



Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes

Psychische Beanspruchung beim Einsatz von
HMDs

BAUA-BERICHT

**Forschung
Projekt F 2288**

M. Wille

**Head-Mounted Displays –
Bedingungen des sicheren und
beanspruchungsoptimalen Einsatzes**

Psychische Beanspruchung beim Einsatz von HMDs

Dortmund/Berlin/Dresden 2016

Diese Veröffentlichung ist der Abschlussbericht zum Schwerpunkt Psychische Beanspruchung beim Einsatz von HMDs des Projekts „Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes“ – Projekt F 2288 – im Auftrag der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei dem Autor.

Autor: Dr. Matthias Wille
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Titelfoto: Elena Meyer
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Umschlaggestaltung: Susanne Graul
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin

Herausgeber: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA)
Friedrich-Henkel-Weg 1-25
44149 Dortmund
Postanschrift: Postfach 17 02 02, 44061 Dortmund
Telefon: 0231 9071-2071
Telefax: 0231 9071-2070
E-Mail: info-zentrum@buaa.bund.de
Internet: www.buaa.de

Berlin:
Nöldnerstr. 40 – 42, 10317 Berlin
Telefon: 030 51548-0
Telefax: 030 51548-4170

Dresden:
Fabricestr. 8, 01099 Dresden
Telefon: 0351 5639-50
Telefax: 0351 5639-5210

Nachdruck und sonstige Wiedergabe sowie Veröffentlichung, auch auszugsweise, nur mit vorheriger Zustimmung der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.



www.buaa.de/dok/6833146

ISBN 978-3-88261-163-2

Inhaltsverzeichnis

	Seite
Kurzreferat	5
Abstract	6
1 Einleitung	7
2 Laborstudie	9
2.1 Zielsetzung der Studie	9
2.2 Versuchsaufgaben	9
2.2.1 Konstruktionsaufgabe	9
2.2.2 Monitoringaufgabe	10
2.3 Apparatur	11
2.3.1 Verwendetes Head-Mounted Display	11
2.3.2 Verwendeter Tablet-PC	12
2.3.3 Eingabegerät und Positionierung	12
2.4 Versuchsplan	13
2.4.1 Unabhängige Variablen	13
2.4.2 Abhängige Variablen	14
2.5 Versuchsablauf	16
2.6 Probanden	18
2.7 Hypothesen	19
2.7.1 Leistung in der Konstruktionsaufgabe	20
2.7.2 Trefferquote bei den Monitoringaufgaben	20
2.7.3 Reaktionszeit bei den Monitoringaufgaben	20
2.7.4 Subjektive Beanspruchung	20
2.7.5 Objektive Beanspruchung	21
2.7.6 Visuelle Ermüdung	21
2.8 Datenaufbereitung	22
2.8.1 Konstruktions- und Monitoringaufgaben	22
2.8.2 Fragebogendaten	23
2.8.3 Herzrate	23
2.9 Inferenzstatistische Verfahren	27
3 Ergebnisse	28
3.1 Leistung in der Konstruktionsaufgabe	28
3.2 Trefferquote in den Monitoringaufgaben	29
3.3 Reaktionszeit in den Monitoringaufgaben	31
3.4 Subjektive Beanspruchung	32
3.4.1 NASA-TLX	32
3.4.2 RSME	34
3.4.3 Vergleich von NASA-TLX und RSME-Ergebnissen	35
3.5 Objektive Beanspruchung	35
3.5.1 Herzrate	36
3.5.2 Herzratenvariabilität	40
3.6 Vergleich zwischen subjektiver und objektiver Beanspruchung	44
3.7 Visuelle Ermüdung	45
3.8 Interview	51
3.9 Zusammenfassung	57

4	Replikation mit Google Glass	59
4.1	Ausgangslage und Zielsetzung der Studie	59
4.2	Versuchsaufbau	60
4.2.1	Versuchsaufgaben	61
4.2.2	Apparatur	63
4.2.3	Versuchsplan	64
4.2.4	Versuchsablauf	66
4.2.5	Probanden	68
4.2.6	Hypothesen	69
4.2.7	Datenaufbereitung	70
4.2.8	Inferenzstatistische Verfahren	70
4.3	Ergebnisse	71
4.3.1	Leistung in der Konstruktionsaufgabe	71
4.3.2	Leistung in den Monitoringaufgaben	72
4.3.3	Subjektive Beanspruchung	76
4.3.4	Visuelle Ermüdung	78
4.3.5	Interview	82
4.4	Vergleich der Befundlage beider HMDs	83
4.4.1	Vergleich der Leistung in der Konstruktionsaufgabe	85
4.4.2	Vergleich der Leistung in der Monitoringaufgabe	85
4.4.3	Vergleich der subjektiven Beanspruchungsparameter	86
4.4.4	Vergleich des Fragebogens zur visuellen Ermüdung	86
4.4.5	Fazit	87
5	Diskussion	88
5.1	Interpretation der Ergebnisse	88
5.1.1	Vergleich zwischen HMD und Tablet-PC	88
5.1.2	Gewöhnung an die HMD-Technologie	89
5.2	Resultierende Handlungsempfehlungen	90
6	Literatur	92
6.1	Literaturangaben zu diesem Projektbericht	92
6.2	Veröffentlichungen zu den beschriebenen Studien	93
	Abbildungsverzeichnis	94
	Anhang	98
Anhang 1	Unterlagen der Hauptstudie	98
Anhang 2	Unterlagen der Replikationsstudie	118

Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes

Psychische Beanspruchung beim Einsatz von HMDs

Kurzreferat

Der Bericht beschreibt zwei Laborstudien, die im Rahmen des Projektes F 2288 „Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes“ der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) durchgeführt worden. Der Fokus dieser Studien lag auf der mentalen Beanspruchung bei mehrstündiger Arbeit mit Head-Mounted Displays (HMDs). Zur Simulation wurden eine Konstruktionsaufgabe und parallel dazu eine Monitoringaufgabe gestellt. Zum Vergleich waren die Arbeitsinhalte sowohl auf einem HMD als auch auf einem Tablet-PC dargestellt. In einer Studie kam ein industrielles HMD zum Einsatz und in einer kürzeren Replikationsstudie ein HMD aus dem Konsumentenbereich.

Die vorliegende Arbeit verzichtet auf eine thematische Einführung und Beschreibung von Anwendungsfällen, da diese bereits im Projektbericht „Physische Beanspruchung beim Einsatz von HMDs“ (www.baua.de/dok/6833052) vorliegt. Nach einer kurzen Beschreibung des Ziels der Studie werden die Methoden im Einzelnen vorgestellt. Es folgt eine Darstellung des Versuchsablaufes und für den Kennwert der Herzratenvariation eine detaillierte Beschreibung der Datenaufbereitung und der verwendeten Filter. In den Ergebnissen werden die vorher aufgestellten Hypothesen einzeln geprüft. Im Anschluss an die zweite Studie folgt ein Vergleich der Ergebnisse beider Studien. Abschließend werden die Ergebnisse hinsichtlich der Faktoren „Vergleich mit einem Tablet-PC“ und „Gewöhnungseffekte an das HMD“ diskutiert und Handlungsempfehlungen für die Gestaltung und den Einsatz von HMDs mit direktem Bezug auf die Untersuchungsergebnisse abgeleitet.

Schlagwörter:

Head-Mounted Displays, HMDs, Datenbrillen, Wearable und Ubiquitäres Computing, Belastung und Beanspruchung, Konstruktionsaufgabe, Monitoringaufgabe, Ergonomie, Human Factors, Rating Scale of Mental Effort, NASA-TLX, Herzratenvariabilität, Visual Fatigue Questionnaire

Head-mounted displays – How to increase safety and reduce physical strain

Psychological strain while using HMDs

Abstract

In this report two laboratory studies are described in detail which took place in the framework of the project F 2288 "Head-Mounted Displays - conditions for safe and strain-optimal use" conducted by the Federal Institute for Occupational Safety and Health in Germany (BAuA). Focus of these studies was the mental strain during several hours of work with head-mounted displays (HMDs). For this purpose, a construction task and a parallel monitoring task was performed, while the work content was presented either on an HMD or tablet-PC for comparison. In the first study an industrial HMD was used and in the shorter replication study a HMD from the consumer sector.

The present work waives a thematic introduction and description of use cases, as this took place in the previous report „Physical strain while using HMDs“ (www.baua.de/dok/6833052). After a brief description of the aim of the study methods are presented in detail. The following is a description of test process and especially for the characteristic value of heart rate variation, a detailed description of the data processing and filters used. In the results the previous formulated hypotheses are tested individually. Following the second study, the results of both studies are compared. Finally, the results are discussed in terms of the factors "compared with a tablet PC" and "habituation effects on the HMD" and recommendations for the design and use of HMDs with direct reference to the findings are given.

Key words:

Head-mounted displays, HMDs, data glasses, Wearable and Ubiquitous Computing, stress and strain, construction task, monitoring task, ergonomics, human factors, Rating Scale of Mental Effort, NASA-TLX, heart rate variability, Visual Fatigue Questionnaire

1 Einleitung

Dieser Teil des Forschungsberichtes beschreibt zwei Studien, die im Rahmen des Forschungsprojektes „Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes“ der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) durchgeführt wurden. Diese Studien sowie dieser Teil des Berichtes sind als Teilleistung des Gesamtprojektes zu verstehen und müssen auch in diesem Zusammenhang interpretiert werden.

Unter Head-Mounted Displays (HMDs) versteht man kleine Monitore, die direkt vor den Augen platziert sind. Hierbei gibt es verschiedene Aufbautypen und Anwendungsfelder (für einen detaillierten Überblick siehe Kapitel 2 in Teil 2 dieser Berichtreihe oder Wille et al. 2011). Die Diversität der einzelnen HMD-Modelle macht es schwierig, von einer einheitlichen Produktklasse zu sprechen, für die z. B. feste Regeln im Umgang zu beachten sind. Hierfür sind die Modelle zu unterschiedlich hinsichtlich der verwendeten Displays, der Sichtverdeckung, des Gewichts, des Kopfträgers und von weiteren Faktoren.

Allgemein kann man HMD-Systeme hinsichtlich folgender Gesichtspunkte unterscheiden:

- Zunächst stellt sich die Frage, wie vielen Augen ein Display vorgesetzt wird. Monokulare HMDs haben nur ein Display vor einem der Augen. Hierbei ist bei manchen Herstellern das Display fest auf einer Seite, während bei anderen Herstellern die Seite des Displays frei wählbar ist. In beiden Fällen kann mit einem monokularen Display die Information aber nur mit einem Auge wahrgenommen werden. Binokulare HMDs verfügen über 2 getrennte Displays vor jedem Auge. In vielen Fällen lassen sich so auch 3D Inhalte darstellen.
- Als nächstes stellt sich die Frage, ob man durch das Display noch die Umwelt sehen kann. Dies ist möglich bei sogenannten „see-through“ Modellen, bei denen die Inhalte auf halbtransparenten Displays dargestellt werden. Hingegen ist bei „look-around“ Modellen eine Durchsicht auf die Umwelt nicht möglich. Der Vorteil von „see-through“ Modellen liegt darin, die Sichtverdeckung im Displaybereich durch halbtransparentes Durchscheinen der dahinter liegenden Umwelt zu minimieren. Der Nachteil liegt in einem wechselnden Kontrast, der gerade bei hellen Hintergrundsflächen eine Informationsaufnahme beschwert. Der Vorteil von „look-around“ Modellen liegt in einem besseren Kontrast, der Nachteil aber in einer stärkeren Verdeckung der Umwelt. Die Bezeichnung „look-around“ kann hierbei irreführend sein, da je nach Displaygröße manchmal ein erheblicher Teil des Sehbereiches verdeckt ist. Viele „look-around“-HMDs lassen sich aber bei Bedarf wegklappen. Bei „look-around“ Systemen, die den gesamten Sichtbereich umfassen, spricht man auch häufig von „geschlossenen Systemen“. Diese eignen sich vor allem für Anwendungen in der virtuellen Realität (VR), da hier eine größere Immersion des Trägers bewirkt wird.
- Ein weiteres Unterscheidungsmerkmal ist das Gewicht des HMDs und damit oft zusammenhängend auch die Art des Kopfträgers. Während schwerere HMD-Systeme – weil sie älterer Bauart sind oder einen größeren Funktions-

umfang haben – oft einen Kopfträger verwenden, der an einen Helm erinnert, haben leichtere HMDs oft nur eine Art Brillengestell als Kopfträger.

- Auch der Funktionsumfang eines HMD-Systems kann sich unterscheiden. So beinhalten manche Systeme zusätzlich noch ein Headset für die Audiokommunikation, während dies bei anderen nicht der Fall ist. Auch hinsichtlich der Interaktionsmöglichkeiten mit dem Display gibt es große Unterschiede zwischen den Systemen. Während manche Systeme auf Spracheingabe reagieren, können andere dies nicht und benötigen eine manuelle Interaktion über Schalter am HMD oder über eine externe Box.
- Ebenso können HMDs hinsichtlich ihres Anwendungsfeldes unterschieden werden. So gibt es „industrielle HMDs“ die für den Arbeitseinsatz gedacht sind und Konsumentenprodukte, die eher für den Privatgebrauch bestimmt sind. Industrielle HMDs zeichnen sich hierbei meist durch eine höhere Robustheit, eine Kombinationsmöglichkeit mit Sicherheitskleidung (vor allem Schutzbrillen) und höhere Akkulaufzeit aus. Konsumentenprodukte hingegen sind oft leichter, haben weniger Akkuleistung und sind auch weniger robust. Allerdings sind die Grenzen hier fließend.
- „Augmented Reality“ (AR) ist ein technisches Verfahren, bei dem computergenerierte Inhalte lagegerecht über die Umwelt eingeblendet werden. Dieses Verfahren wird zwar oft im Zusammenhang mit HMDs gesehen, ist aber kein zwingender Bestandteil eines HMD-Systems. Denn zum Einen können Inhalte auf HMDs auch ohne AR dargestellt werden, wie z. B. Arbeitsanweisungen oder Schaltpläne, und zum Anderen ist AR auch auf Tablet-PC und anderen Geräten realisierbar.

Diese Zusammenstellung von HMD-Merkmalen und Eigenheiten soll dem Leser eine gewisse Übersicht darüber geben, was man alles unter dem Sammelbegriff HMD verstehen kann, erhebt aber keinen Anspruch auf Vollständigkeit. Gerade in den letzten Jahren ist die Anzahl von verschiedenen HMDs stetig gewachsen und die technische Entwicklung hat zu einer Fortführung der Diversität beigetragen.

Im Rahmen des o. g. Forschungsprojektes untersucht die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin den Einsatz von HMDs als adaptives Arbeitsassistenzsystem. Der Vorteil eines solchen Assistenzsystems liegt hier vor allem darin, dass der Nutzer bei der Ausführung der Arbeitsaufgabe beide Hände frei hat, während er zeitgleich arbeitsrelevante Informationen aus dem Display entnehmen kann. Im Projekt werden hierfür ausschließlich monokulare HMDs verwendet, da diese durch die geringere Sichtverdeckung als binokulare HMDs besser für die mobile Arbeitsunterstützung geeignet sind.

Der Fokus der hier berichteten Studie sowie der in Teilprojekt 2 beschriebenen Studie zur physischen Beanspruchung liegt auf einem länger andauernden Einsatz über mehrere Stunden, da es hier noch Erkenntnislücken bei der Anwendung dieser Technologie gibt. Neben der Leistung interessiert bei einem solchen länger andauernden Einsatz von HMDs vor allem der Beanspruchungsverlauf. Dieser ist in Relation zu setzen mit der Beanspruchung durch andere Möglichkeiten der Informationsanzeige (hier ein Tablet-PC). Und vor dem Hintergrund des demographischen Wandels sind altersabhängige Einflüsse ebenfalls von Interesse.

2 Laborstudie

2.1 Zielsetzung der Studie

Um Beanspruchung und Leistung beim länger andauernden Einsatz von HMDs zu untersuchen, wurde in den Laborräumen der BAuA eine Studie durchgeführt. Hierbei arbeitete jeder Proband in drei Sessions jeweils vier Stunden, entweder mit einem HMD oder vergleichsweise mit einem Tablet-PC. Neben der erbrachten Arbeitsleistung in einer Konstruktionsaufgabe und einer parallel dargebotenen Monitoringaufgabe interessiert vor allem die mentale Beanspruchung der Nutzer bei den beiden unterschiedlichen Informationsanzeigetechnologien. Hierzu wurde die Beanspruchung mit mehreren subjektiven Verfahren zu verschiedenen Zeitpunkten erhoben, um so einen Beanspruchungsverlauf nachvollziehen zu können. Des Weiteren wurde die Herzrate über den gesamten Versuchsverlauf erhoben, um so die objektiven Beanspruchungsparameter „Herzrate“ (emotionale Beanspruchung) und „Herzratenvariabilität“ (mentale Beanspruchung) abzuleiten.

2.2 Versuchsaufgaben

Während des Versuches waren zwei Aufgaben parallel zu bearbeiten, die beide als gleich wichtig instruiert wurden ("Dual-task" Paradigma) und deren Inhalte auf dem HMD oder Tablet-PC dargestellt wurden.

2.2.1 Konstruktionsaufgabe

Zum einen sollten die Probanden Modellfahrzeuge aus einzelnen Bausteinen zusammenbauen (Lego-Technic). Hierbei wird mittels schematischer Zeichnung pro Konstruktionsfolie der aktuelle Zustand des Modells dargestellt, neu hinzuzufügende Bausteine in einer Kopfleiste angezeigt und auf deren Positionierung im Gesamtmodell meist mit Pfeilen hingewiesen. So wird Schritt für Schritt ein komplexes Modell aufgebaut (siehe Abb. 2.1). Die ursprünglichen Lego-Papieranleitungen wurden hierzu eingescannt und nachbearbeitet. Dabei wurden einige in der Papierform sehr kleine Details mittels Grafikprogramm größer gemacht. Das Gesamtkonzept der Lego-Technic Anleitung wurde aber übernommen.

Es wurden verschiedene Modelle mit unterschiedlicher Komplexität verwendet (vgl. Anhang 1); wobei die Komplexität der einzelnen Konstruktionsfolien immer vergleichbar ist und komplexere Modelle lediglich aus mehr Konstruktionsfolien bestehen. Zu Beginn des Versuches bearbeiteten die Probanden eher einfache Modelle, in späteren Sessions wurden die Modelle komplexer. Wann immer ein Modell beendet war, bekamen die Probanden das nächste Modell gereicht, so dass während der gesamten Versuchszeit kontinuierlich gebaut wurde. Eventuelle Fehler in der Konstruktion zeigen sich oft erst in späteren Aufbaufolien, da hier dann neue Bausteine nicht mehr passen. In diesem Fall waren die Probanden angehalten, den Fehler zu identifizieren und zurück zu bauen. Hierzu konnten sie zu jedem Zeitpunkt die Konstruktionsfolien vor- und zurück schalten. Sollten Probanden ihre eigenen Fehler nicht gesehen haben, wurden sie vom Versuchsleiter darauf aufmerksam gemacht.

Diese Aufgabe mit der bildbasierten, schrittweisen Aufbauanleitung wurde gewählt, da sie industriellen Montageanleitungen ähnelt, wie sie z. B. in der Instandhaltung verwendet werden.

2.2.2 Monitoringaufgabe

In einer parallel dargebotenen Monitoring-Aufgabe sollten die Probanden auf drei vertikale Balken am äußeren Bildschirmrand achten und bestimmte Zustandsänderungen mit Knopfdruck quittieren. Die Balken variierten kontinuierlich, aber sehr langsam in ihrer Länge und ab und zu in ihrer Farbe (blau/rot).

Es sollte mit Knopfdruck zum einen ein Farbwechsel bestätigt werden, der durch die großflächige Veränderung in der Darstellung einen visuellen Popout-Effekt beinhaltet. Zum anderen sollte ein Positionswechsel des längsten Balkens, der durch die kontinuierliche Längenvariation der Balken entsteht, mit einem anderen Knopf quittiert werden. Hierbei bekamen die Probanden in der Hälfte der Blöcke eine Rückmeldung über ihre letzte Bestätigung in Form einer schriftlichen Mitteilung über den Balken (z. B. "Mitte - Rot").

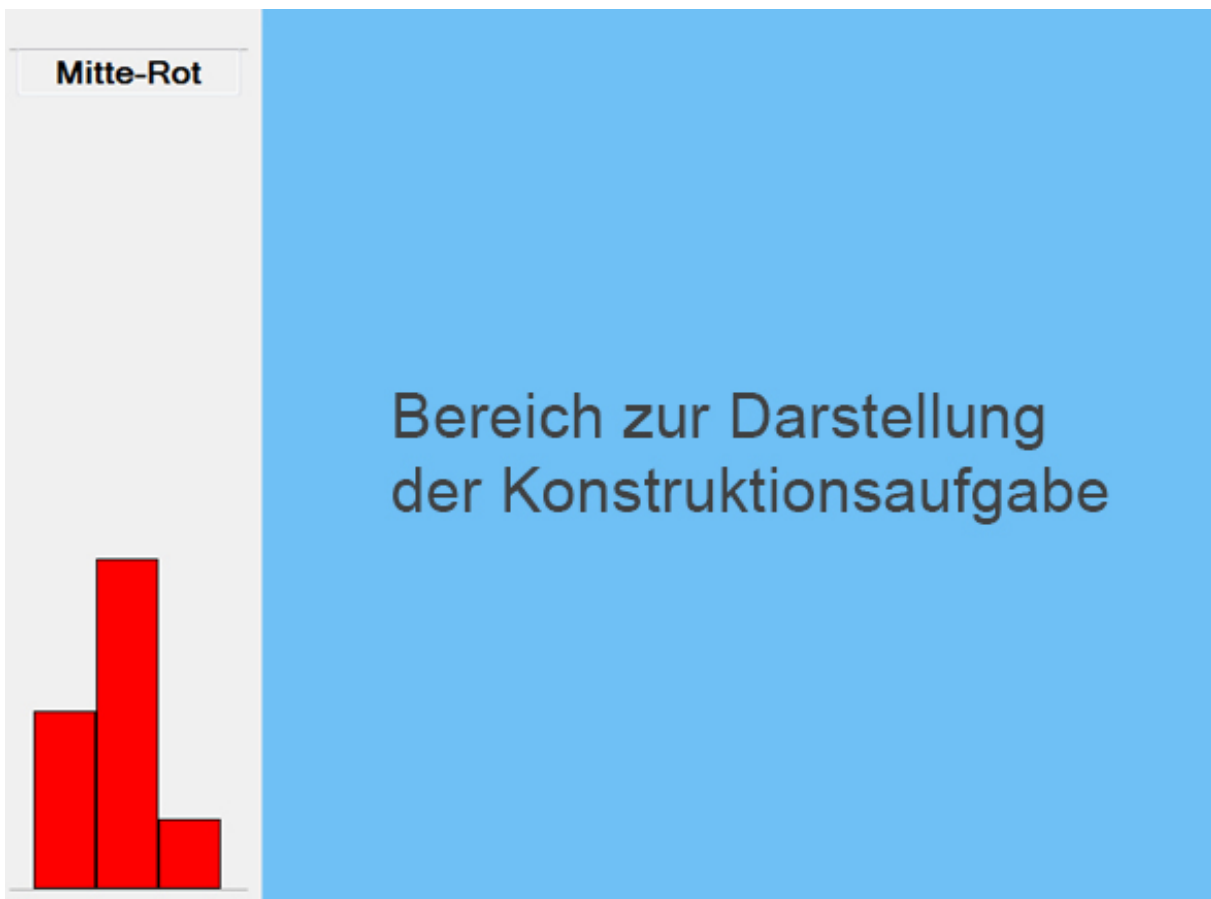


Abb. 2.1 Darstellungsaufteilung der beiden Arbeitsaufgaben. Auf der rechten Seite Platz für die Konstruktionsanleitung, auf der linken Seite die Monitoringaufgabe mit drei Balken und Rückmeldung über die letzten bestätigten Zustände. (Hier Darstellung links montiertes HMD, bei rechts montiertem HMD seitenverkehrt: Die Monitoringaufgabe war immer peripher)

Ein Farbwechsel fand zufallsgesteuert alle 40-240 Sekunden statt, wobei der anhand der späteren Daten errechnete Durchschnitt tatsächlich auch bei 140 Sekunden lag. Die Geschwindigkeit der Längenvariation betrug einen Bildpunkt pro 400 Millisekunden, wobei die Balken unabhängig von einander bis zu ihrem zufälligen Zielwert wuchsen oder schrumpften und dann mit gleichbleibender Geschwindigkeit zu ihrem neuen Zielwert in entgegengesetzter Richtung variierten. Die Länge der Balken ging dabei von Pixel 10 bis Pixel 520 in vertikaler Richtung und eine spätere Analyse der Daten zeigt, dass im Schnitt alle 95 Sekunden reagiert werden musste.

2.3 Apparatur

2.3.1 Verwendetes Head-Mounted Display

Bei dem verwendeten HMD handelt es sich um ein MAVUS-System der Firma Heitec (www.mavus.info), wie es derzeit auch in der industriellen Instandhaltung eingesetzt wird. Es ist ein monokulares "look around"-Display mit einer Auflösung von 800 x 600 Pixeln. Die Technik ist an einem Kopfträger befestigt, der aus dem Medizinbereich stammt und sich auf die individuelle Kopfform einstellen lässt. Am Kopfträger sind auch eine Frontkamera und ein Headset montiert, welche allerdings bei dem Versuch keine Funktion erfüllten. Im industriellen Einsatz sind Headset und Kamera für die Kommunikation mit Experten vorgesehen.

Der Kopfträger wiegt 380 Gramm und ist über ein Kabel mit einer Weste verbunden, die die Funktechnologie zur Übertragung der Daten und den Akkumulator für die Stromversorgung beinhaltet (siehe Abb. 2.2).



Abb. 2.2 Das verwendete MAVUS®¹-HMD-System. Links: Der Kopfträger mit Display, Kamera, Headset. Mitte: Das Display in Nahaufnahme. Rechts: Das Komplettsystem inklusive Weste und Tragekoffer.

¹ mavus® ist eine eingetragene Marke der HEITEC Innovations GmbH

Die Displayeinheit des HMDs kann wahlweise vor dem rechten oder linken Auge montiert werden. Im Rahmen des Versuches wurde das Display immer vor dem dominanten Auge des Probanden angebracht.

Die Fokussierung der Anzeige – das „Scharfstellen“ – erfolgt bei diesem HMD einzig über die Positionierung der Displayeinheit, da das vor dem Display angebrachte Prisma eine feste Brennweite hat. Hierzu kann das Display in der vertikalen und horizontalen Richtung am Kopfträger verschoben werden. Des Weiteren befindet sich am Display und am Verbindungstück zum Kopfträger jeweils ein Kugelgelenk. Über diese beiden Kugelgelenke kann das Display näher an das Auge heran- oder weiter weg gedrückt werden, bis ein optimaler Punkt gefunden ist. Der Spielraum hierbei beträgt oft nur wenige Millimeter, so dass bei der Anpassung des Displays an den Probanden eine gewisse Feinarbeit nötig ist.

2.3.2 Verwendeter Tablet-PC

Als Tablet-PC kam ein CL900 der Firma Motion (www.motioncomputing.de) zum Einsatz. Dieser Tablet-PC wurde speziell für den Arbeitseinsatz entwickelt und hat im Gegensatz zu bekannten Tablet-PCs aus dem Consumerbereich (z. B. iPad von Apple) eine größere Toleranz bezüglich Stößen, Erschütterungen oder Spritzwasser. Prozessor ist ein Intel Atom Z 670 und das Gewicht beträgt ca. 1 Kilogramm. Der Bildschirm hat eine Größe von 10,1 Zoll im 16:9 Format mit 1366x768 Pixeln, wobei im Rahmen des Versuches nur ein Fenster von 800 x 600 Pixeln dargestellt und die übrige Fläche abgeklebt wurde, um zu gewährleisten, dass die Darstellung der Arbeitsinhalte gleich ausfällt (siehe Abb. 2.3 links).

2.3.3 Eingabegerät und Positionierung

Ebenfalls im Sinne der vergleichbaren Bedingungen zwischen Tablet-PC und HMD wurden sämtliche im Versuchsverlauf nötigen Interaktionen (Vor- und Zurückschalten der Montagefolien, Tasten zur Quittierung der Balkenaufgaben) über ein umgebautes Nummernpad der Firma Logitech eingegeben (siehe Abb. 2.3 rechts). Der Tablet-PC war somit nur Darstellungseinheit und die Touchscreenfunktionen wurden nicht genutzt.

Tablet-PC und Nummernpad waren auf der Arbeitsfläche frei positionierbar. Für den Tablet-PC gab es weiterhin verschiedene Stützen, so dass unterschiedliche Neigungswinkel realisiert werden konnten. Die Probanden waren aufgefordert, sich vor Versuchsbeginn ihren Arbeitsbereich einzurichten.



Abb. 2.3 Links: Der Tablet-PC Motion CL900 mit abgeklebter Darstellungsfläche von 800x600 Pixeln und Legoanleitung. Rechts: Das umgebaute Nummernpad mit Reaktionstasten für die Monitoringaufgabe (oben/Mitte) und Tasten zum vor- und zurückschalten der Aufbaufolien.

2.4 Versuchsplan

Es liegt ein Messwiederholungsplan mit mehreren abhängigen und unabhängigen Variablen vor. Die unabhängigen Variablen sind dabei im mixed Design, so dass einige über die Messwiederholung innerhalb der Probanden variiert werden (within subject), und bei einigen wird die Population in Gruppen eingeteilt (between subject).

2.4.1 Unabhängige Variablen

2.4.1.1 Versuchssession

Durch die Messwiederholung nahm jeder Proband an drei Sessions an drei verschiedenen Tagen teil, wobei zwischen den Sessions mindestens zwei und maximal zehn Tage lagen. Die erste Session wurde immer mit dem HMD durchgeführt. Für die zweite und dritte Session wurde die Vergleichsmessung mit dem Tablet-PC und die Wiederholungsmessung mit dem HMD permutiert.

Im Rahmen der Auswertung interessiert vor allem der Vergleich zwischen zweiter HMD-Messung und Tablet-PC, um unterschiedliche Auswirkungen der Informationsanzeigetechnologie herauszustellen. Die erste Messung mit dem HMD wird für diesen Vergleich bewusst nicht heran gezogen, da bei einem ersten Umgang mit der für die Probanden ungewohnten Technologie Lern- und Gewöhnungseffekte auftreten können, die nicht Teil des Technologievergleiches sein sollen.

Um die Gewöhnung an die HMD-Technologie zu untersuchen, wird hingegen der Vergleich zwischen erster und zweiter HMD-Messung herangezogen.

Die drei Sessions werden also nicht als Stufen eines Faktors behandelt, sondern es werden Einzelvergleiche zwischen den Sessions gezogen.

2.4.1.2 Rückmeldung in der Monitoringaufgabe

Bei der parallel zur Konstruktionsaufgabe dargebotenen Monitoringaufgabe sollten die Probanden drei Balken am Bildschirmrand beobachten und auf gewisse Zustandsänderungen mit Knopfdruck reagieren (siehe 2.2 und Abb. 2.1). Hierbei bekamen sie in der Hälfte der Blöcke eine über den Balken geschriebene Rückmeldung über ihren letzten bestätigten Zustand, und in der anderen Hälfte der Blöcke gab es diese Rückmeldung nicht. Die Rückmeldung wurde in den vier Blöcken jeweils abwechselnd gezeigt oder nicht gezeigt und die Variation durchpermutiert, so dass die Hälfte der Probanden Rückmeldung in Block eins und drei und die andere Hälfte in Block zwei und vier hatte.

2.4.1.3 Alter

Um Effekte des Alters auf Leistung und Beanspruchung zu überprüfen, wurde die Probandengruppe mit einem Mediansplit in zwei Gruppen unterteilt und das Alter somit als Zwischensubjektfaktor eingeführt. Die jüngere Gruppe war hierbei zwischen 18 und 31 Jahren die ältere Gruppe zwischen 38 und 67 Jahren. Diese Unterteilung wurde positiv auf die Normalverteilungsannahme getestet. Für eine Übersicht über die Verteilung der beiden Altersgruppen siehe auch Abb. 2.7.

2.4.1.4 Technikaffinität

Die Präferenz und Hingabe zu neuen Technologien unterscheidet sich je nach individuellem Probanden und kann Einfluss auf die Ergebnisse haben. Dies ist vor allem bei subjektiven Werten und sehr neuen, für den Probanden ungewöhnlichen Technologien zu beachten. Deswegen wurde mittels eines Fragebogens (Fragebogen TA-EG, Karrer, K., Glaser, C., Clemens, C., & Bruder, C., 2009, siehe auch Anhang 1) auch die Technikaffinität der Probanden erfasst. Diese wurden dann ebenfalls mittels Mediansplit in zwei Gruppen eingeteilt und als Zwischensubjektfaktor verwendet, um Einflüsse der Technikaffinität auf die Ergebnisse zu untersuchen.

2.4.2 **Abhängige Variablen**

2.4.2.1 Anzahl der bearbeiteten Konstruktionsfolien pro Session

Als Indikator für die erbrachte Leistung wird die Anzahl der pro Versuchssession abgearbeiteten Konstruktionsfolien verwendet. Dies ist möglich, da die einzelnen Konstruktionsfolien von vergleichbarer Komplexität hinsichtlich der nötigen Arbeitsschritte und hinzuzufügenden Teile sind und komplexere Modelle lediglich aus mehr Konstruktionsfolien bestehen.

2.4.2.2 Trefferquote in den Monitoringaufgaben

Bei den beiden Monitoringaufgaben (Reaktion auf Farbwechsel oder Wechsel der Position des längsten Balkens) wird die Trefferquote getrennt berechnet. Die Trefferquote ist der prozentuale Anteil an richtigen Reaktionen von allen möglichen Reaktionen. Als richtige Reaktion wird die Bestätigung eines Farb- oder Längenwechsels angesehen, solange die Reaktion noch erfolgte, bevor in dieser Dimension ein weite-

rer Wechsel stattfand. Es gab also keine zeitliche Begrenzung. Eine falsche Reaktion kann das Unterlassen einer geforderten Reaktion nach oben genannten Prinzipien sein, oder die erneute Reaktion, obwohl zwischenzeitlich kein Statuswechsel erfolgt ist.

2.4.2.3 Reaktionszeit in den Monitoringaufgaben

Die Reaktionszeit ist die vergangene Zeit zwischen Stimulus onset (Veränderung der Balkenfarbe oder der neuen Positionierung des längsten Balkens) und der erfolgten richtigen Reaktion. Auch hier gibt es keine zeitliche Begrenzung. Einzig die erneute Variation des Stimulus begrenzt die mögliche Reaktionszeit. Die Reaktionszeiten werden getrennt für die Reaktion auf Farb- und Längenwechsel ausgewertet.

2.4.2.4 RSME-Werte zu 8 Zeitpunkten

Die Rating Scale of Mental Effort (RSME nach Zijlstra, 1993) ist eine eindimensionale Skala von 0-150, auf der Probanden ihre Beanspruchung beurteilen (siehe Abb. 2.4). Diese subjektive Beanspruchungserhebung kann direkt während der Arbeitszeit abgefragt werden („Wie anstrengend ist es gerade?“) und von den Probanden mit einem spontanen Rating und der Antwort als Zahlenwert (z. B. „68“) benannt werden. Hierzu empfiehlt es sich, die Probanden vor Versuchsbeginn mit der Skala bekannt zu machen, indem z. B. verschiedene Alltagssituationen vorgestellt werden, die dann auf der Skala eingeschätzt werden.

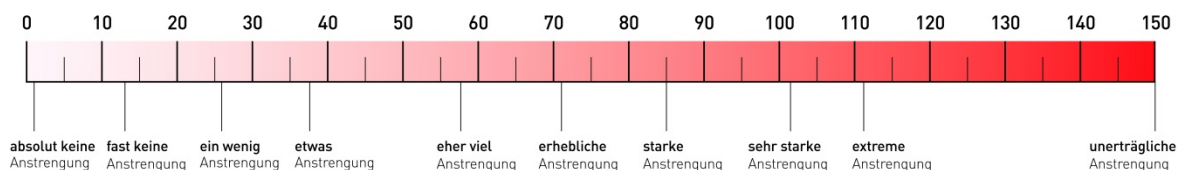


Abb. 2.4 Die Rating Scale of Mental Effort (RSME) nach Zijlstra, 1993. Während des Versuches wurde die RSME zu acht Zeitpunkten erhoben: In der Mitte und gegen Ende von jedem der vier jeweils ca. 50 Minuten langen Arbeitsblöcke. Dies ermöglicht eine detaillierte Einsicht in die Beanspruchungsentwicklung über die Gesamtzeit.

2.4.2.5 NASA-TLX zu zwei Zeitpunkten

Der NASA-Task Load Index (NASA-TLX nach Hart und Staveland, 1988) ist eines der bekanntesten Verfahren zur subjektiven Beanspruchungseinschätzung. Im Gegensatz zur RSME verfügt der NASA-TLX über sechs Subskalen, auf denen die Probanden unterschiedliche Dimensionen von Beanspruchung auf einer Skala von 0-100 einschätzen. Die Subskalen sind: geistige, körperliche und zeitliche Anforderungen, Leistung, Anstrengung und Frustration. Im Gegensatz zur ursprünglichen Intention der Autoren erweist es sich als schwierig, diese sechs Subskalen in einer Arbeitssituation abzufragen, da die Ablenkung von der Arbeitsaufgabe durch das sechsfache Rating deutlich höher ist als z. B. bei der RSME. Die Unterteilung in Subskalen bietet auf der anderen Seite den Vorteil, dass man Einsicht in bestimmte Aspekte von Beanspruchung hat und so feststellen kann, ob z. B. eher zeitliche oder

körperliche Anforderungen einen größeren Anteil an der empfundenen Beanspruchung haben. Da die Gesamtbeanspruchung eine Kombination aus diesen Teilaspekten ist und diese Kombination individuelle Schwerpunkte hat, wird beim NASA-TLX auch immer ein Paarvergleich der einzelnen Skalen durchgeführt und dieser geht bei der Errechnung des Gesamtwertes als gewichtetes Mittel mit ein.

Der NASA-TLX wurde in der Computerform zu zwei Zeitpunkten durchgeführt: Zu Beginn einer großen Pause nach ca. zwei Stunden Arbeitszeit und direkt nach Ende des Versuches nach ca. vier Stunden Arbeitszeit. Der Paarvergleich wurde am Ende jeder Session mit den Probanden vorgenommen.

2.4.2.6 Herzrate und Herzratenvariabilität

Parallel zu den oben erwähnten Verfahren der subjektiven Beanspruchungsmessung wurde während der gesamten Versuchszeit auch der Herzschlag mit einem Polar Brustgurt aufgezeichnet. Ausgewertet wird dieser später nach dem Freiburger Monitoring System (Myrtek, M., Foerster, F., Brügger, G., 2001). Aus dem Herzschlag lassen sich die beiden objektiven, physiologischen Beanspruchungsparameter der Herzrate (Herzschläge pro Minute - Anstieg bei emotionaler Beanspruchung) und Herzratenvariabilität (Variation der Herzschläge pro Minute - Absinken bei mentaler Beanspruchung) ableiten. Hierzu ist es wichtig, dass andere Einflussfaktoren auf die Herztätigkeit minimiert werden. Deswegen wurde beim Versuchssetting z. B. darauf geachtet, dass die Probanden keine größeren Bewegungen machen müssen.

2.4.2.7 Visual Fatigue Questionnaire zu fünf Zeitpunkten

Der Visual Fatigue Questionnaire (VFQ, nach Bangor, 2000, siehe auch Anhang 1) fragt 16 Items auf einer eindimensionalen Skala von 0-10 ab. Alle Items haben einen mehr oder weniger direkten Bezug zu visueller Ermüdung. Bei Items wie „schwere Augen“, „Schwierigkeit, scharf zu sehen“ oder „Blendung durch Lichter“ ist das sofort ersichtlich, aber auch „Kopfschmerzen“, „Nackenschmerzen“ und „geistige Ermüdung“ werden abgefragt. Bei diesem Fragebogen gibt es keinen Gesamt- oder Summenscore, so dass jedes Item für sich selbst interpretiert werden muss.

Der VFQ wurde vor Versuchsbeginn erhoben, um die Ausgangslage festzustellen und dann immer direkt am Ende jedes der vier ca. 50-minütigen Blöcke. Insgesamt liegen also fünf Messzeitpunkte vor.

2.5 **Versuchsablauf**

Die Untersuchungsreihe wurde in den Laborräumen der BAuA in Dortmund durchgeführt. Die Probanden saßen dabei in einem gut ausgeleuchteten Raum vor einem höhenverstellbaren Tisch. Zu Beginn des Versuches wurde der Arbeitsplatz auf die Körpermaße des Probanden eingestellt. Die Versuchsleiter saßen hinter einer Trennwand und überwachten den Versuch, der auch mittels Video komplett aufgezeichnet wurde (siehe Abb. 2.5).

Nach Begrüßung der Probanden erhielten diese zunächst eine Probandeninformation in schriftlicher Form (siehe Anhang 1). Hier wurden die Probanden darüber aufgeklärt, dass sie freiwillig an dem Versuch teilnehmen und ihn jederzeit ohne Angabe von Gründen abbrechen können. Des Weiteren wurde erwähnt, dass der Versuch nicht der medizinischen Diagnostik gilt und dass während des Versuches eine Vide-

aufzeichnung läuft, die allerdings nur zur Kontrolle dient und nicht veröffentlicht wird. Die daran anschließende Einwilligungserklärung mussten die Probanden unterschreiben und anschließend wurden noch einige mögliche medizinische Hindernisgründe für die Untersuchung erfragt. Danach wurde den Probanden der Fragebogen zur Technikaffinität gereicht (siehe Anhang 1). Diese beiden bisher erwähnten Verfahren wurden nur in der ersten Versuchssession durchgeführt. Alle weiteren folgenden Verfahren wurden in jeder Versuchssession durchgeführt.



Abb. 2.5 Versuchsräume der BAuA. Links: Arbeitsplatz der Probanden mit höhenverstellbarem Tisch und Stuhl vor der Trennwand, im Hintergrund das HMD mit Weste. Rechts: Arbeitsplatz der Versuchsleiter hinter der Trennwand.

Dann wurde die RSME-Skala (Zijlstra, 1993) zur mentalen Beanspruchung vorgestellt, die später zu verschiedenen Zeitpunkten im Versuch abgefragt wurde. Damit die Probanden die Skala ein wenig kennenlernen, wurden dann sieben Alltagssituationen beschrieben, die sich die Probanden vorstellen und auf der RSME-Skala hinsichtlich ihrer Beanspruchung bewerten sollten. Die Werte wurden dann in den Beobachtungsbogen eingetragen (siehe Anhang 1). Die RSME-Skala lag den Probanden dabei als einzelnes Blatt im DIN-A4 Format vor und war auch während des gesamten Versuches als Ausdruck auf der Trennwand einsehbar.

Im Anschluss daran wurde der Brustgurt zur Herzratenmessung angelegt und die Baseline der Herzrate über sechs Minuten in absoluter Ruhe aufgezeichnet. Diese Ruhemessung ist notwendig, um die individuell sehr verschiedenen Herzraten auswerten zu können.

Nach der Ruhephasenmessung wurde den Probanden die Instruktion in schriftlicher Form gegeben (siehe Anhang 1). Sobald die Probanden die Instruktion gelesen hatten, wurde nachgefragt, ob sie alles verstanden haben und ob es noch Fragen gibt. Eventuelle Fragen wurden vom Versuchsleiter beantwortet.

Anschließend wurde den Probanden das HMD gezeigt und die Funktionen erläutert. Der Kopfträger wurde vom Versuchsleiter an die Probanden angepasst und das HMD eingestellt. Hierbei wurde sich genügend Zeit genommen, das HMD einzustellen. Während dies bei manchen Probanden sehr einfach und schnell ging, dauerte es bei anderen Probanden deutlich länger, bis eine zufriedenstellende Einstellung gefunden war, die sowohl komfortabel in Bezug auf den Kopfträger war, als auch einen scharfen Blick auf das Display ermöglichte (vgl. 2.3.1). Es wurde den Probanden gegenüber betont, dass die Einstellung des Displays sehr wichtig sei und dass sie sich melden sollten, falls während des Versuches eine Nachjustierung (z. B. durch verrut-

schen des Kopfträgers) nötig sei. Hierbei wurde das monokulare HMD immer vor dem individuellen Führungsauge des Probanden montiert und die Balken immer im äußeren Bereich des Blickfeldes dargestellt. In der ersten Session wurde dann die Kombination der beiden Aufgaben für ca. 10 Minuten geübt.

Der anschließende Versuch erfolgte in vier Blöcken von jeweils ca. 50 Minuten. Zwischen den ersten und letzten beiden Blöcken gab es eine kurze Pause von fünf Minuten, in der das HMD aber anbehalten wurde. Zwischen Block 2 und 3 – in der Mitte des Versuches – gab es eine größere Pause von ca. 15 Minuten, in der das HMD abgenommen wurde und die Probanden aufgefordert waren, sich zur Belastungsvarianz kurz zu bewegen.

Prozessbegleitend wurden mehrere subjektive Beanspruchungsparameter erhoben (siehe auch 2.4.2): Der RSME wurde jeweils in der Mitte eines Blockes und am Ende während der Arbeit abgefragt, so dass insgesamt acht Messzeitpunkte vorliegen. Der NASA-TLX wurde nicht direkt in der Arbeitssituation, sondern unmittelbar zu Beginn der großen Pause und nach Ende des Versuches erhoben, so dass hier zwei Messzeitpunkte zur Verfügung stehen. Der Paarvergleich zur individuellen Gewichtung der Skalen wurde ebenfalls im Anschluss an den Versuch durchgeführt. Des Weiteren wurde der Fragebogen zur visuellen Ermüdung zu fünf Zeitpunkten erhoben: Unmittelbar vor Beginn des Versuches, um die individuelle Ausgangslage festzustellen, und jeweils direkt nach Ende eines jeden Blockes.

Als objektiver Beanspruchungsparameter wurde die Herzrate während des gesamten Versuches erhoben. Im Anschluss an jede Versuchssession wurde mit den Probanden ein Interview durchgeführt.

Abb. 2.6 zeigt eine schematische Darstellung des Versuchsablaufes einer Session im Überblick.

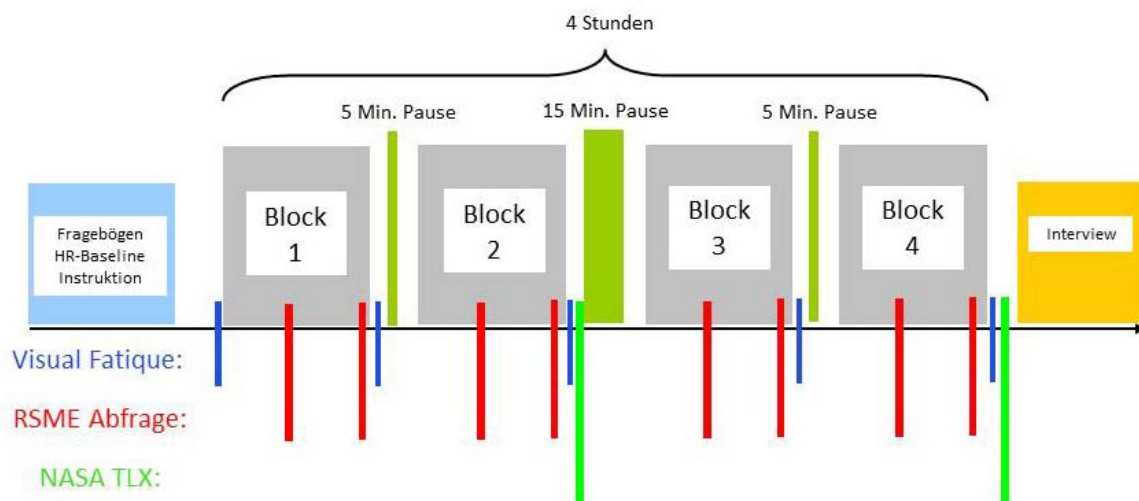


Abb. 2.6 Schematische Darstellung einer Session im Versuchsablauf.

2.6 Probanden

Insgesamt nahmen 44 Probanden an der Untersuchung teil. Drei Probanden haben den Versuch aus unterschiedlichen Gründen abgebrochen: Ein Proband gab während der Untersuchung plötzlich Terminprobleme bekannt und verließ die Studie. Ei-

ne Probandin brach nach einer Stunde ab, da sie die Konstruktionsaufgabe nicht mochte. Neben diesen beiden motivationalen Abbrüchen gab es auch einen Abbruch, der auf den Versuchsaufbau zurückzuführen ist: Eine Probandin bekam Kreislaufprobleme nach ca. einer halben Stunde Arbeitszeit mit dem HMD. Der Versuch wurde abgebrochen und die Probandin ruhte sich auf einer bereitgestellten Liege aus. Die Probandin gab an, auch unter Reisekrankheit zu leiden, und es ist anzunehmen, dass ein Fall von Simulatorkrankheit vorlag.

Bei den verbleibenden 41 Probanden gab es keine Probleme und sie konnten an allen drei Sitzungen teilnehmen. Das Alter der Probanden lag zwischen 18 und 67 Jahren. Die Stichprobe wurde später mittels Mediansplit in zwei Altersgruppen unterteilt, um so möglichen Alterseffekten nachzugehen. In der Gruppe der Jüngeren waren 21 Probanden im Alter zwischen 18-31 Jahren ($M_w = 23,48$; $S_d = 3,341$), wovon 9 männlich und 12 weiblich waren. In der Gruppe der Älteren waren 20 Probanden im Alter von 38 bis 67 Jahren ($M_w = 50,55$; $S_d = 9,214$), wovon 11 männlich und 9 weiblich waren. Keiner der Probanden hatte vorher schon mit einem HMD gearbeitet.

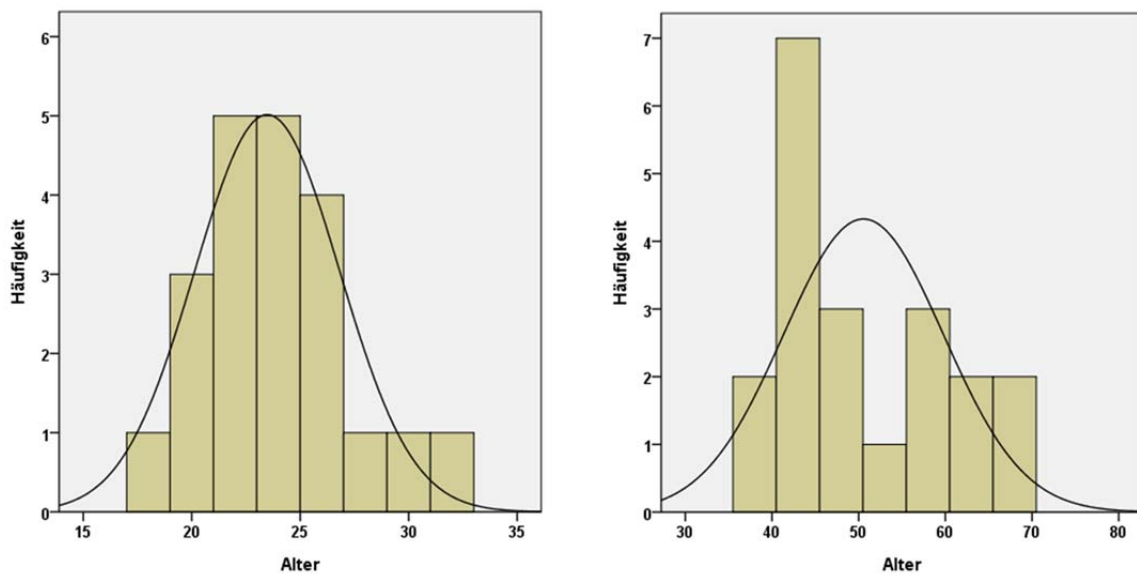


Abb. 2.7 Häufigkeitsverteilung des Alters der Probanden, unterteilt nach gebildeten Gruppen.

