverser Videoplattformen, Videotutorials immer mehr an Bedeutung gewonnen. Die mittlerweile geringen technischen Hürden bei der Produktion von Videos sowie dem einfachen und oft kostenlosen Zugang über Plattformen wie YouTube haben stark zu ihrem Erfolg beigetragen. Vorteile gegenüber papierbasierten Tutorials ergeben sich zum einen aus der auf Paivios *Dual Coding Theory* basierenden Annahme der CTML, dass die parallele Informationsaufnahme aus auditiven und visuellen Modalitäten das Lernen erleichtern. Praktische Vorteile ergeben sich daraus, dass die Nutzenden synchron Systemzustände zwischen dem Tutorial und ihrer Software abgleichen und exakt vorgeführt bekommen können, wie sie einen Handlungsschritt ausführen können (vgl. van der Meij & van de Meij 2014: 151). Studien von Palmiter und Elkerton (1993), Loyd und Robertson (2012) sowie van der Meij und van der Meij (2014) haben diesbezüglich gezeigt, dass mit videobasierten Tutorials weniger Fehler gemacht wurden als mit papierbasierten Tutorials und zum Teil höhere Lerneffekte erzielt werden konnten (van der Meij & van der Meij 2014: 157). Zu einem gegenteiligen Ergebnis kam Maestre (2012). Seine Proband\*innen bevorzugten zudem mit großer Mehrheit eine statische Variante des Tutorials, da hier die Struktur des Tutorials einfacher erfassbar war und so schneller die richtigen Handlungsschritte gefunden werden konnten (vgl. Maestre 2012: 258). Ebenso konnte Lang (2016) keine Vorteile von Videotutorials gegenüber textbasierten Tutorials nachweisen. Seine Ergebnisse beziehen sich sowohl auf die Bearbeitungszeit, Effektivität, Motivation und wahrgenommene Schwierigkeit des Tutorials (vgl. Lang 2016: 33). Generelle Vorteile von textbasierten Tutorials betreffen darüber hinaus die Struktur und das damit verbundene einfachere Auffinden von spezifischen Aufgaben durch einfaches Überfliegen der Inhalte. Auch die individuell festlegbare Bearbeitungsgeschwindigkeit zeigt

sich von Vorteil. Des Weiteren laden textbasierte Tutorials zum aktiveren Handeln ein, da das Gelesene vor der Ausführung erst selbst interpretiert und verstanden werden muss (vgl. van der Meij & van der Meij 2014: 151).

Unabhängig von der medialen Ausgabeart des Software-Tutorials ist es jedoch, dass in den meisten Software-Tutorials Abbildungen der Benutzeroberfläche enthalten sind. Sie sollen die Ausführung der Handlungsanweisungen visualisieren und damit erleichtern. In den folgenden Unterkapiteln wird genauer auf verschiedene Abbildungsarten in Software-Tutorials eingegangen werden. Dabei werden zuerst Screenshots und ihre Auszeichnung durch visuelle Signalisierung näher beleuchtet. Im Anschluss folgt eine nähere Betrachtung des Forschungsstandes zu Simplified User Interfaces sowie dem Umgang mit visuellen Details in Dokumentationen.

### **2.3.2 Screenshots**

Die am meisten verwendete Abbildungsart in Software-Dokumentationen der letzten Jahrzehnte sind sogenannte Screenshots (vgl. Meng 2019: 397, vgl. van der Meij & Gellevij 1998: 529). Der Begriff Screenshot bezeichnet laut Duden (2022) eine „fotoähnliche Abbildung oder Speicherung dessen, was auf dem Bildschirm zu sehen ist“. Dabei kann sowohl der gesamte Bildschirm als auch nur ein bestimmtes Fenster oder Ausschnitt gemeint sein. Zudem kann mit einem Screenshot sowohl der Ausgangszustand, Zwischenzustände oder das Gesamtergebnis eines Prozesses dargestellt werden. Unabhängig ihres Formates oder Inhaltes können Screenshots grundsätzlich in Sekundenschnelle mit einem Druck auf der Tastatur oder mit einfach bedienbarer Software erstellt werden (vgl. Farkas 1999: 42 f., vgl. van der Meij & Gellevij 2004, vgl. Meng & Ullmann 2015: 41).

Screenshots haben vier Funktionen, die Nutzer\*innen dabei unterstützen können, eine Software erfolgreich zu bedienen. Zum einen helfen Screenshots dabei, den Wechsel der Aufmerksamkeit zwischen Bildschirm, Software-Dokumentation und der Benutzung des Eingabegerätes zu koordinieren. Screenshots helfen dabei dem *Split-Attention-Effekt* (siehe Kapitel 2.1.1) entgegenzuwirken (vgl. van der Meij 2000: 295). So legen sie den optimalen Zeitpunkt fest, wann auf den Bildschirm geschaut werden soll und bieten gleichzeitig Orientierung, wenn der Blick sich zurück auf die Software-Dokumentation richtet. Doch nicht nur hierfür bietet ein Screenshot Orientierung. Ebenso hilft er dabei relevante Objekte (z. B. Schaltflächen oder bestimmte Menüeinträge) leichter zu identifizieren. Besonders komplexe Software, wie z. B. Bildbearbeitungs- oder Animationsprogramme, verwirren Nutzer\*innen oft zu Beginn der Nutzung mit einer Unzahl von Symbolen und Fenstern. Mithilfe eines Screenshots kann hier ein genauer Abgleich

getroffen werden, wo sich das gesuchte Element befindet, welches angeklickt werden soll (vgl. Gellevij & van der Meij 2004: 225). Genauso kann mit dem Screenshot schließlich das Endergebnis eines Handlungsschrittes überprüft werden. Es muss keine Transferleistung erbracht werden, geschriebenen Text aus der Software-Dokumentation mit dem visuellen Ergebnis am Bildschirm abzugleichen, stattdessen kann schlichtweg der Screenshot verglichen werden. (vgl. Gellevij & van der Meij 2004: 232, vgl. Horton 1993: 146 ff.). Eine vierte, wichtige Funktion eines Screenshots ist die Erleichterung, zu verstehen, wie das Programm funktioniert und aufgebaut ist, also ein mentales Modell der zu benutzenden Software zu entwickeln. Mentale Modelle spielen sowohl bei der Identifizierung eine wichtige Rolle als auch bei der Lösung von Aufgaben ebenso wie im Umgang mit Transferaufgaben (vgl. Gellevij & van der Meij 2004: 225).

Wie effektiv der Einsatz von Screenshots in der Software-Dokumentation ist, haben bereits eine Reihe an Studien untersucht, angefangen mit einer Studie aus dem Jahr 1996 von Hans van der Meij. Dabei sollten die Proband\*innen Aufgaben mithilfe einer Software-Dokumentation lösen. Die Hälfte der Proband\*innen hatte hierfür eine rein textbasierte Dokumentation zur Verfügung, die andere Hälfte konnte auf zusätzliche Screenshots in ihren jeweiligen Dokumentationen zurückgreifen. Das Ergebnis der Studie besagt, dass ein signifikanter Zeitunterschied in der Bearbeitungsdauer festgestellt werden konnte, jedoch keine Unterschiede in der Fehleranzahl während der Bearbeitung (vgl. van der Meij 1996: 371 ff.). Auch Studien von Gellevij und van der Meij (2004) sowie Sweller und Chandler (1994) konnten den positiven Effekt von Screenshots bestätigen. In einer ähnlichen Studie kommen Gellevij und Kolleg\*innen zusätzlich zu den Ergebnissen, dass Screenshots die Bildung eines mentalen Modells beim Lernen mit komplexen Anleitungen positiv unterstützen und dabei helfen, relevante Bildelemente schneller zu identifizieren (vgl. Gellevij et al. 2002: 215). Sweller und Chandler (1994) legen die verstärkte Komplexität von Software-Dokumentationen auch als Randbedingung für positive Effekte durch Screenshots fest. Bei einfachen Aufgabenstellungen zeigen sich demnach keine positiven Effekte. Weitere Randbedingungen liegen in einer eindeutig erkennbaren Funktion, also der Relevanz des Screenshots im Kontext der Dokumentation, sowie dem darauf angepassten Design des Screenshots hinsichtlich Ausschnitt, Größe und Signalisierung (vgl. Meng & Ullmann 2015: 42). Generell scheinen sich also viele Forschungsergebnisse zu lernförderlichen Abbildungen (siehe Kapitel 2.2) auch auf Screenshots anwenden zu lassen.

Hinsichtlich der optimalen Gestaltung von Screenshots können vier Parameter angeführt werden, in denen sich Screenshots in Software-Dokumentationen unterscheiden. Der abgebildete Bereich der Software-Oberfläche, die Position des Screenshots im Dokument,

die Größe des Screenshots sowie der Einsatz von Signalisierung (vgl. van der Meij & Gellevij 1998: 529, vgl. van der Meij 2000: 296). Über alle vier Parameter muss sich bei der Konzeption einer Software-Dokumentation oder eines Tutorials Gedanken gemacht werden, da sie die effektive Nutzung des Dokuments beeinflussen können. Der abgebildete Bereich der Software-Oberfläche stellt entweder den gesamten Bildschirm oder nur ein Fenster oder gar einen kleineren Ausschnitt der Oberfläche dar. Die Abbildung des gesamten Bildschirms bietet einen Überblick über die gesamte Software-Oberfläche und erleichtert damit den Nutzer\*innen Zusammenhänge zwischen mehreren Bildbereichen zu erkennen, um ein mentales Modell zu entwickeln und sich in der Software zu orientieren. Kleinere Ausschnitte können dahingegen dazu dienen, eine bestimmte Funktion der Software zu identifizieren, kleinteilige Bereiche der Software erkennbar zu machen und sich zu fokussieren. Unabdingbar ist es jedoch, den Ausschnitt groß genug zu wählen, damit die Nutzer\*innen den relevanten Bereich in der Software lokalisieren können (vgl. Meng & Ullmann 2015: 44). Horton (1993) weist daraufhin, dass nur der jeweils relevante Teilbereich einer Bildschirmoberfläche im Screenshot abgebildet werden sollte, damit die Nutzer\*innen nicht mit irrelevanten Informationen kognitiv belastet werden (vgl. Horton 1993: 148). Eine Studie von Gellevij und Kollegen hat jedoch empirisch untersucht, wie sich unterschiedliche Bildschirmausschnitte von Screenshots auf die Effektivität der Anleitung auswirken und kam zu einem gegenteiligen Ergebnis. Hierfür haben sie drei inhaltlich identische Anleitungen getestet, die sich lediglich hinsichtlich der darin enthaltenen Abbildungsarten unterschieden: Vollbild-Screenshots, Screenshots aus relevanten Teilbereichen sowie eine Anleitung ohne Screenshots. Die rein textuelle Anleitung führte zu ähnlichen Ergebnissen, wie die Variante mit Vollbild-Screenshots. Beide Varianten zeigten sich jedoch effektiver als die Teilausschnitt-Variante. Alle drei Bedingungen zeigten keine Unterschiede bezüglich der Motivation der Nutzer\*innen bei deren Nutzung (vgl. Gellevij et al. 1999: 86).

Die Position des Screenshots innerhalb der Anleitung kann sich ebenfalls auf die effektive Nutzung der Anleitung auswirken. Nach Mayers *räumlichem Kontiguitätsprinzip* lernen Menschen besser, wenn zusammengehörige Texte und Abbildungen nahe zusammenstehen und nicht entfernt voneinander (vgl. Mayer 2021: 207). Auf diese Weise müssen keine kognitiven Kapazitäten dafür aufgebraucht werden, um zusammengehörige Inhalte einander zuzuordnen. Zudem können nahe zusammenstehende Inhalte auch leichter im Arbeitsgedächtnis behalten werden. Nachgewiesen wurde dieser Effekt durch diverse Experimente von Mayer und Kolleg\*innen. In einem Transfertest zur Funktionsweise von Autobremsen schnitten demnach Proband\*innen sehr viel besser ab, wenn Diagramme und erklärender Text im Lernmaterial direkt nebeneinander anstatt auf

verschiedenen Seiten platziert waren (vgl. Mayer 1989: 240 ff.). Ein weiterer relevanter Parameter bezieht sich auf die Größe des abgebildeten Screenshots im Dokument. Van der Meij (2000) bezieht sich hier auf eine ältere Quelle, welche darauf hinweist, dass ein Screenshot im Dokument 50-75 % der realen Bildschirmgröße betragen sollte. Im Hinblick auf die heutigen Bildschirmgrößen lässt sich diese Aussage jedoch eher nicht auf die aktuelle Zeit übertragen. Vielmehr sollte sich die Größe des abgebildeten Screenshots danach richten, dass alle relevanten Elemente gut lesbar und lokalisierbar sind. Auch sollte die Auflösung des angefertigten Screenshots im Blick behalten werden. Ein Screenshot besteht normalerweise aus Text und grafischen Elementen, die sich auf dem Bildschirm wiederum aus Reihen von einzelnen Pixeln zusammensetzen. Hierdurch ergeben sich einige Herausforderungen bei dem Einsatz in Software-Dokumentationen oder Tutorials. Zum einen müssen auch kleinere Elemente und besonders der Text deutlich erkennbar sein. Dies kann besonders bei starker Vergrößerung bzw. Verkleinerung des angefertigten Screenshots problematisch werden, ebenso wie bei dem Einsatz von Screenshots in einem responsiven Design auf mobilen Endgeräten. Darüber hinaus ist es aus Gründen der Barrierefreiheit notwendig, wenn Texte – auch in Abbildungen – vorlesbar sind. Kurzfristige Änderungen am abgebildeten Inhalt führen zu der zusätzlichen Anforderung, dass der zu ändernde Screenshot im Dokument leicht auffindbar und austauschbar ist. Ebenso werden Software-Dokumentationen häufig in verschiedenen Sprachen erstellt, weshalb die Screenshots ebenso angepasst werden müssen (vgl. Mitschke 2017: 46 f.). Hinzu kommt der Einsatz von Dokumentationen für verschiedene Betriebssysteme oder Änderungen in der Software-Benutzeroberfläche nach Updates oder bereits während der Software-Entwicklung. All diese Änderungen haben zum Teil aufwändige Organisationsprozesse zur Folge, damit die Screenshots stets den aktuellen und richtigen Systemzustand abbilden, wofür sie unter Umständen häufig neu erstellt und ausgetauscht werden müssen.

Der letzte wichtige Punkt bei der Einbindung von Screenshots in eine Software-Dokumentation betrifft die Entscheidung über den Einsatz von Signalisierungen zur Hervorhebung relevanter Bereiche.

### **2.3.3 Visuelle Signalisierung**

Das in Kapitel 2.2.3. erwähnte *Signalisierungsprinzip* der CTML besagt, dass Menschen besser lernen, wenn die wesentlichen Strukturen des Lernmaterials hervorgehoben werden. Somit wird die Aufmerksamkeit auf die relevanten Bereiche gelenkt und extrinsische Belastung verringert (vgl. Mayer 2021: 167, vgl. Alpizar et al. 2020: 2097). Die Grundlagen zu dem Prinzip und empirische Beweise von Mayer und Kolleg\*innen wurden

im genannten Kapitel bereits diskutiert und sollen daher an dieser Stelle nicht wiederholt werden.

Dieses Prinzip wird häufig auch auf Screenshots in Software-Tutorials angewendet, um die Nutzer\*innen dabei zu unterstützen, die für den auszuführenden Handlungsschritt relevanten Elemente einfacher identifizieren zu können. Dies geschieht durch das Einfügen von Pfeilen, Rahmen, farblichen Hervorhebungen oder Beschriftungen. Wichtig ist hierbei zu beachten, dass zwar Details durch die Auszeichnung hinzugefügt werden, jedoch keine zusätzlichen Screenshot-Inhalte. In der Praxis bedeuten Signalisierungen für Technische Redakteur\*innen jedoch auch zusätzliche Kosten und zusätzlichen Zeitaufwand. Es muss nicht mehr nur entschieden werden an welcher Stelle ein Screenshot sinnvoll ist und welcher Ausschnitt jeweils gewählt werden soll. Stattdessen muss jeder Screenshot, der signalisiert werden soll, einzeln bearbeitet werden (vgl. Meng 2019: 397). Während es durch Mayer und Kolleg\*innen viele Hinweise für positive Effekte durch Signalisierungen in Lerntexten gibt, ist die Forschungslage bei handlungsorientierten Texten, wie Tutorials, gering.

Hinweise auf einen positiven Effekt gibt es jedoch durch die Forschungsergebnisse eines Eyetracking-Experiments von Meng (2019). Untersucht wurden zwei Tutorial-Versionen, welche inhaltlich identisch waren, sich aber hinsichtlich ihrer Abbildungsarten unterschieden. Eine Version enthielt reine Screenshots, während in der anderen Variante signalisierte Screenshots verwendet wurden. Die Ergebnisse zeigen einen positiven Effekt von signalisierten Screenshots gegenüber reinen Screenshots bezüglich der während der Bearbeitung des Tutorials gemachten Fehler. Die Bearbeitungszeit unterschied sich jedoch nicht signifikant. Zudem gibt es durch die Forschungsergebnisse Hinweise darauf, dass auch bei handlungsorientierten Texten Signalisierungen positiv dazu beitragen, relevante Bildelemente zu identifizieren und die Aufmerksamkeit der Nutzer\*innen effektiv zu lenken (vgl. Meng 2019: 396).

### **2.3.4 Simplified User Interfaces**

Ein Simplified User Interface (SUI) ist eine Designtechnik, welche auf herkömmliche Screenshots angewandt werden kann. Die Technik tauchte erstmals 2013 vor allem in der Marketingkommunikation von großen Techunternehmen auf, hat mittlerweile aber auch Einzug in die Technische Kommunikation erhalten. Für eine Simplified User Interface Grafik wird ein Screenshot einer Softwareanwendung oder Website verwendet und so stark reduziert, dass nur noch wesentliche Inhalte übrigbleiben. Irrelevante Texte, Symbole oder Elemente des Screenshot-Inhaltes werden dabei entweder komplett entfernt oder stark vereinfacht, z. B. als Rechteck, dargestellt (siehe Abb.2). Lediglich Inhalte, die der

Orientierung dienen oder für den an dieser Stelle explizit auszuführenden Arbeitsschritt notwendig sind, bleiben bestehen (vgl. Bollen & Saremba 2021, 2020a: 13 f.).

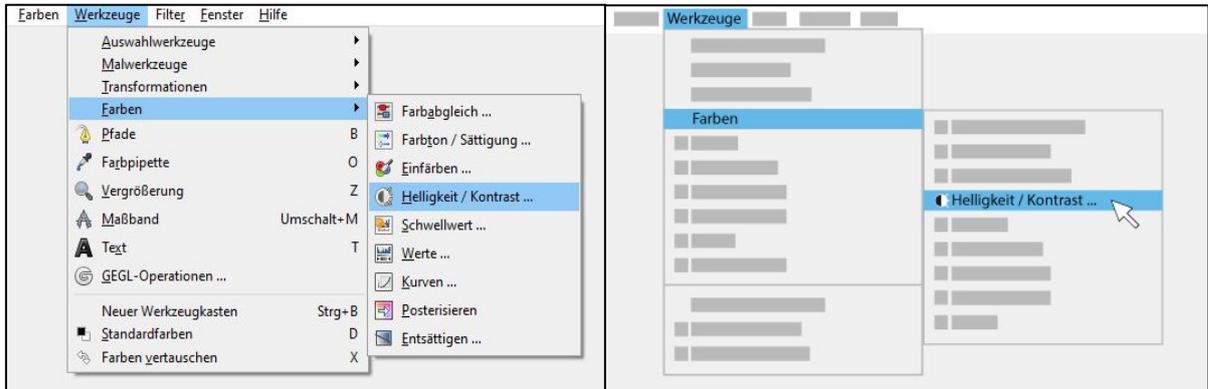


Abb. 2: Ein reiner Screenshot als Grundlage (links), sowie das daraus erstellte Simplified User Interface (rechts) (Quelle: Screenshot aus GIMP (links) sowie Eigene Darstellung (rechts)).

Durch ihren reduzierten Inhalt soll die Aufmerksamkeit der Nutzer\*innen gezielt auf die wesentlichen Informationen gelenkt werden, alles andere kann ausgeblendet werden. Bollen und Saremba (2020a: 14) sprechen SUI-Grafiken aufgrund dieser Aufmerksamkeitslenkung und Detailreduktion das Verursachen einer geringeren kognitiven Belastung im Vergleich zu unbearbeiteten Screenshots zu. SUIs sollen demnach zu einer leichteren Orientierung beitragen und damit einfacher interpretiert werden können. Dies untermauern sie mit einem Predictive Eyetracking-Test, welchen sie anhand einiger beispielhafter Screenshots durchgeführt haben. Predictive Eyetracking bezeichnet hierbei eine auf künstlicher Intelligenz basierende Alternative zu herkömmlichen Eyetracking-Studien. Anstatt Proband\*innen-Tests durchzuführen, werden hier zuvor gesammelte Eyetracking-Daten genutzt, um das menschliche Sehen zu simulieren (vgl. Slivka 2020).

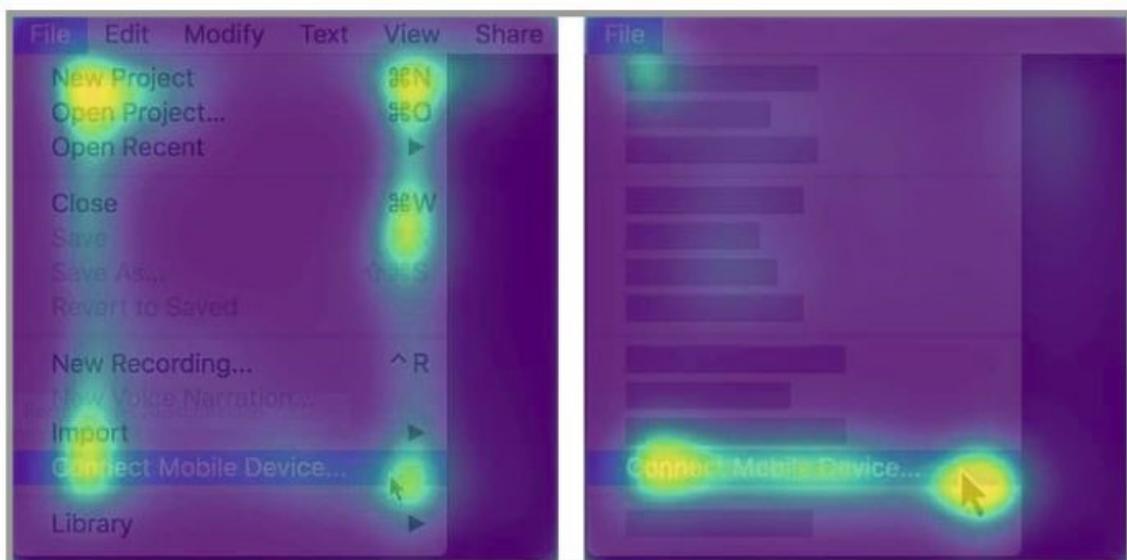


Abb. 3: Heatmap von einem herkömmlichen Screenshot (links) und einer SUI-Grafik (rechts). (Quelle: Bollen & Saremba 2020: 13.)

Das Ergebnis der Tests zeigt bei der SUI-Grafik eine deutliche, direkte Fokussierung auf dem relevanten Bereich, während der Blick bei einem herkömmlichen Screenshot hin- und herwandert und mehrere, auch irrelevante, Bereiche fokussiert hat (siehe Abb. 3).

Abgesehen von diesem kurzen Test gibt es bisher keine weiteren Studien oder Proband\*innen-Tests zum Thema SUI-Grafiken in der Technischen Kommunikation. Speziell zum Thema der Effektivität von SUI-Grafiken gegenüber herkömmlichen Screenshots besteht noch ein großes Forschungsdefizit (vgl. Bollen & Saremba 2020b), welchem mit dieser Arbeit entgegengewirkt werden soll. An dieser Stelle ist darauf hinzuweisen, dass die Gesamtanzahl der verfügbaren Quellen zum Thema SUIs sehr begrenzt ist und von Mitarbeitenden der Firma TechSmith dominiert wird. Sie stellt somit keine vollends wissenschaftlich fundierte, durch Studien belegte Literaturgrundlage dar.

Trotz der geringen Forschungs- und Quellenlage werden Simplified User Interfaces eine Reihe von Vorteilen gegenüber herkömmlichen Screenshots zugesprochen. Neben der bereits genannten geringeren kognitiven Belastung wird vor allem die Langlebigkeit und Lokalisierbarkeit hervorgehoben. Hinzu kommen Vorteile im Bereich des Datenschutzes (vgl. Bollen & Saremba 2020a: 14, vgl. Bollen 2019, vgl. Boatman 2020).

Die Langlebigkeit von SUI-Grafiken erklärt sich in ihrer Wiederverwendbarkeit im Gegensatz zu einem Screenshot. Eine der größten Herausforderungen in der Technischen Kommunikation ist es, den Inhalt stets aktuell zu halten. Während des Softwareentwicklungsprozesses, aber auch bei Updates, fallen häufig Änderungen an der Softwareoberfläche an. Einzelne Elemente werden verschoben, hinzugefügt oder entfernt. Möglicherweise wird die Bezeichnung einer Schaltfläche geändert oder anders angeordnet. Bei der Verwendung von herkömmlichen Screenshots in der Software-Dokumentation muss an dieser Stelle prinzipiell bei jeder Anpassung ein neuer Screenshot angefertigt werden. Nur so kann sichergestellt werden, dass dieser die Software exakt abbildet und die Nutzer\*innen nicht verwirrt. Mit dem Einsatz von SUI-Grafiken können diese Schritte den Technischen Redakteur\*innen häufig erspart bleiben. Da irrelevante Elemente bereits entfernt oder auf eine einfache Form reduziert wurden, müssen in diesen Bereichen keine Änderungen vorgenommen werden, wenn sich der Text, ein Symbol oder die Position eines jenen ändert (vgl. Boatman 2020).

Bei der Verwendung einer Software-Dokumentation für verschiedene Betriebssysteme kann ein SUI aufgrund der Vereinfachung mehrfach verwendet werden, anstatt in jedem System einen neuen Screenshot anfertigen zu müssen (vgl. Bollen & Saremba 2020: 14). Software-Dokumentationen werden aber nicht nur oft für verschiedene Betriebssysteme erstellt, sondern auch in verschiedenen Sprachen. Laut einer Umfrage unter mehr als 600

Mitarbeitenden aus dem Bereich der Technischen Kommunikation gaben 61 % der Befragten an, dass sie ihren Content übersetzen würden. Dabei übersetzen 37 % in zwei bis fünf Sprachen, 34 % sogar in mehr als 25 Sprachen (vgl. Abel 2019: 49).

Schon bei dem Anfertigen einer Software-Dokumentation in nur wenigen unterschiedlichen Sprachen führen Screenshots zu einem großen Zeit-, Organisations- und Verwaltungsaufwand. Die Grafiken müssen in allen Sprachen angefertigt, verwaltet und stets aktuell gehalten werden. Hier zeigt sich ein SUI ebenfalls als vorteilhaft, da ein Text in SUIs größtenteils komplett entfällt. So kann die Lokalisierung der Abbildung stark vereinfacht werden oder kann vollständig entfallen. SUIs können so in vielen Sprachversionen der Dokumentation wiederverwendet werden. Wenn überhaupt notwendig muss bei Softwareaktualisierungen somit nur eine Grafikversion aktualisiert werden. Dies hat wiederum eine große Zeitersparnis, weniger Organisationsaufwand und Kosten zur Folge (vgl. Boatman 2020, vgl. Bollen 2019, vgl. Bollen & Saremba 2020: 14).

Ein weiterer Vorteil der Reduzierung und Vereinfachung von Screenshot-Inhalten findet sich bei dem Thema Datenschutz. Häufig befinden sich persönliche Daten oder Benutzerkonten in dem abgebildeten Screenshot. Dafür müssen Beispiel-Accounts mit erfundenen Daten angelegt werden, um keine Datenschutzrechte zu verletzen. Dies stellt sich jedoch oft – je nach Software – als zeitaufwändig heraus, da viele Daten erfunden werden müssen. Eine Alternative liegt darin, die Daten im herkömmlichen Screenshot unkenntlich zu machen. Jedoch wirkt dieses häufig ablenkend oder gar unprofessionell. Ein SUI bietet hier den Vorteil, dass persönliche Daten vollständig entfernt oder genauso visuell reduziert werden wie andere irrelevante Details. Hierdurch entsteht neben der Zeitersparnis, da keine Daten erfunden werden müssen, vor allem auch ein insgesamt ästhetischeres Gesamtbild (vgl. Bollen 2019).

Einen wesentlichen Nachteil stellt dahingegen die aufwändigere Erstellung eines Simplified User Interfaces dar. Während ein Screenshot meist mit einem Mausklick und ohne notwendige Vorkenntnisse erstellt werden kann, stellt sich der Erstellungsprozess eines SUI's zum Teil langwieriger und komplizierter dar. Von der Firma TechSmith wurde eine kommerzielle Software entwickelt, welche Screenshots und Bildschirmvideos anfertigen und bearbeiten kann (siehe TechSmith o. D.). Nach der Anfertigung eines Screenshots lässt sich mit dieser Software mit wenigen Mausklicks ein individuelles Simplified User Interface aus dem Screenshot erstellen. Farben, Formen und Reduzierungsgrad können dabei individuell angepasst werden. Dafür sind nur geringe Vorkenntnisse notwendig. Eigene Tests der Autorin dieser Arbeit haben jedoch gezeigt, dass ein zufriedenstellendes Ergebnis stark vom Komplexitätsgrad des Screenshots abhängig ist. Wenig komplexe

Screenshots, wie die einer Website, welche lediglich aus einer überschaubaren Anzahl an Text- und Bildelementen bestehen, können vergleichsweise gut automatisch von der Software in ein SUI umgewandelt werden. Je detaillierter der Screenshot aufgebaut ist, beispielsweise das Interface einer komplexen Bildbearbeitungssoftware, desto weniger zufriedenstellend ist oft das Ergebnis. Eine intensive und individuelle Nachbearbeitung ist hier häufig notwendig.

Neben der genannten Software lassen sich SUIs auch mit jeder herkömmlichen Grafik- oder Bildbearbeitungssoftware von Hand erstellen. Hierunter fallen professionelle Software wie Adobe Illustrator, Adobe Photoshop oder CorelDraw. Die Verwendung der genannten Software erfordert dabei jedoch grafische Vorkenntnisse. Aber auch Standard-Software wie Microsoft PowerPoint oder jede andere Software, mit der individuelle Formen und Texte erstellt werden können, kann prinzipiell verwendet werden. Ein Vorteil von einem Nachbau der Screenshots in einem Grafikprogramm ist, dass das SUI nun vektorbasiert und damit beliebig skalierbar ist. Dies kommt dem Einsatz einer digitalen Software-Dokumentation auf verschiedenen Endgeräten zugute. Bei einem herkömmlichen, pixelbasierten Screenshot sind die Technischen Redakteur\*innen an die bei der Screenshot-Erstellung eingestellte Auflösung gebunden. Außerdem lassen sich die vektorbasierten Formen bei Bedarf animieren, sodass GIFs oder ganze Videos aus dem SUI erstellt werden können, welche die Prozesse anschaulich visualisieren (vgl. Stegelin 2018: ab 20:00 min).

Unabhängig davon mit welcher Software gearbeitet wird, zur Erstellung eines Simplified User Interfaces wird zunächst immer ein Screenshot als Grundlage benötigt. Oft ist es sinnvoll, diesen zu Beginn auf den notwendigen Ausschnitt zuzuschneiden. Hierdurch entfallen unnötige und ablenkende Informationen und die auszuführende Aktion wird in den Mittelpunkt gerückt (vgl. Bollen & Saremba 2020a: 14). Dieser zugeschnittene Screenshot wird um alle irrelevanten Informationen reduziert. Alle für die Orientierung und das Verständnis unnötigen Grafiken, Icons und Bilder können entfernt werden. In einem nächsten Schritt werden irrelevante Texte durch einfache Formen wie ein Rechteck überdeckt. Dies kann auch auf grafische Elemente angewendet werden, sofern diese inhaltlich irrelevant sind, aber ihre Position der Orientierung auf dem Screenshot dient (vgl. Bollen 2019). Bei der Erstellung der Formen muss die grundlegende Hierarchie des Screenshot-Layouts wieder aufgegriffen werden. Beispielsweise wird für eine Überschrift ein breiterer Kasten verwendet als für einen untergeordneten Text. Auch empfiehlt es sich, gleichwertige Texte oder Abbildungen mit den immer gleich gestalteten Formen zu überdecken, um die Orientierung zu erleichtern (vgl. Bollen & Saremba 2020a: 16). Darüber hinaus ist es wichtig, den grundlegenden Aufbau des Screenshots nicht zu

verändern. Das Layout der abgebildeten Software oder Website sollte auch mit der Reduzierung noch wiedererkennbar sein (vgl. Bollen 2019).

Bei der Gestaltung der überdeckenden Formen kann sich zu Beginn ein einfaches Farbschema, bestehend aus wenigen Farben überlegt werden. Hier empfiehlt es sich, ein oder zwei schlichte Farben wie Grautöne für den Hintergrund zu verwenden, wobei sich hier am originalen Screenshot-Hintergrund orientiert werden sollte. Für den Vordergrund kann eine geringe Anzahl kontrastreicher Farben verwendet werden, welche ebenfalls an das Original angepasst gewählt wird (vgl. Bollen & Saremba 2021 / 2020: 16). Zur Erleichterung der Arbeit können, je nach verwendeter Erstellungssoftware, einzelne Bausteine, Farbpaletten oder Vorlagen zur Wiederverwendung abgespeichert werden. So muss nicht bei jedem zu erstellenden SUI von Neuem begonnen werden. Ebenso kann es hilfreich sein, einen Styleguide zu erstellen, in dem die wichtigsten Gestaltungsrichtlinien festgelegt werden. Zum Beispiel Abstände, Farbschemata, Strichstärken, welche Form für welches Element verwendet wird und vieles weitere (Bollen & Saremba 2020: 16).

Neben der Wiederverwendbarkeit von SUI-Elementen, welche die Erstellung künftiger SUIs verkürzen, bringt Bollen (2020: ab 19:00 min) hervor, dass ein SUI nicht jeden Screenshot in einer Software-Dokumentation ersetzen sollte. Stattdessen können SUIs selektiv an bestimmten Stellen eingesetzt werden, wo sie Sinn ergeben. Eine grundlegende Rolle dabei die oben genannten Vorteile von SUIs spielen. Werden Abbildungen in vielen verschiedenen Sprachen oder Update-Versionen benötigt oder beinhalten sie persönliche Daten, welche geschützt werden müssen? Ist dies nicht der Fall muss individuell abgewogen werden, ob sich eine Erstellung lohnt oder nicht. Hierzu bedarf es jedoch weiterer Forschung.

Generell sind zum Einsatz von Simplified User Interfaces in Software-Dokumentationen oder Tutorials noch viele Fragen offen. Für welche Einsatzzwecke und Ausgabemedien sind SUIs möglicherweise besser oder schlechter geeignet? Gibt es durch den Einsatz von SUIs einen nachhaltigeren Lerneffekt? Offen bleiben Fragen, inwieweit ein SUI reduziert werden kann, beispielsweise bis hin zu einem kompletten Verzicht von Text in einem SUI, welches eine Lokalisierung komplett hinfällig machen würde. Oder welche Rolle überhaupt die Reduzierung von Detailinformationen in einem SUI auf den Lernprozess spielt. Wird die kognitive Belastung der Betrachtenden wirklich reduziert, oder wirkt ein SUI aufgrund fehlender Informationen verwirrend und erschwert den Bearbeitungsprozess? SUIs zeichnen sich auf den ersten Blick durch viele Vorteile aus. Die grundlegende Frage ist jedoch, wie effizient sie im Vergleich zu einfacher zu erstellenden Screenshots sind und wie effektiv Lernende damit in der Praxis umgehen können.

## 2.4 Hypothesen

Die zuvor vorgestellten theoretischen Grundlagen haben gezeigt, dass Abbildungen in Lernmaterialien und Software-Dokumentationen bzw. Tutorials grundsätzlich als effektiv angesehen werden können. Einflussfaktoren und Randbedingungen, wie das Vorwissen der Nutzer\*innen, die Relevanz der jeweiligen Abbildung, mögliche Auszeichnungen durch Signalisierungen sowie der Komplexitätsgrad der Abbildungen, können sich entscheidend auf die Effektivität auswirken.

Die optimale Gestaltung eines Screenshots ist bisher jedoch wenig erforscht. Einerseits können sich zusätzliche, relevante Details, wie Signalisierungen, zwar positiv auf den Lerneffekt auswirken, gleichzeitig kann eine Reduktion von unnötigen Details dies jedoch unter Umständen ebenfalls. Simplified User Interfaces stellen eine im Detailgrad stark reduzierte Form der Screenshots dar, die eine geringere kognitive Belastung des Arbeitsgedächtnisses zur Folge haben soll (vgl. Bollen & Saremba 2020a: 14). Neben den vielen bereits vorgestellten praktischen Vorteilen durch ihren Einsatz in Dokumentationen für Technische Redakteur\*innen bleiben bisher empirische Beweise für die positive Lerneffekte für die Nutzer\*innen aus. Die vorliegende Arbeit soll einen ersten Beitrag dazu leisten, dies zu untersuchen. Hierfür sollen folgende Forschungsfragen untersucht werden:

F01: Wie effizient sind Simplified User Interfaces in Software-Tutorials im Vergleich zu herkömmlichen Abbildungsarten?

F02: Hat die Reduktion von Detailinformationen bei SUIs einen zusätzlichen, positiven Effekt?

Da in der vorhandenen Literatur zum Thema Simplified User Interfaces ein positiver Effekt von dieser Abbildungsart suggeriert wird, werden folgende Hypothesen zur Untersuchung der Forschungsfragen aufgestellt:

H01: SUI-Grafiken erleichtern die Identifikation relevanter Bildelemente im Vergleich zu anderen Abbildungsarten.

H02: Die Bearbeitung eines Software-Tutorials mit SUI-Grafiken benötigt weniger Zeit als eine solche mit herkömmlichen Abbildungsarten.

H03: Bei der Bearbeitung eines Software-Tutorials mit SUI-Grafiken kommt es zu weniger Fehlern als bei der Bearbeitung eines Tutorials mit herkömmlichen Abbildungsarten.

### **3 Methodisches Vorgehen**

Die Untersuchung der oben genannten Hypothesen und Fragestellungen sollte durch ein Eyetracking-Experiment mit Proband\*innen stattfinden. Hierzu wurde eine Studie konzipiert, in der drei Software-Tutorials miteinander verglichen wurden. Inhaltlich identisch unterscheiden sich die drei Versionen lediglich hinsichtlich der darin enthaltenen Abbildungsarten: unbearbeitete Screenshots, signalisierte Screenshots und Simplified User Interface. Ziel war es herauszufinden, mit welcher Variante die Proband\*innen am effizientesten arbeiten können.

Die folgenden Kapitel bieten zu Beginn einen Einblick in die für die Studie verwendete technische Methode des Eyetrackings. Im Anschluss wird die eigentliche Studie genauer vorgestellt. Das Studiendesign sowie das konzipierte Material werden präsentiert und dabei ein Einblick in das allgemeine Vorgehen bei der Konzipierung, Durchführung und Auswertung der Studie gegeben.

#### **3.1 Methode: Eyetracking**

Viele Studien zum multimedialen Lernen greifen auf die Methodik zurück, den Proband\*innen Lernmaterial zu präsentieren und im Anschluss den Lerneffekt durch eine Wissensabfrage oder Transferaufgaben zu messen (siehe z. B. Mayer & Gallini 1990, Mayer et al. 1996). Dieses Forschungsdesign zeigt sich zwar in den jeweiligen Studien als nützlich, für die Untersuchung eines Software-Tutorials ist es jedoch weniger geeignet. Denn das Lesen des anleitenden Textes ist mit einer direkten Handlungsausführung verbunden. Duggan und Payne (2001: 297) zeigen darüber hinaus auf, dass sich besser an gelernte Instruktionsanweisungen erinnert werden kann, wenn deren Ausführung zeitgleich zum Lesen stattfindet. Eine alternative Methodik, bei der diese zeitgleiche Handlungs- und Lesefreiheit gegeben ist, stellt die in der vorliegenden Studie verwendete Technik des Eyetrackings dar. Dabei werden die Blickbewegungen der Proband\*innen aufgezeichnet, während sie individuell zwischen Tutorial und Softwareoberfläche hin- und herwechseln können. Im Folgenden wird die Methode des Eyetrackings erläutert.

Unter Eyetracking wird eine apparative Methode verstanden, mit der Blickbewegungen und -positionen von Proband\*innen bei der Betrachtung von Stimuli aufgezeichnet werden. Dabei messen die Eyetracking-Systeme die genaue Position der Reflexion von Infrarotlicht auf der Hornhaut der Proband\*innen relativ zur Pupille. Stimuli können dabei sowohl digitale wie analoge Medien, aber auch die Umwelt darstellen. Durch dieses Verfahren kann analysiert werden, welche Bereiche des Stimulus eine Person wie häufig und lange

anschaut und in welcher Reihenfolge diese betrachtet werden (vgl. Blake 2013: 367; vgl. Carter & Luke 2020: 50).

