
Pupillometrie als Methode zur Erfassung mentaler Beanspruchungen im automotiven Kontext

Dissertation

zur Erlangung des akademischen Grades eines Doktors der Philosophie
der Philosophischen Fakultäten der Universität des Saarlandes

vorgelegt von Maximilian Schwalm
aus Heidelberg

Saarbrücken 2009

Dekan: Prof. Dr. Rainer Krause

Berichterstatter: 1. Prof. Dr. Hubert Zimmer
2. Prof. Dr. Dirk Wentura

Tag der Disputation: 9 Dezember 2008

Danksagung

Die vorliegende Arbeit ist das Ergebnis meiner Forschungstätigkeit am Forschungs- und Innovationszentrum der BMW Group in München.

Mein besonderer Dank gilt an dieser Stelle meinem Doktorvater und akademischen Lehrer Prof. Dr. Hubert Zimmer, der mich während meiner Promotionszeit stets nach allen Möglichkeiten unterstützt und gefördert hat. Das gegenseitige Vertrauen und der persönliche Kontakt ermöglichten diese Arbeit.

Auch bei Herrn Dr. Andreas Keinath möchte ich mich an dieser Stelle besonders bedanken. Er hat als Betreuer dieser Arbeit vor Ort bei BMW in München einen ganz wesentlichen Teil zu deren Gelingen beigetragen. Bei ihm habe ich immer Rat und Antworten in einem persönlichen Kontakt finden können. Seine Umsicht und seine Unterstützung haben mir ein Arbeiten an dieser Methode überhaupt erst ermöglicht.

Auch möchte ich mich hiermit bei Herrn Dr. Lutz Eckstein für sein entgegengebrachtes Vertrauen bedanken. Er ermöglichte mir mit einer Anstellung als Doktorand bei BMW diese Arbeit anzufertigen. Mein Dank gilt an dieser Stelle auch Herrn Dr. Klaus Bengler. Er stellte mir vertrauensvoll alle für diese Arbeit benötigten technischen sowie räumlichen Mittel zur Verfügung.

Ferner möchte ich mich bei Herrn Prof. Dr. Dirk Wentura für das Interesse und die Zweitbegutachtung meiner Arbeit bedanken.

Und schließlich möchte ich mich bei all jenen bedanken, die mich durch Ihre Teilnahme an den Versuchen, deren Auswertung oder auch deren Vorbereitung unterstützt und gefördert haben. Mein Dank gilt auch jenen die mir bei der Anfertigung dieser Arbeit mit Rat und Tat beiseite standen.

Zusammenfassung

Die vorliegende Arbeit untersucht die Anwendung einer neuen Methode der Pupillometrie zur Erfassung von mentaler Beanspruchung bei der Optimierung von Fahrerinformations- und Fahrerassistenzsystemen im automotiven Kontext. Die in der vorliegenden Arbeit eingesetzte Methode erlaubt es, über charakteristische Veränderungen der Größe der menschlichen Pupille Rückschlüsse auf die aktuelle Beanspruchung eines Fahrers zu ziehen und diese mittels eines Index, dem „Index of Cognitive Activity“ (ICA, Marshall, Davis & Knust, 2004) abzubilden. Es werden insgesamt vier Experimente beschrieben. Es zeigt sich, dass der hier beschriebene Index sowohl in einem experimentell streng kontrollierten Kontext, als auch in realistischen, anwendungsnahen Fahr- und Bediensituationen auftretende mentale Beanspruchungen für den Fahrer abbilden kann. Es konnte auch gezeigt werden, dass es mittels der hier beschriebenen Methode der Pupillometrie und deren hoher zeitlicher Auflösung der Messung erstmals möglich ist, den tatsächlichen Verlauf von Beanspruchungen bei der Bearbeitung einer Aufgabe abzubilden. Die hier beschriebenen Ergebnisse zeigen eine Möglichkeit der Optimierung zukünftiger Mensch-Maschine Schnittstellen hinsichtlich der für den Benutzer auftretenden Beanspruchungen mit einer bis lang unerreichten Genauigkeit.

Abstract

In this study a new method for measuring mental demands is discussed. Its relevance in an automotive context is the optimisation of new driver information as well as driver-communication systems. This method allows to identify mental demands through observing changes in the size of the human pupil and to display the actual demand through an index called ‘Index of Cognitive Activity’ (ICA, Marshall, Davis & Knust, 2004). In all, four experiments are being described. It could be shown that the index described here may identify mental demands not only in a highly controlled experimental context, but also in the context of a realistic and applied driving-and-interaction situation. It was shown for the first time that with this method and its high timescale resolution it is possible to display the actual characteristic course of mental demands while performing a task. The study thus describes a highly effective way of optimising new human-machine interfaces with respect to the occurring mental demands to the user with an accuracy that has not been reached so far.

Zusammenfassung.....	4
Abstract.....	5
Teil 1	10
1 Einleitung	10
2 Zum Begriff der mentalen Beanspruchung	13
2.1 Die limitierte Prozessressource	15
2.2 Zur Definition der mentalen Beanspruchung	19
2.3 Faktoren der Fahrer Beanspruchung	23
3 Messmethoden mentaler Beanspruchung	29
3.1 Subjektive Maße.....	33
3.2 Leistungsmaße.....	36
3.3 Physiologische Maße.....	40
3.3.1 Elektrokardiogramm (EKG).....	40
3.3.2 Elektroencephalogramm (EEG)	43
3.3.3 Elektrodermale Aktivität	45
3.3.4 Weitere psychophysiologischen Maße	46
4 Pupillometrie.....	48
4.1 Physiologische Grundlagen.....	48
4.2 Empirische Evidenzen Pupillometrie und mentale Beanspruchung.....	54
4.3 Index of Cognitive Activity (ICA)	61
4.3.1 Beschreibung des Verfahrens	63
4.3.2 Datenanalyse.....	65
4.3.3 Experimentelle Evidenz.....	71
4.3.4 Zusammenfassung Index of Cognitive Activity.....	75