2.7 Hypothesen

Sämtliche Hypothesen, die in dieser Studie untersucht werden, sind ungerichtete Unterschiedshypothesen. Es wird also geprüft, ob es einen signifikanten Unterschied macht, ob mit HMD oder Tablet-PC gearbeitet wird, aber es gibt im Vorfeld keine konkrete Aussage über die erwartete Richtung des Effektes. Dies deckt sich auch gut mit dem verwendeten inferenzstatistischen Verfahren der Varianzanalyse im Rahmen des allgemeinen linearen Modells, das immer zweiseitig testet.

Um den Einfluss des Displaytypus zu ermitteln, wird die Session mit dem Tablet-PC (Tab) immer mit der zweiten Messung mit dem HMD (HMD2) verglichen, da die Probanden dann schon eine gewisse Vorerfahrung mit der neuen Technologie haben.

Um Effekten der Gewöhnung nachzugehen, wird immer die erste Messung mit dem HMD (HMD1) mit der zweiten Messung mit dem HMD (HMD2) verglichen.

Alter und Technikaffinität werden als betweensubject Faktor mit einem Mediansplit der Population betrachtet, aber für eine bessere Übersicht hier nicht in den Hypothesen aufgeführt, da sie zwar als potentiell beeinflussende Faktoren anerkannt werden, aber nicht Teil der Hauptintention dieser Studie sind.

2.7.1 Leistung in der Konstruktionsaufgabe

Einfluss des Displaytypus:

H1a: Die Anzahl der bearbeiteten Konstruktionsfolien pro Session unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

Gewöhnung an das HMD:

H1b: Die Anzahl der bearbeiteten Konstruktionsfolien pro Session unterscheidet sich zwischen den Bedingungen HMD1 und HMD2.

2.7.2 Trefferquote bei den Monitoringaufgaben

Einfluss des Displaytypus:

H2a: Die Trefferquote bei den Monitoringaufgaben unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

Gewöhnung an das HMD:

H2b: Die Trefferquote bei den Monitoringaufgaben unterscheidet sich zwischen den Bedingungen HMD1 und HMD2.

Einfluss der Rückmeldung:

H2c: Die Trefferquote bei den Monitoringaufgaben unterscheidet sich zwischen den Blöcken mit und ohne Rückmeldung.

Einfluss des Aufgabentypus:

H2d: Die Trefferquote bei den Monitoringaufgaben „Farbwechsel“ und „Positionswechsel des längsten Balkens“ unterscheidet sich.

2.7.3 Reaktionszeit bei den Monitoringaufgaben

Einfluss des Displaytypus:

H3a: Die Reaktionszeit bei den Monitoringaufgaben unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

Gewöhnung an das HMD:

H3b: Die Reaktionszeit bei den Monitoringaufgaben unterscheidet sich zwischen den Bedingungen HMD1 und HMD2.

Einfluss der Rückmeldung:

H3c: Die Reaktionszeit bei den Monitoringaufgaben unterscheidet sich zwischen den Blöcken mit und ohne Rückmeldung.

Einfluss des Aufgabentypus:

H3d: Die Reaktionszeit bei den Monitoringaufgaben „Farbwechsel“ und „Positionswechsel des längsten Balkens“ unterscheidet sich.

2.7.4 Subjektive Beanspruchung

Einfluss des Displaytypus:

H4a: Die NASA-TLX-Werte unterscheiden sich zwischen den Bedingungen HMD2 und Tab.

H4b: Die RSME-Werte unterscheiden sich zwischen den Bedingungen HMD2 und Tab.

Einfluss der Dauer der Arbeit:

H4c: Die NASA-TLX-Werte unterscheiden sich zwischen den Erhebungszeitpunkten.

H4d: Die RSME-Werte unterscheiden sich zwischen den Erhebungszeitpunkten.

Gewöhnung an das HMD:

H4e: Die NASA-TLX-Werte unterscheiden sich zwischen den Bedingungen HMD1 und HMD2.

H4f: Die RSME-Werte unterscheiden sich zwischen den Bedingungen HMD1 und HMD2.

2.7.5 Objektive Beanspruchung

Einfluss des Displaytypus:

H5a: Die mittlere Herzrate der Probanden unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

H5b: Die Herzratenvariabilität der Probanden unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

Gewöhnung an das HMD:

H5c: Die mittlere Herzrate der Probanden unterscheidet sich zwischen den Bedingungen HMD1 und HMD2.

H5d: Die Herzratenvariabilität der Probanden unterscheidet sich zwischen den Bedingungen HMD2 und Tab.

2.7.6 Visuelle Ermüdung

Da der Visual Fatigue Questionnaire (siehe 0) keinen Gesamtmittelwert hat, müssen alle Items einzeln auf die Frage der Displayunterschiede und der Gewöhnung untersucht werden. Eine Ausformulierung der 32 Hypothesen (16 items x 2 Fragestellungen) wäre hier nicht zielführend und würde dem Überblick nicht dienen. Stattdessen werden hier die Items des VFQ aufgelistet:

- Trockene Augen
- Tränende Augen
- Die Augen sind gereizt oder brennen
- Schmerzen in oder um die Augen
- Schwere Augen
- Schwierigkeiten, Linien folgen zu können
- Schwierigkeiten, scharf zu sehen
- Springender Text
- Verschwommene Buchstaben
- Blendung durch Lichter
- Verschwommene Sicht

- Doppelsehen
- Kopfschmerzen
- Nackenschmerzen
- Übelkeit
- Geistige Ermüdung

Das Fehlen eines Gesamtwertes führt hier auch zu einer methodischen Diskussion: Da nun mit vielen einzelnen Items auf ein Gesamtphänomen geschlossen wird, erhöht sich die Wahrscheinlichkeit, dass durch Zufall Signifikanzen auftreten. Von daher muss das Signifikanzniveau angepasst werden. Dies geschieht mit der Bonferroni-Methode, die besagt, dass, wenn man n unabhängige Hypothesen an einem Datensatz testet, die statistische Signifikanz, die für jede Hypothese getrennt benutzt werden soll, $1/n$ der Signifikanz ist, die sich beim Test nur einer Hypothese ergeben würde. Bei einem Standardsignifikanzniveau von 5% würde sich somit ein angepasstes Signifikanzniveau von 0,3125% ergeben oder ein p von .003.

Auf der anderen Seite ist fragwürdig, ob wirklich alle Items des Fragebogens auf das selbe Gesamtphänomen schließen. Denn z. B. „geistige Ermüdung“ und „Schwierigkeit, scharf zu sehen“ sind bei Weitem nicht deckungsgleich, während „Schwierigkeit, scharf zu sehen“ und „Schwierigkeit, Linien folgen zu können“ durchaus auf ein vergleichbares Phänomen („Schärfe der Darstellung“) abzielen. Methodisch wäre es schön, wenn der Autor des Fragebogens hier eine Faktorenanalyse durchgeführt hätte, um einzelne Items zu gruppieren.

Im Sinne einer konservativen Hypothesentestung wird in der Auswertung des Fragebogens zur visuellen Ermüdung das Alphaniveau deutlich angehoben, so dass nur Ergebnisse als relevant gelten, die ein Alpha-Risiko von 0,3% unterschreiten.

2.8 Datenaufbereitung

2.8.1 Konstruktions- und Monitoringaufgaben

Das von der BAuA geschriebene Programm, das die Konstruktionsfolien und die variierenden Balken für die Monitoringaufgaben darstellt, gibt alle Ereignisse mit Zeitstempel wieder. Hier finden sich jeder Tastendruck der Probanden (Start, Folie vor, Folie zurück, Farbwechsel bei den Balken, Positionswechsel des längsten Balkens) und jede maschinelle Veränderung (Farbwechsel zu Rot, Farbwechsel zu Blau, linker Balken ist der Längste, mittlerer Balken ist der Längste, rechter Balken ist der Längste) als Zahlencode mit exakter Zeit in Millisekunden. Diese Datenmatrizen sind immer an das Fahrzeugmodell gebunden, d. h. bei einem neuen Fahrzeug entsteht eine neue Datenmatrize.

In einem ersten Schritt wurde der Legocode für das entsprechende Fahrzeug ergänzt und die Einzelmatrizen zu einer Gesamtmatrize pro Proband und Versuchssession zusammengeführt. Diese Gesamtmatrize wurde dann als Ausgangslage für die Ermittlung verschiedener Kennwerte benutzt.

Um die Anzahl der abgearbeiteten Konstruktionsfolien zu ermitteln, wurden von der Summe der „Taste vor“ Ereignisse, die Summe der „Taste zurück“ Ereignisse abgezogen und so die Nettoschritte – die Anzahl der wirklich bearbeiteten Folien - abgeleitet.

Um die Trefferquote in den Monitoringaufgaben zu ermitteln wurden zunächst zwei neue Matrizen erstellt in denen entweder nur Aktionen der Farbwechselmonitoring-

aufgabe (FaWe) oder der Balkenlängemonitoringaufgabe (BaLe) enthalten waren. Hier wurden nun mittels SPSS-Skript die Aktionen als Treffer gewertet, wenn auf eine Zustandsänderung ein entsprechender Knopfdruck als nächstes erfolgte. Sollte der nächste Event in der Zeitreihe aber wieder eine Zustandsänderung und kein Knopfdruck sein, so wird dieser Trial als Fehler gewertet. Fanden mehrere Knopfdrücke hintereinander ohne zwischenzeitliche Zustandsänderung statt, so wurden die Wiederholungen ebenfalls als Fehler gewertet. Wiederholungen, die aber in einem Zeitraum von unter einer Sekunde stattfanden, wurden nicht als Fehler gewertet, sondern sind eventuell auf Tastenhänger oder motorisch bedingtes Fehlverhalten zurückzuführen und wurden von der Auswertung ausgeschlossen. Die Trefferquote errechnete sich dann jeweils als prozentualer Anteil der richtigen Reaktionen von allen gegebenen Reaktionen.

Bei den Reaktionszeiten, die ebenfalls getrennt für beide Monitoringaufgaben errechnet wurden, wurden in der Matrize zunächst alle Fehler herausgefiltert und dann die Folgewertdifferenz der Zeitachse zwischen Stimulus onset (Veränderung) und Reaktion für alle richtigen Reaktionen gebildet. Hierbei gab es keine Maximalzeit für die Reaktion (z. B. fünf Sekunden), sondern das mögliche Zeitfenster wurde ausschließlich vom nächsten Event bestimmt.

2.8.2 Fragebogendaten

Die Fragebogendaten wurden handschriftlich in SPSS übertragen. Dies gilt sowohl für die RSME- und die VFQ-Daten, aber auch für die eigentlich schon in Computerform erhobenen NASA-TLX-Werte. Denn eine automatische Übertragung aus dem vorhandenen Datenformat wäre zu aufwendig gewesen.

Die Daten wurden vom Versuchsleiter persönlich übertragen und die Richtigkeit stichprobenartig überprüft.

2.8.3 Herzrate

Die mit einem Polar Brustgurt aufgezeichneten Herzratendaten zeigen im Rohformat ausschließlich das Inter Beat Intervall (IBI), also die Zeit zwischen zwei Herzschlägen, in Millisekunden an. Für die geplante Auswertung der objektiven Beanspruchung nach dem Freiburger Monitoring System (Myrtek et al., 2001) müssen zunächst einige Umformungen stattfinden.

Um die Kennwerte „Herzrate“ (HR) und „Herzratenvariabilität“ (HRV) zu erlangen, ist eine aufwendige Aufbereitung der Rohdaten notwendig, die hier kurz dargestellt wird:

- Zunächst wurden die Daten in SPSS eingelesen und mittels Syntax um jeweils VP, Session, Block ergänzt.
- Eine durchlaufende Gesamtzeit wurde aus den aufaddierten IBIs errechnet
- Eine Einteilung in Fünf-Minuten-Abschnitte fand statt
- Die Herzrate in Beats per Minute (bpm) wurde aus den IBIs errechnet
- Die Folgewertdifferenz von HR (bpm) und RR (ms) wurde errechnet.
- Somit ließ sich das mittlere Quadrat sukzessiver Differenzen (MQSD) errechnen, das im Rahmen des ausgewählten Auswertungsverfahrens nötig ist, um die HRV zu bestimmen.

Diese Daten wurde dann für jedes Individuum und jede Versuchssession und Block zunächst in Fünf-Minuten-Abschnitten geplottet, um so eine erst visuelle Inspektion

der Daten vorzunehmen (vgl. Abb. 2.8). Die geplotteten Rohdaten zeigen noch viele Artefakte, die z.B. durch Kontaktprobleme der Elektroden entstehen können. Solche Artefakte sind bei der Aufzeichnung physiologischer Werte durchaus normal. In diesem speziellen Versuchsssetup entstehen sie zum einen durch die lange Tragezeit der Elektroden über eine ganze Versuchssession von vier Stunden und zum anderen durch die sitzende Position der Probanden, die je nach Körperumfang und Haltung ein Ablösen der Elektroden bewirken kann. Eine Korrektur der Artefakte ist zwingend erforderlich, da durch die Quadrierung der Folgewertdifferenz selbst einzelne solcher Abweichungen ganze Messreihen verfälschen können.

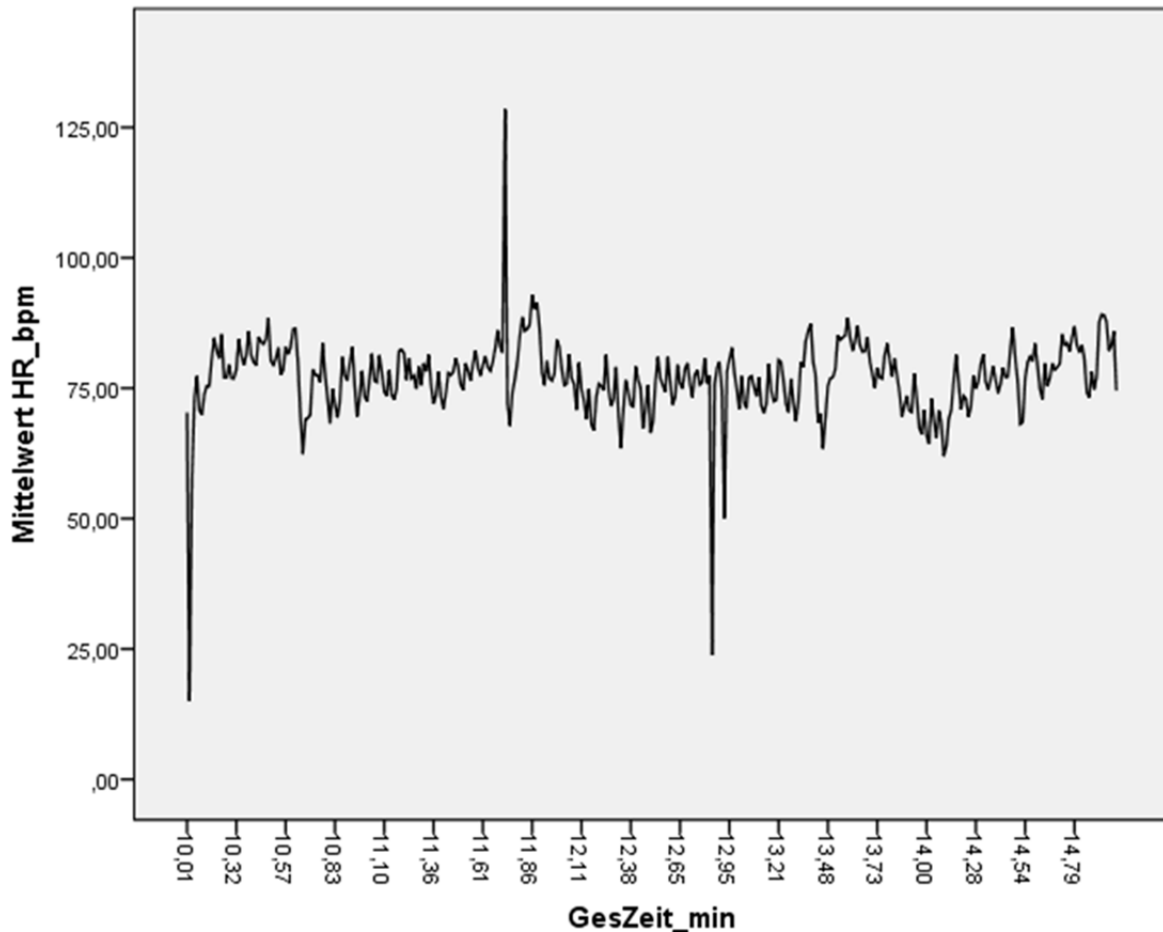


Abb. 2.8 Herzrattendaten eines bestimmten fünfminütigen Abschnittes ohne Filter. Deutliche Artefakte mit unrealistischen Werten sind erkennbar.

Als Artefaktkorrektur wurde eine modifizierte Version des Filters nach Mulder (1988) verwendet. Die Regeln der Artefaktdetektion nach Mulder (1988, S. 33, zitiert nach: www.walterpiechulla.de/seminarPapers/Herzratenindikatoren.pdf, S. 14-15) lauten:

- Keine Korrekturen durchführen, wenn nicht völlig klar ist, ob ein Artefakt aufgetreten ist oder nicht.
- Wenn dasselbe abweichende IBI-Muster wiederholt auftritt, ohne dass es entweder als Messfehler oder als Rhythmusstörung interpretiert werden kann, dann sollte ernsthaft erwogen werden, die Aufzeichnung von der Auswertung auszuschließen.

- „Löcher“ in der IBI-Reihe größer als fünf Sekunden sollten nicht korrigiert werden. Die Zeitreihe sollte entweder in zwei Reihen aufgespalten werden, die getrennt analysiert werden oder ganz von der Auswertung ausgeschlossen werden.
- Reihen, in denen die summierte Artefaktzeit fünf Prozent oder mehr der Aufzeichnungszeit erreicht, sollten von der Analyse ausgeschlossen werden.
- Aus dem mittleren IBI der vergangenen 30 Sekunden wird der Mittelwert EIBI (expected IBI) berechnet, ein IBI ist ein „Loch“, wenn es größer als zwei mal EIBI ist und ein „verkürztes IBI“, wenn es entweder kürzer als 300 msec oder 0,6 mal EIBI ist. Verkürzte IBIs sollten mit dem Vorgänger und dem Nachfolger verglichen werden, um die Art der Rhythmusstörung festzustellen.

Die Regeln der Artefaktkorrektur nach Mulder lauten:

- Verkürzte IBIs werden einfach entfernt, egal ob es sich um einen *false alarm* der R-Zacken-Erkennung oder eine eingestreute Extrasystole handelt.
- Löcher werden linear interpoliert.
- Kompensierte Extrasystolen werden entfernt, das entstehende Loch wird linear interpoliert.
- Phasenverschiebungen werden entweder gar nicht korrigiert oder das Ereignis, das das verkürzte IBI verursacht, wird entfernt und das entstehende Loch linear interpoliert.

Bei einem erneuten Plotten der Daten nach Anwendung der Mulder-Filter zeigte sich aber, dass einige Anpassungen notwendig sind. Mulder toleriert z. B: bis zu zweimal größere IBIs. Das führt aber dazu, dass Ausfälle einzelner Herzschläge, die in dem vorliegenden Setup häufig vorkommen, teilweise nicht erkannt werden. Daher wurde das maximale IBI auf 1,6 des Durchschnitts-IBI herunter genommen. Des Weiteren gibt Mulder 30 Sekunden für den zentrierten gleitenden Durchschnitt an, was im Rahmen der Datenaufbereitung mit SPSS so nicht möglich ist, weswegen ein gleitender Durchschnitt von 50 Werten genommen wurde, der den 30 Sekunden nur ungefähr entspricht. Ausgeschlossen werden also folgende IBIs:

- Kleiner als 300 Millisekunden
- Größer als 1,6 mal bzw. kleiner als 0,6 mal des zentrierten gleitenden Durchschnitts über 50 Werte ist

Dieser Filter wurde auf alle Datensätze angewendet und die Daten erneut geplottet (vgl. Abb. 2.9). Hierbei zeigte sich eine deutliche Verringerung der Artefakte. Einzelne Artefakte blieben aber bestehen. Zumeist handelte es sich hier um Herzschläge, die am Anfang oder Ende einer automatisch ausgefilterten Datensektion im Übergang zu reliablen Daten waren. Diese Artefakte wurden über die neuen Grafiken zunächst identifiziert und dann direkt in der Datenmatrix gefiltert. Hierbei wurde konservativ vorgegangen, so dass im Zweifel, ob der entsprechende Wert eine normale Variation des Herzschlages ist oder es sich um ein messungsbedingtes Artefakt handelt, der Wert aus der Datenreihe herausgenommen wurde. Die so aufbereiteten Rohdaten wurden dann erneut für jeden einzelnen fünfminütigen Abschnitt jedes Blockes und jeder Versuchssession für alle Individuen geplottet (vgl. Abb. 2.10). Dann wurde eine finale Sichtung aller Daten durch den Versuchsleiter vorgenom-

men, um so auszuschließen, dass noch Artefakte verblieben sind, die die Messreihen verfälschen könnten.

Aus den so bereinigten Daten wurden dann für jeden fünfminütigen Abschnitt die mittlere, minimale und maximale Herzrate sowie die Standardabweichung als Kennwert für die weitere Analyse berechnet. Für die Herzratenvariabilität wurde für jeden Abschnitt das mittlere Quadrat sukzessiver Differenzen (MQSD) errechnet, daraus die Wurzel gezogen und dies auf die individuelle Baseline der Herzrate, die vor Versuchsbeginn erhoben wurde, bezogen. So dass der eigentliche Kennwert die prozentuale Abweichung der Herzratenvariabilität von der Ruhephase ist.

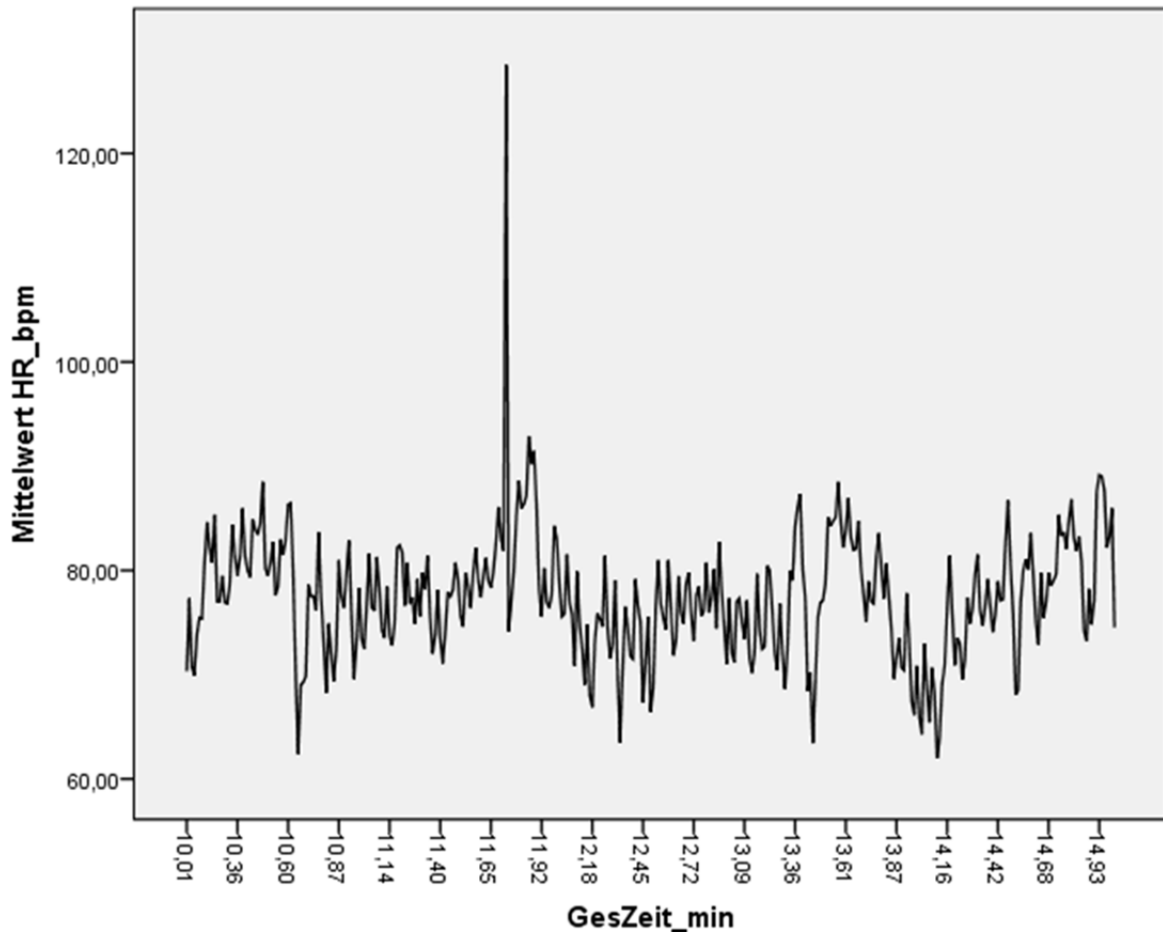


Abb. 2.9 Verlauf des gleichen fünfminütigen Abschnittes der Herzrate nach automatischer Artefaktkorrektur mit modifiziertem Filter nach Mulder. Ein einzelnes Artefakt ist noch erkennbar.

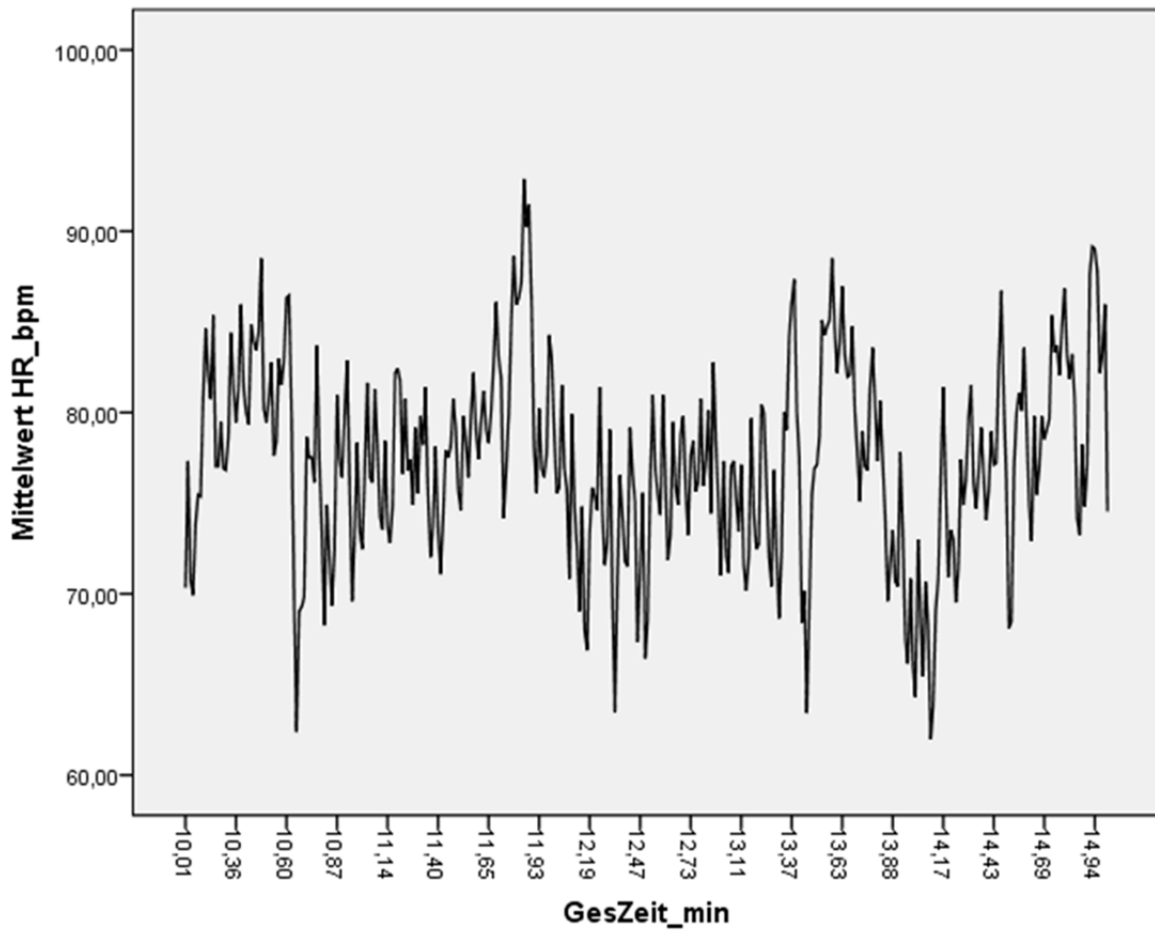


Abb. 2.10 Verlauf des gleichen fünfminütigen Abschnittes der Herzrate nach zusätzlicher manueller Artefaktkorrektur. Kein Artefakt ist mehr erkennbar.

2.9 Inferenzstatistische Verfahren

Als inferenzstatistische Verfahren wurde bei allen Vergleichen eine Varianzanalyse mit Messwiederholung im Mixdesign in SPSS21 verwendet. Hierbei waren die Versuchssession (und somit auch der Displaytyp) und die verschiedenen Zeitpunkte der jeweiligen Erhebung within-subject Faktoren, und die Altersgruppe und die Gruppe der Technikaffinität waren between-subject Faktoren.

3 Ergebnisse

3.1 Leistung in der Konstruktionsaufgabe

Als Leistungsindikator in der Konstruktionsaufgabe wird die Anzahl der bearbeiteten Konstruktionsfolien pro Versuchssession verwendet. Dies ist möglich, obwohl unterschiedlich komplexe Modelle von den Probanden aufgebaut wurden, da sich die einzelnen Konstruktionsfolien in der Komplexität ähneln und komplexere Modelle aus einfach mehr Konstruktionsfolien bestehen.

Bei grafischer Betrachtung der Ergebnisse (vgl. Abb. 3.1) fällt sofort auf, dass die Gruppe der älteren Probanden weniger Folien bearbeiten als die jüngeren Probanden. Der Faktor Alter zeigt auch bei allen Vergleichen einen hochsignifikanten Einfluss auf die Anzahl der bearbeiteten Konstruktionsfolien [$F(1,37) = 23.647$; $p < .001$]. Hierbei ist zu erwähnen, dass dies nicht auf die Vorerfahrung der Gruppen mit Lego-Technic zurückzuführen ist, da hier viele Ältere über aktuelle Vorerfahrung verfügten, weil sie mit ihren Enkeln Lego bauen. Vielmehr ist dieser Effekt wahrscheinlich auf nachlassende Schnelligkeit oder feinmotorische Fähigkeiten im Alter allgemein zurückzuführen.

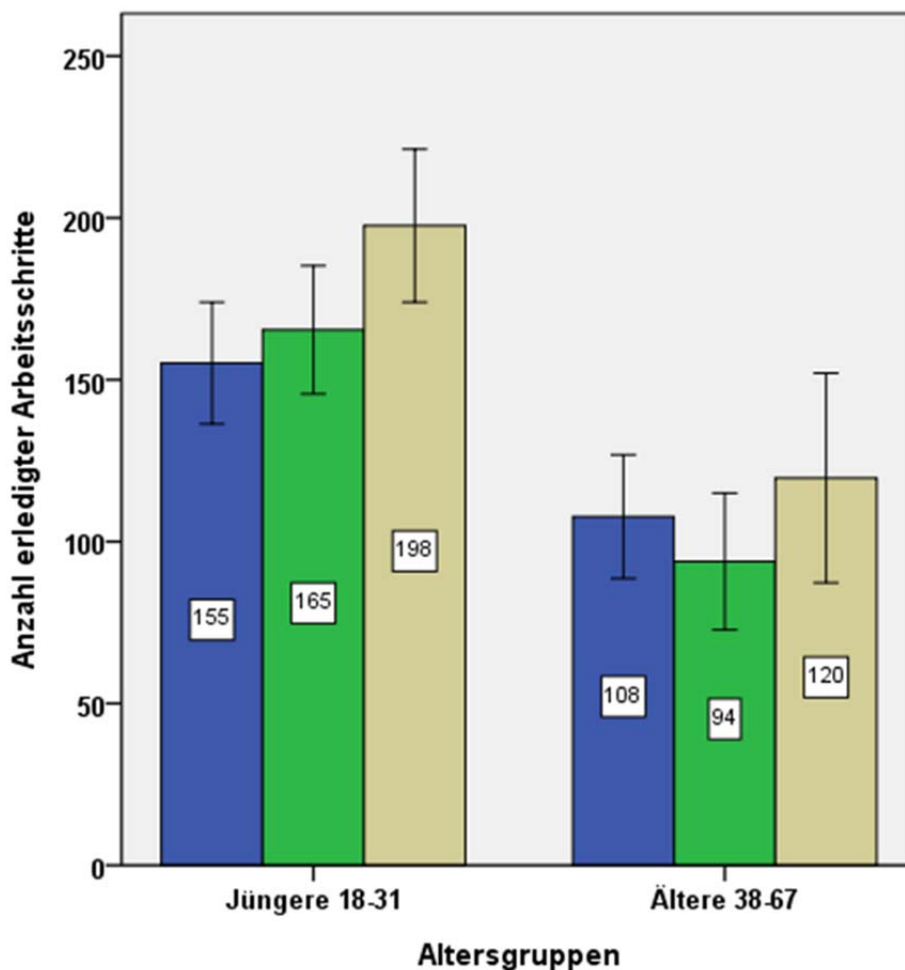


Abb. 3.1 Anzahl bearbeiteter Konstruktionsfolien pro Versuchssession, unterteilt nach Altersgruppen. Blau = 1. Messung HMD, Grün = 2. Messung HMD, Beige = Tablet-PC. Die Whiskers zeigen das 95%-Konfidenzintervall.

Der Einfluss der verwendeten Anzeigentechnologie zeigt sich beim Vergleich zwischen der zweiten Messung mit dem HMD (grüne Balken) und dem Tablet-PC (beige Balken). Hier ist festzustellen, dass mit dem Tablet-PC signifikant mehr Folien abgearbeitet werden, als mit dem HMD [$F(1,37) = 24.285$; $p < .001$]. Der Unterschied beträgt bei den Älteren 27%, bei den Jüngeren, die allgemein fast die doppelte Anzahl an Folien bearbeiten, immerhin noch 20%. Die Hypothese 1a kann also bestätigt werden und es ist festzuhalten, dass in der Konstruktionsaufgabe signifikant weniger Arbeitsschritte mit dem HMD erledigt werden, als wenn dieselben Inhalte auf einem Tablet-PC präsentiert werden.

Eine Gewöhnung an die Displaytechnologie zeigt der Vergleich zwischen erster und zweiter Messung mit dem HMD. Hier sind keine signifikanten Unterschiede in der Anzahl der bearbeiteten Folien zu finden [$F(1,37) = .208$; $p = .651$]. Die Hypothese 1b muss also zurückgewiesen werden: Es ist kein Effekt der Gewöhnung an die HMD-Technologie auszumachen.

Bei diesem statistischen Vergleich zeigt sich auch eine signifikante Wechselwirkung zwischen Alter und Anzahl der bearbeiteten Folien [$F(1,37) = 9.444$; $p = .004$]: Ältere machen weniger Arbeitsfolien bei der zweiten HMD-Session, während jüngere Probanden mehr Folien im zweiten HMD-Durchgang bearbeiten als während der ersten Session. Dieser Effekt ist nur schwer zu interpretieren. Er könnte auf ein bestimmtes Lego-Modell zurückzuführen sein, das im zweiten HMD-Durchgang bearbeitet wurde: Hierbei handelte es sich um einen Traktor mit Anhänger und Pneumatikfunktion, bei der u. a. viele kleine Plastikschräuche auf kleine Ansatzstücke aufgebracht werden mussten. Diese Arbeitsschritte stellen zwar nur einen kleinen Teil des Modells dar, erfordern aber noch mehr Feinmotorik als bei allen anderen Modellen, weswegen der Sachverhalt hier kurz erwähnt wird.

Ein Einfluss der Technikaffinität konnte bei der Leistung in der Konstruktionsaufgabe nicht ausgemacht werden.

3.2 Trefferquote in den Monitoringaufgaben

Die Trefferquote bei den Monitoringaufgaben zeigt an, in wieviel Prozent der Fälle richtig reagiert wurde. Auslasser in der Reaktion oder erneutes Drücken, obwohl sich der Zustand nicht geändert hat, werden als Fehler angesehen.

Bei zunächst grafischer Betrachtung der Ergebnisse (vgl. Abb. 3.2) fällt auf, dass die Trefferquote bei dem zu entdeckenden Farbwechsel aller Balken (blaue Balken in Abb. 3.2) immer größer war, als die Trefferquote bei einem Wechsel der Position des Längsten von den drei Balken (grüne Balken in Abb. 3.2). Dieses Ergebnis wird als Haupteffekt „Aufgabentyp“ signifikant [$F(1,36) = 84.690$; $p < .001$] und war auch so zu erwarten, bedingt durch den Popouteffekt der sich verändernden großen Farbfläche. Die Hypothese H2d kann also angenommen werden, und es ist festzuhalten, dass auf die Veränderung einer Farbfläche zuverlässiger reagiert wird als auf die Veränderung von Längenverhältnissen in einer Grafik.

Ebenfalls gut zu sehen ist, dass mit einem gegebenen Feedback die Trefferquote immer höher ist als ohne Feedback. Dieser Effekt wird ebenfalls hoch signifikant [$F(1,37) = 58.257$; $p < .001$], und die Hypothese H2c kann somit angenommen werden: Die Rückmeldung über die letzte bestätigte Position erhöht die Trefferquote.

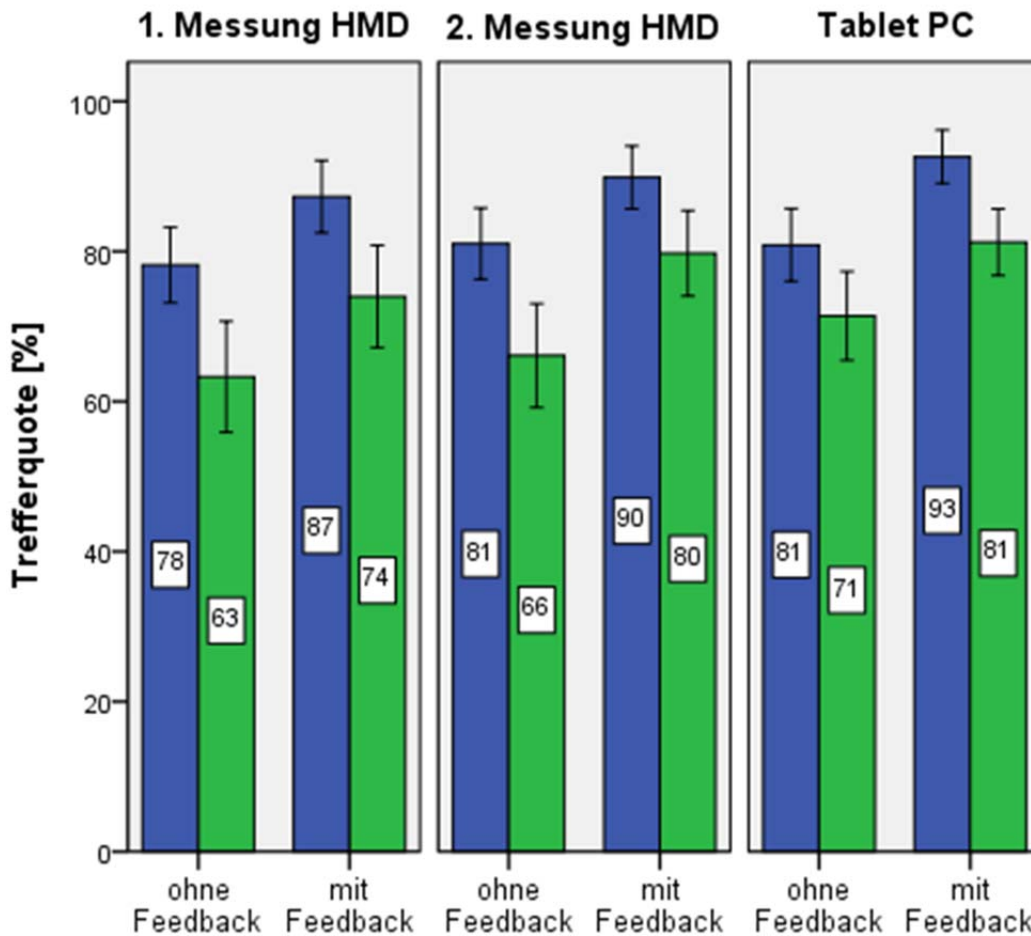


Abb. 3.2 Trefferquote bei den beiden Monitoringaufgaben, unterteilt nach Versuchssession und gegebener Rückmeldung. Blau = Reaktion auf Farbveränderung; Grün = Reaktion auf Wechsel der Position des längsten Balkens. Die Whiskers zeigen das 95%-Konfidenzintervall.

Ein Effekt der Displaytechnologie beim Vergleich zwischen zweiter Messung mit dem HMD und dem Tablet-PC zeigt keinen signifikanten Unterschied [$F(1,37) = 3.635$; $p = .065$], wohl aber statistische Tendenz, da das Signifikanzniveau nur knapp verfehlt wurde. Die Richtung geht hier zu einer besseren Trefferquote beim Tablet-PC. Dies ist erstaunlich, da das HMD ja immer vor den Augen und somit die Anzeige im Sichtfeld des Probanden war, während der Tablet-PC auf dem Tisch abgelegt wurde und somit nicht zwangsläufig immer im Sichtfeld des Probanden war. Die Hypothese 2a muss somit zurückgewiesen werden, und es ist festzustellen, dass es keinen eindeutigen signifikanten Unterschied in der Trefferquote der Monitoringaufgabe zwischen HMD und Tablet-PC gibt, wobei es aber eine interpretationswürdige Tendenz zur besseren Trefferquote beim Tablet-PC gibt. Hier zeigt sich auch eine signifikante Wechselwirkung zwischen Aufgabe und Display [$F(1,37) = 4.286$; $p = .046$]: Der Unterschied in der Trefferquote zwischen beiden Monitoringaufgaben ist beim HMD größer als beim Tablet-PC.

Ein Effekt der Gewöhnung, wie es sich aus dem Vergleich zwischen erster und zweiter Messung mit dem HMD ergibt, wird ebenfalls signifikant [$F(1,37) = 5.395$; $p = .027$]. Die Hypothese 2b kann also angenommen werden, und es ist festzustellen,

len, dass die Trefferquote bei einer zweiten Erhebung besser ist als bei der ersten Erhebung mit der HMD-Technologie.

Ein Effekt des Alters oder der Technologieaffinität lässt sich bei der Trefferquote in den Monitoringaufgaben nicht finden.

3.3 Reaktionszeit in den Monitoringaufgaben

Bei den Reaktionszeiten zeigt sich wieder ein deutlicher Unterschied zwischen den beiden Monitoringaufgaben (vgl. Abb. 3.3): Die Reaktionszeit auf den Farbwechsel der Balken (blaue Balken in Abb. 3.3) ist immer deutlich kürzer als auf den Positionswechsel des längsten Balkens (grüne Balken in Abb. 3.3). Dieser Effekt wird auch signifikant [$F(1,36) = 17.560$; $p < .001$], so dass die Hypothese 3d angenommen werden kann: Auf einen großflächigen Farbwechsel wird also nicht nur akkurater, sondern auch schneller reagiert als auf einen Größenwechsel der einzelnen Elemente.

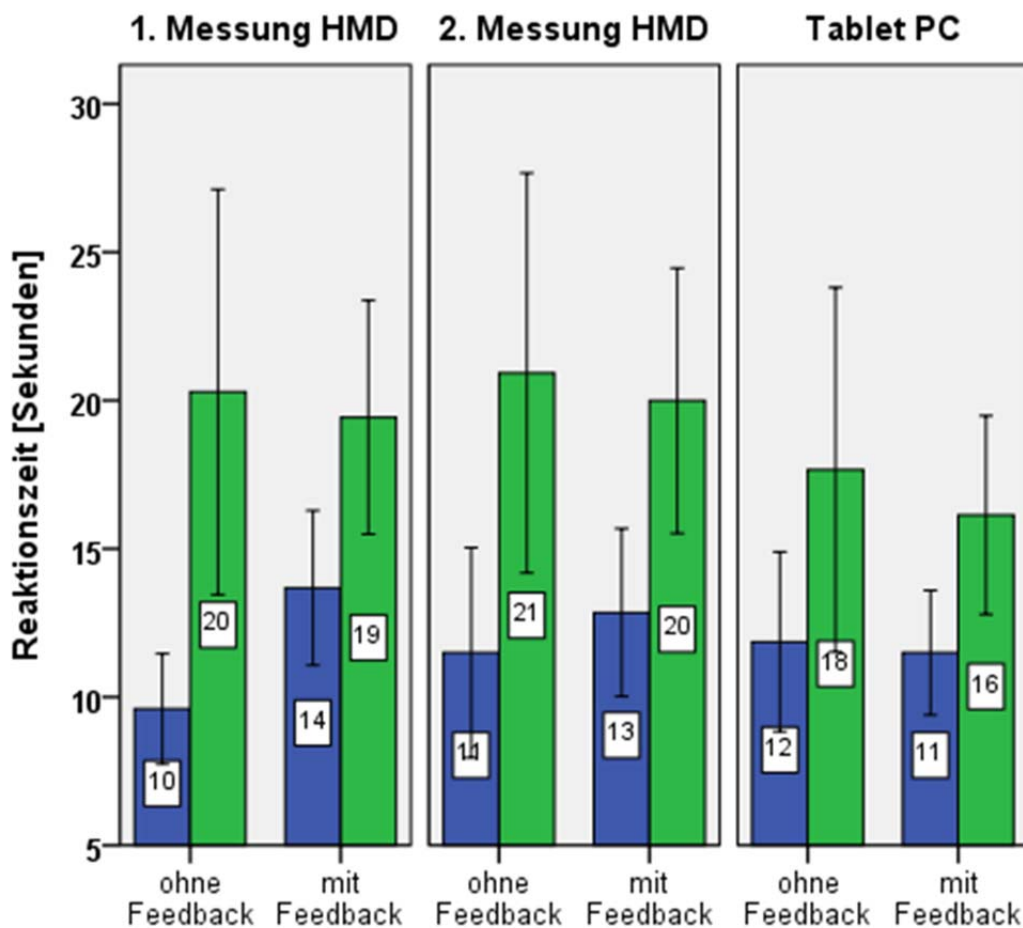


Abb. 3.3 Reaktionszeiten bei den beiden Monitoringaufgaben, unterteilt nach Versuchssession und gegebener Rückmeldung. Blau = Reaktion auf Farbveränderung; Grün = Reaktion auf Wechsel der Position des längsten Balkens. Die Whiskers zeigen das 95%-Konfidenzintervall.

Einen Einfluss des Feedbacks wie bei der Trefferquote gibt es bei der Reaktionszeit nicht [$F(1,37) = .078$; $p = .782$]. Die Hypothese 3c muss also zurückgewiesen werden.

Auch beim Einfluss des Displaytypus auf die Reaktionszeit zeigt sich kein signifikanter Effekt [$F(1,37) = 2,858$; $p = .100$]. Allerdings liegt das Alpha-Risiko mit 10% noch im Bereich der statistischen Tendenz. Bei einseitiger Hypothesentestung (die mit einer Varianzanalyse aber nicht möglich ist) würde hier ein signifikanter Effekt der Displaytechnologie vorliegen. Interessanterweise ist die Richtung des Effektes die gleiche wie bei der Trefferquote (vgl. 3.2): Schlechtere Reaktionszeit beim HMD verglichen mit dem Tablet-PC. Die Hypothese 3a muss trotzdem zurückgewiesen werden: Es zeigt sich kein signifikanter Einfluss des Displaytypus auf die Reaktionszeit in den Monitoringaufgaben. Allerdings zeigt sich auch hier, wie schon bei der Trefferquote, eine signifikante Wechselwirkung zwischen Display und Aufgabe [$F(1,37) = 4.880$; $p = .034$]: Bei dem HMD gibt es einen größeren Aufgabenunterschied als beim Tablet-PC.

Ein Effekt der Gewöhnung an die HMD-Technologie durch den Vergleich zwischen erster und zweiter Session mit dem HMD kann hier aber nicht ausgemacht werden [$F(1,37) = .196$; $p = .661$]. Die Hypothese 3b muss also zurückgewiesen werden. Dieser Vergleich zeigt aber auch eine signifikante Wechselwirkung zwischen Aufgabe und Feedback [$F(1,37) = 6.492$; $p = .015$]: Mit Feedback ist der Aufgabenunterschied kleiner.

Des Weiteren zeigen sich bei der Reaktionszeit keine Effekte des Alters oder der Technikaffinität.

3.4 Subjektive Beanspruchung

Die subjektive Beanspruchung wurde mit zwei Verfahren parallel erfasst: Zum einen der bekannte NASA-TLX und zum anderen die Rating Scale of Mental Effort (RSME). Beide Verfahren erfragen die Beanspruchung direkt vom Probanden, unterscheiden sich aber in ihrer Konzeption (vgl. 2.4.2.4 und 2.4.2.5).

Die parallele Verwendung beider Verfahren kann zum einen der gegenseitigen Validierung der Ergebnisse dienen, zum anderen ist ein Vergleich beider Verfahren bezüglich der Sensitivität interessant.

3.4.1 NASA-TLX

Beim NASA-TLX (siehe Abb. 3.4) zeigt sich ein signifikanter Einfluss des Displaytypus bei einem Vergleich zwischen zweiter HMD-Erhebung und der Session mit dem Tablet-PC [$F(1,37) = 26.952$; $p < .001$]. Die Hypothese 4a kann also angenommen werden, und es ist festzuhalten, dass die Nutzung eines HMDs zu höheren Beanspruchungswerten beim NASA-TLX führt als die Verwendung eines Tablet-PCs.

Eine Steigerung der empfundenen Beanspruchung über die Zeit (innersubject factor: Zeitpunkt, hier nach zwei oder vier Stunden Arbeit) wird ebenso signifikant [$F(1,37) = 14.267$; $p = .001$], allerdings liegt keine Wechselwirkung zwischen Displaytyp und Zeitpunkt vor [$F(1,37) = 1.375$; $p = .253$]. Die Hypothese 4c kann also angenommen werden: Es zeigt sich eine Steigerung der empfundenen Beanspruchung über die Zeit. Allerdings ist dies gleichermaßen für das HMD und den Tablet-PC der Fall, da diesbezüglich keine Wechselwirkung vorliegt.

Eine andere Wechselwirkung zwischen Displaytyp und Alter erreicht aber statistische Tendenz [$F(1,37) = 3.172$; $p = .083$]. Hier ist es so, dass die Beanspruchungswerte bei älteren Probanden zwar allgemein auf einem höheren Niveau sind (HMD: 72.64; Tablet-PC: 64.55), der Anstieg zwischen Tablet-PC und HMD aber geringer ausfällt als bei den Jüngeren (HMD: 69.51; Tablet-PC: 52.98).

Und erstmals zeigt sich auch ein Einfluss von Technikaffinität in Wechselwirkung mit dem Displaytyp [$F(1,37) = 5.000$; $p = .031$] in der Form, dass nicht technikaffine Probanden nicht nur allgemein einen höheren Beanspruchungsscore haben (HMD: 73.18; Tablet-PC: 66.17) als technikaffine Nutzer (HMD: 68.96; Tablet-PC: 51.35), sondern auch weniger von der Entlastung durch den Tablet-PC profitieren.