Der technologische Fortschritt hat es vor allem in den letzten 20 Jahren möglich gemacht, die Hard- und Software für Eyetracking-Systeme zu einem kommerziellen Produkt zu entwickeln, welches sich in der Wissenschaft bewährt und immer häufiger genutzt wird. Es kommt häufig in Studien aus Disziplinen zum Einsatz. Beispielsweise Neurowissenschaften und Psychologie, UX-Design und Usability, Übersetzung, Marketing, Informatik, (Psycho-)Linguistik, Bildung, Ingenieurwesen, Leseforschung und vielen weitere (vgl. Duchowski 2007, vgl. Carter & Luke 2020: 50, vgl. Holmqvist & Andersson 2017: 3).

Entwickelt wurde die Methode und erste Technik jedoch schon vor Jahrzehnten. Noch viel weiter reicht die Forschung zu den theoretischen Grundlagen zurück. Aristoteles beschäftigte sich bereits in der Antike mit der visuellen Wahrnehmung und den dazugehörigen Blickbewegungen des Menschen. Im Mittelalter folgten erste medizinische Studien hierzu (vgl. Geise 2011: 154). Ab wann genau konkrete Forschungen zu Blickbewegungen stattfand, ist heute schwer zu belegen. Als wichtige Persönlichkeiten in diesem Zusammenhang können Johannes Müller (1801-1858) und Charles Bell (1774-1842) angesehen werden. Bell (1823) untersuchte beispielsweise in einem Experiment die aktiven und passiven Augenbewegungen und deren Auswirkungen auf die Blickrichtung (vgl. Wade & Tatler 2005: 33). Im darauffolgenden Jahrhundert standen vorrangig die Grundlagen der Augenbewegungsanalyse im Vordergrund der Forschung (vgl. Rayner 1998: 372). Hierbei kamen jedoch noch stark invasive Methoden zum Einsatz, die die Wahrnehmung der Proband\*innen massiv eingeschränkt und negativ beeinflusst haben. Delabarre (1898) führte eine Selbststudie durch, in der er einen Draht als Blickrichtungsanzeiger nutzte, der mit Gips auf seiner Hornhaut befestigt war. In den folgenden Jahren wurden schließlich weniger invasive Verfahren entwickelt, welche vermehrt auf Foto- und Filmtechnik setzten (vgl. Dodge & Cline 1901).

Seit den 1930er Jahren lässt sich eine stärkere Anwendungsorientierung in der Entwicklungsgeschichte des Eyetrackings feststellen. Ab etwa 1970 folgte dann eine Zeit, in der die bestehenden Systeme hinsichtlich der Genauigkeit, Methodik und Bedienbarkeit entscheidend weiterentwickelt und verbessert wurden. Die ersten videobasierten Eyetracking-Systeme wurden konstruiert, welche an einen Computer angeschlossen werden konnten. So können nicht nur herkömmliche analoge Medien, sondern ebenso digitale Medien untersucht werden. Viel entscheidender ist an dieser Stelle die Möglichkeit, die erhobenen Daten direkt aufzeichnen und sehr viel einfacher analysieren zu können als zuvor (vgl. Rayner 1998: 372, vgl. Geise 2011: 156). Wo zuvor noch Proband\*innen direkt

beobachtet und ihr Verhalten beschrieben werden musste, konnte nun eine einfache Synchronisation der Blickbewegungen mit den Stimuli am Computer durchgeführt und in Echtzeit verfolgt werden. Dies hat neben der Zeit- und Kostenersparnis vor allem deutliche Verbesserungen in der Benutzerfreundlichkeit sowohl für die Proband\*innen als auch für die Forscher\*innen zur Folge und ermöglicht die Entwicklung vieler neuer Forschungsmöglichkeiten. Dies führte zu einem massiven Anstieg der Verwendung von Eyetracking-Systemen in der Wissenschaft (vgl. Carter & Luke 2020: 50, vgl. Geise 2011: 156).

Trotz der neuen Möglichkeiten und Anwendungsfelder hat die Methodik des Eyetrackings besonders in medien- und kommunikationswissenschaftlichen Forschungsfeldern bisher nur bedingt Einfluss. Ausgenommen sind Forschungsfelder wie (Web-)Usability, E-Learning und die Übersetzung. Auch die Technische Kommunikation ist nach bisherigem Stand nur wenig geprägt von Eyetracking-Studien (vgl. Geise 2011: 158, vgl. Meng 2020: 107). Als veröffentlichte Studien können die Untersuchungen von Balling (2013, 2018) und Meng (2019) genannt werden. Während Balling sich vorrangig mit der Wirkung von Textoptimierungen beim Leseprozess beschäftigt, legt Meng (2019) den Fokus seiner Untersuchungen auf die Effekte von visueller Signalisierung von Screenshots (siehe Kapitel 2.3.3).

Eine Herausforderung im Einsatz von Eyetracking-Systemen in der Technischen Kommunikation liegt darin, dass Nutzer\*innen Technische Dokumentationen eher handlungs als lernorientiert lesen (siehe Kapitel 2.3.1, vgl. Redish 1989). Dementsprechend sollte auch der Versuchsaufbau und das Studiendesign erfolgen. Beispielsweise sollte die Versuchsperson selbst entscheiden können, wann sie was lesen und wann sie handeln möchte. Eine Einschränkung in der Bewegungsfreiheit der Proband\*innen durch ein hoch präzises Eyetracking-System, welches eine Stabilisierung des Kopfes der Proband\*in benötigt, kann hier in Konflikt zur freien Handlungsfähigkeit stehen. Zudem muss der Versuchsaufbau so konstruiert sein, dass sowohl die Stimuli (die Technische Dokumentation) als auch die Soft- bzw. Hardware parallel und individuell von der Versuchsperson benutzt werden kann (vgl. Meng 2020: 108f.). Wenn Versuchsaufbau und Studiendesign und das Eyetracking-System gut aufeinander abgestimmt werden, bietet die Methode jedoch vielfältige Möglichkeiten, Forschungen zu effizienteren Technischen Dokumentationen voranzubringen.

Um die Funktionsweise eines Eyetracking-Systems zu verstehen, wird ein grober Überblick über die visuelle Wahrnehmung des Menschen mit seinen Augen und entsprechenden Blickbewegungen benötigt.

### 3.1.1 Visuelle Wahrnehmung und Blickbewegungen

Das menschliche Auge funktioniert, in dem Licht durch die Pupille aufgenommen und mithilfe der Hornhaut und Linse so auf die Netzhaut fokussiert wird, dass ein auf dem Kopf stehendes Bild erzeugt wird. Die nur 1,5 mm breite Fovea Centralis in der Netzhautmitte ermöglicht es Details und Farben zu erfassen. Die so erfassten visuellen Informationen werden über den Sehnerv weitergeleitet. Daraus erfolgt die Zusammensetzung zu einem Bild und dessen Interpretation im Gehirn. Die geringe Größe der Fovea führt dazu, dass Blickbewegungen nötig sind, um genügend visuelle Informationen aufnehmen zu können (vgl. Carter & Luke 2020: 51).

„We have the subjective impression of an immediate, full detail, pictorial view of the world. We are prone to forget that this impression is, in a very real sense, an illusion.“ (Findlay & Gilchrist 2003: 2). Dieses Zitat verdeutlicht, dass der Mensch der Meinung ist, eine durchgehende und vollständige Wahrnehmung seiner Umgebung zu haben, dies aber aus physiologischer Sicht nicht der Fall ist. Stattdessen setzt sich der Blickverlauf aus einzelnen, schnell hintereinander folgenden Fixationen zusammen. Die Fixation beschreibt die Zeitspanne, in der das auf einen bestimmten Punkt fixierte Auge nahezu stillsteht. In dieser Zeit nimmt es visuelle Informationen auf, bevor es schließlich zum nächsten Fixationspunkt weiterspringt. Die Übergänge zwischen den einzelnen Fixationen werden als Sakkaden bezeichnet (vgl. Rayner 2009: 1458, vgl. Rötting 2001: 68). Hinzu kommen ungerichtete Mikro- und Anpassungsbewegungen, welche eine Stabilisierung des Blicks in der Bewegung ermöglichen (vgl. Geise 2011: 167 f.).

Die Fixationen gehören zu den wichtigsten Parametern, welche in der Analyse von Eyetracking-Aufzeichnungen ausgewertet werden. Die Mikrobewegungen und Sakkaden werden bei der Blickbewegungserfassung durch Eyetracking dahingegen nicht aufgezeichnet. Letztere können jedoch rekonstruiert werden. (vgl. Geise 2011: 168, 203). Visuell darstellen lässt sich der Blickverlauf einer Person durch sogenannte Scan Paths (siehe Kapitel 3.1.3.).

Die Länge einer Fixation wird in der Literatur unterschiedlich angegeben. Sie bewegt sich jedoch meist zwischen mindestens 100 ms und 300 ms. Sie variiert abhängig vom Stimulus und individuellen Faktoren wie der Konzentrationsfähigkeit der Person oder Komplexität der Aufgabe (vgl. Rayner 2009: 1459, vgl. Rötting 2001, S. 16 f.). Die Sakkaden variieren dahingegen zwischen nur etwa 30 bis 50 ms. Während dieser kurzen Zeitspanne werden keine visuellen Informationen aufgenommen. Der Mensch ist quasi blind. Dieses Phänomen wird auch als Sakkadenblindheit bezeichnet. Trotzdem finden zu dieser Zeit meist kognitive Prozesse statt (vgl. Rayner 2009: 1458, vgl. Carter & Luke 2020: 51).

### **3.1.2 Eyetracking-Systeme und ihre Funktionsweise**

Es gibt mittlerweile verschiedene Eyetracking-Systeme, welche meist videobasiert arbeiten. Videobasiert bedeutet an dieser Stelle, dass der corneale Reflex und die Pupillenposition durch eine Kamera aufgezeichnet und in Beziehung zueinander gesetzt werden (vgl. Blake 2013: 371). Hierunter zählen zum einen das Remote Eyetracking, ein stationäres System ohne direkten Kontakt zwischen Versuchsperson und Gerät. Zum anderen sind mobile Eyetracking-Systeme auf dem Markt, welche wie eine Brille oder eine Kopfbedeckung von der Versuchsperson aufgesetzt werden können. Für besonders präzise Messungen gibt es Eyetracking-Systeme, bei denen der Kopf der Proband\*innen durch eine Kinnauflage stabilisiert werden muss, damit möglichst viele Bewegungen vermieden werden können. Die schnellsten kommerziellen Eyetracker weisen eine Abtastfrequenz von bis 2000 Hz auf, es wird also 2000 mal pro Sekunde die genaue Blickposition der Versuchsperson aufgezeichnet. Mobile Eyetracker sind hier mit einer Frequenz von etwa 50 Hz wesentlich langsamer. Je mobiler ein Eyetracker einsetzbar ist, desto weniger präzise ist er zudem meist (vgl. Carter & Luke 2020: 51). Darüber hinaus gibt es mittlerweile die Möglichkeit des webbasierten Eyetrackings, bei dem die Webcam eines Computers als Eyetracker verwendet wird.

Da in der dieser Arbeit zugrunde liegenden Studie das Remote Eyetracking angewendet wurde, wird im Folgenden zunächst hierauf eingegangen und die Funktionsweise eines Eyetrackers anhand dessen erläutert. Der Vollständigkeit halber werden im Anschluss auch das mobile und das webbasierte Eyetracking grundlegend vorgestellt.

#### Remote Eyetracking

Remote Eyetracking bezeichnet eine non-invasive, passive Methode des Eyetrackings. Dabei berührt die Versuchsperson das Eyetracking-Gerät selbst nicht. Stattdessen ist das System in denselben Computerbildschirm integriert oder daran montiert, auf welchem die Stimuli für das Experiment betrachtet werden. Vorteile dieser Methode sind meist eine relativ große Bewegungsfreiheit der Proband\*innen während der Durchführung des Experiments. Zudem können auch Proband\*innen mit korrigiertem Sehvermögen durch eine Brille oder Kontaktlinsen teilnehmen (vgl. Blake 2013: 273 f.).

Die Versuchsperson wird auf einem Stuhl in etwa 60 cm Entfernung vom Computerbildschirm platziert. Um die Messgenauigkeit nicht zu beeinträchtigen, sollten unnötige Bewegungen vermieden werden und daher möglichst kein Drehstuhl oder Stuhl mit Rollen verwendet werden. Zu Beginn des Experiments wird der Eyetracker kalibriert. Dies muss für jede Versuchsperson individuell durchgeführt werden, da der menschliche Augapfel von Person zu Person physiologische Unterschiede aufweist und der Eyetracker

darauf eingestellt werden muss. Besonders die Fovea, der Teil des Auges, der für das zentrale Sehen verantwortlich ist, variiert in seiner Position stark und beeinflusst damit die Messgenauigkeit des Eyetrackers. Ebenso können die Hornhautkrümmung und Pupillenposition variieren (vgl. Lindén 2021: 28).

Für die Versuchsperson bedeutet dies, dass sie zur individuellen Kalibrierung einen sich über den Computerbildschirm bewegendem Punkt fixiert und mit ihren Blickbewegungen verfolgen muss. Der Eyetracker misst dabei per Infrarotlicht den cornealen Reflex (d. h. die Spiegelung einer Lichtquelle auf der Hornhaut) der Proband\*innen. Dieser ist relativ zur Position des Mittelpunkts der Pupille (vgl. Hammoud 2008: 3, vgl. Duchowski 2007: 54 ff.). Gleichzeitig werden die x- und y-Koordinaten des zu fixierenden Punktes aufgezeichnet und schließlich in Relation zur Position des cornealen Reflexes und der des Mittelpunktes der Pupille gesetzt (vgl. Blake 2013: 372). Anhand dieser Daten kann dann rekonstruiert werden, welche Bereiche des Monitors zu welchem Zeitpunkt fixiert wurden. Damit kann der individuelle Blickverlauf einer Person rekonstruiert und vorhergesagt werden (vgl. Geise 2011: 183, vgl. Lindén 2021: 28).

Die Aufzeichnung der Daten erfolgt computergestützt über verschiedene Software zur Steuerung des Eyetrackers, zur Konzeption von Experimenten, Kalibrierung, Aufzeichnung und Analyse der erhobenen Daten. Die Eyetracking-Software zeichnet dabei ein schwarz-weißes Video auf, welches zur Analyse dient. Schwarz dargestellt wird dabei die Pupille der Versuchsperson, weiß der corneale Reflex. Dies wird auf die aufgezeichnete Blickposition auf dem Stimulusmaterial übertragen und in Fixationen und Sakkaden übersetzt, so dass der Blickverlauf rekonstruiert werden kann. (vgl. Geise 2013: 185). Dabei enthält jeder Eyetracking-Datensatz die rekonstruierte Position des Blickpunktes für ein oder beide Augen sowie die x- und y-Koordinaten der entsprechenden Position auf dem Stimulusmaterial (vgl. Carter & Luke 2020: 53).

Die Messgenauigkeit von Remote Eyetracking-Geräten ist relativ hoch. Von Herstellern wird sie durchschnittlich mit  $0,4^\circ - 0,5^\circ$  Abweichung von der wirklichen Blickposition angegeben. Dies bedeutet bei einem herkömmlichen Versuchsaufbau, bei dem eine Versuchsperson etwa 60 cm vom Bildschirm entfernt sitzt, einen Unterschied von 0,5 cm zwischen der tatsächlichen und gemessenen Blickpositionen (vgl. Blake 2013: 374).

Die Auswertung von stationär gewonnenen Eyetracking-Daten ist vergleichsweise einfach. Mithilfe der passenden Software werden Blickbewegungen und Stimuli synchronisiert und können so auf vielfältige Art und Weise grafisch und statistisch ausgewertet werden (vgl. Geise 2011: 180, vgl. Blake 2013: 375).

### Mobiles Eyetracking

Mobile Eyetracking-Systeme haben im Vergleich zu der stationären Variante einen direkten Kontakt zum Körper der Versuchsperson. War die Benutzerfreundlichkeit invasiver Eyetracker in der Vergangenheit noch sehr unbefriedigend, hat die technologische Entwicklung hier zu einer komfortableren Verwendungsmöglichkeit beigetragen. Mobile Eyetracker gibt es in verschiedenen Formen. Zum einen wurden Eyetracking-Brillen entwickelt, zum anderen gibt es Systeme die wie ein Helm auf dem Kopf getragen werden können (vgl. Geise 2011: 180). Unabhängig von der äußeren Form der Apparatur, sind an ihr jeweils zwei Kameras montiert. Eine zeichnet das Blickfeld der Versuchsperson in der Umgebung auf, die andere Kamera ist auf das Auge der Person gerichtet. So können die genauen Blickbewegungen aufgezeichnet und mit der Aufzeichnung des Blickfeldes synchronisiert werden (vgl. Hammoud 2008, vgl. Duchowski 2007).

Im Unterschied zu stationären Eyetrackern eröffnen mobile Eyetracker vollkommen andere Möglichkeiten. Während der Versuchsaufbau mit dem stationären System an ein Labor bzw. einen festen Schreibtisch mit Computer gebunden ist, können sich Versuchspersonen mit einem mobilen Eyetracker frei in ihrer Umgebung bewegen. Völlig andere Studenumgebungen und -kontexte sind so möglich zu untersuchen, z. B. die Blickbewegungen bei einem Spaziergang durch die Stadt oder beim Autofahren. Auch können beispielsweise analoge Medien, welche eine Interaktion wie das Umblättern benötigen, untersucht werden.

Schwieriger ist dahingegen die Auswertung der auf diese Art und Weise erhobenen Daten. Es liegen keine festen Stimuli vor, welche von der Versuchsperson betrachtet wurden. Stattdessen gibt es Videoaufzeichnungen mit von Proband\*in zu Proband\*in unterschiedlichen Sichtfeldern und ihren darin jeweiligen Blickverläufen. Eine Synchronisation ist hier nur nach einer aufwändigen manuellen Codierung möglich (vgl. Blake 2013: 375).

### Webbasiertes Eyetracking

Neben den stationären und mobilen Eyetracking-Systemen lässt die moderne Technik es mittlerweile zu, onlinebasierte Eyetracking-Studien durchzuführen. Den Hintergrund bilden hierfür maschinelle Lernalgorithmen, die es einer handelsüblichen Webcam ermöglichen als Eyetracker eingesetzt zu werden. Diese übernimmt die Kalibrierung und Aufzeichnung der Blickbewegungen in Relation zum Stimulus. Ein Onlineportal übernimmt dann die Aufzeichnung und Synchronisation der Daten (vgl. Woytal 2018: 50).

Klare Vorteile dieses Verfahrens sind die vereinfachten Datenerhebungsmöglichkeiten. Eyetracking-Studien können mit diesem System unabhängig von Ort und Zeit durchgeführt

werden. Zugleich ermöglichen sie somit eine wesentlich größere Proband\*innenanzahl und sind lokal unabhängig vom eigentlichen Forschungsort. Die Proband\*innen können unter realistischen Studienbedingungen direkt von zu Hause teilnehmen und mit ihren vertrauten Geräten in gewohnter Umgebung arbeiten. Hier kommt auch eine große Kostenersparnis zum Tragen, da nicht in teure Eyetracking-Systeme investiert und ein Labor zur Verfügung stehen muss. Die Aufzeichnung durch ein Onlineportal macht es zudem möglich, mehrere Proband\*innen zur gleichen Zeit das Experiment bearbeiten zu lassen. Nachteilig ist die vergleichsweise niedrige Präzision der Webcam-Aufzeichnung im Vergleich zu Eyetracking-Geräten. Zudem können nur statische Stimuli auf einem Bildschirm präsentiert werden. Außerdem findet in der Regel kein Kontakt zwischen Proband\*innen und Forscher\*innen statt. Es kann bei Problemen nicht eingegriffen oder die korrekte Ausführung überwacht werden. Aufgrund der zum Teil schwerwiegenden Nachteile bietet das Webcam-basierte Eyetracking zwar eine interessante Ergänzung zum stationären oder mobilen. Eine Studie mit hochwertigen Eyetracking-Systemen kann sie jedoch nicht vollständig ersetzen (vgl. Woytal 2018: 50 f., vgl. Semmelmann & Weigelt 2018: 462).

### **3.1.3 Messwerte und Analysemöglichkeiten**

Durch Eyetracking kann eine Vielzahl an Messwerten erhoben und ausgewertet werden. Hier muss für jedes Experiment individuell entschieden werden, welche als relevant betrachtet werden können. An dieser Stelle soll lediglich eine begrenzte, für diese Arbeit relevante Auswahl vorgestellt werden. Für eine umfassende Übersicht über weitere Messwerte und Analysemöglichkeiten, siehe beispielsweise Holmqvist & Anderson (2017) oder Blascheck et al. (2014).

Eyetracking-Daten können zum einen statistisch und damit quantitativ ausgewertet werden. Zum anderen bieten Visualisierungstechniken das Potential einer qualitativen und explorativen Analyse. Durch die Visualisierung können die Beziehungen zwischen den räumlichen und zeitlichen Daten hervorgehoben werden und dafür genutzt werden, Hypothesen aufzustellen, welche statistisch überprüft werden können (vgl. Blascheck et al. 2014: 63, vgl. Blake 2013: 377).

Eine visuelle Auswertung kann zum Beispiel über *Scanpaths*, *Heat-Maps* oder *Focus-Maps* (auch *Lumination Maps* genannt) stattfinden. Bei allen drei Methoden liegt das zuvor aufgezeichnete Video des Stimulusmaterials zugrunde. Auf diesem wird bei dem *Scanpath* der Blickverlauf der Versuchsperson abgebildet. Je länger ein Bereich fixiert wurde, desto größer wird auf dieser Position ein Kreis dargestellt. Die einzelnen Kreise sind durch Linien miteinander verbunden, welche die Sakkaden visualisieren. Zeitlich analog zum Videoverlauf des Stimulus werden die Fixationen und Sakkaden so auf dem Stimulus

abgebildet, so dass der genaue Blickverlauf der Proband\*innen nachvollzogen werden kann (vgl. Blake 2013: 377). *Focus Maps* und *Heat Maps* folgen einem ähnlichen Prinzip. Beide stellen die Intensität dar, mit der bestimmte Bereiche des Stimulus fixiert werden (siehe Abb. 4). Bei der *Focus Map* werden dabei nur die Bereiche des Stimulus sichtbar dargestellt, auf denen der Fokus des Betrachtenden gelegen hat. Alle nicht betrachteten Bereiche werden abgedunkelt. Auf einer *Heat Map* wird die visuelle Darstellung entgegengesetzt ausgeführt. Hier ist der gesamte Stimulus sichtbar. Je länger und öfter ein Bereich auf dem Stimulus fixiert wurde, desto rötlicher ist dieser Bereich markiert (vgl. Blake 2013: 377).

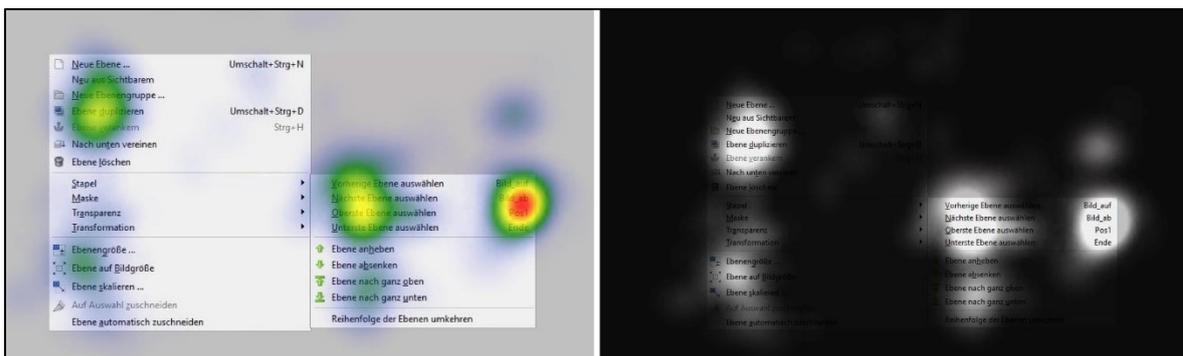


Abb. 4: Links: Heat Map Visualisierung eines Screenshots. Rechts: Identischer Stimulus als Focus Map Darstellung (Quelle: Eigene Darstellung aus Screenshot von GIMP.)

*Areas of Interest (AOI)* bieten vielfältiges Potenzial für eine weiterführende statistische Auswertung. Sie beschreiben die für den Forschenden interessanten Bereiche eines Stimulus. Diese werden nach der Datenerhebung über eine Analysesoftware auf den Stimuli eingezeichnet und somit definiert. Im Anschluss können statistische Daten für exakt diese Bereiche des Stimulus ausgegeben und statistisch analysiert werden. Beispielsweise die *Fixation Count*, *Revisits* oder die *Dwell Time* (vgl. Blascheck et al. 2014: 65).

Die *Fixationsanzahl (Fixation Count)* beschreibt, wie oft ein bestimmter Bereich von den Proband\*innen mit den Augen fixiert wurde. Wird ein Bereich häufig fixiert, kann es einerseits auf ein erhöhtes Interesse an diesem Bereich hinweisen. Andererseits kann es auch darauf hindeuten, dass an dieser Stelle ein hoher Komplexitätsgrad vorliegt und etwas nicht gleich verstanden wird (vgl. Geise 2011: 201). Die *Dwell Time (Verweildauer)* setzt sich aus allen Fixationen und Sakkaden zusammen und beschreibt die gesamte Verweilzeit, die in einem bestimmten Bereich verbracht wurde (vgl. SMI 2014: 230). Ebenso wie die *Fixationsanzahl* kann hier eine hohe *Verweildauer* sowohl auf verstärktes Interesse als auch auf eine erhöhte Schwierigkeit hindeuten (vgl. Geise 2011: 201).

In dieser Arbeit werden unter anderem auch die *Revisits* analysiert. Wie der Name bereits sagt, gibt dieser Parameter an, wie oft der Blick der Versuchsperson auf einen bestimmten Bereich, eine AOI, nach dem ersten Verlassen des Bereichs wieder zu ihm zurückkehrt.

### **3.2 Studienüberblick**

Da die Studie den Einsatz von Grafiken in Software-Tutorials untersuchen soll, wurde zuerst ein Software-Tutorial konzipiert und gestaltet. Nach der allgemeinen Konzeption des Tutorials wurde dieses in drei unterschiedliche Varianten umgewandelt (für eine genauere Beschreibung siehe auch Kapitel 3.3.4):

1. Tutorial mit Simplified User Interfaces als Grafiken (SUI)
2. Tutorial mit signalisierten Screenshots als Grafiken (SIG)
3. Tutorial mit herkömmlichen Screenshots (ohne jede weitere Bearbeitung) als Grafiken (SCR)

Im Anschluss erfolgte die Konzeption des Eyetracking-Experimentes im SMI ExperimentCenter und das damit verbundene Anlegen der Stimuli (siehe 3.3 Studiendesign). Um die Verständlichkeit und Korrektheit der Aufgaben und Abläufe zu gewährleisten, wurde nun folgend ein Pretest mit einem Testprobanden durchgeführt. Nach kleineren Anpassungen am Experiment und dem Tutorial wurden schließlich Proband\*innen gesucht.

Die Durchführung der Studie erfolgte im Usability Labor der Hochschule Merseburg. Zu vereinbarten Terminen erschienen die Proband\*innen einzeln im Labor. Zuerst wurde der allgemeine Ablauf der Studie erklärt, Hinweise zum Verhalten bei einer Eyetracking-Aufzeichnung gegeben und mögliche Fragen beantwortet. Hierbei wurde darauf geachtet, nicht näher auf thematische Inhalte oder Ziele bzw. Fragestellungen der Studie einzugehen, um die Proband\*innen nicht zu beeinflussen. Daraufhin folgte die Unterzeichnung der Datenschutz- und Einwilligungserklärung zur Teilnahme.

Im Anschluss begann die eigentliche Studie (siehe Kapitel 3.3), in der die Proband\*innen die gestellten Aufgaben am Computer bearbeiteten. Dabei erfolgte eine Eyetracking-Aufzeichnung sowie ein Screen Recording. Am Ende des Experiments erhielt jede\*r Proband\*in eine Aufwandsentschädigung. Nach Abschluss aller Datenaufzeichnungen erfolgte die statistische Auswertung der erhobenen Daten (siehe Kapitel 3.4).

#### **3.2.1 Proband\*innenauswahl und -zuordnung**

Als Proband\*innen sollten aktuell und ehemalige Studierende dienen. Aufgrund des begrenzten zeitlichen Umfangs der Masterarbeit und damit dieser Studie, musste eine

realistisch umsetzbare Stichprobenanzahl gewählt werden. Diese wurde auf 24 festgelegt, welches acht Proband\*innen pro Gruppe entspricht.

Die Proband\*innen wurden aus dem privaten, sowie über den E-Mailverteiler der Hochschule Merseburg, aus dem hochschulinternen Umfeld eingeladen. Aufgrund ihres akademischen Hintergrunds konnte dabei davon ausgegangen werden, dass die Proband\*innen zumindest grundlegend gute Vorerfahrung im Umgang mit Computern haben. Dies soll zusätzliche, technische Schwierigkeiten während der Durchführung der Studie vermeiden. Außerdem wurden mindestens gute Deutschkenntnisse vorausgesetzt, welche es erlauben, auf Deutsch verfasste Anweisungen problemlos verstehen zu können. Ebenso war ein normales oder korrigiertes Sehvermögen notwendig, um an der Studie teilnehmen zu können. Der weitere persönliche Hintergrund sowie die Vorerfahrungen der Proband\*innen mit Bildbearbeitung im Allgemeinen und dem in der Studie verwendeten Bildbearbeitungsprogramm GIMP im Speziellen, werden in Kapitel 4.1 näher beleuchtet.

Da in dem Experiment drei verschiedene Grafikvarianten gegeneinander getestet werden sollten, wurden die Proband\*innen (für sie unwissentlich) vor Studienbeginn jeweils einer der drei Gruppen zugeordnet. Dabei entschied die zufällige Reihenfolge der Teilnahme über die Gruppenzuordnung. Proband\*in 1 erhielt alle Stimuli im Experiment mit Simplified User Interfaces als Grafiken, Proband\*in 2 alle Stimuli mit signalisierten Screenshots, Proband\*in 3 alle Stimuli mit herkömmlichen Screenshots ohne weitere Bearbeitung. Danach begann die Reihenfolge immer wieder von vorn. Alle Texte, Aufgabenstellungen und sonstige Inhalte waren in allen drei Testvarianten identisch.

### **3.2.2 Versuchsaufbau und Technik**

Während der Studie saßen die Proband\*innen an einem Schreibtisch sowie auf einem stabilen Stuhl ohne Rollen, um unnötige Bewegungen zu vermeiden. Der Raum wurde halb abgedunkelt, so dass kein Sonnenlicht direkt auf den Monitor fallen konnte. Außerdem wurde dafür gesorgt, dass das Experiment möglichst nicht durch laute Umgebungsgeräusche oder andere Personen gestört werden konnte.

Alle Stimuli wurden von den Proband\*innen von dem 24-Zoll-Monitor aus etwa 60 cm Entfernung rezipiert. Am unteren Bildschirmrand wurde der Eyetracker montiert. Bei dem verwendeten Eyetracker handelt es sich um das RED 250mobile Eyetracking-System der Firma SMI. Dies besteht aus einem stationären Remote Eyetracker (siehe Kapitel 3.1.2 und 3.2.2) und der dazu passenden Software (siehe Tab. 3).

Während die empirischen Fragen, einleitende Texte sowie der Orientierungstest im Vollbildmodus durchgeführt wurden, war für den Hauptteil eine Teilung des Monitors

notwendig. Wie bereits dargestellt, sollte sich bei Eyetracking-Studien in der Technischen Kommunikation an einem zum reading-to-learn-to-do passenden Arbeiten der Versuchspersonen orientiert werden. Die Proband\*innen sollten frei entscheiden können, wann, wie lange und wie oft sie das Tutorial lesen und zu welchem Zeitpunkt im Grafikprogramm sie die entsprechende Aktion ausführen. Aus diesem Grund mussten Tutorial und Programm parallel individuell nutzbar sein. Außerdem sollte eine Teilung der Aufmerksamkeit (Split-Attention Effekt, siehe Kapitel 2.1.1) vermieden werden, weshalb sich dafür entschieden wurde, Tutorial und Software auf demselben Monitor darzustellen. Auf der linken Seite des Monitors wurde auf etwa 2/3 der Fläche das Bildbearbeitungsprogramm angezeigt. Im rechten Drittel wurde das Tutorial in der Windows-Fotoanzeige geöffnet (siehe Abb. 5). Zur Betrachtung des Software-Tutorials wurde dieses Programm gewählt, da es eine einfache vor und zurück Navigation anbietet und leicht bedienbar ist. Die Taskleiste am Proband\*innen-Monitor wurde für die Durchführung der Studie ausgeblendet, um unnötige Aktionen zu vermeiden.

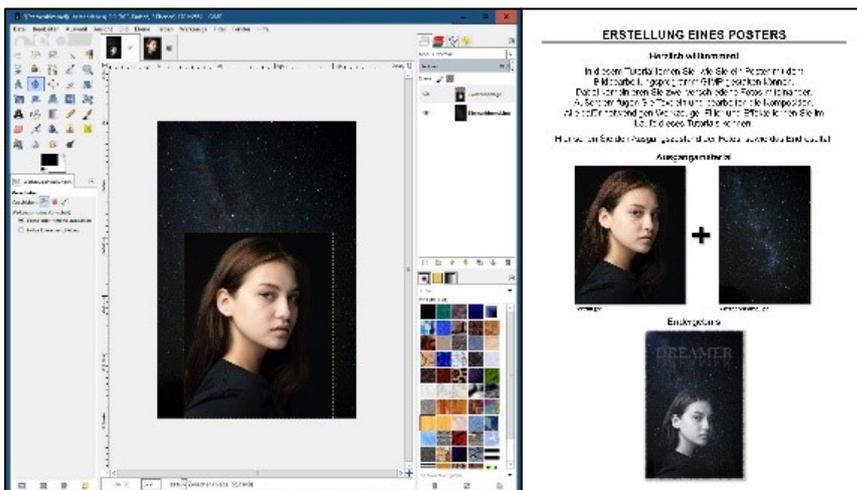


Abb. 5: Bildschirmordnung während der Bearbeitung des Software-Tutorials. (Quelle: Eigene Darstellung.)

Neben dem Eyetracking-System wurde auch eine Reihe von Softwareanwendungen im Rahmen dieser Studie verwendet. Im Folgenden wird eine genaue Auflistung dieser aufgeführt (siehe Tab. 3).

Software	Beschreibung / Verwendungszweck
GIMP 2.8.18	Open-Source-Grafikprogramm zur Bearbeitung des Tutorials
Adobe Illustrator CS6	Vektorbasiertes Grafikprogramm; zur Erstellung der SUI-Grafiken verwendet
iViewRED, iViewETG	Treibersoftware für Kommunikation zwischen Eyetracking-Hardware & Computer
SMI ExperimentCenter 3.7.68	Software zur Konzeption und Durchführung von Eyetracking-Experimenten; Datenaufzeichnung für das RED System
SMI BeGaze 3.7	Analysesoftware zur Auswertung von Eyetracking-Daten
RStudio 2022.07.2	Grafische Benutzeroberfläche für die Statistik-Programmiersprache R; zur Auswertung der erhobenen Daten verwendet

Tab. 3: Für die Studie verwendete Software. (Quelle: Eigene Darstellung.)

### 3.3 Studiendesign und Material

Der Aufbau der Studie zur Untersuchung der Effizienz von Simplified User Interfaces in Software-Tutorials gliedert sich in vier Teile (siehe Abb. 6).

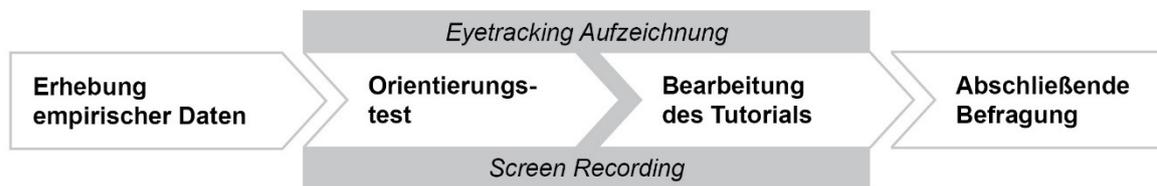


Abb. 6: Schematischer Studienaufbau. (Quelle: Eigene Darstellung.)

Das Experiment begann mit der Erhebung einiger empirischer Daten zu den Proband\*innen. Dabei wurden allen Proband\*innen – unabhängig von ihrer Gruppe – die identischen Fragestellungen präsentiert (siehe Kapitel 3.3.1). Während der Beantwortung der Fragen wurden noch keine Eyetracking-Daten aufgezeichnet und kein Screen Recording durchgeführt, da dies beides nicht notwendig für die spätere Auswertung ist. Da die Eyetracking-Aufzeichnung zudem mit einer gewissen Bewegungseinschränkung der Proband\*innen verbunden ist, sollte zudem möglichst lange darauf verzichtet werden. Des Weiteren eignen sich leicht zu beantwortende Fragen, wo nichts „falsch“ gemacht werden kann gut für eine entspannte Eingewöhnung der Proband\*innen an das Studiensetting.

Im Anschluss begann die eigentliche Eyetracking-Aufzeichnung. Hierzu wurde vorab der Eyetracker kalibriert und validiert (siehe Kapitel 3.3.2).

#### 3.3.1 Erhebung empirischer Daten

Zu Beginn der Studie sollten alle Proband\*innen identische empirische Fragen beantworten. Die Fragen wurden dabei als Stimuli im ExperimentCenter angelegt und den Proband\*innen auf dem Computerbildschirm präsentiert. Dies erleichtert die spätere Auswertung, da die Antworten der Proband\*innen im Anschluss über das ExperimentCenter exportiert werden konnten, um sie im Anschluss statistisch auszuwerten.

Frage eins bezog sich auf das Alter der Proband\*innen. Bei der Frage „Wie alt sind Sie?“ sollte die Altersspanne angegeben werden, in der sich jede\*r Proband\*in befindet. Die Entscheidungsmöglichkeiten lagen hier bei <20, 21-30, 31-40, 41-50, 51-60 oder >60. Aus logischen Gründen konnte dabei jeweils nur eine Antwort ausgewählt werden. Da zur Studie aktuell Studierende bzw. ehemalige Studierende / aktuell Berufstätige eingeladen wurden, konnte auf eine genauere Einteilung bei Minderjährigen und Senior\*innen verzichtet werden. Durch die Frage sollte nach Möglichkeit später analysiert werden, ob es

Unterschiede im Umgang mit den verschiedenen Grafikarten bei Personen in unterschiedlichen Altersspannen gibt.

Bei Frage zwei wurde gefragt, welchem Geschlecht sich die Proband\*innen zugehörig fühlt. Zur Auswahl standen: weiblich, männlich und/oder divers. An dieser Stelle war es möglich mehrere Antworten zu wählen, um allen Möglichkeiten der empfundenen Geschlechtsidentität gerecht zu werden.

Frage drei lautete: „Was machen Sie hauptberuflich?“ Antwortmöglichkeiten waren: Student\*in, Mitarbeiter\*in/Dozent\*in/Professor\*in der HoMe [Hochschule Merseburg] oder ein Freitextfeld, in dem individuelle Antworten gegeben werden konnten. Spezifiziert wurden die Antworten dann in Frage vier. Diese lautete: „Welchen Studiengang studieren Sie?“. Als Antwort konnte eine individuelle Eingabe gemacht werden oder angegeben werden, dass nicht studiert wird. Es wurden keine Antwortmöglichkeiten im Sinne von spezifischen Studiengängen der Hochschule Merseburg vorgegeben, da die Einladung auch an Studierende außerhalb der Hochschule Merseburg verschickt wurde und somit nicht alle Möglichkeiten im Vorhinein erfasst werden konnten. Da Berufstätige bereits bei Frage drei ihren Beruf angegeben haben, mussten sie in Frage vier ihre Antwort nicht weiter spezifizieren. Mit den Fragen drei und vier sollte dabei der fachliche Hintergrund der Proband\*innen erfasst werden, um diesen später analysieren und miteinander vergleichen zu können. Dieser kann ggfs. Rückschlüsse auf die Erfahrung mit Tutorials oder Eyetracking geben. Ebenso kann so herausgefiltert werden, ob Proband\*innen aus demselben Studiengang wie die Studienleiterin kommen und somit möglicherweise Erfahrung mit Usability Studien besteht.