Teil 2	77
5 Empirischer Teil	77
5.1 Allgemeines	77
5.2 Verwendeter Eyetracker	78
5.3 Datenaufbereitung.....	80
6 Experiment 1: Untersuchung zur grundlegenden Validierung des experimentellen Settings und des ICA	81
6.1 Hintergrund.....	81
6.2 Fragestellung.....	82
6.3 Methode	83
6.3.1 Verwendetes Material	83
6.3.2 Ablauf	84
6.3.3 Stichprobe	85
6.4 Ergebnisse.....	86
6.4.1 ICA.....	86
6.4.2 Leistung.....	88
6.4.3 Reaktionszeiten	89
6.4.4 Subjektive Daten	90
6.5 Zusammenhang subjektive Schwierigkeit und ICA	91
6.6 Diskussion.....	91
7 Experiment 2: Untersuchung zum grundlegenden Nachweis von Beanspruchungseffekten mittels des ICA bei einer simulierten Fahraufgabe	93
7.1 Hintergrund.....	93
7.2 Fragestellung.....	95
7.3 Methode	96
7.3.1 Verwendete Aufgaben	96
7.3.2 Versuchsaufbau.....	99
7.3.3 Ablauf	100
7.3.4 Stichprobe	101

7.4	Ergebnisse	101
7.5	Diskussion	107
8	Experiment 3: Untersuchung zur weiteren Validierung der Methode im Fahrkontext sowie der Nachweis strategischer Beanspruchungseffekte in einer Fahraufgabe	111
8.1	Hintergrund	111
8.2	Fragestellung	116
8.3	Methode.....	119
8.3.1	Verwendete Aufgaben	119
8.3.2	Ablauf	123
8.3.3	Stichprobe.....	125
8.4	Ergebnisse	125
8.4.1	Fahrleistung	125
8.4.2	Index of Cognitive Activity.....	127
8.4.2.1	Auswertung Bedingungen mit Fahraufgabe	128
8.4.2.2	Bedingungen ohne Fahraufgabe	142
8.4.3	Ergebnisse Nebenaufgaben	143
8.4.3.1	Surrogate Reference Task.....	144
8.4.3.2	CTT	148
8.4.4	Subjektive Daten	154
8.5	Diskussion	156
9	Experiment 4: Einsatz der Methode für die Bewertung und Vergleich zweier realer Bediensysteme sowie Interaktionsmodalitäten.....	163
9.1	Hintergrund	164
9.2	Fragestellung	166
9.3	Methode.....	167
9.3.1	Versuchsaufbau	167
9.3.2	Verwendete Aufgaben und Systeme	168
9.3.2.1	Fahraufgabe	168
9.3.2.2	IDrive System	168
9.3.2.3	System HMMI	170

9.3.3	Alphanumerische Eingabe	173
9.3.3.1	Verwendete Bedienelemente.....	173
9.3.3.2	Eingabe über Speller	173
9.3.3.3	Eingabe mittels Handschrifterkennung	175
9.3.4	Beschreibung der Bedienvorgänge	176
9.3.5	Ablauf	177
9.3.6	Stichprobe	178
9.3.7	Erhobene Daten.....	178
9.4	Ergebnisse.....	179
9.4.1	Index of Cognitive Activity	179
9.4.1.1	Ergebnisse ohne Fahraufgabe	183
9.4.1.2	Ergebnisse mit Fahraufgabe	190
9.4.1.3	Zusammenfassung Ergebnisse Pupillometrie	195
9.4.2	Bearbeitungsdauern.....	196
9.4.3	Fahrleistung.....	200
9.4.4	Blickdaten	202
9.4.5	Subjektive Daten.....	208
9.5	Diskussion.....	211
10	Allgemeine Zusammenfassung und Diskussion	216
11	Kritische Diskussion	225
12	Ausblick	228
13	Literatur	230
14	Anhang.....	247

Teil 1

1 Einleitung

Die Entwicklung neuer Fahrerinformations- und Fahrerassistenzsysteme im Automobilbereich steht heutzutage vor einer neuen Herausforderung. Während bis vor einigen Jahren oftmals nur die technische Umsetzbarkeit wie beispielsweise die Verfügbarkeit von bestimmten Informationen für ein entsprechendes Fahrerinformationssystem den entscheidenden Faktor bei der Entwicklung neuer Systeme darstellte, so steht heute mit zunehmenden technischen Möglichkeiten noch ein ganz anderer Faktor bei der Entwicklung solcher Systeme im Vordergrund: Der einer optimalen Bedienbarkeit der Funktion durch den Fahrer. Bei der Komplexität der heutzutage verfügbaren Fahrerinformations- sowie Fahrerassistenzsysteme muss gewährleistet sein, dass diese nicht nur eine maximale Funktionalität aufweisen, sondern auch sicher und effizient vom Fahrer genutzt werden können.