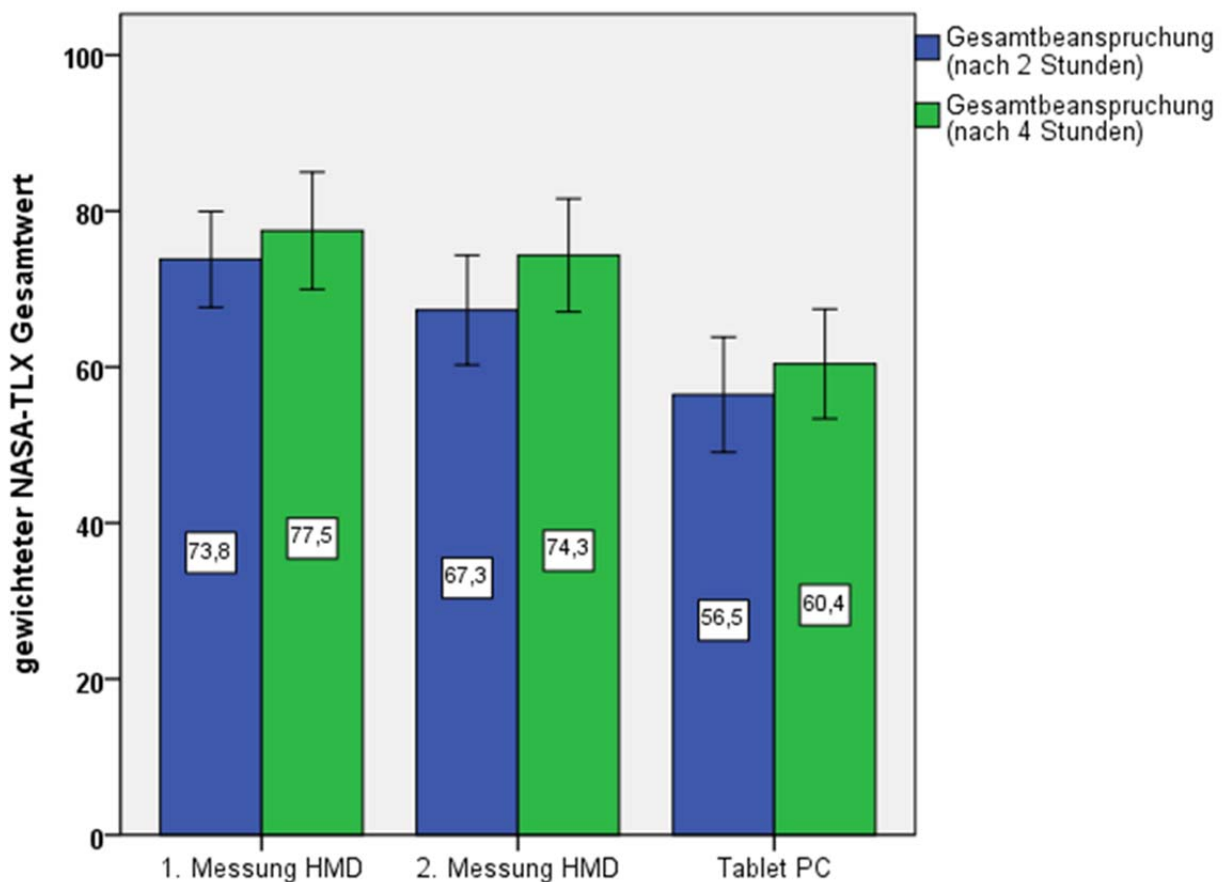


Abb. 3.4 Gewichteter NASA-TLX Gesamtwert nach zwei und vier Stunden unterteilt nach Versuchssession. Die Whiskers zeigen das 95%-Konfidenzintervall.

Ein Gewöhnungseffekt in der Beanspruchung zwischen erster und zweiter Messung mit dem HMD erreicht zwar nicht das Signifikanzniveau, befindet sich aber im Bereich statistischer Tendenz [$F(1,37) = 3.409$; $p = .073$]: Die Beanspruchung beim zweiten Einsatz des HMDs ist also tendenziell geringer, obwohl die Hypothese H4e zurückgewiesen werden muss. Des Weiteren zeigt sich bei diesem Vergleich ein signifikanter Haupteffekt des Zeitpunktes [$F(1,37) = 16.270$; $p < .001$] und eine signifikante Wechselwirkung zwischen Zeitpunkt und Alter [$F(1,37) = 10.258$; $p = .003$]: Bei beiden Probandengruppen steigt die Beanspruchung über die Zeit, bei älteren Probanden aber mehr.

3.4.2 RSME

Auch bei den RSME-Werten zeigt sich ein hochsignifikanter Einfluss des Displaytypus auf die Beanspruchung [$F(1,35) = 84.435$; $p < .001$]. Die Hypothese 4b kann also angenommen werden: Die subjektive Beanspruchung gemessen durch den RSME ist beim HMD höher als beim Tablet-PC.

Ebenso gibt es einen signifikanten Anstieg der Beanspruchung über die Zeit [$F(1,29) = 6.956$; $p < .001$], womit die Hypothese 4d ebenfalls angenommen werden kann. Eine Wechselwirkung zwischen Displaytypus und Zeitpunkt existiert nicht [$F(1,29) = 1.763$; $p = .133$], so dass die Steigerung der Beanspruchung über die Zeit bei beiden Displaytypen als gleichförmig angesehen werden muss.

Ein auftretender Haupteffekt des Alters [$F(1,35) = 6.684$; $p = .014$] gibt wieder, dass Ältere allgemein beanspruchter sind als jüngere Probanden.

Abb. 3.5 zeigt die RSME-Werte im Verlauf für die jeweiligen Versuchssessions, unterteilt nach Altersgruppen.

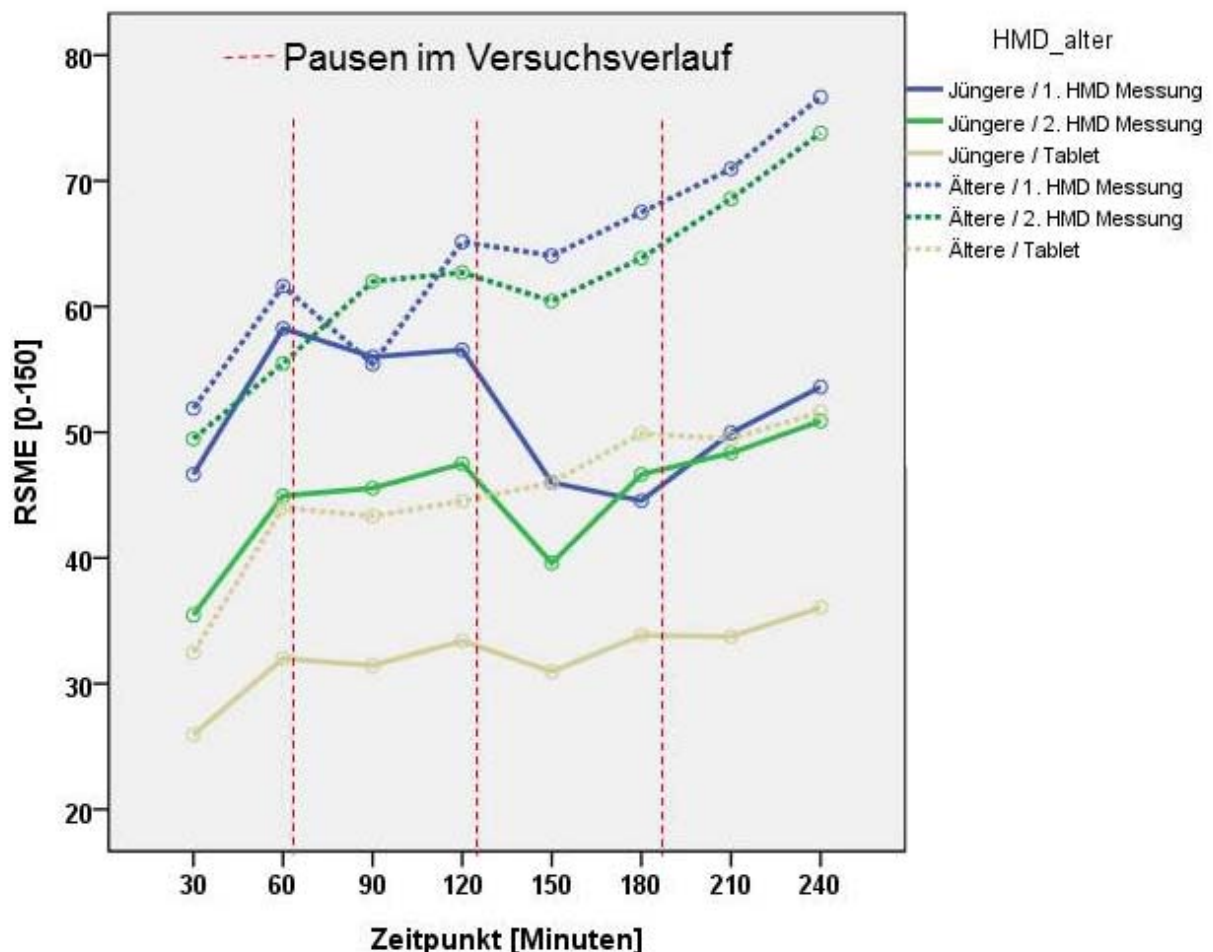


Abb. 3.5 RSME-Werte über den Verlauf der einzelnen Versuchssessions, unterteilt nach Alter und mit grafischem Hinweis auf die Pausen.

Der Gewöhnungseffekt an die HMD-Technologie verpasst auch hier das Signifikanzniveau [$F(1,34) = 3.083$; $p = .088$], erreicht aber ebenso wie beim NASA-TLX statistische Tendenz. Die Hypothese 4f muss also zurückgewiesen werden, obwohl eine

Tendenz zu einer Verringerung der Beanspruchungswerte bei einer zweiten Erhebung mit dem HMD im Vergleich zur ersten Erhebung auszumachen ist.

Erneut signifikant wird ein Haupteffekt des Zeitpunktes [$F(7,28) = 8,432$; $p < .001$] sowie eine Wechselwirkung zwischen Zeitpunkt und Alter [$F(7,28) = 4,276$; $p = .003$], die anzeigt, dass die empfundene Beanspruchung bei Älteren stärker über die Zeit steigt verglichen mit jüngeren Probanden.

3.4.3 Vergleich von NASA-TLX und RSME-Ergebnissen

Die beiden parallel eingesetzten Verfahren zur subjektiven Beanspruchungsmessung zeigen nahezu identische Ergebnisse: Bei beiden gibt es einen signifikanten Haupteffekt des Displaytypus, der besagt, dass die Beanspruchung allgemein beim HMD höher ist als beim Tablet-PC (Hypothesen 4a und 4b). Beide zeigen ebenso eine signifikante Zunahme der Beanspruchung über die Zeit (Hypothesen 4c und 4d), wobei beide aber keine Wechselwirkungen zwischen Displaytyp und Zeitpunkt ausfindig machen können, was bedeutet, dass die Zunahme der Beanspruchung über die Zeit bei HMD und Tablet-PC als gleichförmig zu interpretieren ist. Ebenfalls zeigen beide Verfahren eine statistische Tendenz zur Gewöhnung an die Displaytechnologie bei einem Vergleich zwischen erster und zweiter Erhebung mit dem HMD (Hypothesen 4e und 4f).

In einigen Wechselwirkungen gibt es aber Unterschiede in den Ergebnissen. Während der NASA-TLX eine Tendenz zur Wechselwirkung zwischen Alter und Displaytypus aufzeigt, tut die RSME dies nicht. Ebenso zeigt sich beim NASA-TLX eine Wechselwirkung zwischen Displaytypus und Technikaffinität, die der RSME nicht nachweisen kann. Dahingegen zeigt die RSME einen Haupteffekt des Alters, der im NASA-TLX keine Signifikanz erreicht. Die Wechselwirkung zwischen Alter und Zeitpunkt geben aber wiederum beide Verfahren wieder.

Trotz der Detailunterschiede in den Wechselwirkungen gleichen sich die Ergebnisse beider Verfahren doch erstaunlich. Zwar erfragen beide Verfahren den gleichen Sachverhalt (die subjektive Beanspruchung) und wurden mit den gleichen Probanden durchgeführt, so dass die Ähnlichkeit auch in Bezug auf individuelle Antworttendenzen zu erwarten war; die Exaktheit der gleichen Ergebnisse, inklusive gleichem Ausmaß bei den statistischen Tendenzen, vermag aber schon zu erstaunen, zumal die Konzeption beider Verfahren doch recht unterschiedlich ist.

Da der NASA-TLX viel aufwendiger in der Durchführung ist, ist nach Einschätzung des Autors im Zweifelsfall der RSME vorzuziehen. Es sei denn, man interessiert sich auch für die Subskalen des NASA-TLX, die einzelne Aspekte der Beanspruchung näher beleuchten.

3.5 Objektive Beanspruchung

Objektive Beanspruchungsparameter werden aus physiologischen Reaktionen abgeleitet. Sie unterliegen so im Gegensatz zu den subjektiven Beanspruchungsparametern nicht den individuellen Antworttendenzen der Probanden. Ein Vergleich zwischen subjektiven und objektiven Beanspruchungsparametern kann Hinweise geben, welche Beanspruchungssituationen eher durch die mentale Bewertung der Situation entstehen und welche eher von der Situation an sich geprägt sind (vgl. 3.6).

Die in diesem Versuchssetup gewählten objektiven Beanspruchungsparameter der Herzrate und Herzratenvariabilität (vgl. 2.4.2.6) wurden zunächst sorgsam aufbereitet

(vgl. 2.8.3) und individuelle Kennwerte für jeden Fünfminutenabschnitt über den gesamten Versuchsablauf gebildet. Bei der Hypothesenprüfung werden aber nur bestimmte Fünfminutenabschnitte getestet, da bei zu großer Zellenanzahl im Vergleich die Freiheitsgrade der Varianzanalyse bei gegebener Stichprobengröße zu klein werden. Bevor die Hypothesenprüfung berichtet wird, wird eine Übersicht über den Verlauf der Kennwerte über den gesamten Versuchsablauf gegeben.

3.5.1 Herzrate

Die Herzrate gibt – soweit andere Einflussgrößen wie z. B. die körperliche Bewegung ausgeschlossen werden – die emotionale Beanspruchung wieder: Sie steigt bei höherer Beanspruchung.

Abb. 3.6 zeigt den Verlauf der mittleren Herzrate für alle Fünfminutenabschnitte über den gesamten Versuchsablauf.

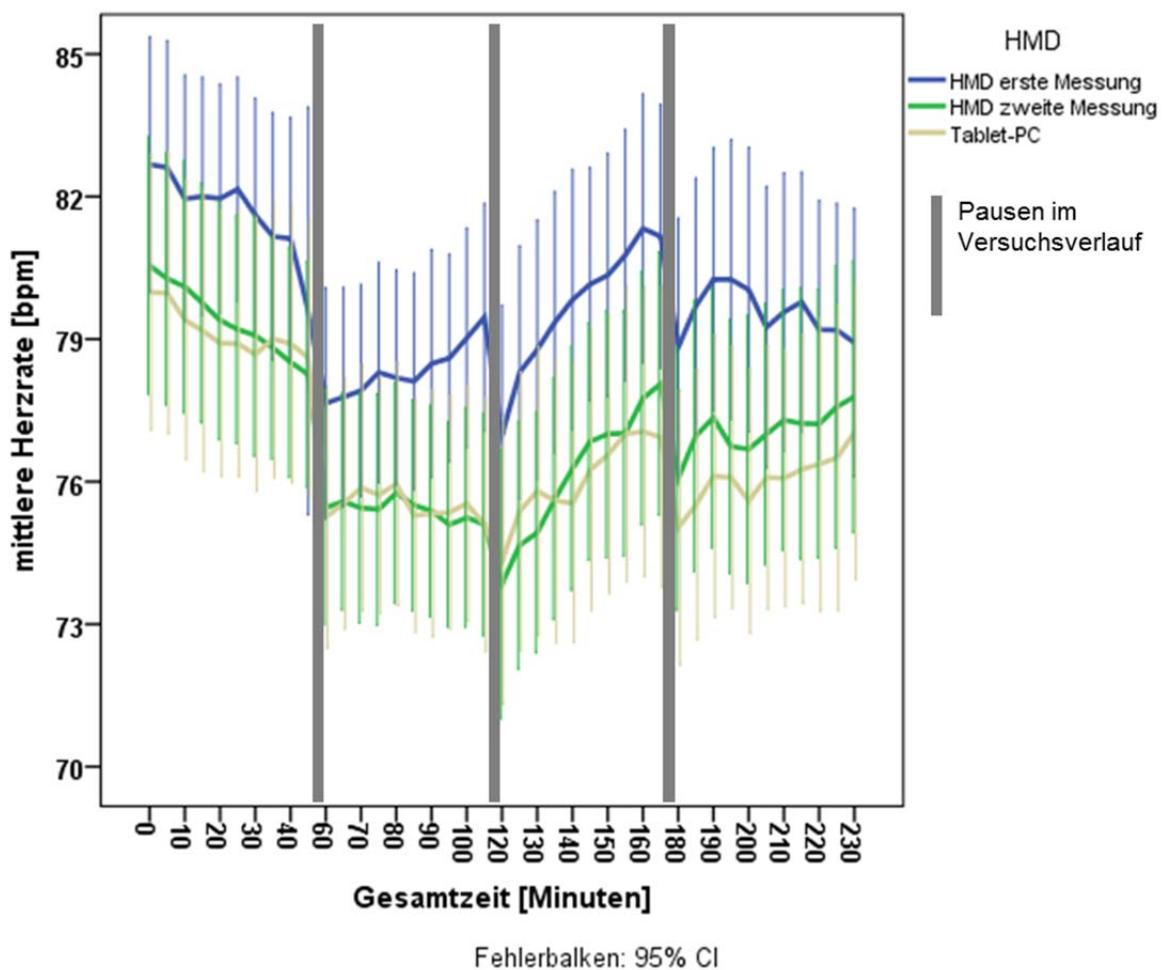


Abb. 3.6 Mittlere Herzrate aller Fünfminutenabschnitte über den gesamten Versuchsablauf. Fehlerbalken geben das 95%-Konfidenzintervall wieder. Pausen wurden vor den Minuten 60, 120 und 180 gemacht und sind durch graue Balken markiert.

Zu erkennen ist ein nahezu paralleler Verlauf der drei Versuchssessions, der sich in einem Abfallen der Herzrate während des ersten Blockes und einem Anstieg der

Herzrate während aller anderen Blöcke zeigt. Das Abfallen während des ersten Blockes, das eine abnehmende emotionale Beanspruchung wiedergibt, könnte auf eine Gewöhnung an die Versuchssituation zurückzuführen sein. Die Zunahme während der folgenden Blöcke, die auch eine Zunahme der emotionalen Beanspruchung wiedergibt, entspricht der zunehmenden Beanspruchung über die Zeit bei anstrengenden Tätigkeiten. Die Pausen zwischen den Blöcken bei den Minuten 60, 120 und 180 zeigen sich deutlich in den Datenreihen und führen immer zu einer Abnahme der Herzrate als Reaktion auf Entspannung in der Pause. Des Weiteren fällt das höhere Niveau der Herzrate während der ersten HMD-Messung auf, während die zweite HMD-Messung und die Messung mit Tablet-PC auf einem vergleichbaren, niedrigeren Niveau stattfinden. Dies kann als ein Hinweis auf höhere emotionale Beanspruchung im ersten Durchgang - eventuell wegen mangelnder Gewöhnung - gesehen werden, während nach der Gewöhnung HMD und Tablet-PC vergleichbar sind.

Für die Überprüfung der Hypothese H5a, die durch einen Vergleich zwischen den Bedingungen HMD2 und Tab den Einfluss des Displaytypus nach Gewöhnung untersucht, wurden acht Messzeitpunkte herangezogen: Die Fünfminutenabschnitte ab den Minuten 20, 45, 80, 105, 140, 165, 200 und 225 (vgl. Abb. 3.7). Diese acht Abschnitte sind jeweils kurz vor den in 3.4.2 ausgewerteten RSME-Abfragen, so dass kein Einfluss der Sprechaktivität oder kognitiver, subjektiver Bewertung der Situation durch die parallele RSME-Abfrage anzunehmen ist.

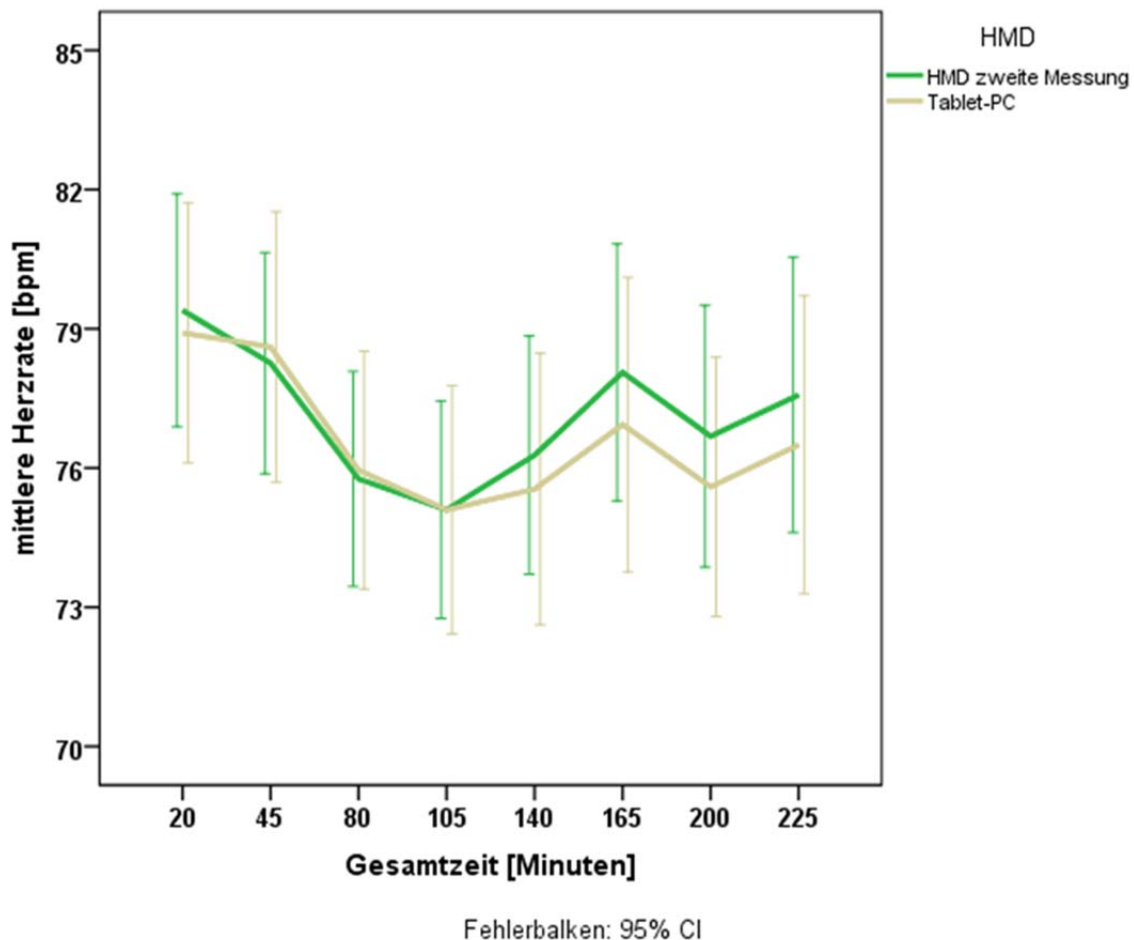


Abb. 3.7 Mittlere Herzrate der zweiten HMD-Messung und des Tablet-PCs zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Ein Haupteffekt des Displaytypus liegt hier nicht vor [$F(1,34) = 0.503$; $p = .483$]. Aus physiologischer Sicht ist die emotionale Beanspruchung also nicht unterschiedlich beim Umgang mit HMD oder Tablet-PC und die Hypothese H5a muss zurückgewiesen werden. Auch Wechselwirkungen, die den Displaytyp mit einschließen, liegen nicht vor, wie dies ist auch in der Grafik durch den parallelen Verlauf zu erkennen ist. Die Variation der Herzrate über die Zeit wird signifikant [$F(7,28) = 18.891$; $p < .001$], wobei hier aber keine einfache Steigerung über die Zeit vorliegt, sondern ein komplexerer Verlauf mit einem Abfallen zu Beginn und einem Anstieg zum Ende. Andere Interaktionen und Wechselwirkungen werden nicht signifikant, so zeigen weder das Alter [$F(1,34) = 1.096$; $p = .308$] noch die Technikaffinität der Probanden [$F(1,34) = 0.054$; $p = .818$] Einfluss.

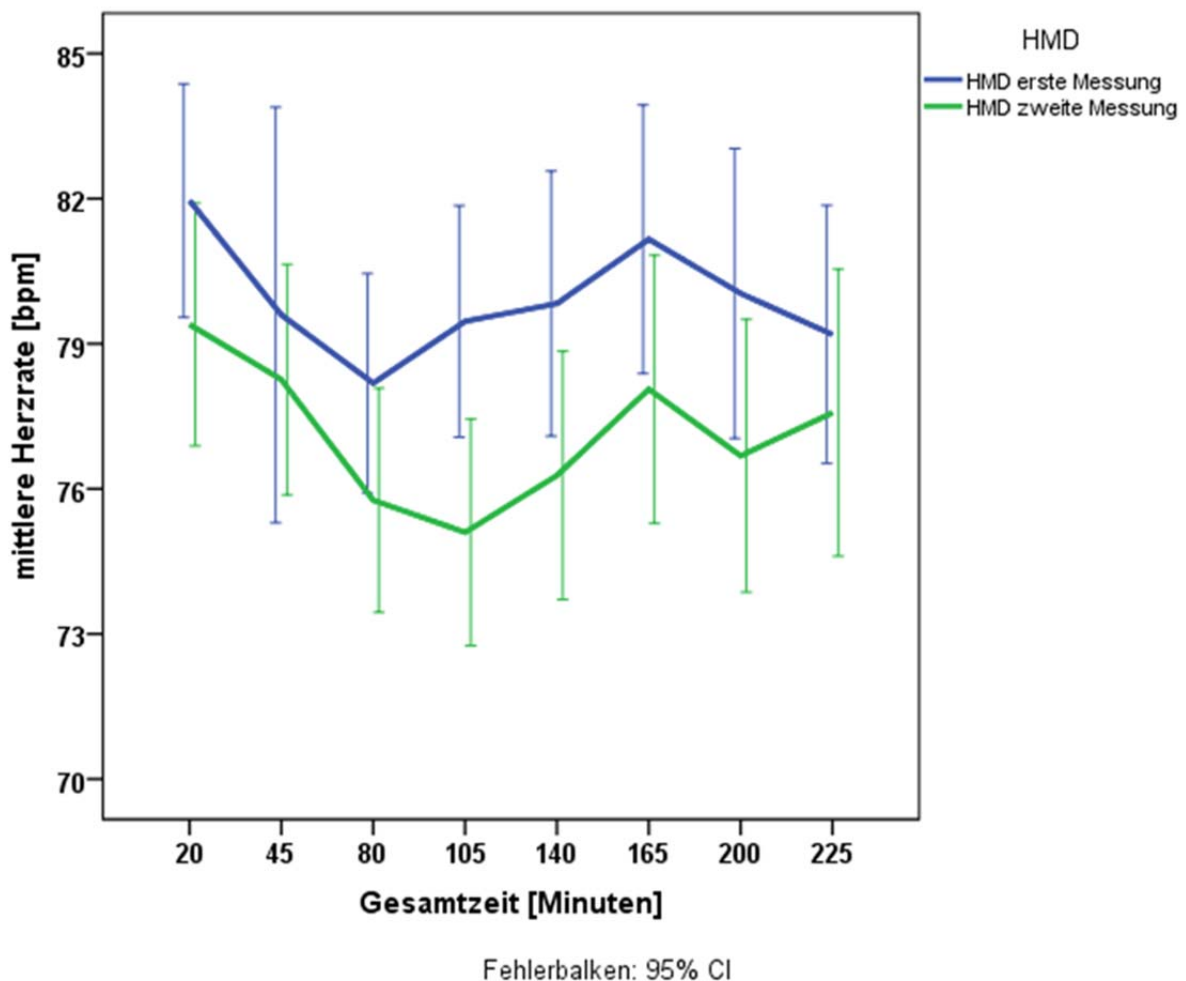


Abb. 3.8 Mittlere Herzrate der ersten und zweiten HMD-Messung zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Bei Überprüfung der Hypothese 5c, die einen Gewöhnungseffekt durch Vergleich der Bedingungen HMD1 und HMD2 untersucht, zeigt sich ein anderes Bild, wie in Abb. 3.8 schon erkennbar ist: deutlich fällt der parallele, aber auf unterschiedlichen Niveau stattfindende Verlauf auf. Es liegt ein signifikanter Haupteffekt des Displaytypus vor [$F(1,34) = 7.840$; $p = .008$], der besagt, dass die erste Messung mit dem HMD als emotional beanspruchender erlebt wird. Die Hypothese H5c ist also anzunehmen und es kommt beim Umgang mit einem HMD mit wachsender Erfahrung zu einer Verringerung der Herzrate und damit einhergehend zu einer Verringerung der

emotionalen Beanspruchung. Ebenfalls signifikant wird die Variation über die Zeit [F (1,34) = 21.670 ; p < .001], wobei auch hier kein linearer Verlauf vorliegt. Bei den Wechselwirkungen erreichen Zeit * Alter [F (7,28) = 2.257 ; p = .059] und Display * Technikaffinität [F (1,34) = 3.830; p = .059] statistische Tendenz. Hierbei sind jüngere Probanden auf einem niedrigeren Niveau und zeigen weniger Variation über die Zeit. Technikaffine Probanden haben bei der ersten HMD-Messung ein höheres Niveau als nichttechnikaffine Probanden, während bei der zweiten HMD-Messung beide auf dem gleichen Niveau sind. Es zeigt sich auch eine signifikante Triplet Wechselwirkung Display * Zeit * Technikaffinität [F(7, 28) = 2.404 ; p = .046], die so zu interpretieren ist, dass der erste HMD-Durchgang vor allem von technikaffinen Probanden als emotional beanspruchend erlebt wird. Dies könnte auf Begeisterung oder Engagement der Probanden hindeuten. Weder Alter [F (1,34) = 2.939 ; p = .096] noch Technikaffinität [F (7,28) = 1.136 ; p = .294] werden aber als einzelne Zwischensubjektoren signifikant.

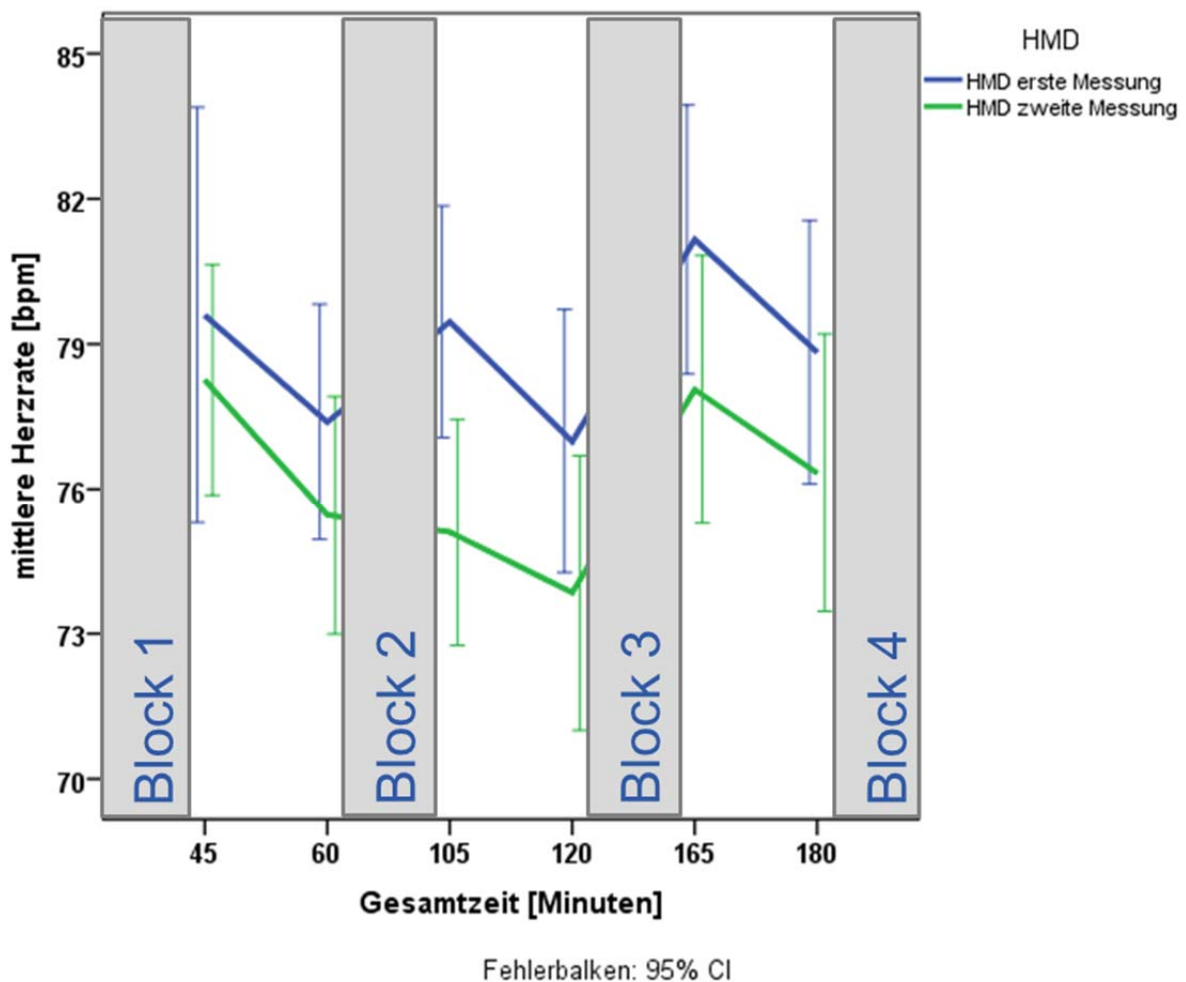


Abb. 3.9 Mittlere Herzrate jeweils kurz vor und nach den drei Pausen. Fehlerbalken geben das 95%-Konfidenzintervall wieder.

Eine Analyse der Wirkung von Pausen auf die Herzrate kann vorgenommen werden, indem man jeweils den Abschnitt kurz vor einer Pause mit dem kurz danach zu Beginn des nächsten Blockes vergleicht. Abb. 3.9 illustriert dies. Hierbei zeigt sich eine signifikante Abnahme der Herzrate durch die Pause [F(1,35) = 20.476 ; p < .001]. Dabei unterscheidet sich die Wirkung der Pausen nicht, obwohl die mittlere Pause

länger war und das HMD abgenommen wurde, während die erste und zweite Pause kürzer waren und das HMD aufbehalten wurde.

3.5.2 Herzratenvariabilität

Die Herzratenvariabilität gibt – bei Gleichhaltung anderer Einflussgrößen wie z. B. körperliche Bewegung – die mentale Beanspruchung wieder. Hierbei ist die Herzratenvariabilität um so geringer je höher die mentale Beanspruchung ist. Im entspannten Ruhezustand schlägt das Herz also unregelmäßiger, während es bei zunehmender mentaler Beanspruchung immer regelmäßiger schlägt.

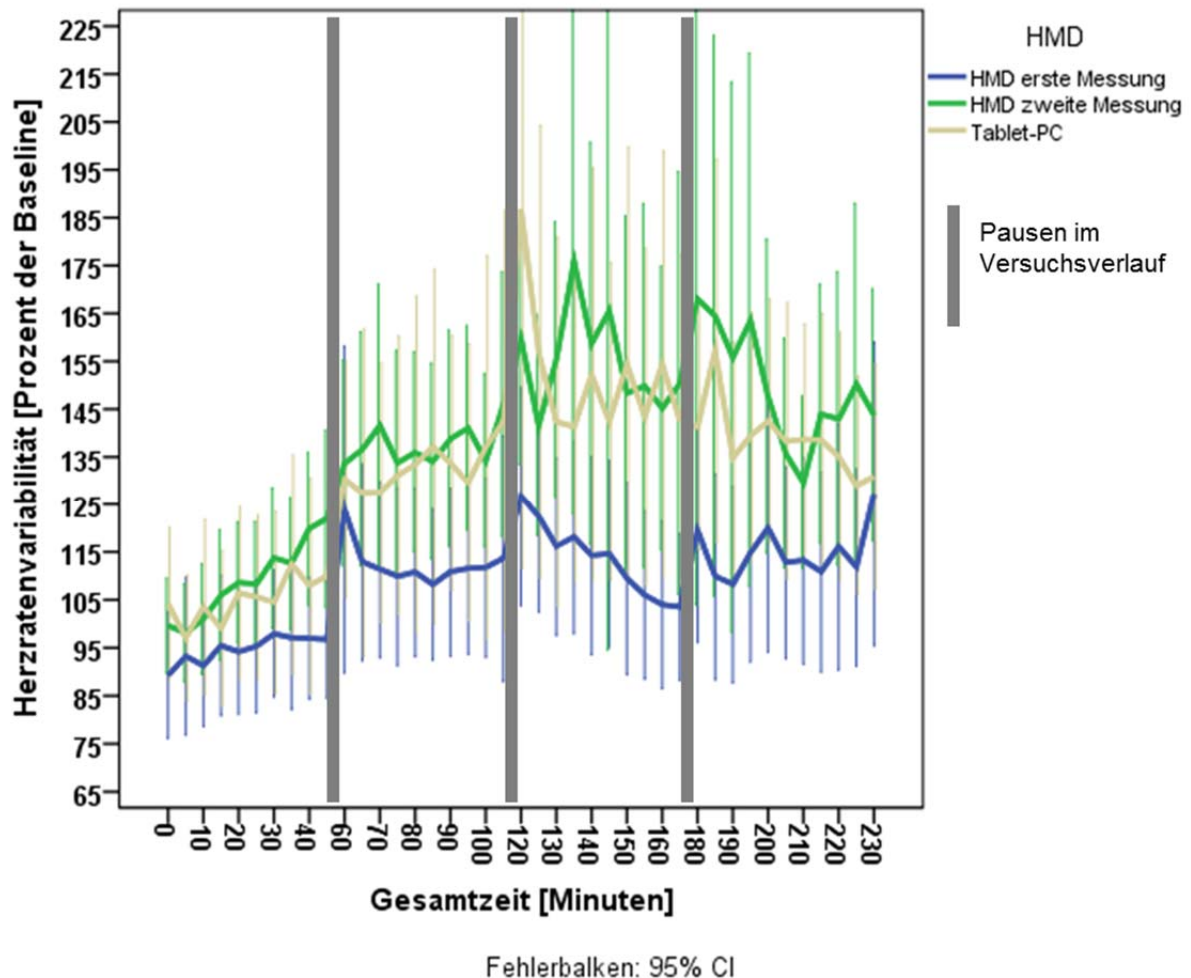


Abb. 3.10 Mittlere Herzratenvariabilität aller Fünfminutenabschnitte über den gesamten Versuchsablauf. Fehlerbalken geben das 95%-Konfidenzintervall wieder. Pausen wurden vor den Minuten 60, 120 und 180 gemacht und sind durch graue Balken markiert.

Auch hier werden zunächst in einem Überblick alle Fünfminutenabschnitte des gesamten Versuches dargestellt (vgl. Abb. 3.10). Dabei wird die Herzratenvariabilität immer als prozentuale Veränderung in Bezug auf die individuelle Baseline wiedergegeben. So werden die großen allgemeinen individuellen Unterschiede von den Variationen im Experiment getrennt. Auch hier zeigt sich wie schon bei der Herzrate ein ähnlicher Verlauf der drei Sessions, wobei die zweite HMD-Messung und der Tablet-

PC auf vergleichbarem Niveau sind, während die erste HMD-Messung ein deutlich niedrigeres Niveau hat und somit höhere mentale Beanspruchung signalisiert. Es fällt auf, dass fast alle Messzeitpunkte auf einem höheren Niveau als die Baseline liegen und somit den Versuchszeitraum selbst als weniger beanspruchend charakterisieren als die Ruhephase zu Beginn des Versuches. Dieser ungewöhnliche Fakt spiegelt wahrscheinlich aber eher eine hohe mentale Beanspruchung während der Ruhephase wieder, als eine niedrige Beanspruchung während des Versuches. Ein Grund hierfür könnte sein, dass die Probanden in der Ruhephase besorgt, aufgeregt oder nervös waren bezüglich des anschließenden langandauernden Experimentes. Dabei wurde bei der Versuchsdurchführung darauf geachtet, dass die Erhebung der Ruhephase noch vor der Vorstellung des HMDs und der Arbeitsinhalte des Experimentes stattfand. Es mag in diesem Zusammenhang empfehlenswert sein, die Ruhephasenmessung an einem Vortag zu machen, was aber organisatorisch einige Probleme mit sich bringt. Eine Ruhephasenmessung nach dem Experiment wäre ungewöhnlich und es bestände die Gefahr der Übertragung von Beanspruchung aus dem vorhergehenden Experiment in die Ruhephase. Die Reliabilität der Daten für die anstehenden Vergleiche ist mit der niedrigen Herzratenvariabilität in der Ruhephase aber nicht gefährdet.

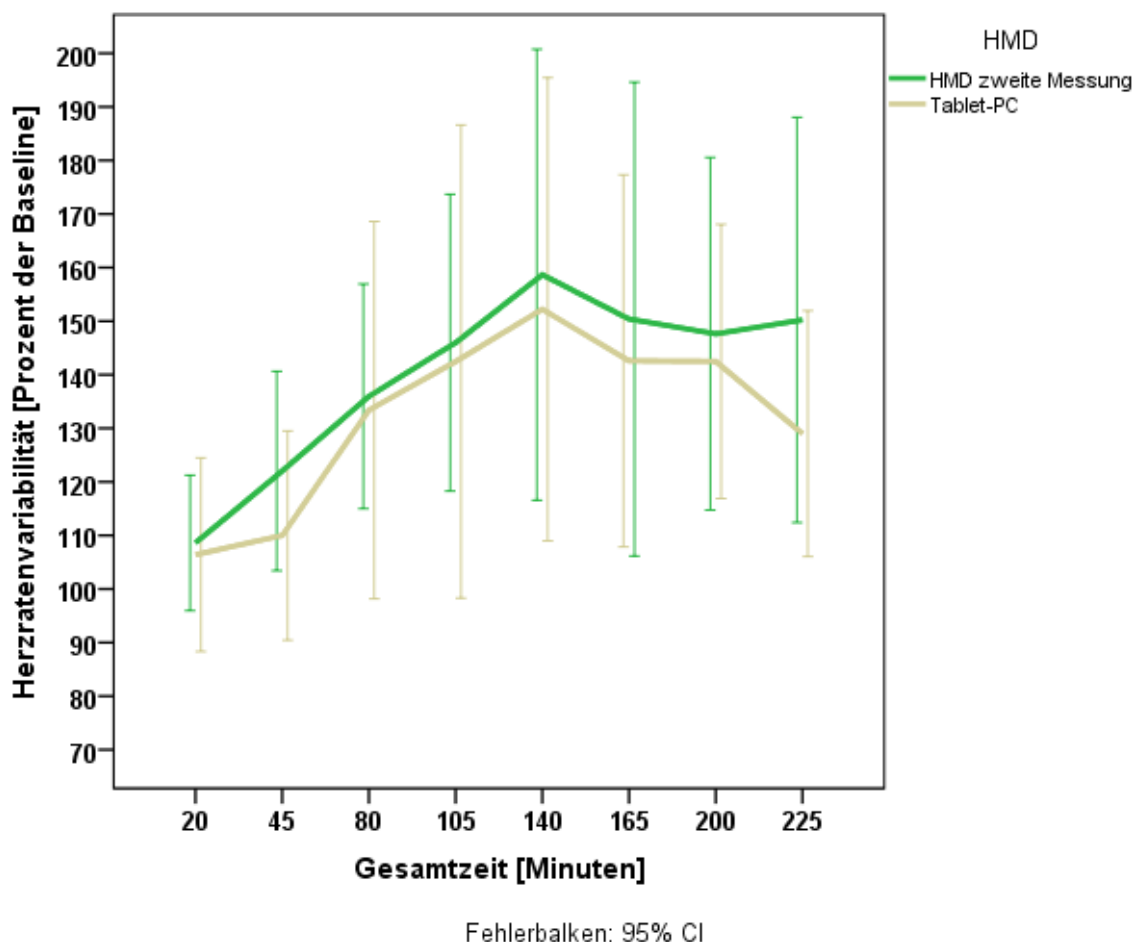


Abb. 3.11 Herzratenvariabilität der zweiten HMD-Messung und des Tablet-PCs zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Bei Betrachtung des Verlaufes fällt weiterhin ein Anstieg über die Zeit auf. Diese abnehmende mentale Anspannung könnte als Hinweis auf zunehmende Gewöhnung an die Situation interpretiert werden. Die Pausenstrukturen zeigen sich im Verlauf ebenfalls gut. Jeweils zu den Minuten 60, 120 und 180 direkt nach den Pausen kommt es zu einem Anstieg der Herzratenvariabilität und somit zu weniger mentaler Beanspruchung als vor den Pausen.

Für die Überprüfung der Hypothese H5b, die durch einen Vergleich zwischen den Bedingungen HMD2 und Tab den Einfluss des Displaytypus nach Gewöhnung untersucht, wurden wieder die acht Messzeitpunkte herangezogen, die bereits bei Auswertung der Herzrate herangezogen wurden: Die Fünfminutenabschnitte ab den Minuten 20, 45, 80, 105, 140, 165, 200 und 225 (vgl. Abb. 3.11). Diese acht Abschnitte sind jeweils kurz vor den in 3.4.2 ausgewerteten RSME-Abfragen, so dass kein Einfluss der Sprechaktivität oder kognitiver, subjektiver Bewertung der Situation durch die parallele RSME-Abfrage anzunehmen ist. Auch hier zeigt sich kein signifikanter Haupteffekt des Displaytypus [$F(1,34) = 0.249$; $p = .621$]. Aus physiologischer Sicht ist also das Arbeiten mit HMD und Tablet-PC nach einer eventuellen Gewöhnung während des ersten HMD-Durchganges gleich mental beanspruchend. Signifikant wird einzig die Variation über die Zeit [$F(7,28) = 4.519$; $p = .002$], die in beiden Bedingungen einen Anstieg und somit ein Abfallen der mentalen Beanspruchung verzeichnet, was wahrscheinlich einer Gewöhnung an die Situation zuzuschreiben ist. Wechselwirkungen zeigen sich bei dem Vergleich keine.

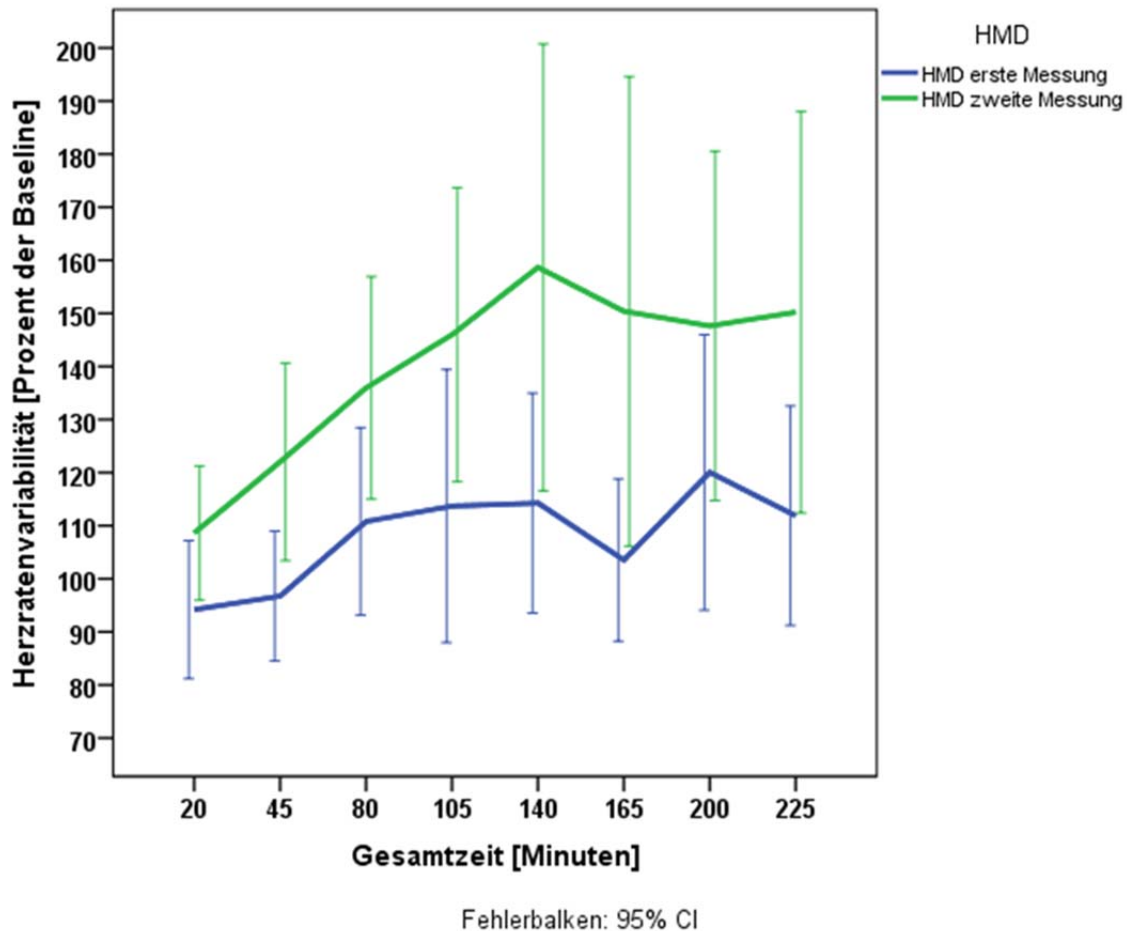


Abb. 3.12 Herzratenvariabilität der ersten und zweiten HMD-Messung zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Hypothese 5d überprüft die Gewöhnung an die Displaytechnologie, wobei die erste und die zweite Versuchssession mit dem HMD verglichen werden (vgl. Abb. 3.12). Hierbei zeigt sich ein signifikanter Haupteffekt der Versuchssession [$F(1,33) = 4.858$; $p = .035$]: Wie auch in der Grafik zu sehen, findet die erste HMD-Messung auf deutlich geringerem Niveau statt, was auf eine höhere mentale Beanspruchung schließen lässt. Die Hypothese 5d kann also bestätigt werden. Weiterhin wird auch wieder die gleichförmige Variation über die Zeit signifikant [$F(7,27) = 5.817$; $p < .001$]. Statistische Tendenz zeigt eine Wechselwirkung Zeit * Alter [$F(7,27) = 2.273$; $p = .059$], die wiedergibt, dass Ältere auf einem niedrigeren Niveau starten, der Anstieg über die Zeit dann aber bei Älteren größer ist und somit weniger mentale Beanspruchung gegen Ende der Session vorliegt. Eine Triplet Wechselwirkung Display * Alter * Technikaffinität wird signifikant [$F(1,33) = 4.735$; $p = .037$] und ist so zu interpretieren, dass der zweite HMD-Durchgang vor allem von den nicht technikaffinen älteren Probanden als mental weniger beanspruchend erlebt wird. Ein Grund hierfür könnte Desinteresse sein.