Frage fünf und sechs bezogen sich auf die Vorerfahrung der Proband\*innen. In Frage fünf sollte die eigene Vorerfahrung mit Bildbearbeitungssoftware allgemein eingeschätzt werden. Wie die Proband\*innen ihre Vorerfahrung einschätzen, konnte dabei aus den Antwortmöglichkeiten einer Likert Skala (Sehr gut, Gut, Durchschnittlich, Nicht gut, Gar keine Erfahrung) angegeben werden. Dieselbe Likert Skala wurde bei Frage sechs verwendet. Diese Frage bezog sich speziell auf die Vorerfahrung mit dem Bildbearbeitungsprogramm GIMP. Beide Fragen sind relevant für eine spätere Auswertung der Eyetracking-Daten. Die Vorerfahrung mit Bildbearbeitung und speziell mit der in der Studie verwendeten Software GIMP beeinflusst unter Umständen massiv die Bearbeitungszeit oder Fehlerrate während des Bearbeitens des Tutorials. Proband\*innen mit sehr viel Vorerfahrung in GIMP werden vermutlich bei einem GIMP Tutorial wesentlich schneller sein und weniger Fehler machen als Personen, welche noch nie zuvor mit dem Programm gearbeitet haben.

### **3.3.2 Kalibrierung / Validierung des Eyetrackers**

Nach Abschluss der empirischen Fragen begann der sogenannte Orientierungstest, bei dem bereits Eyetracking-Daten aufgezeichnet wurden. Hierfür wurde erst einmal der Eyetracker durch eine 9-Punkt Kalibrierung kalibriert und im Anschluss validiert.

Das Prozedere wurde den Proband\*innen an dieser Stelle vor Kalibrierungsbeginn genau erklärt. Dabei wurde darum gebeten, nun eine bequeme Sitzposition etwa 60 cm vom Bildschirm entfernt, einzunehmen, bei der die Tastatur und Maus gut erreichbar sind. Ebenso wurde noch einmal darauf hingewiesen, dass die Versuchsperson ab diesem Zeitpunkt möglichst keine Sitzpositionswechsel vornehmen, generell starke (Kopf-)Bewegungen vermeiden und den Blick nach Möglichkeit auf den Computerbildschirm beschränkt lassen sollen, damit der Eyetracker korrekte und vollständige Aufzeichnungen tätigen kann. Fühlte sich die Person bereit, wurde die Kalibrierung und anschließende Validierung der Daten gestartet.

Wenn die Kalibrierungswerte der beiden Augen  $\leq 0,5^\circ$  Abweichung aufwiesen, wurde die Eyetracking-Aufzeichnung begonnen. Betrug die Abweichung mehr als  $0,5^\circ$  wurde die Kalibrierung einmalig wiederholt. Im Falle einer erneuten zu großen Abweichung wurde auf weitere Kalibrierungsversuche verzichtet, um die zu teilnehmende Person nicht zu verunsichern und zu belasten. An dieser Stelle wurde dann ein Vermerk im Proband\*innen Protokoll getätigt. Daraufhin wurden der Test und die Aufzeichnung der Eyetracking-Daten gestartet.

### **3.3.3 Orientierungstest**

Bei einem Software-Tutorial ist üblicherweise neben den jeweiligen Abbildungen beschreibender Text für die durchzuführenden Arbeitsschritte enthalten. Dieser kann die Betrachtungshäufigkeit oder -intensität der Abbildungen unter Umständen beeinflussen. Bei einer ausführlichen textlichen Beschreibung eines Arbeitsschrittes in einem Tutorial könnte gegebenenfalls weniger genau oder überhaupt nicht auf enthaltene Abbildungen geschaut werden. Wie effizient sind die drei Grafikvarianten Screenshot / signalisierter Screenshot / Simplified User Interface jedoch, wenn kein zusätzlicher Text vorhanden ist? Wie schnell finden Proband\*innen einen gesuchten Parameter, wenn lediglich eine Abbildung ohne weiteren Text vorhanden ist? Fällt hier die Reduktion von Detailinformationen mehr ins Gewicht, als wenn zusätzlich gelesen werden kann, was zu tun ist? Um diese Fragen zu beantworten, wurde der eigentlichen Hauptaufgabe der Studie ein weiterer Test vorangestellt. Ziel war es dabei, die Orientierung auf einer einzelnen Grafik (unabhängig von Aufgabenbeschreibungen o. Ä.) analysieren zu können.

Hierzu wurden den Proband\*innen nacheinander fünf einzelne Abbildungen aus einem Bildbearbeitungsprogramm präsentiert. Jede Versuchsperson erhielt dabei alle Abbildungen ausschließlich in der Grafikvariante entsprechend der zuvor zugeordneten Gruppe (SUI/SIG/SCR). Ziel der Aufgabe war es, auf diesen Abbildungen schnellstmöglich einen bestimmten Parameter oder Wert zu identifizieren.

Bei der Auswahl der Abbildungen wurde darauf geachtet, möglichst unterschiedliche Darstellungen zu wählen. Bei der ersten Abbildung wurde ein Listenmenü gewählt, aus dem eine bestimmte Tastenkombination gefunden werden musste (siehe Abb. 7). Bei der zweiten Abbildung wurde dahingegen ein Fenster mit diversen Einstellungsmöglichkeiten und Reglern dargestellt, bei der dritten die gesamte Benutzeroberfläche der Software. Bei Aufgabe vier wurde erneut ein Listenmenü verwendet, jedoch in diesem Fall eine komplexere Variante mit Untermenü. Die letzte Abbildung stellte erneut ein Fenstermenü dar, jedoch mit Gliederung und anderen Einstellungsoptionen.

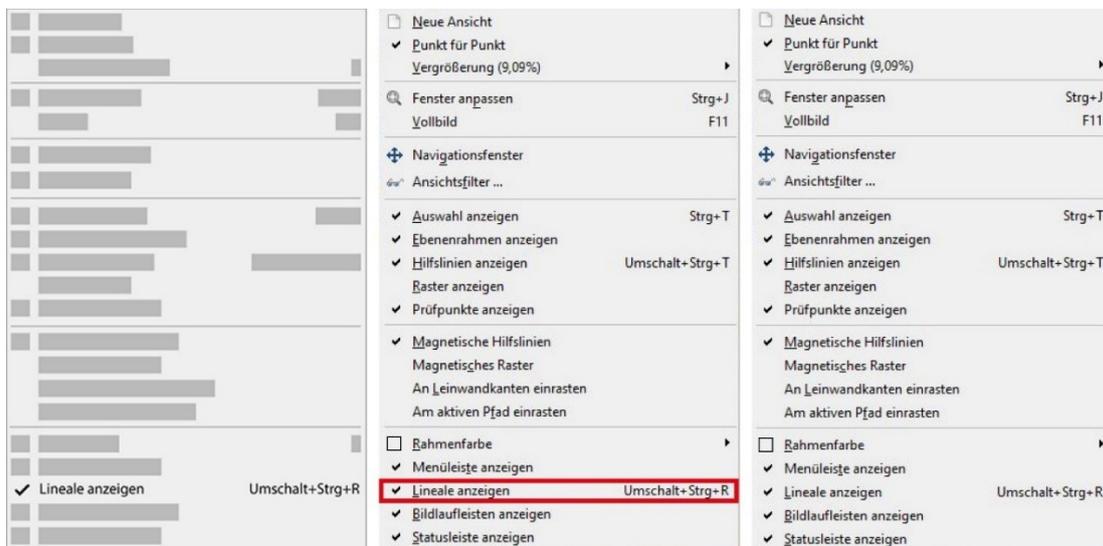


Abb. 7: Stimuli Aufgabe 01: „Welche Tastenkombination benötigen Sie für die Aktion Lineale anzeigen?“ im Orientierungstest. Links: Stimuli für SUI-Gruppe, Mitte: Stimuli für SIG-Gruppe, rechts: Stimuli für SCR-Gruppe. (Quelle: Eigene Darstellung auf Basis von Screenshots aus GIMP.)

Eine Übersicht aller angezeigten Stimuli und Fragestellungen befindet sich aus Platzgründen im Anhang (siehe Anhang: A01).

Bei der Durchführung dieses Orientierungstests wurden Eyetracking-Daten sowie ein Screen Recording aufgezeichnet. Analysiert werden soll anhand dieser Aufgabe im Anschluss die benötigte Zeit und Anzahl an Fixationen bis zur erfolgreichen Identifizierung des gesuchten Parameters. Ebenso werden mögliche Fehler ausgewertet. Die Ergebnisse zu diesen Fragestellungen werden zwischen den drei Gruppen SUI / SIG / SCR miteinander verglichen, um herauszufinden, auf welcher Abbildungsvariante sich effizienter orientiert werden kann.

### 3.3.4 Software-Tutorial

Der Hauptteil der Studie besteht aus der Bearbeitung eines Software-Tutorials für ein Bildbearbeitungsprogramm. Während der Bearbeitung werden Eyetracking-Daten der Proband\*innen aufgezeichnet. Zu dem findet ein Screen Recording statt. Für diese Aufgabe wurde der Computerbildschirm der Proband\*innen geteilt. Auf der linken Seite des Bildschirms befindet sich das Bildbearbeitungsprogramm und das notwendige Bildmaterial. Beides ist zu Untersuchungsbeginn bereits geöffnet. Auf der rechten Seite ist die Startseite des Tutorials wiederzufinden. Da es sich bei dem verwendeten Eyetracker um ein stationäres Gerät handelt, welches Daten von lediglich einem Monitor aufzeichnen kann, war eine Verwendung der Software im Vollbildmodus mit gleichzeitig geöffnetem Tutorial nicht möglich.

Mit diesem Versuchsaufbau bekamen die Proband\*innen die Aufgabe, das Tutorial möglichst selbstständig zu bearbeiten. Um die Ergebnisse der Aufzeichnungen nicht zu beeinflussen, sollte möglichst wenig von der Studienleiterin eingegriffen oder geholfen werden. Dies war lediglich bei technischen Problemen der Fall oder wenn eine Versuchsperson gar nicht weiterkam. In letzterem Fall wurden jedoch auch nur dann so wenige Hinweise wie möglich gegeben.

Grund dafür, dass hier ein Tutorial für ein Bildbearbeitungsprogramm gewählt wurde lag darin, dass diese Software vielfältige und abwechslungsreiche Möglichkeiten liefert. Zum einen ist das Erstellen unterschiedlich schwieriger Handlungssequenzen möglich. Andererseits bieten Bildbearbeitungsprogramme oft eine komplexe Benutzeroberfläche mit vielfältigen Untermenüs und Fenstern, wodurch viele verschiedenartige Abbildungen (Listenmenüs, diverse Fenstermenüs etc.) in der Studie getestet werden konnten. Das spezielle Bildbearbeitungsprogramm GIMP wurde verwendet, da ein Programm gesucht wurde, mit dem möglichst wenige Proband\*innen Vorerfahrungen haben. Da an der Hochschule Merseburg in vielen Kursen im Bereich Bildbearbeitung mit Adobe Photoshop gearbeitet wird, sollte hierzu eine Alternative gefunden werden. Daher war die Wahl des Open-Source-Programms GIMP eine passende Alternative.

#### Tutorial

Im Tutorial haben die Proband\*innen die Aufgabe, ein Poster zu erstellen. Hierfür müssen zwei vorgegebene Fotos miteinander kombiniert und im Anschluss mithilfe diverser Funktionen und Effekten bearbeitet werden.

Das Tutorial besteht aus insgesamt 16 Seiten (siehe beigefügte CD). Es beginnt mit einer Begrüßung und einem Überblick über die zu bearbeitende Aufgabe. Die grobe Aufgabenstellung wird erläutert. Zugleich wird das Ausgangsbildmaterial sowie das

Ergebnis präsentiert. So können sich die Proband\*innen bereits einen ersten Eindruck von der Aufgabe machen und wissen somit was sie erwartet. Auf Seite zwei des Tutorials folgt eine Übersicht über die Benutzeroberfläche der Software. Für das Tutorial relevante Bereiche werden hier hervorgehoben und benannt. So können sich die Proband\*innen bereits mit der Terminologie vertraut machen, welches die korrekte Ausführung der späteren Handlungsanweisungen vereinfacht.

Die folgenden 13 Seiten beinhalten den Hauptteil des Tutorials. Acht Handlungssequenzen (A-H) bestehend aus insgesamt 42 Handlungsschritten bilden das Software-Tutorial. Jede Handlungssequenz hat eine kurze, sachliche Überschrift in Infinitivform, welche das Handlungsziel dieser Sequenz zusammenfasst. Daraufhin folgt eine kurze Einleitung mit Beschreibung dessen, was in dieser Sequenz passieren wird.

Im Anschluss folgen einzelne Handlungsschritte, welche für die Ausführung notwendig sind. Die Aufteilung in einzelne, durchnummerierte Schritte erleichtert den Proband\*innen die Orientierung und Informationsaufnahme, während sie zwischen Benutzeroberfläche der Software und dem Tutorial hin- und herwechseln. Zur einfachen Bearbeitung sind die Handlungsschritte mit orientierenden Komponenten versehen, die darauf hinweisen, wo sich die gesuchte Funktion befindet. Hierfür wurde auf der zweiten Seite (Benutzeroberflächenübersicht) das notwendige Wording eingeführt. Einige Handlungsschritte enthalten zudem bereits eine Zielangabe. Alle enden jedoch jeweils mit einer Resultatangabe, damit die Versuchsperson kontrollieren kann, ob der Schritt korrekt ausgeführt wurde. An einigen Stellen wurden zudem mit einem Ausrufezeichen versehene Hinweisboxen eingefügt. Diese enthalten zusätzliche Informationen, um beispielsweise einen Fehler im Handlungsschritt zu vermeiden oder selbstständig korrigieren zu können.

Wird in dem Handlungsschritt ein neues Menü geöffnet oder eine bestimmte Funktion gesucht, wird dies anhand einer Abbildung bildlich dargestellt. An dieser Stelle unterscheiden sich die drei Varianten des Tutorials. Ein Drittel der Proband\*innen arbeitet mit einem Tutorial mit herkömmlichen, unbearbeiteten Screenshots. Diese sind lediglich auf den entsprechenden Ausschnitt bzw. das entsprechende Fenster zugeschnitten. Das Tutorial von einem weiteren Drittel der Proband\*innen enthält signalisierte Screenshots. Dies sind die identischen Screenshots der ersten Proband\*innen Gruppe, jedoch wurden die für diesen Handlungsschritt relevanten Bereiche durch eine rote Umrahmung optisch hervorgehoben. Bei mehreren Schritten in einem Screenshot wurden die einzelnen Schritte zudem durchnummeriert. Die dritte Proband\*innen Gruppe arbeitet mit dem ebenfalls inhaltlich identischen Tutorial. Jedoch sind in diesem Simplified User Interfaces anstatt der Screenshots enthalten. Diese weisen den identischen Inhalt bzw. Ausschnitt wie die beiden

zuvor genannten Screenshotvarianten auf, sind jedoch stark auf die wesentlichen Inhalte reduziert. (siehe Abb. 08, die vollständigen Tutorials können auf der beigefügten CD in digitaler Form eingesehen werden).

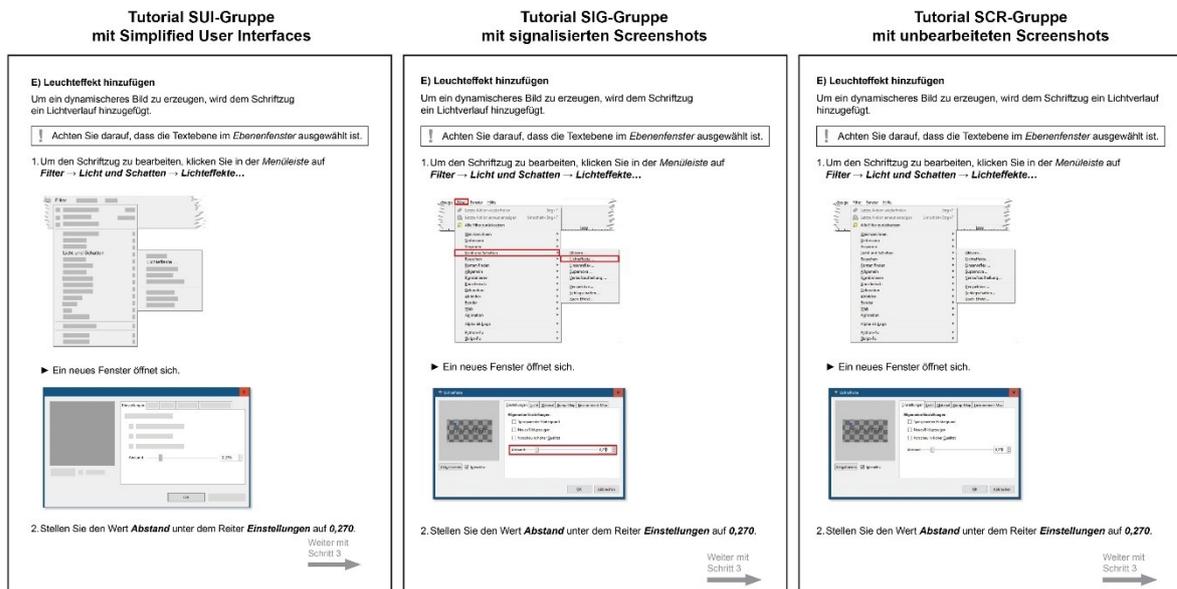


Abb. 8: Beispielhafte Seite des Tutorials in den drei verschiedenen Grafikvarianten (SUI, SIG, SCR). (Quelle: Eigene Darstellung).

Erstellt wurden die SUIs in Adobe Illustrator. Der Original-Screenshot wurde in den Hintergrund gelegt und im Anschluss zunächst nachgezeichnet. Alle irrelevanten Texte, Symbole, Funktionen etc., welche aber zur Orientierung auf der Abbildung notwendig sind, wurden durch einen grauen Kasten verdeckt und ausgegraut. Elemente, welche weder für den Handlungsschritt noch für die Orientierung relevant sind, wurden zudem vollständig entfernt. Lediglich die Texte / Symbole, welche direkt für den Schritt benötigt werden, blieben bestehen. Beendet wird das Tutorial auf der letzten Seite mit einer Darstellung des Endergebnisses. Hier können die Proband\*innen noch einmal abgleichen, ob ihr Resultat dem originalen Endergebnis entspricht. Mit dem Beenden des Tutorials wird gleichzeitig auch die Eyetracking-Datenaufzeichnung beendet.

Ziel von diesem Hauptteil der Studie ist es, vergleichen zu können, wie effizient die einzelnen Proband\*innengruppen mit ihrem jeweiligen Tutorial arbeiten konnten. Lassen sich Unterschiede in der Bearbeitungszeit feststellen? Werden mit der SUI-Variante des Tutorials mehr oder weniger Fehler gemacht, als mit den anderen? Wie (häufig) werden die Abbildungen jeweils überhaupt betrachtet?

### 3.3.5 Abschließende Befragung

Den Proband\*innen wurden im Anschluss noch einige abschließende Fragen zur persönlichen Einschätzung des Tutorials und der jeweils darin enthaltenen Abbildungsvariante gestellt.

Dafür wurden insgesamt sechs Fragen gestellt bzw. Aussagen getroffen, die anhand einer vierstufigen Skala bewertet werden sollten. Es wurde bewusst eine gerade Skala verwendet, um von den Proband\*innen eine eindeutige positive oder negative Tendenz in ihren Antworten ableiten zu können. Die Fragen / Aussagen wurden dabei teils positiv teils negativ formuliert, damit die Proband\*innen in kein gewohntes Antwortschema verfallen, sondern sich die Fragen und Antwortmöglichkeiten genau durchlesen müssen.

#### 01: Wie fanden Sie das Tutorial?

Antwortmöglichkeiten: Leicht, eher leicht, eher schwer, schwer

Mit dieser Frage soll überprüft werden, wie schwierig die Proband\*innen das Tutorial empfunden haben. Wie bereits erwähnt, kann die Schwierigkeit eines Tutorials die Effektivität von Screenshots in Software-Tutorials beeinflussen. Daher ist es an dieser Stelle interessant zu sehen, ob dies auch andere Abbildungsvarianten betrifft bzw. unterschiedliche Ergebnisse bei unterschiedlichen Abbildungsvarianten verursacht.

#### 02: Denken Sie, dass sie alle Aufgaben erfolgreich lösen konnten?

Antwortmöglichkeiten: Ja, eher ja, eher nein, nein

Mit dieser Frage sollte erfasst werden, ob die Proband\*innen selber das Gefühl haben, erfolgreich gearbeitet zu haben oder ob sie Schwierigkeiten mit dem Tutorial gehabt haben. Dies kann nach der Auswertung in Bezug zur Fehlerrate gesetzt werden, um herauszufinden, ob die Proband\*innen mögliche gemachte Fehler selbst bemerkt haben. Zudem beeinflusst es die empfundene Schwierigkeit des Tutorials.

#### 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der Aufgaben.

Antwortmöglichkeiten: Stimme zu, stimme eher zu, stimme eher nicht zu, stimme nicht zu

Mit dieser Frage wurde die Aufmerksamkeit der Proband\*innen zum ersten Mal bewusst auf die Abbildungen gelenkt. Da sie zuvor nicht über das Ziel der Studie – die Untersuchung der unterschiedlichen Grafikvarianten – informiert wurden, sollte diese und die folgenden Fragen dies auflösen. Da die Effizienz der einzelnen Grafiken analysiert werden soll, ist eine subjektive Einschätzung dieser von den Proband\*innen an dieser Stelle äußerst relevant.

#### 04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können.

Antwortmöglichkeiten: Stimme zu, stimme eher zu, stimme eher nicht zu, stimme nicht zu

In der Studie soll analysiert werden, inwieweit die jeweiligen Abbildungen im Tutorial überhaupt betrachtet werden und ob es hier Unterschiede zwischen den Varianten gibt. Eine subjektive Einschätzung der Proband\*innen gibt einen ersten Anhaltspunkt auf diese Frage, welche durch die Analyse der Eyetracking-Daten bestätigt oder widerlegt werden kann.

05: Die Abbildungen im Tutorial haben mich verwirrt.

Antwortmöglichkeiten: Stimme zu, stimme eher zu, stimme eher nicht zu, stimme nicht zu  
Abbildungen in Tutorials sollen die Nutzer\*innen bei der Bearbeitung unterstützen. Da bei den SUIs Detailinformationen zum Teil stark reduziert werden, könnte dies unter Umständen jedoch auch verwirrend auf sie wirken. Genauso könnten Tutorials mit reinen Screenshots ohne jegliche Hervorhebung von relevanten Bereichen ebenso verwirren, da kein Orientierungspunkt gegeben ist. Ob die Proband\*innen durch ihre jeweilige Abbildungsvariante verwirrt wurden, sollten sie bei dieser Aussage einschätzen.

06: Die Abbildungen im Tutorial waren optisch ansprechend.

Antwortmöglichkeiten: Stimme zu, stimme eher zu, stimme eher nicht zu, stimme nicht zu  
Vorteile eines SUI's sind unter Anderem, die klare, moderne Optik. Häufig werden sie aus diesen Gründen in der Marketingkommunikation verwendet, um ein ansprechenderes Erscheinungsbild zu erzeugen. Wie wird diese Grafikvariante also bei ihrem Einsatz in einem Software-Tutorial eingeschätzt? Beeinflusst die Optik das effiziente Arbeiten mit den Grafiken?

Nach Beantwortung der obenstehenden Fragen war die Studie beendet. An dieser Stelle wurde über die eigentlichen Forschungsziele aufgeklärt und mögliche Fragen beantwortet.

### **3.4 Datenanalyse**

Dieses Unterkapitel bietet einen Überblick über die Vorgehensweise bei der Analyse der zuvor erhobenen Daten. Alle Datensätze wurden zunächst im Programm SMI BeGaze gesichtet. Bei den Eyetracking-Daten des Orientierungstests und der Bearbeitung des Tutorials waren zunächst einige Vorbereitungen zur Analyse notwendig. Die empirischen Daten zu den Proband\*innen sowie ihrer Bewertung des Tutorials und der Abbildungen konnten direkt aus dem Programm heraus als CSV-Datei exportiert werden, um im Anschluss statistisch ausgewertet zu werden.

#### **3.4.1 Empirie**

Die empirischen Daten, bestehend aus den Datensätzen der zu Beginn sowie am Ende der Studie gestellten Fragen, wurden über den Metrics Export in BeGaze exportiert. Da es sich

um eine vergleichsweise geringe Anzahl an Datensätzen handelt, wurden diese zunächst in einer Tabelle zusammengefasst und manuell gesichtet. Bei den Fragen nach Beruf, Fachbereich und Studiengang konnte aufgrund der individuellen und damit sehr heterogenen Antworten auf eine weitere statistische Auswertung verzichtet werden. Die Ergebnisse hierzu werden in Kapitel 4.1 aufgelistet.

Die Vorerfahrungen der Proband\*innen mit Bildbearbeitung allgemein und dem verwendeten Bildbearbeitungsprogramm GIMP im Speziellen wurden zunächst codiert, um sie im Anschluss analysieren zu können. Die Codierung erfolgte dabei anhand der Skala: Gar keine Erfahrung = 0; Nicht gut = 1; Durchschnittlich = 2; Gut = 3; Sehr gut = 4. Auf diese Weise können die Mittelwerte und Standardabweichungen der gegebenen Antworten im Anschluss berechnet und zwischen den Gruppen verglichen werden. Bei starken Abweichungen in der Vorerfahrung zwischen den drei Gruppen sollen die Werte nach der Analyse aller weiterer Parameter (Bearbeitungszeit, Fehlerrate etc.) genutzt werden, um diese in Korrelation zur Vorerfahrung setzen zu können. Dies ist notwendig, da eine hohe Vorerfahrung zu einer schnelleren Bearbeitungszeit oder weniger Fehlern führen kann als keine oder eine geringe Vorerfahrung.

Die im Anschluss an das Eyetracking-Experiment erhobenen empirischen Daten zur Bewertung des Software-Tutorials und der enthaltenen Abbildungsarten werden wie die Fragen zur Vorerfahrung codiert und statistisch ausgewertet.

Erfasst wurden in diesem Arbeitsschritt zudem die Kalibrierungswerte des Eyetrackers (Left & Right Eye deviation (X/Y)), um die Genauigkeit der Messungen beurteilen zu können. Ebenso erfasst wurde in diesem Zusammenhang, welche Proband\*innen Sehhilfen (Brille oder Kontaktlinsen) verwendet haben, da dies die Messung ebenfalls beeinflussen kann.

### **3.4.2 Analyse der Eyetracking-Daten**

Zur Untersuchung der drei Hypothesen dieser Arbeit sollten vorrangig die Bearbeitungszeit, die Fehlerrate, sowie besonders relevante Bereiche des Tutorials genauer untersucht werden. Hinsichtlich dieser drei Faktoren wurde die Analyse der erhobenen Eyetracking-Daten des Orientierungstests und des Hauptteils (Bearbeitung des Software-Tutorials) durchgeführt. Abgesehen von der statistischen Analyse erfolgten alle weiteren Schritte in der Analysesoftware zur Auswertung von Eyetracking-Daten BeGaze.

#### Analyse der Bearbeitungszeit

Zur Überprüfung der ersten aufgestellten Hypothese, soll die Bearbeitungszeit der Proband\*innen analysiert werden. Die Hypothese lautete dabei: „Die Nutzung eines

Software-Tutorials mit SUI-Grafiken benötigt weniger Zeit als eine solche mit herkömmlichen Abbildungsarten.“

Für den Orientierungstest zu Beginn der Studie sollte analysiert werden, wie viel Zeit die Proband\*innen benötigen, um die gesuchte Lösung zu finden. Hierfür wurde ein Zeitstempel an der Stelle der Aufzeichnung gesetzt, als die Lösung gefunden wurde. Aus der Differenz zwischen dem Einblenden des Stimulus und dem gesetzten Zeitstempel der Lösungsfindung kann dann die entsprechende Zeitspanne berechnet werden. Diese wurde im Anschluss unter den drei Gruppen miteinander verglichen und statistisch ausgewertet (siehe Kapitel 3.4.3, 4.2).

Um die Bearbeitungszeiten für den Hauptteil der Studie, die Bearbeitung des Software-Tutorials, zu analysieren, wurden die aufgezeichneten Screen Recordings zunächst gesichtet. Die Gesamtbearbeitungszeit jeder Versuchsperson kann bereits im Analyseprogramm BeGaze direkt abgelesen werden. Interessant war für die Untersuchung jedoch nicht der gesamte im Tutorial verbrachte Zeitraum inklusive Begrüßung und Vorstellung der Benutzeroberfläche. Analysiert werden sollten die Bearbeitungszeiten der einzelnen Handlungssequenzen und Handlungsschritte. Folgende Schritte wurden für jede einzelne Versuchsperson durchgeführt: Das Screen Recording wurde mit der Software BeGaze gesichtet. Jeweils der erste Frame, in dem eine neue Handlungssequenz eingeblendet wurde, wurde mit einem Zeitstempel markiert. Dies wurde so für alle acht Handlungssequenzen durchgeführt. Zu beachten war, dass viele der Sequenzen aus zwei Seiten des Tutorials bestehen. Aus der zeitlichen Differenz zwischen dem ersten Frame von Handlungssequenz A und dem ersten Frame der Endresultatseite kann nun die Gesamtbearbeitungszeit des Tutorials berechnet werden. Die gleiche Berechnung wurde angewendet, um die Bearbeitungsdauer der einzelnen Handlungssequenzen zu berechnen.

Die auf diese Weise erhobenen Bearbeitungszeiten werden in einer Tabelle aufgeführt und können so im Anschluss unter den drei Gruppen verglichen und statistisch genauer analysiert werden (siehe Kapitel 3.4.3).

#### Analyse der Fehlerrate

Die Fehlerrate der drei Proband\*innen-Gruppen sollte analysiert werden, um die zweite aufgestellte Hypothese beantworten zu können. Diese lautete: „Bei der Nutzung eines Software-Tutorials mit SUI-Grafiken kommt es zu weniger Fehlern als bei der Nutzung eines Tutorials mit herkömmlichen Abbildungsarten.“

Zur Fehlerrate des Orientierungstests lässt sich sagen, dass an dieser Stelle keine weiteren Auswertungen diesbezüglich stattgefunden haben. Grund dafür ist, dass bei dieser Aufgabe keine Versuchsperson Fehler gemacht hat.

Für die Auswertung der Bearbeitung des Tutorials wurde folgender Prozess durchgeführt. Zunächst wurden dafür die aufgezeichneten Videosequenzen einzeln gesichtet. Dabei wurde für jede Versuchsperson ein Fehlerprotokoll angefertigt, in dem für jeden Handlungsschritt einzeln notiert wurde, ob dieser korrekt ausgeführt wurde. Hatte es Abweichungen gegeben, wurden diese genau vermerkt und definiert, worin die Abweichung bestanden hat. Um mögliche Fehler der Proband\*innen bei der Bearbeitung der Studie analysieren und untereinander vergleichen zu können, musste zunächst eine möglichst genaue Definition des Begriffs „Fehler“ festgelegt werden. Es muss dabei möglichst spezifisch festgelegt werden, welche Handlung unter welchen Umständen als Fehler betrachtet wird, um eine möglichst objektive Betrachtung und Bewertung zu gewährleisten. Hierfür wurden folgende Definitionen für die acht Handlungssequenzen und deren 42 Handlungsschritte festgelegt:

Ein einzelner Handlungsschritt wird mit einer 1 codiert, wenn er vollständig und korrekt ausgeführt wurde. Ausschlaggebend dafür ist der erste Mausklick, welcher im Bildbearbeitungsprogramm getätigt wurde. Wird dieser an der im Tutorial beschriebenen Stelle korrekt getätigt, wird der Handlungsschritt als korrekt ausgeführt bewertet. Wird dieser erste Mausklick bereits falsch ausgeführt, wird der Handlungsschritt mit einer 0 codiert, auch wenn er im Nachhinein noch korrigiert werden sollte. Ebenso als falsch bewertet wurde das vollständige Auslassen eines Handlungsschrittes. Sollte die Versuchsperson anstatt dem im Tutorial beschriebenen Weg eine Alternativmöglichkeit nutzen (beispielsweise das Öffnen eines bestimmten Fensters über die Menüleiste anstatt per Rechtsklick), wird dies trotzdem als korrekt angesehen und damit eine 1 vergeben. Ebenso als korrekt wurde es bewertet, wenn die Aktion anfänglich korrekt ausgeführt, dann aber fälschlicherweise Korrekturen ausgeführt wurden. Auch Folgefehler wurden als korrekt = 1 bewertet. Wenn also die eigentliche Aktion richtig ausgeführt wurde, aber aufgrund eines vorhergegangenen Fehlers das Ergebnis zwangsweise trotzdem falsch war. Nach der Codierung der einzelnen Handlungsschritte (mit 0 = falsch und 1 = korrekt anhand der oben genannten Definition) wurde dies auf die acht Handlungssequenzen übertragen. Jede Handlungssequenz wurde mit einer 1 als korrekt codiert, wenn alle einzelnen darin enthaltenen Handlungsschritte zuvor ebenfalls mit einer 1 als korrekt bewertet wurden. Die Handlungssequenz erhält eine 0 als Bewertung, wenn einer oder mehrere ihrer Handlungsschritte als Fehler bewertet wurden.

Nach Abschluss aller Codierungen der Handlungsschritte und Handlungssequenzen wurden die entsprechenden Codierungswerte in einer Tabelle, sortiert nach Proband\*innen-Gruppe und Handlungssequenzen bzw. -schritten aufgelistet. Diese Werte können nun statistisch analysiert werden (siehe Kapitel 3.4.3).

#### Auswertung durch Areas of Interest (AOI)

Zur Validierung der dritten Hypothese: „Relevante Bereiche in Tutorials mit SUI-Grafiken werden häufiger und schneller fixiert als in Tutorials mit anderen Abbildungsarten.“ wurde das erhobene Eyetracking-Material mithilfe von Areas of Interests (AOIs) analysiert. Areas of Interest sind für den Forschenden besonders relevante Bereiche eines Stimulus, die individuell ausgewählt werden können (siehe Kapitel 3.1.3). Im Anschluss können anhand dieser Segmente spezifische Analysen für diese Bereiche durchgeführt werden (vgl. Geise 2011: 209).

Für den Orientierungstest macht die Auswertung von Areas of Interest methodisch nur wenig Sinn, da es in der Aufgabe um eine möglichst schnelle Lösungsfindung ging und nicht darum, wie viel Zeit in einem relevanten Bereich verbracht wurde. Aufgrund dessen wurde hier auf die Auswertung via AOIs verzichtet. Anstatt dessen sollten bei dieser Aufgabe die einzelnen Fixationen und Blickverläufe der Proband\*innen näher betrachtet werden. Hierfür wurden die Aufzeichnungen der Proband\*innen zu dieser Aufgabe einzeln in BeGaze gesichtet und die Scan Paths der Proband\*innen visuell dargestellt. Hieraus konnte abgelesen werden, wie viele Fixationen benötigt wurden, um die gesuchte Lösung zu finden. Zu Vergleichszwecken wurden außerdem Heat Maps und Focus Maps der einzelnen Stimuli angefertigt, um vergleichen zu können, welche Bereiche der Abbildungen wie intensiv betrachtet wurden.

Für die Erstellung von Areas of Interest bei dem Hauptteil der Studie, der Bearbeitung des Bildbearbeitungstutorials, war eine gewisse Vorarbeit notwendig. Denn Areas of Interest können nur auf statischen Bildern angelegt werden. Die Bearbeitung des Tutorials wurde jedoch als Screen Recording aufgezeichnet, da es sich bei der Bearbeitung um einen interaktiven Prozess handelt. Mit dem Grafikprogramm wird von der Versuchsperson aktiv gearbeitet. Gleichzeitig wird im Software-Tutorial immer wieder umgeblättert, um jede der 16 Seiten und die darauf enthaltenen Handlungssequenzen bearbeiten zu können. Aus diesem Grund wurde das Videomaterial zuerst in einzelne, statische Bilder segmentiert. Diese sollten im Anschluss mit den Eyetracking-Daten der einzelnen Proband\*innen verknüpft werden. Diese Aufgabe wurde im Custom Trail Editor der Analysesoftware BeGaze durchgeführt. Für jede einzelne Seite des Tutorials wurde dabei ein Custom Trail als Snapshot aus dem Screen Recording erstellt, welches neben dem Grafikprogramm die

entsprechende Seite aus dem Tutorial zeigt. Die einzelnen Custom Trails wurden der Übersichtlichkeit halber entsprechend der darin abgebildeten Handlungssequenz benannt.

Das folgende Prozedere wurde dann 24-mal, für das Screen Recording von jeder Versuchsperson, manuell durchgeführt. Die Screen Recording Datei einer Versuchsperson wurde im Custom Trail Editor in BeGaze geöffnet. Der Custom Trail Snapshot von der ersten Seite des Tutorials (Handlungssequenz A) wurde ausgewählt. Dann konnten die zuvor zur Analyse der Bearbeitungszeit gesetzten Zeitstempel verwendet werden, um das genaue Zeitfenster im Screen Recording zu markieren, in dem die Versuchsperson an dieser Handlungssequenz A gearbeitet hat. Im Anschluss wurde der Prozess mit den Custom Trails der anderen Seiten des Tutorials entsprechend wiederholt. Das Ergebnis stellt eine genaue zeitliche Segmentierung der Videodatei nach Handlungssequenzen dar. Diese Segmente werden dann auf die Custom Trail Snapshots übertragen. Wurde dieser Prozess für alle Proband\*innen durchgeführt, liegen nun zur Analyse nicht mehr 24 Screen Recording Dateien des gesamten Tutorials vor. Stattdessen liegt für jede einzelne Seite des Tutorials eine Videodatei vor, auf der jeweils die Blickverläufe aller Proband\*innen dargestellt werden können.

Für jeden Custom Trail können nun Areas of Interest angelegt werden. Hierfür wurden vier spezifische Bereiche ausgezeichnet (siehe Abb. 9). Eine Area of Interest stellt die gesamte Benutzeroberfläche des Bildbearbeitungsprogramms dar. Eine weitere AOI entspricht der gesamten Seite des Software-Tutorials. Durch diese Aufteilung kann im Anschluss ausgewertet werden, wie viel der Gesamtbearbeitungszeit die Proband\*innen jeweils im Tutorial oder dem Bildbearbeitungsprogramm verbracht haben. Von weiterem Interesse war eine genauere Analyse des Tutorials an sich. Dafür wurden zusätzliche AOIs erstellt. Jeder Textabschnitt auf der Tutorialsseite des Custom Trails wurde als eine AOI definiert, genauso wie jede Abbildung mit einer weiteren AOI markiert wurde. Die einzelnen Textabschnitt-AOIs wurden gruppiert, ebenso die Abbildungs-AOIs. Dies wurde für jede Seite des Tutorials entsprechend durchgeführt.

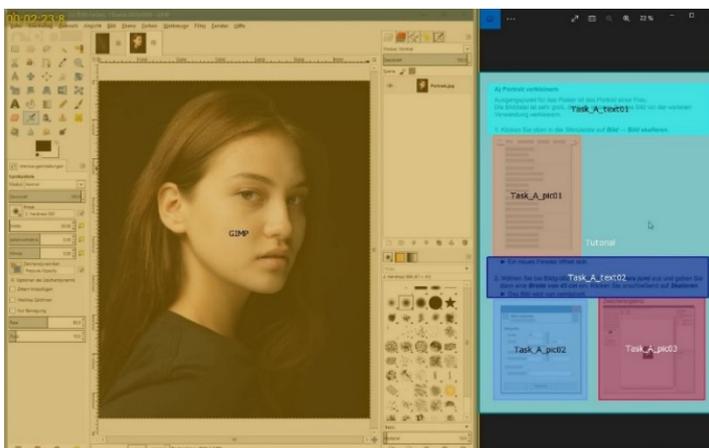


Abb. 9: Beispielhafte Darstellung der angelegten AOIs für GIMP, die gesamte Tutorialsseite sowie für die einzelnen Bild- und Textpassagen im Tutorial. (Quelle: Bearbeiteter Screenshot aus SMI BeGaze und GIMP, Eigene Darstellung.)

Im Anschluss konnten die Eyetracking-Daten der Proband\*innen anhand dieser gesetzten Areas of Interest exportiert und statistisch ausgewertet werden.

### **3.4.3 Export und statistische Auswertung**

Nachdem alle Eyetracking-Daten vorbereitet waren, konnten auch sie aus BeGaze exportiert werden.

Zum Export ausgewählt wurden dabei die Datensätze aller Proband\*innen nach AOI Name und Gruppe anhand ihrer Custom Trails. Als interessante Indikatoren wurden: *Dwell Time [ms] + [%]*, *Revisits* und *Fixation Count* definiert und sollten somit näher ausgewertet werden. Laut dem BeGaze Manual (SMI 2014: 218-220) werden diese Indikatoren wie folgt definiert: Die *Dwell Time* (dt. Verweilzeit) beschreibt die Zeit, die der Blick einer Versuchsperson in einer AOI verweilt, also die Summe aller Fixationen und Sakkaden in diesem Bereich. Sie kann sowohl in Millisekunden (ms) als auch in % ausgegeben werden. *Revisits* beschreibt die Anzahl, wie oft eine bestimmte AOI betrachtet wurde bzw. der Blick einer Versuchsperson nach dem Verlassen der AOI noch einmal zu dieser zurückgesprungen ist. Der *Fixation Count* definiert die Anzahl der Fixationen, die eine Person für eine AOI aufgebracht hat.