Informations- und Assistenzsysteme, die vom Fahrer selbst bedient werden können oder müssen, erhöhen maßgeblich die operationalen Anforderungen für den Fahrer. Verwendet ein Fahrer diese Systeme während des Fahrens, so muss er seine Aufmerksamkeit zwischen dem von ihm verwendeten System und der eigentlichen Fahraufgabe aufteilen (De Waard, 1996). Um dabei nicht das Risiko für einen Fahrfehler zu erhöhen, muss die durch ein solches System erzeugte zusätzliche Beanspruchung für den Fahrer so gering wie möglich gehalten werden. Die Integration dieser Systeme in das Umfeld des Fahrers muss somit in einer Art und Weise geschehen, dass diese bei ihrer Verwendung nur zu einer minimalen und somit für die Fahrleistung unkritischen Erhöhung der Beanspruchung des Fahrers führen und damit in allen Fahrsituationen sicher und effizient nutzbar sind.

Folgt man nun der Forderung nach einer Minimierung der für den Fahrer auftretender Beanspruchungen während der Nutzung eines solchen Systems, so stellt sich die Frage, wie solche auftretenden Beanspruchungen prinzipiell quantitativ messbar gemacht werden können. Hierzu wurden bereits eine ganze Reihe verschiedener Methoden vorgeschlagen, um die für den Fahrer auftretenden Beanspruchungen bei der Nutzung eines zusätzlichen Informations- oder Assistenzsystems zu messen (Breuer, Bengler, Heinrich & Reichelt, 2003; O'Donnell & Eggemeier, 1986; De Waard, 1996). Diese Verfahren basieren sowohl auf den subjektiven Auskünften des Fahrers über die von ihm empfundene Beanspruchung sowie auf Methoden, die aufgrund der gezeigten Leistung bei der Lösung einer Aufgabe Rückschlüsse auf die auf den Fahrer einwirkenden Beanspruchungen erlauben. Schließlich werden aber auch Methoden angewandt, bei denen mittels Aufzeichnung und Interpretation von physiologischen Reaktionen Rückschlüsse auf den aktuellen Zustand der Beanspruchung des Fahrers möglich sind. All diese in der Literatur diskutierten Methoden bieten unterschiedliche Vor- und Nachteile bei der Erfassung der bei einem Fahrer auftretenden Beanspruchungen und somit der Identifikation von Optimierungspotentialen bei der Entwicklung neuer Fahrerinformations- sowie Fahrerassistenzsysteme.

Im ersten Teil der vorliegenden Arbeit soll zunächst das Konzept der mentalen Beanspruchung aus einer kognitiv psychologischen Sicht diskutiert und beschrieben werden. Anschließend sollen verschiedene Methoden zur Messung von mentaler Beanspruchung mit ihren jeweiligen Vor- und Nachteilen diskutiert und analysiert werden. Darauf folgend wird die Methode der Pupillometrie vorgestellt. Ein physiologisches Verfahren für die Erfassung mentaler Beanspruchung über Reaktionen der menschlichen Pupille. Die Objektivität, die hohe zeitliche Auflösung der Messung, und die in Echtzeit verfügbare Information über den Zustand der Beanspruchung des Fahrers haben diese Methode im Kontext neuer Analysemöglichkeiten in jüngster Zeit in den Fokus einer anwendungsorientierten Forschung im automotiven Kontext gerückt.

Im zweiten Teil der vorliegenden Arbeit wird ein neues Verfahren der Pupillometrie, der „Index of Cognitive Activity“ (ICA) als Methode zur Erfassung der Fahrerbeanspruchung beschrieben und experimentell validiert. Dabei wird die Methode des ICA auf ihre Anwendbarkeit bei der Entwicklung und Optimierung von neuen Anzeige- und Bedienkonzepten im automobilen Bereich überprüft und analysiert.

2 Zum Begriff der mentalen Beanspruchung

Die Begriffe der Belastung und Beanspruchung an sich entstammen der Arbeitspsychologie der 40iger Jahre (Kerkau, 2005). Seither jedoch wurden die Begriffe in verschiedenen Kontexten sowie mit unterschiedlichen inhaltlichen Bedeutungen verwendet (Schönflug, 1987). Neuere Richtungen haben hier verwandte Begriffe wie etwa „mental workload“ oder „mentale Arbeitslast“ hervorgebracht.

Zunächst tauchten insbesondere im Zusammenhang mit einzelnen technischen Systemen, bei denen eine Vielzahl hochspezialisierter Geräte gleichzeitig zum Einsatz kamen, Fragen in Bezug auf die auftretende mentale Beanspruchungen für den Anwender auf. So etwa im Kontext der Flugsicherung, in dem in den vergangenen dreißig Jahren zunehmend Fragen der mentalen Beanspruchung von Piloten adressiert wurden wie etwa: „Wie beschäftigt ist der Pilot“ oder „Wie viele Aufgaben kann der Pilot in einer kritischen Situation sicher und effizient lösen“ (De Waard, 2006). Während diese sicherheitskritischen Fragestellungen zunächst nur einige wenige Systeme betraf, welche allein durch Spezialisten bedient wurden, wurden solche Fragestellungen mit zunehmender Technisierung auch in alltäglicheren und weniger spezialisierten Gebieten zunehmend relevant. Die Forderung einer auf die Anforderungen des menschlichen Benutzers angepasste Schnittstelle zwischen Mensch und technischem System wurde immer wichtiger.

Das wohl am weitesten verbreitete, technisch hoch komplexe System dieser Art stellt heutzutage das Kraftfahrzeug dar. Kaum ein anderes technische System steht einer so breiten Masse von Benutzern aller Schichten und jeden Alters zu Verfügung und wird so häufig und intensiv genutzt. Dabei stellt die zunehmende Verfügbarkeit zusätzlicher durch den Fahrer zu bedienender technischer Systeme, angefangen bei einem Radio, über Navigationssysteme bis hin zu modernen Fahrerassistenzsystemen, einen immer wichtigeren Faktor bei der sicheren und an die Limitierungen des Fahrers angepasste Integration solcher Systeme dar. Eine solchermaßen optimierte Anpassung von Fahrerassistenz- sowie Fahrerinformations-

systeme an die Fähigkeiten des Fahrers setzt zunächst eine genauere konzeptuelle Betrachtung des Beanspruchungsbegriffes sowie der Messbarkeit von Beanspruchung voraus. Im Folgenden soll daher zunächst eine kurze Einführung in die wichtigsten Konzepte und ein Überblick über bestehenden Methoden zur Erfassung mentaler Beanspruchung gegeben werden.