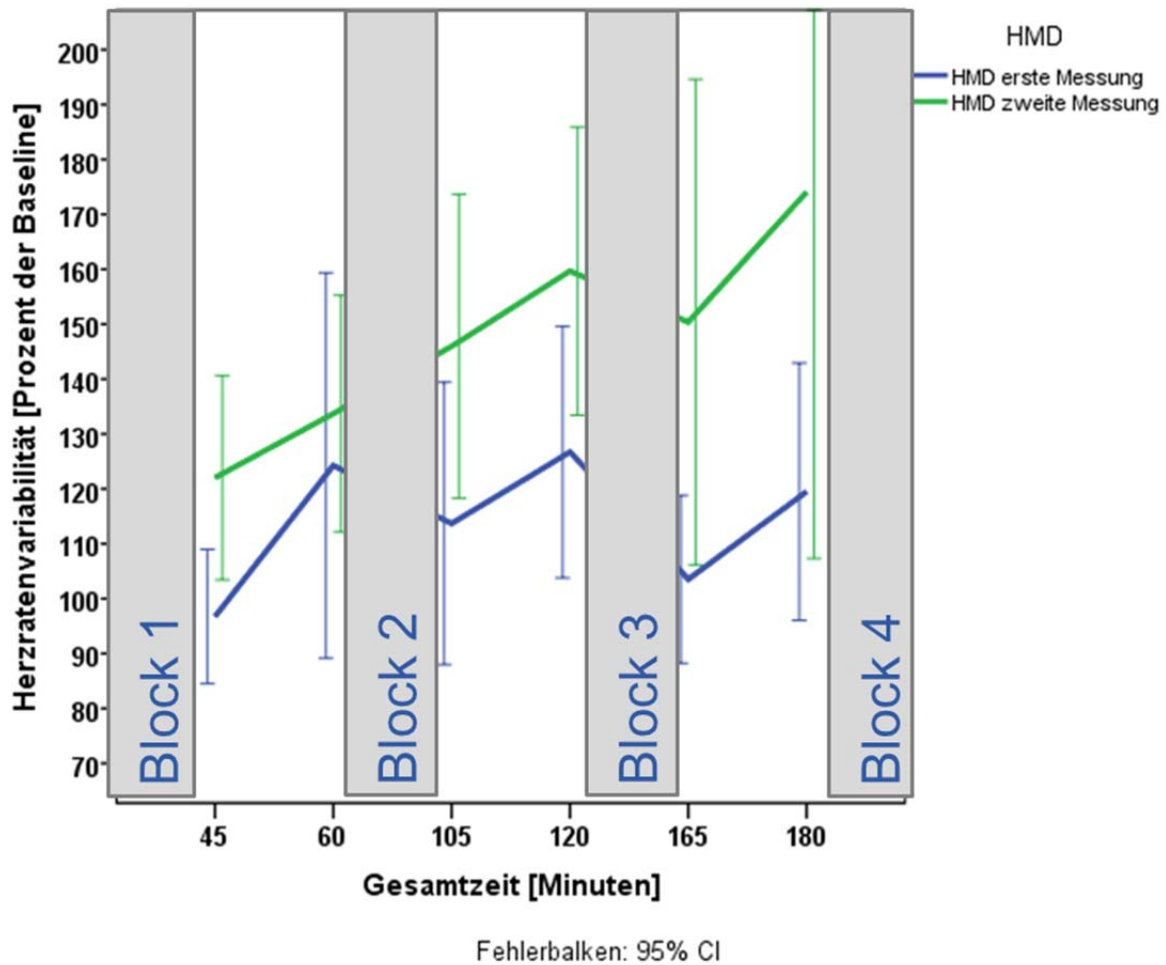


Abb. 3.13 Herzratenvariabilität jeweils kurz vor und nach den drei Pausen. Fehlerbalken geben das 95%-Konfidenzintervall wieder.

Eine Analyse der Wirkung von Pausen auf die Herzratenvariabilität wurde analog zu 36 vorgenommen, indem jeweils der Abschnitt kurz vor einer Pause mit dem kurz danach zu Beginn des nächsten Blockes verglichen wurde, wie Abb. 3.13 illustriert. Auch hier zeigte sich eine signifikante Beanspruchungsreduktion durch die Pausen [$F(1,34) = 5.988$; $p = .020$], die diesmal mit der Zunahme der Werte nach den Pausen einhergeht. Eine unterschiedliche Wirkung der Pausen ist auch hier nicht zu verzeichnen, obwohl die erste und dritte Pause kürzer waren und das HMD nicht abgenommen wurde. Signifikante Wechselwirkungen, die z. B. wiedergeben könnten, ob die Pausen im ersten HMD-Durchgang, der ja als anstrengender erlebt wurde, einen größeren Effekt haben, liegen nicht vor. Es ist also nur von einer allgemeinen Entspannung in den Pausen auszugehen, nicht aber von spezifischen Effekten.

3.6 Vergleich zwischen subjektiver und objektiver Beanspruchung

Subjektive und objektive Beanspruchungsparameter unterscheiden sich in einem wichtigen Punkt: Während die subjektiven Beanspruchungsparameter eine höhere Beanspruchung beim HMD und keine Gewöhnung ausmachen, zeigen die objektiven Beanspruchungsparameter keine höhere Beanspruchung beim HMD-Einsatz dafür aber eine Gewöhnung an den Umgang damit.

Zwar zeigen beide subjektiven Beanspruchungsparameter (NASA-TLX und RSME) eine deutlich geringere Beanspruchung beim zweiten HMD-Durchgang, diese erreicht aber nur statistische Tendenz. Die Richtung einer Gewöhnung wird aber eingehalten.

Bei der Frage nach dem Displayeffekt sieht das anders aus: Hier zeigen die subjektiven Beanspruchungsparameter einen deutlich signifikanten Effekt, während die objektiven Beanspruchungsparameter keinen Effekt und auch keine Tendenz in dieser Richtung ausmachen.

Diese Unterschiede können auf der unterschiedlichen Herangehensweise beider Verfahren basieren: Während die objektiven Beanspruchungsparameter direkt die Beanspruchung aus körperlichen Reaktionen ableiten, unterliegen die subjektiven Beanspruchungsparameter moderierenden Variablen der Introspektion. Denn statt der direkten Reaktion auf eine externe Belastung messen sie die bewusste Bewertung der Belastung, in die noch Dinge wie individuelle Antworttendenzen, aber auch die derzeitige Situation – abseits der Belastungssituation – einfließen. Das Tragen eines schweren, ungewohnten Gerätes und die überwachte Versuchssituation mögen so bei den subjektiven Beanspruchungsparametern zu einer schlechteren Bewertung geführt haben, als dies alleine von der Beanspruchungssituation ausgegangen ist. Trotzdem kann diese subjektive Einschätzung nicht als irrelevant gelten, da sie das Empfinden der Person wiedergibt und sich dies auf z. B. die Arbeitsmotivation niederschlagen kann.

3.7 Visuelle Ermüdung

Bei dem Fragebogen zur visuellen Ermüdung wurde das Signifikanzniveau auf 0,3% angehoben, wie in 2.4.2.7 ausführlich diskutiert. Es werden in Ausnahmefällen aber auch Ergebnisse erwähnt, die ein höheres Alpha-Risiko darstellen.

Allgemein lässt sich feststellen, dass es deutliche Veränderungen fast aller Itemscores gibt, die Mittelwerte aber auf einem relativ geringen Niveau von 0,5 bis 3,5 auf der zehnstufigen Skala liegen. Die Itemscores nehmen dabei über die Zeit zu und haben eine höhere Ausprägung beim HMD. In einigen Fällen kommt es auch zu einer stärkeren Zunahme der Itemscores über die Zeit beim HMD.

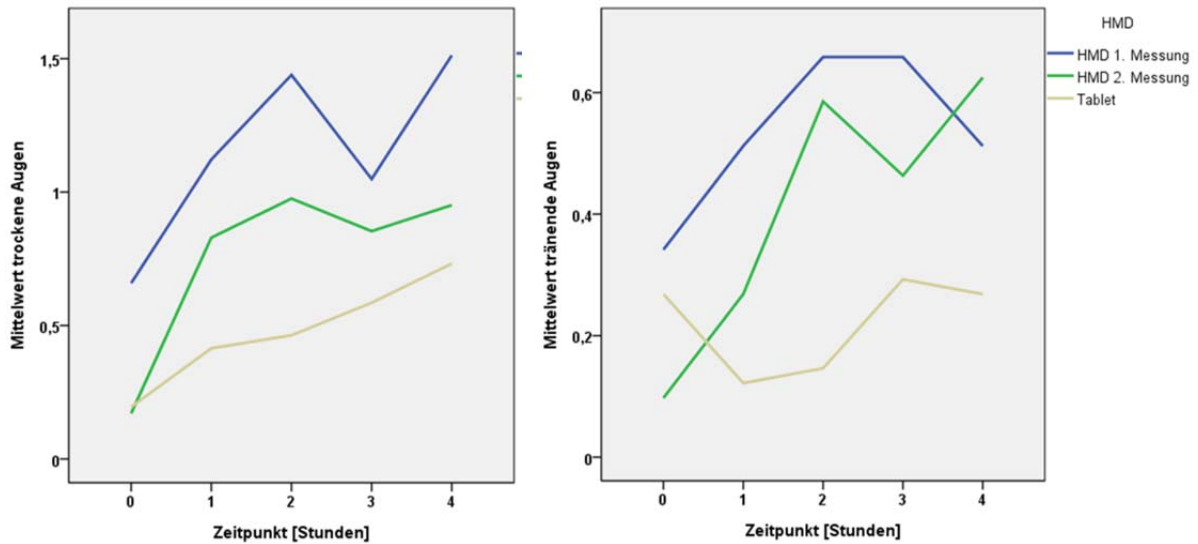


Abb. 3.14 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „trockene Augen“, rechts: „trärende Augen“.

Bei den „trockenen Augen“ gibt es einen signifikanten Effekt der Displaytechnologie [$F(1,37) = 11.440$; $p = .002$], wie sich bei einem Vergleich zwischen zweiter Messung mit dem HMD und Tablet-PC zeigt. Dabei liegt das Mittelwertniveau aber unter 1,5 auf der zehnstufigen Skala. Der Effekt ist also vorhanden, aber das Ausmaß nicht so intensiv. Ein Gewöhnungseffekt bei dem Vergleich zwischen erster und zweiter Messung mit dem HMD zeigt auf dem korrigierten Alphaniveau nur statistische Tendenz [$F(1,37) = 9.280$; $p = .004$], geht aber in die erwartete Richtung von geringeren Werten bei der zweiten Erhebung.

Bei den „trärenden Augen“ zeigen sich keine signifikanten Unterschiede, was wohl auch dem allgemein niedrigen Score der Werte zuzuordnen ist.

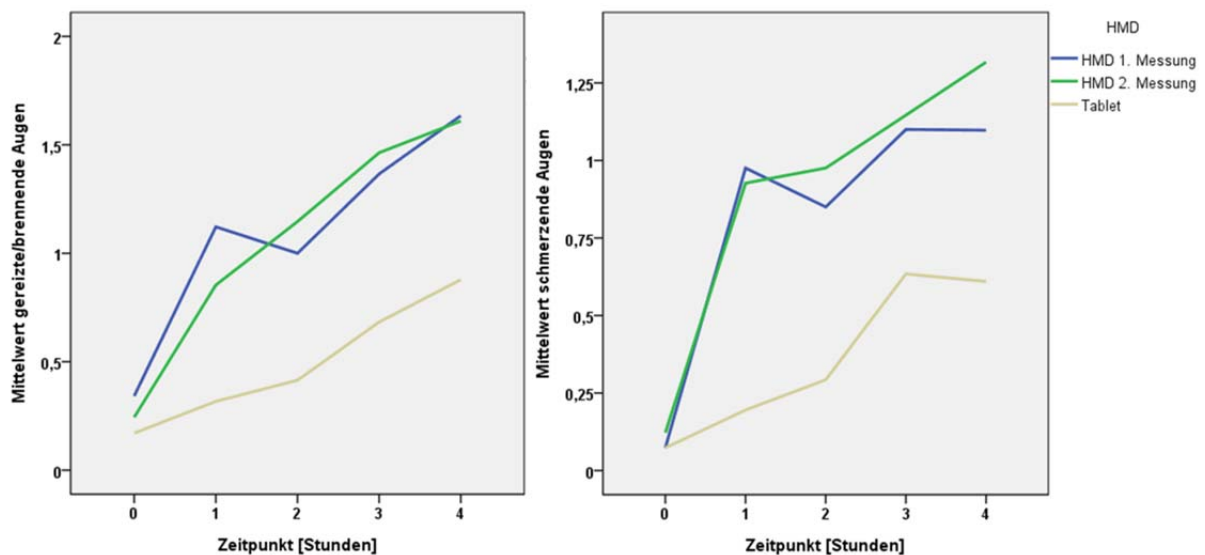


Abb. 3.15 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „gereizte/brennende Augen“, rechts: „schmerzende Augen“.

Bei den „gereizten/brennenden Augen“ zeigen sich erneut ein Haupteffekt der Displaytechnologie [$F(1,37) = 19.291$; $p < .001$] und ein Haupteffekt des Zeitpunktes [$F(1,37) = 5.021$; $p = .003$]. Das HMD führt also zu stärker gereizten Augen und über die Zeit nehmen die brennenden Augen zu, wobei dieser Anstieg bei beiden Displaytechnologien als gleich zu werten ist, da keine Wechselwirkungen zwischen Displaytypus und Zeitpunkt vorliegen [$F(4,34) = 2.959$; $p = .034$]. Zu bedenken ist aber, dass die Wechselwirkung bei normalem Alphaniveau signifikant wäre und auch in die zu erwartende Richtung ginge, dass die Steigerung über die Zeit beim HMD stärker ist. Ein Gewöhnungseffekt an die HMD-Technologie lässt sich nicht nachweisen [$F(1,37) = .002$; $p = .964$].

Bei den „schmerzenden Augen“ lassen sich bei angepasstem Alphaniveau keine Effekte nachweisen. Die Unterschiede in der Displaytechnologie [$F(1,37) = 7.691$; $p = .009$] und im Zeitpunkt [$F(4,34) = 4.238$; $p = .007$] gehen aber wieder in die erwartete Richtung. Ein Gewöhnungseffekt an die HMD-Technologie lässt sich auch hier nicht nachweisen [$F(1,37) = .103$; $p = .750$].

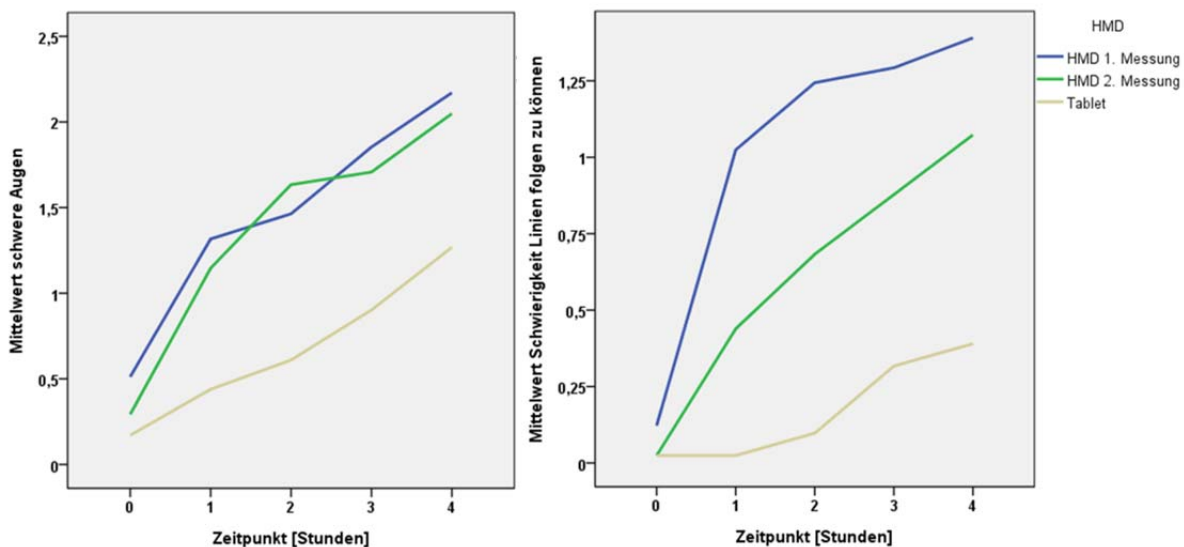


Abb. 3.16 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „schwere Augen“, rechts: „Schwierigkeit, Linien folgen zu können“.

Bei den „schweren Augen“ zeigen sich ein signifikanter Effekt der Displaytechnologie [$F(1,37) = 16.622$; $p < .001$] und ein signifikanter Effekt des Zeitpunktes [$F(4,34) = 6.807$; $p < .001$]. Auch hier führt das HMD zu höheren Werten, und es gibt eine allgemeine Steigerung über die Zeit. Ein Gewöhnungseffekt an die Displaytechnologie lässt sich nicht nachweisen [$F(1,37) = .136$; $p = .715$].

Bei dem Item „Schwierigkeit, Linien zu folgen“ verpassen die Haupteffekte Displaytyp [$F(1,37) = 9.460$; $p = .004$] und Zeitpunkt [$F(4,34) = 4.107$; $p = .008$] zwar die hochgesetzte Signifikanzgrenze, sind aber aufgrund ihres geringen Alphaniveaus trotzdem erwähnenswert. Eine Gewöhnung an die HMD-Technologie verpasst ebenfalls die angepasste Signifikanzgrenze [$F(1,37) = 6.917$; $p = .012$]. Die Tendenz geht aber in die erwartete Richtung mit niedrigeren Werten bei der zweiten Erhebung mit dem HMD.

Bei der „Schwierigkeit, scharf zu sehen“ gibt es einen Haupteffekt des Displaytypus [$F(1,37) = 28.662$; $p < .001$] und des Zeitpunktes [$F(4,33) = 8.510$; $p < .001$]. Wei-

terhin tritt hier auch eine signifikante Wechselwirkung zwischen Displaytypus und Zeitpunkt auf [$F(4,33) = 7.668$; $p < .001$]. Diese ist in der Richtung zu interpretieren, dass der Anstieg über die Zeit beim HMD steiler ausfällt. Der Gewöhnungseffekt an die Displaytechnologie verpasst die hochgesetzte Signifikanzgrenze knapp [$F(1,37) = 9.319$; $p = .004$], geht aber auch wieder in Richtung niedrigerer Werte beim zweiten HMD-Durchgang.

Bei dem Item „springender Text“ zeigen sich keine signifikanten Effekte, was wohl auch auf das allgemein niedrige Niveau der Mittelwerte zurückzuführen ist, die zwischen 0 und 0,8 liegen.

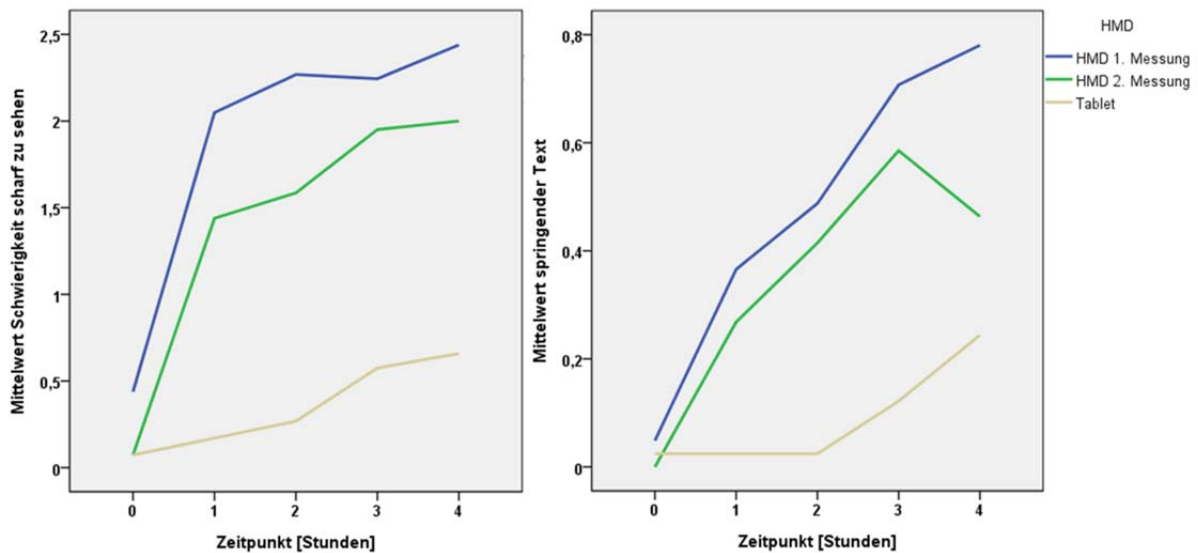


Abb. 3.17 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „Schwierigkeit, scharf zu sehen“, rechts: „springender Text“.

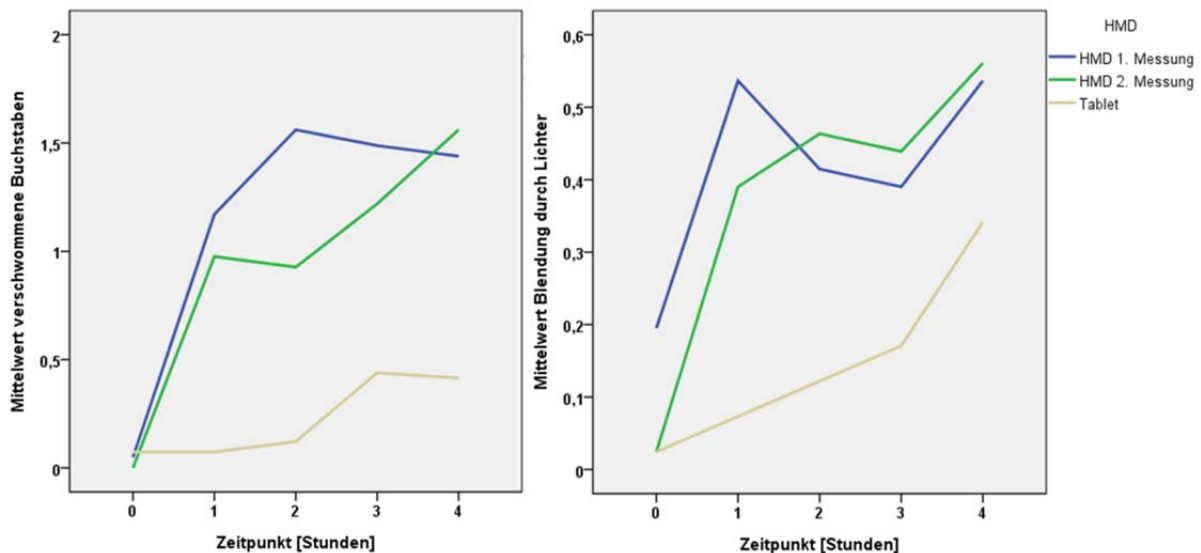


Abb. 3.18 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „verschwommene Buchstaben“, Rechts: „Blendung durch Lichter“.

Bei dem Item „verschwommene Buchstaben“ zeigt sich ein Haupteffekt des Displaytypus [$F(1,37) = 14.626$; $p < .001$] und des Zeitpunktes [$F(4,34) = 5.627$; $p = .001$]

sowie eine Wechselwirkung zwischen Displaytypus und Zeit [$F(4,34) = 4.673$; $p = .004$]. Die Werte sind beim HMD also allgemein höher, sie steigen bei beiden Displaytypen über die Zeit, der Anstieg ist beim HMD allerdings steiler. Ein Gewöhnungseffekt an die Displaytechnologie lässt sich nicht nachweisen.

Beim Item „Blendung durch Lichter“ zeigen sich keine signifikanten Effekte, was wiederum dem allgemein niedrigen Niveau der Mittelwerte zwischen 0 und 0,6 zuzuschreiben ist.

Auch beim Item „verschwommene Sicht“ zeigt sich ein Haupteffekt des Displaytypus [$F(1,37) = 22.944$; $p < .001$] und des Zeitpunktes [$F(4,34) = 7.996$; $p < .001$] sowie eine Wechselwirkung zwischen Displaytypus und Zeit [$F(4,34) = 5.912$; $p = .001$]. Die Werte sind also wieder beim HMD allgemein höher, sie steigen bei beiden Displaytypen über die Zeit, der Anstieg ist beim HMD allerdings steiler. Ein Gewöhnungseffekt an die Displaytechnologie lässt sich auch hier nicht nachweisen.

Bei dem Item „Doppelsehen“ zeigen sich keinerlei signifikante Effekte, und erneut ist das Niveau der Mittelwerte sehr niedrig und der Wert „1“ wird nie überschritten.

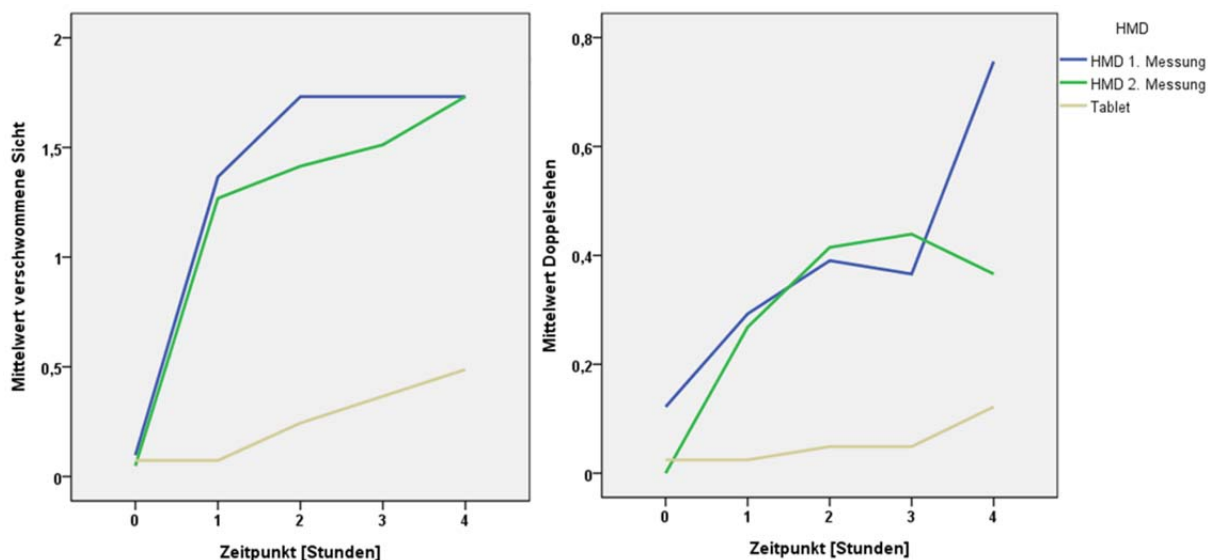


Abb. 3.19 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „verschwommene Sicht“, rechts: „Doppelsehen“.

Beim Item „Kopfschmerzen“ zeigt sich das bekannte Muster eines Haupteffektes des Displaytypus [$F(1,37) = 21.435$; $p < .001$] und des Zeitpunktes [$F(4,34) = 10.489$; $p < .001$] sowie einer Wechselwirkung zwischen Displaytypus und Zeit [$F(4,34) = 5.546$; $p = .002$]. Das HMD verursacht also allgemein mehr Kopfschmerzen, die Kopfschmerzen steigen bei beiden Displaytypen über die Zeit, der Anstieg beim HMD ist aber wesentlich größer. Ein Gewöhnungseffekt an die Displaytechnologie lässt sich auch bei den Kopfschmerzen nicht nachweisen [$F(1,37) = 1.914$; $p = .175$], obwohl sämtliche Mittelwerte der zweiten Erhebung mit dem HMD unter denen der ersten Erhebung bleiben.

Und auch beim Item „Nackenschmerzen“ zeigt sich dieses Muster. Es liegt ein Haupteffekt des Displaytypus [$F(1,37) = 22.000$; $p < .001$] und des Zeitpunktes [$F(4,34) = 11.900$; $p < .001$] sowie einer Wechselwirkung zwischen Displaytypus und Zeit [$F(4,34) = 5.341$; $p = .002$] vor. Das HMD verursacht also allgemein mehr

Nackenschmerzen, die Nackenschmerzen steigen bei beiden Displaytypen über den Untersuchungszeitraum und dieser Anstieg ist beim HMD wesentlich größer. Ein Gewöhnungseffekt an die Displaytechnologie findet ebenfalls nicht statt und ist im Falle der Nackenschmerzen auch nicht im grafischen Verlauf zu erkennen.

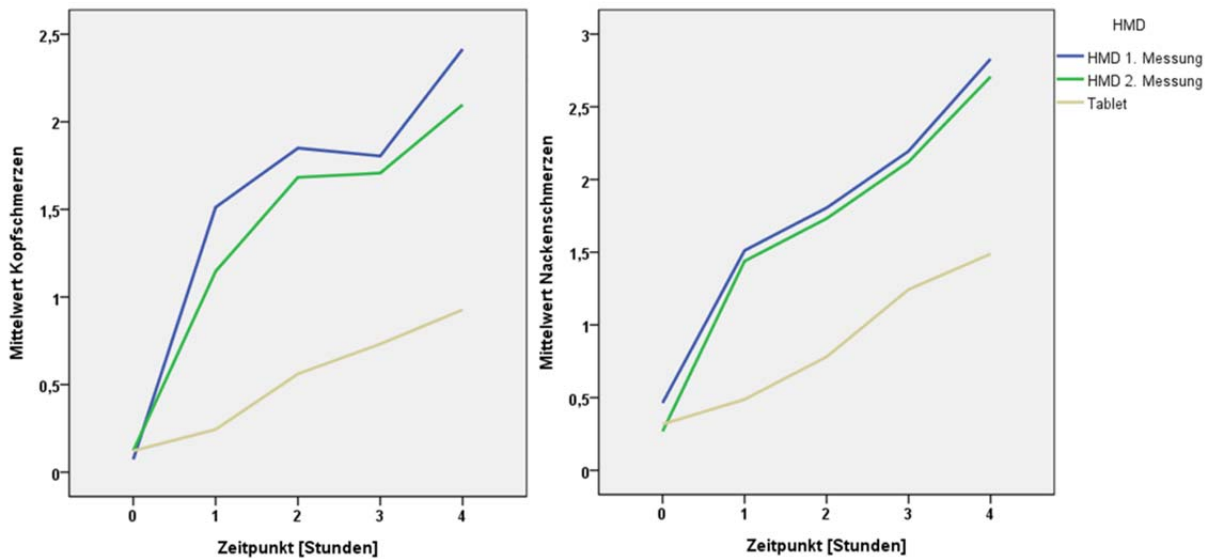


Abb. 3.20 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „Kopfschmerzen“, Rechts: „Nackenschmerzen“.

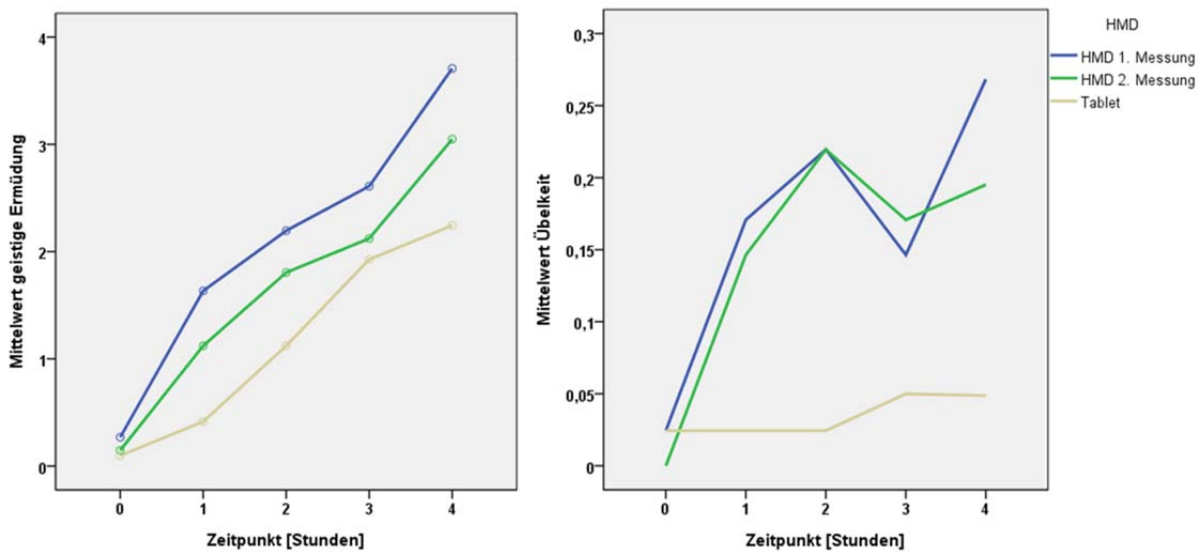


Abb. 3.21 VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „geistige Ermüdung“, Rechts: „Übelkeit“.

Bei dem Item „geistige Ermüdung“ verpasst der Haupteffekt des Displaytypus die hochgesetzte Signifikanzgrenze [$F(1,37) = 7.956$; $p = .008$], die Tendenz geht aber auch eindeutig in Richtung einer höheren geistigen Ermüdung beim HMD. Die allgemeine Steigerung über die Zeit wird signifikant [$F(4,34) = 16.665$; $p < .001$], und es tritt eine Wechselwirkung Zeitpunkt mal Alter auf [$F(4,34) = 5.857$; $p = .001$] in der Richtung, dass die Steigerung über die Zeit bei Älteren stärker ist. Die Wechselwirkung zwischen Displaytypus und Zeit verpasst das hochgesetzte Signifikanzniveau [$F(4,34) = 3.688$; $p = .013$], zeigt aber Tendenz in Richtung eines stärkeren An-

stiegs der geistigen Ermüdung über die Zeit beim HMD. Ein Gewöhnungseffekt an die Displaytechnologie lässt sich bei der geistigen Ermüdung aufgrund des hochgesetzten Signifikanzniveaus nicht nachweisen [$F(1,37) = 4.175$; $p = .048$], obwohl auch hier sämtliche Mittelwerte der zweiten Erhebung mit dem HMD unter denen der ersten Erhebung bleiben und beide Linien annähernd parallel verlaufen.

Bei dem Item „Übelkeit“ zeigen sich keinerlei signifikante Effekte und die Mittelwerte sind wieder auf sehr niedrigem Niveau von unter 0,3.

3.8 Interview

Am Ende jeder Session wurde mit den Probanden ein Interview durchgeführt (siehe Anhänge). Hier wurden den Probanden mehrere Fragen zum Umgang mit der Konstruktionsaufgabe und der Monitoringaufgabe gestellt und auch zu den Eigenschaften der Technologie und wo sie noch Verbesserungspotential sehen. Die Antworten waren dabei offen und wurden vom Versuchsleiter notiert.

Zunächst wurde gefragt, ob die Probanden mit der Konstruktionsaufgabe zurecht gekommen sind, was die meisten bejahten. Sollten die Probanden Probleme mit der Konstruktionsaufgabe gehabt haben, so wurde ihnen die Papierversion der Aufbauanleitung gezeigt und gefragt, ob es damit besser gegangen wäre. Die Probanden, die mit der Konstruktionsaufgabe aber leichte Probleme gehabt hatten, führten dies nur in äußerst wenigen Fällen auf die Darstellungsform zurück, sondern auf prinzipielle Probleme bei der Interpretation technischer Zeichnungen.

Als Nächstes wurden die Probanden gefragt, ob sie beim Bearbeiten der Balkenaufgabe eine bestimmte Strategie verwendet hätten, was viele verneinten. Diejenigen, die eine Strategie hatten, haben sich entweder durch subvokales Sprechen die letzte Position der Balken gemerkt oder auch teilweise Legoteile als Markierung verwendet. Die Rückmeldung in der Balkenaufgabe wurde allgemein als hilfreich empfunden.

Auf die Frage, ob sie alles scharf sehen konnten, antworteten die Probanden je nach Versuchbedingung recht unterschiedlich (vgl. Abb. 3.22). Während bei den Tablet-Durchgängen dies meist bejaht wurde, war dies bei den HMD-Durchgängen oft nicht der Fall. Hierbei unterscheiden sich die Antworten in der Hinsicht, dass manche betonen, dass sie nur zeitweise alles scharf sehen konnten, weil z. B. der Kopfträger verrutschte und zwischenzeitlich wieder neu justiert werden musste und Personen, die nur Teile des Displays scharf erkennen konnten, wobei hier meist der obere oder untere Rand als unscharf angesehen wurden. Bei den Antworten finden sich auch Kombinationen aus beiden Antwortkategorien. Eine Veränderung der Antworttendenzen zwischen erster und zweiter Messung mit dem HMD als Effekt der Gewöhnung lässt sich nicht ausmachen.

Bei der Interpretation dieser Ergebnisse ist zu beachten, dass wie in Kapitel 2.5 erwähnt, am Beginn der Sessionen sich jeweils viel Zeit dafür genommen wurde, das HMD an die Probanden anzupassen und die Probanden auch darauf hingewiesen wurden, sich zu melden, sollte das HMD sich während der Session verstellen, was viele Probanden auch in Anspruch nahmen. Dass trotzdem die Schärfe der Darstellung kritisiert wird – wobei diese hauptsächlich bei dem verwendeten HMD auf der

Positionierung des Displays basiert – hängt größtenteils mit dem Kopfträger als solchem zusammen.

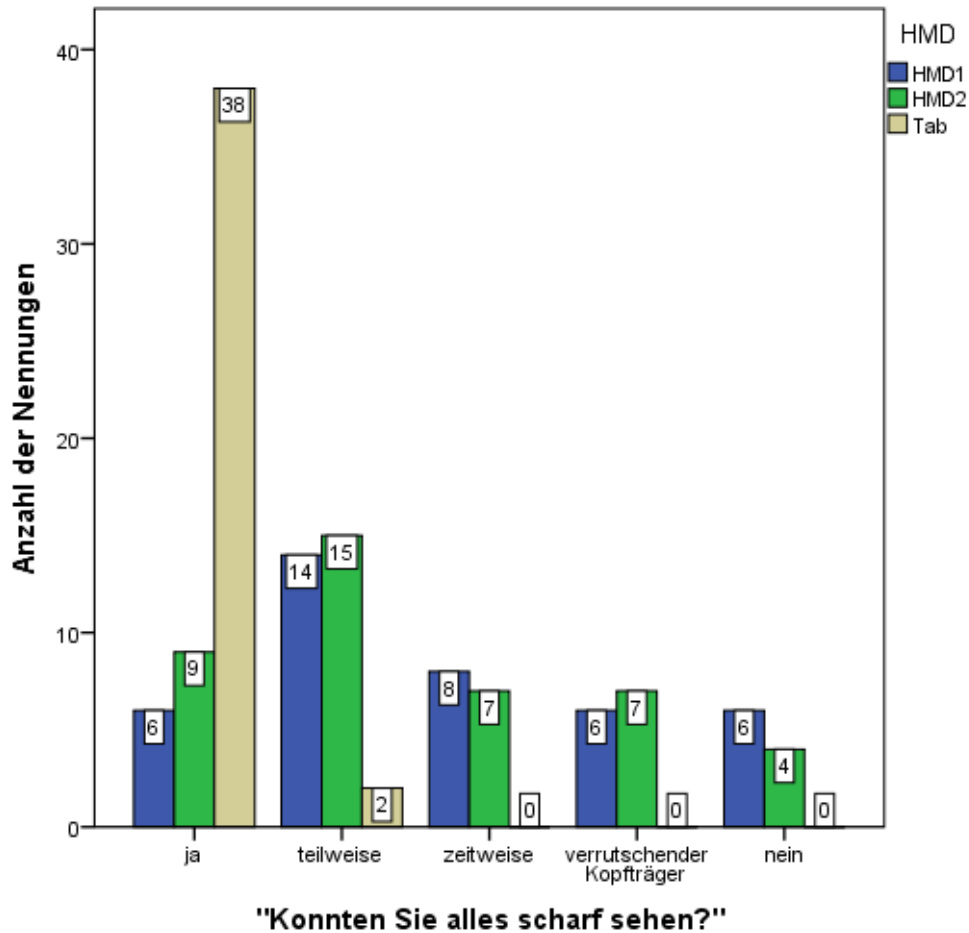


Abb. 3.22 Antworten auf die Frage „Konnten Sie alles scharf sehen?“

Wie Abbildung 3.23 zeigt, wurde der Kopfträger von den Probanden als unangenehm empfunden. Hierbei wurde am häufigsten der zu hohe Druck bemängelt, der von dem fest eingestellten Kopfträger ausging. Dieser Sachverhalt hängt eng zusammen mit dem zweithäufigst genannten Kritikpunkt, dem hohen Gewicht und dem dritthäufigst genannten Problem, der Tatsache, dass der Kopfträger oft verrutscht. Durch das hohe Gewicht und die Tendenz zu verrutschen musste der Kopfträger oft fester eingestellt werden als dies bei einem leichteren Exemplar nötig gewesen wäre. Hierbei ist dies wiederum darauf zurückzuführen, dass bei einem Verrutschen des Kopfträgers das Display nicht mehr scharf abzulesen war. So bedingen sich die einzelnen unergonomischen Faktoren gegenseitig. Es ist aber auch zu erkennen, dass ca. 25% der Probanden den Kopfträger als ganz o.k. empfanden und keine Kritik äußerten. Der Tragekomfort variiert also individuell sehr stark. Aus der Erfahrung der Versuchsleiter geht hierbei hervor, dass dies vor allem mit der Kopfform und der Haarfrisur der Probanden zusammenhing: Je fülliger die Haare, desto schlechter hielt der Kopfträger und musste oft fester eingestellt werden, damit er nicht verrutscht.

Gewöhnungseffekte an den Kopfträger lassen sich bei einem Vergleich zwischen erster und zweiter Messung mit dem HMD nicht feststellen.

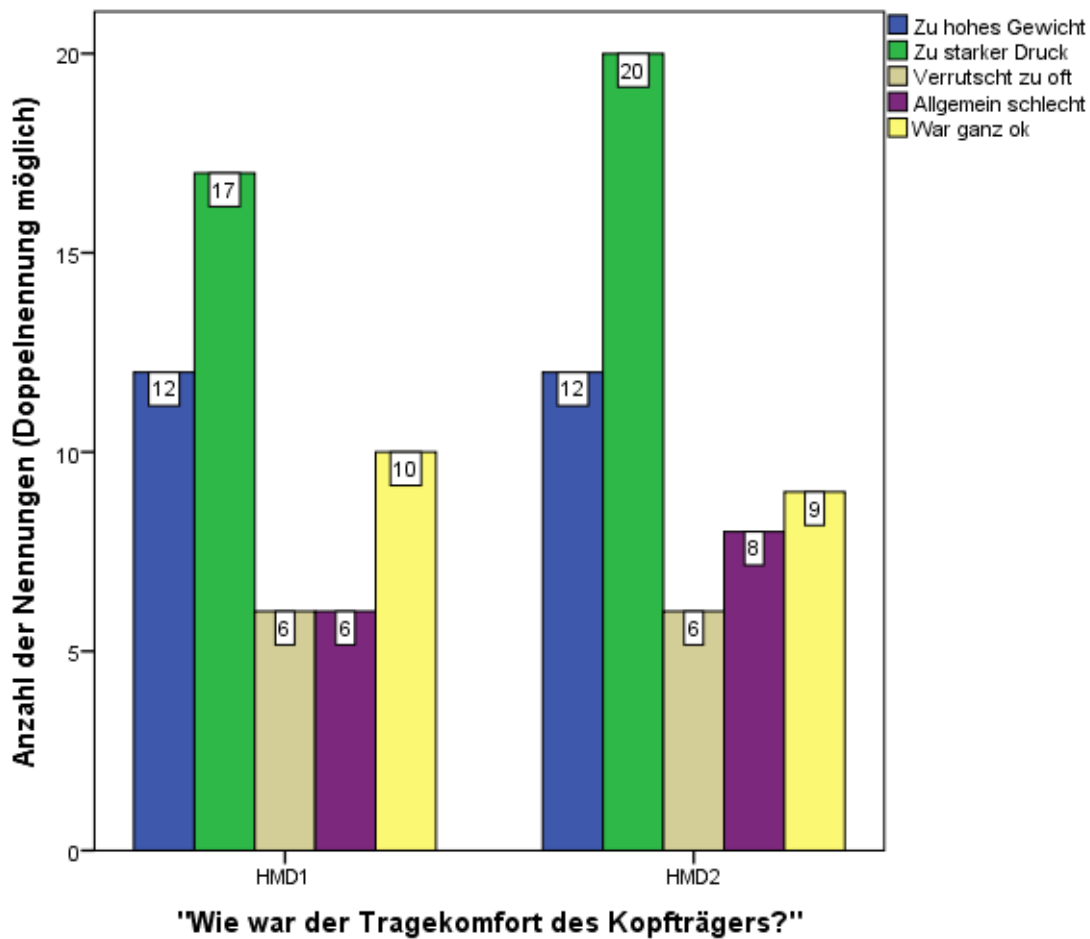


Abb. 3.23 Antworten auf die Frage „Wie war der Tragekomfort des Kopfträgers?“

Die monokulare Darstellung vor nur einem Auge (vgl. Abb. 3.24) charakterisierten die meisten Probanden als anstrengend, wobei sie überwiegend auch eine Gewöhnung an die monokulare Darstellung angaben. Oft wird vor allem die erste halbe Stunde als ungewohnt und anstrengend erlebt, während danach der Umgang mit dieser Darstellungsart gut funktioniert. Nur eine sehr kleine Anzahl fand die monokulare Darstellung allgemein und dauerhaft als schlecht. Hingegen hatten 20% der Probanden mit der monokularen Darstellung von Anfang an gar kein Problem. Auch hier zeigen sich, wie schon beim Tragekomfort des Kopfträgers, individuelle Unterschiede.

Eine Gewöhnung an die monokulare Darstellung von der ersten zur zweiten Erhebung mit dem HMD lässt sich aber nicht ausmachen: In beiden Sessions gab es zunächst eine Gewöhnung über die Zeit, die als anstrengend empfunden wurde.

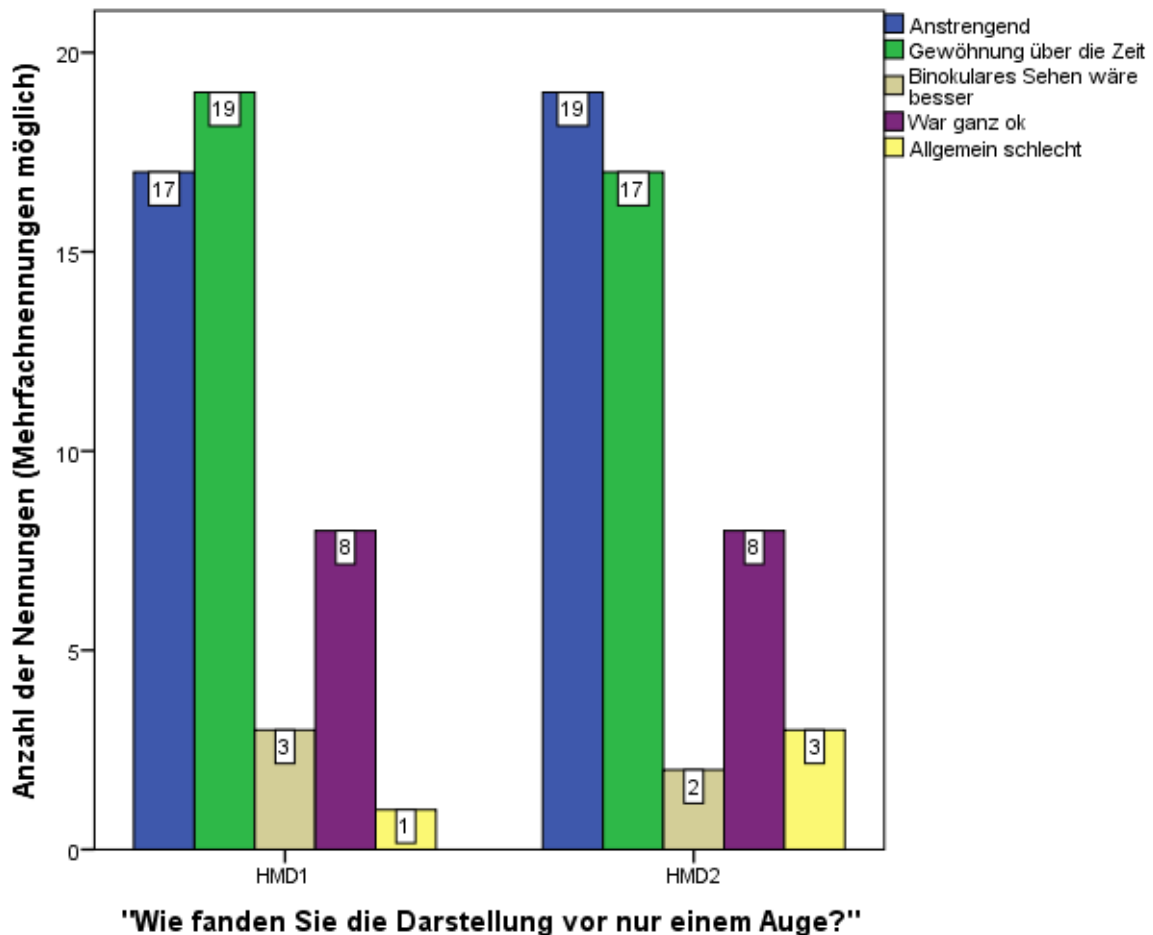


Abb. 3.24 Antworten auf die Frage „Wie fanden Sie die Darstellung vor nur einem Auge?“

In einem weiteren Punkt wurde gefragt, wo die Probanden noch Verbesserungsmöglichkeiten in der Technologie sehen (vgl. Abb. 3.25). Hierbei fällt auf, dass sich die meisten Verbesserungsvorschläge beim HMD auf den Kopfträger beziehen. Hier wünschten sich die Probanden eine Gewichtsreduzierung und verbesserte Einstellmöglichkeiten, die dann auch zu einem besseren Sitz des Kopfträgers, zu weniger Verrutschen und damit auch zu einem schärferen Display führen würden, was ebenfalls ein häufig genannter Wunsch beim HMD war.

Ein größeres Display und die Möglichkeit, in die technischen Konstruktionsanweisungen hineinzuzoomen war ein Wunsch, der unabhängig vom Displaytypus auch beim Tablet-PC von einigen Probanden geäußert wurde. Auffällig ist, dass beim Tablet-PC die Hälfte der Probanden keine Verbesserungswünsche hatte, während beim HMD dies nur bei zwei Probanden der Fall war.

Eine prinzipielle Veränderung zwischen erster und zweiter Messung mit dem HMD als Gewöhnungseffekt lässt sich nicht ausmachen.

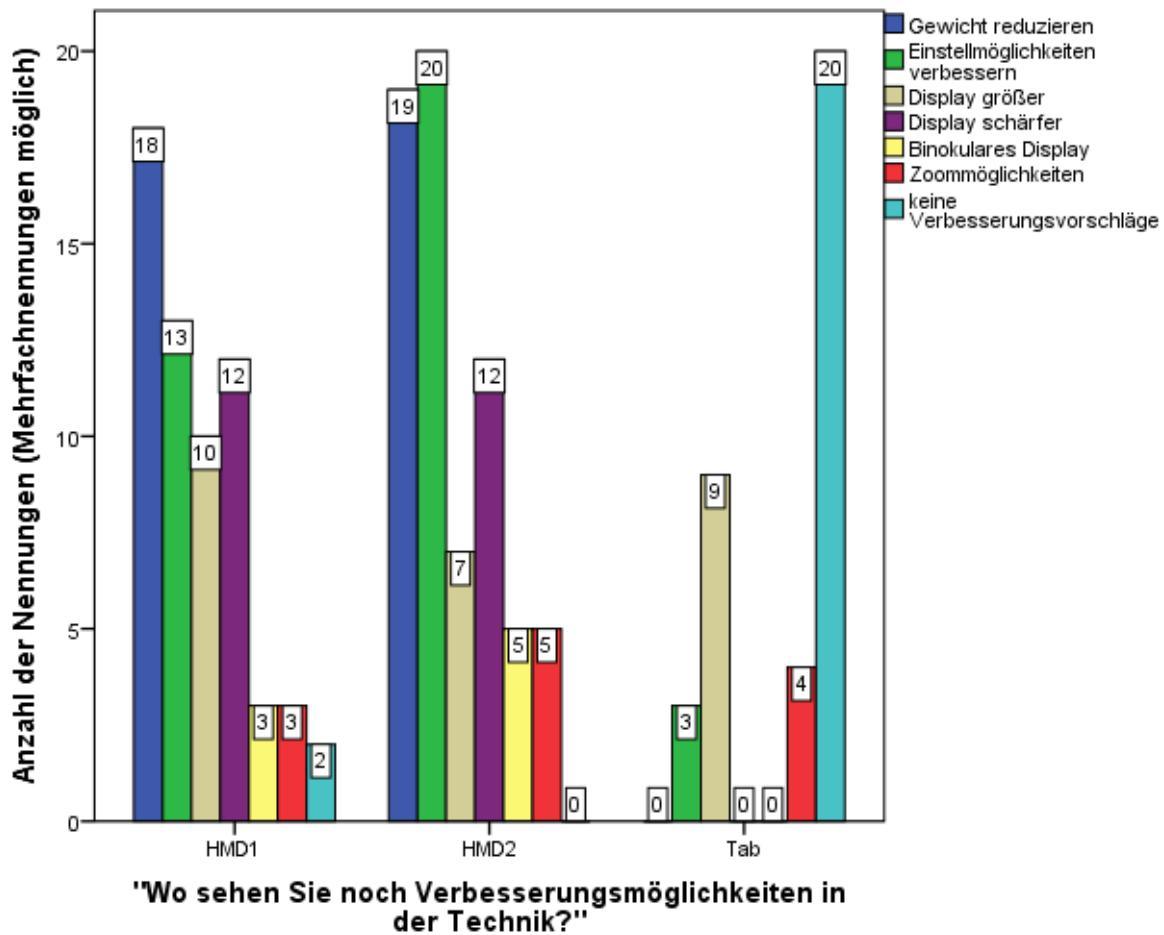


Abb. 3.25 Antworten auf die Frage „Wo sehen Sie noch Verbesserungsmöglichkeiten in der Technologie?“

Des Weiteren wurde nachgefragt, ob die Probanden mit fortlaufender Zeit eine Veränderung in ihrem Empfinden/Sehen bemerkt haben und ob sich dies gegebenenfalls nach den Pausen wieder verändert hätte (vgl. Abb. 3.26). Hierbei wird die steigende Anstrengung über die Zeit und eine zunehmende Unschärfe deutlich mehr beim HMD angegeben, kommt aber auch beim Tablet-PC vor. Ebenso wird ein Erholungseffekt der Pause beim HMD deutlich häufiger attestiert als beim Tablet-PC, was sich mit den Ergebnissen der subjektiven Beanspruchung (vgl. 3.4) deckt. Die Hälfte der Probanden macht beim Tablet-PC gar keine Variation ihrer Empfindung über die Zeit aus, während die beim HMD nur sehr wenige Probanden so sehen.