Als nächstes wurden die Daten sortiert. Wie bereits erwähnt, wurde bei der Kalibrierung des Eyetrackers darauf geachtet, Abweichungen von mehr als  $0,5^\circ$  zu vermeiden. War dies auch bei einer wiederholten Kalibrierung nicht möglich, wurde die Aufzeichnung trotzdem durchgeführt und ein Vermerk getätigt. Da zur Auswertung der Eyetracking-Daten nur die Aufzeichnung eines Auges notwendig ist, wurden die Kalibrierungswerte zunächst noch einmal betrachtet. Die Datensätze für das Auge mit der geringen Abweichung wurden weiterverwendet, die des anderen Auges verworfen (siehe Tabelle 23 im Anhang A02). Auf diese Weise gibt es lediglich bei zwei der Proband\*innen etwas schlechtere Kalibrierungswerte als gewünscht.

Die Datensätze wurden im Anschluss mithilfe der statistischen Programmiersprache *R* unter Zuhilfenahme der grafischen Benutzeroberfläche *RStudio* statistisch ausgewertet.

*An dieser Stelle sei darauf hingewiesen, dass die statistische Auswertung von der Autorin bewusst nur eingeschränkt erfolgt ist. Unter Beachtung des zeitlichen Umfangs der Masterarbeit und der Tatsache, dass im zugrundeliegenden Masterstudium keinerlei statistische Vorkenntnisse erlangt werden konnten, mussten an dieser Stelle auf eine tiefere statistische Auswertung der erhobenen Daten verzichtet werden.*

Für den Orientierungstest werden die Reaktionszeiten bis zur Identifikation der richtigen Lösung und die Anzahl der Fixationen ausgewertet. Für das Tutorial werden die

Bearbeitungszeit für das gesamte Tutorial sowie pro Handlungsanweisung untersucht. Außerdem werden die korrekt gelösten Handlungsanweisungen und Handlungsschritte betrachtet. Mithilfe der AOIs wird die Verweilzeit (Dwell Time) im Tutorial und GIMP verglichen. Zudem wird die Fixationsanzahl (Fixation Count) der Bilder im Tutorial, sowie den erneuten Rücksprüngen (Revisits) zu diesen analysiert.

Alle auszuwertenden Datensätze wurden anhand ihrer Mittelwerte verglichen. Nach der Berechnung der Mittelwerte erfolgte hierzu zunächst die Berechnung der Standardabweichung und der Standardfehler. Diese sollen als Maß für die Abweichung der einzelnen Werte vom Mittelwert bzw. der Grundgesamtheit dienen. Im Anschluss wurden die Daten auf ihre Normalverteilung geprüft. Aufgrund der kleinen Stichprobengröße ( $n=24$ ) fand die Überprüfung mithilfe des Shapiro-Wilk Tests statt. Um die Varianzhomogenität zwischen den drei Gruppen zu testen, wurde der Levene-Test angewendet.

In einem letzten Schritt wurde untersucht, ob die Daten signifikante Mittelwertsunterschiede aufweisen. Waren die Voraussetzungen für parametrische Tests aufgrund einer fehlenden Normalverteilung ( $p < 0.5$ ) nicht erfüllt, fand die Analyse mithilfe des Kruskal Wallis Tests statt. Bei dem Vorliegen von Mittelwertsunterschieden wurde im Anschluss mit dem Dunn Kruskal Wallis Multiple Comparison Test ausgewertet, zwischen welchen der drei Gruppen Unterschiede vorliegen. Für diese wurde dann die Effektstärke der Unterschiede auf Basis von Cohen (1992: 157) berechnet. Lag eine Normalverteilung der Daten vor, wurde eine analoge Analyse mithilfe eines ANOVA-Tests durchgeführt. Im Anschluss erfolgte ein paarweiser nachgelagerter t-Test, um herauszufinden, zwischen welchen Gruppen Mittelwertsunterschiede vorliegen.

Nach Abschluss der statistischen Analyse wurden die Ergebnisse durch diverse Diagramme (Boxplots, Histogramme, Q-Q-Plots) visualisiert, um die Interpretation zu erleichtern.

## 4. Ergebnisse

Im folgenden Kapitel werden zunächst die Ergebnisse der durchgeführten Eyetracking-Studie aufgeführt. Die Diskussion der Ergebnisse erfolgt im darauffolgenden Kapitel. Ergänzende Visualisierungen zu den ausgewerteten Eyetracking-Daten finden sich der Übersichtlichkeit halber im Anhang A03-A05. Die Datensätze sowie R-Scripts finden sich auf der beiliegenden CD.

### 4.1 Persönlicher Hintergrund

Insgesamt haben 24 Proband\*innen an der Eyetracking-Studie teilgenommen. Von diesen definieren sich 15 selbst als weiblich, 9 als männlich, niemand als divers. Dreiviertel der Proband\*innen waren zwischen 21 und 30 Jahre. Lediglich zwei Versuchspersonen wiesen ein Alter von unter 20 Jahren und vier von über 30 Jahren auf.

Die Versuchspersonen stammen dabei fast ausschließlich aus dem Umfeld der Hochschule Merseburg. Alle Proband\*innen waren zum Zeitpunkt der Studie eingeschriebene Studierende oder hatten bereits ein Studium abgeschlossen. Die Proband\*innen haben damit einen fachlichen Hintergrund aus 17 verschiedenen Studiengängen aus diversen Fachdisziplinen. Darunter fallen sowohl naturwissenschaftliche, wirtschaftliche, kulturelle als auch soziale Studiengänge (1x Angewandte Informatik, 1x Angewandte Medien- und Kulturwissenschaften, 1x Automatisierungstechnik, 2x Chemie- und Umweltingenieurwesen, 2x Green Engineering, 2x aktuell + 1x ehemals Informationsdesign und Medienmanagement, 1x ehemals Kommunikationswissenschaften, 1x Kompass Orientierungssemester, 1x Kultur- und Medienpädagogik, 1x Maschinenbau, 1x Physiktechnik, 1x Politikwissenschaften, 2x Projektmanagement, 3x Soziale Arbeit, 1x Technische Redaktion & E-Learning, 1x Wirtschaftsinformatik, 1x Wirtschaftsingenieurwesen).

Während der Durchführung der Studie trugen 37.5 % der Proband\*innen eine Brille bzw. Kontaktlinsen.

#### 4.1.1 Vorerfahrungen

Auf einer Skala von 0 = Gar keine Vorerfahrung bis 4 = Sehr gute Vorerfahrung weist die gesamte Proband\*innengruppe eine Vorerfahrung mit Bildbearbeitung im Allgemeinen von 1.63 (SD=0.97) auf (siehe Tab. 4). Dies entspricht insgesamt einer geringen bis durchschnittlichen Vorerfahrung.

	Gesamt		SUI		SIG		SCR	
	MW	SD	MW	SD	MW	SD	MW	SD
Vorerfahrung mit Bildbearbeitung	<b>1.63</b>	0.97	<b>1.88</b>	1.13	<b>1.75</b>	0.89	<b>1.25</b>	0.89
Vorerfahrung mit GIMP	<b>0.7</b>	0.91	<b>0.75</b>	1.16	<b>0.75</b>	1.04	<b>0.63</b>	0.52

Tab. 4: Vorerfahrungen der Proband\*innen mit Bildbearbeitung im Allgemeinen und GIMP im Speziellen. Mittelwert (MW) und Standardabweichungen (SD); (0 = Gar keine, 1 = Nicht gut, 2 = Durchschnittlich, 3 = Gut, 4 = Sehr gut. (Quelle: Eigene Darstellung)

Sehr gute Vorerfahrung konnte von den 24 Proband\*innen niemand aufweisen. Etwa 16 % der Proband\*innen gaben an, gute Vorerfahrung mit Bildbearbeitung zu haben. Der Großteil der Probanden\*innen schätzte sich selbst jedoch als durchschnittlich erfahren ein (ca. 46 %). Weitere 5 Personen weisen geringe Vorerfahrung auf, 4 Proband\*innen hatten bisher gar keine Berührungspunkte mit Bildbearbeitung.

Zwischen den drei Gruppen zeigen sich nur geringfügige Unterschiede in der Vorerfahrung mit Bildbearbeitung (siehe Abb. 10). Werden die Mittelwerte betrachtet, liegen alle drei Gruppen auf der Skala zwischen einer *nicht guten* und *durchschnittlichen* Vorerfahrung. Während die SUI-Gruppe einen Mittelwert von 1.88 in ihrer Vorerfahrung aufweist und damit zu einer durchschnittlichen Vorerfahrung tendiert, liegt der Wert bei der SIG-Gruppe mit 1.75 minimal darunter. Die SCR-Gruppe weist mit 1.25 einen wiederum etwas niedrigeren Wert auf.

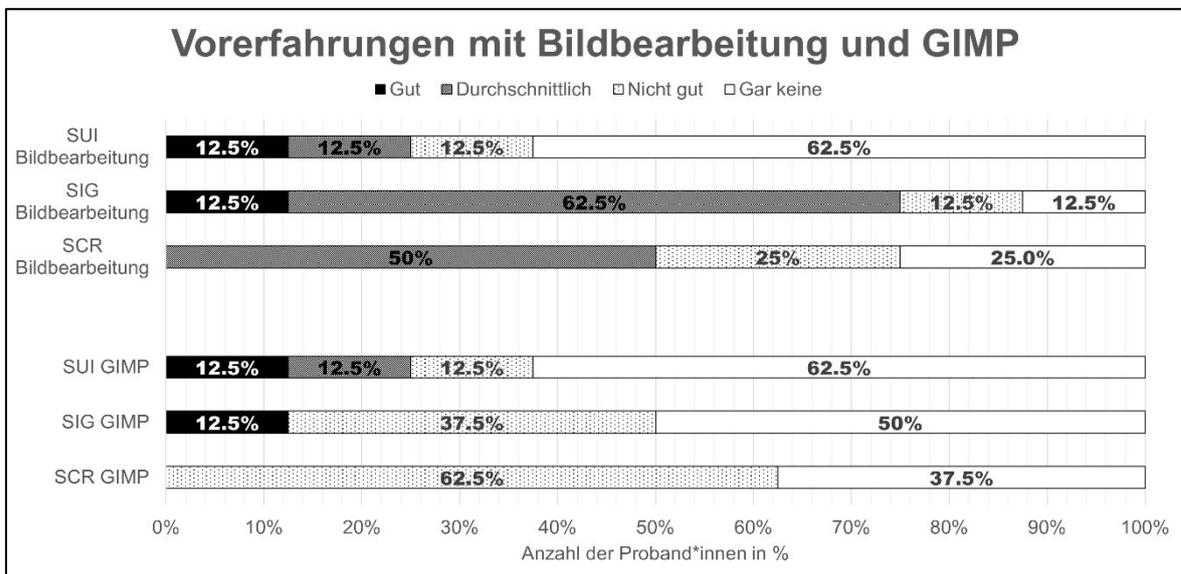


Abb. 10: Verteilungen der Vorerfahrungen der Proband\*innen mit Bildbearbeitung und GIMP in %. (Quelle: Eigene Darstellung).

Die Auswertung der Vorerfahrungen mit dem in der Studie verwendeten Bildbearbeitungsprogramm GIMP zeigt ein ähnlich homogenes Ergebnis (siehe Tab. 4). Die Vorerfahrung der gesamten Gruppe liegt hier jedoch mit einem Mittelwert von 0.7 noch

niedriger. Die Hälfte der Proband\*innen gaben an, bisher gar keine Erfahrung mit GIMP zu haben, weitere 38 % nur sehr geringe. Lediglich eine Versuchsperson schätzte sich als durchschnittlich erfahren ein und zwei weitere hatten bereits gute Vorerfahrung. Niemand gab an, sehr erfahren mit dem Programm zu sein. Zwischen den einzelnen drei Gruppen liegen kaum Unterschiede. Sowohl die SUI-Gruppe als auch die SIG-Gruppe wiesen einen sehr geringen Erfahrungswert von 0.75 auf. Die SCR-Gruppe liegt mit 0.63 minimal dahinter.

Insgesamt lässt sich also feststellen, dass die Proband\*innen zwar grundlegende Vorerfahrung mit Bildbearbeitung mitbringen, jedoch kaum Erfahrung mit dem in der Studie verwendeten Programm GIMP. Die drei Untersuchungsgruppen weisen dabei so geringfügige Differenzen in ihren Vorerfahrungswerten auf, dass diese für die weitere Analyse nicht weiter bedacht werden müssen.

## 4.2 Orientierungstest

Der Orientierungstest wurde hinsichtlich der benötigten Zeit und der Fixationsanzahl ausgewertet. Verwendet wurden für diese Aufgabe 23 von 24 Datensätzen. Eine Versuchsperson (SUI) wurde für die Analyse ausgeschlossen, da die Aufgabe aufgrund von Sprachproblemen nicht korrekt verstanden worden ist. Auf die ursprünglich geplante Auswertung der Fehler wird an dieser Stelle verzichtet, da keine Fehler gemacht wurden.

### 4.2.1 Reaktionszeit im Orientierungstest

Die hier ausgewertete Reaktionszeit entspricht dem Zeitraum von der Einblendung des Stimulus bis zur Identifikation der korrekten Lösung in der Abbildung. Die Proband\*innen benötigten im Durchschnitt etwa 3470 ms ( $\pm 3.47$  s) pro Aufgabe, um die gesuchte Lösung im angezeigten Stimulus zu finden (siehe Tab. 5, Abb. 11).

	MW [ms]	SD	SE
<b>SUI</b> [n=7]	2044.14	914.093	154.510
<b>SIG</b> [n=8]	2076.45	986.849	156.034
<b>SCR</b> [n=8]	6251.71	4260.322	691.115
<b>Ges.</b>	3470.51	3247.019	305.453

Tab. 5: Mittelwerte (MW), Standardabweichungen (SD) und Standardfehler (SE) der benötigten Zeiten pro Aufgabe, unterschieden nach Gruppen. (Quelle: Eigene Darstellung.)

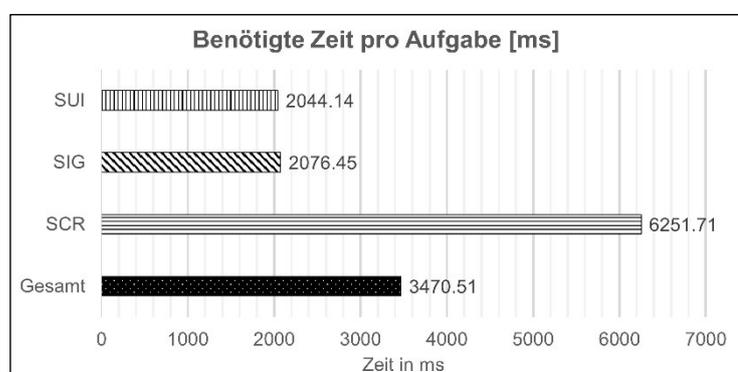


Abb. 11: Benötigte Zeit bis zur Lösungsfindung im Orientierungstest. Mittelwerte pro Aufgabe. (Quelle: Eigene Darstellung.)

Ein Vergleich der drei Gruppen zeigt, dass Proband\*innen mit Simplified User Interfaces (SUI) und signalisierten Screenshots (SIG) mit im Durchschnitt rund 2 s ähnlich viel Zeit benötigt haben. Die Standardabweichung (SD) und Standardfehler (SE) liegen ebenfalls nah beieinander. Die Gruppe, die mit unbearbeiteten Screenshots (SCR) gearbeitet hat, brauchte etwa das dreifache der Zeit und weist zudem höhere Standardabweichungen und -fehler auf. Auch ist hier eine breitere Streuung der Mittelwerte gegeben (siehe A03.1: Abb. 32).

Bei den Datensätzen der drei untersuchten Gruppen liegt jeweils keine Normalverteilung vor ( $p < 0.05$ ) (SUI:  $p = 0.00002275$ , SIG:  $p = 0.0000001813$ , SCR:  $p = 0.0000004665$ ). Mithilfe des Levene-Tests wurde festgestellt, dass signifikante Unterschiede ( $p < 0.05$ ) zwischen den Varianzen der Gruppen bestehen ( $F(2,110) = 9.8959$ ,  $p = 0.0001117$ ).

Neben den signifikanten Varianzunterschieden liegen zudem auch signifikante Mittelwertsunterschiede  $p < 0.05$  vor. Festgestellt wurde dies mithilfe des Kruskal-Wallis-Tests mit dem Ergebnis  $p = 5.08E-14$ . Da das Ergebnis nur aussagt, dass Unterschiede vorliegen, jedoch nicht zwischen welchen Gruppen, wurde dies mittels des Dunn Kruskal Wallis Multiple Comparison Tests spezifiziert (siehe Tab. 6). Hier zeigt sich, dass signifikante Unterschiede ( $p < 0.05$ ) zwischen der Screenshot- (SCR) und der SUI-Gruppe sowie zwischen der Screenshot- (SCR) und Signalisierungs-Gruppe bestehen. Diese weisen jeweils einen starken Effekt von  $p = 0.77$  auf (Bewertung auf Basis von Cohen 1992: 157). Lediglich zwischen SIG und SUI konnte kein signifikanter Unterschied festgestellt werden. Weitere Visualisierungen zur Verteilung der Mittelwerte finden sich im Anhang (A03.1, Abb. 32).

	Z	P unadjust.	P adjust.	Effektstärke
<b>SCR-SIG</b>	6.693544	2.18E-11	4.36E-11	0.7776201
<b>SCR-SUI</b>	6.8375855	8.05E-12	2.42E-11	0.7741368
<b>SIG-SUI</b>	0.3699347	0.71143	0.71143	-

Tab. 6: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.)

#### 4.2.2 Fixationsanzahl im Orientierungstest

Die Fixationsanzahl spiegelt die Anzahl an Fixationen wider, die von den Proband\*innen getätigt wurden, bis sie den gesuchten Parameter im Stimulus identifiziert haben. Die Proband\*innen insgesamt benötigten im Durchschnitt etwa 7.2 Fixationen pro Aufgabe, um die gesuchte Lösung im angezeigten Stimulus zu finden (siehe Tab. 7, Abb. 12).

	MW [ms]	SD	SE
<b>SUI</b> [n=7]	4.028571	1.580873	0.26721
<b>SIG</b> [n=8]	4.475	2.037941	0.32222
<b>SCR</b> [n=8]	13.15789	10.38326	1.68438
<b>Ges.</b>	7.256637	7.460346	0.70181

Tab. 7: Mittelwerte (MW), Standardabweichungen (SD) und Standardfehler (SE) der benötigten Fixationen, unterschieden nach Gruppen. (Quelle: Eigene Darstellung)

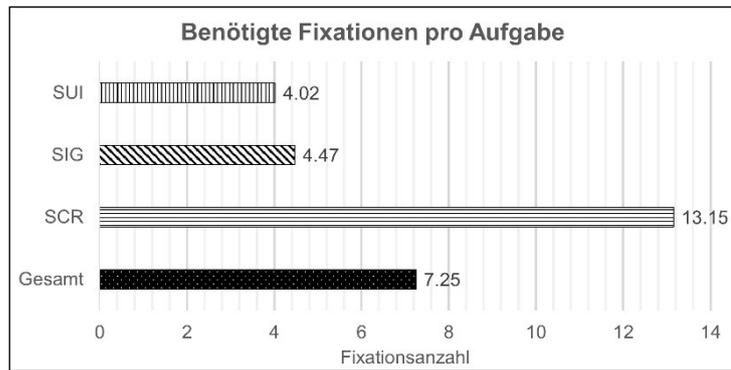


Abb. 12: Benötigte Fixationen bis zur Lösungsfindung im Orientierungstest. Mittelwerte pro Aufgabe. (Quelle: Eigene Darstellung.)

Ein Vergleich der drei Gruppen zeigt auch hier, dass die SUI- und SIG-Gruppe ähnliche Mittelwerte aufweisen. Beide benötigten etwas mehr als 4 Fixationen pro Aufgabe zur korrekten Lösungsfindung. Die Standardabweichungen und -fehler liegen ebenfalls nah beieinander. Große Abweichungen zeigt die Gruppe, die mit reinen Screenshots gearbeitet hat. Diese benötigten im Durchschnitt 13.15 Fixationen, bis sie den richtigen Bereich im Stimulus identifizieren konnten. Die Standardabweichung und Standardfehler sind ebenfalls ausgeprägter, was auf eine breitere Streuung der Mittelwerte hinweist (siehe A03.2 – Abb. 33). Wieder liegt bei den Untersuchten Stichproben jeweils keine Normalverteilung ( $p < 0.05$ ) der Daten vor (SUI:  $p = 0.007672$ , SIG:  $p = 0.0007$ , SCR:  $p = 0.000000077$ ). Der Levene-Tests zeigt für die Fixationsanzahl, dass signifikante Unterschiede ( $p < 0.05$ ) zwischen den Varianzen der Gruppen bestehen ( $F(2,110) = 9.1388$ ,  $p = 0.000213$ ). Zudem liegen signifikante Mittelwertsunterschiede ( $p < 0.05$ ) vor. Der Kruskal-Wallis-Test liefert hier ein Ergebnis von  $p = 2.25E-13$ . Mithilfe des Dunn Kruskal Wallis Multiple Comparison Tests werden diese Unterschiede, ebenso wie bei der Reaktionszeit, zwischen der SCR- und SIG-Gruppe sowie zwischen der SCR- und SUI-Gruppe festgestellt (siehe Tab. 8). Die Effektstärke in beiden Gruppen ist ebenfalls stark ausgeprägt. Ein signifikanter Mittelwertsunterschied zwischen Signalisierung und SUI liegt nicht vor. Weitere Visualisierungen zur Verteilung der Mittelwerte finden sich im Anhang (A03.2, Abb. 33).

	Z	P unadjust.	P adjust.	Effektstärke
<b>SCR-SIG</b>	6.3808189	1.76E-10	3.52E-10	0.728226
<b>SCR-SUI</b>	6.8002876	1.04E-11	3.13E-11	0.7803535
<b>SIG-SUI</b>	0.6382521	5.23E-01	0.523310	-

Tab. 8: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.)

Insgesamt zeigt sich, dass sowohl SUIs als auch signalisierte Screenshots zu einer schnelleren Identifikation relevanter Bildbereiche führen. Zudem benötigen beide Gruppen

weniger Fixationen, um die Lösung zu finden. Die Gruppe, die mit unbearbeiteten Screenshots gearbeitet hat, hat etwa das dreifache der Zeit und an Fixationen benötigt. Für beide untersuchten Parameter können signifikante Mittelwertsunterschiede mit starkem Effekt zwischen den Gruppen SCR und SUI sowie zwischen SCR und SIG festgestellt werden. Ein Vergleich der Gruppen SIG und SUI zeigt keine Mittelwertsunterschiede. Die Varianzen weisen jeweils signifikante Unterschiede auf.

Anhand von Heat Maps lässt sich visualisieren, wie intensiv ein Bereich auf einem Stimulus fixiert wurde. In Abb. 34 / 35 (Anhang A04) zeigt sich auf den meisten Simplified User Interfaces eine geringe Streuung der Fixationen im Vergleich zu den anderen Gruppen. Besonders auf Stimulus 01, dem einfachen Listenmenü, ist zu erkennen, dass bei der SCR-Gruppe zunächst die gesamte Liste im Menü mit den Augen abgescannt wird. Währenddessen findet bei der SUI-Gruppe eine direktere und intensivere Fixierung der relevanten Bildbereiche statt. Die Ergebnisse der SIG-Gruppe liegen jeweils dazwischen. Bei den anderen untersuchten Menüarten zeigt sich hier ein ähnliches, wenn auch etwas weniger ausgeprägtes Ergebnis.

### **4.3 Tutorial**

Die Bearbeitung Software-Tutorials stellt den Hauptteil der Eyetracking-Studie dar. Aufgrund dessen werden für die Analyse diverse Parameter analysiert. Die Bearbeitungszeit des Tutorials wird sowohl für den gesamten Bearbeitungszeitraum als auch für die einzelnen Handlungsschritte betrachtet. Die Fehlerrate soll für die acht Handlungsanweisungen sowie für die 42 einzelnen Handlungsschritte untersucht werden. Mithilfe der AOs findet zudem eine detaillierte Auswertung statt. Statistisch berechnet wird sowohl wie lange der Blick der Proband\*innen im Tutorial bzw. im Bildbearbeitungsprogramm verweilt hat. Ebenso wird untersucht, wie häufig die Bilder im Tutorial fixiert wurden und wie oft der Blick zu ihnen zurückgekehrt ist. Die Ergebnisse werden zwischen den drei Untersuchungsgruppen verglichen, um signifikante Unterschiede in der effizienten Nutzung eines Tutorials mit den jeweiligen Abbildungsarten aufzudecken. Für die folgenden Analysen konnten die Datensätze von allen Proband\*innen (n=24) analysiert werden. Die Gruppengröße entspricht jeweils n=8.

#### **4.3.1 Bearbeitungszeit des Tutorials**

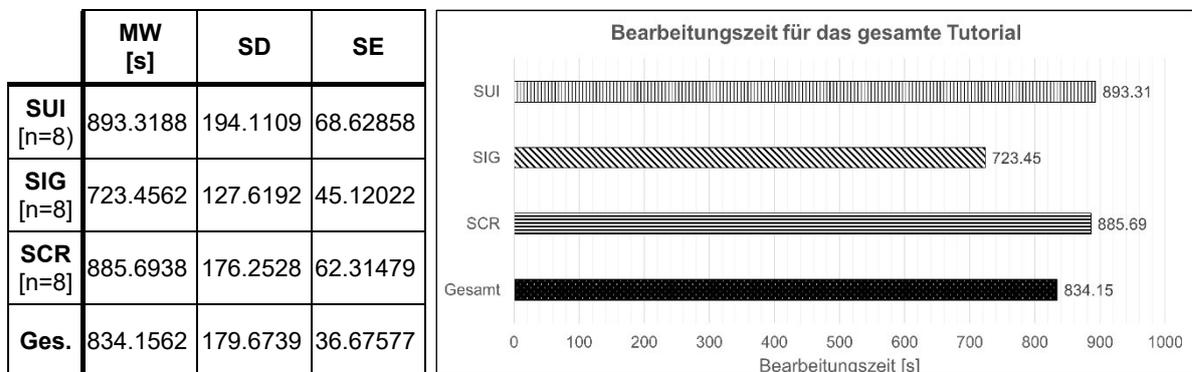
Zur Analyse der Bearbeitungszeit des Tutorials wird zunächst die Gesamtbearbeitungszeit betrachtet. Im Anschluss findet eine Unterteilung in die 8 Handlungsschritte statt, um zu einem differenzierteren Ergebnis zu gelangen. Zur einfacheren Interpretation der Daten

wurden die in Millisekunden aufgezeichneten Werte vor der weiteren Analyse in Sekunden umgerechnet.

### Analyse der Bearbeitungszeit für das gesamte Tutorial

Die Bearbeitungszeit für das gesamte Tutorial betrachtet den Zeitraum beginnend mit dem Zeitpunkt in dem Aufgabe A im Tutorial geöffnet wurde. Den Endpunkt stellte das Verlassen der letzten Aufgabenseite (H) dar. Demnach fließen Begrüßung, die Übersicht über die Softwareoberfläche und die Präsentation des Endergebnisses im Tutorial nicht in die Berechnungen mit ein.

Im Durchschnitt benötigten alle Proband\*innen ca. 13:54 min ( $\hat{=}$  834.15 s) zur Bearbeitung des Tutorials (siehe Tab. 9 / Abb. 13). Die Bearbeitungszeit zwischen den Gruppen SUI und SCR liegen sehr nah beieinander. Die SUI-Gruppe benötigte im Durchschnitt ca. 14:53 min ( $\hat{=}$  893.31 s) und die Screenshot-Gruppe ca. 14:46 min ( $\hat{=}$  885.69 s). Schneller als beide genannten Gruppen zeigte sich die Gruppe, welche mit signalisierten Screenshots gearbeitet hat mit einem Mittelwert von ca. 12:03 min ( $\hat{=}$  723.45 s). In einem ähnlichen Verhältnis wie die Mittelwerte der Gruppen, liegen auch deren Standardabweichungen bzw. Standardfehler. Visualisierungen zu der Verteilung der Daten finden sich zusätzlich im Anhang (A05.1, Abb. 36).



Tab. 9: Bearbeitungszeit für das gesamte Tutorial nach Proband\*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

Abb. 13: Bearbeitungszeit für das gesamte Tutorial in Sekunden. (Quelle: Eigene Darstellung.)

Eine Normalverteilung der Stichproben ( $p > 0.05$ ) ist jeweils gegeben (SUI:  $p = 0.2411$ , SIG:  $p = 0.07098$ , SCR:  $p = 0.4032$ ). Die Überprüfung der Varianzen der erhobenen Daten zeigt, dass zudem eine Varianzhomogenität ( $p > 0.05$ ) zwischen den Gruppen besteht ( $F(2,21) = 0.3494$ ,  $p = 0.7091$ ). Die Unterschiede zwischen den Stichproben können also nur zufällig entstanden sein.

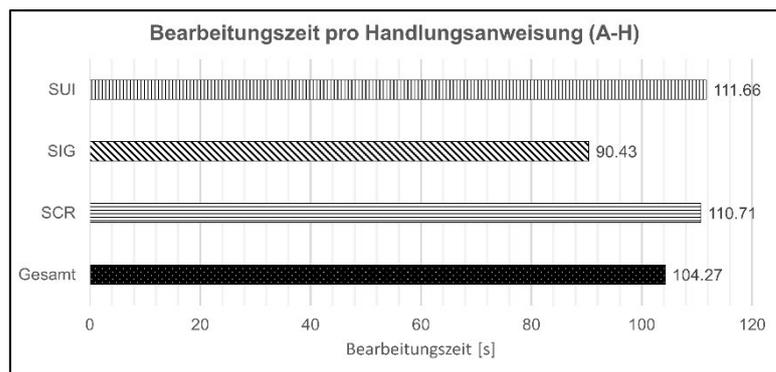
Zur weiteren Untersuchung der Mittelwertsunterschiede wurde mithilfe einer ANOVA festgestellt, dass keine, bzw. nur sehr marginal signifikante Unterschiede zwischen den drei

Gruppen bestehen ( $F(2,21)=2.598, p=0.0981$ ). Demnach lässt sich also kein klarer Vorteil für SUIs erkennen. Die Erhebung zusätzlicher Daten wäre hier vorteilhaft, um die Ergebnisse zu verifizieren.

#### Analyse der Bearbeitungszeit nach einzelnen Handlungsanweisungen (A-H)

Neben der Betrachtung der Gesamtbearbeitungszeit wird auch die durchschnittliche Bearbeitungszeit der einzelnen Handlungsanweisungen analysiert. Hier zeigt sich, dass im Durchschnitt 1:44 min ( $\hat{=} 104.27$  s) pro Handlungsanweisung benötigt wurde.

	MW [s]	SD	SE
<b>SUI</b> [n=8]	111.6659	61.69364	7.711705
<b>SIG</b> [n=8]	90.4325	54.37876	6.797345
<b>SCR</b> [n=8]	110.7120	59.72384	7.465480
<b>Ges.</b>	104.2702	59.19207	4.27182



Tab. 10: Bearbeitungszeit pro Handlungsschritt nach Proband\*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

Abb. 14: Durchschnittliche Bearbeitungszeit pro Handlungsschritt in Sekunden.

Proband\*innen welche mit SUI-Grafiken im Tutorial gearbeitet haben, benötigten im Durchschnitt ca. 1:52 min ( $\hat{=} 111.66$  s) pro Handlungsanweisung, nahezu identisch zeigt sich die Screenshot-Gruppe mit durchschnittlich ca. 1:51 min ( $\hat{=} 110.81$  s) pro Handlungsanweisung (siehe Tab. 10 / Abb. 14). Der Einsatz von signalisierten Screenshots führte zu einer etwa 35 s schnelleren Bearbeitungszeit von ca. 1:30 min ( $\hat{=} 90.43$  s). Die Standardabweichung ist bei der SIG-Gruppe am geringsten ausgeprägt.

Die Datensätze aller Gruppen sind nicht normalverteilt ( $p < 0.05$ ; SUI:  $p=3.443E-05$ , SIG:  $p=3.153E-07$ , SCR:  $3.813E-05$ ). Sie weisen dabei, ähnlich der Gesamtbearbeitungszeit, keine signifikanten Unterschiede hinsichtlich ihrer Varianzen ( $p > 0.05$ ) auf ( $F(2,189)=09.486, p=0.3891$ ). Bei der Betrachtung der Mittelwertsunterschiede mithilfe eines Kruskal-Wallis-Tests lässt sich ein Grenzfall feststellen (siehe Tab. 11). Ein p-Wert von  $p=0.03714$  zeigt zunächst einen Mittelwertsunterschied zwischen den Gruppen. Werden die bereinigten p-Werte betrachtet, zeigen sich zwischen keiner der Gruppen ein Unterschied. Zwischen der Signalisierungs- und SUI-Gruppe ist der p-Wert hierfür jedoch mit  $p=0.054$  grenzwertig. Bei den nicht bereinigten Werten zeigt sich sowohl bei dem Vergleich der SCR- und SIG-Gruppe als auch bei der SIG- und SUI-Gruppe ein Unterschied

im Mittelwert. Visualisierungen zu der Verteilung der Daten finden sich zusätzlich im Anhang (A05.2, Abb. 37).

	Z	P unadjust.	P adjust.	Effektstärke
<b>SCR-SIG</b>	2.23556647	0.02538018	0.07614055	0.198606349
<b>SCR-SUI</b>	0.02624464	0.97906221	0.97906221	0.001053056
<b>SIG-SUI</b>	-2.2093218	0.02715226	0.05430453	0.193762014

Tab. 11: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.)

Eine weitere Datenerhebung und tieferegehende Auswertung sind an dieser Stelle unerlässlich, um zu eindeutigeren Ergebnissen zu gelangen.

### 4.3.2 Fehlerrate im Tutorial

Im Folgenden wird analysiert, wie viel Prozent der Handlungsanweisungen und Handlungsschritte von der jeweiligen Proband\*innen-Gruppe jeweils korrekt gelöst wurde.

#### Fehler-Analyse der 8 Handlungsanweisungen

Zunächst werden die 8 übergeordneten Handlungsanweisungen analysiert. Durchschnittlich wurden unter allen Proband\*innen 80.2 % der Handlungsanweisungen korrekt gelöst. Bei einer Betrachtung der drei Gruppen zeigen sich nur sehr geringfügige Unterschiede (siehe Tab. 12 / Abb. 15). Während die Proband\*innen welche mit signalisierten Screenshots gearbeitet haben 82.81 % der Handlungsanweisungen korrekt bearbeitet hat, liegt die SUI-Gruppe mit 79.68 % nur knapp auf Platz zwei. Die Screenshot-Gruppe schneidet zwar am schlechtesten bei der Bearbeitung ab, jedoch nur mit einem Unterschied von 1.56 % im Vergleich zur SUI-Gruppe.

	MW [%]	SD	SE
<b>SUI</b> [n=8]	79.6875	11.45157	4.048740
<b>SIG</b> [n=8]	82.8125	11.45157	4.048740
<b>SCR</b> [n=8]	78.1250	14.56206	5.148465
<b>Ges.</b>	80.20833	12.17885	2.485996

Tab. 12: Korrekt gelöste Handlungsanweisungen nach Proband\*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

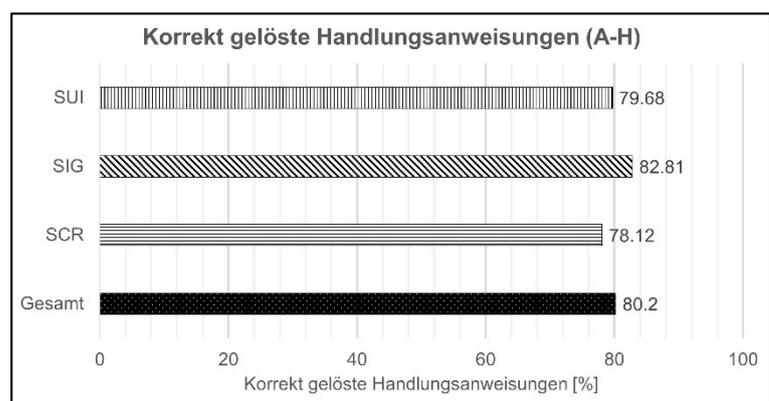


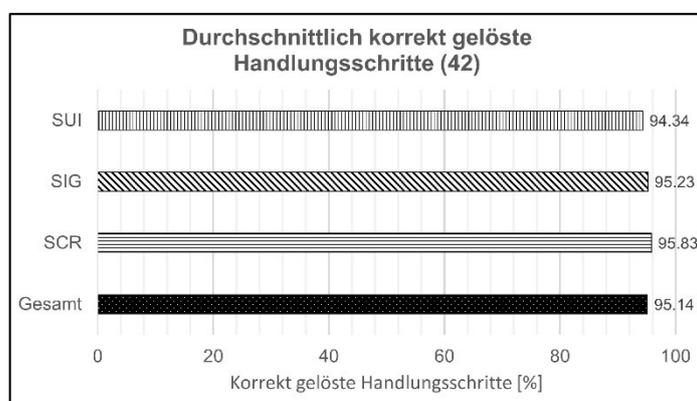
Abb. 15: Durchschnittlich korrekt gelöste Handlungsanweisungen in %. (Quelle: Eigene Darstellung.)

Der aufgrund der geringen Stichprobengröße angewandte Shapiro-Wilk Test zeigt bei sowohl der SUI-Gruppe als auch der SIG-Gruppe eine Normalverteilung ( $p > 0.05$ ) der Daten (SUI:  $p = 0.3241$ , SIG:  $p = 0.3241$ ). Die Werte der Screenshot-Gruppe sind dahingegen nicht normalverteilt ( $p < 0.05$ ; SCR:  $p = 0.03177$ ). Es liegt zudem auch kein signifikanter Unterschied ( $p > 0.05$ ) in den Varianzen vor ( $F(2,21) = 0.066$ ,  $p = 0.9363$ ). Da bei einer Gruppe keine Normalverteilung vorliegt, wurden die Mittelwertunterschiede im Folgenden mit dem nicht parametrischen Kruskal-Wallis-Test berechnet. Die Ergebnisse zeigen jedoch keinen Mittelwertsunterschied ( $p > 0.5$ ) zwischen den Gruppen an ( $p = 0.6216$ ). Visualisierungen zu der Verteilung der Daten finden sich zusätzlich im Anhang (A05.3, Abb. 38).

### Fehler-Analyse der 42 Handlungsschritte

Ein ähnliches Ergebnis wie bei der Betrachtung der 8 Handlungsanweisungen zeigt sich auch, wenn die Ausführung der 42 einzelnen Handlungsschritte näher untersucht wird. Insgesamt konnten durchschnittlich 95.14 % der Handlungsschritte korrekt gelöst werden. Mittelwertsunterschiede zwischen den einzelnen Gruppen sind kaum feststellbar (siehe Tab. 13 / Abb. 16). Sowohl die SIG-Gruppe als auch die SCR-Gruppe konnte durchschnittlich rund 95 % der Handlungsschritte korrekt lösen. Die Werte der Screenshot-Gruppe weisen dabei jedoch eine sehr viel geringere Standardabweichung und Standardfehler auf. Der SUI-Gruppe gelang die korrekte Bearbeitung von 94.32 % der Schritte.

	MW [%]	SD	SE
<b>SUI</b> [n=8]	94.34625	6.606112	2.335613
<b>SIG</b> [n=8]	95.23875	4.589816	1.622745
<b>SCR</b> [n=8]	95.835	2.772616	0.980267
<b>Ges.</b>	95.14	4.735363	0.966601



Tab. 13: Korrekt gelöste Handlungsschritte nach Proband\*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

Abb. 16: Durchschnittlich korrekt gelöste Handlungsschritte in %. (Quelle: Eigene Darstellung.)

In allen Datensätzen konnte sich eine Normalverteilung der Daten nicht bestätigen ( $p < 0.05$ ; SUI:  $p = 0.0006751$ , SIG:  $p = 0.04485$ , SCR:  $p = 0.03177$ ). Wie die in Tab. 13 präsentierten Mittelwerte bereits zeigen, lassen sich auch anhand des Levene-Tests keine signifikanten Unterschiede zwischen den Varianzen ( $p > 0.05$ ) feststellen ( $F(2,21) = 0.2749$ ,  $p = 0.7623$ ). Demnach liegt eine Varianzhomogenität vor. Mittelwertsunterschiede lassen sich für keine

der drei Gruppen feststellen ( $p > 0.05$ ), da im Kruskal-Wallis-Test  $p = 0.9628$  entspricht. Die visualisierte Verteilung der Mittelwerte (Abb. 16) bestätigt hier das Ergebnis. Weitere Visualisierungen zu der Verteilung der Daten finden sich zusätzlich im Anhang (A05.4, Abb. 39).