2.1 Die limitierte Prozessressource

Ein wichtiger Baustein bei der Diskussion des Begriffes der mentalen Beanspruchung ist die Beschreibung einer „Ressource“ oder „Kapazität“ für die menschliche Informationsverarbeitung und deren Limitierung. Die Idee einer limitierten Prozessressource findet sich in vielen frühen grundlegenden Theorien der Aufmerksamkeit (Broadbent, 1958; Kahnemann, 1973; Posner 1978; Wickens, 1984). Gemeint ist damit eine Metapher wie sie Kahnemann 1973 beschrieben hat. In diesem Modell wird zugrunde gelegt, dass die menschliche Informationsverarbeitung Ressourcen benötigt, die nur in begrenztem Maße zur Verfügung stehen. Eine Verteilung der zur Verfügung stehenden Ressourcen geschieht je nach Aufgabenschwierigkeit und der Motivation die einzelnen Aufgaben zu lösen. Kahnemann definiert Beanspruchung in diesem Zusammenhang als „[...] the proportion of the capacity an operator spends on task performance.“ (Kahnemann, 1973). Die Auslastung dieser einen zentralen Prozessressource kann, nach Kahnemann, mit einem geeigneten Maß abgebildet werden. So schlägt Kahnemann bereits Anfang der 70iger Jahre die Validierung dieses Modells mithilfe physiologischer Maße und insbesondere der Pupillometrie vor (Kahnemann, 1973).

Eine Weiterentwicklung des Ressourcenbegriffes stellt die Multiple Ressourcentheorie (MRT) nach Wickens dar (Wickens & Hollands, 2000). Die MRT Theorie beschreibt eine orthogonale Trennung der mentalen Prozesse zwischen Perzeption und Antwortauswahl, sowie zwischen den Modalitäten (auditiv / visuell) und dem Code der Information (räumlich / verbal). Diese Struktur der Prozessressourcen lässt sich nach Wickens wie in der folgenden Abbildung darstellen:

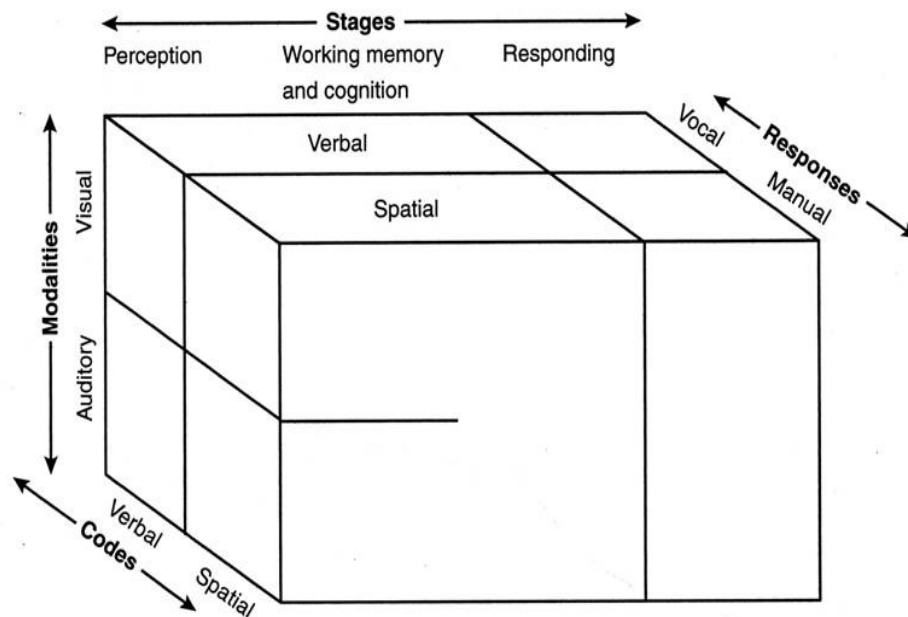


Abbildung 1: Das MRT Modell nach Wickens und seine postulierte Struktur der Prozesskomponenten (Wickens & Hollands, 2000).

Evidenz für eine solche Konzeption von Prozessressourcen sieht Wickens in der Tatsache, dass es einfacher ist, Aufmerksamkeit zwischen zwei verschiedenen Modalitäten wie zum Beispiel Auge und Ohr zu trennen, als innerhalb eines modalitätsspezifischen Kanals. Verteilt auf zwei unterschiedliche Modalitäten wird ein „time-sharing“ für zwei parallel durchgeführte Aufgaben deutlich effizienter. So kann gezeigt werden, dass durch die Darbietung von Informationen mittels crossmodaler Displays (also Displays, welche relevante Informationen auf mehrere Modalitäten verteilt darbieten können) eine deutliche Performanzsteigerung gegenüber uni-modalen Displays erreicht (Wickens & Hollands, 2000).