Eine Gewöhnung an die HMD-Technologie bei einem Vergleich zwischen erster und zweiter Erhebung lässt sich auch hier nicht ausmachen.

Nach der zweiten HMD-Session wurde auch immer gefragt, ob die Probanden einen Unterschied zur ersten Erhebung bemerkt haben (vgl. Abb. 3.27). Diese Frage hat einen gewissen Suggestivcharakter, so dass sich Probanden ermuntert sehen konnten, einen Unterschied zu berichten, den es vielleicht gar nicht gab. Die Antworten hierauf sind recht unterschiedlich. Ungefähr ein Viertel der Probanden sah keinen Unterschied. Die Hälfte der Probanden gab an, dass die zweite Erhebung angeneh-

mer gewesen sei, da sie sich an das HMD und die Tätigkeit gewöhnt hätten. Fünf Probanden gaben aber auch an, dass die zweite Erhebung unangenehmer war als die erste Erhebung, da nun der Reiz des Neuen gefehlt hätte und somit die Belastungen durch das HMD als höher empfunden würden.

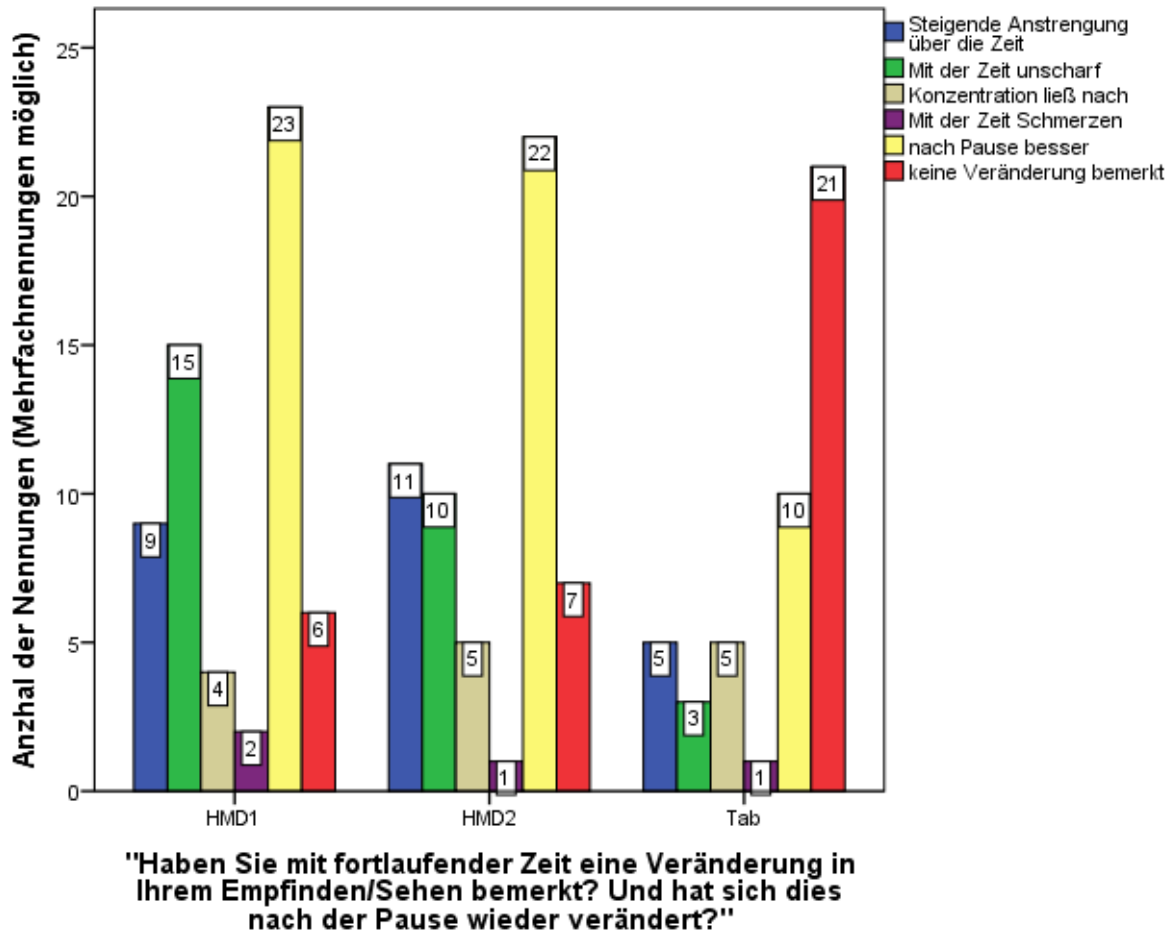


Abb. 3.26 Antworten auf die Frage „Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden/Sehen bemerkt? Und hat sich dies nach der Pause wieder verändert?“

Bei der Frage, welche der beiden Technologien sie für die Arbeitsaufgabe vorziehen würden, nannten die meisten Probanden den Tablet-PC. Nur ein Proband von 41 zieht das HMD dem Tablet-PC vor.

Abschließend wurde in der letzten Sitzung die Frage gestellt, ob die Probanden den Versuch noch einmal machen würden. Dies bejahten die meisten Probanden und konnten so als Probandenpool für mögliche zukünftige HMD-Studien gewonnen werden.

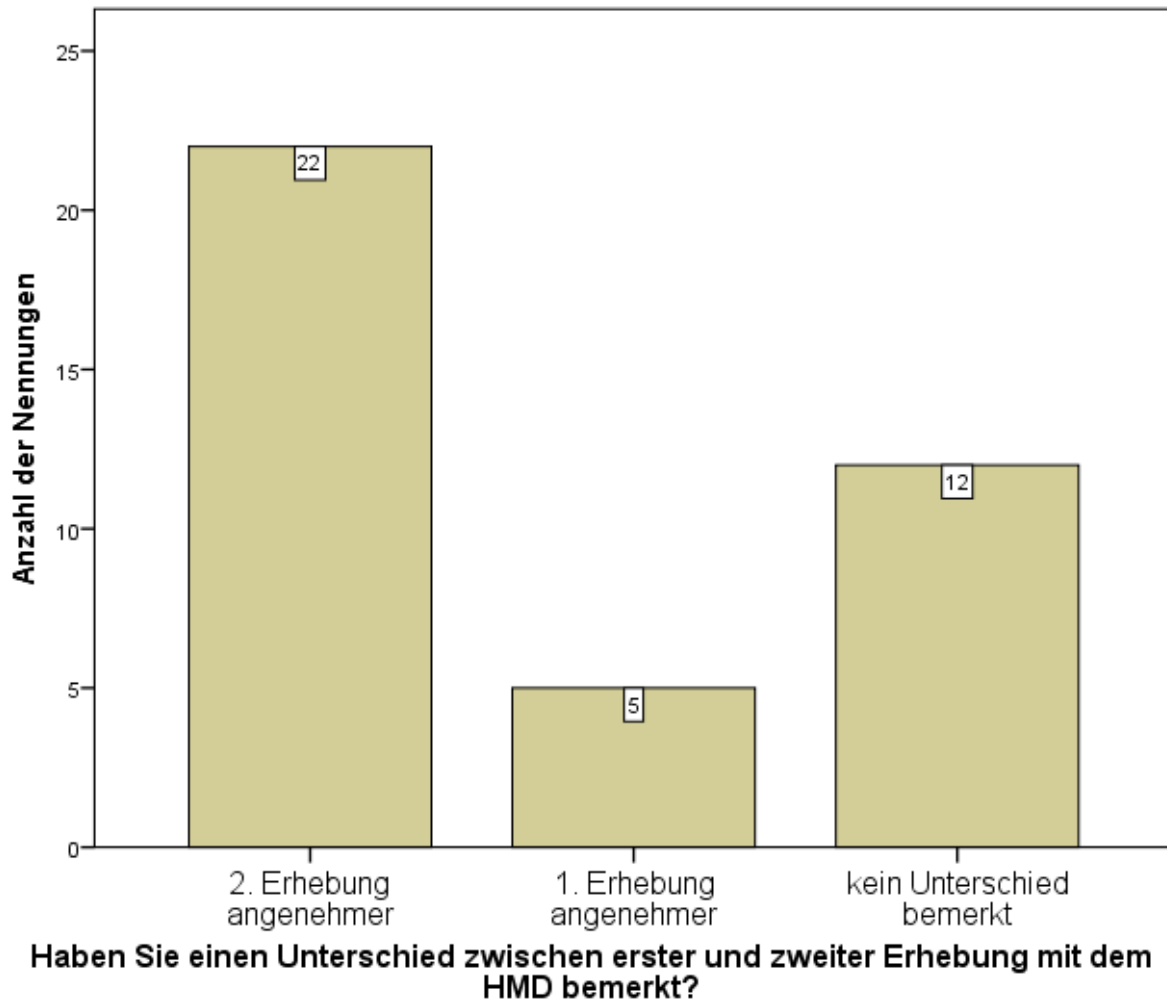


Abb. 3.27 Antworten auf die Frage „Haben Sie einen Unterschied zwischen erster und zweiter Erhebung mit dem HMD bemerkt?“

3.9 Zusammenfassung

In einer Laborstudie zu einem länger andauernden Einsatz von HMDs wurde die Beanspruchung der Probanden mit verschiedenen Verfahren und zu mehreren Zeitpunkten erfasst, um so einen Einblick in den Beanspruchungsverlauf zu bekommen. Als subjektive Verfahren wurden der NASA-TLX und der RSME eingesetzt, die beide einen zu erwartenden Anstieg der Beanspruchung über die Zeit wiedergeben. Ebenso zeigen beide Verfahren eine höhere subjektive Beanspruchung beim HMD im Vergleich zum Tablet-PC. Wechselwirkungen zwischen den Zeitpunkten und den verwendeten Displaytypen tauchen aber nicht auf, so dass der Anstieg der Beanspruchung über die Zeit bei HMD und Tablet-PC als gleichförmig anzusehen ist. Ältere werden durch das HMD teilweise mehr beansprucht. Die objektiven Beanspruchungsparameter Herzrate und Herzratenvariabilität zeigen im Gegensatz dazu keinen Unterschied in den Displaytypen, wohl aber eine Gewöhnung mit Abfallen der Beanspruchung bei der zweiten HMD-Erhebung. Die visuelle Ermüdung, erhoben durch einen Fragebogen, zeigt in vielen Items höhere Werte für das HMD. Dies trifft auch für Kopf- und Nackenschmerzen zu.

Bei der Leistung zeigt sich ein langsames Arbeiten in der bildbasierten Konstruktionsaufgabe mit dem HMD, das sich auch in der zweiten Erhebung bestätigt und somit nicht auf Gewöhnungseffekten basiert. In der parallel zu bearbeitenden Monitoring-Aufgabe, in der die Probanden auf bestimmte Zustandsänderungen einer Anzeige reagieren sollten, zeigt sich eine tendenziell bessere Performanz auf dem Tablet-PC im Vergleich zum HMD. Dies ist als ungewöhnlich zu betrachten, da die Anzeige bei einem HMD immer im Sichtfeld ist, während sie es beim Tablet-PC nicht zwangsläufig ist. Die in manchen Forschungsprojekten geführte Argumentationslinie, wichtige Anzeigen auf einem HMD zu zeigen, damit sie zuverlässiger erkannt werden, muss somit in Frage gestellt werden.

In Interviews wurde der mangelnde Tragekomfort des verwendeten industriellen HMDs beanstandet. Da der Kopfträger und andere Teile des Aufbaus (z. B. Gewicht, Kabel usw.) aber nur Teile der jeweiligen Hardwaregestaltung sind und nicht unmittelbarer Teil des Konzeptes HMD, stellt sich die Frage, wieviel der gefundenen und beschriebenen Effekte auf der Hardwareumsetzung und wie viel auf dem Konzept HMD basieren. So könnten z. B. die Kopfschmerzen beim HMD-Gebrauch vom Gewicht des Kopfträgers kommen oder aber von der monokularen visuellen Beanspruchung über mehrere Stunden. Auch die subjektiven Beanspruchungsparameter können von der unbequemen Hardware mit verursacht werden. Die Hardware von HMDs erfährt aber aktuell einen großen Wandel, da im Jahr 2014 z. B. mehrere große Konzerne HMDs auf den Konsumentenmarkt gebracht haben. Durch die Massenanfertigung von HMDs, die bisher nur ein Nischenprodukt waren, wurden die Displays besser, das Gewicht geringer und der Tragekomfort deutlich verbessert. Die Frage stellt sich, ob die gefundenen Effekte mit einem solchen HMD repliziert werden können. Deswegen wurde eine Replikation des Versuchssetups mit Google Glass durchgeführt, über das im nächsten Kapitel berichtet wird.

4 Replikation mit Google Glass

4.1 Ausgangslage und Zielsetzung der Studie

Die Produktklasse der Head-Mounted Displays (HMDs) befindet sich derzeit in einem großen Umbruch. Bis vor wenigen Jahren waren es Nischenprodukte, die nur für spezielle Anwendungen und Anwender gemacht wurden. Mit der damit einhergehenden kleinen Stückzahl von HMDs und den dadurch bedingten kleinen Budgets der Entwicklung wurden zwar immer Fortschritte hinsichtlich Displayqualität und Tragekomfort gemacht, diese waren aber im Vergleich zu der Entwicklung der letzten Jahre recht gering. Denn in den letzten Jahren haben verschiedene große Firmen das HMD als potentielltes Konsumentenprodukt entdeckt (z. B. Google Glass). Damit einher geht eine massive Entwicklung hinsichtlich höherer Displayqualität und verbessertem Tragekomfort, denn ein unbequemes HMD lässt sich Konsumenten zur freiwilligen Nutzung nur schwer verkaufen. Diese noch recht neue Produktklasse der Konsumenten-HMDs zeichnet sich nicht nur durch höheren Tragekomfort, sondern auch durch einen geringeren Preis aus, so dass sich HMDs im Moment an der Schwelle zum Massenmarkt befinden. Diese Konsumenten-HMDs unterscheiden sich aber (im Moment noch) in einigen Punkten von den industriellen HMDs: Die Akkuleistung ist meist sehr gering, so dass diese Geräte nicht für die dauerhafte Anzeige von Inhalten zur Arbeitsunterstützung gemacht sind. Der Akku von Google Glass hält bei einer Daueranzeige z. B. keine Stunde durch. Das Gerät ist eigentlich für sogenannte Mikrointeraktionen gedacht, bei denen die Anzeige nur wenige Sekunden an ist, eine erfragte Information anzeigt und danach wieder ausgeht. Bei permanenter Anzeige von Informationen leert sich nicht nur der Akku mit hoher Geschwindigkeit, sondern das Gerät wird am Brillenbügel auch warm, bzw. bei intensiver Nutzung sogar heiss. Auch die Robustheit von Industrie-HMDs ist wesentlich höher als die von Konsumenten-HMDs, ebenso die Toleranz hinsichtlich Umgebungstemperatur und Luftfeuchtigkeit. Nichts desto trotz wurden HMDs allgemein durch den Entwicklungsschub im Konsumentenbereich der letzten Jahre massiv verbessert und es ist anzunehmen, dass diese Verbesserungen in kurzer Zeit auch in die industriellen HMDs übergehen.

Vor diesem Hintergrund wurde die in den vorangegangenen Kapiteln beschriebene Studie mit Google Glass - als einem typischen Vertreter der neueren HMDs aus dem Konsumentenbereich – repliziert. Die Frage stellt sich, welche der gefundenen Effekte auf der Hardwareumsetzung und welche auf dem Konzept HMD an sich beruhen. So wiegt das in der Hauptstudie verwendete Mavus HMD 340 Gramm, während Google Glass nur 50 Gramm wiegt. Das Mavus HMD hat einen Kopfträger, während Google Glass nur ein etwas breiteres Brillengestell hat. Beim Mavus HMD muss zusätzlich eine schwere Weste getragen werden, die sowohl den Akku als auch die Kommunikationstechnologie enthält, während bei Google Glass all dies in einem Bügel der Brille untergebracht ist. Bei diesen massiven Unterschieden ist es kaum zu glauben, dass zwischen beiden Modellen nur fünf Jahre Entwicklungszeit liegen. Dies demonstriert zum einen den rasanten Fortschritt in der HMD-Entwicklung und zum anderen die finanzielle und technologische Stärke von Großunternehmen, die ein neues Massenprodukt auf den Markt bringen wollen.

Im Rahmen der Studie interessiert vor allem, ob sich die Effekte in der Leistung, Beanspruchung und visuellen Ermüdung zwischen Mavus HMD und Google Glass unterscheiden und somit auf die Hardware zurückzuführen sind, oder ob sie repliziert werden und somit unabhängig von der Hardwareumsetzung ein allgemeines Phänomen von HMDs sind.

4.2 Versuchsaufbau

Die Replikationsstudie fand in Zusammenarbeit mit der Gruppe „Embedded Sensing Systems“ (ESS) der Technischen Universität Darmstadt statt (www.ess.tu-darmstadt.de). Hierbei stellte ESS Google Glass zur Verfügung (das zu diesem Zeitpunkt am europäischen Markt noch nicht frei verfügbar war) und programmierte die Umsetzung der Versuchsaufgaben in der Programmiersprache Android und ein Programm zur Aufbereitung der Daten in eine SPSS-kompatible Matrize. Die Studie wurde in den Laboratorien der Bundesanstalt durchgeführt und auch die Probandenakquise lag in den Händen der BAuA. Hierzu wurden möglichst viele der Probanden aus der Originalstudie wieder eingeladen.

Bei der Replikation der Studie wurden die Versuchsaufgaben – die Konstruktionsaufgabe und parallele Monitoringaufgabe - übernommen. Die ursprüngliche Versuchsdauer von vier Stunden konnte aber aus organisatorischen Gründen nicht realisiert werden. Statt dessen dauerte der Versuch nur eine halbe Stunde und war in eine Studienreihe mit zwei anderen Experimenten integriert. Bei einem dieser Experimente handelte es sich um eine Studie zur Arbeitsunterstützung durch Google Glass im Biologielabor. Hier wurden Arbeitsanweisungen auf dem HMD gezeigt, die sequentiell abgearbeitet werden mussten, um die DNA aus Zwiebeln und Tomaten zu gewinnen. Diese zu Beginn der Reihe durchgeführte Studie dauerte circa eine Stunde und dabei gab es noch eine Variation von Sicherheitshinweisen, die der Untersuchung von persuasiven Technologien diente. Danach folgte eine Studie zur Fehlersuche an einem Wandschrank, die circa eine halbe Stunde dauerte. Anschließend wurde die Replikationsstudie für ebenfalls eine halbe Stunde durchgeführt, so dass eine Gesamttragezeit von Google Glass von zwei Stunden realisiert wurde. Genaueres zum Ablauf findet sich in Kapitel 4.2.4.

Im Gegensatz zur Originalstudie wurde hier ein between-subject Design gewählt. Die Probanden arbeiteten also entweder mit Google Glass oder einem Tablet-PC. Der Tablet-PC stammte hier ebenfalls aus dem Konsumentenbereich, so dass die gesamte Apparatur diesmal aus dem Konsumentenbereich stammte, während in der ursprünglichen Studie sowohl HMD als auch Tablet-PC aus dem Industriebereich stammten.

Während der gesamten Versuchsreihe wurden die Handbewegungen der dominanten Hand mit einer mobilen Lageerkennung aufgezeichnet. Diese Daten sind für den Studienpartner ESS wichtig, da probiert werden soll, aus bestimmten Bewegungsmustern auf bestimmte Arbeitsschritte zu schließen. Das dafür notwendige Gerät gleicht einer Armbanduhr und dürfte keinen hinderlichen Einfluss auf die Probanden gehabt haben.

4.2.1 Versuchsaufgaben

Während des Versuches waren wieder zwei Aufgaben parallel zu bearbeiten, die beide als gleich wichtig instruiert wurden ("Dual-task" Paradigma) und deren Inhalte auf dem HMD oder Tablet-PC dargestellt wurden.

4.2.1.1 Konstruktionsaufgabe

Die Probanden sollten wieder ein Modellfahrzeug von Lego-Technic aus einzelnen Bausteinen zusammenbauen. Die Konstruktionsfolien zeigen dabei eine schematische Zeichnung des Modells und der dazu zuzufügenden Bausteine an. Die originale Papieranleitung wurde dazu bereits für den Hauptversuch digitalisiert und bearbeitet. Für die Umsetzung auf Google Glass fand eine erneute Bearbeitung statt, um die Darstellungen dem neuen Displayformat von 640 x 360 Pixeln (statt 800 x 600 Pixel beim MAVUS-HMD) anzupassen.

Als Legomodell wurde diesmal ausschließlich der Baukasten 9392 verwendet, bei dem ein Quadbike aufgebaut wird (vgl. Anhang 1). Die Verwendung mehrerer Modelle war nicht notwendig, da die Fertigstellung dieses Modells ca. 45 Minuten dauert und nicht in der gegebenen Versuchszeit möglich war. Bei Fehlern in der Konstruktion - die sich oft erst in späteren Aufbaufolien zeigen, da hier dann neue Bausteine nicht mehr passen - waren die Probanden angehalten, den Fehler zu identifizieren und zurück zu bauen. Hierzu konnten sie zu jedem Zeitpunkt die Konstruktionsfolien vor- und zurück schalten. Sollten Probanden ihre eigenen Fehler nicht gesehen haben, wurden sie vom Versuchsleiter darauf aufmerksam gemacht.



Abb. 4.1 Darstellungsaufteilung der Arbeitsinhalte in der Replikationsstudie mit Google Glass. Links: Konstruktionsfolie für Lego-Technic ; rechts: Monitoringaufgabe mit drei Balken und Rückmeldung über letzte bestätigte Farbe und Position.

Die Darstellung der Arbeitsinhalte und das Format waren auf dem Tablet-PC und auf dem HMD exakt gleich. Hierzu wurden die Inhalte für den Tablet-PC auf dessen höhere Auflösung hochgerechnet. Bei Google Glass gab es wegen des kleinen Displays noch eine zusätzliche Zoomfunktion: Über Sprachbefehle konnten die Probanden in die Konstruktionsfolie hinein zoomen und hatten dann eine zweifache Vergrößerung. Dieser Vergrößerungsfaktor reicht aus, um auch kleinste Details komfortabel zu erkennen, wie sich in Voruntersuchungen erwiesen hat. Den Zoomausschnitt – also welchen Teil des Bildes man dann sieht – konnten die Probanden über die Kopfbewegung steuern. Hierzu wurde der interne Lagesensor von Google verwendet. Dabei mussten die Kopfbewegungen nicht groß sein und haben circa 15 Grad Bewegungswinkel nicht überschritten.

4.2.1.2 Monitoringaufgabe

In der parallel dargebotenen Monitoringaufgabe sollten die Probanden wieder auf drei vertikale Balken am äußeren Bildschirmrand achten und auf bestimmte Zustandsänderungen reagieren. Die Balken variierten kontinuierlich, aber sehr langsam in ihrer Länge und ab und zu in ihrer Farbe (blau/rot). Reagiert werden sollte auf einen Farbwechsel und auf einen Wechsel der Position des längsten Balkens, der durch die kontinuierliche Längenvariation der Balken entstehen kann. Im Durchschnitt war alle 106 Sekunden (SD = 48,639) ein Farbwechsel zu bestätigen und alle 94 Sekunden (SD = 61,968) ein Positionswechsel des längsten Balkens. Die Probanden bekamen diesmal während des gesamten Versuches eine Rückmeldung über ihre letzte Bestätigung in Form einer schriftlichen Mitteilung über den Balken (z. B. "red - right"). Diese Rückmeldung war in englischer Sprache, da auch die gesamten Interaktionen (vgl. Kapitel 4.2.1.3) in Englisch waren.

Für die Farbwechsel- und Längenwechselreaktion gab es diesmal keine zwei unterschiedlichen Reaktionen oder Tasten; auf beide wurde in gleicher Weise reagiert. Dies vereinfacht die Situation für die Probanden, führt aber bei der Datenauswertung dazu, dass bei einem zufälligen Überschneiden beider Reaktionsanforderungen nicht klar ist, auf was reagiert wurde. Während der kurzen Vorbereitungszeit auf diesen Versuch war es aber nicht möglich, eine Version mit zwei unterschiedlichen Reaktionsantworten zu realisieren.

Im Unterschied zur vorangegangenen Studie wurde diesmal auch eine Baseline in der Monitoringaufgabe erhoben, bei der die Probanden ausschließlich auf diese achten sollten und keine parallele Konstruktionsaufgabe stattfand (vgl. Kapitel 4.2.4).

4.2.1.3 Interaktion

Die Interaktionen fanden diesmal nicht mit einem umgebauten Nummernpad statt. Vielmehr waren sie bei beiden Displaytypen unterschiedlich und dem jeweiligen Gerät angepasst:

Bei dem Tablet-PC wurden alle Interaktionen mittels Berührungsgesten durchgeführt. Diese Art der Interaktion ist Standard bei Tablet-PCs. Mit einem Wischen nach links wurde die nächste Konstruktionsfolie und mit einem Wischen nach rechts die vorherige Konstruktionsfolie aufgerufen. Als Reaktion auf die Monitoringaufgaben wurde ein Doppeltippen an beliebiger Stelle des Displays eingesetzt.

Bei Google Glass wurden alle Interaktionen mit Sprachbefehlen durchgeführt. Diese waren in Englisch, da es sich bei dem Gerät um einen Prototypen handelt und noch dazu eine Verbindung mit dem Internet, die eine deutsche Spracherkennung verbes-

sert hätte, jedoch während des Versuches aus Datenschutzgründen unterlassen wurde. Die Anzahl der Sprachbefehle war aber sehr gering und so von den Probanden gut erlernbar. In Vortests wurden Sprachbefehle ausgewählt, die leicht verständlich sind und vom System auch bei einer deutschgefärbten Aussprache gut verstanden werden. Ausserdem wurden die Sprachbefehle zu Beginn des Experiments trainiert und die Probanden hatten sie immer in ausgedruckter Form vorliegen. Der Sprachbefehl für die nächste Konstruktionsfolie war „next slide“, der für eine Folie zurück war „previous“. Als Reaktion auf die Monitoringaufgaben war der Befehl „bar changed“. Für die bei Google Glass zusätzlich realisierte Zoomfunktion waren die Sprachbefehle „zoom Image“ zum Hereinzoomen und „scale down“ zum Herauszoomen. Da diese beiden Begriffe manchmal vom System nicht verstanden wurden, wurde allen Probanden auch eine manuelle Möglichkeit zum Zoomen erläutert: Ein Tippen auf den Brillenbügel zoomt hinein und mit dem nächsten Slide zoomt das System automatisch wieder hinaus.

4.2.2 Apparatur

Im Gegensatz zur Originalstudie stammten diesmal beide Geräte aus dem Konsumentenbereich.

4.2.2.1 Google Glass

Google Glass ist ein monokulares HMD in Brillenformat, dessen Bildschirm sich immer am rechten oberen Rand befindet. Im Gegensatz zum MAVUS-HMD kann die Seite des Bildschirms nicht gewechselt werden. Das HMD besitzt eine Auflösung von 640 x 360 Pixeln und wiegt 50 Gramm. Während des Versuches war es mit einem Akkumulator-Erweiterungspack verbunden, damit eine dauerhafte Anzeige über die Gesamtzeit von zwei Stunden möglich war.

4.2.2.2 Tablet-PC

Bei dem Tablet-PC handelt es sich um einen Samsung Galaxy SM-T210 mit einer Auflösung von 1024 x 600 Pixeln. Das Display hat einen Durchmesser von 17,8 cm (7“) und das Gerät wiegt 300 Gramm.

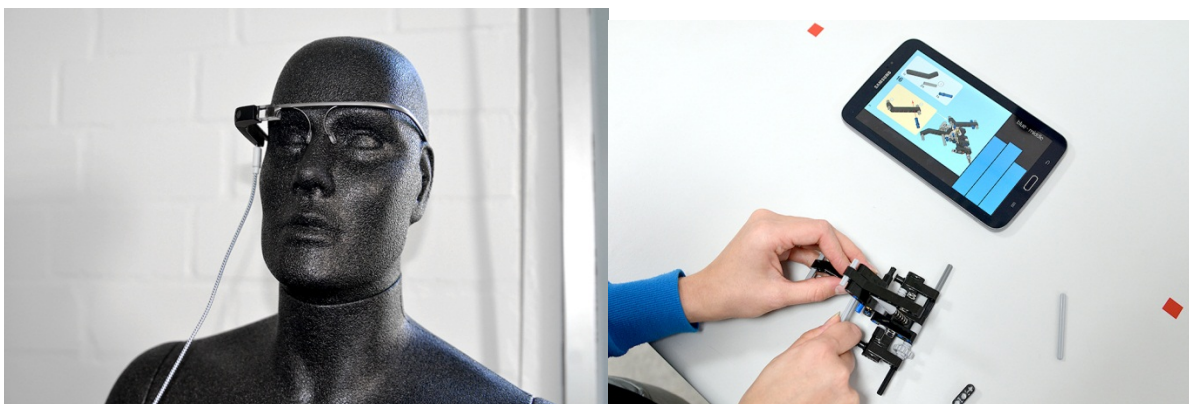


Abb. 4.2 Links: Google Glass mit angeschlossenem Akkumulatorexpansionspack. Rechts: der verwendete Tablet-PC.

4.2.3 Versuchsplan

4.2.3.1 Unabhängige Variablen

- Displaytyp

Die Probanden arbeiteten entweder mit dem HMD oder dem Tablet-PC als Informationsanzeigegerät. Hierbei wurde in dieser Replikationsstudie im Gegensatz zur Originalstudie ein between-subject Design gewählt, so dass die Hälfte der Probanden in der einen und die andere Hälfte in der anderen Bedingung war.

- Alter der Probanden

Die Probanden wurden wieder nach dem Alter in zwei Gruppen eingeteilt. Hierzu wurde ein Mediansplit der Gesamtpopulation vorgenommen. Eine theoretische Herleitung der Altersgruppen fand nicht statt. Bei der Verteilung auf die Displaygruppen wurde darauf geachtet, dass das Alter sich möglichst gleich verteilt.

- Technikaffinität

Die Technikaffinität wurde ebenfalls wieder erhoben, bzw. es wurde auf die Werte der vorangegangenen Untersuchung zurückgegriffen (Fragebogen TA-EG, Karrer et al., 2009, vgl. Anhang 1). Diese wurden dann ebenfalls mittels Mediansplit in zwei Gruppen eingeteilt und als Zwischensubjektfaktor verwendet, um Einflüsse der Technikaffinität auf die Ergebnisse zu untersuchen.

Die unabhängige Variable „Rückmeldung in der Monitoringaufgabe“, die in der Originalstudie blockweise variiert wurde, fällt hier weg. Statt dessen wird die Rückmeldung immer gegeben, da sie sich in der vorherigen Studie als positiv für bessere Reaktionen und Reaktionszeiten erwiesen hat, aber keine Wechselwirkung mit dem Displaytyp oder anderen Faktoren aufgetreten ist.

4.2.3.2 Abhängige Variablen

- Anzahl der bearbeiteten Konstruktionsfolien

Als Indikator für die Arbeitsleistung wird wieder die Anzahl der bearbeiteten Konstruktionsfolien innerhalb eines bestimmten Zeitfensters gesehen. Hierbei wurde ein Zeitfenster von 25 Minuten genommen.

- Trefferquote in den Monitoringaufgaben

Die Trefferquote ist der prozentuale Anteil der richtigen Reaktionen von allen Reaktionen. Falsche Reaktionen können entweder das Unterlassen einer Reaktion trotz Veränderung in der Monitoringaufgabe sein, oder die Reaktion, obwohl sich nichts in der Monitoringaufgabe verändert hat. Das Zeitfenster für eine Reaktion auf einen veränderten Stimuli ist zeitlich nicht fest begrenzt (wie z. B. durch eine maximale Reaktionszeit von fünf Sekunden), sondern wird nur

durch die nächste Stimuliveränderung, die eine erneute Reaktion bedingt, begrenzt.

- (Reaktionszeit in den Monitoringaufgaben)

Die Reaktionszeit in den Monitoringaufgaben für die Fälle der richtigen Reaktionen wurde ebenfalls ausgewertet. Die Interpretation der Ergebnisse ist aber nur eingeschränkt möglich, da sich bei diesem Versuchssetup nicht nur die Displays zur Informationsanzeige unterscheiden, sondern diesmal auch die Arten der Interaktion. Es ist davon auszugehen, dass die Berührungsinteraktion auf dem Tablet-PC immer schneller ist als die Sprachinteraktion bei dem HMD. Dies ist schon alleine deshalb der Fall, weil für die Sprachinteraktion immer erst der gesamte Sprachbefehl gesprochen und vom System interpretiert werden muss, was zusammengefasst einige Sekunden dauern kann. So ist anzunehmen, dass die Reaktionszeiten beim Tablet-PC schon alleine aufgrund der Interaktion immer schneller sind als beim HMD. Da nun nicht unterschieden werden kann, ob die Effekte auf der unterschiedlichen Informationsdarstellung oder auf der unterschiedlichen Interaktion beruhen, ist eine Interpretation nur mit Vorbehalt möglich.

- NASA-Task Load Index (NASA-TLX)

Der NASA-TLX wurde am Ende des Gesamtversuches nach zwei Stunden erhoben. Der Zeitpunkt ist somit vergleichbar mit dem Zeitpunkt der ersten NASA-TLX-Erhebung in der Originalstudie während der zweiten Pause. Obwohl beim Durchführen des NASA-TLX betont wurde, dies auf die Beanspruchung der letzten halben Stunde zu beziehen, in der die Replikationsstudie durchgeführt wurde, ist nicht auszuschließen, dass manche Probanden auch unbewusst das Gesamtexperiment inklusive der anderen beiden Studien in die Bewertung mit einfließen lassen. Dies muss bei der Interpretation der Ergebnisse bedacht werden.

- Rating Scale of Mental Effort (RSME) zu vier Zeitpunkten.

Der RSME als weiterer subjektiver Beanspruchungsparameter wurde wieder jede halbe Stunde erhoben, so dass insgesamt vier Messungen vorliegen, die zu vergleichbaren Zeiten wie die Messzeitpunkte in der Originalstudie sind. Aber auch hier wurden in den ersten eineinhalb Stunden andere Aufgaben bearbeitet, was eine Übertragbarkeit dieser Messzeitpunkte einschränkt. Für einen Überblick über die Beanspruchung beim längeren Arbeiten mit einem HMD sind die Werte aber zulässig, auch wenn die Aufgabentypen wechseln.

- Visual Fatigue Questionnaire (VFQ) zu drei Zeitpunkten

Der VFQ (vgl. Anhang 1) wurde vor Beginn des Experimentes zum Ermitteln der Ausgangslage ausgefüllt. Anschließend in der Pause nach einer Stunde, in der davor die Studie im Biologiesetup stattfand, und am Ende des Experimentes nach zwei Stunden, in denen davor die Studie zur Fehlersuche und die Replikationsstudie stattfanden. Auch hier sind die Zeitpunkte denen der Originalstudie angeglichen. Die Arbeitsaufgaben variieren aber systematisch

über die Zeit. Auch ist zu beachten, dass in der ersten Stunde alle Probanden das HMD trugen und erst in der zweiten Stunde die Population in Tablet-PC- und HMD- Nutzer aufgeteilt wurde.

- Interview

Am Ende der Studie fand wieder ein Interview statt. Die Fragen des Interviews orientieren sich an denen der Originalstudie. Sie wurden aber noch um Fragen des speziellen neuen Setups ergänzt (z. B. zur Spracherkennung), und bei Probanden, die bereits in der Originalstudie mitwirkten, wurde auch ein Vergleich beider HMDs abgefragt.

4.2.4 Versuchsablauf

Die Untersuchungsreihe wurde wie die Originalstudie in den Laborräumen der BAUA in Dortmund durchgeführt. Der Raum war gut ausgeleuchtet und mögliche Blendquellen, wie etwa das Fenster, wurden verdeckt. Die Replikationsstudie fand im Gegensatz zur Originalstudie im Stehen statt an einem auf die Standhöhe der einzelnen Probanden angepassten Tisch. Hinter dem Tisch befand sich eine weiße Wand, um mögliche Irritationen durch Muster bei dem see-through-HMD zu vermeiden.

Nach der Begrüßung der Probanden erhielten diese zunächst wieder eine Probandeninformation in schriftlicher Form (vgl. Anhang 2). Diese ähnelt der Probandeninformation aus der Originalstudie (vgl. Anhang 1), enthält aber auch ein paar abgewandelte Passagen. Die Probanden wurden darüber aufgeklärt, dass sie freiwillig an dem Versuch teilnehmen und ihn jederzeit ohne Angabe von Gründen abbrechen können. Weiterhin wurde erklärt, dass während der gesamten Versuchszeit die Bewegungen des dominanten Handgelenkes aufgezeichnet werden und dazu ein Uhren-ähnliches Gerät zu tragen ist. Hiermit sollen in einer späteren Auswertung des ersten Abschnittes der Studienreihe zur Laborerleichterung die Handgelenksbewegungen einzelne Handlungen zugeordnet werden. Des Weiteren wurde erwähnt, dass der Versuch nicht der medizinischen Diagnostik gilt und dass während des Versuches eine Videoaufzeichnung läuft. Diese dient zunächst ausschließlich zur Kontrolle und Zuordnung der Daten und wird nicht veröffentlicht, es sei denn die Probanden geben dazu ihre ausdrückliche Einwilligung, was am Ende des Versuches getrennt abgefragt wurde. Der Hintergrund einer möglichen Videoveröffentlichung bezieht sich auf den ersten Teil der Studienreihe zur Laborunterstützung (die nicht Teil dieses Berichtes ist): Hier sollen die Videodaten anderen Forschern zur Verfügung gestellt werden, damit sie die Übereinstimmung von Handgelenksbewegungen zu bestimmten Arbeitsschritten nachvollziehen können. Die daran anschließende Einwilligungserklärung mussten die Probanden unterschreiben. Danach wurde den Probanden, die noch nicht an der Originalstudie teilgenommen hatten, der Fragebogen zur Technikaffinität gereicht (vgl. Anhang 1) und eine Abfrage der Personenkenntwerte vorgenommen (vgl. Anhang 1). Anschließend wurde den Probanden wieder die RSME-Skala vorgestellt und zur Einübung mehrere Alltagssituationen abgefragt (vgl. Anhang 1). Vor Versuchsbeginn wurde dann noch der Fragebogen zur visuellen Ermüdung als Ausgangslage erhoben (vgl. Anhang 1).

Die Versuchsreihe begann immer mit der Studie zur Laborunterstützung (vgl. Abb. 4.3). Hier wurden die Probanden zunächst in Google Glass, seine Funktionsweise und Bedienung eingeführt. Dann wurde in einem Laborsetting die DNA aus Zwiebeln und Tomaten extrahiert, wobei die Anweisungen auf dem HMD in schriftlicher Form

erfolgten. Eine spätere Analyse soll den Handbewegungen bestimmte Arbeitsschritte zuordnen. Zusätzlich in dieser Teilstudie wurde ein Warnhinweis variiert, der Probanden dazu bringen sollte, eine bestimmte Flüssigkeit immer in den Kühlschrank zu legen. Während dieses Teilversuches, den alle Probanden mit dem HMD durchführten, wurde nach 25 und nach 50 Minuten der RSME abgefragt. Auf diesen Versuchsteil wird hier nicht weiter eingegangen, da er nicht Teil der Replikation ist.

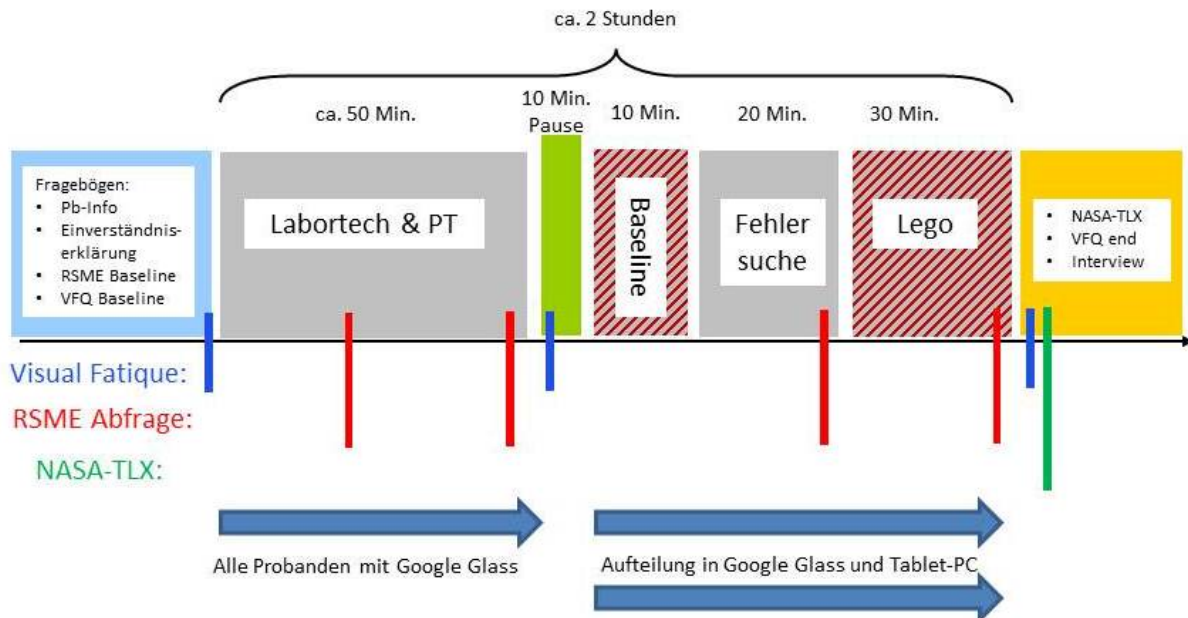


Abb. 4.3 Schematischer Versuchsablauf der gesamten Studienreihe. Die hier berichtete Replikation fand in der zweiten Stunde in den rot gestreiften Abschnitten statt. Weitere Erläuterungen im Text.

Es folgte eine zehnmünütige Pause, zu deren Beginn der VFQ abgefragt wurde. Nach der Pause wurden die Probanden in zwei Gruppen eingeteilt, wobei auf eine angemessene Altersverteilung geachtet wurde. Die Gruppen arbeiteten im Anschluss entweder mit dem HMD oder dem Tablet-PC weiter. Entsprechend erhielten sie die Instruktion für die nächste Stunde (vgl. Anhang 2), in der die Interaktionsformen (Bewegungssteuerung beim Tablet-PC, Sprachbefehle beim HMD) noch einmal erläutert wurden. Im Anschluss daran wurden die Interaktionen im Beisein des Versuchsleiters geübt. Hierbei wurde bei der Spracheingabe auch erfasst, in wievielen von fünf Wiederholungen der Proband vom System verstanden wurde. Da die Sprachbefehle in Englisch waren, zeigte sich in den Pretests, dass manche Probanden schlecht verstanden wurden. Deshalb waren die Sprachbefehle mehrfach vereinfacht worden. Eine Evaluation der Sprachverständlichkeit der einzelnen Probanden wurde dennoch als wichtig angesehen. Im Anschluss daran wurde die Monitoringaufgabe für fünf Minuten geübt. Hier sollten die Probanden ausschließlich auf die Monitoringaufgabe achten, so dass diese im Singel-task Modus durchgeführt wurde. Neben der Übung kann ein Vergleich dieser Baseline mit der später im Legoversuch durchgeführten Monitoringaufgabe im Dual-task Modus verglichen werden, um so Performanzeinbußen, die auf dem Dual-task Paradigma beruhen, aufzuzeigen. Dieser „Baseline“ genannte Block der Studienreihe dauerte ca. zehn Minuten.

Diesem Block folgte dann eine Studie zur Fehlersuche an einem Schaltschrank, die ca. 25 Minuten dauerte und von den Probanden je nach Gruppeneinteilung entweder mit Google Glass oder dem Tablet-PC durchgeführt wurde. Auf diese Teilstudie wird hier ebenfalls nicht eingegangen.

In dem letzten Block – nach ca. eineinhalb Stunden anderer Aufgaben - folgte dann die Replikation der Hauptstudie dieses Berichtes: Für 25 Minuten bauten die Probanden ein Lego-Technic-Fahrzeug zusammen (Modell 9293, siehe Anhang 1) und achteten parallel auf die Monitoringaufgaben (Farbwechsel und Positionswechsel des längsten Balkens). Hierbei wurden die Inhalte je nach Gruppeneinteilung entweder auf Google Glass oder auf dem Tablet-PC gezeigt. Zu Beginn erhielten die Probanden erneut eine schriftliche Instruktion, die die Aufgaben und Interaktionen beschrieb (siehe Anhänge). Die Versuchsanordnung unterschied sich von der Originalstudie nicht nur in der kürzeren Dauer, sondern auch darin, dass die Probanden diesmal standen und die Lego-Bausteine von einem ca. vier Meter entfernten Punkt aufgenommen werden mussten und dann am Arbeitstisch zusammengebaut werden sollten. Diese Vorgehensweise wurde gewählt, um den Mobilitätsaspekt zu betonen: Tablet-PC-Nutzer mussten diesen manuell mitnehmen, während HMD-Nutzer die nötigen Informationen automatisch immer mitnahmen. Am Ende dieses Studienteils, kurz vor Schluss, wurde dann in der Arbeitssituation noch einmal der RSME abgefragt. Im Anschluss an diese letzte Teilstudie der Studienreihe wurde erneut der VFQ vorgelegt und der NASA-TLX wurde am Computer abgefragt. Es folgte dann noch ein Interview, das sich an das Interview der Originalstudie anlehnt, aber auch weitere Fragen ergänzt (siehe Anhänge). Das Interview stellte das Ende der Studienreihe dar. Es wurden dann noch die Kontaktdaten der Probanden für die Ausbezahlung der Aufwandsentschädigung (10,- Euro pro Stunde, ca. 30 Euro für die Studienreihe) aufgenommen.



Abb. 4.4 Links: Aufnahme der Lego-Technic-Bausteine an einem Ort (hier mit Tablet-PC). Rechts: Zusammenbau des Fahrzeuges an einem vier Meter entfernten Ort (hier exemplarisch mit Google Glass, inkl. Akkuerweiterung).

4.2.5 Probanden

An dieser Replikationsstudie nahmen insgesamt 37 Probanden teil. Keiner der Probanden hat den Versuch abgebrochen, so dass alle Probanden in die Auswertung mit einfließen konnten. Im Rahmen des between-subject Designs wurden die Probanden den beiden Gruppen „HMD“ und „Tablet-PC“ zugeteilt, wobei auf eine ver-

gleichbare Altersverteilung geachtet wurde. 20 Probanden waren in der HMD-Gruppe mit einem durchschnittlichen Alter von 37,25 Jahren ($SD = 13,803$) und 17 Probanden in der Tablet-PC-Gruppe mit einem durchschnittlichen Alter von 37,06 Jahren ($SD = 16,006$). Abbildung 4.5 gibt die Verteilung für beide Gruppen wieder.

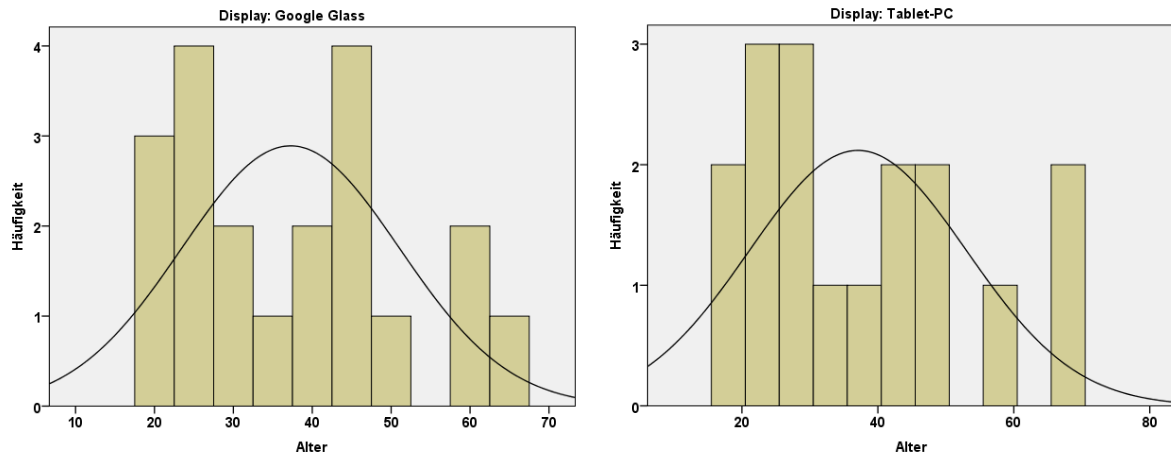


Abb. 4.5 Altersverteilung der HMD- und Tablet-PC-Gruppe der Replikationsstudie. Links: Google Glass. Rechts: Tablet-PC.

Die Geschlechtsverteilung ist diesmal in beiden Gruppen nicht ausgewogen: In der HMD-Gruppe waren 12 Frauen und 8 Männer und in der Tablet-PC-Gruppe 7 Frauen und 10 Männer. Da in der vorherigen Studie aber keine Geschlechtsunterschiede in den abhängigen Variablen festgestellt werden konnten, sollte diese Tatsache keinen weiteren Einfluss haben.

Durch einen Mediansplit der Altersverteilung wurden in beiden Gruppen die Probanden wieder in Jüngere und Ältere unterteilt. Die 10 jüngeren Probanden der HMD-Gruppe waren durchschnittlich 25,6 Jahre alt ($SD = 4,971$), die 10 älteren Probanden 48,9 Jahre ($SD = 8,711$). Die 8 jüngeren Probanden der Tablet-PC-Gruppe waren im durchschnitt 23,38 Jahre alt ($SD = 3,543$) und die 9 älteren Versuchsteilnehmer 49,22 Jahre ($SD = 12,153$).

4.2.6 Hypothesen

4.2.6.1 Leistung in der Konstruktionsaufgabe

H6: Die Anzahl der bearbeiteten Konstruktionsfolien unterscheidet sich zwischen in den Bedingungen HMD und Tablet-PC.

4.2.6.2 Trefferquote in der Monitoringaufgabe

H7: Die Trefferquote in der Monitoringaufgabe unterscheidet sich zwischen in den Bedingungen HMD und Tablet-PC.

4.2.6.3 Subjektive Beanspruchung

H8: Die Werte des gewichteten Gesamtmittelwertes des NASA-TLX unterscheiden sich zwischen den Bedingungen HMD und Tablet-PC.

H9: Die Werte des RSME unterscheiden sich zwischen den Bedingungen HMD und Tablet-PC.