### 4.3.3 AOI-Auswertung

Anhand der zuvor festgelegten Areas of Interest werden an dieser Stelle die Unterschiede in der Verweilzeit (Dwell Time) im Tutorial und dem Bildbearbeitungsprogramm analysiert. Darüber hinaus werden Unterschiede in der Fixationsanzahl der Bilder im Tutorial zwischen den Gruppen untersucht. Einen weiteren Untersuchungsparameter stellen die Revisits dar, also welche Gruppe häufiger mit ihrem Blick zu einer Abbildung zurückgekehrt ist.

#### Verweilzeit (Dwell Time) im Tutorial oder GIMP

Die Verweilzeit (Dwell Time) zeigt hier die Zeit an, die der Blick einer Versuchsperson pro Aufgabe im Durchschnitt entweder im Tutorial oder im Bildbearbeitungsprogramm GIMP verbracht hat. Es soll untersucht werden, welche Gruppe schneller mit dem Tutorial arbeiten konnte.

Durch die Betrachtung der Mittelwerte lässt sich feststellen, dass alle Gruppen rund 5 s mehr Zeit pro Handlungsanweisung im Tutorial verbracht haben als mit der aktiven Bearbeitung im Bildbearbeitungsprogramm (siehe Abb. 17, Tab. 14). Am wenigsten Zeit hat die SIG-Gruppe mit 42.68 s im Tutorial und 38.72 s in GIMP benötigt. Die anderen beiden Gruppen liegen in etwa gleich auf. Im Verhältnis der Verweilzeiten zwischen den beiden AOIs zueinander zeigt sich jedoch, dass dies prozentual nahezu identisch ist (siehe Abb. 18, Tab. 14). Unabhängig von ihrer Gesamtbearbeitungszeit, haben alle Gruppen ca. 52% der Zeit im Tutorial und 47 % der Zeit in GIMP verbracht. Auch die Standardfehler und -abweichungen zeigen sich ähnlich stark ausgeprägt.

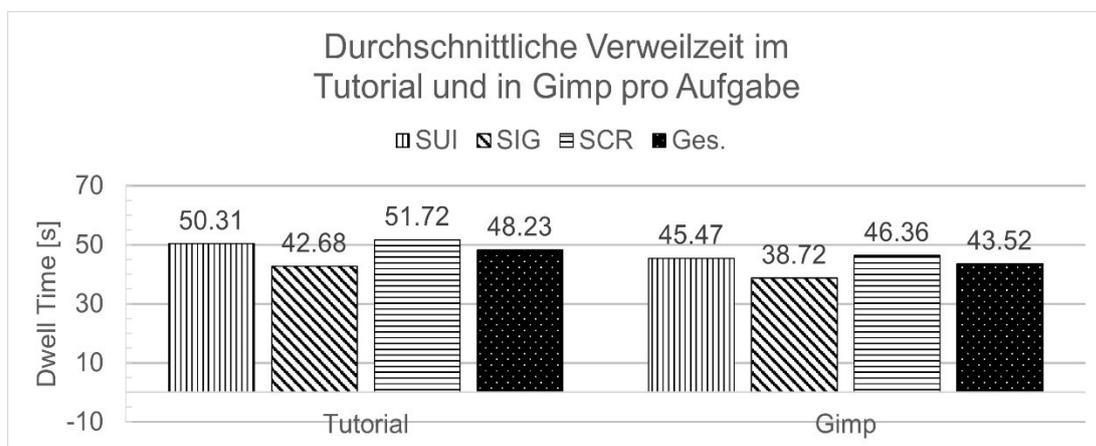


Abb. 17: Durchschnittliche Verweilzeit der Proband\*innen im Tutorial und in GIMP pro Aufgabe. (Quelle: Eigene Darstellung.)

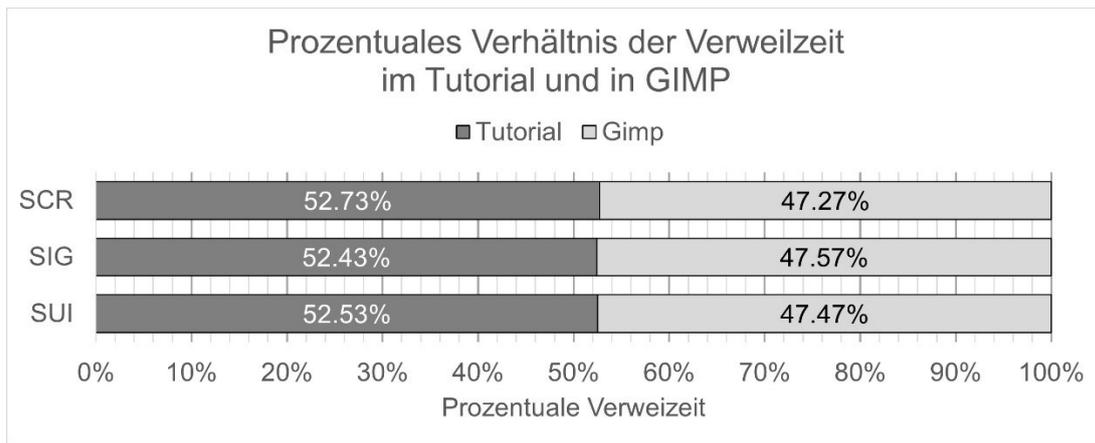


Abb. 18: Prozentuales Verhältnis der Verweilzeiten im Tutorial und in GIMP. (Quelle: Eigene Darstellung.)

	MW (s)			SD		SE	
	Tutorial	Gimp	%-Verhältnis	Tutorial	Gimp	Tutorial	Gimp
<b>SUI</b>	50.3182	45.4778	52.73:47.27	27.5927	27.3059	3.44909	3.413241
<b>SIG</b>	42.6804	38.7218	52.43:47.57	23.8456	26.2386	2.98070	3.279823
<b>SCR</b>	51.7209	46.3696	52.53:47.47	26.1171	27.7787	3.26464	3.472339
<b>Ges.</b>	48.2398	43.5230	52.57:47.42	26.0679	27.1894	1.88129	1.962228

Tab. 14: Verweilzeiten im Tutorial und in GIMP nach Proband\*innen-Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

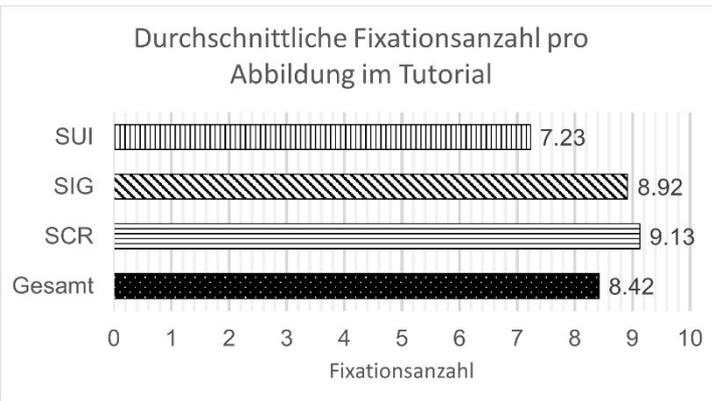
Die Datensätze aller Gruppen weisen keine Normalverteilung auf ( $p < 0.05$ ; SUI:  $p = 0.0007921$ , SIG:  $p = 6.305E-6$ , SCR:  $p = 0.002785$ ). Wie die visuelle Betrachtung der Daten bereits vermuten lässt, liegen auch nach statistischer Berechnung keine Mittelwertsunterschiede vor ( $p > 0.05$ ). Für das Tutorial zeigt sich dies durch  $p = 0.08513$  und für GIMP noch deutlicher durch  $p = 0.1263$ . Es lässt sich auch kein signifikanter Unterschied zwischen den Varianzen (Tutorial:  $F(2,189) = 1.2732$ ,  $p = 0.2823$ , GIMP:  $F(2,189) = 0.3871$ ,  $p = 0.6795$ ) feststellen. Weitere Visualisierungen zu der Verteilung der Daten finden sich zusätzlich im Anhang (A05.5, Abb. 40).

#### Fixationsanzahl in den Abbildungen

Die Fixationsanzahl in den Bildern soll einen Hinweis darauf geben, wie viele Fixationen notwendig waren, um alle nötigen Informationen aus der Abbildung entnehmen zu können. Hierfür wurden zunächst die Mittelwerte aller Fixationen im gesamten Tutorial pro Proband\*in gebildet, um diese im Anschluss unter den Gruppen vergleichen zu können.

Insgesamt haben die Proband\*innen durchschnittlich 8.42 Fixationen für eine Abbildung aufgebracht. Die SUI-Gruppe zeigt dabei mit 7.23 Fixationen die geringste Anzahl, während die SIG-Gruppe mit 8.92 und die SCR-Gruppe mit 9.13 Fixationen etwas mehr, jedoch ähnlich viele Fixationen benötigt hat (siehe Tab. 15, Abb. 19). Die Standardabweichung und der Standardfehler sind bei der SIG-Gruppe jedoch ausgeprägter.

	MW	SD	SE
<b>SUI</b> [n=8]	7.226562	2.059911	0.824039
<b>SIG</b> [n=8]	8.917969	3.500516	1.237619
<b>SCR</b> [n=8]	9.128906	2.330734	0.824039
<b>Ges.</b>	8.424479	2.7259	0.556422



Tab. 15: Durchschnittliche Fixationsanzahl pro Abbildung im Tutorial nach Proband\*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.)

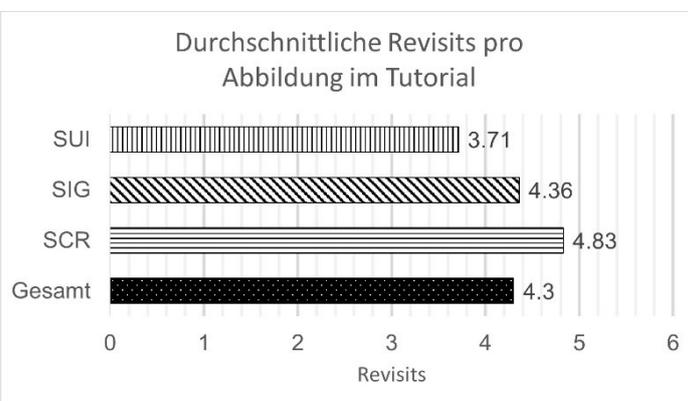
Abb. 19: Durchschnittlich Fixationsanzahl pro Abbildung im Tutorial. (Quelle: Eigene Darstellung.)

Bei den betrachteten Daten liegt jeweils eine Normalverteilung ( $p > 0.05$ ) vor (SUI:  $p = 0.2835$ , SIG:  $p = 0.5666$ , SCR:  $p = 0.06678$ ). Das Ergebnis des Levene-Test zeigt, dass kein signifikanter Unterschied ( $p > 0.05$ ) zwischen den Varianzen der Gruppen besteht ( $F(2,21) = 1.5749$ ,  $p = 0.2305$ ). Zur weiteren Untersuchung der Mittelwertsunterschiede wurde mithilfe einer ANOVA festgestellt, dass zwischen den Gruppen keine signifikanten Mittelwertsunterschiede bestehen ( $F(2,21) = 1.19$ ,  $p = 0.324$ ). Auf eine Untersuchung der genauen Vergleiche der Gruppen kann daher an dieser Stelle verzichtet werden. Die Durchschnittliche Fixationsanzahl pro Abbildung im Tutorial unterscheidet sich also statistisch nicht signifikant zwischen den drei Gruppen SUI, SIG und SCR. Weitere Visualisierungen zu den Mittelwertsunterschieden finden sich im Anhang (A05.6, Abb. 41).

#### Rücksprünge zu den Bildern (Revisits)

Mithilfe der Revisits soll verglichen werden, welche Gruppe am häufigsten, nach dem erstmaligen Verlassen der Abbildung, mit dem Blick zu dieser zurückkehrt.

	MW	SD	SE
<b>SUI</b> [n=8]	3.710938	1.026402	0.362888
<b>SIG</b> [n=8]	4.359375	1.84462	0.652172
<b>SCR</b> [n=8]	4.830452	0.838707	0.296528
<b>Ges.</b>	4.300255	1.337942	0.273106



Tab. 16: Durchschnittliche Revisits pro Abbildung im Tutorial nach Gruppe. Mittelwerte (MW), Standardabweichung (SD) und -fehler (SE). (Quelle: Eigene Darstellung.)

Abb. 20: Durchschnittlich Revisits pro Abbildung im Tutorial. (Quelle: Eigene Darstellung.)

Die Ergebnisse zeigen, dass zu jeder Abbildung im Durchschnitt 4,3-mal mit dem Blick zurückgekehrt wurde. Auch hier liegen die Mittelwerte der einzelnen Gruppen wieder nahe beieinander (siehe Tab. 16, Abb. 20), die Standardabweichung und der Standardfehler sind jedoch bei der SIG-Gruppe mehr als doppelt so groß, gegenüber der SCR-Gruppe. Sie weisen also eine weitere Streuung auf. Während die SUI-Gruppe mit durchschnittlich 3.71 Rücksprüngen die geringste Anzahl dieser zeigt, liegen die anderen beiden Gruppen nur knapp dahinter (SIG: 4.34, SCR: 4.83).

Der Shapiro-Wilk-Test auf Normalverteilung zeigt, dass die Stichprobe der SCR-Gruppe relativ knapp keine Normalverteilung aufweist ( $p=0.04635$ ). Die Stichproben der Gruppen, die mit signalisierten Screenshots und SUIs gearbeitet haben, zeigen jedoch eine Normalverteilung (SUI:  $p=0.5346$ , SIG:  $p=0.08545$ ). Da die Normalverteilung bei einer Gruppe nicht gegeben ist, werden die Bedingungen für parametrische Tests nicht erfüllt. Daher werden die weiteren Mittelwertsunterschiede mit dem Kruskal-Wallis-Test analysiert. Zunächst zeigt jedoch der Levene-Test, dass keine signifikanten Varianzunterschiede zwischen den Gruppen bestehen ( $F(2,21)=0.9167$ ,  $p=0.4153$ ). Die statistische Berechnung der Mittelwertsunterschiede mit dem Kruskal-Wallis-Test zeigt, dass mit  $p=0.09842$  kein Unterschied zwischen den Mittelwerten der einzelnen Gruppen besteht. Demnach macht die Abbildungsart also keinen Unterschied darauf, wie oft zu einer Abbildung mit dem Blick zurückgesprungen wird. Weitere Visualisierungen zu den Mittelwertsunterschieden finden sich im Anhang (A05.7, Abb. 42.).

## **4.4 Bewertung des Tutorials und der Abbildungen**

Nach Abschluss des Tutorials wurden die Proband\*innen zu einer persönlichen Einschätzung des Tutorials und der darin enthaltenen Abbildungen gebeten. Damit soll analysiert werden können, wie schwierig das Tutorial empfunden wurde und wie hilfreich die Abbildungen empfunden wurden. Besonders der Schwierigkeitsgrad einer Aufgabe stellt einen wichtigen Einflussfaktor auf die Effizienz von Abbildungen in multimedialem Material da. Zudem ist es interessant herauszufinden, ob die subjektiven Eindrücke der Proband\*innen mit den statistischen Ergebnissen der Studie übereinstimmen. Die Ergebnisse der Befragung werden im Folgenden aufgelistet.

### **4.4.1 Subjektive Bewertung des Tutorials**

Mit den ersten beiden gestellten Fragen sollte die subjektive Wahrnehmung des gesamten Tutorials von den Proband\*innen eingeschätzt und bewertet werden.

**01: Wie fanden Sie das Tutorial?**

Die Schwierigkeit des gesamten Tutorials wird insgesamt als *eher leicht* bis *leicht* bewertet (95,8 % / 23 v. 24). Innerhalb der drei Proband\*innengruppen lassen sich dabei auch nur geringfügige Unterschiede feststellen. In der SUI- und SIG-Gruppe bewerten jeweils 5 von 8 Proband\*innen das Tutorial als leicht, bei der SCR-Gruppe 4 von 8. Alle weiteren Proband\*innen schätzen es als eher leicht ein. Lediglich eine Person aus der SUI-Gruppe weicht von dieser Einschätzung ab und empfand das Tutorial als schwer (siehe Abb. 21 / Tab. 17).

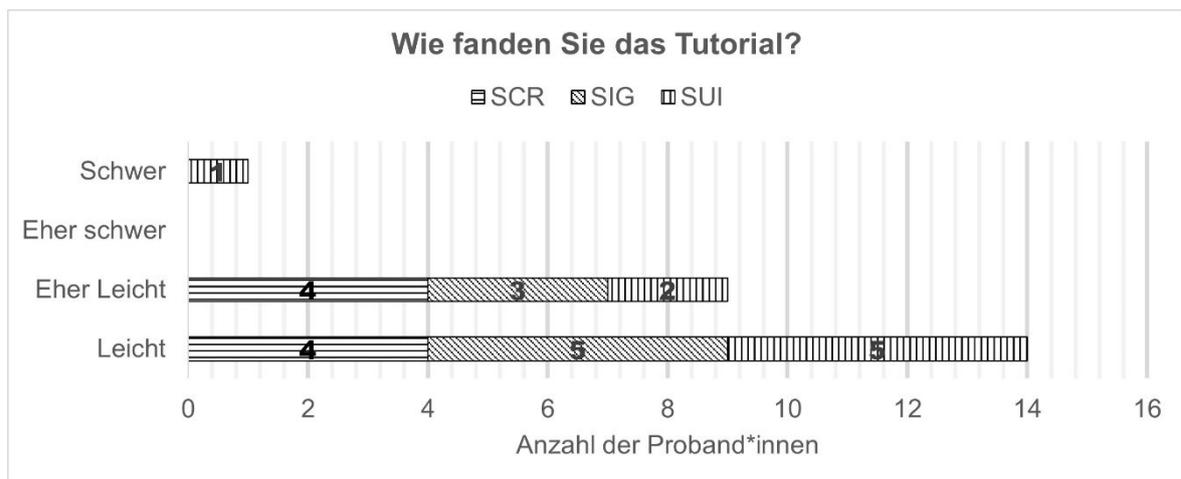


Abb. 21: Antwortverhalten der Gruppen zu Frage 01: Wie fanden Sie das Tutorial? (Quelle: Eigene Darstellung)

	Leicht		Eher leicht		Eher schwer		Schwer	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	58,3	14	37,5	9	0	0	4,2	1
SUI (8)	62,5	5	25	2	0	0	12,5	1
SIG (8)	62,5	5	37,5	3	0	0	0	0
SCR (8)	50	4	50	4	0	0	0	0

Tab. 17: Antwortverhalten zu Frage 01: Wie fanden Sie das Tutorial? (Quelle: Eigene Darstellung.)

**02: Denken Sie, dass Sie alle Aufgaben erfolgreich lösen konnten?**

Das Antwortverhalten auf die zweite Frage fällt ähnlich zur vorherigen aus. 95,8 % (23 v. 24) der Proband\*innen stimmen der Aussage *zu* oder *eher zu*, dass Sie alle Aufgaben im Tutorial erfolgreich lösen konnten. Dabei kommen jeweils 5 von 8 Proband\*innen der SUI- und SIG-Gruppe und 4 von 4 Proband\*innen der SIG-Gruppe zu dieser Einschätzung. Ebenso wie bei Frage 01 gibt es auch hier eine Gegenstimme, welche nicht denkt, dass sie alle Aufgaben erfolgreich lösen konnte (siehe Abb. 22 / Tab. 18).

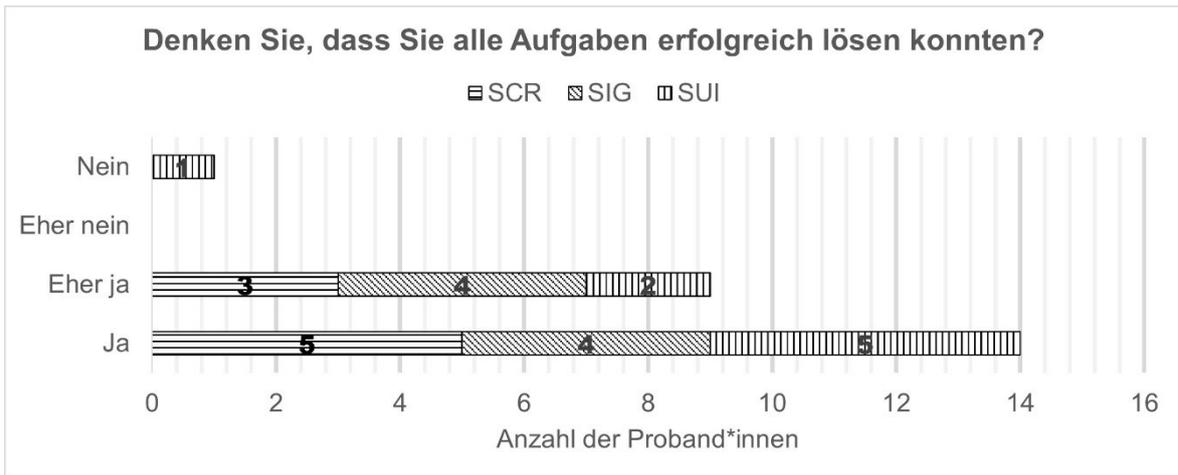


Abb. 22: Antwortverhalten zu Frage 02: Denken Sie, dass Sie alle Aufgaben erfolgreich lösen konnten?  
(Quelle: Eigene Darstellung.)

	Ja		Eher ja		Eher nein		nein	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	58.3	14	37.5	9	0	0	4.2	1
SUI (8)	62.5	5	25	2	0	0	12.5	1
SIG (8)	50	4	50	4	0	0	0	0
SCR (8)	62.5	5	37.5	3	0	0	0	0

Tab. 18: Antwortverhalten zu Frage 02: Denken Sie, dass Sie alle Aufgaben erfolgreich lösen konnten?  
(Quelle: Eigene Darstellung.)

#### 4.4.2 Subjektive Bewertung der Abbildungen

Die folgenden Fragen und Aussagen zielten darauf ab herauszufinden, wie die Proband\*innen die jeweilige Abbildungsart subjektiv empfunden haben.

##### 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der Aufgaben.

Alle Proband\*innen stimmen – unabhängig ihrer Gruppe – der Aussage, dass ihre jeweiligen Abbildungen im Tutorial hilfreich waren, *eher zu* oder *zu*. Während zwei Proband\*innen aus der SUI-Gruppe *eher zustimmen*, gibt es bei den Gruppen SIG und SCR eine 100 % Zustimmung zur Aussage (siehe Abb. 23 / Tab. 19).

	Stimme zu		Stimme eher zu		Stimme eher nicht zu		Stimme nicht zu	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	91,6	22	8,3	2	0	0	0	0
SUI (8)	75	6	25	2	0	0	0	0
SIG (8)	100	8	0	0	0	0	0	0
SCR (8)	100	8	0	0	0	0	0	0

Tab. 19: Antwortverhalten zu Frage 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der Aufgaben.  
(Quelle: Eigene Darstellung.)

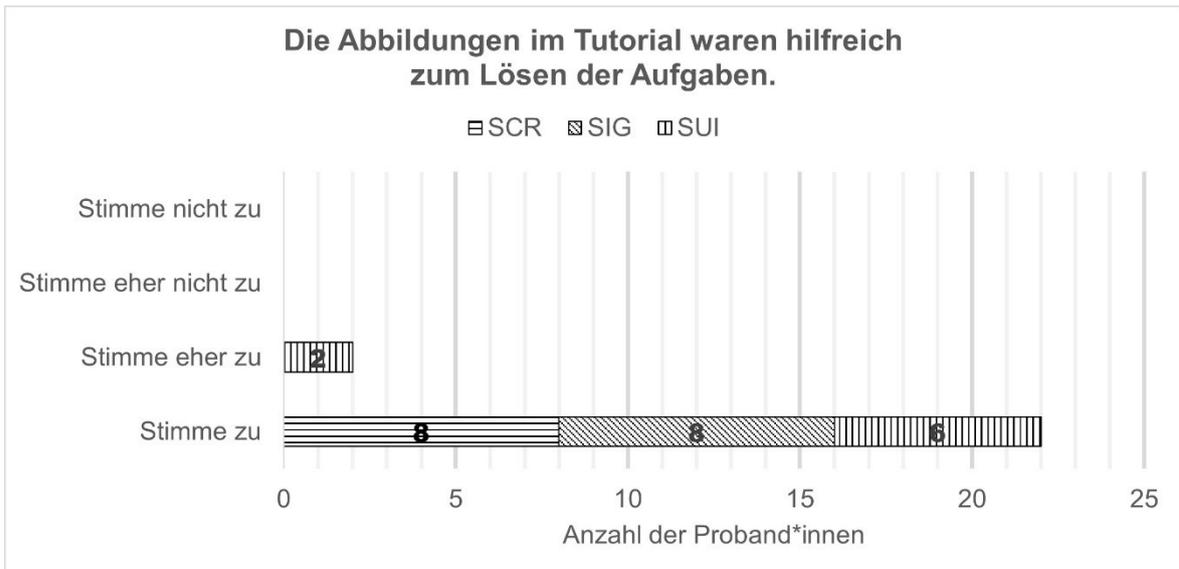


Abb. 23: Ergebnisse zu Frage 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der Aufgaben. (Quelle: Eigene Darstellung.)

**04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können.**

Auf die Frage, ob die Proband\*innen das Tutorial ihrer Einschätzung nach auch ohne Abbildungen hätten lösen können, ist das Antwortverhalten etwas heterogener. 66.6 % (16 v. 24) der Proband\*innen stimmen der Aussage eher oder gar nicht zu. Demnach glauben sie nicht, dass sie das Tutorial ohne Abbildungen hätten lösen können. Das restliche Drittel der Proband\*innen (33.3 % / 8 v. 24) stimmt der Aussage dahingegen zumindest *eher zu*. Hier lässt sich vorrangig in der SIG-Gruppe eine positivere Tendenz feststellen. Die Hälfte der SIG-Proband\*innen denkt demnach, dass sie die Abbildungen nicht unbedingt benötigt hätten. In den anderen beiden Gruppen ist es jeweils nur ein Viertel (siehe Abb. 24 / Tab. 20).

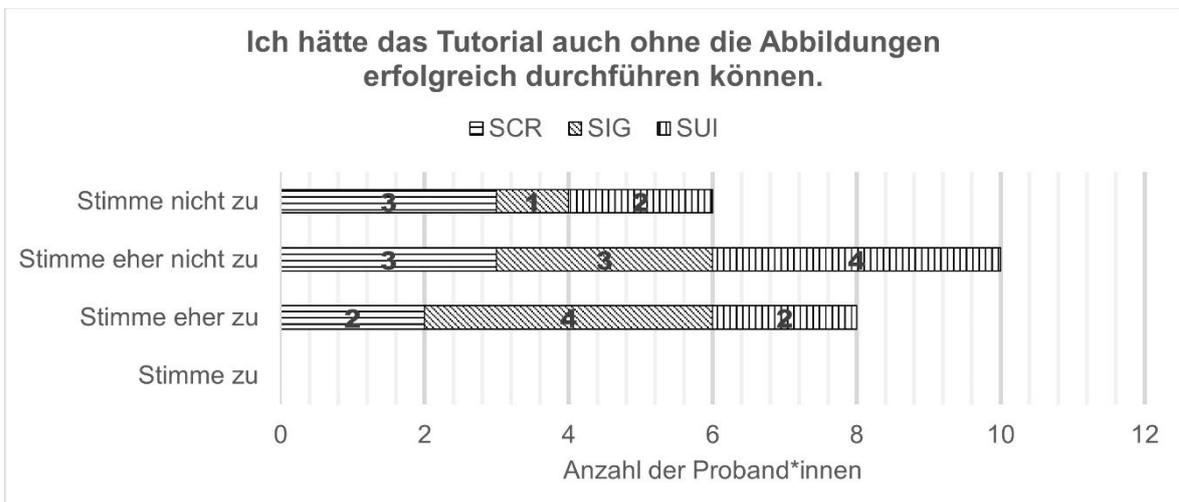


Abb. 24: Antwortverhalten zu Frage 04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können. (Quelle: Eigene Darstellung)

	Stimme zu		Stimme eher zu		Stimme eher nicht zu		Stimme nicht zu	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	0	0	33.3	8	41.6	10	25	6
SUI (8)	0	0	25	2	50	4	25	2
SIG (8)	0	0	50	4	37.5	3	12.5	1
SCR (8)	0	0	25	2	37.5	3	37.5	3

Tab. 20: Antwortverhalten zu Frage 04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können. (Quelle: Eigene Darstellung.)

#### 05: Die Abbildungen im Tutorial haben mich verwirrt.

Die fünfte Frage beschäftigt sich damit, ob die Proband\*innen das Gefühl haben, durch die jeweiligen Abbildungen verwirrt worden zu sein. Dieser Aussage stimmten insgesamt fast alle Proband\*innen *nicht* (75 %) oder *eher nicht* (20.8 %) zu. Lediglich eine Versuchsperson aus der SIG-Gruppe, fühlt sich *eher* verwirrt. Insgesamt lässt sich aber eine stärkere Präferenz feststellen, der Aussage gar nicht zuzustimmen. Demnach also nicht durch die jeweiligen Abbildungen verwirrt worden zu sein. Am stärksten lässt sich dies bei der Gruppe SCR feststellen, am wenigsten bei der SIG-Gruppe, jedoch sind die Unterschiede aufgrund der Proband\*innenanzahl gering (siehe Abb. 25 / Tab. 21).

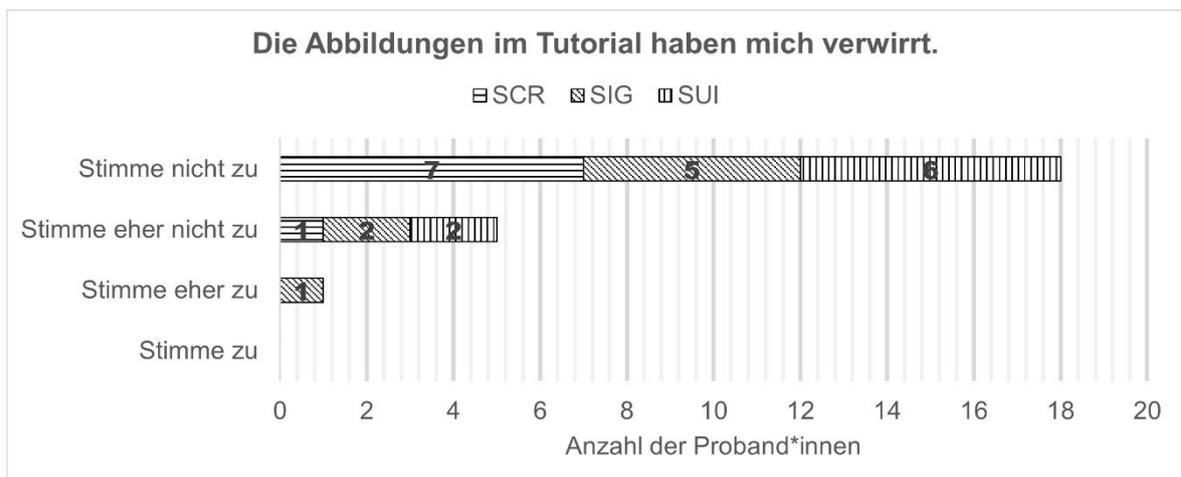


Abb. 25: Antwortverhalten zu Frage 05: Die Abbildungen im Tutorial haben mich verwirrt. (Quelle: Eigene Darstellung)

	Stimme zu		Stimme eher zu		Stimme eher nicht zu		Stimme nicht zu	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	0	0	4.2	1	20.8	5	75	18
SUI (8)	0	0	0	0	25	2	75	6
SIG (8)	0	0	12.5	1	25	2	62.5	5
SCR (8)	0	0	0	0	12.5	1	87.5	7

Tab. 21: Antwortverhalten zu Frage 05: Die Abbildungen im Tutorial haben mich verwirrt. (Quelle: Eigene Darstellung.)

#### 06: Die Abbildungen im Tutorial waren optisch ansprechend.

Auch bei Frage 06 lässt sich ein klares Antwortbild erkennen. Alle Proband\*innen stimmen der Aussage, ob die jeweiligen Abbildungen im Tutorial optisch ansprechend waren, *eher* oder *vollständig* zu. Im Vergleich der Gruppen lässt sich die leichte Tendenz erkennen,

dass die reinen Screenshots am optisch ansprechendsten wahrgenommen wurden (siehe Abb. 26 / Tab. 22). 75 % (6 v. 8) stimmen hier der Aussage völlig zu. Die SUIs wurden im Tutorial als etwas weniger ansprechend wahrgenommen, nur 2 von 8 Proband\*innen (25 %) stimmen der Aussage *völlig zu*, die anderen 75 % jedoch zumindest *eher*.

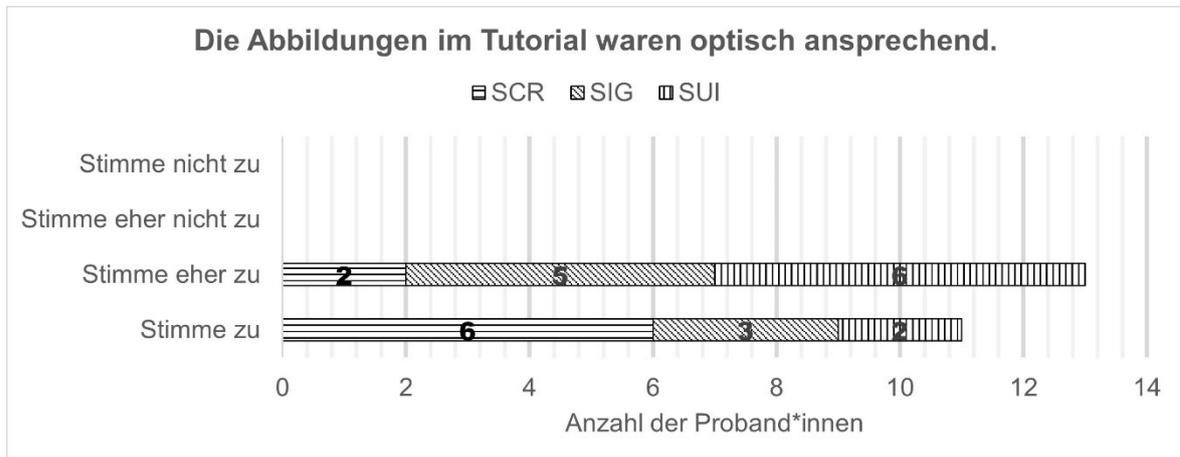


Abb. 26: Antwortverhalten zu Frage 06: Die Abbildungen im Tutorial waren optisch ansprechend. (Quelle: Eigene Darstellung)

	Stimme zu		Stimme eher zu		Stimme eher nicht zu		Stimme nicht zu	
	%	Total	%	Total	%	Total	%	Total
Gesamt (24)	45.8	11	54.2	13	0	0	0	0
SUI (8)	25	2	75	6	0	0	0	0
SIG (8)	37.5	3	62.5	5	0	0	0	0
SCR (8)	75	6	25	2	0	0	0	0

Tab. 22: Antwortverhalten zu Frage 06: Die Abbildungen im Tutorial waren optisch ansprechend. (Quelle: Eigene Darstellung.)

Zusammenfassend lässt sich sagen, dass die Proband\*innen sich – unabhängig von ihrer Gruppe – meistens in der groben Tendenz einig bei der Beantwortung der Fragen waren. 95 % aller Proband\*innen fanden das Tutorial leicht oder eher leicht. Darüber hinaus werden die Abbildungen als (eher) hilfreich, (eher) nicht verwirrend und (eher) optisch ansprechend bewertet. Ohne Abbildungen wäre das Tutorial für die SUI- und SCR-Gruppe (eher) nicht lösbar gewesen, die SCR-Gruppe ist hier unentschiedener Meinung.

Bei den Abstufungen zwischen vollkommener und eher Zustimmung gibt es zwischen den Gruppen leichte Unterschiede. So bewertet jeweils eine Person das Tutorial als schwer und glaubt nicht, alle Aufgaben erfolgreich gelöst zu haben. Optisch werden die reinen Screenshots zudem als etwas attraktiver bewertet als die anderen Abbildungsarten.

## 5. Diskussion

Zu Beginn der Diskussion soll an dieser Stelle darauf hingewiesen werden, dass die Ergebnisse der durchgeführten Studie nicht vollends belastbar und aussagekräftig sind. Aufgrund der geringen Proband\*innen-Anzahl ( $n=24 \pm 8$  pro Gruppe) und weiterer Umstände, die im Folgenden noch diskutiert werden, dienen die Ergebnisse eher als ein erster Anhaltspunkt für die Effizienz von Simplified User Interfaces in Software-Tutorials. Eine Fortführung und Erweiterung der Studie werden daher als sinnvoll erachtet, um diese Ergebnisse verifizieren zu können. Trotz der nur als vorläufig zu betrachtenden Ergebnisse lassen sich jedoch bereits Tendenzen zum Einsatz von SUIs in Tutorials ableiten, die an dieser Stelle genauer betrachtet werden sollen.

Zur näheren Analyse der zuvor vorgestellten Ergebnisse werden zunächst der persönliche Hintergrund der Proband\*innen und deren subjektive Einschätzungen zum Tutorial und den Abbildungsarten diskutiert. Insgesamt nahmen 24 Personen aus dem Umfeld der Hochschule Merseburg an der Eyetracking-Studie teil. Sie weisen durch ihr laufendes oder bereits abgeschlossenes Studium einen höheren Bildungsstand auf, unterscheiden sich jedoch stark in ihrem fachlichen Hintergrund. Aufgrund dessen können keine Rückschlüsse über den fachlichen Hintergrund auf die Verwendung der einzelnen Abbildungsarten in Tutorials gezogen werden. Die allgemeine Vorerfahrung mit Bildbearbeitung hat sich unter den Proband\*innen als eher gering herausgestellt (1.63; siehe Kapitel 4.1.1). Die SUI-Gruppe ist zwar geringfügig erfahrener als die SIG-Gruppe, welche wiederum geringfügig erfahrener ist als die SCR-Gruppe. Jedoch liegen alle drei Gruppen auf der Einschätzungsskala zwischen einer *nicht guten* (1) und *durchschnittlichen* (2) *Vorerfahrung* und können somit als etwa gleichwertig angesehen werden. Relevanter für die Analyse ist die Vorerfahrung mit dem in der Studie verwendeten Bildbearbeitungsprogramm GIMP. Sehr erfahrene Versuchspersonen liegen bei der Bearbeitung des Tutorials klar im Vorteil gegenüber unerfahrenen Personen. Sie würden vermutlich aufgrund ihres Vorwissens zu schnelleren Ergebnissen mit weniger Fehlern gelangen, unabhängig von der Abbildungsart. Die Auswahl der Bildbearbeitungssoftware GIMP hat sich an dieser Stelle als vorteilhaft bewiesen. Wie vor Beginn der Studie vermutet, werden im Umfeld der Proband\*innen, sofern überhaupt, meist andere Bildbearbeitungsprogramme verwendet. Damit zeigt sich insgesamt eine nur sehr geringfügig vorhandene Vorerfahrung mit GIMP unter den Proband\*innen. Alle Gruppen liegen nahe beieinander zwischen einer Vorerfahrung von 0.75 (SUI und SIG) und 0.63 (SCR). Eine 0 bedeutet an dieser Stelle *gar keine Vorerfahrung*, eine 1 *nicht gute Vorerfahrungen*. Insgesamt kann also von einer geringen Vorerfahrung der Proband\*innen ausgegangen werden, womit sie als Anfänger\*innen bezeichnet werden können.

Wie in Kapitel 2.2.1 gezeigt, beeinflusst das Vorwissen der Lernenden den effektiven Einsatz von Abbildungen in Lernmaterialien. Besonders Anfänger\*innen profitieren häufig von Grafiken, während diese sich auf die Performance von Expert\*innen nicht oder gar negativ auswirkt (vgl. Mayer & Gallini 1990: 724, vgl. Kalyuga et al. 1998: 1). An dieser Stelle kann also davon ausgegangen werden, dass die Proband\*innen in der vorliegenden Studie grundsätzlich vom Einsatz der Abbildungen, unabhängig von ihrer Gestaltung, profitiert haben können. Dies spiegelt sich auch in der subjektiven Bewertung des Tutorials und der Abbildungsarten durch die Proband\*innen wider (Kapitel 4.4.1 / 4.4.2). Der Frage, ob die Abbildungen im Tutorial hilfreich zum Lösen der Aufgaben waren, stimmten 22 von 24 Proband\*innen voll und die anderen beiden (beide aus der SUI-Gruppe) eher zu (siehe Abb. 23). Auch bei der persönlichen Einschätzung, ob das Tutorial ohne Abbildungen hätte gelöst werden können, zeigen sich große Tendenzen dem zu widersprechen und damit die Abbildungen zu befürworten. Die Proband\*innen der SIG-Gruppe zeigten hier jedoch ein differenzierteres Bild. Während die Hälfte der Gruppe der Meinung ist, das Tutorial ohne Abbildungen (eher) nicht erfolgreich hätte lösen können, empfindet die andere Hälfte das Gegenteilige (siehe Abb. 24). Neben diesen Ergebnissen zeigt sich auch, dass alle Abbildungsarten als (eher) optisch ansprechend und (eher) nicht verwirrend wahrgenommen wurden. Die Proband\*innen der Screenshot-Gruppe zeigten sich hier am wenigsten verwirrt durch die Abbildungen. Um den möglichen positiven Effekt der Abbildungen allgemein zu verifizieren ist es notwendig die Studie fortzuführen. Mit einer zusätzlichen Kontrollbedingung, einem Tutorial ohne jegliche Abbildungsart, könnte die Effizienz von Abbildungen generell in diesem speziellen Tutorial überprüft werden.