Auch wenn ein Modell wie das MRT nach Wickens zunächst eine vergleichsweise intuitive Beschreibung der psychologischen Prozesse zu bieten scheint und in vielen aktuellen Ansätzen zur Usability Untersuchungen sowie der Entwicklung von neuen Konzepten zu einer Mensch-Maschine-Interaktion (MMI) zu finden ist, welche sich auf genau diese Konzeption der Prozessressourcen berufen, zeigt sich in neueren Studien, dass ein solches multimodales Modell der Prozessressourcen

immer noch zu einfach ist, um alle beobachtbaren Prozesse beschreiben zu können (Haigney & Westermann, 2001). So deuten Untersuchungen nach Spence und Driver (Spence & Driver, 1994, 2003 Calvert, Spence & Stein, 2004; Driver & Spence, 2004; Driver & Spence, 1998, Spence & Walton, 2005) darauf hin, dass die Effizienz der Performanz bei einer Aufgabe mit mehreren Informationsquellen in unterschiedlichen Modalitäten durch die relative räumliche Trennung der Informationsquellen beeinflusst werden kann (Spence & Read, 2003; Spence & Driver, 1994). Konkret besagen diese Studien, dass im Fall, dass etwa die Aufmerksamkeit der visuellen Modalität auf eine bestimmte Position im Raum gerichtet ist, die auditive Aufmerksamkeit eine Ausrichtung auf dieselbe Position im Raum erfährt. So sollten in einer experimentellen Situation zwei auditive Sprachstreams überwacht werden, während gleichzeitig eine Quelle sich schnell ändernder (von auditiver Information unabhängigen) visuellen Information überwacht werden sollte (Spence & Driver, 1994). Die Ergebnisse zeigen, dass hier das Überwachen der relevanten auditiven Quelle schwieriger ist, wenn die auditive und visuelle Information von unterschiedlichen räumlichen Positionen aus präsentiert wurden.

Spence und Read diskutieren diese Ergebnisse so, dass es neben modalitätsspezifischen Pools der Verarbeitungskapazitäten auf Ebenen höherer koordinativer kognitiven Funktionen modalitätsübergreifende Verbindungen geben muss, die beispielsweise aufgrund einer modalitätsübergreifenden Konstruktion des mentalen Raumes zu den Effekten wie oben beschrieben führen. Beanspruchung wird somit nicht als ein modalitätsspezifisches Konstrukt gesehen, sondern vielmehr als ein in höheren ressourcenübergreifenden koordinativen Systemen wirkende Limitierung.

Aber unabhängig davon, ob man Ressourcen als modalitätsspezifisch oder aber als crossmodales Konstrukt betrachtet, so bleibt jedoch ein Gedanke in allen Theorien der gleiche: die menschliche kognitive Leistungsfähigkeit bei der Verarbeitung

von Informationen ist begrenzt. Die Beanspruchung einer solchen begrenzten mentalen Ressource und die Messbarkeit einer solchen mentalen Beanspruchung stehen im Fokus der vorliegenden Arbeit und sollen im Folgenden genauer beschrieben werden.

2.2 Zur Definition der mentalen Beanspruchung

Neben der Diskussion, welche mentalen Ressourcen für eine Verarbeitung von Informationen zur Verfügung stehen und wie diese in einem einheitlichen psychologischen Modell beschrieben werden können, muss an dieser Stelle auch diskutiert werden, was unter dem Begriff „Beanspruchung“ aus einer definitorischen Sichtweise heraus zu verstehen ist. Eine einfache Definition von Beanspruchung ist die, dass es sich hierbei um die Gesamtheit aller Anforderungen handelt, die zu einem bestimmten Zeitpunkt an das Individuum gestellt werden.

Jedoch attribuiert eine solche Definition Beanspruchung einzig auf den Einfluss externer Faktoren (De Waard, 1996), die keine Aussage darüber machen, wie dieser Einfluss auf das einzelne Individuum wirkt. Es erscheint daher notwendig, Beanspruchung auch unter dem Aspekt des subjektiven Erlebens zu betrachten (Rouse et al., 1993; De Waard, 1996; Staal, 2004).

Bei der Definition des Begriffes „mentale Beanspruchung“ sind daher zwei Aspekte zu unterscheiden (De Waard, 1996; Staal, 2004; Hilburn & Jorna, 2001):

- **Task-Load:** das heißt die Anforderung, die durch die Aufgabe an sich erzeugt wird. Diese ist external bestimmt und nicht durch die Person beeinflusst.
- **Work-Load:** diese beschreibt die Prozesse, die die externalen Anforderungen (Task-Load) der Aufgabe bei dem Individuum selbst erzeugt. Sie bildet somit die ganzheitliche subjektive Erfahrung des Individuums mit den Anforderungen der Aufgabe, sowie die verwendeten Strategien, persönliche Ziele und Motivationen der Person ab.

Hilburn und Jorna (2001) beschreiben diese Trennung in zwei Faktoren am Beispiel einer Air Traffic Control (ATC) Situation in folgendem Modell:

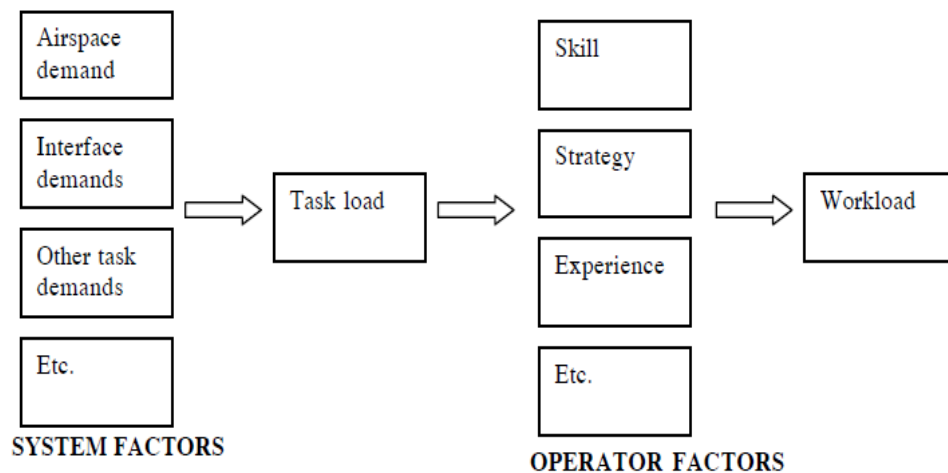


Abbildung 2: Die Abbildung zeigt das Modell nach Hilburn und Jorna (2001, p. 386).