4.2.6.4 Visuelle Ermüdung

Die Werte der einzelnen Items des VFQ werden verglichen und Unterschiede zwischen den beiden Displaygruppen herausgearbeitet. Auf eine Aufführung sämtlicher Unterschiedshypothesen für die 16 Items wird hier aus Übersichtsgründen verzichtet. Die durch das Fehlen eines Gesamtmittelwertes eventuell nötige Korrektur des Alphaniveaus oder Einschränkungen bei der Interpretation der Signifikanzen werden in Abschnitt 2.7.6 bereits diskutiert.

4.2.7 **Datenaufbereitung**

Alle Veränderungen der den abhängigen Variablen zugrunde liegenden Ereignisse (Sprach- bzw. Touchbefehle, angezeigte Folie in der Konstruktionsaufgabe, Farbe und Position des längsten Balkens in der Monitoringaufgabe) wurden auf dem jeweiligen Anzeigegerät mit Zeitstempel in eine Datenmatrize geschrieben, die im Anschluss an den Versuch von einem Laptop ausgelesen wurde. Die weitere Datenaufbereitung zur Ermittlung der Kennwerte (z. B. Differenz zwischen Ereignis und Reaktion als Reaktionszeit) wurde mit Hilfe von Filtern und Syntax in SPSS 21 vorgenommen.

Hierbei zeigte sich, dass die Trefferquote beim Farbwechsel der Balken in der Monitoringaufgabe nicht hundertprozentig vertrauenswürdig ist: Ein versteckter Prozess in Android schrieb fehlerhaften Farbwechsel von Ereignissen in die Ergebnismatrix, die nicht auf dem Display dargestellt wurden und auf die somit die Probanden nicht reagieren konnten. Das Problem trat bei beiden Displaytypen in gleicher Form auf und betrifft circa 10% der Fälle. In der Ergebnismatrize ist zwischen diesen „Ghostvalues“ und wirklichen verpassten Reaktionen nicht zu unterscheiden und sie können nur grob auf Basis der Wiederholwahrscheinlichkeit des Farbwechsels identifiziert werden. Dies führt dazu, dass die Anzahl der verpassten Reaktionen („Misses“) bei der Farbwechselmonitoringaufgabe in diesem Versuch leicht erhöht ist. Die Monitoringaufgabe zum Wechsel des längsten Balkens ist davon nicht betroffen.

4.2.8 **Inferenzstatistische Verfahren**

Als inferenzstatistisches Verfahren wurde eine Varianzanalyse im Rahmen des allgemeinen linearen Modells im mixed design angewendet. Hierbei sind diesmal (im Gegensatz zu der Studie mit dem MAVUS-HMD) die beiden Displaytypen between-subject Faktoren. Die verschiedenen Messzeitpunkte über den Untersuchungsverlauf, wie sie bei den abhängigen Variablen RSME und VFQ vorkommen, sind within subject Faktoren.

4.3 Ergebnisse

4.3.1 Leistung in der Konstruktionsaufgabe

Bei der Anzahl der bearbeiteten Konstruktionsfolien während der Versuchszeit (siehe Abb. 4.6) zeigt sich ein vergleichbares Muster wie bei der Untersuchung mit dem industriellen MAVUS-HMD: Mit dem Tablet-PC werden signifikant mehr Folien in der gleichen Zeit abgearbeitet als mit dem HMD [$F(1, 33) = 8.691$; $p = .006$], ältere Probanden bearbeiten allgemein weniger Folien als Jüngere [$F(1, 33) = 15.789$; $p < .001$], wobei es aber keine Wechselwirkung mit dem Displaytyp gibt [$F(1, 33) = .091$; $p = .765$], was bedeutet, dass Performanzeinbußen im Alter unabhängig vom Displaytyp sind.

Die gefundenen statistischen Effekte der Replikationsstudie sind also die gleichen wie bei der vorherigen Untersuchung, was die Befundlage untermauert. Die Hypothese 6 kann also angenommen werden. In diesem Fall werden ca. 33% mehr Folien mit dem Tablet-PC in der Versuchszeit bearbeitet.

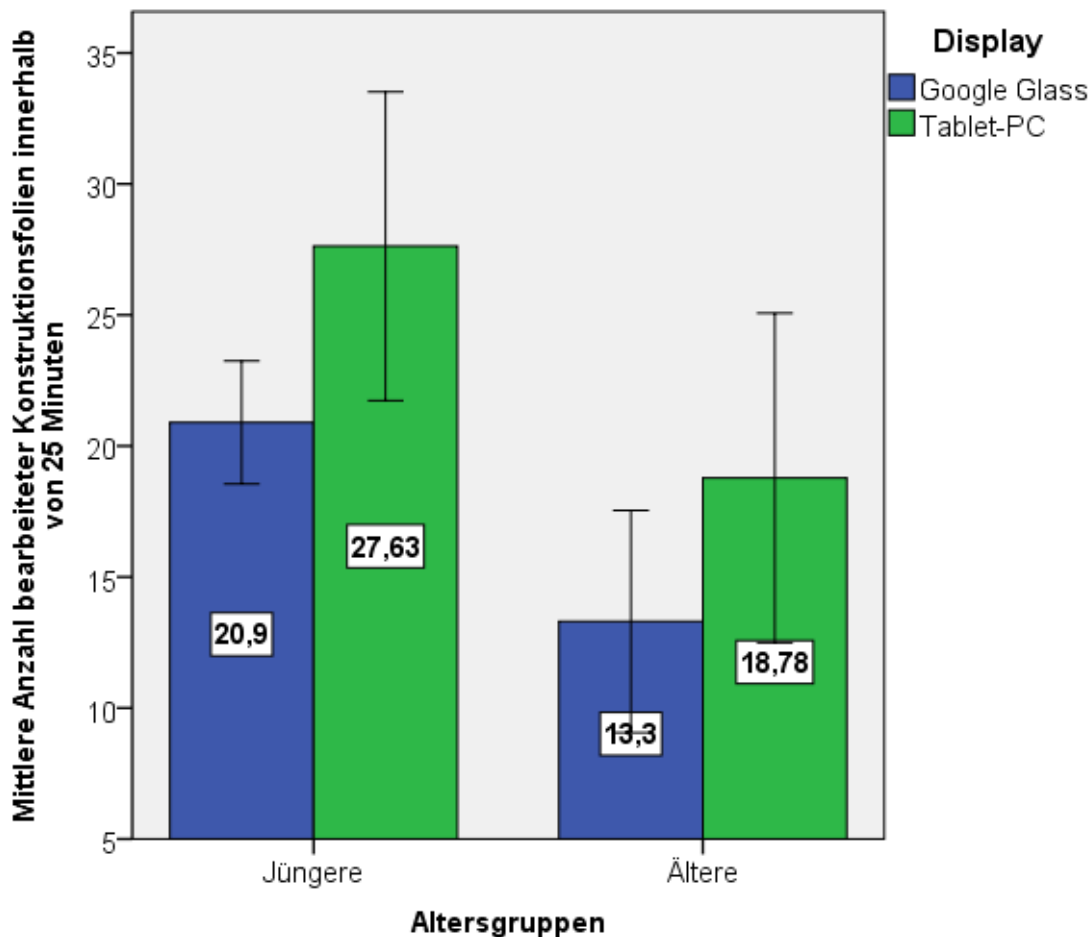


Abb. 4.6 Mittlere Anzahl bearbeiteter Konstruktionsfolien innerhalb von 25 Minuten Versuchszeit unterteilt nach Altersgruppen und Displaytyp. Fehlerbalken zeigen das 95%-Konfidenzintervall.

4.3.2 Leistung in den Monitoringaufgaben

Bei den beiden Monitoringaufgaben (Reaktion auf Farbwechsel und Reaktion auf Wechsel der Position des längsten Balkens) gibt es im Gegensatz zur Originalstudie mit dem industriellen HMD diesmal auch eine Baseline-Bedingung, bei der die Monitoringaufgaben für ca. fünf Minuten alleine ohne zusätzliches Legobauen getestet wurden. Diese Bedingung ist als Kontrolle anzusehen, ob die zu erwartenden Leistungseinbußen bei paralleler Bearbeitung (Dualtask) zutreffen.

4.3.2.1 Trefferquote

Bei den Reaktionen auf den Farbwechsel der Balken (Abb. 4.7) zeigt sich in der Trefferquote eine zu erwartende hoch signifikant bessere Reaktion unter der Baselinebedingung als unter der Dualtask-Bedingung, bei denen gleichzeitig noch Lego aufgebaut wird [$F(1, 33) = 21.171$; $p < .001$]. Im Gegensatz zur Originalstudie zeigt sich hier auch ein signifikanter Effekt des Alters [$F(1, 33) = 5.383$; $p = .027$], der wiedergibt, dass die Reaktion älterer Probanden weniger korrekt ist. Des Weiteren wird hier auch der Unterschied zwischen den Displays, der in der Originalstudie eine Tendenz zu besseren Reaktionen auf dem Tablet-PC aufweist signifikant [$F(1, 33) = 7.827$; $p = .009$]. Die besseren Reaktionen finden auf dem Tablet-PC statt.

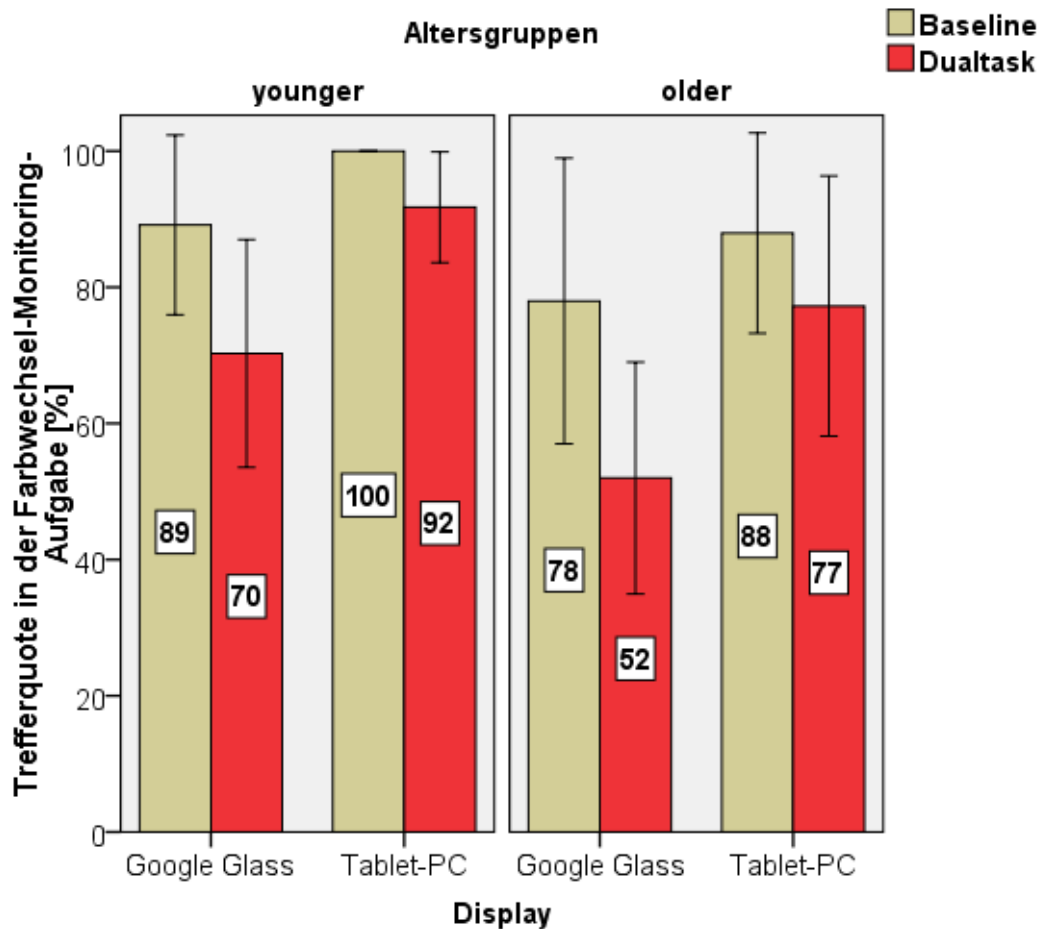


Abb. 4.7 Trefferquote in Prozent in der Farb-Monitoring-Aufgabe unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Zusätzlich gibt es noch eine Wechselwirkung zwischen den Faktoren Dualtask und Display: Die mit der Dualtask einhergehenden Performanzeinbußen sind bei Google Glass tendenziell höher als bei dem Tablet-PC [$F(1, 33) = 3.487$; $p = .071$], was unterstreicht, dass HMDs für paralleles Monitoring weniger gut geeignet sind. Bei der Interpretation der Daten sei noch einmal darauf hingewiesen, dass die Trefferquoten beim Farbwechsel einige Fehlerwerte enthalten, die durch den technischen Aufbau entstanden sind (siehe 4.2.7), sich aber über alle Bedingungen gleich verteilen.

Bei der Trefferquote zum Positionswechsel des längsten Balkens (Abb. 4.8) zeigt sich ebenfalls ein signifikanter Effekt von Dualtask [$F(1, 33) = 5.529$; $p = .025$] sowie ein signifikanter Effekt des Alters [$F(1, 33) = 5.277$, $p = .028$], der aufzeigt, dass ältere Probanden allgemein schlechter reagieren. Der Unterschied zwischen den Displays erreicht hier wie in der Originalstudie nur statistische Tendenz [$F(1, 33) = 3.140$; $p = .086$] in Richtung besserer Werte für den Tablet-PC. Eine Wechselwirkung Dualtask*Display liegt hier nicht vor.

Die Hypothese 7, die besagt, dass es Unterschiede in den Trefferquoten je nach Displaytyp gibt, kann also nur bedingt angenommen werden: Die Unterschiede gehen in Richtung besserer Werte für den Tablet-PC, erreichen aber nur bei der Farbwechsel-Monitoring-Aufgabe Signifikanz, während sonst eine statistische Tendenz vorliegt.

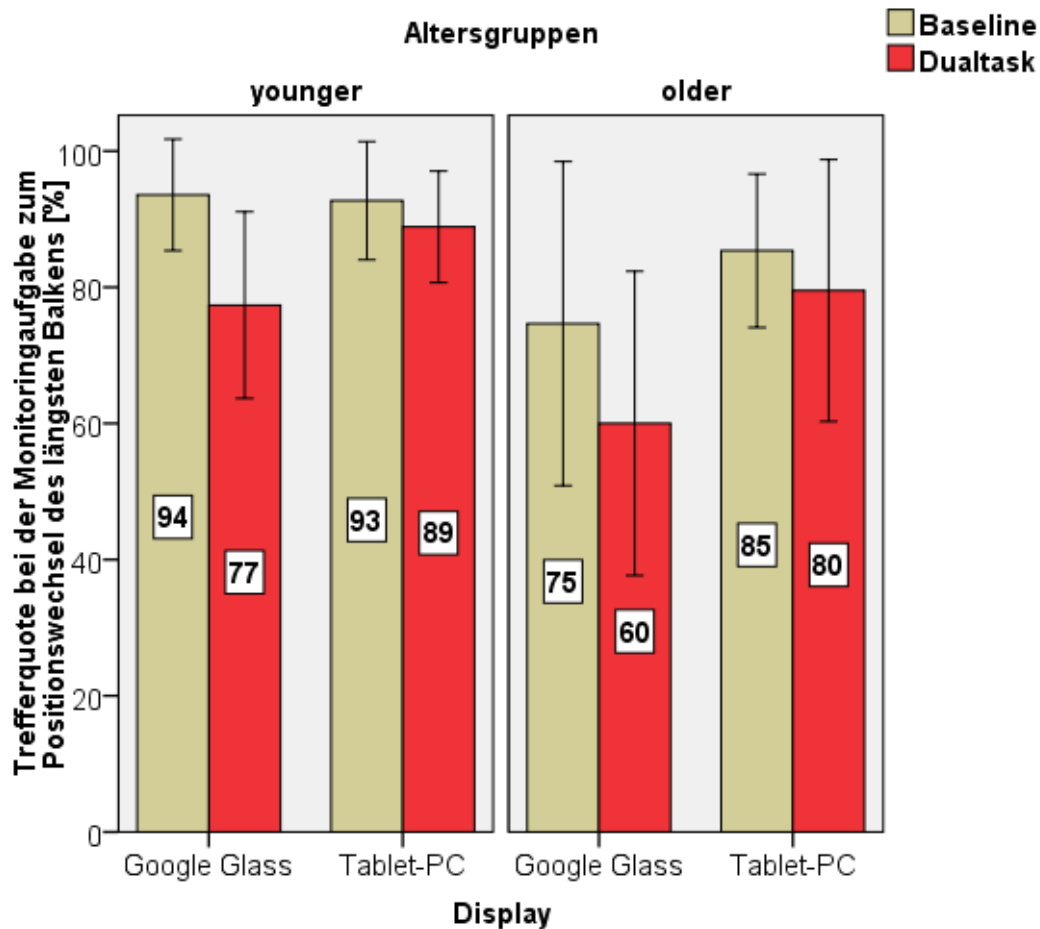


Abb. 4.8 Trefferquote in Prozent in der Monitoringaufgabe zum Positionswechsel des längsten Balkens unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.

4.3.2.2 Reaktionszeit

Bei einem Vergleich der Reaktionszeiten muss beachtet werden, dass sich nicht nur das Display zur Informationsdarstellung unterscheidet (HMD oder Tablet-PC), sondern auch die Interaktionsformen unterschiedlich sind. So wird die Sprachinteraktion immer zu höheren Reaktionszeiten führen als die Touch-Interaktion, da hier erst der Sprachbefehl entkodiert werden muss, was erst nach Beendigung des Sprachkommandos möglich ist, während bei der Berührungsinteraktion die Reaktionszeit aber sofort erhoben wird. Von daher ist ein Vergleich der Reaktionszeiten hier nicht fair und deswegen wurden keine Hypothesen dazu gebildet. Eine Darstellung der Reaktionszeiten erfolgt aber trotzdem für einen umfangreichen Gesamtüberblick.

Abb. 4.9 zeigt die Reaktionszeiten in der Farbwechsel-Monitoring-Aufgabe. Bei einem statistischen Vergleich würde sich ein tendenzieller Effekt des Dualtask beim Vergleich mit der Baseline zeigen [$F(1, 33) = 3.192$; $p = .083$], der längere Reaktionszeiten unter Dualtaskbedingungen wiedergibt. Des Weiteren würde sich ein signifikanter Effekt des Displaytyps zeigen [$F(1, 33) = 5.112$; $p = .030$] mit längeren Reaktionszeiten für das HMD, aber kein Alterseffekt [$F(1, 33) = 1.724$, $p = .198$].

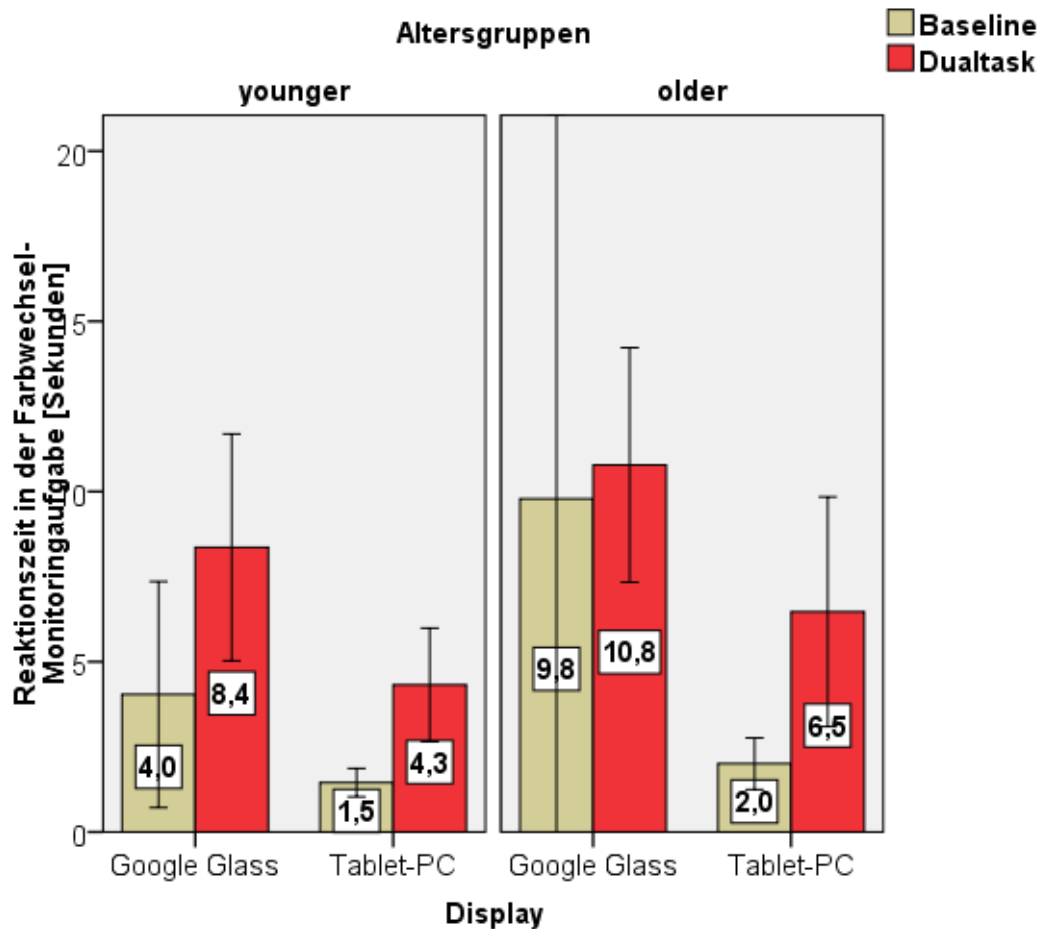


Abb. 4.9 Reaktionszeiten in Sekunden in der Farbwechsel-Monitoringaufgabe unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Abbildung 4.10 zeigt Reaktionszeit in Sekunden in der Monitoringaufgabe zum Positionswechsel des längsten Balkens. Bei einem statistischen Vergleich würde sich eine signifikante Verlängerung der Reaktionszeit für Dualtask im Vergleich zur Baseline zeigen [$F(1, 32) = 16.604$; $p < .001$], aber keine Unterschiede bezüglich des Displays [$F(1, 32) = .053$; $p = .819$]. Die schnelleren Reaktionszeiten für jüngere Probanden im Vergleich zu älteren hätte nur statistische Tendenz [$F(1, 32) = 3.175$; $p = .084$].

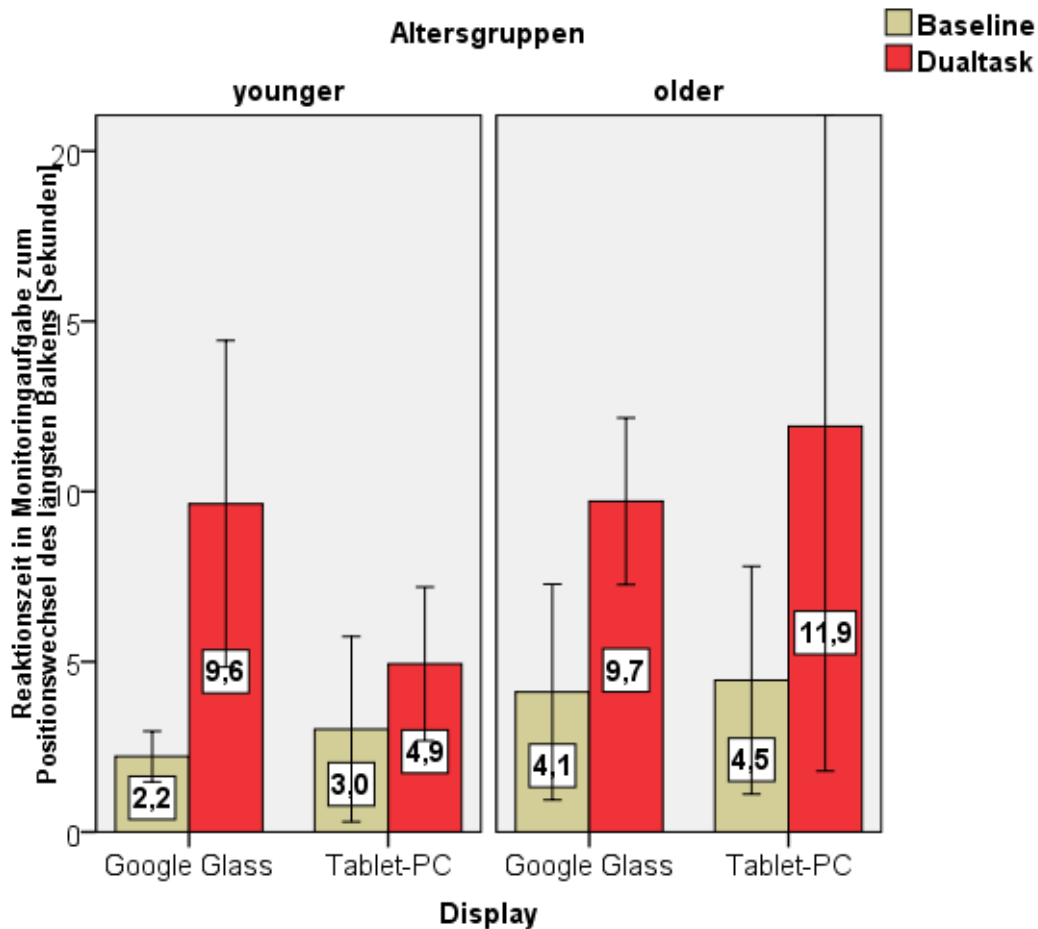


Abb. 4.10 Reaktionszeit in Sekunden in der Monitoringaufgabe zum Positionswechsel des längsten Balkens unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.

Wie bereits zu Beginn dieses Abschnittes erwähnt sind die Reaktionszeiten hier aber nicht nur von dem Displaytyp, sondern auch von der unterschiedlichen Interaktionsform bestimmt. Diese Konfundierung der Faktoren macht eine Interpretation schwierig bis unmöglich. Orientiert man sich an der Originalstudie, bei der die Interaktion von HMD und Tablet-PC auf dem gleichen umgebauten Nummernpad stattfand, so gingen dort die Reaktionszeitunterschiede zwar in die gleiche Richtung (schnellere Reaktionen auf dem Tablet-PC), erreichten aber gerade noch statistische Tendenz.

4.3.3 Subjektive Beanspruchung

4.3.3.1 NASA-TLX

Der gewichtete Summenscore des NASA-TLX (Abb. 4.11) zeigt signifikant höhere Werte für die HMD-Nutzung im Vergleich zur Tablet-PC-Nutzung [$F(1, 32) = 6.645$; $p = .015$]. Ebenso zeigt sich ein signifikanter Effekt des Alters [$F(1, 32) = 5.952$; $p = .020$], der wiedergibt, dass ältere Probanden mehr subjektive Beanspruchung empfinden. Dies ist aber unabhängig vom Displaytypus, da keine Wechselwirkung zwischen Alter und Displaytyp vorliegt [$F(1, 32) = 1.380$; $p = .249$]. Die Hypothese 8

kann also angenommen werden: Die NASA-TLX-Werte unterscheiden sich signifikant zwischen den Displaybedingungen.



Abb. 4.11 Gewichteter Summenscore des NASA-TLX, unterteilt nach Altersgruppen und Displaytyp. Fehlerbalken zeigen das 95%-Konfidenzintervall.

4.3.3.2 Rating Scale of Mental Effort (RSME)

Die Werte des RSME wurden zu vier Zeitpunkten im Verlauf des Gesamtversuches erhoben. Die Replikationsstudie mit Google Glass fand in der letzten halben Stunde des Gesamtversuches statt (siehe Abb. 4.3), so dass hier ein Vergleich der Messzeitpunkte „Minute 90“ und „Minute 120“ herangezogen werden, die den Beginn und das Ende der hier beschriebenen Studie repräsentieren.

Abbildung 4.12 zeigt die mittleren Werte der RSME-Skala unterteilt nach Altersgruppen, Displaytyp und Zeitpunkt. Es zeigt sich zwar eine signifikante Steigerung der RSME-Werte über die 30 Minuten Studienzeit hinweg [$F(1, 32) = 23.191$; $p < .001$], aber kein Effekt des Displaytyps [$F(1, 32) = 2.472$; $p = .126$]. Ein Effekt der Altersgruppen erreicht statistische Tendenz [$F(1, 32) = 3.252$; $p = .081$], eine Wechselwirkung zwischen Alter und Display liegt aber nicht vor [$F(1, 32) = 1.335$; $p = .257$].

Die Hypothese 9 muss also zurückgewiesen werden: Die RSME-Werte unterscheiden sich bei diesem Vergleich nicht hinsichtlich der Displaytypen. Eine Ursache da-

für, dass der Vergleich der RSME-Werte in dieser Studie keine Unterschiede bringt, während in der Originalstudie deutliche Unterschiede zu finden sind, könnte das unterschiedliche Versuchsdesign sein. In der Originalstudie wurde ein within-subject Design gewählt, während hier ein between-subject Design vorliegt, bei dem gerade bei geringer Stichprobengröße individuelle Antworttendenzen der Probanden einen großen Anteil an den Ergebnissen haben. Der RSME ist eher für das within-subject Design geeignet.

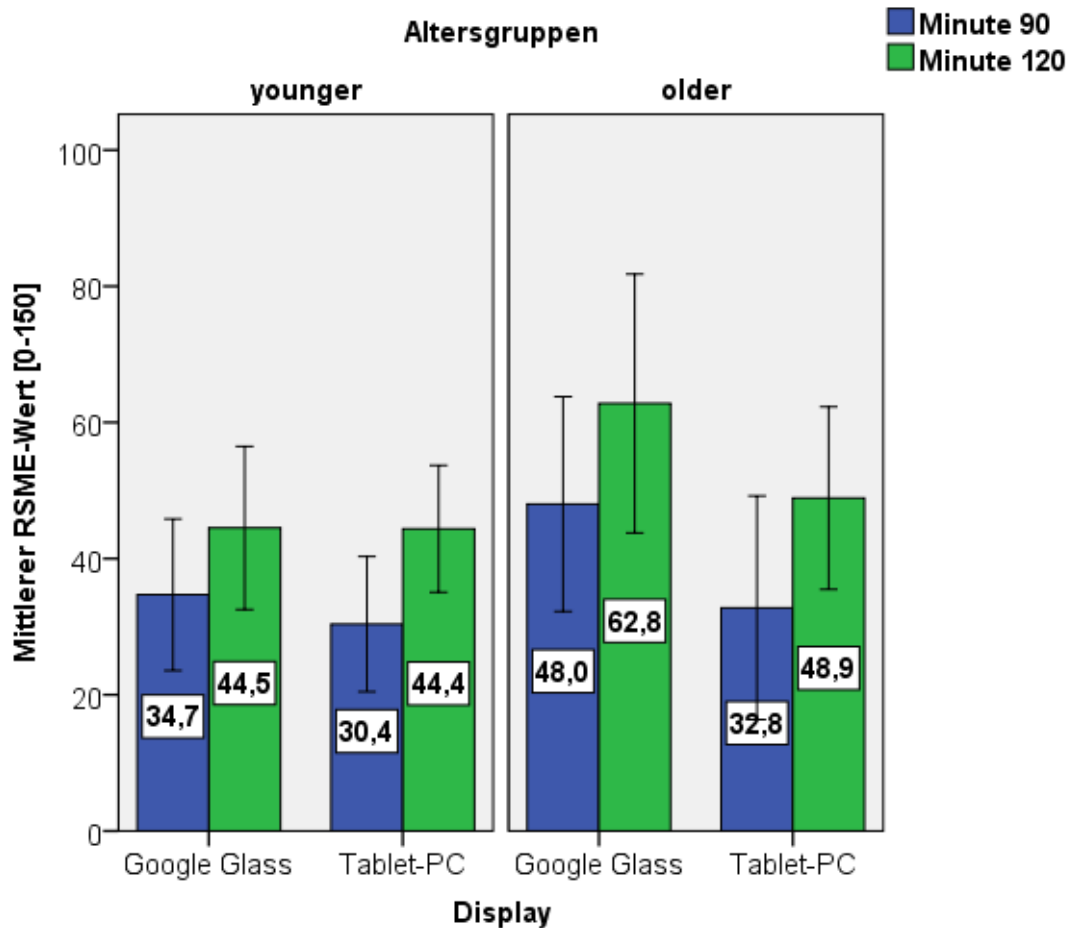


Abb. 4.12 Mittlere Werte der RSME-Skala, unterteilt nach Altersgruppen, Displaytyp und Zeitpunkt (zu Beginn der Studie nach 90 Minuten und gegen Ende nach 120 Minuten Gesamtversuchszeit). Fehlerbalken zeigen das 95%-Konfidenzintervall.

4.3.4 Visuelle Ermüdung

Die Auswertungen in diesem Abschnitt geben immer zwei Messzeitpunkte wieder, die im Abstand von einer Stunde liegen. Messzeitpunkt „1“ ist nach der ersten Stunde und Messzeitpunkt „2“ nach der zweiten Stunde. Wie in Abschnitt 4.2.4 dargestellt wurde in der ersten Stunde aber ein anderer Teilversuch der Studienreihe durchgeführt, der hier nicht ausgewertet wird. In der ersten Hälfte der zweiten Stunde wurde ebenfalls eine andere Teilstudie durchgeführt (Studie zur Fehlersuche) und erst in der letzten halben Stunde die Replikation mit Lego-Technik. Da aber in der gesamten

zweiten Stunde die Unterteilung der Probanden in HMD- und Tablet-PC-Nutzer schon stattfand, wird hier die einstündige Arbeit mit dem Displaytyp bewertet.

Von den einzelnen Items des VFQ werden hier nur relevante Items gezeigt. Bei den Items „tränennde Augen“, „schmerzende Augen“, „springender Text“, „Blendung durch Lichter“, „Verschwommene Buchstaben“, „Doppelsehen“ und „Übelkeit“ zeigten sich keine Effekte und die Mittelwerte blieben meist unter einem Wert von Eins, so dass hier auf eine Ergebnisdarstellung aus Gründen der Übersicht verzichtet wird. Allgemein lässt sich feststellen, dass wie auch in der Originaluntersuchung die Mittelwerte der Items auch bei statistisch relevanten Unterschieden meist im unteren Drittel der Skala verbleiben. Bei der Interpretation der Effekte der einzelnen Items sei noch einmal auf den fehlenden Summenscore dieses Fragebogens hingewiesen, der eine Korrektur des Alpha Risikos von 5% auf 0,3% erforderlich macht ($p = .003$ statt $p = .05$), wie in 2.4.2.7 ausführlich diskutiert.

Abbildung 4.13 zeigt die Entwicklung der Mittelwerte für die Items „trockene Augen“ und „gereizte/brennende Augen“. Bei den trockenen Augen gab es eine deutliche Steigerung über die Zeit [$F(1, 33) = 6.427$; $p = .016$], aber keinen Effekt des Displays [$F(1, 33) = 1.823$; $p = .186$] oder Wechselwirkungen zwischen den beiden Faktoren. Ebensovienig liegen signifikante Effekte oder Wechselwirkungen mit dem Alter vor.

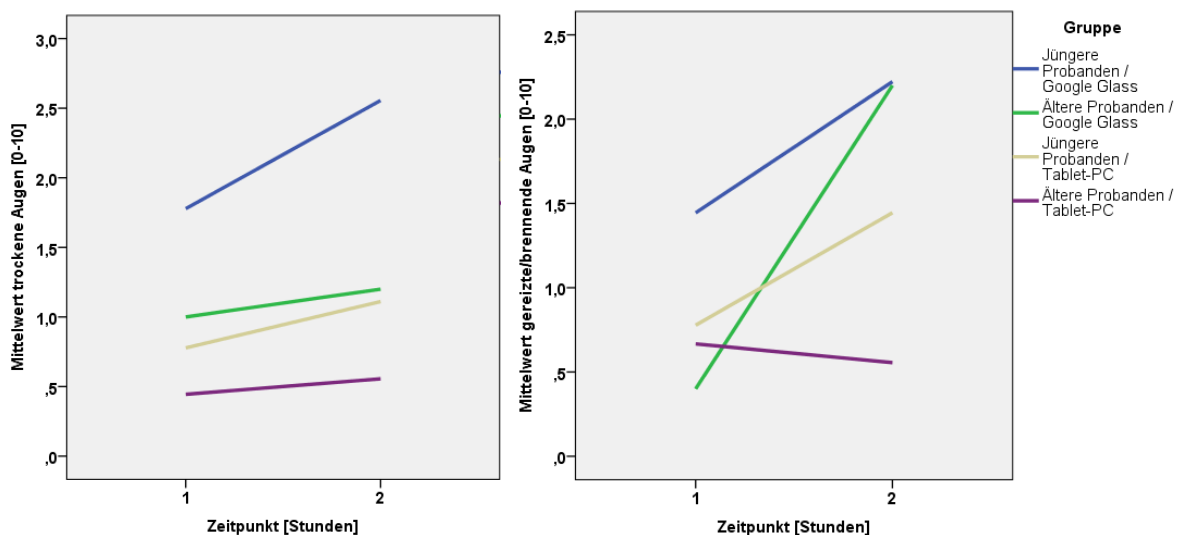


Abb. 4.13 Mittelwerte für die VFQ-Items „trockene Augen“ (links) und „gereizte/brennende Augen“ (rechts), unterteilt nach Display und Altersgruppe.

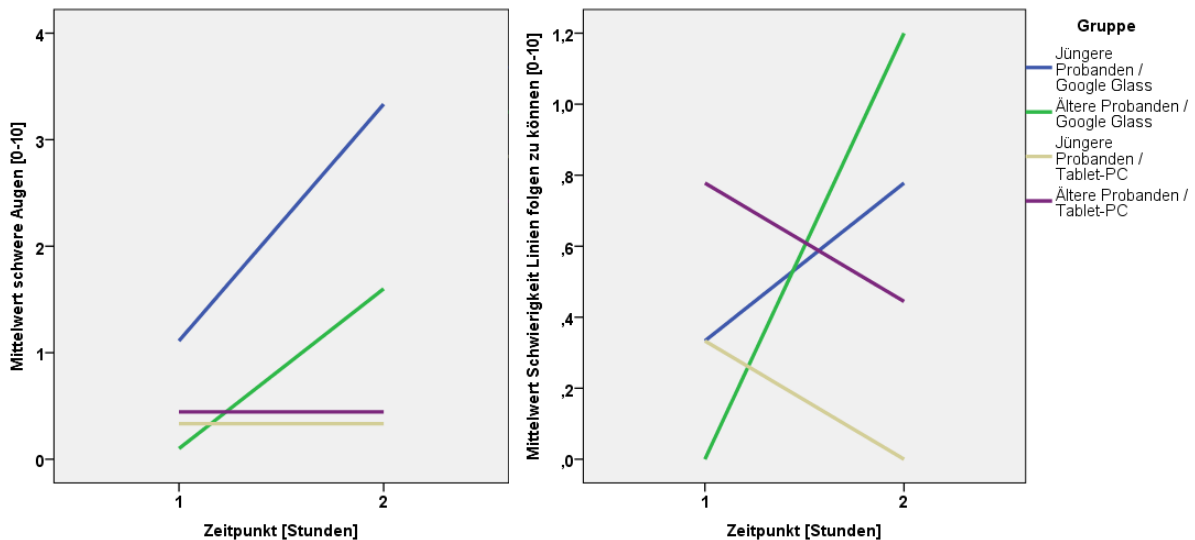


Abb. 4.14 Mittelwerte für die VFQ-Items „schwere Augen“ (links) und „Schwierigkeit Linien zu folgen“ (rechts), unterteilt nach Display und Altersgruppe.

Beim Item „gereizte/brennende Augen“ (Abb. 4.13, rechts) gab es ebenfalls eine allgemeine Steigerung über die Zeit [$F(1, 33) = 9.798$; $p = .004$] und keinen Effekt des Displays [$F(1, 33) = 1.114$; $p = .299$]. Die Wechselwirkung zwischen Zeit und Display erreicht einen Wert, der bei unkorrigiertem Alphaniveau einer statistischen Tendenz entspräche [$F(1, 33) = 3.445$; $p = .072$]. Alterseffekte oder Wechselwirkungen mit diesen liegen nicht vor.

Bei dem Item „schwere Augen“ (Abb. 4.14, links) liegt eine signifikante Steigerung über die Zeit vor [$F(1, 33) = 11.236$; $p = .002$] sowie im selben Ausmaß eine Wechselwirkung zwischen Zeit und Displaytyp [$F(1, 33) = 11.236$; $p = .002$]. Dieses gleiche Ausmaß lässt sich bereits in der Grafik erkennen, da die Werte der Tablet-PC-Nutzer bei beiden Zeitpunkten auf dem selben Niveau bleiben, während die Werte der HMD-Nutzer ansteigen. Ein Effekt des Displays liegt bei Betrachtung mit korrigiertem Alphaniveau nicht vor [$F(1, 33) = 7.235$; $p = .011$], wäre aber unter normalen Bedingungen signifikant.

Das Item „Schwierigkeit, Linien folgen zu können“ (Abb. 4.14, rechts) zeigt Variationen der Mittelwerte nur am unteren Ende der Skala zwischen 0 und 1,2. Trotzdem zeigt sich hier eine interessante Wechselwirkung zwischen Zeit und Displaytyp [$F(1, 33) = 8.422$; $p = .007$]: Während die Werte für beide Altersgruppen der Tablet-PC-Bedingung abfallen, steigen die Werte für beide Altersbedingungen des HMDs an. Ansonsten zeigen sich keine Haupteffekte oder Wechselwirkungen.

Das Item „Schwierigkeit, scharf zu sehen“ (Abb. 4.15, links) zeigt ebenfalls eine interessante Wechselwirkung zwischen Zeit und Display [$F(1, 33) = 10.751$; $p = .002$], die erneut einen Anstieg der Werte über die Zeit beim HMD und ein Abfallen der Werte beim Tablet-PC zeigt. Andere Effekte werden hier nicht signifikant. Das Gleiche gilt für das Item „verschwommene Sicht“ (Abb. 4.15, rechts). Auch hier zeigt sich eine Wechselwirkung zwischen Zeit und Display [$F(1, 31) = 7.674$; $p = .009$], bei der die Werte für das HMD wieder ansteigen und die des Tablet-PCs wieder abfallen. Weitere Effekte können auch hier nicht aufgezeigt werden.

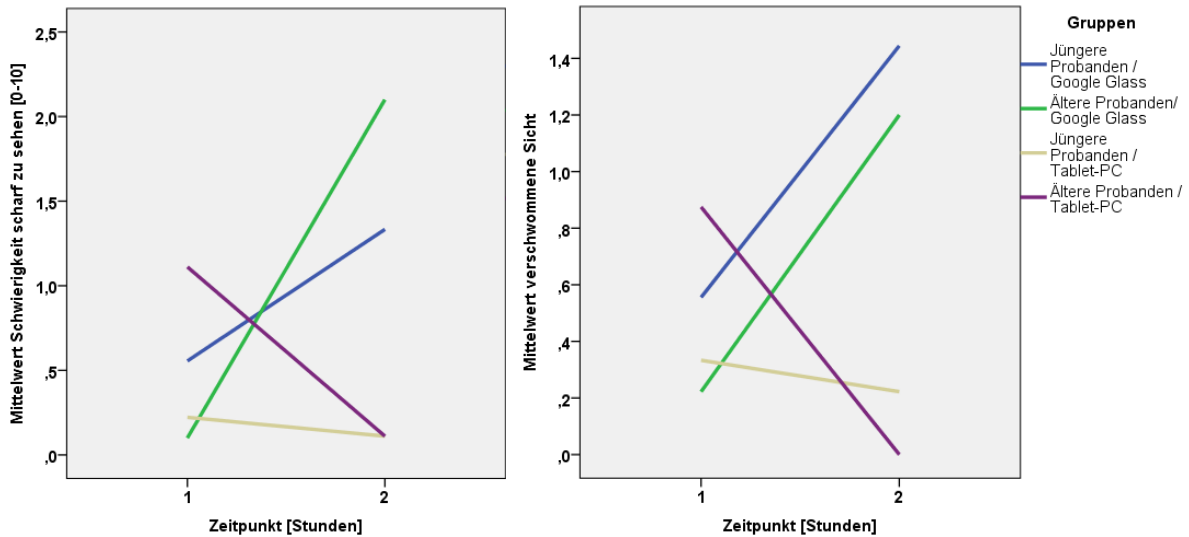


Abb. 4.15 Mittelwerte für die VFQ-Items „Schwierigkeiten, scharf zu sehen“ (links) und „verschwommene Sicht“ (rechts), unterteilt nach Display und Altersgruppe.

Die Abbildung 4.16 zeigt auf der linken Seite das Item „Kopfschmerzen“ und auf der rechten Seite das Item „Nackenschmerzen“. In beiden Fällen zeigt sich zwar ein deutlicher Anstieg über die Zeit ($[F(1, 31) = 6.634 ; p = .015]$ für Kopfschmerzen und $[F(1, 31) = 8.508 ; p = .007]$ für Nackenschmerzen), aber es gibt keine Wechselwirkung mit dem Displaytyp oder einen Haupteffekt des Displaytypes wie es in der Originalstudie war. Dies repräsentiert den größeren Tragekomfort von Google Glass im Vergleich zum MAVUS-HMD. Effekte des Alters liegen erneut nicht vor, auch wenn in beiden Fällen die jüngeren Probanden unabhängig vom verwendeten Display höhere Rankings zu diesen Items abgegeben haben.

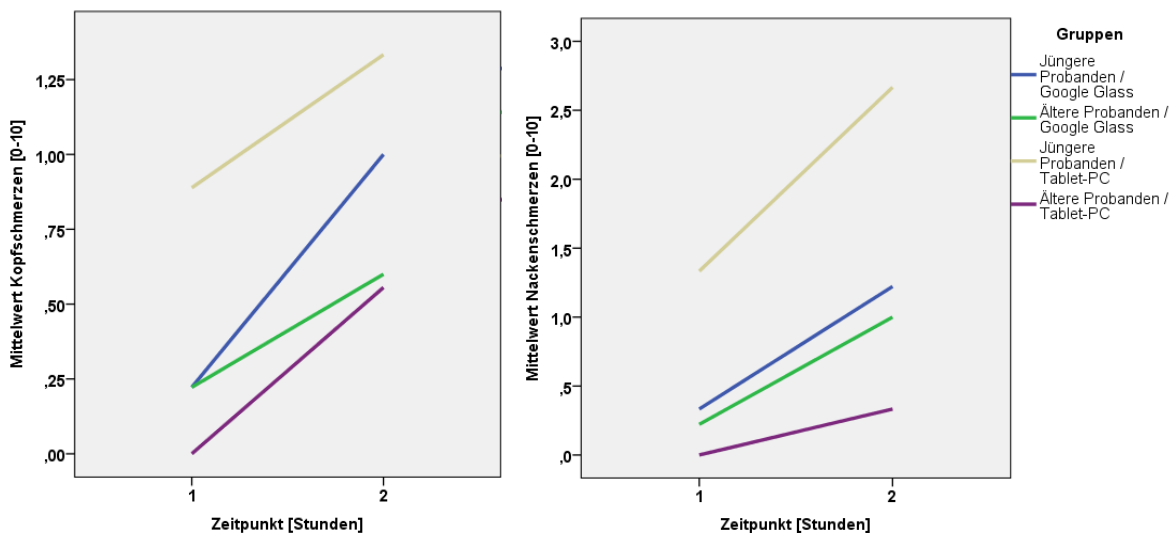


Abb. 4.16 Mittelwerte für die VFQ-Items „Kopfschmerzen“ (links) und „Nackenschmerzen“ (rechts), unterteilt nach Display und Altersgruppe.

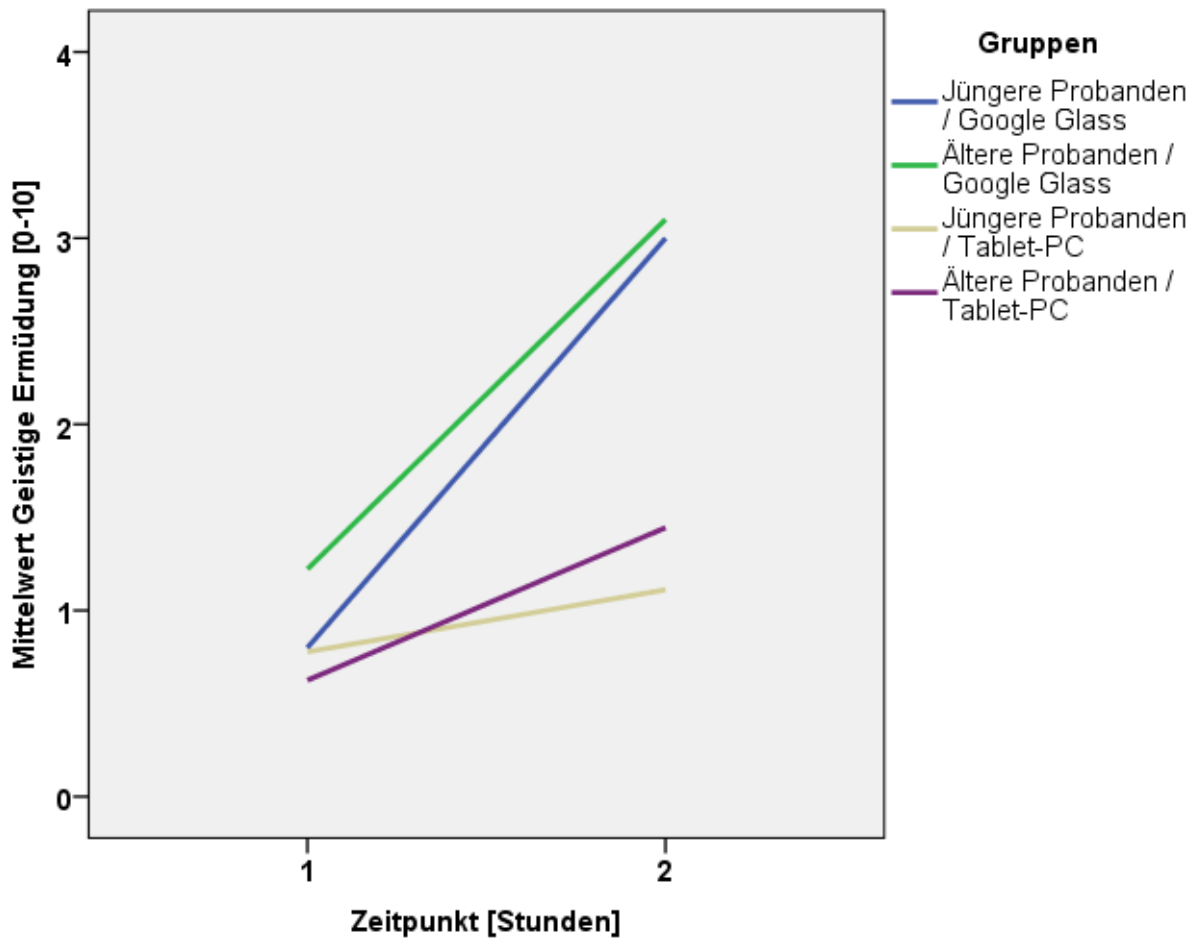


Abb. 4.17 Mittelwerte für das VFQ-Item „Geistige Ermüdung“, unterteilt nach Display und Altersgruppe.

Abbildung 4.17 zeigt die „geistige Ermüdung“. Hierbei zeigt sich ein hochsignifikanter Effekt des Zeitpunktes [$F(1, 31) = 20.871$; $p < .001$] der mit höheren Werten zu späteren Zeitpunkten einhergeht. Eine Wechselwirkung zwischen Zeit und Displaytyp wäre bei unkorrigiertem Alphaniveau ebenfalls signifikant [$F(1, 31) = 4.312$; $p = .046$]. In der Grafik ist zu sehen, dass alle Probandengruppen bei Stunde eins auf ungefähr dem selben Niveau sind. Während die Tablet-PC-Nutzer nur einen leichten Anstieg zu Stunde zwei verzeichnen, fällt dieser bei den HMD-Nutzern unabhängig von den Altersgruppen deutlich steiler aus. Alterseffekte liegen hier ebensowenig wie ein Haupteffekt des Displays vor.