Insgesamt wurde das Tutorial als (eher) leicht wahrgenommen. Lediglich eine Person, welche mit SUIs gearbeitet hat, empfand es als schwer. Die Einfachheit des Software-Tutorials darf an dieser Stelle nicht außer Acht gelassen werden. Einige Studien haben bereits gezeigt, dass sowohl der Einsatz von Screenshots allgemein als auch von signalisierten Screenshots vorrangig bei komplexem Lernmaterial von Vorteil ist. Bei einfachen Aufgaben zeigen sich oft keine Effekte (siehe Kapitel 2.3.2, vgl. Gellevij et al. 2002: 215, vgl. Sweller & Chandler 1994). Die Ergebnisse der vorliegenden Studie können daher von dieser Randbedingung beeinflusst worden sein. Eine Ausweitung der Studie mit komplexeren Aufgabenstellungen wäre an dieser Stelle interessant und würde gegebenenfalls zu aussagekräftigeren Ergebnissen für die Abbildungen führen.

Der durchgeführte Orientierungstest zu Beginn der Studie hatte den Hintergrund, die drei Abbildungsarten (SUI, SIG, SCR) separiert von einer zusätzlichen verbalen Beschreibung analysieren zu können. Hier ist der einzige Stimulus die jeweilige Abbildungsart. Die Proband\*innen werden also nicht von zusätzlichem, beschreibendem Text, wie in einem

Tutorial, beeinflusst. Sie können lediglich die Abbildung nutzen, um die Aufgabenstellung zu lösen.

Die Ergebnisse des Orientierungstests zeigen, dass mit SUIs wesentlich schneller und mit weniger Fixationen ein relevanter Bildbereich in der Abbildung identifiziert werden kann, als mit einem unbearbeiteten Screenshot (SCR). Jedoch zeigt die Verwendung eines signalisierten Screenshots (SIG) einen nahezu identischen Effekt wie der des SUIs. Beide Abbildungsarten, SUI und SIG, ermöglichten es den Proband\*innen in durchschnittlich 2 s die gesuchte Lösung zu finden. Proband\*innen, welche einen unbearbeiteten Screenshot als Stimulus präsentiert bekamen, benötigten etwa das dreifache an Zeit und Fixationen. Dies hat sich auch in der statistischen Berechnung verifiziert. Für beide Parameter, Reaktionszeit und Fixationsanzahl, zeigen sich signifikante Mittelwertsunterschiede zwischen den Gruppen SCR und SUI sowie SCR und SIG mit einem starken Effekt. Die Varianzanalyse hat zudem gezeigt, dass die Werte keine Varianzhomogenität aufweisen. Demnach liegt es nahe, dass die Ergebnisse sich nicht durch Zufall ergeben haben.

Diese Ergebnisse decken sich auch mit dem in Kapitel 2 erläuterten theoretischen Grundlagen zur Effektivität von Abbildungen. Der Vorteil eines SUIs kann sich an dieser Stelle durch seinen reduzierten Detailgrad erklären. Der Mensch tendiert dazu, einfach erfassbare Strukturen in einer Abbildung zu erkennen und diese einfacher interpretieren zu können. Ein hoher Detailgrad erschwert es dabei jedoch, diese Muster wahrzunehmen und sich zu orientieren, wodurch die kognitive Belastung steigt (vgl. Ries 2020: 26). Ein SUI mit seinem klaren, strukturierten Aufbau mit wenig Details scheint daher die Identifikation der relevanten Bildelemente zu erleichtern. Auch hat sich in Studien über den Detailgrad von Strichzeichnungen gezeigt, dass eine optimierte Detailmenge den positivsten Effekt auf die Bearbeitung der Anleitung zeigt (vgl. Alexander et al. 2016, siehe Kapitel 2.2.4). Die Ergebnisse des Orientierungstests spiegeln dies ebenso wider, da sich nicht nur das optimierte, reduzierte SUI effizient gezeigt hat, sondern auch der optimierte signalisierte Screenshot. Bei diesem kommen zusätzliche Informationen hinzu, in dem vorliegenden Fall die rote Umrahmung der zu findenden Lösung. Diese Signalisierung dient der Aufmerksamkeitslenkung und unterstützt die Identifikation relevanter Bildbereiche (vgl. Alpizar et al. 2020: 2097). Dieser Effekt hat sich in den Ergebnissen des Orientierungstests bestätigt. Anhand der Heat Maps (siehe Anhang A04: Abb. 34 / 35) wird dies auch visuell deutlich. Besonders bei textlastigen Abbildungen zeigt sich, wie die Fixationen der Proband\*innen bei nicht optimierten Screenshots einen großen Teil der textlichen Informationen im Bild absキャンen, bevor die richtige Lösung gefunden wird. Bei signalisierten Screenshots ist dieser weitläufige Blickverlauf bereits geringer ausgeprägt. Bei SUIs hingegen zeigt sich oft eine nahezu direkte Fixierung der relevanten Bildbereiche.

Hier wirkt sich die Reduktion der Detailinformationen also positiv aus. Zeigt der Stimulus jedoch die gesamte Bildschirmoberfläche mit vielen Details, aber ohne die vorgegebene Struktur eines Listenmenüs, stellt sich ein anderes Ergebnis dar. Anhand der Heat Maps lassen sich hier kaum Unterschiede zwischen den Abbildungsarten feststellen. An dieser Stelle wäre ebenfalls eine Ausweitung des Tests von Vorteil. Dabei sollte eine größere Anzahl unterschiedlicher Stimuli verwendet werden, um die Ergebnisse zu verifizieren. Da Simplified User Interfaces in diesem Test nahezu identische statistische Ergebnisse zu signalisierten Screenshots erzielt haben, kann die Frage aufgeworfen werden, ob SUIs möglicherweise einfach eine andere Art der Signalisierung mit ähnlichen Effekten darstellen. Wo durch Signalisierung die Aufmerksamkeit durch zusätzliche Elemente gesteuert wird, geschieht dies bei SUIs durch die Reduktion irrelevanter Elemente. Dadurch liegt der Fokus der Betrachtenden auf den relevanten Informationen. Diese werden quasi dadurch „signalisiert“ hervorgehoben, da alles Irrelevante auf unauffällige Farben und Formen reduziert wurde und so in den Hintergrund tritt.

Im Kontrast zu dem effizienten Einsatz von SUIs bei einem Test, wo es lediglich um die schnelle Identifikation von relevanten Bildbereichen geht, zeigt die Designtechnik im untersuchten Tutorial keine klaren Vorteile gegenüber anderen Abbildungsarten. Erst einmal zeigt sich bei den Datensätzen des Software-Tutorials jedoch, dass bei allen untersuchten Parametern keine Unterschiede hinsichtlich der Varianzen bestehen, diese also homogen sind. Demnach können die Stichproben also von einer Grundgesamtheit mit gleicher Varianz abstammen. Dies hat zur Folge, dass die Unterschiede gegebenenfalls nur auf Zufall basieren. Auch Mittelwertsunterschiede konnten statistisch nicht nachgewiesen werden. Trotzdem sollen die Ergebnisse hier unter Vorbehalt diskutiert werden, da sich erste Tendenzen zeigen.

Betrachtet man die Gesamtzeit für die Bearbeitung des vollständigen Tutorials, zeigt sich hier erst einmal die signalisierte Screenshot-Variante als am effektivsten. In der vorliegenden Studie haben die Proband\*innen mit dieser Abbildungsart im Durchschnitt 12:03 min zur Bearbeitung des Tutorials benötigt. Die SCR-Gruppe benötigte 14:46 min, ähnlich wie die SUI-Gruppe mit 14:53 min Bearbeitungszeit im Durchschnitt. Dieses Ergebnis spiegelt sich auch bei einer genaueren Analyse der Bearbeitungszeit pro einzelner Handlungsanweisung wider. Hier benötigt die SIG-Gruppe durchschnittlich 1:30 min pro Handlungsanweisung, während die anderen beiden Gruppen nahezu gleichauf liegen und jeweils gut 21 s mehr Zeit brauchen. Im Gegensatz zur Gesamtbearbeitungsdauer liegt hier zwischen der SIG- und SUI-Gruppe ein Grenzfall vor, wenn man die Mittelwertsunterschiede statistisch analysiert. Während zwischen den anderen Gruppen deutlich kein Mittelwertsunterschied nachweisbar ist, ist dies bei der Kombination SIG-SUI

nur sehr knapp kein Unterschied ( $p=0.054$  gerichtet,  $p=0.0271$  ungerichtet). Auch wenn die erhobenen Daten einen Vorteil für die signalisierte Screenshot-Variante zeigen, und keinen für SUIs, kann das Ergebnis aus den genannten Gründen jedoch nicht uneingeschränkt als nachgewiesen angesehen werden. Stattdessen erfordert es zum einen eine zusätzliche, erweiterte statistische Auswertung. Zum anderen liefert auch dieser Fall einen weiteren Grund die Datenerhebung fortzusetzen, um zu eindeutigeren Ergebnissen zu gelangen. Im Rahmen dieser Masterarbeit kann dies jedoch nicht mehr ausgeführt werden.

Sofern sich das derzeitige Ergebnis in einer Erweiterung der Studie bestätigen sollte, stellt es einen weiteren Nachweis für das Signalisierungsprinzip nach Mayer und die Anwendung von Multimedialen Lernprinzipien auf handlungsorientierte Texte dar. An dieser Stelle bleibt jedoch die Frage offen, warum SUIs an dieser Stelle einen weniger positiven Effekt liefern. Zur Identifikation relevanter Bildelemente haben sich SUIs im Orientierungstest als gleich effizient zu Signalisierungen gezeigt. Im realen Anwendungskontext eines Software-Tutorials scheinen jedoch zusätzliche Detailinformationen durch Signalisierungen möglicherweise vorteilhafter zu sein. Die Tendenz des Orientierungstests hin zu einem Vorteil von SUIs gegenüber anderen Abbildungsarten spiegelt sich jedoch auch in der Fixationsanzahl sowie den Rücksprüngen des Blicks zum Bild nach dem erstmaligen Verlassen wider. Bei beiden untersuchten Parametern zeigen SUIs die geringste Fixationsanzahl (7.23 Fixationen pro Bild, 3.71 Rücksprünge). Die SIG- und SCR-Gruppe liegen hier nahe beieinander bei einer durchschnittlich höheren Fixationszahl. Die Unterschiede zu SUI sind jedoch auch hier wieder insgesamt nur sehr geringfügig ausgeprägt. Zudem sind statistisch weder Varianz- noch Mittelwertsunterschiede nachweisbar, weshalb auch diese Ergebnisse mit Vorsicht betrachtet werden müssen. Hinzu kommt, dass eine geringere Fixationsanzahl sowohl positive als auch negative Rückschlüsse auf die Effizienz einer Abbildung zulässt. Wenige Fixationen und wenige Rücksprünge zum Bild können einerseits bedeuten, dass die relevanten Bildelemente schnell identifiziert werden konnten und kein weiterer Abgleich notwendig war ( $\hat{=}$  positiver Effekt). Gleichzeitig können sie jedoch auch auf mangelndes Interesse oder Irrelevanz von der Abbildung hindeuten ( $\hat{=}$  negativer Effekt). Zieht man die bereits erläuterten subjektiven Bewertungen der Proband\*innen in diese Betrachtung mit ein, kann hier von einem eher positiven Effekt ausgegangen werden. Der Großteil der Proband\*innen hatte angegeben, dass die Abbildungen als hilfreich und notwendig zum Lösen der Aufgaben eingeschätzt werden.

Betrachtet man die bei der Bearbeitung des Tutorials gemachten Fehler, zeigt sich hier kein signifikanter Unterschied zwischen den drei Proband\*innen-Gruppen. Von den acht Handlungsanweisungen wurden mit der SIG-Variante 82.81 % aller Aufgaben korrekt

gelöst. Mit SUIs dahingegen 79.68 % und mit SCRs 78.12 %. Auch wenn die Unterschiede nicht groß sind, zeigt sich hier wieder eine leichte Tendenz zur effizienteren Nutzung von Signalisierungen. Wird die Fehlerzahl separierter, auf Basis aller 42 Handlungsschritte betrachtet, relativiert sich das Ergebnis jedoch wieder. Alle drei Gruppen konnten demnach rund 95 % der Handlungsschritte korrekt lösen, womit kein Unterschied feststellbar ist. Mit SUI-Grafiken sind Proband\*innen also in der Lage genauso wenige Fehler zu machen, wie mit den Abbildungsalternativen. Neben der Fehlerzahl gilt dies auch für die Verweilzeit (Dwell Time) der Proband\*innen im Tutorial und im Bildbearbeitungsprogramm. Basierend auf der gemessenen kürzeren Gesamtbearbeitungszeit des Tutorials für die Signalisierungsgruppe, verbringen die Proband\*innen dieser Gruppe auch grundlegend weniger Zeit sowohl im Tutorial als auch in GIMP, während die anderen Gruppen gleichauf liegen. Interessanter ist es daher, das prozentuale Verhältnis der Verweilzeit zwischen Tutorial und GIMP zu betrachten. Hier zeigt sich ein nahezu exakt identisches Ergebnis. Alle Gruppen haben rund 53 % der Bearbeitungszeit im Tutorial verbracht sowie rund 47 % im Bildbearbeitungsprogramm.

Diese uneindeutigen und statistisch nicht vollends belastbaren Ergebnisse werfen die Frage auf, ob die Abbildungen im Tutorial überhaupt relevant sind. Wie bereits erwähnt zeigen sich Abbildungen generell, als auch visuelle Signalisierungen vor allem bei komplexem Material als nützlich. Das Tutorial wurde jedoch als (sehr) einfach bewertet. Vielleicht waren die Abbildungen daher an einigen Stellen gar nicht notwendig und wurden nur sporadisch betrachtet. Dafür spricht auch, dass der Text des Tutorials teilweise ebenfalls aus der Abbildung ablesbare Inhalte wiedergegeben hat. Somit könnten sich die Bilder an einigen Stellen als redundant und nicht notwendig gezeigt haben. Sung und Mayer haben hierzu bereits nachgewiesen, dass sich nur relevante Grafiken als effizient auszeichnen. Irrelevante Grafiken dahingegen nicht (vgl. Sung & Mayer 2012: 1618). Auch wenn die Abbildungen im Tutorial sich nicht völlig irrelevant zeigen, könnte die Kombination aus Redundanz und Einfachheit der Aufgaben zu einem ähnlichen Effekt beigetragen haben. Dies könnte insgesamt auch als Erklärung dienen, warum viele der gemessenen Werte zufällig entstanden sein können.

Insgesamt lässt sich also festhalten, dass eine Fortführung der Studie dringend notwendig ist, um zu eindeutigen Ergebnissen zu gelangen und die Ergebnisse dieser Studie verifizieren (oder widerlegen) zu können. Neben einer wesentlich größeren Stichprobengröße, sollte vor allem komplexeres Material untersucht werden. Zudem sollte bei dem Tutorial darauf geachtet werden, dass sich Bild und Text nicht zu redundant zeigen. Grundsätzlich wäre auch eine weitere Kontrollbedingung, das rein textbasierte Tutorial ohne Abbildungen, eine denkbare Option für eine Erweiterung der Studie. Somit könnte der

grundlegende Effekt von Abbildungen im Tutorial untersucht werden. Außerdem würde es eine Möglichkeit bieten leichter zu beurteilen ob der Effekt einer Abbildungsart auf sein Naturell als Abbildung zurückgeht, oder sich stattdessen ein Effekt aus der Designtechnik ableiten lässt.

Ob sich der Aufwand, ein Simplified User Interface zu erstellen, lohnt, ist anhand der vorhandenen Ergebnisse nur eingeschränkt zu beurteilen. SUIs haben sich als effizient dafür herausgestellt, relevante Bildbereiche schnell und mit wenig Aufwand identifizieren zu können. Jedoch bieten signalisierte Screenshots dasselbe Potenzial, verlangen aber weniger Aufwand in der Erstellung. Im realen Anwendungskontext in einem Tutorial konnte durch ein SUI kein messbarer, positiver Effekt gegenüber anderen Abbildungsvarianten festgestellt werden. Es werden zwar tendenziell etwas weniger Fixierungen und Rücksprünge zu den Bildern getätigt, dies hat jedoch in der vorliegenden Studie keine Auswirkungen auf die Bearbeitungszeit oder andere Parameter. Aufgrund dessen lässt sich daran kein positiver Effekt auf die Effizienz ausmachen. Nicht zu vernachlässigen ist hingegen das Potenzial welches SUIs hinsichtlich der einfacheren Lokalisierung und Wiederverwendbarkeit im praktischen Arbeitskontext bieten. Wird eine Software-Dokumentation in viele verschiedene Sprachen übersetzt, spart die Erstellung eines SUIs viel Übersetzungsaufwand und -kosten. Irrelevante Texte werden entfernt oder auf einfache Formen reduziert, so dass die Sprache an dieser Stelle keine Rolle spielt. Stattdessen kann das SUI in vielen Sprachvarianten eingesetzt werden, wobei nur der relevante Teil des SUIs übersetzt werden muss. Ebenso sind nur geringe Anpassungen für die Wiederverwendung in unterschiedlichen Betriebssystemen oder nach Software-Updates notwendig. Einmal erstellte SUI-Elemente können zudem immer wieder verwendet werden. Aus diesen Gründen bietet es sich an, individuell abzuwägen, an welchen Stellen einer Dokumentation sie sinnvoll eingesetzt werden können, so dass sich der Erstellungsaufwand lohnt. Ein Beispiel für einen sinnvollen Einsatzzweck von SUIs könnten Bereiche in der Dokumentation sein, in denen bestimmte relevante Elemente identifiziert und sich grundlegend in der Software orientiert werden soll. Ein selektiver Einsatz von SUIs in Kombination mit anderen Abbildungsvarianten wäre zudem denkbar.

Die angewendete Methodik des Eyetrackings als Instrument zur Untersuchung von Software-Tutorials hat sich als effektiv und nützlich gezeigt. Neben der Bearbeitungszeit und der Fehlerrate konnten durch individuelle Areas of Interest auch die einzelnen Abbildungen für sich stehend analysiert und zwischen den Gruppen verglichen werden. Durch den Versuchsaufbau, bei dem Tutorial und Bildbearbeitungsprogramm nebeneinander auf demselben Bildschirm dargestellt werden, war es zudem möglich, beide Stimuli separat als Areas of Interest auszuwerten. So konnte ein Vergleich zwischen der

Nutzung vom Tutorial und dem Programm gezogen werden. Insgesamt bietet Eyetracking damit viel Potenzial, um zu einer Verbesserung der Forschungslage zur effizienten Nutzung von Technischen Dokumentationen beizutragen. Durch die Beobachtung und Analyse der Blickbewegungen und Fixationen auf individuellen Areas of Interest ergibt sich eine Vielzahl an Möglichkeiten: Abbildungsarten, Layouts, Textauszeichnungen und viele weitere Gestaltungsmittel können je nach Bedarf individuell analysiert werden. Bei der Untersuchung von multimedialen Lernmaterialien wird häufig als Indikator für eine effiziente Gestaltung im Anschluss an den Lernprozess Wissen abgefragt oder Transfertests durchgeführt. Für Software-Tutorials eignet sich dies weniger, da Tutorials nicht bearbeitet werden, um sie auswendig zu lernen, sondern um sie aktiv zu nutzen. Hierfür ist ein Studiendesign notwendig, was es den Proband\*innen erlaubt, individuell zwischen Material und Software agieren zu können. Die in dieser Studie durchgeführte Zweiteilung des Bildschirms für Tutorial und Bildbearbeitungsprogramm in Kombination mit Eyetracking-Aufzeichnungen hat diese Möglichkeit effektiv genutzt. Gleichzeitig konnte ein Split-Attention-Effekt durch diesen Versuchsaufbau so gering wie möglich gehalten werden, da sich die Versuchspersonen nur auf einen Monitor konzentrieren müssen.

## 6. Fazit

Die vorliegende Arbeit hatte zum Ziel, die Effizienz von Simplified User Interfaces in Software-Tutorials zu untersuchen. Hierfür wurden anhand der Cognitive Load Theory und der Cognitive Theory of Multimedia Learning zunächst die grundlegenden Prozesse beim Lernen mit Texten und Bildern erläutert. Der Fokus lag dabei auf Prinzipien und Einflussfaktoren für die lernförderliche Gestaltung von Abbildungen. Grundlegend kann Lernmaterial positiv von dem Einsatz von Abbildungen profitieren, sofern bestimmte Randbedingungen gegeben sind. Hierzu zählt eine ausreichende Komplexität des Materials, sowie ein eher geringes Vorwissen der Lernenden. Ebenso spielt die Relevanz der Abbildung eine entscheidende Rolle. Aufmerksamkeitssteuernde Elemente wie Signalisierungen beeinflussen die Effizienz einer Abbildung ebenfalls. Jedoch hat sich in bisheriger Forschung auch hier vorrangig bei komplexerem Material ein positiver Effekt eingestellt. Hinsichtlich der Komplexität und des Detaillierungsgrades einer Abbildung kann sich sowohl ein erhöhter Detailgrad, beispielsweise durch Signalisierungen, positiv auswirken. Gleichzeitig zeigt jedoch auch ein reduzierter Detailgrad auf die wesentlichen Informationen einen positiven Effekt auf die kognitive Belastung und kann zu einer leichteren Identifikation von relevanten Bildelementen führen. Dieser Effekt kommt auch bei Simplified User Interfaces zum Tragen.

Als weiteres Fazit sollen die Forschungsfragen und Hypothesen dieser Arbeit noch einmal betrachtet und beantwortet werden.

### ***H01: SUI-Grafiken erleichtern die Identifikation relevanter Bildelemente im Vergleich zu anderen Abbildungsarten.***

Hypothese 01 kann an dieser Stelle nicht abgelehnt werden. Im Orientierungstest haben SUI-Grafiken zu einer schnelleren Identifikation von relevanten Bildelementen im Vergleich zu unbearbeiteten Screenshots beigetragen. Zudem waren bei SUI-Stimuli deutlich weniger Fixationen notwendig. Auch eine Betrachtung der Heat Maps zeigt, zumindest für strukturierte und textlastige Abbildungen, dass eine intensivere Fixation von relevanten Bildbereichen mit SUIs erzielt werden kann. Dahingegen wird vor allem bei unbearbeiteten Screenshots das gesamte Bild von oben nach unten abgescannt, um die Lösung zu finden. Zu bedenken ist jedoch, dass signalisierte Screenshots einen nahezu identischen Effekt gezeigt haben. Im Tutorial konnte dieser positive Effekt in der vorliegenden Studie nur in geringfügigen Tendenzen anhand der Fixierungsanzahl der Bilder nachgewiesen werden. Demnach haben Proband\*innen mit SUIs zwar geringfügig weniger Fixierungen benötigt, um die jeweilige Abbildung zu erfassen. Zu einem weiteren Effekt auf die Bearbeitungszeit oder andere Parameter hat dies jedoch nicht geführt.

***H02: Die Bearbeitung eines Software-Tutorials mit SUI-Grafiken benötigt weniger Zeit als eine solche mit herkömmlichen Abbildungsarten.***

Hypothese 02 kann nicht bestätigt werden. Um sie vollständig abzulehnen sind jedoch weitere Forschungen zur Validierung der Ergebnisse notwendig. Sowohl in der Bearbeitungszeit des gesamten Tutorials als auch in der durchschnittlichen Bearbeitungszeit pro Handlungsanweisung hat sich die SUI-Gruppe als die langsamste Gruppe herausgestellt. Jedoch sind die Unterschiede zur SCR-Gruppe jeweils nur minimal, so dass beide Gruppen als ähnlich effizient angesehen werden können. Die SIG-Gruppe hat dahingegen durchschnittlich wesentlich schneller gearbeitet. Die Unterschiede sind aufgrund fehlender statistisch nachgewiesener Varianz- und Mittelwertsunterschiede nicht vollends belastbar. Auch hier ist eine Fortführung der Forschung von Relevanz.

**H03: Bei der Bearbeitung eines Software-Tutorials mit SUI-Grafiken kommt es zu weniger Fehlern als bei der Bearbeitung eines Tutorials mit herkömmlichen Abbildungsarten.**

Hypothese 03 kann an dieser Stelle tendenziell abgelehnt werden. Die Ergebnisse der Studie zeigen, dass die Proband\*innen mit SUIs in etwa genauso viele Fehler machen, wie mit unbearbeiteten oder signalisierten Screenshots. Zu beachten ist bei der Einschätzung jedoch, dass das Tutorial insgesamt als sehr einfach bewertet wurde. Damit kann ein positiver Effekt von Abbildungen grundsätzlich, unabhängig von ihrer Gestaltung, in Frage gestellt werden.

**F01: Wie effizient sind Simplified User Interfaces in Software-Tutorials im Vergleich zu herkömmlichen Abbildungsarten?**

**F02: Hat die Reduktion von Detailinformationen bei SUIs einen zusätzlichen, positiven Effekt?**

Unter Vorbehalt bezüglich der in der Diskussion erläuterten Kritikpunkte an der durchgeführten Studie, zeigen sich Simplified User Interfaces nur bedingt effizient in Software-Tutorials. Bezüglich der korrekten Bearbeitung des Tutorials sind SUIs genauso effizient wie unbearbeitete und signalisierte Screenshots. Wird die Bearbeitungszeit betrachtet, sind SUIs weniger effizient als signalisierte Screenshots. Mit diesen wurde in der vorliegenden Studie deutlich schneller bei gleicher Fehlerzahl gearbeitet. Die Reduktion der Detailinformationen in einem Simplified User Interface wirkt sich vor allem positiv bei der Identifizierung von relevanten Bildbereichen aus. Diese konnte im Orientierungstest mit signifikanten Unterschieden zur SCR-Gruppe festgestellt werden. Während der Bearbeitung des Tutorials haben sich die reduzierten Details möglicherweise positiv auf die

Anzahl der Fixierungen und Rücksprünge zur jeweiligen Grafik ausgewirkt. Die Unterschiede zu den Vergleichsgruppen sind jedoch nur minimal und statistisch nicht vollends belegbar. Vorteile bieten SUIs vor allem durch ihre Wiederverwendbarkeit sowie der einfacheren Lokalisierung. Die Reduktion der Detailinformationen wirkt sich zudem positiv auf die Identifikation von relevanten Bildbereichen aus. Aufgrund dessen lässt sich der Einsatz von SUIs an ausgewählten Stellen in einem Software-Tutorial empfehlen. Relevante Einsatzbereiche können Stellen des Tutorials sein, wo sich auf der Benutzeroberfläche orientiert und bestimmte relevante Bereiche schnell und einfach lokalisiert werden sollen. Wird das Tutorial in viele Sprachen übersetzt oder sind viele kleine Änderungen in der Softwareoberfläche absehbar, können zum Beispiel Technische Redakteur\*innen ebenfalls von SUIs profitieren.

Eine Übertragbarkeit der CTML auf handlungsorientierte Texte zeigt sich in der vorliegenden Studie bedingt. Erste Hinweise darauf liefert der als positiv zu bewertende Effekte des Signalisierungsprinzips. Das Multimedia-Prinzip konnte nur durch die subjektive Einschätzung der Proband\*innen bestätigt werden, dass die Abbildungen im Tutorial hilfreich zum Lösen der Aufgabe waren. Ein Beweis durch die Eyetracking-Aufzeichnungen steht noch aus. Hierfür wäre eine weitere Kontrollbedingung, ein Tutorial ohne Abbildungen, notwendig.

Aufgrund der geringen Anzahl an Proband\*innen in der vorliegenden Studie wäre es notwendig, die Studie mit einer größeren Anzahl fortzuführen, um die bisherigen Ergebnisse zu verifizieren. Zudem wurde das untersuchte Software-Tutorial von den Proband\*innen als sehr leicht bis leicht wahrgenommen. Bilder haben sich generell in bisheriger Forschung jedoch vor allem bei komplexen Aufgaben als vorteilhaft gezeigt. Dies könnte erklären, warum sich in der durchgeführten Studie keine positiven Effekte für SUIs hinsichtlich der Bearbeitungszeit und Genauigkeit gezeigt haben. Weitere Forschung sollte dieser Frage nachgehen und die Effizienz von SUIs in komplexen Aufgabenstellungen untersuchen. Auch die mögliche Redundanz von einigen Abbildungen zum Text des Tutorials sollten bei einer Fortführung der Studie eliminiert werden.

## **6.1 Ausblick**

Die vorliegende Studie bietet zwar bereits Hinweise darauf, dass Simplified User Interfaces die Orientierung auf einem Bild und die Identifizierung relevanter Bildelemente erleichtern können. Weitere Forschung ist jedoch notwendig, um herauszufinden, wie SUIs optimal gestaltet und eingesetzt werden können. Wie weit kann man SUIs verändern bzw. reduzieren, um den Arbeitsprozess bestmöglich zu unterstützen? Welche Auswirkungen macht der komplette Verzicht von Texten oder irrelevanten Elementen, anstatt der

Unkenntlichmachung? Insbesondere die Entfernung von Text hätte weitere Vorteile bezüglich der Lokalisierung und Wiederverwendbarkeit zu bieten. Auch sollte weiter untersucht werden, für welche Abbildungen einer Dokumentation oder eines Tutorials SUIs möglicherweise positivere Effekte zeigen und wie sich eine Mischung aus SUIs und signalisierten Screenshots in einem Dokument auswirkt. Außerdem stellt sich die Frage, ob die geringere Detailmenge zu einem langanhaltenden positiven Lerneffekt beiträgt. Zudem kann der Einsatz von SUIs auf verschiedenen Geräten untersucht werden. Auf einem kleineren Display, wie einem Tablet oder Smartphone, sind Details aufgrund der geringeren Größe an sich schwerer erkennbar. Hier könnte die reduzierte Screenshot-Variante einfacher erkennbar sein. Auch der Einsatz von SUIs in (teil-)immersiven Medien stellt ein weiteres interessantes Forschungsfeld dar.

Im Hinblick auf die durchgeführte Eyetracking-Studie ist die bereits erwähnte Ausweitung auf mehr Proband\*innen und komplexeres Lernmaterial notwendig. Dadurch erzielte, statistisch belastbarere Ergebnisse, könnten zudem weitere Hinweise darauf geben, ob multimediale Lernprinzipien auch auf handlungsorientierte Texte übertragbar sind.

# Anhang

## A01: Orientierungstest Stimuli

Frage 01: Welche Tastenkombination benötigen Sie für die Aktion: Lineale anzeigen?

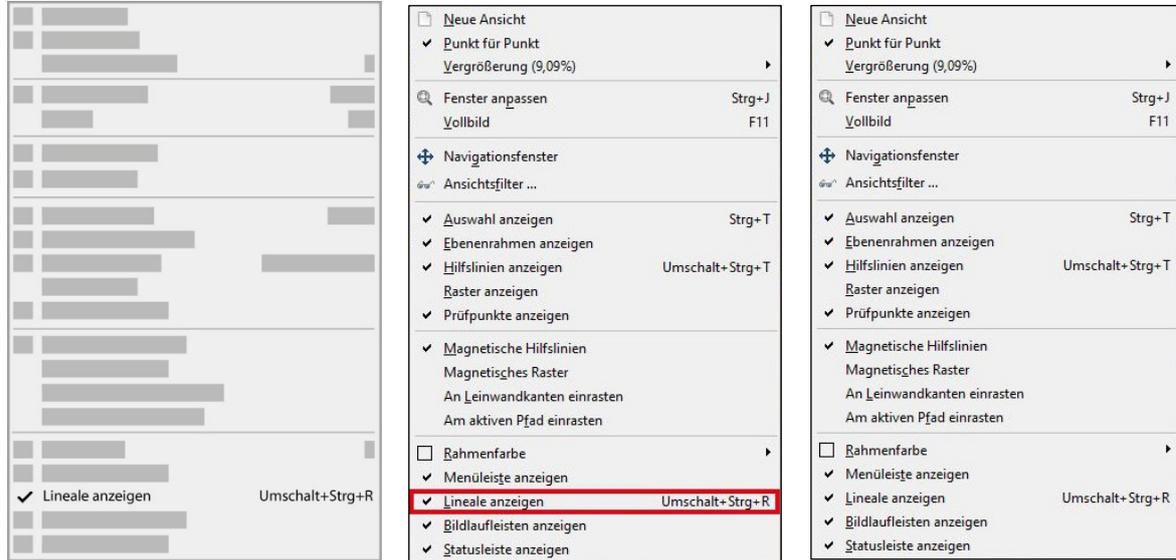


Abb. 27: Gezeigter Stimulus: Einfaches Listenmenü. Je nach Gruppenzugehörigkeit entweder SUI Variante (links), signalisierter Screenshot (Mitte) oder unbearbeitete Screenshot (rechts). (Quelle: Eigene Darstellungen nach Screenshots aus GIMP.)

Frage 02: Welchen Wert hat der Parameter Drehung?

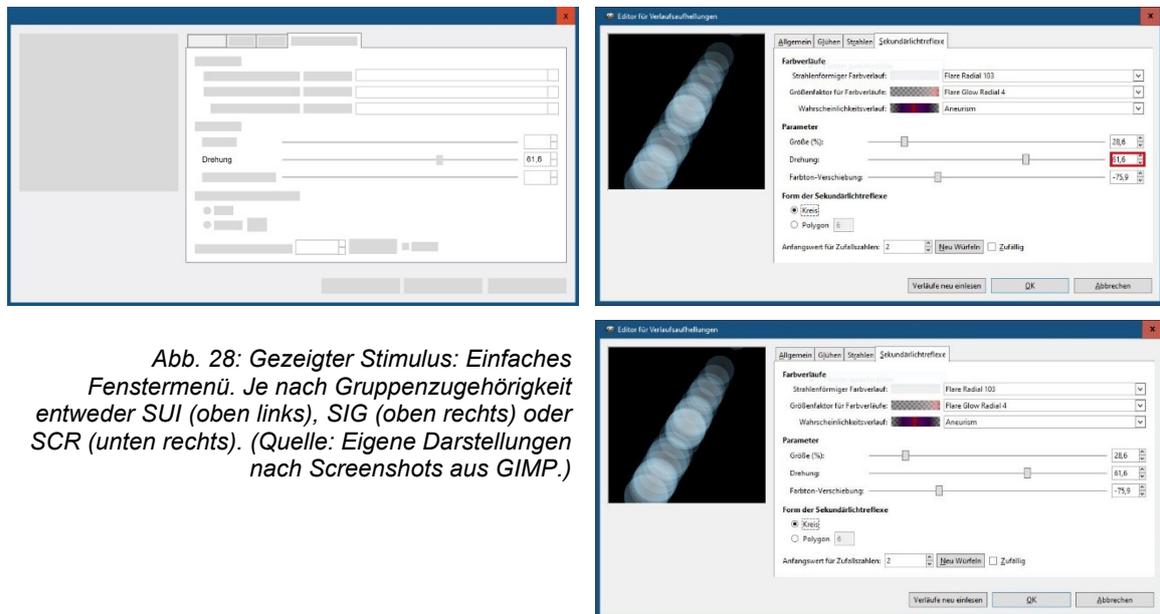


Abb. 28: Gezeigter Stimulus: Einfaches Fenstermenü. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellungen nach Screenshots aus GIMP.)

Frage 03: Welchen Wert hat die Deckkraft des Bildes?

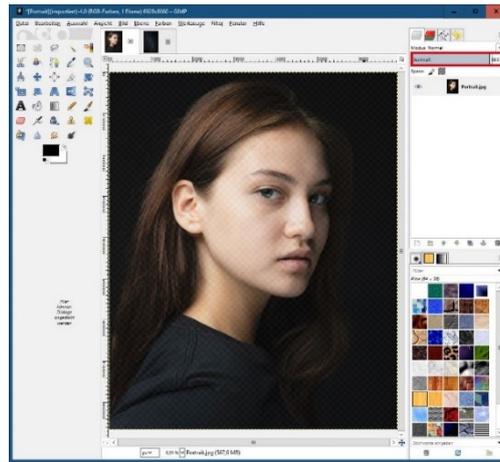
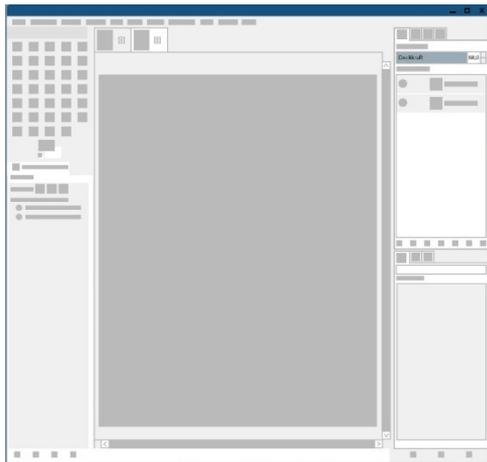
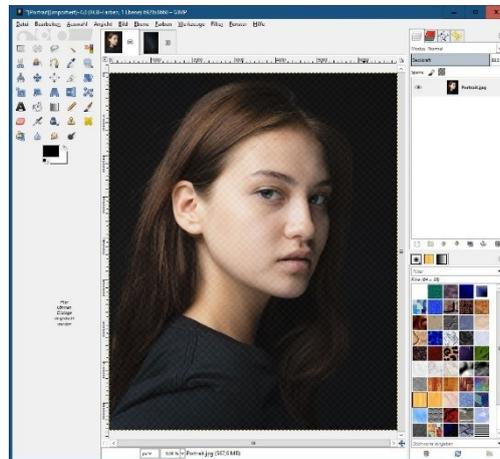


Abb. 29: Gezeigter Stimulus: Gesamte Softwareoberfläche. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.)



Frage 04: Welche Taste benötigen Sie für die Aktion: Oberste Ebene auswählen?

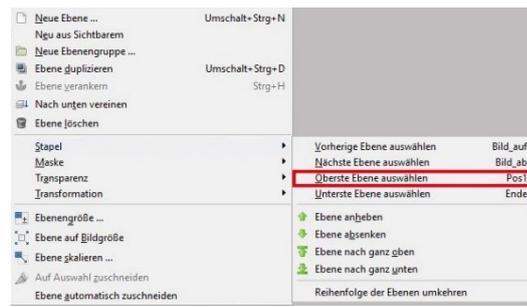
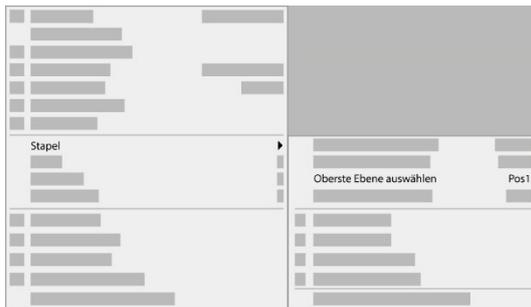


Abb. 30: Gezeigter Stimulus: Doppeltes Listenmenü. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.)



Frage 05: Welches Drucksimulationsprofil ist hier eingestellt?

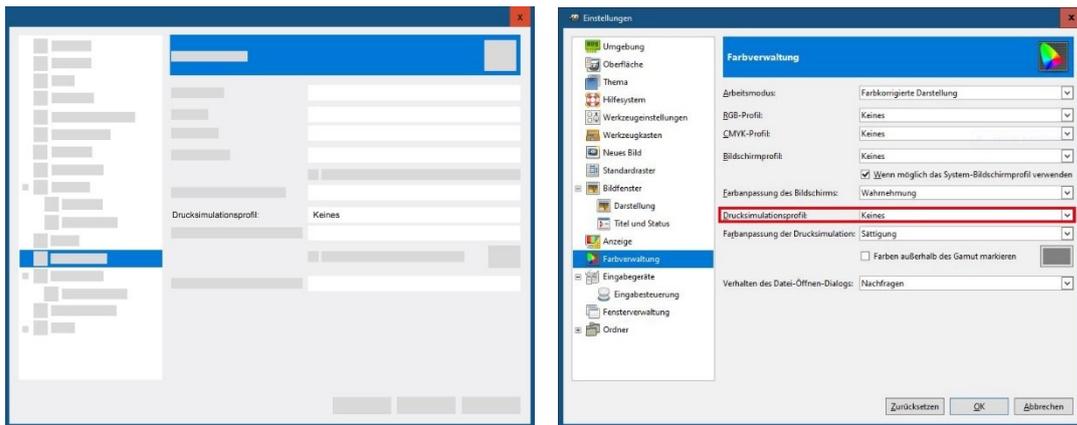
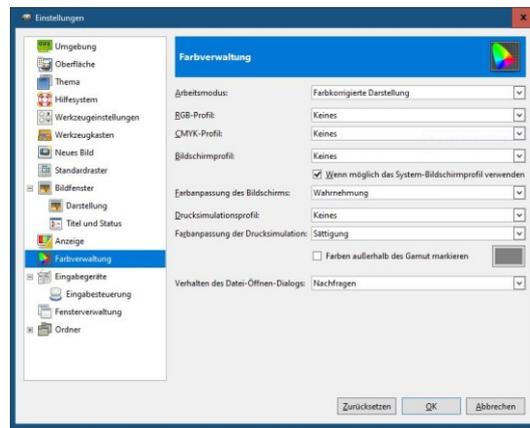


Abb. 31: Gezeigter Stimulus: Komplexes Fenstermenü. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.)