Das Modell zeigt eine Trennung zwischen den Anforderungen seitens des Systems (System Factors) sowie den Faktoren, die die Eigenschaften des Benutzers selbst betreffen (Operator Factors). Das Produkt aus den Anforderungen der Umwelt und den Möglichkeiten, die das Individuum mitbringt auf diese Anforderungen zu reagieren, ergibt die letztendlich erlebte Beanspruchung (Workload).

Eine ähnliche Begriffsbildung findet sich bei Parasuraman und Hancock (2001): „Workload may be driven by the task load imposed on human operators from external environmental sources but not deterministically so, because workload is also mediated by the individual response of human operators to the load and their skill levels, task management strategies, and other personal characteristics.” (S. 306).

Bezieht man diese beiden Aspekte in die Definition von Beanspruchung mit ein, so wird deutlich, dass die Höhe der Beanspruchung von der Aufgabe an sich und dem Individuum selbst, welches die Aufgabe löst, abhängig ist. Daraus folgt aber auch, dass die gleichen Aufgabenanforderungen bei verschiedenen Personen zu unterschiedlich hohen Beanspruchungen führen können. Verdeutlicht wird dieser

Zusammenhang bei De Waard (1996) anhand einer Mathematik Prüfung. Hier gibt es zum einen die Anforderungen, die die mathematischen Fragen selbst stellen (task load). Diese sind dieselben für jeden Teilnehmer und beispielsweise davon abhängig wie viele Berechnungsschritte notwendig sind etc. Den anderen Aspekt stellen die Ziele jedes einzelnen Individuums in dieser Prüfung dar (nur bestehen vs. gute Note) sowie die individuellen Fähigkeiten, die Aufgaben zu lösen (Erfahrung sowie Können). Die auf ein Individuum wirkende Beanspruchung in dieser Situation ist somit ein Ergebnis aus den externen Anforderungen, die die Situation an das Individuum stellt, und den Prozessen, die bei dem Individuum durch diese externen Anforderungen ausgelöst werden.

Eine solche definitorische Betrachtung des Beanspruchungsbegriffes zeigt deutliche Parallelen zu einem weiteren eng verwandten psychologischen Konstrukt auf, das in diesem Zusammenhang ebenfalls kurz diskutiert werden muss: Stress. Transaktionale Konzeptionen des Stressbegriffes sehen Stress als ein Produkt der Interaktion zwischen Umgebung und Individuum, „[...] the result of a mismatch between individuals' perception of the demands of the task or situation and their perception of the resources for coping with them.“ (Strokes & Kite, 2001; S. 116). Stress stellt hier einen weiteren wichtigen Faktor dar, der untrennbar mit dem zweiten, individuellen Anteil der Beanspruchung verknüpft ist und psychologische Prozesse beschreibt, die die Bindung von Prozessressourcen für psychologische (emotionale) Prozesse beinhalten.

Eine Definition von Beanspruchung, die an die im vorangegangenen Abschnitt diskutierten Ressourcentheorien anknüpft und die in diesem Abschnitt diskutierten Ideen aufgreift, geben O'Donnell und Eggemeier (1986). Sie definieren Beanspruchung zum einen aus der Anforderung, die eine bestimmte Aufgabe an die limitierten Prozessressourcen der Person stellen, sowie den Anteil dieser Ressourcen, die die handelnde Person tatsächlich bereit und in der Lage ist für diese Aufgabe einzusetzen. In dieser Definition spiegelt sich sowohl die Idee einer ressour-

cen-gebundenen Prozesskapazität, als auch die Idee interindividueller Faktoren wider, die den Einsatz dieser Kapazitäten moderieren.

Diese Definition fasst die bereits eingeführten Konzepte zusammen, nennt aber auch einen weiteren wichtigen Punkt, der die Brücke schlägt zwischen den zur Verfügung stehenden Kapazitäten und den interindividuellen Faktoren, die den Einsatz dieser Kapazitäten bestimmen: Anstrengung (engl. effort). Unter Anstrengung ist dabei der freiwillige Einsatz von zur Verfügung stehenden Ressourcen bei der Lösung einer Aufgabe zu verstehen (Staal, 2004). Die investierte Anstrengung hängt dabei nicht nur von den Anforderungen der Aufgabe selbst ab, sondern stark von inneren Zielen und Strategien, sowie der Übung, Erfahrung und dem Zustand der Person (Vincente et. al., 1987).

Aus der vorangegangenen Diskussion wird deutlich, dass es sich bei dem Begriff der mentalen Beanspruchung um ein multidimensionales Konstrukt handelt (Jex, 1988) bei dem viele Faktoren – die wichtigsten darunter die Anforderungen der Aufgabe, sowie die Möglichkeiten, Strategien und Motivationen diesen Anforderungen zu begegnen – untrennbar über Konzepte wie beispielsweise Anstrengung miteinander verknüpft sind. Eine genaue empirische Trennung dieser Faktoren in einer komplexen, anwendungsorientierten Situation erscheint somit weder möglich noch zielführend (Staal, 2004).