4.3.5 Interview

Nach der Versuchsreihe wurde mit den Probanden wieder ein Interview durchgeführt (vgl. Anhang 2). Hierbei zeigte sich, dass 50 Prozent der Probanden die Umsetzung der Arbeitsaufgaben auf Google Glass mochten. 25 Prozent der Probanden hätten es bevorzugt, die Legoanleitung auf dem Papier zu haben, und die verbliebenen 25 Prozent mochten die Arbeitsaufgabe an sich nicht. 40 Prozent der Befragten fanden die Displayposition von Google Glass zu hoch für diese Aufgabe. Das Display sitzt bei Google Glass sowohl sehr seitlich als auch recht weit oben im Blickfeld. Dies ist wahrscheinlich dem eigentlichen Zweck dieses HMDs geschuldet, das seinem Nut-

zer, während er läuft, kurze Informationen einspielen soll. Wäre das Display im zentralen Sichtbereich, würde es eventuelle Hindernisse auf dem Weg verdecken und somit der ungehinderten Bewegung im Wege stehen. Für die hier vorliegende Arbeitsaufgabe, die kontinuierliche Informationsaufnahme verlangt, ist die Position des Displays aber zu peripher, um anstengungsfrei Informationen zu entnehmen.

85 Prozent der Probanden mochten die Zoom-Funktion. Dabei ist erstaunlich, dass diese Funktion allgemein nur sehr selten genutzt wurde – im Schnitt gab es gerade vier Zoomvorgänge während des gesamten Versuches pro Proband. Wobei einige Probanden den Zoom gar nicht genutzt haben.

90 Prozent der Probanden waren mit dem Tragekomfort von Google Glass zufrieden. Kritik gab es in einigen Fällen über die Wärme, die das Gerät abstrahlt und die bei längerer Anzeige- und Tragezeit unangenehm werden kann. Auch hierbei ist zu betonen, dass die Wärme hauptsächlich durch die zwei Stunden andauernde Informationsanzeige zustande kam. Im für das Gerät ursprünglich gedachten Anwendungskontext sind Informationen immer nur wenige Sekunden im Display einsehbar, danach schaltet sich das Gerät automatisch ab. Bei so einer Nutzung entsteht keine nennenswerte Wärme und der Akkumulator des Gerätes hält auch länger.

Trotz dieser kleinen Einschränkungen bevorzugten alle Probanden Google Glass gegenüber dem MAVUS-HMD. Diese Frage wurde an alle Probanden gestellt, die bereits an der Originalstudie teilgenommen hatten, was ungefähr 80 Prozent der Probanden ausmacht. Auch die Tablet-PC Nutzer wurden dies gefragt, da sie ja in der ersten Stunde der Studienreihe auch mit Google Glass gearbeitet hatten.

4.4 Vergleich der Befundlage beider HMDs

Ein Vergleich der Befundlage beider HMDs kann Hinweise darauf geben, welche Effekte eher durch das Konzept der monokularen HMDs an sich verursacht sind und welche Effekte eher durch die jeweilige Hardwarerealisation des HMD-Konzeptes bedingt sind.

Die beiden HMDs – das industrielle MAVUS-HMD und Google Glass, das aus dem Konsumentenbereich stammt – unterscheiden sich in einer Menge von Faktoren (vgl. Tab. 4.1), so z. B. im Gewicht des Kopfträgers oder in der Positionierung des Displays. Wenn nun Effekte in der Originalstudie gefundenen wurden, die sich in der Replikation nicht mehr nachweisen lassen, so ist dies ein Hinweis darauf, dass die Effekte auf der Hardwareumsetzung des industriellen HMDs beruhen. Zeigen sich aber vergleichbare Effekte, so sind diese eher dem Prinzip des monokularen HMDs an sich zuzuordnen. Ein solcher Vergleich ist möglich geworden, da hier der seltene Fall vorliegt, dass die gleichen Versuchsaufgaben auf unterschiedlichen HMDs realisiert wurden. Allerdings ist bei so einem Vergleich auch Vorsicht geboten, da sich die HMDs in vielen Faktoren unterscheiden und es somit schwierig ist, eventuelle Unterschiede in Leistung oder Beanspruchung auf einen einzelnen dieser Faktoren zurückzuführen.

Des Weiteren muss bei einem Vergleich der Befundlage beider HMDs – und somit beider Studien - auch bedacht werden, dass das Versuchsdesign der beiden Studien ein paar wesentliche Unterschiede aufweist:



- Die Ausführungszeit der Ausgaben war in der Originalstudie viel länger (vier Stunden) als in der Replikation (halbe Stunde). Diese unterschiedlichen Zeiten führen dazu, dass ein Vergleich streng genommen nur über die erste halbe

Stunde gemacht werden kann, da von Google Glass keine längere Ausführungszeit der Versuchsaufgaben vorliegt. Dies würde dann aber dazu führen, dass das Alleinstellungsmerkmal der Originalstudie – der Beanspruchungsverlauf über mehrere Stunden – nicht beachtet werden kann.

- Die Reihenfolge der Studien war nicht variiert. Dies ist bei einer Replikationsstudie zwar durchaus normal, da aber 70% der Probanden bereits an der Originalstudie teilgenommen hatten, liegt bei diesen Probanden eine Vorerfahrung sowohl mit der Technologie als auch mit der Aufgabe vor, die die Ergebnisse beeinflussen kann.
- Die Originalstudie wurde im „within-subject Design“ durchgeführt, so dass jeder Proband an drei Sessions mit jeweils vier Stunden teilnahm, während bei der Replikation ein „between-subject Design“ vorliegt, bei dem die Probanden an nur einer Session teilnahmen und entweder mit dem HMD oder dem Tablet-PC arbeiteten.

Diese Unterschiede wirken sich mit großer Wahrscheinlichkeit neben der Erfahrung auch auf die Motivation der Probanden aus. Hiervon können auch die Antworttendenzen der Probanden bei subjektiven Parametern - wie sie in Fragebögen vorliegen – betroffen sein.

Tab. 4.1 Unterschiede im Aufbau der beiden verwendeten HMDs (Auszüge)

	MAVUS-HMD	Google Glass
Bild		
Aufbautyp	„Look around“	„See through“
Auflösung	800 x 600 Pixel	640 x 360 Pixel
Position des Displays	zentral	peripher
Auge	Links oder rechts (in der Studie immer das Führungsauge)	rechts
Gewicht des Kopfträgers	380 Gramm	50 Gramm

4.4.1 Vergleich der Leistung in der Konstruktionsaufgabe

In beiden Studien wurden signifikant weniger Konstruktionsfolien in der Versuchszeit mit dem HMD abgearbeitet als mit dem Tablet-PC. In der Originalstudie mit einer Laufzeit von vier Stunden waren es 22,5% weniger Folien (vgl. 3.1 und siehe Abb. 3.1). In der Replikationsstudie waren es 33,9% weniger Folien mit dem HMD (vgl. 4.3.1 und siehe Abb. 4.6). Dabei wurde für diesen Vergleich in der Originalstudie der zweite Durchgang mit dem HMD herangezogen, bei dem die Probanden schon Erfahrung mit der HMD-Technologie durch den ersten Durchgang hatten. Und in der Replikationsstudie waren die meisten Probanden durch vorherige Teilnahme an der Originalstudie sowie den anderen Teilstudien in den ersten eineinhalb Stunden der Versuchsreihe an die Technologie gewöhnt. Somit ist als Ursache für diesen Effekt eine mangelnde Gewöhnung an die neue Technologie unwahrscheinlich. Vielmehr lässt sich vermuten, dass durch die monokulare Informationsaufnahme beim HMD Leistungseinbußen im Vergleich zur binokularen Informationsaufnahme beim Tablet-PC entstehen. Die Tatsache, dass dieser Effekt bei Google Glass im Ausmaß größer ist, könnte zufallsbedingt sein. Er könnte aber auch auf dem Sachverhalt beruhen, dass die Displayposition bei Google Glass wesentlich peripherer ist als beim MAVUS-HMD. Ein Hinweis darauf findet sich auch in den Interviews, in denen einige Probanden die Position des Google Glass-Displays explizit kritisierten. Es ist aber ebenso denkbar, dass das größere Ausmaß der schlechteren Performanz bei Google Glass auf der Tatsache beruht, dass es sich um ein „see through“ HMD handelt, während das MAVUS-HMD ein „look around“ Display hat. Die Probanden könnten bei der Informationsaufnahme durch den halb durchscheinenden Hintergrund irritiert sein. Wobei zu erwähnen ist, dass die Probanden extra vor eine weiße Wand gesetzt wurden und Blend- und Streulicht vermieden wurde, um solchen Effekten entgegenzuarbeiten.

Lässt man aber die Interpretation des Ausmaßes des Effektes beiseite, zeigt sich, dass mit beiden HMDs trotz bereits vorhandener Gewöhnung an die HMD-Technologie langsamer in der Konstruktionsaufgabe gearbeitet wird. Dieser Effekt ist also klar der Technologie der monokularen HMDs an sich und nicht der jeweiligen Hardwareumsetzung zuzuschreiben.

4.4.2 Vergleich der Leistung in der Monitoringaufgabe

Bei der Güte der Reaktion auf eine Monitoringaufgabe zeigen sich in beiden Studien schlechtere Werte für das HMD verglichen mit dem Tablet-PC. Dies war im Vorfeld des Projektes so nicht erwartet worden. Da Anzeigen auf dem HMD immer im Sichtfeld des Nutzers sind, war im Gegenteil erwartet worden, dass auf eine Anzeigenveränderung im HMD besser reagiert würde als auf eine Anzeigenänderung auf einem Tablet-PC, der nicht zwangsläufig immer im Blick des Nutzers ist. Diesem Befund kommt eine gewisse Bedeutung zu, da einige Forschungsprojekte für sicherheitskritische Anzeigen explizit ein HMD verwenden wollen, damit die Botschaft vom Nutzer schnell wahrgenommen wird. Zwar bezieht sich der Befund nur auf das parallele Monitoring bei gleichzeitiger Informationsaufnahme und sagt nichts über erscheinende Warnungen auf einem sonst abgeschalteten Schirm aus, die Befundlage reicht aber, um auch solche Hinweise zu überprüfen.

Der Effekt der schlechteren Reaktion auf die parallele Monitoringaufgabe tritt in der Originalstudie nur als Tendenz auf (vgl. 3.2 und siehe Abb. 3.2) wird aber in der Rep-

likation signifikant (vgl. 4.3.2.1 und siehe Abb. 4.7 und Abb. 4.8). Der Grund für den stärkeren, weil nun signifikanten Unterschied in der Replikationsstudie könnte erneut die größere Peripherie des Displays bei Google Glass sein. Dies wird sich bei den Monitoringaufgaben noch stärker als bei den Konstruktionsaufgaben auswirken, weil die Monitoringaufgaben immer am äußeren Rand dargestellt wurden.

Da beide Studien Effekte in die gleiche Richtung zeigen, ist davon auszugehen, dass die schlechtere Reaktion auf eine parallele Monitoringaufgabe ebenfalls eher dem Konzept des monokularen HMDs an sich als der jeweiligen Hardwareumsetzung zuzuschreiben ist. Das Ausmaß des Effektes scheint aber mit der Positionierung des Displays in der jeweiligen Hardwareumsetzung zu variieren und es ist anzunehmen, dass die Leistung umso schlechter wird, je peripherer das Display angebracht ist. Auch hier könnte zusätzlich der „see-through“ Faktor eine Rolle spielen.

4.4.3 Vergleich der subjektiven Beanspruchungsparameter

Bei der subjektiven Beanspruchung zeigen beide Studien höhere Werte für die HMDs verglichen mit den Tablet-PCs. In der Originalstudie sind diese Unterschiede sowohl beim NASA-TLX als auch beim RSME hochsignifikant (vgl. 3.4.1 und 3.4.2). In der Replikationsstudie ist der Unterschied beim NASA-TLX ebenfalls signifikant (vgl. 4.3.3.1), beim RSME kann aber kein signifikanter Unterschied nachgewiesen werden (vgl. 4.3.3.2). Dies könnte nach Meinung des Autors aber auch auf das verwendete between-subject Design in dieser Studie zurückzuführen sein: Der RSME ist in hohem Maße auch von den individuellen Antworttendenzen der Probanden bestimmt. Bei einem between-subject Design mit kleinen Zellengrößen kann es gut sein, dass die zufällige Verteilung individueller Antworttendenzen eine versuchsbedingte Variation überstimmt. Misst man auf der anderen Seite auch der Tatsache Bedeutung zu, dass der Effekt beim NASA-TLX in der Replikationsstudie zwar signifikant, aber eben nicht hochsignifikant wie in der Originalstudie ist, so könnte dies in Zusammenhang mit den nichtsignifikanten Unterschieden beim RSME auch interpretiert werden als geringerer Unterschied in der mentalen Beanspruchung bei Google Glass im Vergleich zum Tablet-PC als beim MAVUS-HMD. Dies könnte gut erklärbar über den besseren Tragekomfort von Google Glass sein, der auch in den Interviews bestätigt wird. Auf der anderen Seite könnte dies auch in der kürzeren Tragezeit des HMDs in der Replikationsstudie begründet sein.

Da beide Studien teilweise hochsignifikant höhere subjektive Beanspruchung beim Arbeiten mit einem HMD aufzeigen, ist dies wohl auch eher als ein genereller Effekt von monokularen HMDs anzusehen. Da der Effekt aber weniger deutlich ausgeprägt ist bei Google Glass, ist ebenso davon auszugehen, dass die jeweilige Hardwareumsetzung ebenfalls Einfluss auf das subjektive Beanspruchungsempfinden hat. Hier ist zu vermuten, dass das geringere Gewicht und der fehlende Kopfträger bei Google Glass Vorteile im Tragekomfort bringen, die sich positiv auf die empfundene Beanspruchung auswirken.

(Die objektiven Beanspruchungsparameter (vgl. 3.5) liegen nur für das MAVUS-HMD vor und zeigten nach einer Gewöhnung ein vergleichbares Beanspruchungsniveau zwischen HMD und Tablet-PC.)

4.4.4 Vergleich des Fragebogens zur visuellen Ermüdung

Die Ausprägung der Mittelwerte für die 16 Items des VFQ bleiben in beiden Studien im unteren Drittel der Skala (vgl. 3.7 und 4.3.4). Die Werte einzelner Probanden sind

aber durchaus höher. Die Mittelwerte für Items wie „trockene Augen“ oder „gereizte/brennende Augen“ unterscheiden sich nicht grundlegend bei den beiden Studien. Ebenso für das Item „Schwierigkeit, scharf zu sehen“. Die Werte für „schwere Augen“ sind nach zwei Stunden Tragezeit bei Google Glass sogar etwas höher als beim MAVUS-HMD. Das Gleiche trifft zu für das Item „geistige Ermüdung“, das bei Google Glass nach zwei Stunden ungefähr einen Skalenpunkt höher ist als beim MAVUS-HMD. Die gefundenen signifikanten Unterschiede zum jeweiligen Tablet-PC sind aber vergleichbar.

Einen großen Unterschied gibt es im Gegensatz dazu bei den beiden Items „Kopf- und Nackenschmerzen“. Diese kommen mit Google Glass so gut wie gar nicht mehr vor und es gibt hier keine signifikanten Unterschiede zum Tablet-PC, wie das noch in der Originalstudie mit MAVUS-HMD der Fall war. Dies ist im Anwendungsfall nicht zu unterschätzen. Ein Anzeigegerät darf keine Kopf- und Nackenschmerzen verursachen. Es ist davon auszugehen, dass das hohe Gewicht des MAVUS-HMD (380 Gramm) in Kombination mit dem Kopfträger, der eng um die Stirn befestigt werden muss, um das Gewicht zu halten, was in manchen Fällen zu Druckkopfschmerzen führte, dafür verantwortlich sind.

Bezüglich allgemeiner Parameter der visuellen Ermüdung lassen sich also keine wesentlichen Unterschiede bei den beiden HMD-Typen feststellen. Kopf- und Nackenschmerzen aber sind stark von der Hardwareumsetzung beeinflusst und kommen beim leichteren und komfortableren Google Glass nicht mehr vor.

4.4.5 Fazit

Allgemein lässt sich feststellen, dass die meisten hier berichteten Effekte unabhängig von der Hardwareumsetzung sind und allgemein auf monokulare HMDs zuzutreffen scheinen. Die langsamere Performanz in der Konstruktionsaufgabe und auch die schlechtere Leistung in der Monitoringaufgabe scheinen eher das Resultat des Konzeptes an sich zu sein. Wobei die periphere Displayposition von Google Glass anscheinend zu noch schlechteren Werten führt. Die Beanspruchung ist hingegen bei Google Glass tendenziell geringer (wenn auch immer noch höher als beim Tablet-PC), was wahrscheinlich auf den höheren Tragekomfort zurückzuführen ist. Vor allem Kopf- und Nackenschmerzen kommen bei Google Glass nicht mehr vor. Man kann also sagen „Google Glass ist komfortabler, aber nicht besser“.

5 Diskussion

5.1 Interpretation der Ergebnisse

5.1.1 Vergleich zwischen HMD und Tablet-PC

Bei dem Vergleich der beiden Anzeigetechnologien wurde in der Originalstudie bewusst die zweite Messung mit dem HMD zugrunde gelegt, da die Probanden dann schon eine gewisse Eingewöhnungszeit an die neue Technologie absolviert hatten. Trotzdem zeigen sich deutliche Unterschiede in der erbrachten Leistung, der Beanspruchung und der visuellen Ermüdung, wobei das HMD immer schlechter abschneidet als der Tablet-PC:

- Bei der Leistung werden fast ein Viertel weniger Konstruktionsfolien mit dem HMD bearbeitet als mit dem Tablet-PC (vgl. 3.1 und 4.3.1).
- Bei den Monitoringaufgaben zeigt sich eine Tendenz, dass im Vergleich zum Tablet-PC mit dem HMD weniger akkurat und schnell reagiert wird und dass, obwohl die Anzeige beim HMD immer im Sichtfeld der Probanden war, während der Tablet-PC auf dem Tisch lag, so dass bei Blickabwendung eventuelle Veränderungen in der Monitoringaufgabe gar nicht mehr im Sichtfeld gewesen sind (vgl. 3.2 und 3.3 sowie 4.3.2).
- Bei der subjektiv empfundenen Beanspruchung (vgl. 3.4 und 4.3.3) führt das HMD bei beiden verwendeten Verfahren in der Originalstudie zu signifikant höheren Werten als der Tablet-PC. Die objektive Beanspruchung zeigt hingegen einen Gewöhnungseffekt an das HMD und danach keine signifikanten Unterschiede zwischen HMD und Tablet-PC (vgl. 3.5).
- Bei der visuellen Ermüdung zeigen sich beim HMD höhere Werte (vgl. 3.7 und 4.3.4), und hier tritt zusätzlich auch öfter eine Wechselwirkung ein, die wiedergibt, dass die visuelle Ermüdung nicht nur auf einem höheren Niveau beim HMD ist, sondern auch die Zunahme über die Zeit stärker ist als beim Tablet-PC.
- Im Interview der Originalstudie (vgl. 3.8) zieht nur ein Proband von 41 das HMD dem Tablet-PC vor, und es werden die Trageeigenschaften des Kopfträgers, sowie die Schärfe der Darstellung auf dem Display von vielen Probanden kritisiert. Bei der Replikation mit Google Glass bevorzugen alle Probanden dieses im Vergleich zum MAVUS-HMD (vgl. 4.3.5).

Die Ergebnislage spricht also zunächst deutlich gegen die HMD-Technologie. Allerdings gibt es ein paar Faktoren, die bei der Interpretation der Ergebnisse ebenfalls bedacht werden müssen:

- Die HMD-Technik ist für die Probanden neu. Niemand hatte vorher mit einem HMD gearbeitet. Und selbst wenn erst die zweite Session als Vergleich hier herangezogen wurde, und somit eine gewisse Vorerfahrung da war, ist die HMD-Technologie wahrscheinlich immer noch neu und ungewohnt im Vergleich zu einem kleinen Monitor auf dem Tisch.
- Viele Faktoren, wie Schärfe der Darstellung und Trageeigenschaften des Kopfträgers (und deren wechselseitige Abhängigkeit) sind Teil der konkret

verwendeten Hardware und nicht des HMD-Konzeptes an sich. Da Displaytypen, Schärfereinstellmöglichkeiten und Kopfträger bei verschiedenen HMDs stets unterschiedlich sind, können die Ergebnisse dieser Studie nur bedingt auf andere HMDs übertragen werden.

- Die in der Laborstudie verwendete Aufgabe bringt zwei der Hauptvorteile des HMDs – die Mobilität und das Freibleiben beider Hände – nicht zum Tragen. Denn hier wurde sitzend an einem Tisch gearbeitet. In einer solchen Situation würde man intuitiv kein HMD verwenden. Der Grund dafür, dass im Versuchsssetup die Mobilität nicht mehr eingebaut wurde, liegt daran, dass gleichzeitig die Herzrate als objektiver Beanspruchungsparameter erhoben wurde. Wäre hierbei Bewegung Teil des Versuchsssetups, wäre eine Auswertung der Herzrate als Beanspruchungsparameter nicht möglich gewesen, da die eher geringen Veränderungen in der Herzrhythmik, die auf Beanspruchung basieren, dann von Bewegungsartefakten überlagert worden wären. Die so entstandene leicht unnatürliche Versuchssituation kann sich aber auf die subjektiven Einschätzungen der Probanden auswirken und zu schlechteren Resultaten führen.

Für das untersuchte HMD und die untersuchte Tätigkeit kann aber festgehalten werden, dass ein Tablet-PC nicht nur von der überwiegenden Mehrheit der Probanden präferiert wird, sondern allgemein auch mit besserer Leistung und geringerer Beanspruchung und visueller Ermüdung einhergeht.

5.1.2 Gewöhnung an die HMD-Technologie

Um die Gewöhnung an die HMD-Technologie zu untersuchen, wurden die ersten vier Stunden Versuchssession mit HMD mit der zweiten Versuchssession mit HMD verglichen. Dieser Vergleich gibt also nur eine erste Gewöhnung wieder, und es kann vermutet werden, dass die Gewöhnung danach noch weiter geht und sich wahrscheinlich über Wochen oder Monate erstreckt.

- Bei der Leistung zeigt sich kein Effekt der Gewöhnung bei der Anzahl der abgearbeiteten Montagefolien oder bei der Reaktionszeit auf die Monitoringaufgaben. Allerdings zeigt sich ein Gewöhnungseffekt bei der Trefferquote in den Monitoringaufgaben: Im zweiten Durchgang wurde akkurater reagiert. Dies kann natürlich neben einem Effekt der Gewöhnung an die Technologie auch eine Gewöhnung an die im Versuch geforderten parallel auszuführenden Aufgaben sein.
- Bei der subjektiven Beanspruchung zeigt sich bei beiden Messverfahren eine statistische Tendenz zu geringeren Beanspruchungswerten bei der zweiten Erhebung. Dies kann als Hinweis darauf gedeutet werden, dass die empfundene Beanspruchung beim wiederholten Umgang mit der HMD-Technologie abnimmt. Dies ist vor allem vor dem Hintergrund zu sehen, dass die empfundene Beanspruchung beim HMD sonst wesentlich höher liegt als beim Tablet-PC.
- Die objektiven Beanspruchungsparameter zeigen einen starken Gewöhnungseffekt an das HMD und sind beim zweiten Durchgang signifikant geringer. Dies unterstützt die gefundene Tendenz bei den subjektiven Beanspruchungswerten und deutet auf eine deutliche Gewöhnung an die Technologie schon innerhalb der ersten Stunden hin.

- Eine weitere Verringerung der Beanspruchung bei noch längerer Erfahrung mit einem HMD ist zu vermuten, kann aber mit dem vorliegenden Versuch nicht nachgewiesen werden. Diese Verringerung muss auch nicht zwangsläufig darauf hinaus führen, dass die subjektive Beanspruchung beim HMD und Tablet-PC bei genügend Eingewöhnungszeit gleich ist. Denn zum einen würde eine Gewöhnung an den Tablet-PC hier vermutlich auch auf Dauer zu geringeren Beanspruchungswerten führen und zum anderen ist nicht klar, ob das Basisniveau, das sich bei genügend Übung nicht mehr verringert, bei beiden Technologien gleich ist. Es zeigt aber an, dass beim Einsatz von HMDs mit einer starken Gewöhnungsphase, vor allem am Anfang zu rechnen ist.
- Auch einige Items des Fragebogens zur visuellen Ermüdung zeigen Tendenzen oder Signifikanzen zur Verringerung der Werte beim zweiten HMD-Einsatz. So z. B. bei den Items „trockene Augen“, „Schwierigkeit, Linien folgen zu können“, „Schwierigkeit, scharf zu sehen“ und „geistige Ermüdung“. Hier könnte eine bessere Passung des Kopfträgers vorliegen, basierend auf der Erfahrung der Probanden und der damit verbundenen Rückmeldung an den Versuchsleiter, beim Einstellen des Kopfträgers. Zu erwähnen bleibt aber, dass die Kopf- und Nackenschmerzen keinen Effekt der Gewöhnung zeigen.

5.2 Resultierende Handlungsempfehlungen

An dieser Stelle werden erste Handlungsempfehlungen für den HMD-Einsatz aus den Ergebnissen der Studie abgeleitet. Die Auflistung findet getrennt nach Problemfeldern statt.

Zur Aufgabe:

- HMDs sollten nur eingesetzt werden, wenn die Aufgabe es erfordert. Also nur dann, wenn andere Möglichkeiten der Informationsdarstellung nicht praktikabel sind und die Anforderung an Mobilität und freie Hände für manuelle Tätigkeit gegeben ist.
- Obwohl die Anzeige des HMDs baubedingt immer im Sichtfeld ist, kann nicht davon ausgegangen werden, dass das gleichzeitige, parallele Beachten einer Anzeige – z. B. zur Überwachung eines Maschinenstatus – auf dem HMD besser funktioniert als auf einer anderen Anzeige, die nicht immer im Sichtfeld ist. Es zeigten sich im vorliegenden Versuch sogar tendenziell schlechtere Reaktionen in der parallelen Monitoringaufgabe beim HMD.

Zum HMD-Aufbau:

- Der Kopfträger des HMDs ist nicht nur für den Tragekomfort wesentlich, sondern auch für die damit verbundene Beanspruchung, gerade im Langzeiteinsatz. Er sollte möglichst leicht sein und das Gewicht gut verteilt sein. Er muss bequem sein, aber trotzdem fest, um ein Verrutschen zu vermeiden. Hierzu sollte er über gute Einstellmöglichkeiten an die Kopfform des Nutzers verfügen.
- Der Fokus des Displays (das „Scharfstellen“), sollte nicht wie bei dem hier verwendeten HMD einzig von der Positionierung abhängig sein, sondern im gewissen Maße unabhängig über einen Regler einstellbar sein.

- Die Möglichkeit zum Tragen von Brillen ist bei vielen HMDs eingeschränkt. Gleitsichtbrillen stellen ein allgemeines Problem dar, da deren unterschiedliche Brennweite nicht mit dem HMD in Übereinstimmung ist. Aber auch normale Brillen können ein Problem darstellen, vor allem wenn sie baubedingt nicht bequem zwischen Nutzer und HMD-Aufsatz passen. Dies hängt in erster Linie vom Gestell ab. Um möglichst vielen Brillenträgern die Nutzung des HMDs zu ermöglichen, sollte der Abstand der Displayeinheit vom Gesicht genügend Raum bieten. Kontaktlinsen stellten im Versuchsverlauf kein Problem dar.

Zum Umgang mit der HMD-Technologie:

- Der Mitarbeiter muss beim Erstkontakt ausführlich in der Technologie unterwiesen werden. Hier sind vor allem alle Einstellmöglichkeiten mit dem HMD von einem in der Technologie erfahrenen Mitarbeiter zu demonstrieren. Denn eine schlechte Einstellung, die am Anfang vielleicht noch akzeptabel erscheint, führt langfristig zu höherer Beanspruchung.
- Da die Beanspruchung beim HMD-Einsatz zunächst deutlich höher ist als bei anderen Anzeigetechnologien, müssen die Pausen größer sein. Beim langfristigen Arbeiten mit einem HMD über mehrere Stunden empfiehlt es sich, nach regelmäßig kurze Pausen zu machen und das HMD hierbei abzusetzen. Neben der empfundenen Beanspruchung wird diese Empfehlung auch von den Ergebnissen der visuellen Ermüdung untermauert: Hier war das HMD nicht nur auf einem schlechteren Niveau, sondern der Anstieg über die Zeit war auch steiler als bei einem Tablet-PC. Deswegen sind regelmäßige Pausen beim Einsatz von HMDs dringend anzuraten.
- Im vorliegenden Versuch konnte an verschiedenen Stellen eine Gewöhnung an die HMD-Technologie nach der ersten Session festgestellt werden. Bei anderen Parametern zeigten sich zumindest Tendenzen zur Gewöhnung. Dies sollte bei der Einführung der Technologie beachtet werden. Es empfiehlt sich, die Arbeitszeiten mit dem HMD langsam zu steigern und nicht gleich einen mehrstündigen Einsatz zu fordern.
- Ältere Mitarbeiter sind in Bezug auf die visuelle Ermüdung beim HMD-Einsatz noch stärker betroffen als jüngere Mitarbeiter. Die Zeiten zur Gewöhnung und zu Pausen sind entsprechend anzupassen.

6 Literatur

6.1 Literaturangaben zu diesem Projektbericht

- Bangor, A. W. (2000). Display technology and ambient illumination influences on visual fatigue at VDT Workstations. Dissertation an der Virginia Polytechnic Institute and State University, USA. Online available: <http://scholar.lib.vt.edu/theses/available/etd-03072001-091123/unrestricted/vfatigue.pdf> [Januar 2015]
- Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 139-183). Amsterdam: North-Holland.
- Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009). Technikaffinität erfassen – Der Fragebogen TA-EG. In: A. Lichtenstein, C. Stöbel & C. Clemens (Eds.), Der Mensch als Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme (pp. 196-201). Düsseldorf, Germany: VDI Verlag GmbH. Online verfügbar unter: www.tu-berlin.de/fileadmin/f25/dokumente/8BWMMS/13.3-Karrer.pdf [August 2014]
- Mulder, L. J. M. (1988). *Assessment of cardiovascular reactivity by means of spectral analysis*. Unpublished dissertation. Rijksuniversiteit Groningen, Netherlands.
- Myrtek, M., Foerster, F., Brügger, G. (2001). Freiburger Monitoring System. Frankfurt am Main: Peter Lang.
- Wille, M., Grauel, B. & Adolph, L. (2011) Methodische Vorgehensweise zur Beanspruchungsermittlung bei Head-Mounted Displays. In: Reflexionen und Visionen der Mensch-Maschine-Interaktion – Aus der Vergangenheit lernen, Zukunft gestalten. 9. Berliner Werkstatt Mensch-Maschine-Systeme. (S. 84-85). Düsseldorf: VDI Verlag, Fortschrittsberichte, Band 33. ISBN 978-3-18-303322-5.
- Zijlstra, F.R.H. (1993). Efficiency in work behaviour: an approach for modern tools. PHD Thesis, Delft University of Technology. Online verfügbar unter: <http://resolver.tudelft.nl/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f> [Januar 2015].

6.2 Veröffentlichungen zu den beschriebenen Studien

- Wille, M. & Wischniewski, S. (2014) Influence of head mounted display hardware on performance and strain. In: D. de Waard, J. Sauer, S. Röttger, A. Kluge, D. Manzey, C. Weikert, A. Toffetti, R. Wiczorek, K. Brookhuis, and H. Hoonhout (Eds.) (2015). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2014 Annual Conference. ISSN 2333-4959 (online). Available from <http://hfes-europe.org>
- Wille, M., Scholl, P., Wischniewski, S. & Van Laerhoven, K. (2014). Comparing Google Glass with Tablet-PC as Guidance System for Assembling Tasks. Body Sensor Network, Zürich, accepted.
- Wille, M., Grauel, B. & Adolph, L. (2013). Strain caused by head mounted displays. In: De Waard, D., Brookhuis, K., Wiczorek, R., Di Nocera, F., Barham, P., Weikert, C., Kluge, A., Gerbino, W. & Toffetti, A., (Eds.) Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2013 Annual Conference. Online verfügbar unter: www.hfes-europe.org/wp-content/uploads/2014/06/Wille.pdf
- Wille, M., Grauel, B. & Adolph, L. (2013). Beanspruchung bei Head-Mounted Displays im Langzeiteinsatz. In: Brandenburg, E., Doria, L., Gross, A., Günzler, T., & Smieszek, H. (Eds.). (2014). Grundlagen und Anwendungen der Mensch-Maschine-Interaktion: 10. Berliner Werkstatt Mensch-Maschine-Systeme 10.-12. Oktober 2013; Proceedings. Universitätsverlag der TU Berlin, S. 608-618.
- Wille, M. (2013). Head-Mounted Displays - Beanspruchung im Langzeiteinsatz. In: Gesellschaft für Arbeitswissenschaften e.V., Bericht zum 59. Arbeitswissenschaftlichen Kongress. Dortmund: GfA-Press, S. 401-405.
- Wille, M., Grauel, B. & Adolph, L. (2012). Head-Mounted Displays - Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes. In: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (Eds.) Datenbrillen - Aktueller Stand von Forschung und Umsetzung sowie zukünftiger Entwicklungsrichtungen. Paderborn: Bonifatius GmbH, S. 6-11. ISBN 978-3-88261-146-5.
- Wille, M., Grauel, B. & Adolph, L. (2011). Methodische Vorgehensweise zur Beanspruchungsermittlung bei Head-Mounted Displays. In: Reflexionen und Visionen der Mensch-Maschine-Interaktion - Aus der Vergangenheit lernen, Zukunft gestalten. 9. Berliner Werkstatt Mensch-Maschine-Systeme. Düsseldorf: VDI Verlag, Fortschrittsberichte, Band 33, S. 84-85. ISBN 978-3-18-303322-5.
- Völker, K.; Adolph, L.; Pacharra, M.; Windel, A. (2010): Datenbrillen - Aktueller Stand von Forschung und Umsetzung sowie zukünftige Entwicklungsrichtungen. In: Neue Arbeits-und Lebenswelten gestalten. Bericht zum 56. Kongress der Gesellschaft für Arbeitswissenschaft. Dortmund: GfA-Press, S. 61-65

Abbildungsverzeichnis

Abb. 2.1	Darstellungsaufteilung der beiden Arbeitsaufgaben. Auf der rechten Seite Platz für die Konstruktionsanweisung, auf der linken Seite die Monitoringaufgabe mit drei Balken und Rückmeldung über die letzten bestätigten Zustände. (Hier Darstellung links montiertes HMD, bei rechts montiertem HMD seitenverkehrt: Die Monitoringaufgabe war immer peripher)	10
Abb. 2.2	Das verwendete MAVUS®-HMD-System. Links: Der Kopfträger mit Display, Kamera, Headset. Mitte: Das Display in Nahaufnahme. Rechts: Das Komplettsystem inklusive Weste und Tragekoffer.	11
Abb. 2.3	Links: Der Tablet-PC Motion CL900 mit abgeklebter Darstellungsfläche von 800x600 Pixeln und Leguanleitung. Rechts: Das umgebaute Nummernpad mit Reaktionstasten für die Monitoringaufgabe (oben/Mitte) und Tasten zum vor- und zurückschalten der Aufbaufolien.	13
Abb. 2.4	Die Rating Scale of Mental Effort (RSME) nach Zijlstra, 1993.	15
Abb. 2.5	Versuchsräume der BAuA. Links: Arbeitsplatz der Probanden mit höhenverstellbarem Tisch und Stuhl vor der Trennwand, im Hintergrund das HMD mit Weste. Rechts: Arbeitsplatz der Versuchsleiter hinter der Trennwand.	17
Abb. 2.6	Schematische Darstellung einer Session im Versuchsablauf.	18
Abb. 2.7	Häufigkeitsverteilung des Alters der Probanden, unterteilt nach gebildeten Gruppen.	19
Abb. 2.8	Herzratendaten eines bestimmten fünfminütigen Abschnittes ohne Filter. Deutliche Artefakte mit unrealistischen Werten sind erkennbar.	24
Abb. 2.9	Verlauf des gleichen fünfminütigen Abschnittes der Herzrate nach automatischer Artefaktkorrektur mit modifiziertem Filter nach Mulder. Ein einzelnes Artefakt ist noch erkennbar.	26
Abb. 2.10	Verlauf des gleichen fünfminütigen Abschnittes der Herzrate nach zusätzlicher manueller Artefaktkorrektur. Kein Artefakt ist mehr erkennbar.	27
Abb. 3.1	Anzahl bearbeiteter Konstruktionsfolien pro Versuchssession, unterteilt nach Altersgruppen. Blau = 1. Messung HMD, Grün = 2. Messung HMD, Beige = Tablet-PC. Die Whiskers zeigen das 95%-Konfidenzintervall.	28
Abb. 3.2	Trefferquote bei den beiden Monitoringaufgaben, unterteilt nach Versuchssession und gegebener Rückmeldung. Blau = Reaktion auf Farbveränderung; Grün = Reaktion auf Wechsel der Position des längsten Balkens. Die Whiskers zeigen das 95%-Konfidenzintervall.	30
Abb. 3.3	Reaktionszeiten bei den beiden Monitoringaufgaben, unterteilt nach Versuchssession und gegebener Rückmeldung. Blau = Reaktion auf Farbveränderung; Grün = Reaktion auf Wechsel der Position des längsten Balkens. Die Whiskers zeigen das 95%-Konfidenzintervall.	31

Abb. 3.4	Gewichteter NASA-TLX Gesamtwert nach zwei und vier Stunden unterteilt nach Versuchssession. Die Whiskers zeigen das 95%-Konfidenzintervall.	33
Abb. 3.5	RSME-Werte über den Verlauf der einzelnen Versuchssessions, unterteilt nach Alter und mit grafischem Hinweis auf die Pausen.	34
Abb. 3.6	Mittlere Herzrate aller Fünfminutenabschnitte über den gesamten Versuchsablauf. Fehlerbalken geben das 95%-Konfidenzintervall wieder. Pausen wurden vor den Minuten 60, 120 und 180 gemacht und sind durch graue Balken markiert.	36
Abb. 3.7	Mittlere Herzrate der zweiten HMD-Messung und des Tablet-PCs zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.	37
Abb. 3.8	Mittlere Herzrate der ersten und zweiten HMD-Messung zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.	38
Abb. 3.9	Mittlere Herzrate jeweils kurz vor und nach den drei Pausen. Fehlerbalken geben das 95%-Konfidenzintervall wieder.	39
Abb. 3.10	Mittlere Herzratenvariabilität aller Fünfminutenabschnitte über den gesamten Versuchsablauf. Fehlerbalken geben das 95%-Konfidenzintervall wieder. Pausen wurden vor den Minuten 60, 120 und 180 gemacht und sind durch graue Balken markiert.	40
Abb. 3.11	Herzratenvariabilität der zweiten HMD-Messung und des Tablet-PCs zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.	41
Abb. 3.12	Herzratenvariabilität der ersten und zweiten HMD-Messung zu acht Messzeitpunkten. Fehlerbalken zeigen das 95%-Konfidenzintervall.	43
Abb. 3.13	Herzratenvariabilität jeweils kurz vor und nach den drei Pausen. Fehlerbalken geben das 95%-Konfidenzintervall wieder.	44
Abb. 3.14	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „trockene Augen“, rechts: „tränennde Augen“.	46
Abb. 3.15	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „gereizte/brennende Augen“, rechts: „schmerzende Augen“	46
Abb. 3.16	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „schwere Augen“, rechts: „Schwierigkeit, Linien folgen zu können“.	47
Abb. 3.17	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „Schwierigkeit, scharf zu sehen“, rechts: „springender Text“.	48
Abb. 3.18	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „verschwommene Buchstaben“, Rechts: „Blendung durch Lichter“.	48
Abb. 3.19	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „verschwommene Sicht“, rechts: „Doppelsehen“.	49
Abb. 3.20	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „Kopfschmerzen“, Rechts: „Nackenschmerzen“.	50
Abb. 3.21	VFQ-Werte über den gesamten Versuchszeitraum pro Session. Links: „geistige Ermüdung“, Rechts: „Übelkeit“.	50
Abb. 3.22	Antworten auf die Frage „Konnten Sie alles scharf sehen?“	52

Abb. 3.23	Antworten auf die Frage „Wie war der Tragekomfort des Kopfträgers?“	53
Abb. 3.24	Antworten auf die Frage „Wie fanden Sie die Darstellung vor nur einem Auge?“	54
Abb. 3.25	Antworten auf die Frage „Wo sehen Sie noch Verbesserungsmöglichkeiten in der Technologie?“	55
Abb. 3.26	Antworten auf die Frage „Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden/Sehen bemerkt? Und hat sich dies nach der Pause wieder verändert?“	56
Abb. 3.27	Antworten auf die Frage „Haben Sie einen Unterschied zwischen erster und zweiter Erhebung mit dem HMD bemerkt?“	57
Abb. 4.1	Darstellungsaufteilung der Arbeitsinhalte in der Replikationsstudie mit Google Glass. Links: Konstruktionsfolie für Lego-Technic ; rechts: Monitoringaufgabe mit drei Balken und Rückmeldung über letzte bestätigte Farbe und Position.	61
Abb. 4.2	Links: Google Glass mit angeschlossenenem Akkumulator-expansionspack. Rechts: der verwendete Tablet-PC.	63
Abb. 4.3	Schematischer Versuchsablauf der gesamten Studienreihe. Die hier berichtete Replikation fand in der zweiten Stunde in den rot gestreiften Abschnitten statt. Weitere Erläuterungen im Text.	67
Abb. 4.4	Links: Aufnahme der Lego-Technic-Bausteine an einem Ort (hier mit Tablet-PC). Rechts: Zusammenbau des Fahrzeuges an einem vier Meter entfernten Ort (hier exemplarisch mit Google Glass, inkl. Akkuerweiterung).	68
Abb. 4.5	Altersverteilung der HMD- und Tablet-PC-Gruppe der Replikationsstudie. Links: Google Glass. Rechts: Tablet-PC.	69
Abb. 4.6	Mittlere Anzahl bearbeiteter Konstruktionsfolien innerhalb von 25 Minuten Versuchszeit unterteilt nach Altersgruppen und Displaytyp. Fehlerbalken zeigen das 95%-Konfidenzintervall.	71
Abb. 4.7	Trefferquote in Prozent in der Farb-Monitoring-Aufgabe unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.	73
Abb. 4.8	Trefferquote in Prozent in der Monitoringaufgabe zum Positionswechsel des längsten Balkens unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.	74
Abb. 4.9	Reaktionszeiten in Sekunden in der Farbwechsel-Monitoringaufgabe unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.	75
Abb. 4.10	Reaktionszeit in Sekunden in der Monitoringaufgabe zum Positionswechsel des längsten Balkens unterteilt nach Displaytyp, Altersgruppe und Baseline und Dualtask. Fehlerbalken zeigen das 95%-Konfidenzintervall.	76
Abb. 4.11	Gewichteter Summenscore des NASA-TLX, unterteilt nach Altersgruppen und Displaytyp. Fehlerbalken zeigen das 95%-Konfidenzintervall.	77

- Abb. 4.12** Mittlere Werte der RSME-Skala, unterteilt nach Altersgruppen, Displaytyp und Zeitpunkt (zu Beginn der Studie nach 90 Minuten und gegen Ende nach 120 Minuten Gesamtversuchszeit). Fehlerbalken zeigen das 95%-Konfidenzintervall. 78
- Abb. 4.13** Mittelwerte für die VFQ-Items „trockene Augen“ (links) und „gereizte/brennende Augen“ (rechts), unterteilt nach Display und Altersgruppe. 79
- Abb. 4.14** Mittelwerte für die VFQ-Items „schwere Augen“ (links) und „Schwierigkeit Linien zu folgen“ (rechts), unterteilt nach Display und Altersgruppe. 80
- Abb. 4.15** Mittelwerte für die VFQ-Items „Schwierigkeiten, scharf zu sehen“ (links) und „verschwommene Sicht“ (rechts), unterteilt nach Display und Altersgruppe. 81
- Abb. 4.16** Mittelwerte für die VFQ-Items „Kopfschmerzen“ (links) und „Nackenschmerzen“ (rechts), unterteilt nach Display und Altersgruppe. 81
- Abb. 4.17** Mittelwerte für das VFQ-Item „Geistige Ermüdung“, unterteilt nach Display und Altersgruppe. 82

Anhang

Anhang 1 Unterlagen der Hauptstudie

1.1 Probandeninformation

Probandeninformation

zum Forschungsvorhaben

“Datenbrillen - Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes”

Sehr geehrte Versuchsteilnehmerin, sehr geehrter Versuchsteilnehmer, wir danken Ihnen für Ihr Interesse, an der Untersuchung zu Datenbrillen teilzunehmen, und möchten Sie auf folgende Punkte hinweisen:

Die Teilnahme an dieser Untersuchung ist *freiwillig*. Die Teilnahme kann jederzeit ohne Angabe von Gründen widerrufen werden, ohne dass Ihnen daraus Nachteile entstehen. Auch eine laufende Untersuchung wird auf Ihren Wunsch hin jederzeit unterbrochen.

Diese Untersuchung dient *nicht* der medizinischen Diagnostik oder der Beurteilung Ihrer Belastbarkeit und Leistungsfähigkeit. Die Ergebnisse dieser Untersuchung helfen, Funktionen von Datenbrillen im Vergleich zu anderen Darstellungsformen zu bewerten. Auf Grundlage der ermittelten physiologischen Kennwerte und Ihrer subjektiven Einschätzung werden die Datenbrillen bewertet und es können Möglichkeiten zur Optimierung aufgezeigt werden.

Ihre personenbezogenen Daten werden maschinell gespeichert und weiterverarbeitet. Sie unterliegen dem Datenschutz, der eine unbefugte Weitergabe an Dritte verhindert.

Bei wissenschaftlichen Studien werden persönliche Daten und medizinische Befunde über Sie erhoben. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen ohne Namensnennung.

Es folgen nun einige Hinweise zu den bei dieser Studie eingesetzten Untersuchungstechniken. Selbstverständlich erhalten Sie vor Versuchsbeginn weitere ausführliche Informationen von den Untersuchern.

In Simulationsversuchen werden Datenbrillen und andere Darstellungsformen in unterschiedlichen Probandenversuchen zum Zweck der ergonomischen Bewertung getestet. Neben Fragebögen zur Beanspruchung wird die Beanspruchung auch über die Herzrate und die Herzratenvariabilität erhoben. Hierzu wird während des gesamten Versuches ein Brustgurt die Herzaktivität erfassen. Dies dient nicht der medizinischen Diagnostik im Einzelfall, sondern gibt im Vergleich die Anstrengung der Probanden unter verschiedenen Rahmenbedingungen wieder.

Die gesamte Versuchszeit wird videodokumentiert. Dies geschieht, um eventuell ungewöhnlichen Daten ein spezielles Ereignis zuzuordnen. Diese Videodokumentationen werden nur im Bedarfsfall vom Versuchsleiter

eingesehen. Eine Veröffentlichung des Videos oder von Ausschnitten findet nicht statt.

Die im Versuch zu erledigenden Aufgaben erfordern keinerlei Vorkenntnisse und werden Ihnen jeweils vor Beginn des Versuchs erklärt bzw. es wird Ihnen ausreichend Zeit gegeben, sich mit der jeweiligen Aufgabe vertraut zu machen.

Einwilligungserklärung

zum Forschungsvorhaben

“Datenbrillen”

Ich bin über Wesen und Bedeutung der geplanten Untersuchungen ausführlich aufgeklärt worden. Die Probandeninformationen über die Untersuchung habe ich gelesen und verstanden.

Meine Teilnahme an dem Forschungsvorhaben ist ganz und gar freiwillig. Ich wurde darauf hingewiesen, dass ich meine Einwilligung jederzeit ohne Angabe von Gründen widerrufen kann, ohne dass mir dadurch irgendwelche Nachteile entstehen.

Zu meiner eigenen Sicherheit werde ich mich während des Versuchs so verhalten, wie es von den Untersuchern erläutert wird.

Mit den Aufzeichnungen von Untersuchungsdaten bin ich einverstanden.

Ich bin darüber informiert, dass bei wissenschaftlichen Studien persönliche Daten und medizinische Befunde erhoben werden. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen und setzt vor Teilnahme an der Studie meine freiwillige Einwilligung voraus.

Ich bin darauf hingewiesen worden, dass die Bundesrepublik Deutschland im Rahmen der Amtshaftung (Art. 34 Satz 1 GG i.V.m. § 839 BGB) für alle Schäden aufkommt, die ich im Zusammenhang mit der Untersuchung durch das schuldhaft (d. h. vorsätzliche oder fahrlässige) Handeln der Bediensteten erleide. Dabei trage ich als Geschädigter grundsätzlich die Beweislast. Mir ist ferner bekannt, dass ich für Schäden, die mir auf dem Weg von meiner Wohnung zum Untersuchungsort und zurück entstehen (sogenannter Wegeunfall), grundsätzlich keinen Schadensersatz von der Bundesrepublik Deutschland verlangen kann. Ich willige hiermit ein, als Proband/-in am Forschungsvorhaben “Datenbrillen” teilzunehmen.

Name: _____ Vorname: _____

Dortmund, den _____ Unterschrift: _____

Probandenfragebogen

zum Forschungsvorhaben

“Datenbrillen”

Name: _____

Adresse: _____

Telefon: _____

Geburtsdatum: _____

Wir bemühen uns, die Untersuchung für Sie so angenehm wie möglich zu gestalten. Dieser Fragebogen dient dazu, mögliche Risiken vor der Untersuchung abzuschätzen und auszuschließen.

1. Haben oder hatten Sie eine der folgenden Krankheiten?

	Ja	Nein
Allgemeine Körperschwäche	<input type="radio"/>	<input type="radio"/>
Bewusstseins- oder Gleichgewichtsstörungen	<input type="radio"/>	<input type="radio"/>
Anfallsleiden (z.B. Epilepsie)	<input type="radio"/>	<input type="radio"/>
Erkrankungen oder Schäden des zentralen oder peripheren Nervensystems	<input type="radio"/>	<input type="radio"/>
Chronischer Alkoholmissbrauch, Betäubungsmittelsucht oder andere Suchtformen	<input type="radio"/>	<input type="radio"/>
Migräne	<input type="radio"/>	<input type="radio"/>
Cluster- Kopfschmerzen	<input type="radio"/>	<input type="radio"/>
Starke Rückenschmerzen (z.B. durch Bandscheibenvorfall)	<input type="radio"/>	<input type="radio"/>
Stoffwechselkrankheiten, die die Belastbarkeit stark einschränken	<input type="radio"/>	<input type="radio"/>
Rot-Grün Blindheit	<input type="radio"/>	<input type="radio"/>
Mobilitätseinschränkungen der Arme oder Hände	<input type="radio"/>	<input type="radio"/>
Fehlende Einwilligungsfähigkeit des Probanden	<input type="radio"/>	<input type="radio"/>

2. Bei weiblichen Probanden: Sind Sie schwanger?

Ja Nein

3. Nehmen Sie zurzeit zentral wirksame Medikamente ein?

Ja Nein

4. Besteht eine Medikamentenallergie?

Ja	Nein	Falls ja, gegen welche
<input type="radio"/>	<input type="radio"/>	_____

5. Haben Sie akut Kreislauf- oder Atembeschwerden (z. B. Bluthochdruck, Asthma usw.)?

Ja	Nein
<input type="radio"/>	<input type="radio"/>

Ich versichere hiermit die Richtigkeit der Angaben.

Dortmund, den _____

(Unterschrift)

1.2 Fragebogen zu Demographie und Vorerfahrung

Probandennummer	
Geschlecht	männlich / weiblich
Alter	
Händigkeit	rechts / links
Führungsaug	rechts / links
Sehhilfe	Brille / Kontaktlinsen Dioptrienzahl: links rechts

Vorerfahrung mit verschiedenen Technologien

(bezogen auf die letzten 3 Monate, Umrechnung in Stunden pro Tag)

Computer	
Tablet	
Smartphone	
Fernsehen	

Vorerfahrung mit bestimmten versuchsbezogenen Techniken

(allgemein, Details nachfragen)

HMDs / Datenbrillen

Lego-Technik

1.3 Instruktion

Sehr geehrte(r) Proband(in),

vielen Dank, dass Sie sich bereit erklärt haben, an unserer Studie über Datenbrillen teilzunehmen. Diese schriftliche Instruktion soll Ihnen eine Orientierung über die Hintergründe, den Versuchsablauf und Ihre Aufgaben geben.