A02: Kalibrierungswerte der Proband\*innen

Proband*in	Left eye deviation (X/Y)	Right eye deviation (X/Y)
P01	0.91°/0.73°	0.34°/0.48°
P02	0.94°/0.68°	1.20°/0.93°
P03	0.31°/0.55°	0.20°/1.74°
P04	0.27°/0.26°	0.17°/0.21°
P05	0.07°/0.35°	0.4°/0.72°
P06	0.31°/0.21°	0.49°/0.30°
P07	0.67°/0.94°	0.27°/0.50°
P08	0.29°/0.28°	0.28°/0.36°
P09	0.19°/0.21°	0.34°/0.33°
P10	0.51°/0.76°	0.22°/0.31°
P11	0.30°/0.57°	0.31°/0.52°
P12	0.40°/0.73°	0.48°/1.03°
P13	0.33°/0.28°	0.64°/0.4°
P14	1.09°/0.76°	0.82°/0.66°
P15	0.26°/0.25°	0.34°/0.71°
P16	0.56°/0.50°	1.13°/0.82°
P17	0.27°/0.42°	0.30°/0.41°
P18	0.29°/0.12°	0.35°/0.27°
P19	0.32°/0.42°	0.31°/0.50°
P20	0.39°/0.35°	7.39°/1.98°
P21	0.16°/0.39°	0.16°/0.19°
P22	0.60°/0.49°	0.71°/0.58°
P23	0.29°/0.62°	0.29°/0.31°
P24	0.19°/0.18°	0.27°/0.32°

Tab. 23: Kalibrierungswerte der Proband\*innen. (Quelle: Eigene Darstellung.)

## A03: Orientierungstest: Ergänzende Visualisierungen

### A03.1: Reaktionszeiten

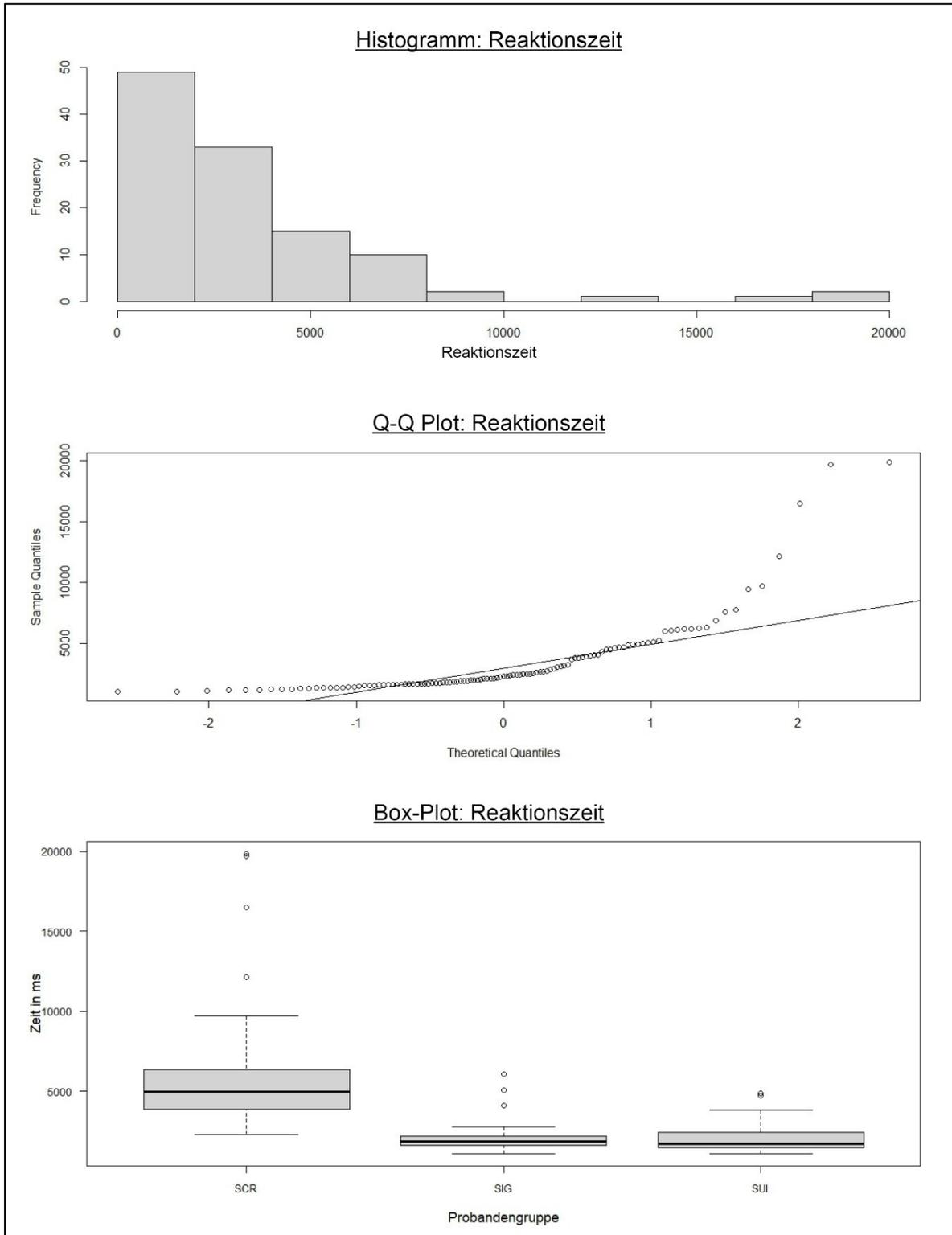


Abb. 32: Verteilungen der Mittelwerte der Reaktionszeiten im Orientierungstest. (Quelle: Eigene Darstellung.)

## A03.2: Fixationsanzahl

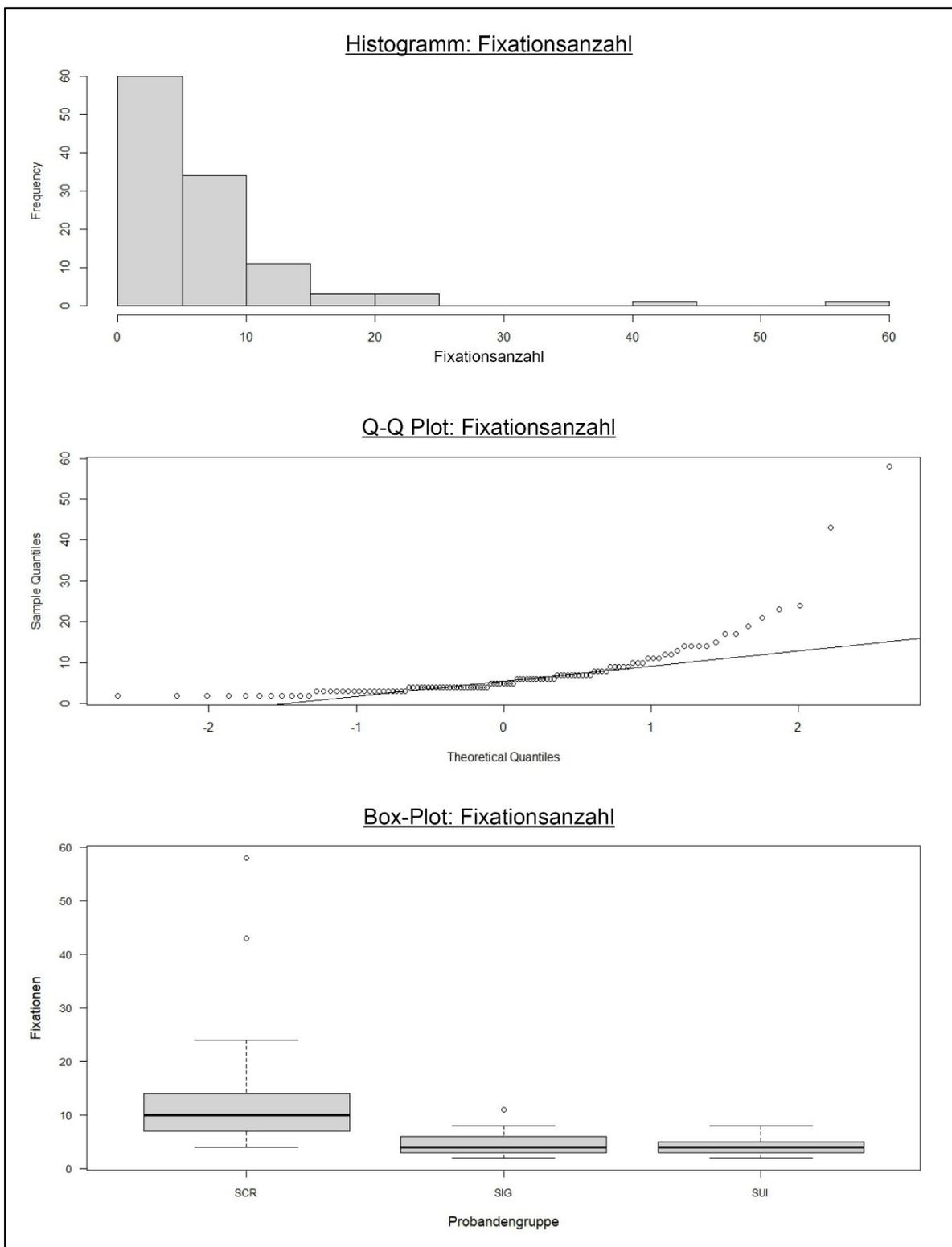


Abb. 33: Verteilungen der Mittelwerte der Fixationsanzahl im Orientierungstest. (Quelle: Eigene Darstellung.)

## A04: Orientierungstest: Heatmaps der Fixationshäufigkeit relevanter Bereiche

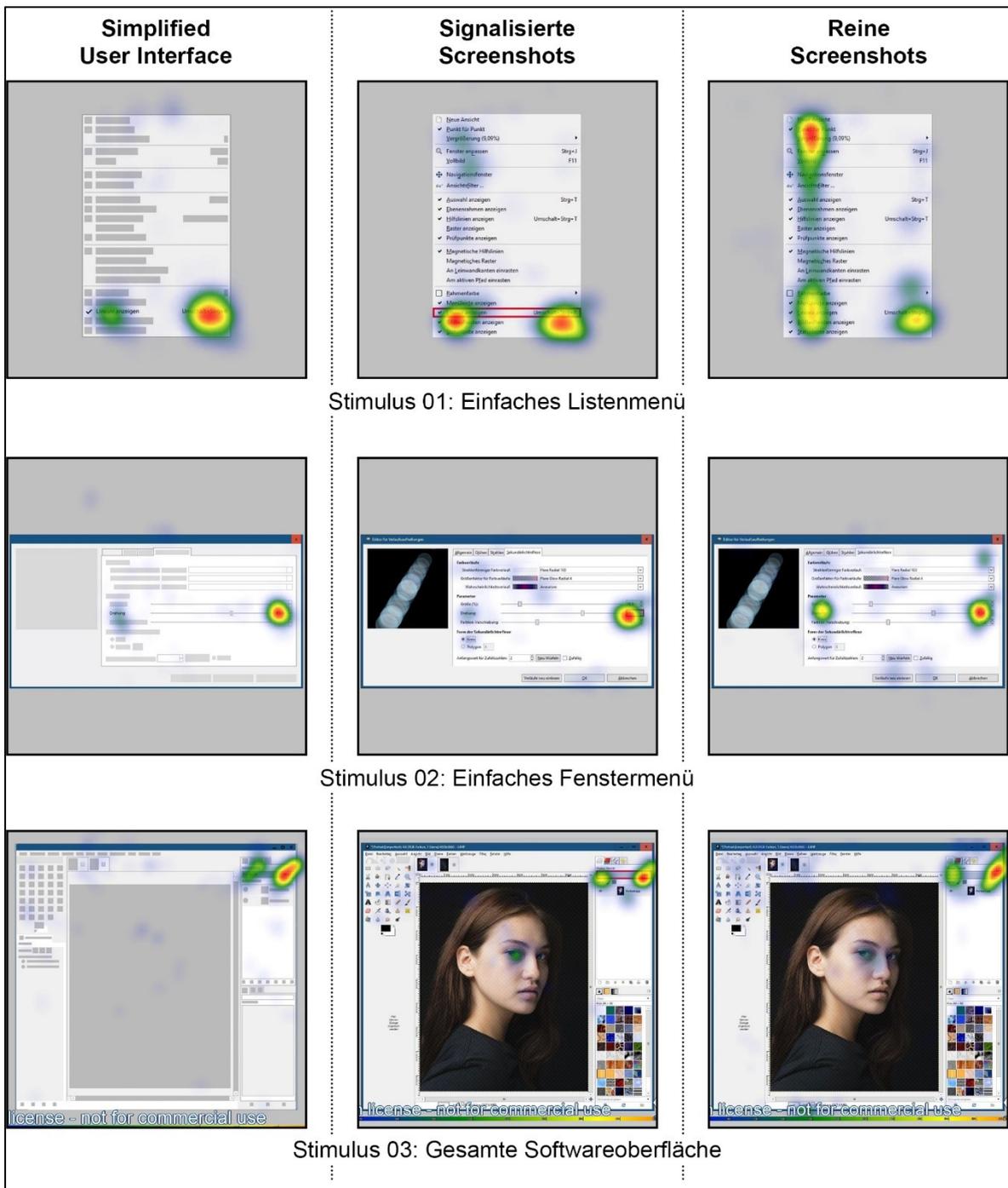


Abb. 34: Heat Maps zur Visualisierung der Fixationen im Orientierungstest im Vergleich der Gruppen – Stimuli 01 - 03. (Quelle: Eigene Darstellung aus Export aus SMI BeGaze.)

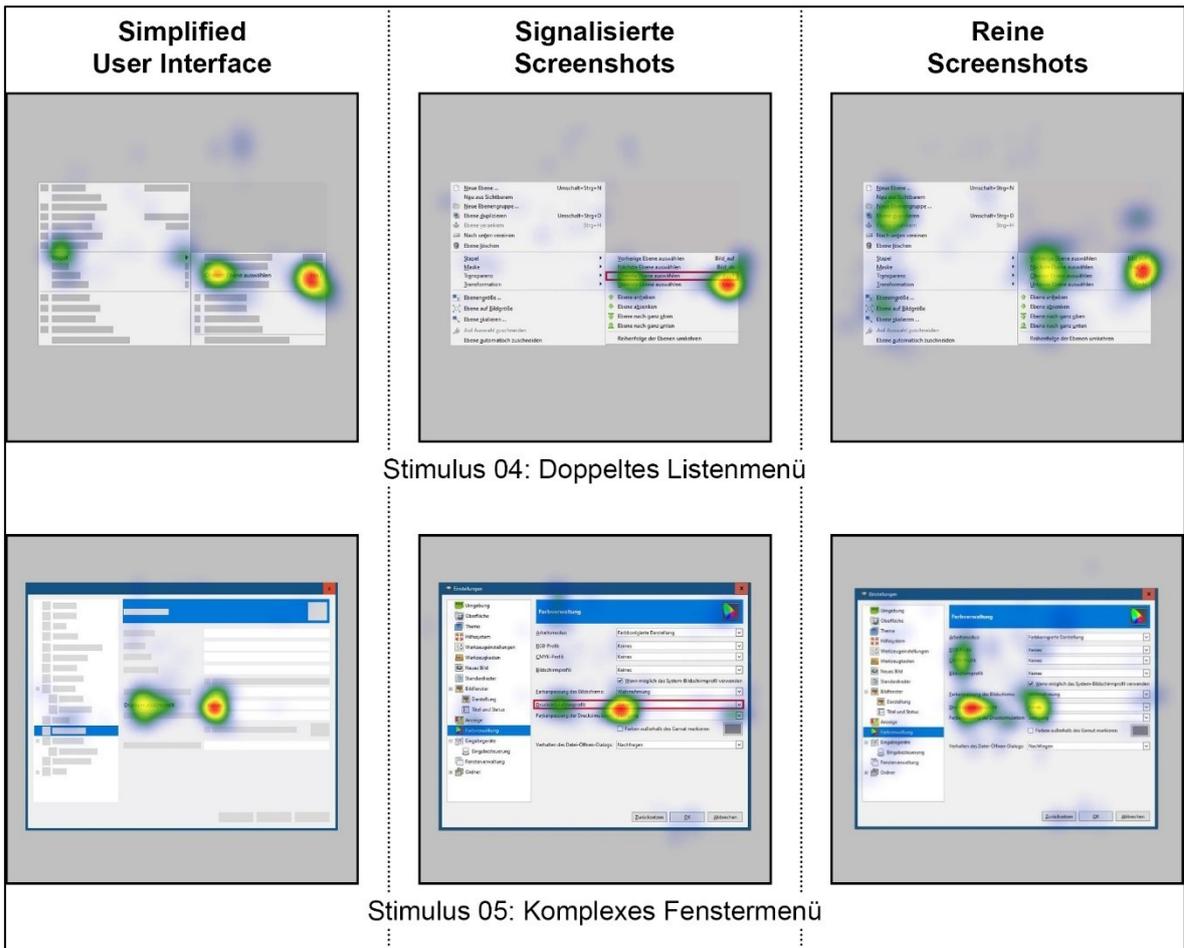


Abb. 35: Heat Maps zur Visualisierung der Fixationen im Orientierungstest im Vergleich der Gruppen – Stimuli 04 - 05. (Quelle: Eigene Darstellung aus Export aus SMI BeGaze)

## A05: Tutorial: Ergänzende Visualisierungen

### A05.1: Tutorial: Gesamte Bearbeitungszeit

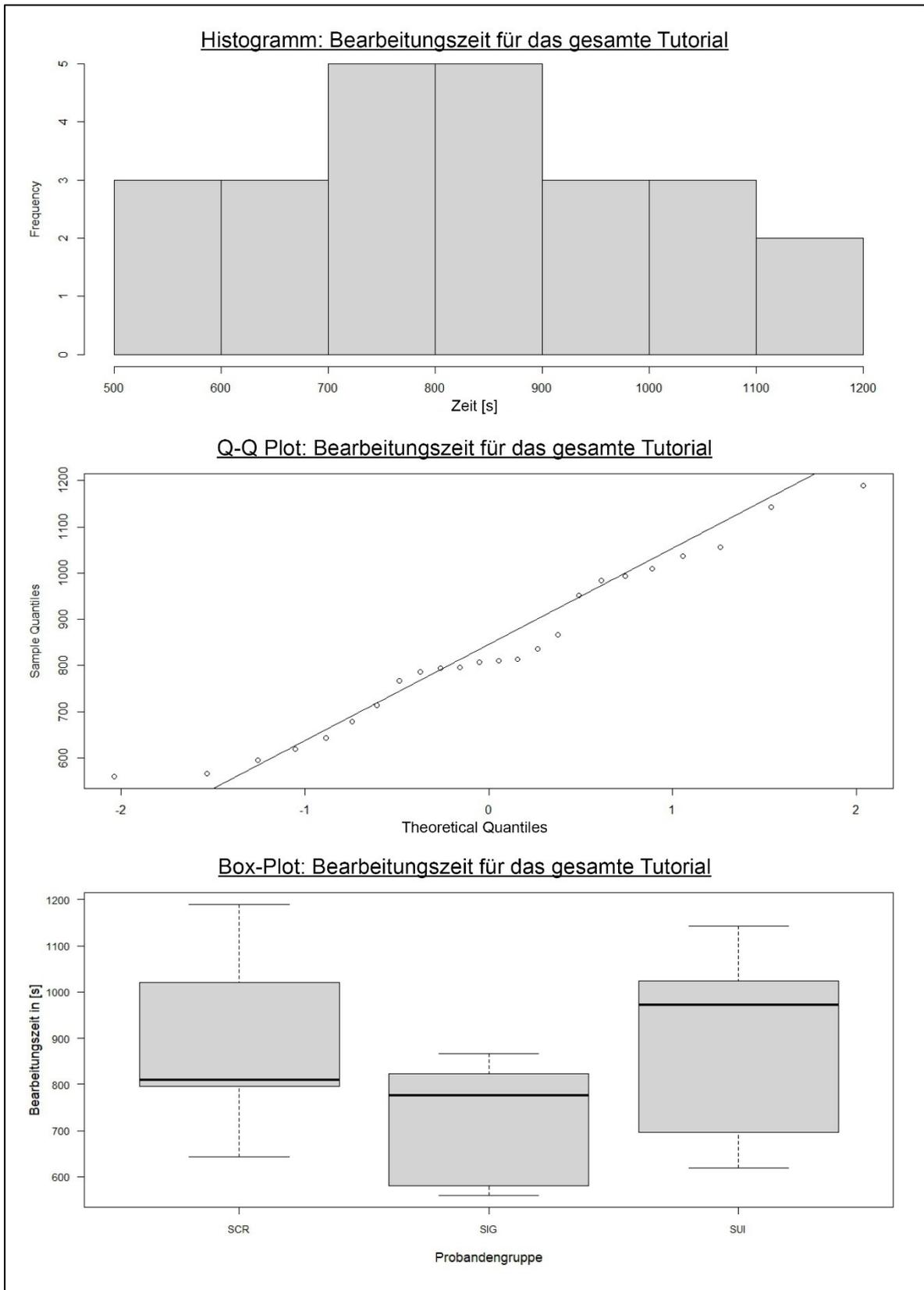


Abb. 36: Verteilungen der Mittelwerte für die Bearbeitungszeit des gesamten Tutorials.  
(Quelle: Eigene Darstellung.)

A05.2: Tutorial: Bearbeitungszeit pro Handlungsanweisung

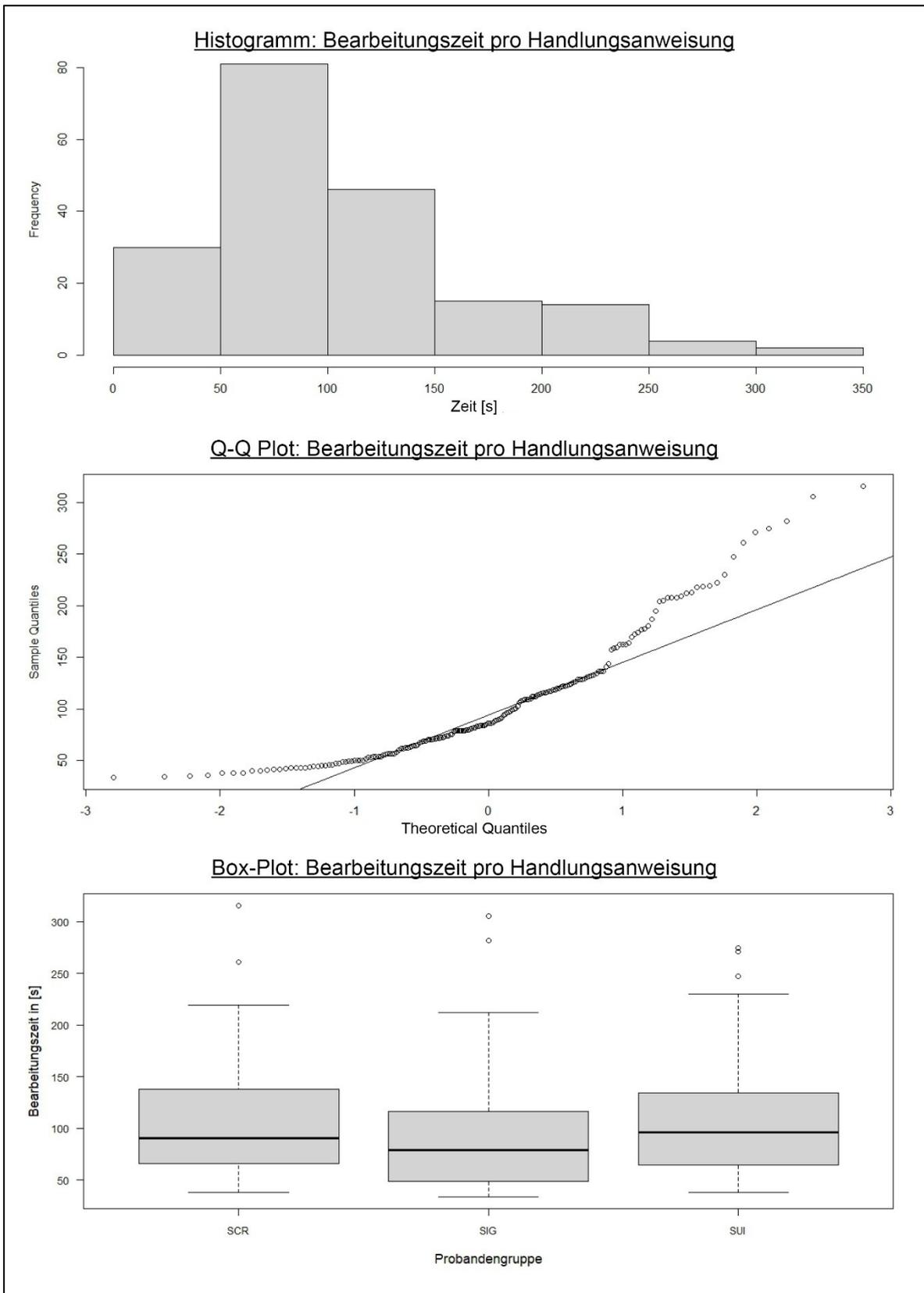


Abb. 37: Verteilungen der Mittelwerte für die Bearbeitungszeit pro Handlungsschritt.  
(Quelle: Eigene Darstellung.)

### A05.3: Tutorial: Korrekt gelöste Handlungsanweisungen

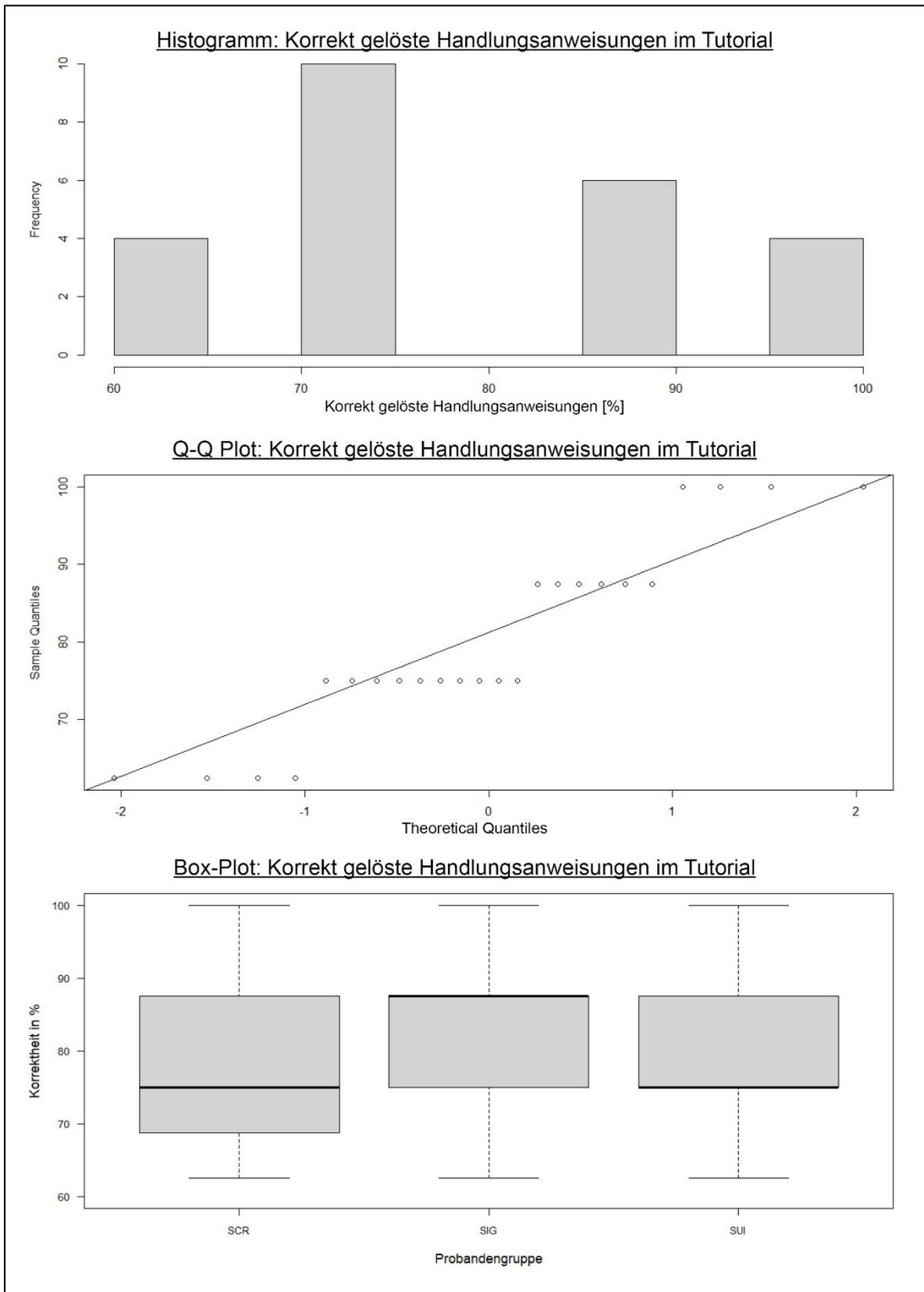


Abb. 38: Verteilungen der Mittelwerte für korrekt gelöste Handlungsanweisungen.  
(Quelle: Eigene Darstellung.)

A05.4: Tutorial: Korrekt gelöste Handlungsschritte

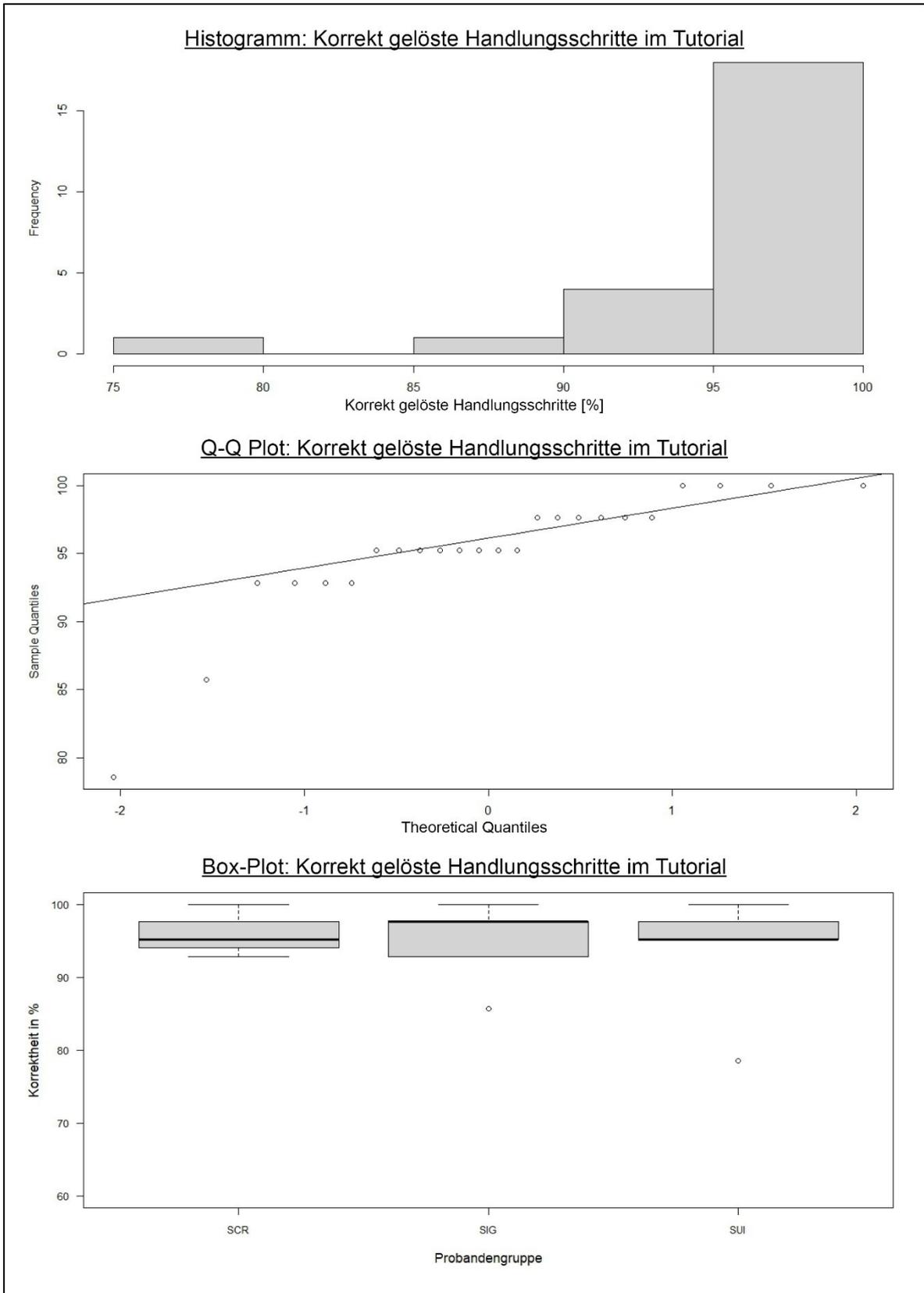


Abb. 39: Verteilungen der Mittelwerte für korrekt gelöste Handlungsschritte.  
 (Quelle: Eigene Darstellung.)

## A05.5: Tutorial: Dwell Time Tutorial vs. GIMP

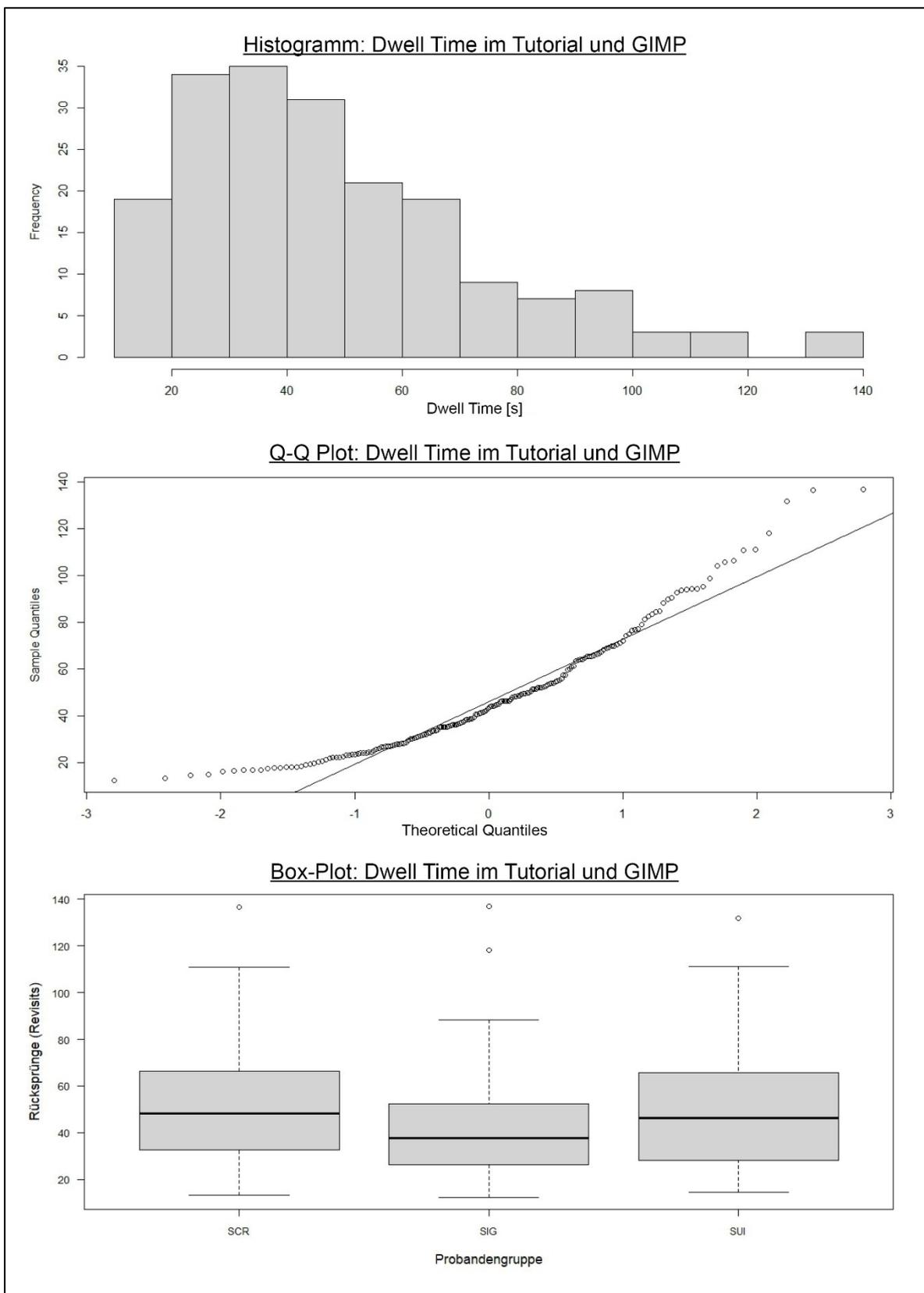


Abb. 40: Verteilungen der Mittelwerte der Dwell Time im Tutorial und GIMP. (Quelle: Eigene Darstellung.)

A05.6: Tutorial: Fixationsanzahl der Bilder

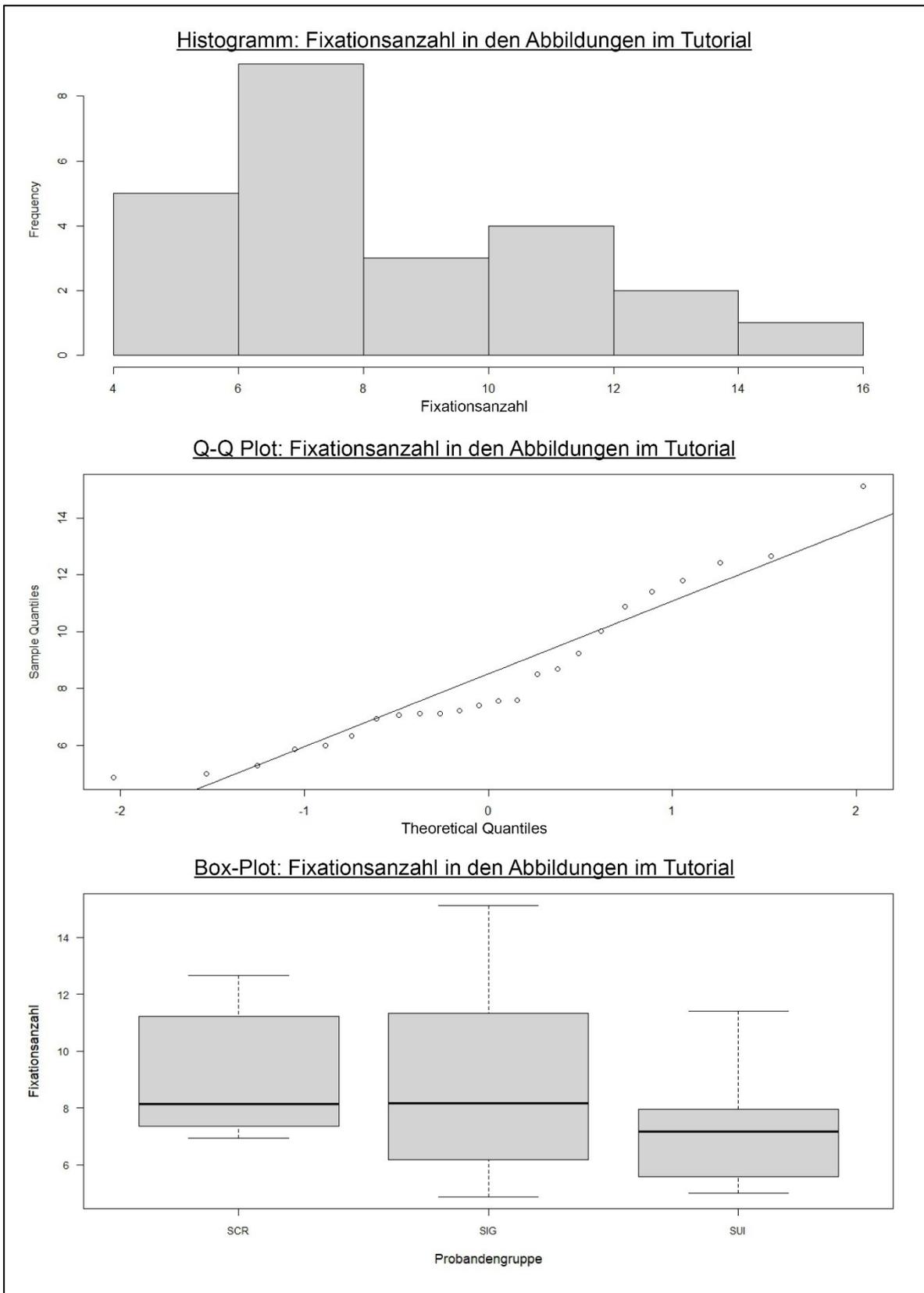


Abb. 41: Verteilungen der Mittelwerte der Fixationsanzahl in den Abbildungen im Tutorial. (Quelle: Eigene Darstellung.)