Ein entsprechendes Messverfahren der mentalen Beanspruchung sollte daher eine möglichst umfassende Messung gestatten, die die Gesamtheit der wirkenden Faktoren auf die Beanspruchung abzubilden vermag.

2.3 Faktoren der Fahrer Beanspruchung

Auch wenn Autofahren allgemein als eine vergleichsweise einfache Tätigkeit betrachtet wird, so stellt sie doch eine Situation dar, in der viele unterschiedliche Faktoren in verschiedenen Modalitäten auf den Fahrer einwirken. Der Fahrer muss all diese Informationen in ein entsprechendes Modell seiner Umwelt integrieren und sich entsprechend diesem Modell adäquat verhalten. Diese Integration wird insbesondere dann kritisch, wenn der Fahrer an die Grenze seiner Kapazität für die Aufnahme und Verarbeitung solcher Informationen stößt. Daraus kann eine Überbeanspruchung resultieren, so dass fahrrelevante Entscheidungen nicht mehr schnell und sicher getroffen werden können. Die Wahrscheinlichkeit für Fahrfehler und Unfälle erhöht sich dadurch. Aufgrund einer solchen sicherheitskritischen Limitierung der Informationsverarbeitung des Fahrers, ist es von besonderem Interesse, solche Quellen der Beanspruchung frühzeitig zu identifizieren und möglichst zu vermeiden.

Beim Autofahren handelt es sich um eine dynamische Kontrollsituation in einer sich kontinuierlich verändernden Umgebung (De Waard, 1996). Dabei ist das Autofahren eine Situation, die nicht nur durch den Fahrer selbst bestimmt wird, sondern auch durch das Verhalten anderer Verkehrsteilnehmer. Michon (1985) beschreibt Autofahren als eine komplexe Aufgabe, die in mindestens drei hierarchische Ebenen unterteilt werden kann. Es wird zunächst eine strategische Ebene beschrieben, auf der strategische Entscheidungen in Bezug auf das Ziel der Fahrt, sowie der zu wählenden Route getroffen werden (strategical Plans). Auf der zweiten Ebene, dem sogenannten „manoeuvring level“, geschehen die Reaktionen auf lokale Situationen, sowie die Reaktionen auf das Verhalten anderer Verkehrsteilnehmer (Controlled action patterns). Auf der untersten Ebene, dem „control level“, werden die grundlegenden Kontrollprozesse wie etwa das Halten der Spur ausgeführt (automatic action pattern). Eine ähnliche Unterscheidung in

drei Ebenen der operationalen Kontrolle findet sich auch in anderen Arbeiten wie beispielsweise bei Van der Molen & Böttcher (1988).

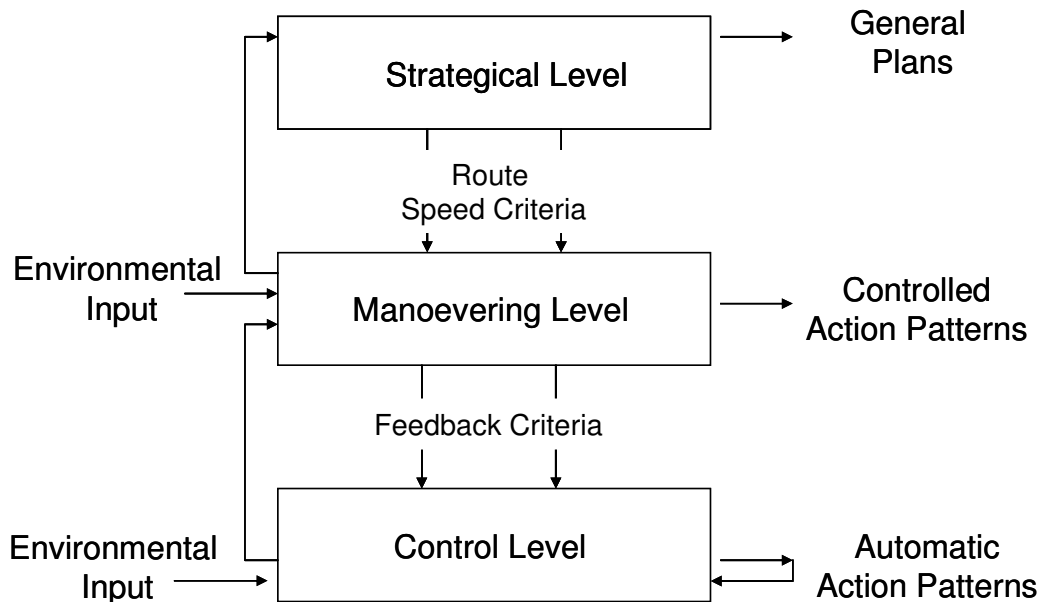


Abbildung 3: Die Abbildung zeigt das hierarchische Modell der Fahrerbeanspruchung nach Michon (1985)

Nach diesem Modell können Anforderungen auf jeder dieser Ebenen direkt zu einer Beeinträchtigung der Fahrleistung führen. Darüber hinaus stehen die einzelnen Ebenen aber auch miteinander in Verbindung. Ein Fahranfänger beispielsweise wird auf der Kontrollebene beim Schalten der Gänge noch so viel Beanspruchung erleben, dass andere höhere Prozesse wie etwa der Blick in den Spiegel beim Abbiegen (manoevering level) vernachlässigt werden. Wie in dem in Abschnitt 2.2 diskutierten Beanspruchungsmodell von Hilburn und Jorna wird auch in dem Modell von Michon zwischen System- und Benutzerfaktoren bei der Charakterisierung der Beanspruchung unterschieden.