Bitte lesen Sie die Instruktion in Ruhe durch. Sollten Sie anschließend noch Fragen haben, wenden Sie sich bitte an den Versuchsleiter.

Über Datenbrillen:

Datenbrillen sind kleine Monitore, die direkt vor den Augen positioniert werden. Manchmal werden Datenbrillen auch „Head Mounted Displays“ (HMDs) genannt, was übersetzt soviel bedeutet wie „am Kopf befestigte Anzeigen“.

Diese allgemeine Definition zeigt, dass es verschiedene Arten von Datenbrillen gibt. So gibt es Datenbrillen die auf beiden Augen Informationen anzeigen oder auch welche, die Informationen vor nur einem Auge anzeigen. Es gibt Datenbrillen bei denen das Display halb transparent ist, so dass neben den Informationen noch die dahinter liegende Umgebung gesehen werden kann, und es gibt Displays die nicht transparent sind und die Umgebung verdecken. Auch der technische Aufbau von Display und Gestell unterscheidet sich stark.

Für unsere Versuche verwenden wir ein Modell das die Informationen auf nur einem Auge anzeigt, nicht transparent ist und einen Kopfträger besitzt, der an einen Fahrradhelm erinnert (*siehe Bild rechts*). Des Weiteren hat dieses Modell ein Headset (Kopfhörer und Mikrophon) und eine Kamera eingebaut, die das Blickfeld des Trägers wiedergibt. *Das Headset und die Kamera finden in unserem Versuchsaufbau aber keine Verwendung.*

Genutzt wird das Modell momentan bei der Instandhaltung von Industrieanlagen: Bei komplexen Maschinen steht der Arbeiter vor Ort mittels dieser Datenbrille mit einem Experten in Kontakt, der irgendwo anders auf der Welt sitzt. Dieser Experte sieht, was der Arbeiter sieht und kann ihn bei der Arbeit unterstützen. Er kann z.B. auf einem Bild der Maschine bestimmte Schrauben einkreisen die entfernt werden müssen und dieses Bild dem Arbeiter auf das Display legen. Außerdem stehen Sie über Funk in Kontakt.

Weitere Anwendungsfelder von Datenbrillen allgemein finden sich z.B. in der Montage oder Kommissionierung von Waren.

Datenbrillen sind eine noch recht neue Technik, die sich stetig weiterentwickelt. Wir testen hier den Stand der Technik, wie er momentan eingesetzt wird. Neuere Modelle von Datenbrillen können z.B. leichter sein,



oder ein größeres Display haben. Oft befinden sich diese Modelle aber noch im Prototypen Status (z.B. derzeit die „Googlebrille“) obwohl Bilder von ihnen manchmal schon große Verbreitung in den Medien finden.

Allgemeines zum Versuchsablauf:

Sie werden gleich an einer Studie zum Langzeiteinsatz von Datenbrillen teilnehmen. Im Rahmen dieser Studie werden Sie an 3 Terminen jeweils 4 Stunden mit einer Datenbrille oder vergleichsweise mit einem Tablet-PC arbeiten.

Die wiederholte Erhebung an 3 Terminen ist notwendig, um so auch Gewöhnung an neue Technologien untersuchen zu können. Probanden die nicht an allen Terminen teilnehmen können später nicht ausgewertet werden.

Während der 4-stündigen Arbeitsphase ist am Ende jeder Stunde eine 5-minütige Pause vorgesehen. Die Pause nach der 2. Stunde ist etwas länger, Sie bekommen die Datenbrille abgenommen und es wird empfohlen sich ein wenig zu bewegen.

In den Pausen erhalten Sie jeweils einen Fragebogen zur visuellen Beanspruchung. In der großen Pause machen Sie zusätzlich noch einen weiteren Fragebogen am Rechner.

Sollten Sie während des Versuches irgendwelche Beschwerden bekommen, teilen Sie dies bitte unmittelbar dem Versuchsleiter mit!

Als Aufwandsentschädigung erhalten Sie 10,- Euro pro Stunde. Die Auszahlung erfolgt nach dem 3. Termin. Bitte bringen Sie zu diesem Termin Ihre Kontodaten mit, damit wir die Überweisung einleiten können.

Ihre Aufgaben während des Versuches:

Während des Versuches werden Sie 2 Aufgaben parallel bearbeiten die beide gleich wichtig sind. Die Arbeitsinhalte werden Ihnen entweder über die Datenbrille, oder über den Tablet-PC angezeigt.

Zum einen werden Sie **Lego-Technik** Fahrzeuge aufbauen. Aus Einzelteilen werden Sie Schritt für Schritt ein komplexes Fahrzeug erstellen. Hierbei werden Ihnen bei jedem Arbeitsschritt in einem oberen Feld die zu verwendenden Teile angezeigt. Im Hauptfeld sehen Sie dann, wie diese Teile an dem Modell zu befestigen sind (*siehe Abbildung auf der nächsten Seite*). Die Arbeitsschritte sind dabei aufeinander folgend und durchnummeriert. Bei umfangreicheren Arbeitsschritten kann es auch Unter-Nummern geben.

Mit zwei Tasten an dem Nummernpad können Sie die Arbeitsschritte vor und zurück blättern. Bedenken Sie, dass Fehler im Aufbau sich an späterer Stelle bemerkbar machen können. Kontrollieren Sie deshalb die exakte Ausführung des Arbeitsschrittes, bevor Sie zum nächsten Schritt blättern. Sollten Sie einen Fehler bemerken, korrigieren Sie ihn oder blättern Sie zurück bis an die entsprechende Stelle in der Aufbauanleitung. Möglicher Weise müssen Sie hierbei auch einige Teile wieder abbauen, um den Fehler beheben zu können. Falls Sie bei einem bestimmten Arbeitsschritt nicht weiter kommen, fragen Sie den Versuchsleiter.

Parallel zum Aufbau der Fahrzeuge sollen Sie eine weitere Aufgabe bearbeiten: Am Rand neben den Lego-Arbeitsschritten werden **drei Balken** angezeigt, die langsam aber kontinuierlich Ihre Länge verändern. Sollte hierbei die Position (rechts-mitte-links) des längsten Balkens wechseln, bestätigen Sie dies bitte durch einen Knopf-

druck. Von Zeit zu Zeit wechseln diese Balken auch die Farbe (blau/rot). Ein Farbwechsel ist durch einen anderen Knopfdruck zu bestätigen.
 In einer Hälfte des Versuches werden Sie hierbei schriftliche Rückmeldung über die letzte von Ihnen bestätigten Position und Farbe erhalten, in der anderen Hälfte des Versuches erhalten Sie diese Rückmeldung nicht.

Bitte bearbeiten Sie beide Aufgaben parallel und möglichst schnell und fehlerfrei!

Wir werden diese Aufgaben zu Beginn der Versuchsreihe kurz üben, damit Sie sie kennenlernen.

Falls es noch Unklarheiten gibt, oder Sie noch weitere Fragen zu den Aufgaben oder dem Versuchsablauf haben, wenden Sie sich bitte jetzt an den Versuchsleiter.

Darstellung der Arbeitsinhalte

The image is divided into two main sections. The left section shows a LEGO assembly instruction for step 22. It features a light blue background with a white box at the top left containing two small grey parts, each labeled '1x'. Below this, the number '22' is displayed in large black font. The main part of the instruction shows a sub-assembly of grey and red LEGO beams and wheels. A white arrow points from the parts in the box to the sub-assembly. A small number '2' is visible in the bottom left corner of the instruction area. The right section of the image shows a bar chart with three vertical green bars of decreasing height from left to right, set against a light blue background.

Lego Aufbauanleitung (links) und 3 Balken (rechts)

Ansonsten wünschen wir Ihnen viel Spaß und gutes Gelingen bei Ihren Aufgaben!

Ihr Laborteam der BAuA



1.4 Fragebogen TA-EG



**ZMMS
TA-EG**



Fragebogen zur Technikaffinität – Einstellung zu und Umgang mit elektronischen Geräten

Vielen Dank, dass Sie an dieser Befragung teilnehmen. Üblicherweise dauert es nicht länger als 10 Minuten, den Fragebogen auszufüllen.

Der Fragebogen befasst sich mit Ihrer persönlichen Meinung zu verschiedenen Aspekten elektronischer Geräte sowie mit der Erfahrung, die Sie im Umgang mit diesen Geräten haben.

Unter dem Begriff „elektronische Geräte“ verstehen wir Geräte, wie:

- Computer, Handys oder Palms/PDAs, Digitalkameras, DVD-Spieler oder Mp3-Spieler,
- Geldautomaten, Ticketautomaten,
- neue Systeme im Auto wie Navigationssysteme.

Nicht gemeint sind:

- Werkzeuge wie Bohrmaschine, Kreissäge, Hammer oder Rasenmäher,
- Haushaltsgeräte wie Toaster, Handrührgerät, Wasserkocher,
- Fahrzeuge und Fahrzeugmotoren.

Ihre Angaben werden selbstverständlich vertraulich behandelt und lediglich zu Forschungszwecken verwendet. Da die Befragung anonym erfolgt, sind keine Rückschlüsse auf Ihre Person möglich.

Aktuelle Quelle:

Karrer, K., Glaser, C., Clemens, C. & Bruder, C. (2009). Technikaffinität erfassen – der Fragebogen TA-EG. In A. Lichtenstein, C. Stößel und C. Clemens (Hrsg.), *Der Mensch im Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme* (ZMMS Spektrum, Reihe 22, Nr. 29, S. 196-201). Düsseldorf: VDI Verlag GmbH.

Im Folgenden finden Sie eine Reihe von Aussagen. Bitte geben Sie für jede der Aussagen an, wie gut sie auf Sie persönlich zutrifft. Kreuzen Sie dazu auf der rechten Seite das Feld an, das Ihrer Meinung am besten entspricht.

	Trifft voll zu	Trifft eher zu	Teils / teils	Trifft eher nicht zu	Trifft gar nicht zu
1. Ich liebe es, neue elektronische Geräte zu besitzen.					
2. Elektronische Geräte machen krank.					
3. Ich gehe gern in den Fachhandel für elektronische Geräte.					
4. Ich habe bzw. hätte Verständnisprobleme beim Lesen von Elektronik- und Computerzeitschriften.					
5. Elektronische Geräte ermöglichen einen hohen Lebensstandard.					
6. Elektronische Geräte führen zu geistiger Verarmung.					
7. Elektronische Geräte machen vieles umständlicher.					
8. Ich informiere mich über elektronische Geräte, auch wenn ich keine Kaufabsicht habe.					
9. Elektronische Geräte machen unabhängig.					
10. Es macht mir Spaß, ein elektronisches Gerät auszuprobieren.					
11. Elektronische Geräte erleichtern mir den Alltag.					
12. Elektronische Geräte erhöhen die Sicherheit.					
13. Elektronische Geräte verringern den persönlichen Kontakt zwischen den Menschen.					
14. Ich kenne die meisten Funktionen der elektronischen Geräte, die ich besitze.					
15. Ich bin begeistert, wenn ein neues elektronisches Gerät auf den Markt kommt.					
16. Elektronische Geräte verursachen Stress.					
17. Ich kenne mich im Bereich elektronischer Geräte aus.					
18. Es fällt mir leicht, die Bedienung eines elektronischen Geräts zu lernen.					
19. Elektronische Geräte helfen, an Informationen zu gelangen.					

Vielen Dank!

1.5 RSME-Werte und Beobachtung

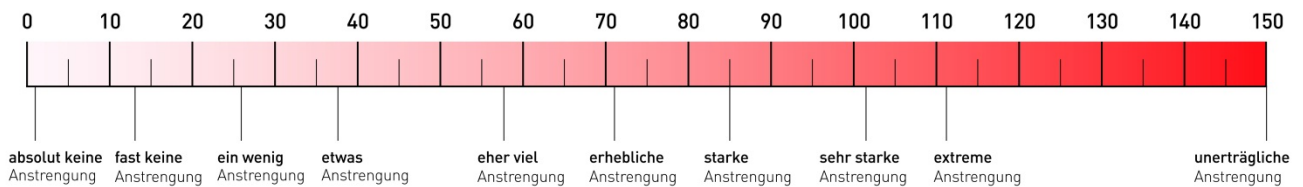
RSME Werte & Beobachtung VPnr _____ Session _____ HMD / Tab

Beispielwerte zur Vorstellung der Skala:

Eine 2-stündige Shoppingtour am Samstag Vormittag in der Innenstadt	
Fernsehabend mit Freunden oder Partner	
Aufblasen von 50 Luftballons für eine Geburtstagsfeier	
Ein Puzzle mit 500 Teilen	
Korrekturlesen eines 30-seitigen Berichtes am Monitor	
Kinobesuch der 4 Stunden Version von „Ben Hur“	
Die Beantwortung dieser Beispielsituationen	

Abfrage der RSME Werte während des Versuches alle halbe Stunde:

Hauptversuch 1. Stunde	30 min	
	1 Stunde	
Hauptversuch 2. Stunde	1 Stunde 30	
	2 Stunden	
Hauptversuch 3. Stunde	2 Stunden 30	
	3 Stunden	
Hauptversuch 4. Stunde	3 Stunden 30	
	4 Stunden	



Einschätzung der manuellen Korrektur am Display von der Versuchsperson

(Häufigkeit des Anfassens):

1. Stunde	----- ----- ----- ----- ----- ----- ----- ----- ----- -----
	nie selten manchmal häufig ständig
2. Stunde	----- ----- ----- ----- ----- ----- ----- ----- ----- -----
	nie selten manchmal häufig ständig
3. Stunde	----- ----- ----- ----- ----- ----- ----- ----- ----- -----
	nie selten manchmal häufig ständig
4. Stunde	----- ----- ----- ----- ----- ----- ----- ----- ----- -----
	nie selten manchmal häufig ständig

Permutationsabfolge der Rückmeldung der Balkenaufgabe:

Stunde	Rückmeldung	Lego-Model <i>(bei Wechsel Videozeit angeben)</i>
1	ja / nein	
2	ja / nein	
3	ja / nein	
4	ja / nein	

Anmerkungen:

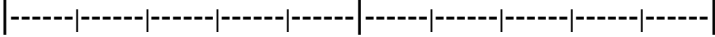
Videozeit	

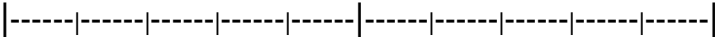
1.6 Fragebogen zur visuellen Ermüdung

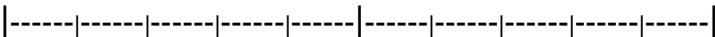
Fragebogen zur visuellen Ermüdung

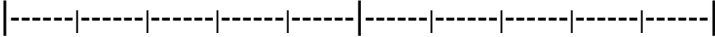
Zur Beantwortung der folgenden Kategorien kreuzen Sie bitte die Skala an der Ihrem Urteil am besten entsprechenden Stelle – auch zwischen den Markierungen – an.

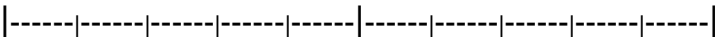
Bemerken Sie, dass Sie im Moment von einem dieser Symptome beeinträchtigt werden?

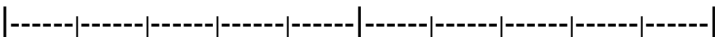
Trockene Augen: 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

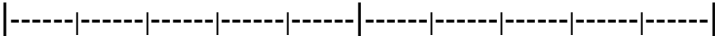
Tränende Augen: 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

Die Augen sind gereizt oder brennen 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

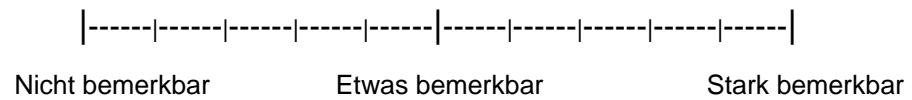
Schmerzen in oder um die Augen 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

Schwere Augen 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

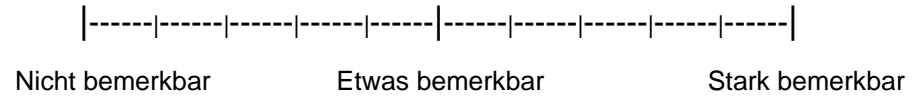
Schwierigkeiten, Linien folgen zu können 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

Schwierigkeiten scharf zu sehen 
 Nicht bemerkbar Etwas bemerkbar Stark bemerkbar

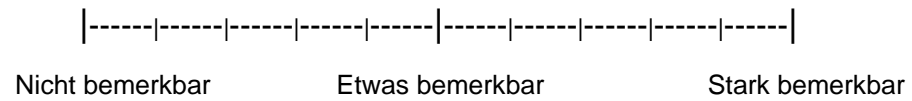
Springender
Text



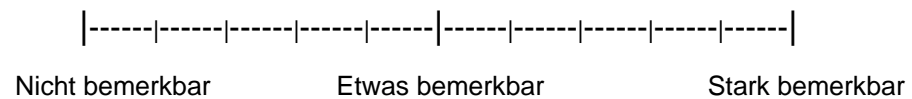
Verschwom-
mene
Buchstaben



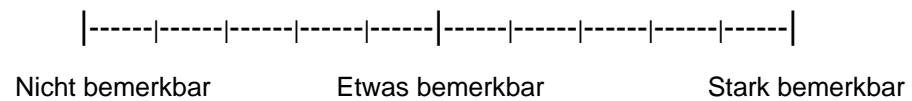
Blendung
durch Lichter



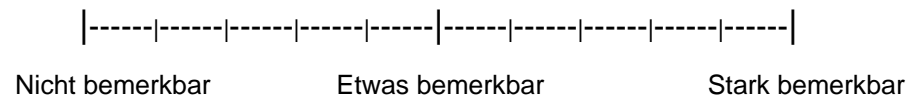
Verschwomme-
ne Sicht



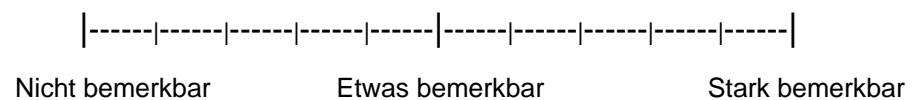
Doppelsehen



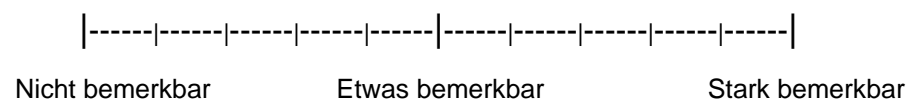
Kopfschmer-
zen



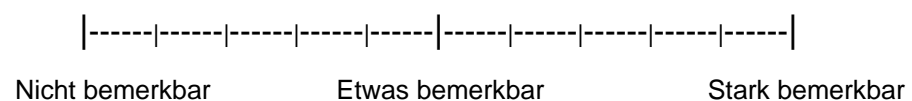
Nacken-
schmerzen



Übelkeit



Geistige
Ermüdung



1.7 HMD-Interview

HMD-Interview

VPNR:

Session:

Kamen Sie mit der Konstruktionsaufgabe gut zurecht?

Falls „nein“: Glauben Sie das lag an der Darstellung im Display, oder wäre das auf Papier das gleiche Problem gewesen?

Hatten Sie beim Bearbeiten der Balkenaufgabe eine bestimmte Strategie?

Fanden Sie die Rückmeldung in der Balkenaufgabe über Ihre letzte gedrückte Antwort hilfreich? Wenn ja, warum?

Wie fanden Sie die Darstellung auf nur einem Auge?

Konnten Sie alles scharf sehen?

Wie waren die Trageeigenschaften des Kopfträgers?

Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden / Sehen bemerkt? Hat sich dies nach den Pausen wieder verändert?

Beim 2.mal HMD: Haben Sie einen Unterschied zur 1. Session bemerkt?

Wo sehen Sie noch Verbesserungspotential in der Technik?

Welche Einsatzgebiete könnten Sie sich für eine Datenbrille vorstellen?

Nur in Session 3: Würden Sie den Versuch noch mal machen?

Haben Sie noch weitere Anregungen oder Hinweise?

1.8 Tablet-PC-Interview

Tab-Interview

VPNR:

Session:

Kamen Sie mit der Konstruktionsaufgabe gut zurecht?

Falls „nein“: Glauben Sie das lag an der Darstellung im Display, oder wäre das auf Papier das gleiche Problem gewesen?

Hatten Sie beim Bearbeiten der Balkenaufgabe eine bestimmte Strategie?

Fanden Sie die Rückmeldung in der Balkenaufgabe über Ihre letzte gedrückte Antwort hilfreich? Wenn ja, warum?

Wie fanden Sie die Darstellung auf dem Tablet-PC?

Konnten Sie alles scharf sehen?

Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden / Sehen bemerkt? Hat sich dies nach den Pausen wieder verändert?

Wie fanden Sie allgemein das Arbeiten mit dem Tablet-PC?
(hat Sie irgendetwas gestört oder fanden Sie etwas besonders gut)

Wo sehen Sie noch Verbesserungspotential in der Technik?

Welche Einsatzgebiete könnten Sie sich für einen Tablet-PC vorstellen?

Fanden Sie das Arbeiten mit dem Tablet-PC oder mit der Datenbrille besser?

Warum?

Nur in Session 3: Würden Sie den Versuch noch mal machen?

Haben Sie noch weitere Anregungen oder Hinweise?

1.9 Verwendete Lego Modelle

8065 „Mini Kipplaster“	<ul style="list-style-type: none"> • 119 Teile • Kipplaster oder Pick-up • Aufbauanleitung: 52 bzw. 32 Seiten
9392 „Quad Bike“	<ul style="list-style-type: none"> • 199 Teile • Quad Bike oder Wüstenbuggy • Aufbauanleitung: 57 bzw. 56 Seiten
8067 „Mobiler Mini-Kran“	<ul style="list-style-type: none"> • 292 Teile • Mobiler Mini-Kran oder Abschleppwagen • Aufbauanleitung: 58 bzw. 59 Seiten
8049 „Traktor mit Forstkran“	<ul style="list-style-type: none"> • 552 Teile • Traktor mit Forstkran oder Tieflader • Aufbauanleitung: 44+91 bzw. 124 Seiten
9394 „Düsenflugzeug“	<ul style="list-style-type: none"> • 499 Teile • Düsenflugzeug oder Propellerflugzeug • Aufbauanleitung: 110 bzw. 104 Seiten

Anhang 2 Unterlagen der Replikationsstudie

2.1 Probandeninformation

Probandeninformation

zum Forschungsvorhaben

“Datenbrillen - Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes”

Sehr geehrte Versuchsteilnehmerin, sehr geehrter Versuchsteilnehmer,
wir danken Ihnen für Ihr Interesse, an der Untersuchung zu Datenbrillen teilzunehmen, und möchten Sie auf folgende Punkte hinweisen:

1. Die Teilnahme an dieser Untersuchung ist *freiwillig*. Die Teilnahme kann jederzeit ohne Angabe von Gründen widerrufen werden, ohne dass Ihnen daraus Nachteile entstehen. Auch eine laufende Untersuchung wird auf Ihren Wunsch hin jederzeit unterbrochen.
2. Diese Untersuchung dient *nicht* der Beurteilung Ihrer Belastbarkeit und Leistungsfähigkeit. Die Ergebnisse dieser Untersuchung helfen, Funktionen von Datenbrillen im Vergleich zu anderen Darstellungsformen zu bewerten. Auf Grundlage der ermittelten Kennwerte und Ihrer subjektiven Einschätzung werden die Datenbrillen bewertet und es können Möglichkeiten zur Optimierung aufgezeigt werden.
3. Ihre personenbezogenen Daten werden maschinell gespeichert und weiterverarbeitet. Sie unterliegen dem Datenschutz, der eine unbefugte Weitergabe an Dritte verhindert.

Bei wissenschaftlichen Studien werden persönliche Daten von Ihnen erhoben. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen ohne Namensnennung.

4. Es folgen nun einige Hinweise zu den bei dieser Studie eingesetzten Untersuchungstechniken. Selbstverständlich erhalten Sie vor Versuchsbeginn weitere ausführliche Informationen von den Untersuchern.

In Simulationsversuchen werden Datenbrillen und andere Darstellungsformen in unterschiedlichen Probandenversuchen zum Zweck der ergonomischen Bewertung getestet. Neben Ihrer Arbeit mit der Datenbrille werden verschiedene Fragebögen zur Beanspruchung erhoben. Dies gibt im Vergleich die Anstrengung der Probanden unter verschiedenen Rahmenbedingungen wieder.

Weiterhin werden während der gesamten Versuchszeit Bewegungen des dominanten Handgelenks aufgezeichnet. Dazu werden sie gebeten ein Uhrenähnliches Gerät zu tragen. Dieses zeichnet die an ihrem Handgelenk wirkenden Beschleunigungen kontinuierlich auf. Dies dient dazu Handgelenkbewegungen einzelnen Handlungen während des Versuchs zuzuordnen.

Die im Versuch zu erledigenden Aufgaben erfordern keinerlei Vorkenntnisse und werden Ihnen jeweils vor Beginn des Versuchs erklärt bzw. es wird Ihnen ausreichend Zeit gegeben, sich mit der jeweiligen Aufgabe vertraut zu machen.

5. Bildmaterial

Die gesamte Versuchszeit wird videodokumentiert. Dies geschieht, um eventuell ungewöhnlichen Daten ein spezielles Ereignis zuzuordnen. Diese Videodokumentationen werden im Bedarfsfall vom Versuchsleiter eingesehen.

Eine Veröffentlichung von Ausschnitten des Videos findet nur im Falle Ihrer ausdrücklichen Einwilligung statt. Hierzu werden Sie am Ende des Versuches getrennt befragt.

Einwilligungserklärung

zum Forschungsvorhaben

“Datenbrillen”

Ich bin über Wesen und Bedeutung der geplanten Untersuchungen ausführlich aufgeklärt worden. Die Probandeninformationen über die Untersuchung habe ich gelesen und verstanden.

Meine Teilnahme an dem Forschungsvorhaben ist ganz und gar freiwillig. Ich wurde darauf hingewiesen, dass ich meine Einwilligung jederzeit ohne Angabe von Gründen widerrufen kann, ohne dass mir dadurch irgendwelche Nachteile entstehen.

Zu meiner eigenen Sicherheit werde ich mich während des Versuchs so verhalten, wie es von den Untersuchern erläutert wird.

Mit den Aufzeichnungen von Untersuchungsdaten bin ich einverstanden.

Ich bin darüber informiert, dass bei wissenschaftlichen Studien persönliche Daten erhoben werden. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen und setzt vor Teilnahme an der Studie meine freiwillige Einwilligung voraus.

Ich bin darauf hingewiesen worden, dass die Bundesrepublik Deutschland im Rahmen der Amtshaftung (Art. 34 Satz 1 GG i.V.m. § 839 BGB) für alle Schäden aufkommt, die ich im Zusammenhang mit der Untersuchung durch das schuldhaft (d. h. vorsätzliche oder fahrlässige) Handeln der Bediensteten erleide. Dabei trage ich als Geschädigter grundsätzlich die Beweislast. Mir ist ferner bekannt, dass ich für Schäden, die mir auf dem Weg von meiner Wohnung zum Untersuchungsort und zurück entstehen (sogenannter Wegeunfall), grundsätzlich keinen Schadensersatz von der Bundesrepublik Deutschland verlangen kann. Ich willige hiermit ein, als Proband/-in am Forschungsvorhaben “Datenbrillen” teilzunehmen.

Name: _____ Vorname: _____

Dortmund, den _____ Unterschrift: _____

Probandenfragebogen

zum Forschungsvorhaben
"Datenbrillen"

Name: _____

Adresse: _____

Telefon: _____

Geburtsdatum: _____

Wir bemühen uns, die Untersuchung für Sie so angenehm wie möglich zu gestalten. Dieser Fragebogen dient dazu, mögliche Risiken vor der Untersuchung abzuschätzen und auszuschließen.

1. Haben oder hatten Sie eine der folgenden Krankheiten?

	Ja	Nein
Allgemeine Körperschwäche	<input type="radio"/>	<input type="radio"/>
Bewusstseins- oder Gleichgewichtsstörungen	<input type="radio"/>	<input type="radio"/>
Anfallsleiden (z.B. Epilepsie)	<input type="radio"/>	<input type="radio"/>
Erkrankungen oder Schäden des zentralen oder peripheren Nervensystems	<input type="radio"/>	<input type="radio"/>
Chronischer Alkoholmissbrauch, Betäubungsmittelsucht oder andere Suchtformen	<input type="radio"/>	<input type="radio"/>
Migräne oder Cluster- Kopfschmerzen	<input type="radio"/>	<input type="radio"/>
Starke Rückenschmerzen (z.B. durch Bandscheibenvorfall)	<input type="radio"/>	<input type="radio"/>
Stoffwechselerkrankungen, die die Belastbarkeit stark einschränken	<input type="radio"/>	<input type="radio"/>
Allergie gegen Zwiebeln oder Tomaten	<input type="radio"/>	<input type="radio"/>
Rot-Grün Blindheit	<input type="radio"/>	<input type="radio"/>
Mobilitätseinschränkungen der Arme oder Hände	<input type="radio"/>	<input type="radio"/>
Fehlende Einwilligungsfähigkeit des Probanden	<input type="radio"/>	<input type="radio"/>

2. Bei weiblichen Probanden: Sind Sie schwanger?

Ja Nein

3. Nehmen Sie zurzeit zentral wirksame Medikamente ein?

Ja Nein

4. Besteht eine Medikamentenallergie?

Ja Nein Falls ja, gegen welche

5. Haben Sie akut Kreislauf- oder Atembeschwerden (z. B. Bluthochdruck, Asthma usw.)?

Ja Nein

6. Haben Sie akut Probleme mit Heuschnupfen?

Ja Nein

Ich versichere hiermit die Richtigkeit der Angaben.

Dortmund, den _____

(Unterschrift)

2.2 Allgemeine Instruktion zu Versuchsbeginn

Sehr geehrte(r) Proband(in),

vielen Dank, dass Sie sich bereit erklärt haben, an unserer neuen Studie über Datenbrillen teilzunehmen.

Diese schriftliche Instruktion soll Ihnen eine Orientierung über die Hintergründe, den Versuchsablauf und Ihre Aufgaben geben. Bitte lesen Sie die Instruktion in Ruhe durch. Sollten Sie anschließend noch Fragen haben, wenden Sie sich bitte an den Versuchsleiter.

Vor einiger Zeit haben Sie bereits an einer Studie zum Langzeiteinsatz von Datenbrillen bei uns teilgenommen. Diese Studie hat sehr dazu beigetragen, wichtige Erkenntnisse im Umgang mit dieser neuen Technologie zu gewinnen. Wir haben diese Erkenntnisse aus der ersten Studie umgesetzt und in unseren Versuchsaufbau einfließen lassen. Wir freuen uns, dass Sie nochmal bereit sind, im Rahmen einer vergleichenden Studie bei uns als Proband mitzuwirken.

Diesmal handelt es sich um eine einmalige Erhebung die ungefähr 3 Stunden dauert. Hierbei werden Sie für 2 Stunden mit einer neuen Datenbrille arbeiten, die weitere Stunde ist für die Einweisung und Fragebögen vorgesehen.

Im Rahmen der Studie werden Sie diesmal verschiedene Aufgaben bearbeiten, wobei die Anweisungen immer auf der Datenbrille sind. Die detaillierten Instruktionen bekommen Sie jeweils vor Beginn des Arbeitspaketes, hier wird nur eine kurze Übersicht gegeben:

- Arbeiten aus dem Biologielabor – extrahieren Sie die DNA aus Zwiebeln und Tomaten (Dauer ca. 50 Minuten)
- Fehlersuche – finden Sie den Fehler an einem Schaltschrank (ca. 25 Minuten)
- Lego – schrittweiser Aufbau eines Modells (ca. 25 Minuten)

In der Mitte des Versuches machen wir eine kurze Pause.

Sollten Sie während des Versuches irgendwelche Beschwerden bekommen, teilen Sie dies bitte unmittelbar dem Versuchsleiter mit!

Zwischendurch wird zu verschiedenen Zeitpunkten Ihre Beanspruchung abgefragt. Hierzu wird eine Skala verwendet, die wir gleich einüben werden. Zusätzlich gibt es einen Fragebogen zur visuellen Ermüdung – am Anfang, in der Pause und am Ende. Am Ende des Versuches und in der Pause führen wir weiterhin ein kleines Interview mit Ihnen durch.

Sie werden heute mit der neuen **Datenbrille „Google Glass“** arbeiten. Sie benötigt keinen Kopfträger und wird wie eine normale Brille getragen. Bei vorhandener Brille kann sie darüber getragen werden. Das Display befindet sich rechts oben und ist halbtransparent.

Die Bedienung der Brille geschieht über Sprachbefehle. Die einzelnen Sprachbefehle werden vor Beginn der einzelnen Versuchsblöcke noch erläutert und geübt. Sie können während des Versuches aber auch „OK Glass“ sagen und bekommen dann alle möglichen Sprachbefehle in der Datenbrille angezeigt. Die Spracheingabe funktio-

niert zurzeit nur auf Englisch. Um die Aussprache der einzelnen Begriffe zu üben werden die Sprachbefehle deshalb im Vorfeld von einem Computer vorgelesen.

Die Umsetzung der Sprachbefehle auf dem System kann eine kleine Zeitverzögerung von 1-2 Sekunden haben, dies ist technisch bedingt.

Im Allgemeinen funktioniert die Spracheingabe gut, kann aber unter Umständen auch manche Begriffe nicht erkennen. Sollte ein Sprachbefehl von der Technik nicht verstanden werden, wiederholen Sie ihn bitte. Machen Sie zwischendurch aber ca. 5 Sekunden Pause, da zu schnell nacheinander wiederholte Begriffe von dem System nicht als einzelne Befehle verstanden werden. Sollte das System Probleme haben, Sie zu verstehen, kann der Versuchsleiter die Eingabe auch manuell vornehmen.

Sie werden im Anschluss die Datenbrille vom Versuchsleiter vorgestellt bekommen.

Viel Spaß und gutes Gelingen!

Ihr Laborteam der BAuA ☺

2.3 Instruktion Google Glass zu Beginn der zweiten Stunde

Instruktion für die 2. Stunde:

Datenbrille

Während der nächsten Stunde des Versuches werden Sie eine Fehlersuche an einem Schaltschrank durchführen und wieder Lego bauen, wie in der letzten Studie. Die Arbeitsinhalte bekommen Sie über die Datenbrille angezeigt.

Die Interaktion mit den Arbeitsinhalten erfolgt wieder mit englischen Sprachbefehlen, so wie auch schon in der ersten Stunde des Versuches. Die Befehle haben sich aber verändert und sollen nun zunächst geübt werden. Hierzu verwenden wir beispielhaft ein weiteres Legomodell, da die Begriffe in der Fehlersuche und bei Lego gleich sind.

Die Sprachbefehle lauten:

„**next slide**“ – ruft die nächste Arbeitsfolie auf

„**previous**“ – schaltet eine Arbeitsfolie zurück

„**zoom Image**“ – zoomt in das Bild hinein

(der Zoomausschnitt wird über Kopfdrehung gewählt)

„**scale down**“ – zeigt wieder das Gesamtbild an

„**bar changed**“ – Reaktion auf Balken (nur bei Lego)

Beispielhafte Darstellung der Arbeitsinhalte

blue / middle

Lego Aufbauanleitung (links) und 3 Balken (rechts)

Zu Beginn der Übung, werden wir zunächst für 5 Minuten die Balkenaufgabe trainieren, die Sie bei dem Legoversuch brauchen:

Am Rand neben den Lego-Arbeitsschritten werden **drei Balken** angezeigt, die langsam aber kontinuierlich Ihre Länge verändern und ab und zu die Farbe wechseln. Sie sollen auf 2 Zustandsänderungen reagieren:

- Wenn die Position des längsten Balkens wechselt (rechts-mitte-links) und
- bei einem Farbwechsel der Balken (rot/blau).

Der Sprachbefehl für beide Zustandsänderungen lautet: „**bar changed**“.

Bei den beiden Balkenaufgaben erhalten Sie eine schriftliche Rückmeldung über die letzte von Ihnen bestätigten Position und Farbe. Diese wird in englischer Sprache über den Balken eingeblendet. Am Anfang vor der ersten Reaktion steht hier ein „XXX – XXX“.

Sie bekommen den Sprachbefehl zunächst vom Computer vorgesagt. Sie starten die Erhebung dann, indem Sie das erste Mal „bar changed“ sagen. Die schriftliche Rückmeldung sollte nun den aktuellen Zustand über den Balken anzeigen. Tut sie dies nicht, hat das System Sie nicht verstanden. In diesem Fall heben Sie bitte die Hand und geben dem Versuchsleiter ein Zeichen. Versuchen Sie dann die Sprach-eingabe nach wenigen Sekunden erneut. Wann immer das System Sie nicht verstehen sollte, zeigen Sie es bitte dem Versuchsleiter an.

Die Erhebung der Balkenaufgabe dauert 5 Minuten, vermeiden Sie während dessen jegliche andere Sprachaktivität, da es sonst zu Fehlerkennungen kommen kann.

Im Anschluss werden die anderen Sprachbefehle geübt:

Zum Üben der Aussprache wird der Computer den Begriff immer vorsagen. Versuchen Sie die Aussprache so genau wie möglich zu imitieren. Die Reaktion des Systems kann einige Sekunden verzögert stattfinden.

Bitte halten Sie sich bei der Übung an die vorgegebene Reihenfolge der Begriffe, da wir später auswerten wollen, wie gut das System Sie verstanden hat. Wiederholen Sie die Begriffe nur auf Anweisung des Versuchsleiters und.

Gutes Gelingen!

Ihr Laborteam der BAuA ☺

2.4 Instruktion Tablet-PC zu Beginn der zweiten Stunde

Instruktion für die 2. Stunde:

Tablet

Während der nächsten Stunde des Versuches werden Sie eine Fehlersuche an einem Schaltschrank durchführen und wieder Lego bauen, wie in der letzten Studie. Die Arbeitsinhalte bekommen Sie über einen Tablet-PC angezeigt. Die Interaktion mit den Arbeitsinhalten erfolgt mittels Berührungsgesten.

- „**wischen nach links**“ – ruft die nächste Arbeitsfolie auf
- „**wischen nach rechts**“ – schaltet eine Arbeitsfolie zurück
- „**Doppelklick**“ – Reaktion auf Balken (nur bei Lego)

Beispielhafte Darstellung der Arbeitsinhalte

blue / middle

Lego Aufbauanleitung (links) und 3 Balken (rechts)

Zu Beginn werden wir zunächst für 5 Minuten die Balkenaufgabe trainieren, die Sie bei dem Legoversuch brauchen:

Am Rand neben den Lego-Arbeitsschritten werden **drei Balken** angezeigt, die langsam aber kontinuierlich Ihre Länge verändern und ab und zu die Farbe wechseln. Sie sollen auf 2 Zustandsänderungen reagieren:

- Wenn die Position des längsten Balkens wechselt (rechts-mitte-links) und
- bei einem Farbwechsel der Balken (rot/blau).

Die Bestätigung für beide Zustandsänderungen ist ein Doppelklick auf dem Display. Bei den beiden Balkenaufgaben erhalten Sie eine schriftliche Rückmeldung über die letzte von Ihnen bestätigten Position und Farbe. Diese wird in englischer Sprache über den Balken eingeblendet. Am Anfang vor der ersten Reaktion steht hier ein „XXX – XXX“.

Sie starten die Erhebung dann, indem Sie das erste mal einen Doppelklick vollführen. Die schriftliche Rückmeldung zeigt nun den aktuellen Zustand über den Balken an. Die Erhebung der Balkenaufgabe dauert 5 Minuten.

Im Anschluss werden die anderen Gesten zum Vor- und Zurückschalten der Folien noch geübt.

Gutes Gelingen!

Ihr Laborteam der BAuA ☺

2.5 Instruktion Google Glass direkt vor der Erhebung

Ihre Aufgaben während des Lego-Versuches:

Während des nächsten Teils des Versuches werden Sie Lego bauen, wie in der letzten Studie. Hierbei sollen Sie parallel Balken am Bildschirmrand beobachten und bestimmte Zustandsänderungen rückmelden. Beide Aufgaben sind gleich wichtig.

Bei **Lego-Technik** werden Sie Schritt für Schritt ein komplexes Fahrzeug aus Einzelteilen erstellen. Hierbei werden Ihnen bei jedem Arbeitsschritt in einem oberen Feld die zu verwendenden Teile angezeigt. Im Hauptfeld sehen Sie dann, wie diese Teile an dem Modell zu befestigen sind (*siehe Abbildung unten*). Die Arbeitsschritte sind dabei aufeinander folgend und durchnummeriert. Bei umfangreicheren Arbeitsschritten kann es auch Unternummern geben.

Bedenken Sie, dass Fehler im Aufbau sich an späterer Stelle bemerkbar machen können. Kontrollieren Sie deshalb die exakte Ausführung des Arbeitsschrittes, bevor Sie zum nächsten Schritt blättern. Sollten Sie einen Fehler bemerken, korrigieren Sie ihn oder blättern Sie zurück bis an die entsprechende Stelle in der Aufbauanleitung. Möglicher Weise müssen Sie hierbei auch einige Teile wieder abbauen, um den Fehler beheben zu können. Falls Sie bei einem bestimmten Arbeitsschritt nicht weiter kommen, fragen Sie den Versuchsleiter.

Beispielhafte Darstellung der Arbeitsinhalte

18

middle / blue

Lego Aufbauanleitung (links) und 3 Balken (rechts)

Die Interaktion mit den Arbeitsinhalten erfolgt mit den englischen Sprachbefehlen, die Sie bereits geübt haben:

- „**next slide**“ – ruft die nächste Arbeitsfolie auf
- „**previous slide**“ – schaltet eine Arbeitsfolie zurück
- „**zoom Image**“ – zoomt in das Bild hinein
- „**scale down**“ – zeigt wieder das Gesamtbild an

Parallel zum Aufbau von Lego sollen Sie eine weitere Aufgabe bearbeiten: Am Rand neben den Lego-Arbeitsschritten werden **drei Balken** angezeigt, die langsam aber kontinuierlich Ihre Länge verändern und ab und zu die Farbe wechseln. Sie sollen auf 2 Zustandsänderungen reagieren: Wenn die Position des längsten Balkens wechselt (rechts-mitte-links) und bei einem Farbwechsel der Balken (rot/blau). Der Sprachbefehl für beide Zustandsänderungen lautet:

„**bar changed**“ – Farbwechsel / Wechsel der Position des längsten Balkens

Bei den beiden Balkenaufgaben erhalten Sie eine schriftliche Rückmeldung über die letzte von Ihnen bestätigten Position und Farbe. Diese wird in englischer Sprache über den Balken eingeblendet.

Bitte bearbeiten Sie beide Aufgaben parallel und möglichst schnell und fehlerfrei!

Falls es noch Unklarheiten gibt, oder Sie noch weitere Fragen zu den Aufgaben oder dem Versuchsablauf haben, wenden Sie sich bitte jetzt an den Versuchsleiter.

Ansonsten wünschen wir Ihnen viel Spaß und gutes Gelingen bei Ihren Aufgaben!

Ihr Laborteam der BAuA ☺

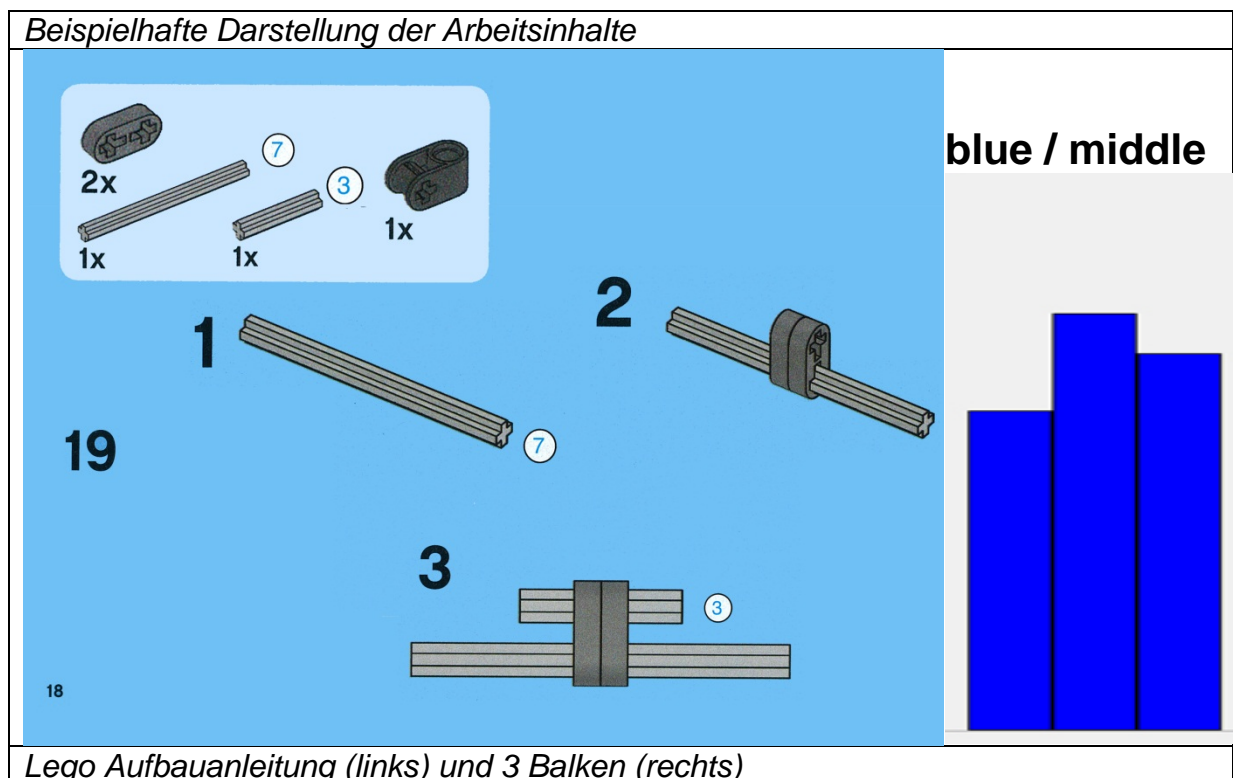
2.6 Instruktion Tablet-PC direkt vor der Erhebung

Ihre Aufgaben während des Lego-Versuches:

Während des nächsten Teils des Versuches werden Sie Lego bauen, wie in der letzten Studie. Hierbei sollen Sie parallel ein paar Balken am Bildschirmrand beobachten und bestimmte Zustandsänderungen rückmelden. Beide Aufgaben sind gleich wichtig.

Bei **Lego-Technik** werden Sie Schritt für Schritt ein komplexes Fahrzeug aus Einzelteilen erstellen. Hierbei werden Ihnen bei jedem Arbeitsschritt in einem oberen Feld die zu verwendenden Teile angezeigt. Im Hauptfeld sehen Sie dann, wie diese Teile an dem Modell zu befestigen sind (*siehe Abbildung unten*). Die Arbeitsschritte sind dabei aufeinander folgend und durchnummeriert. Bei umfangreicheren Arbeitsschritten kann es auch Unternummern geben.

Bedenken Sie, dass Fehler im Aufbau sich an späterer Stelle bemerkbar machen können. Kontrollieren Sie deshalb die exakte Ausführung des Arbeitsschrittes, bevor Sie zum nächsten Schritt blättern. Sollten Sie einen Fehler bemerken, korrigieren Sie ihn oder blättern Sie zurück bis an die entsprechende Stelle in der Aufbauanleitung. Möglicher Weise müssen Sie hierbei auch einige Teile wieder abbauen, um den Fehler beheben zu können. Falls Sie bei einem bestimmten Arbeitsschritt nicht weiter kommen, fragen Sie den Versuchsleiter.



Die Interaktion mit den Arbeitsinhalten erfolgt mittels Berührungsgesten.

„**wischen nach links**“ – ruft die nächste Arbeitsfolie auf

„**wischen nach rechts**“ – schaltet eine Arbeitsfolie zurück

Die Legobausteine müssen diesmal an einer anderen Stelle entnommen werden, als dort, wo der Zusammenbau erfolgt. Dies simuliert eine Arbeitssituation, in der nicht alle Bauteile oder Werkzeuge direkt am Platz gehalten werden können.

Parallel zum Aufbau von Lego sollen Sie eine weitere Aufgabe bearbeiten: Am Rand neben den Lego-Arbeitsschritten werden **drei Balken** angezeigt, die langsam aber kontinuierlich Ihre Länge verändern und ab und zu die Farbe wechseln. Sie sollen auf 2 Zustandsänderungen reagieren: Wenn die Position des längsten Balkens wechselt (rechts-mitte-links) und bei einem Farbwechsel der Balken (rot/blau). Die geforderte Reaktion für beide Zustandsänderungen ist:

„**Doppelclick**“ – Reaktion auf Balken

Bei den beiden Balkenaufgaben erhalten Sie eine schriftliche Rückmeldung über die letzte von Ihnen bestätigten Position und Farbe. Diese wird in englischer Sprache über den Balken eingeblendet.

Bitte bearbeiten Sie beide Aufgaben parallel und möglichst schnell und fehlerfrei!

Falls es noch Unklarheiten gibt, oder Sie noch weitere Fragen zu den Aufgaben oder dem Versuchsablauf haben, wenden Sie sich bitte jetzt an den Versuchsleiter.

Ansonsten wünschen wir Ihnen viel Spaß und gutes Gelingen bei Ihren Aufgaben!

Ihr Laborteam der BAuA ☺

2.7 Interview Google Glass

Interview GG

VPNR:

Datum:

Kamen Sie mit der Konstruktionsaufgabe gut zurecht?

Falls „nein“: Glauben Sie das lag an der Darstellung im Display, oder wäre das auf Papier das gleiche Problem gewesen?

Hatten Sie beim Bearbeiten der Balkenaufgabe eine bestimmte Strategie?

Wie fanden Sie die Darstellung auf nur einem Auge?

Konnten Sie alles scharf sehen?

Wie fanden Sie die Positionierung des Displays?

Wie fanden Sie die Zoom-Funktion?

Wie kamen Sie mit der Spracheingabe zurecht? Hat das System Sie gut verstanden?

Wie waren die Trageeigenschaften des Kopfträgers?

Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden / Sehen bemerkt? Hat sich dies nach der Pause wieder verändert?

Sie haben vor einiger Zeit bereits ein anderes Datenbrillensystem bei uns ausprobiert – wie fanden Sie das eben getestete System im Vergleich?

Wo sehen Sie noch Verbesserungspotential in der eben getesteten Technik?

Welche Einsatzgebiete könnten Sie sich für eine Datenbrille vorstellen?

Haben Sie noch weitere Anregungen oder Hinweise?

2.8 Interview Tablet-PC

Interview GG

VPNR:

Datum:

Kamen Sie mit der Konstruktionsaufgabe gut zurecht?

Falls „nein“: Glauben Sie das lag an der Darstellung im Display, oder wäre das auf Papier das gleiche Problem gewesen?

Hatten Sie beim Bearbeiten der Balkenaufgabe eine bestimmte Strategie?

Wie fanden Sie die Darstellung auf nur einem Auge?

Konnten Sie alles scharf sehen?

Wie fanden Sie die Positionierung des Displays?

Wie fanden Sie die Zoom-Funktion?

Wie kamen Sie mit der Spracheingabe zurecht? Hat das System Sie gut verstanden?

Wie waren die Trageeigenschaften des Kopfträgers?

Haben Sie mit fortlaufender Zeit eine Veränderung in Ihrem Empfinden / Sehen bemerkt? Hat sich dies nach der Pause wieder verändert?

Sie haben vor einiger Zeit bereits ein anderes Datenbrillensystem bei uns ausprobiert – wie fanden Sie das eben getestete System im Vergleich?

Wo sehen Sie noch Verbesserungspotential in der eben getesteten Technik?

Welche Einsatzgebiete könnten Sie sich für eine Datenbrille vorstellen?

Haben Sie noch weitere Anregungen oder Hinweise?