A05.7: Tutorial: Revisits zu den Abbildungen im Tutorial

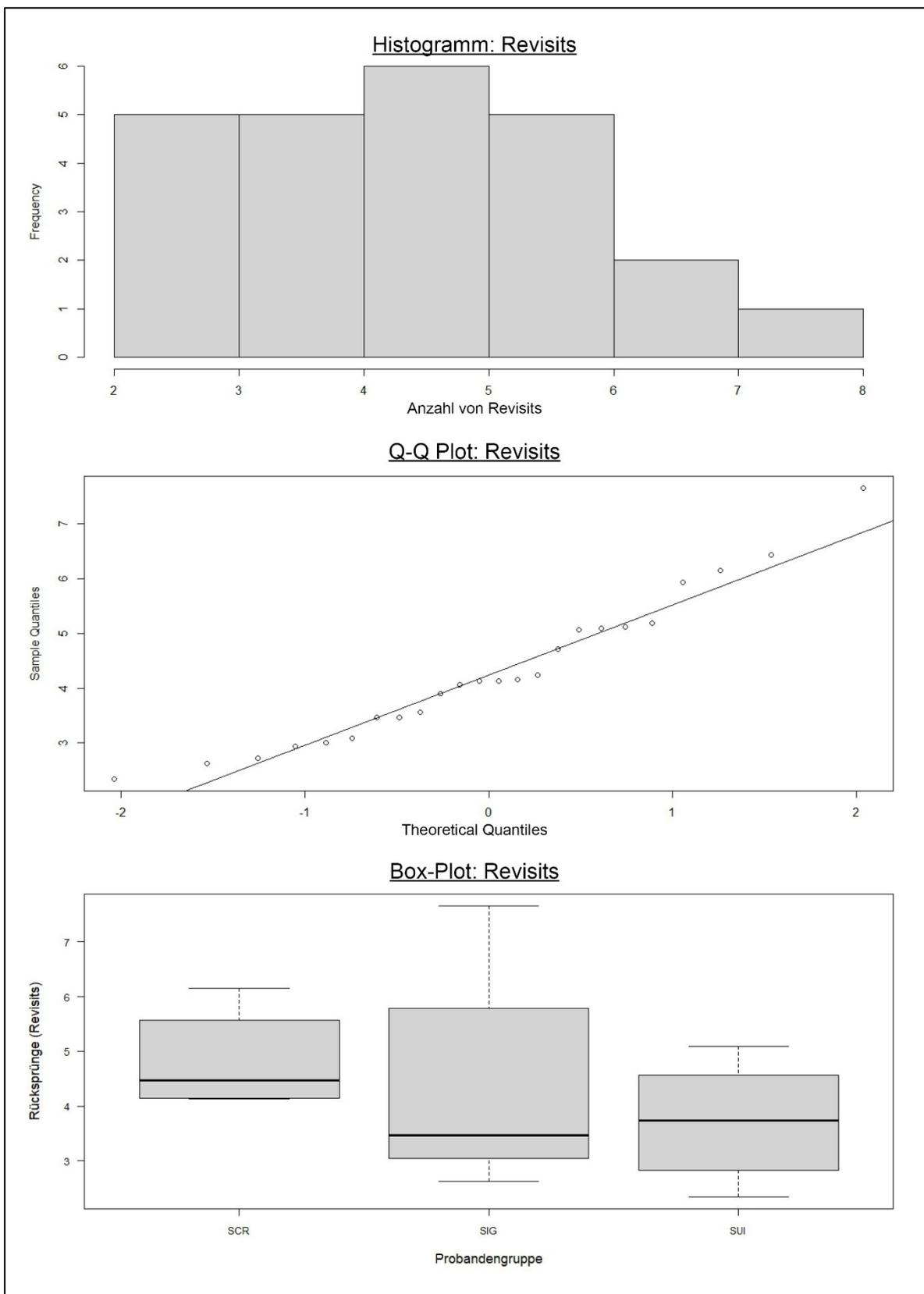


Abb. 42: Verteilungen der Mittelwerte der Revisits zu den Abbildungen im Tutorial. (Quelle: Eigene Darstellung.)

## Literaturverzeichnis

- Abel, S. (The Content Wrangler) (2019): The State of Technical Communication: 2019, <[http://public2.brighttalk.com/resource/core/217857/the-state-of-technicalcommunication\\_474463.pdf](http://public2.brighttalk.com/resource/core/217857/the-state-of-technicalcommunication_474463.pdf)> Letzter Zugriff: 13.01.2023
- Alexander, K. (2013): Kompendium der visuellen Information und Kommunikation. 2. überarb. und erw. Auflage, Springer Vieweg, Berlin/Heidelberg.
- Alexander, K., Schubert, A. & M. Meng (2016): Does detail matter? The effect of visual detail in line drawings on task execution, In: Information Design Journal, 22(1), John Benjamins Publishing Company, S. 49-61.
- Alpizar, D., Adesope, O. O. & R. M. Wong (2020): A meta-analysis of signaling principle in multimedia learning environments, In: Educational Technology Research and Development, 68, S. 2095-2119.
- Ayres, P. & J. Sweller (2014): The Split-Attention Principle in Multimedia Learning, In: Mayer, R. E. (Hrsg.): The Cambridge handbook of multimedia learning, 2. Edition, Cambridge University Press, New York. S. 206-226.
- Baddeley, A. (2002): Working Memory: Looking back and looking forward, In: Nature Reviews Neuroscience, 4, S. 829-839.
- Ballstaedt, S.-P. (1999): Die bildliche Darstellung von Handlungen in technischen Dokumenten, In: Schwender, C. (Hrsg.): Zur Geschichte der Gebrauchsanleitung. Theorien – Methoden – Fakten, Peter Lang, Europäischer Verlag der Wissenschaften, Frankfurt a.M.
- Bell, C. (1823): On the motions of the eye, in illustration of the uses of the muscles and nerves of the orbit, In: Philosophical Transactions, 113, The Royal Society Publishing, London, S. 166–186.
- Blake, C. (2013): Eye-Tracking: Grundlagen und Anwendungsfelder, In: Möhring, W. & D. Schlütz (Hrsg.) (2013): Handbuch standardisierter Erhebungsverfahren in der Kommunikationswissenschaft, Springer Fachmedien, Wiesbaden, S. 367 – 384.
- Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D. & T. Ertl (2014): State-of-the-Art of Visualization for Eye Tracking Data, Eurographics Conference on Visualization 2014, The Eurographics Association 2014, Swansea.
- Bleiel, N. A. & J. Carroll (2021): Minimalism at 30. tcworld magazine. <<https://www.tcworld.info/e-magazine/technical-writing/minimalism-at-30-1103/>> Letzter Zugriff: 22.01.2023.
- Boatman, A. (2020): Simplified User Interface: The Beginner's Guide, TechSmith, <[www.techsmith.com/blog/simplified-user-interface/#:~:text=A%20simplified%20user%20interface%20\(SUI,reduces%20them%20to%20simpler%20shapes](http://www.techsmith.com/blog/simplified-user-interface/#:~:text=A%20simplified%20user%20interface%20(SUI,reduces%20them%20to%20simpler%20shapes)>, Letzter Zugriff: 28.10.2022.
- Bollen, A. (2019): Simplified graphics: Meet the new design style for technical communication, Ursprünglich publiziert in: ISTC Communicator Summer 2019, SUI: revolutionising graphic effects.
- Bollen, A. (2020): Why, When & How of Simplified User Interfaces (SUI) with Anton Bollen, Video-Interview in: The Visual Lounge by TechSmith, < <https://www.youtube.com/watch?v=2L4HwZj8sk4>> Letzter Zugriff: 18.11.2022.
- Bollen, A. & L. Saremba (2020a): Einfach gestalten, In: technische kommunikation, 05/20, Gesellschaft für Technische Kommunikation, tekomp Deutschland e. V., S. 11-16.

- Bollen, A. & L. Saremba (2020b): Your brain doesn't like complicated. Keep your graphics simple, Vortrag in: tc World conference 2020. The online event in TC, 02.11.2020, <[https://www.youtube.com/watch?v=t8K9G\\_gRx6Y](https://www.youtube.com/watch?v=t8K9G_gRx6Y)> , Letzter Zugriff: 26.11.2022.
- Bollen, A. & L. Saremba (2021): Simplified graphics and screenshots in software documentation, Tc world magazine Februar 2021, < <https://www.tcworld.info/e-magazine/technical-writing/simplified-graphics-and-screenshots-in-software-documentation-1102/>>, Letzter Zugriff: 14.01.2023.
- Carter, B. T. & S. G. Luke (2020): Best practices in eye tracking research, In: International Journal of Psychophysiology, 155, Elsevier, Provo, S. 49- 62.
- Clark, J. M. & A. Paivio (1991): Dual Coding Theory and Education, In: Educational Psychology Review, 3(3), S. 149-210.
- Craik F. I. M. & E. Tulving (1975): Depth of processing and the retention of words, In: Journal of Experimental Psychology: General, 104, S. 268-294.
- Cohen, J. (1992): A power primer, In: Psychological Bulletin, 112(1), S. 155–159.
- Delabarre, E. B. (1898): A method of recording eye-movements, In: American Journal of Psychology, 9(4), University of Illinois Press, S. 572-574.
- Dodge, R. & T. S. Cline (1901): The Angle Velocity of Eye Movement, In: Psychological Review, 8(2), S. 145-157.
- Dönmez, M., Doğan, S. & E. Baran (2018): How Signaling Principle Affects Learning: An Eye Tracking Study, In: Mersin University Faculty of Education Journal, 14(2), S. 700-713.
- Duchowski, A. (2007): Eye Tracking Methodology: Theory and practice, Springer London Limited, London.
- Duggan, G. B. & S. J. Payne (2001): Interleaving reading and acting while following procedural instructions, In: Journal of Experimental Psychology: Applied, 7(4), S. 297-307.
- Duden (2022): Screenshot, <<https://www.duden.de/rechtschreibung/Screenshot>>, Letzter Zugriff: 12.11.2022.
- Dwyer, F. M. (1978): Strategies for improving visual learning, In: State College Pennsylvania: Learning Services.
- Eberleh, E. (1990): Komplementarität von Text und Bild, In: Becker, T. et al.: Sprache und Technik. Das Gestalten verständlicher Texte. Konzepte, Probleme, Erfahrungen, Alano Rader Publikationen, Aachen.
- Farkas, D. K. (1999): The logical and rhetorical construction of procedural discourse, In: Technical Communication, 46, S. 42-54.
- Findlay, J. M. & I. D. Gilchrist (2003): Active Vision. The Psychology of Looking and Seeing, Oxford University Press, New York.
- Forsythe, A. (2009): Visual Complexity: Is That All There Is?, In: Harris, D. (2009): Engineering Psychology and Cognitive Ergonomics, HCII 2009 LNAI 5639, San Diego, S. 158-166.
- Friedewald, M. (2000). Die geistigen und technischen Wurzeln des Personal Computers, In: Naturwissenschaftliche Rundschau, 53(4), Wissenschaftliche Verlagsgesellschaft (heute Konradin Mediengruppe), Stuttgart.
- Garcia, M., Badre, A.N. & J. T. Stasko, (1994): Development and validation of icons varying in their abstractness, In: Interacting with Computers, 6(2), S 191–211.

- Garner, R., Brown, R. Sanders, S. & D. Menke (1992): Seductive details and learning from text, In: Renninger, K.A., Hidi, S. & A. Krapp (Hrsg.): The role of interest in learning and development, Hillsdale, NJ: Erlbaum, S. 239-254.
- Garner, R., Alexander, P. A., Gillingham, M. G., Kulikowich, J. M. & R. Brown (1991): Interest and learning from text, In: American Educational Research Journal, 28, S. 643-659.
- Gellevij, M., & H. van der Meij, (2004): Empirical proof for presenting screen captures in software documentation, In: Technical Communication, 51, S. 224–258.
- Gellevij, M., van der Meij, H., de Jong, T. & J. M. Pieters (1999): The Effects of Screen Captures in Manuals: A Textual and Two Visual Manuals Compared, In: IEEE Transactions on professional communication, 42(2), S. 77-91.
- Gellevij, M., van der Meij, H., de Jong, T. & J. M. Pieters (2002): Multimodal versus Unimodal Instructions in a Complex Learning Context, In: Journal of experimental education, 70(3), S. 215-243.
- Geise, S. (2011): Eyetracking in der Kommunikations- und Medienwissenschaft: Theorie, Methode und kritische Reflexion, Studies in Communication / Media, 0. Jahrgang, Nomos, Baden-Baden, S. 149 - 263
- Geiselman, R. E., Landee, B. M., & F. G. Christen (1982): Perceptual discriminability as a basis for selecting graphic symbols, In: Human Factors, 24, S. 329-337.
- Ginns, P. (2006): Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects, In: Learning and Instruction, 16(6), S. 511-525.
- Goll, J. (2019): Keep it simple, stupid, In: Entwurfsprinzipien zur Reduktion der Komplexität., In: Entwurfsprinzipien und Konstruktionskonzepte der Softwaretechnik. Strategien für schwach gekoppelte, korrekte und stabile Software, 2. aktual. Auflage, Springer Vieweg, Wiesbaden.
- Hammoud, R. I. (2008). *Passive Eye Monitoring*. Algorithms, Applications and Experiments, Springer, Berlin.
- Harp, S. F. & R. E. Mayer (1997): The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest, In: Journal of Educational Psychology, 89, S. 92-102.
- Harp, S. F. & R. E. Mayer (1998): How seductive details do their damage: A theory of cognitive interest in science learning, In: Journal of Educational Psychology, 90, S. 414-434.
- Heaps, C. & S. Handel (1999): Similarity and Features of Natural Textures, In: Journal of Experimental Psychology: Human Perception and Performance, 62(8), S. 299-320.
- Holmqvist, K. & R. Andersson (2017): Eye tracking: A comprehensive guide to methods, paradigms and measures, 2. Edition, Lund Eye-Tracking Research Institute, Lund.
- Horton, W. (1993): Dump the the dumb screen dumps, In: Technical Communication, 40, S. 146-148.
- Interaction Design Foundation (2021): KISS (Keep it Simple, Stupid) – A Design Principle, <<https://www.interaction-design.org/literature/article/kiss-keep-it-simple-stupid-a-design-principle>> , Letzter Zugriff: 17.01.2023.
- Irrazabal, N., Saux, G. & D. Burin (2016): Procedural Multimedia Presentations: The Effect of Working Memory and Task Complexity on Instruction Time and Assembly Accuracy, In: Applied Cognitive Psychology, 30(6), S. 1052-1060.

- Jeung, H., Chandler, P. & J. Sweller (1997): The role of visual indicators in dual sensory mode instruction, In: *Educational Psychology*, 17, S. 329-433.
- Kalyuga, S. (2014): The expertise reversal principle in multimedia learning, In: Mayer, R. E. (Hrsg.): *The Cambridge handbook of multimedia learning*, 2. Edition, Cambridge University Press, New York, S. 576-597.
- Kalyuga, S., Chandler, P. & J. Sweller (1998): Levels of expertise in instructional design, In: *Human Factors*, 40(1), S. 1-17.
- Kalyuga, S., Chandler, P. & J. Sweller (1999): Managing Splitt-attention and Redundancy in Multimedia Instructions, In: *Applied Cognitive Psychology*, 13, S. 351-371.
- Kalyuga, S., Chandler, P. & J. Sweller (2000): Incorporating learner expertise into the design of multimedia instruction, In: *Journal of Educational Psychology*, 92, S. 126-136.
- Kalyuga, S., Chandler, P. & J. Sweller (2003): The Expertise Reversal Effect, In: *Educational Psychologist*, 38(1), S. 23-31.
- Kalyuga, S. & A. Renkl (2010): Expertise reversal effect and its instructional implications: introduction to the special issue, In: *Instructional Science*, 38, S. 209-215.
- Koffka, K. (1935): *Principles of Gestalt Psychology*, Harcourt, Brace and Company, New York.
- Lang, G. (2016): The Relative Efficacy of Video and Text Tutorials in Online Computing Education, In: *Information Systems Education Journal*, 14(5), S. 33.
- Lindén, E. (2021): Calibration in deep-learning eye tracking, Dissertation, KTH Royal Institute of Technology, Stockholm.
- Maestre, L. S. (2012): Student preference for tutorial design: a usability study, In: *Reference Service Review*, 40(2), S. 258 - 276.
- Mayer, R. E. (1989): Systematic thinking fostered by illustrations in scientific text, In: *Journal of Educational Psychology*, 81, S. 240-246.
- Mayer, R. E. (2005): *Cognitive Theory of Multimedia Learning*, In: Mayer, R. E. (Hrsg.): *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press, New York, S. 31-48.
- Mayer, R. E. (2012): *Multimedia Learning*, 2. Auflage, Cambridge University Press, New York.
- Mayer, R. E. (2021): *Multimedia Learning*, 3. Auflage, Cambridge University Press, Cambridge / New York.
- Mayer, R. E. & R. B. Anderson (1991): Animations need narrations: An experimental test of a dual coding hypothesis, In: *Journal of Educational Psychology*, 83, S. 484-490.
- Mayer, R. E. & R. B. Anderson (1992): The instructive animation: Helping students build connections between words and pictures in multimedia learning, In: *Journal of Educational Psychology*, 84, S. 444-452.
- Mayer, R. E., Bove, W., Bryman, A., Mars, R. & L. Tapangco (1996): When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons, In: *Journal of Educational Psychology*, 88, S. 64-73.
- Mayer, R. E. & J. K. Gallini (1990): When is an illustration worth ten thousand words?, In: *Journal of Educational Psychology*, 82, S. 715-726.
- Mayer, R. E., Griffith, E., Naftaly, I. & D. Rothman (2008): Increased interestingness of extraneous details lead to decreased learning, In: *Journal of Experimental Psychology: Applied*, 14, S. 329-339.

- Mayer, R. E., Heiser, J. & S. Lonn (2001): Cognitive constraints on multimedia learning: When presenting more materials results in less understandings, In: *Journal of Educational Psychology*, 93, S. 187-198.
- Mayer, R. E. & R. Moreno (2002): Animation as an aid to Multimedia Learning, In: *Educational Psychology Review*, 14, S. 87-99.
- Meng, M. (2019): Effects of Visual Signaling in Screenshots: An Eye Tracking Study, In: *Technical Communication*, 66(4), S. 396-411.
- Meng, M. (2020): Using eye tracking to study information selection and use: A research perspective, In: Ahrens, B.; Beaton-Thome, M.; Krein-Kühle, M.; Krüger, R.; Link, L. & U. Wienen (Hrsg.): *Interdependenzen und Innovationen in Translation und Fachkommunikation*, Frank & Timme GmbH, Berlin, S. 93-122.
- Meng, M. & J. Ullmann (2015): Erfolgreich mit Screenshots, In: *technische kommunikation*, Gesellschaft für Technische Kommunikation, 15, S. 41-47.
- Miller, G. A. (1956): The magical number seven, plus or minus two: Some limits on our capacity for processing information, In: *Psychological Review*, 63, S. 81-97.
- Mitschke, U. (2017): Screenshots mit Effekt, In: *technische kommunikation*, 2/17, Gesellschaft für Technische Kommunikation, Stuttgart.
- Moreno, R. & S. Abercrombie (2010): Promoting awareness of learner diversity in prospective teachers: Signaling individual and group differences within virtual classroom case, In: *Journal of Technology and Teacher Education*, 18(1), S. 111-130.
- Moreno, R. & R. E. Mayer (1999): Multimedia-supported metaphors for meaning making in mathematics, In: *Cognition and Instruction*, 17, S. 215-248.
- Moreno, R. & R. E. Mayer (2007): Interactive Multimodal Learning Environments. Special Issue on Interactive Learning Environments: Contemporary Issues and Trends, In: *Educational Psychology Review*, 19, S. 309-326.
- Paivio, A. (1990): *Mental representations: A dual coding approach*, Oxford University Press, New York.
- Pedzek, K. & H.-C. Chen (1982): Developmental differences in the role of detail in picture recognition memory, In: *Journal of Experimental Child Psychology*, 33, S. 207-215.
- Pedzek, K., Maki, R., Valencia-Laver, D., Whetstone, T., Stoeckert, J. & T. Dougherty (1988): Picture memory: Recognizing added and deleted details, In: *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), S. 468-476.
- Piehl, J. (2002): *Gebrauchsanleitungen optimal gestalten. Über sinnvolle und verständliche Gestaltung*, x.media.press, Springer-Verlag, Berlin/Heidelberg.
- Ponce, H. & R. E. Mayer (2014): An eye-movement analysis of highlighting and graphic organizer study aids for learning from expository text, In: *Computers in Human Behavior*, 41, S. 21-32.
- Rao, A. R. & G. L. Lohse (1993): Identifying High Level Features of Texture Perception, In: *CVGIP: Graphical Models and Image Processing*, 55(3), S. 218-233.
- Rayner, K. (1998): Eye Movements in Reading and Information Processing: 20 Years of Research, In: *Psychological Bulletin*, 124(3), American Psychological Association, Amherst, S. 372-422.
- Reder, L. M. & J. R. Anderson (1980): A comparison of texts and their summaries: Memorial consequences, In: *Journal of Verbal Learning & Verbal Behavior*, 19, S. 121-134.

- Redish, J. C. (1989): Reading to Learn to Do, In: IEEE Transactions on Professional Communication, 32(4), S. 289-293.
- Rey, G. D. (2007): Lernen mit Multimedia. Die Gestaltung interaktiver Animationen, Dissertation, Universität Trier.
- Rey, G. D. (2014): Seductive details and attention distraction – An eye tracker experiment, In: Computers in Human Behavior, 32, S. 133-144.
- Richter, J., Scheiter, K., & A. Eitel (2016): Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis, In: Educational Research Review, 17, S. 19-36
- Ries, F. (2020): Visual Complexity in Human-Machine Interaction, Dissertation, Karlsruher Institut für Technologie.
- Rötting, M. (2001): Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen, In: Schriftenreihe Rationalisierung und Humanisierung, 34, Shaker Verlag, Aachen.
- Sanchez, C. A. & J. Wiley (2006): An examination of the seductive details effect in terms of working memory capacity, In: Memory & Cognition, 34, S. 344-355.
- Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V. & D. Leutner (2014): Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect, In: Contemporary Educational Psychology, 39, S. 275-286.
- Schneider, S., Beege, M., Nebel, S. & G. D. Rey (2018): A meta-analysis of how signaling affects learning with media, In: Educational Research Review, 23, S. 1-24.
- Schnotz, W. & M. Bannert (2003): Construction and interference in learning from multiple representations, In: Learning and Instruction, 13, S. 141-156.
- Schraw, G. (1998): Processing and recall differences among seductive details, In: Journal of Educational Psychology, 90, S. 3-12.
- Schraw, G. & S. Lehman (2001): Situational Interest: A Review of the Literature and Directions for Future Research, In: Educational Psychology Review, 13(1), S. 23-52.
- Schumacher, P. (2007): Creating effective Illustrations for pictorial assembly instructions, In: Information Design Journal, 15(2), S. 97-109.
- Schwender, C. (Hrsg.) (1999): Die Gebrauchsanleitung – eine Anleitung zum Gebrauch. – In: Schwender, C. (Hrsg.): Zur Geschichte der Gebrauchsanleitung. Theorien – Methoden – Fakten, Peter Lang, Europäischer Verlag der Wissenschaften, Frankfurt a.M.
- Semmelmann, K. & S. Weigelt (2018): Online webcam-based eye tracking in cognitive science: A first look, In: Behavior Research Methods, 50, S. 451-465.
- Slivka, M. (2020): Predictive Eye Tracking vs Regular Eye Tracking, <<https://attentioninsight.com/eye-tracking-vs-predictive-eye-tracking/>>, Letzter Zugriff: 18.01.2023
- SMI (2014): BeGaze Manual, Version 3.4, SensoMotoricInstruments GmbH, o.O.
- Snodgrass, J. G. & M. Vanderwart (1980): A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity and Visual Complexity, In: Journal of Experimental Psychology: Human Learning and Memory, 6(2), S. 174-215

- Stegelin, S. (2018): Graphic Content Warning: The Pros, Cons, and Alternatives to Screenshots. Write the Docs Portland 2018, Konferenz, <<https://www.writethedocs.org/videos/portland/2018/graphic-content-warning-the-pros-cons-and-alternatives-to-screenshots-steve-stegelin/>>, Letzter Zugriff: 12.12.2022.
- Sticht, T. G. & D. Welty-Zapf (1976): Reading and Readability Research in the Armed Services, In: Human Resources Research Organization, Alexandria.
- Sticht, T. (1985): Understanding Readers and Their Uses of Text, In: Duffy, T. M. & R. Waller (Hrsg.): Designing Usable Texts, Academic Press, Orlando.
- Sung, E. & R. E. Mayer (2012): When graphics improve liking but not learning from online lessons, In: Computers in Human Behavior, 28, S. 1618-1625.
- Sweller, J. (1988): Cognitive Load During Problem Solving: Effects on Learning, In: Cognitive Science, 12, S. 257-285.
- Sweller, J. (2005): Implications if cognitive load theory for multimedia learning, In: Mayer, R. (Hrsg.): The Cambridge handbook of multimedia learning, Cambridge University Press, New York, S. 19-33.
- Sweller, J. (2010): Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load, In: Educational Psychology Review, 22, S. 123-138.
- Sweller, J. & P. Chandler (1994): Why some material is difficult to learn, In: Cognition and Instruction, 12(3), S. 185-233.
- Sweller, J., van Merriënboer, J. J. G. & F. G. W. C. Paas (1998): Cognitive Architecture and Instructional Design, In: Educational Psychology Review, 10(3), S. 251-296.
- TechSmith (o.D.): Aufnehmen. Aufbereiten. Austauschen. Bildschirmaufnahmen und mehr für bessere Kommunikation, < <https://www.techsmith.de/snagit.html> >, Letzter Zugriff: 13.01.2023.
- van Gog, T. (2014): The signaling (or Cueing) principle in multimedia learning, In: R. E. Mayer (Hrsg.): The Cambridge Handbook of Multimedia Learning, Cambridge University Press, 2. Ausgabe, New York, S. 263-278.
- van der Meij, H. (1996). A closer look at visual manuals, In: Journal of Technical Writing and Communication, 26, S. 371-383.
- van der Meij, H. (1998): Optimizing the joint handling of manual and screen, In: Carroll, J. M. (Hrsg.): Minimalism beyond the Nurnberg Funnel, MIT Press, Cambridge, S. 275-309.
- van der Meij, H. (2000): The role and design of screen images in software documentation, In: Journal of Computer Assisted Learning, 16, S. 294-306.
- van der Meij, H. & J. M. Carroll (1998): Principles and Heuristics for Designing Minimalist Instruction, In: Carroll, J. M. (Hrsg.): Minimalism Beyond the Nurnberg Funnel, MIT Press, Cambridge.
- van der Meij, H. & M. Gellevis (1998): Screen Captures in Software Documentation, In: Technical Communication, 4/98.
- van der Meij, H., Karreman, J. & M. Steehouder (2009): Three Decades of Research and Professional Practice on Printed Software Tutorials for Novices, In: Technical Communication, 56(3), S. 265-292.
- van der Meij, H. & J. van der Meij (2014): A comparison of paper-based and video tutorials for software learning, In: Computers & Education, 78, S. 150-159.

- van Genuchten E., van Hooijdonk, C., Schüler, A. & K. Scheiter (2014): The Role of Working Memory when “Learning How” with Multimedia Learning Material, In: Applied Cognitive Psychology, 28(3), S. 327-335.
- Vermeulen, F. & H. van der Meij (2019): Using the Principles of Minimalism to Create Better User Support, with Hans van der Meij, Instrktive Podcast, < <https://instrktiv.com/en/blog/podcast/minimalism/>>, Letzter Zugriff: 22.01.2023.
- Wade, N. J. & B. W. Tatler (2005): The Moving Tablet of the Eye. The Origins of Modern Eye Movement Research, Oxford University Press, New York.
- Wade, S. E. & B. Adams (1990): Effects of importance and interest on recall of biographical text, In: Journal of Reading Behavior, 22, S. 331-353.
- Wang, Z. & O. Adesope (2017): Do focused self-explanations prompts overcome seductive details? A multimedia study, In: Educational Technology & Society, 20(4), S. 47-57.
- Woytal, I. (2018): Im Auge des Betrachters, In: technische kommunikation, 01/18, S. 49-53.

# Abbildungsverzeichnis

Abb. 1: Modell der Cognitive Theory of Multimedia Learning nach R. E. Mayer. (Quelle: Mayer 2021: 40.).....	10
Abb. 2: Ein reiner Screenshot als Grundlage (links), sowie das daraus erstellte Simplified User Interface (rechts) (Quelle: Screenshot aus GIMP (links) sowie Eigene Darstellung (rechts). ....	30
Abb. 3: Heatmap von einem herkömmlichen Screenshot (links) und einer SUI Grafik (rechts). (Quelle: Bollen & Saremba 2020: 13.).....	30
Abb. 4: Links: Heat Map Visualisierung eines Screenshots. Rechts: Identischer Stimulus als Focus Map Darstellung (Quelle: Eigene Darstellung aus Screenshot von GIMP.) .....	44
Abb. 5: Bildschirmanordnung während der Bearbeitung des Software-Tutorials. (Quelle: Eigene Darstellung.) .....	47
Abb. 6: Schematischer Studienaufbau. (Quelle: Eigene Darstellung.) .....	48
Abb. 7: Stimuli Aufgabe 01: „Welche Tastenkombination benötigen Sie für die Aktion Lineale anzeigen?“ im Orientierungstest. Links: Stimuli für SUI-Gruppe, Mitte: Stimuli für SIG-Gruppe, rechts: Stimuli für SCR-Gruppe. (Quelle: Eigene Darstellung auf Basis von Screenshots aus GIMP.).....	51
Abb. 8: Beispielhafte Seite des Tutorials in den drei verschiedenen Grafikvarianten (SUI, SIG, SCR). (Quelle: Eigene Darstellung). ....	54
Abb. 9: Beispielhafte Darstellung der angelegten AOIs für GIMP, die gesamte Tutorialseite sowie für die einzelnen Bild- und Textpassagen im Tutorial. (Quelle: Bearbeiteter Screenshot aus SMI BeGaze und GIMP, Eigene Darstellung.).....	61
Abb. 10: Verteilungen der Vorerfahrungen der Proband*innen mit Bildbearbeitung und GIMP in %. (Quelle: Eigene Darstellung). ....	65
Abb. 11: Benötigte Zeit bis zur Lösungsfindung im Orientierungstest. Mittelwerte pro Aufgabe. (Quelle: Eigene Darstellung.).....	66
Abb. 12: Benötigte Fixationen bis zur Lösungsfindung im Orientierungstest. Mittelwerte pro Aufgabe. (Quelle: Eigene Darstellung.) .....	68
Abb. 13: Bearbeitungszeit für das gesamte Tutorial in Sekunden. (Quelle: Eigene Darstellung.) .....	70
Abb. 14: Durchschnittliche Bearbeitungszeit pro Handlungsschritt in Sekunden. ....	71
Abb. 15: Durchschnittlich korrekt gelöste Handlungsanweisungen in %. (Quelle: Eigene Darstellung.) .....	72
Abb. 16: Durchschnittlich korrekt gelöste Handlungsschritte in %. (Quelle: Eigene Darstellung.) .....	73
Abb. 17: Durchschnittliche Verweilzeit der Proband*innen im Tutorial und in GIMP pro Aufgabe. (Quelle: Eigene Darstellung.).....	74
Abb. 18: Prozentuales Verhältnis der Verweilzeiten im Tutorial und in GIMP. (Quelle: Eigene Darstellung.) .....	75

Abb. 19: Durchschnittlich Fixationsanzahl pro Abbildung im Tutorial. (Quelle: Eigene Darstellung.) .....	76
Abb. 20: Durchschnittlich Revisits pro Abbildung im Tutorial. (Quelle: Eigene Darstellung.) .....	76
Abb. 21: Antwortverhalten der Gruppen zu Frage 01: Wie fanden Sie das Tutorial? (Quelle: Eigene Darstellung) .....	78
Abb. 22: Antwortverhalten zu Frage 02: Denken Sie, dass Sie alle Aufgaben erfolgreich lösen konnten? (Quelle: Eigene Darstellung) .....	79
Abb. 23: Ergebnisse zu Frage 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der aufgaben. (Quelle: Eigene Darstellung.).....	80
Abb. 24: Antwortverhalten zu Frage 04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können. (Quelle: Eigene Darstellung).....	80
Abb. 25: Antwortverhalten zu Frage 05: Die Abbildungen im Tutorial haben mich verwirrt. (Quelle: Eigene Darstellung).....	81
Abb. 26: Antwortverhalten zu Frage 06: Die Abbildungen im Tutorial waren optisch ansprechend. (Quelle: Eigene Darstellung).....	82
Abb. 27: Gezeigter Stimulus: Einfaches Listenmenü. Je nach Gruppenzugehörigkeit entweder SUI Variante (links), signalisierter Screenshot (Mitte) oder unbearbeitete Screenshot (rechts). (Quelle: Eigene Darstellungen nach Screenshots aus GIMP.) ....	i
Abb. 28: Gezeigter Stimulus: Einfaches Fenstermenü. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellungen nach Screenshots aus GIMP.).....	i
Abb. 29: Gezeigter Stimulus: Gesamte Softwareoberfläche. Je nach Gruppen- zugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts).(Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.).....	ii
Abb. 30: Gezeigter Stimulus: Doppeltes Listenmenü. Je nach Gruppen-zugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.).....	ii
Abb. 31: Gezeigter Stimulus: Komplexes Fenstermenü. Je nach Gruppenzugehörigkeit entweder SUI (oben links), SIG (oben rechts) oder SCR (unten rechts). (Quelle: Eigene Darstellung auf Grundlage von Screenshots aus GIMP.).....	iii
Abb. 32: Verteilungen der Mittelwerte der Reaktionszeiten im Orientierungstest. (Quelle: Eigene Darstellung.) .....	iv
Abb. 33: Verteilungen der Mittelwerte der Fixationsanzahl im Orientierungstest. (Quelle: Eigene Darstellung.) .....	v
Abb. 34: Heat Maps zur Visualisierung der Fixationen im Orientierungstest im Vergleich der Gruppen – Stimuli 01 - 03. (Quelle: Eigene Darstellung aus Export aus SMI BeGaze.) .....	vi
Abb. 35: Heat Maps zur Visualisierung der Fixationen im Orientierungstest im Vergleich der Gruppen – Stimuli 04 - 05. (Quelle: Eigene Darstellung aus Export aus SMI BeGaze) .....	vii

Abb. 36: Verteilungen der Mittelwerte für die Bearbeitungszeit des gesamten Tutorials. (Quelle: Eigene Darstellung.) .....	viii
Abb. 37: Verteilungen der Mittelwerte für die Bearbeitungszeit pro Handlungsschritt. (Quelle: Eigene Darstellung.) .....	ix
Abb. 38: Verteilungen der Mittelwerte für korrekt gelöste Handlungsanweisungen. (Quelle: Eigene Darstellung.) .....	x
Abb. 39: Verteilungen der Mittelwerte für korrekt gelöste Handlungsschritte. (Quelle: Eigene Darstellung.) .....	xi
Abb. 40: Verteilungen der Mittelwerte der Dwell Time im Tutorial und GIMP. (Quelle: Eigene Darstellung.) .....	xii
Abb. 41: Verteilungen der Mittelwerte der Fixationsanzahl in den Abbildungen im Tutorial. (Quelle: Eigene Darstellung.) .....	xiii
Abb. 42: Verteilungen der Mittelwerte der Revisits zu den Abbildungen im Tutorial. (Quelle: Eigene Darstellung.) .....	xiv

# Tabellenverzeichnis

Tab. 1: Gestaltungsempfehlungen der Cognitive Load Theory. (Quelle: Übernommen aus Rey 2007: 19.).....	7
Tab. 2: 15 Prinzipien des multimedialen Lernens nach Mayer (2021). (Eigene Darstellung, übersetzt nach Mayer 2021: 399-402) * Für diese Arbeit relevante Prinzipien, welche im Folgenden genauer vorgestellt werden.....	11
Tab. 3: Für die Studie verwendete Software. (Quelle: Eigene Darstellung.) .....	47
Tab. 4: Vorerfahrungen der Proband*innen mit Bildbearbeitung im Allgemeinen und GIMP im Speziellen. Mittelwert (MW) und Standardabweichungen (SD); (0 = Gar keine, 1 = Nicht gut, 2 = Durchschnittlich, 3 = Gut, 4 = Sehr gut. (Quelle: Eigene Darstellung)	65
Tab. 5: Mittelwerte (MW), Standard-abweichungen (SD) und Standardfehler (SE) der benötigten Zeiten pro Aufgabe, unterschieden nach Gruppen. (Quelle: Eigene Darstellung.) .....	66
Tab. 6: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.) .....	67
Tab. 7: Mittelwerte (MW), Standard-abweichungen (SD) und Standard-fehler (SE) der benötigten Fixationen, unterschieden nach Gruppen. (Quelle: Eigene Darstellung)..	68
Tab. 8: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.) .....	68
Tab. 9: Bearbeitungszeit für das gesamte Tutorial nach Proband*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.) .....	70
Tab. 10: Bearbeitungszeit pro Handlungs-schritt nach Proband*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.) .....	71
Tab. 11: Ergebnisse der Tests auf Mittelwertunterschiede und deren Effektstärke. (Quelle: Eigene Darstellung.) .....	72
Tab. 12: Korrekt gelöste Handlungs-anweisungen nach Proband*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.) .....	72
Tab. 13: Korrekt gelöste Handlungs-schritte nach Proband*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standard-fehler (SE). (Quelle: Eigene Darstellung.) .....	73
Tab. 14: Verweilzeiten im Tutorial und in GIMP nach Proband*innen-Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.) .....	75
Tab. 15: Durchschnittliche Fixations-anzahl pro Abbildung im Tutorial nach Proband*innen Gruppe. Betrachtung der Mittelwerte (MW), Standardabweichung (SD) und Standardfehler (SE). (Quelle: Eigene Darstellung.) .....	76

Tab. 16: Durchschnittliche Revisits pro Abbildung im Tutorial nach Gruppe. Mittelwerte (MW), Standard-abweichung (SD) und -fehler (SE). (Quelle: Eigene Darstellung.)...	76
Tab. 17: Antwortverhalten zu Frage 01: Wie fanden Sie das Tutorial? (Quelle: Eigene Darstellung.) .....	78
Tab. 18: Antwortverhalten zu Frage 02: Denken Sie, dass Sie alle Aufgaben erfolgreich lösen konnten? (Quelle: Eigene Darstellung.) .....	79
Tab. 19: Antwortverhalten zu Frage 03: Die Abbildungen im Tutorial waren hilfreich zum Lösen der Aufgaben. (Quelle: Eigene Darstellung.) .....	79
Tab. 20: Antwortverhalten zu Frage 04: Ich hätte das Tutorial auch ohne die Abbildungen erfolgreich durchführen können. (Quelle: Eigene Darstellung.).....	81
Tab. 21: Antwortverhalten zu Frage 05: Die Abbildungen im Tutorial haben mich verwirrt. (Quelle: Eigene Darstellung.).....	81
Tab. 22: Antwortverhalten zu Frage 06: Die Abbildungen im Tutorial waren optisch ansprechend. (Quelle: Eigene Darstellung.).....	82
Tab. 23: Kalibrierungswerte der Proband*innen. (Quelle: Eigene Darstellung.).....	iii

# Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Masterarbeit mit dem Titel:

*Effizienz von Simplified User Interfaces in Software-Tutorials – Eine Eyetracking-Studie*

selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Stellen sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Die digitalisierte Fassung der Arbeit stimmt mit der in Schriftform vorgelegten Fassung wörtlich überein.

Leipzig, 06. Februar 2023

---

Ort, Datum

Unterschrift (Pia Bsdurrek)