Eine direkte Übertragung des Beanspruchungsmodells von Hilburn und Jorna auf die Situation des Autofahrens zeigt Abbildung 4. Das Modell beinhaltet hier zusätzlich eine rekursive Beeinflussung der Systemfaktoren durch die Benutzerfak-

toren. Darunter ist zu verstehen, dass bestimmte Benutzerfaktoren wie etwa die strategische Planung des Fahrers auch eine bestimmte Selektion der Systemfaktoren beeinflussen kann. So bestimmt zum Beispiel auch die Erfahrung und strategische Planung des Fahrers, in welche Verkehrssituationen der Fahrer kommen wird und welche er bereits im Vornhinein vermeiden kann.

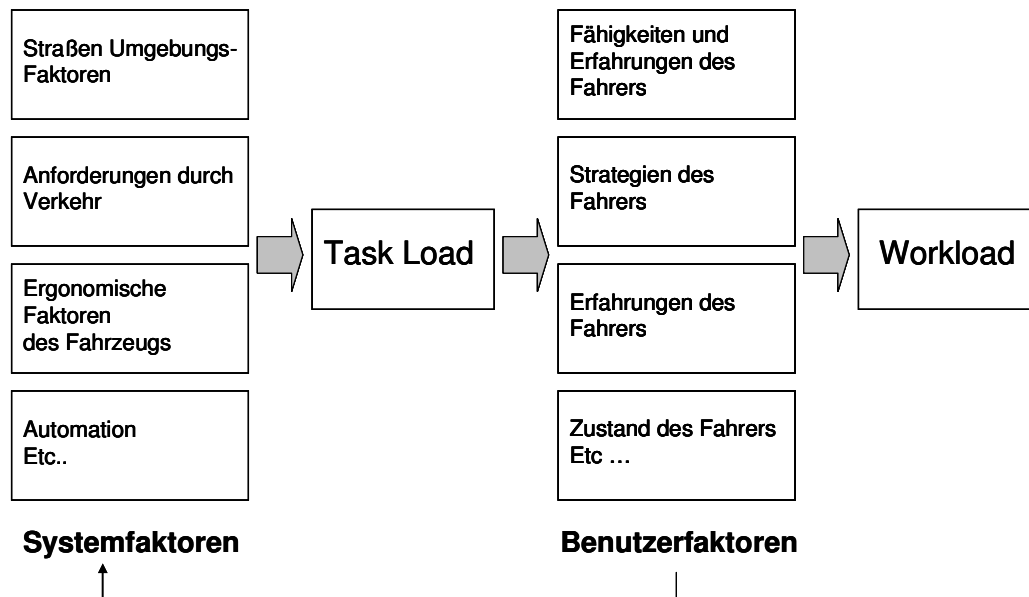


Abbildung 4: Die Abbildung zeigt die Übertragung des Beanspruchungsmodells nach Hilburn & Jorna mit Systemfaktoren und Benutzerfaktoren auf die Situation des Autofahrens.

Bei beiden Modellen wird deutlich, dass die bei dem Fahrer auftretende Beanspruchung ein Produkt aus einer Vielzahl interagierender Faktoren ist. Zunächst produzieren Systemfaktoren wie etwa das Verhalten anderer Verkehrsteilnehmer oder ergonomische Faktoren im Fahrzeug eine Aufgabe, die – durch die Benutzerfaktoren moderiert - zu einer Beanspruchung des Fahrers führt. Zusätzlich beeinflussen Benutzerfaktoren, wie etwa die Erfahrungen oder Strategien des Fahrers, wie stark bestimmte Systemfaktoren, z.B. Anforderungen durch den Verkehr, zu der aufgabenspezifischen Beanspruchung beitragen. Während das Modell nach Michon auf einer allgemeinen strategischen Ebene das Verhalten des Fahrers zu beschreiben versucht, zeigt das zuvor in Abbildung 4 skizzierte Modell in Anlehn-

nung an Hilburn & Jorna eher ein ganzheitliches, auf einzelnen Komponenten beruhendes kognitives Modell der Fahrerbeanspruchung, welches strategische Elemente seitens des Fahrers nur als einen Teilaspekt integriert.

Bei der Betrachtung von Theorien zur Fahrerbeanspruchung lässt sich insgesamt festhalten, dass auch wenn die zugrunde liegenden Prozesse wie die modalitätsspezifischen Anforderungen seitens der Aufgabe, sowie die zur Verfügung stehenden Verarbeitungskapazitäten seitens der Person zwar theoretisch trennbar sind, sie es jedoch empirisch, in einer realen Situation, nur in den seltensten Fällen sind.

In einer realistischen, komplexen anwendungsnahen Situation wie dem Autofahren müssen Informationen auf einer höheren kognitiv - koordinativen Ebene zusammengefügt und in eine effiziente Handlungsplanung umgesetzt werden. Hier fließen alle relevanten Informationen zusammen die die Aufgabe betreffen (System- und Benutzerfaktoren), und bestimmen somit gemeinsam die resultierende Beanspruchung. Es erscheint daher fraglich, ob hier eine vollständige empirische Trennung dieser Faktoren möglich bzw. zielführend ist. Dieses ist ein wichtiger Punkt bei der Betrachtung geeigneter Messverfahren für die mentale Beanspruchung in einem anwendungsorientierten Kontext. Hier erscheint es sinnvoll - ja sogar notwendig - diese generelle, multifaktorielle Beanspruchung des Fahrers in ihrer Gesamtheit abzubilden.

Ein weiterer wichtiger und besonders interessanter Punkt bei der Diskussion von Messverfahren für mentale Beanspruchung, insbesondere in einer solch komplexen Situation wie der das Autofahrens, ergibt sich aus der Diskussion der bereits zuvor erwähnten internen Strategien des Fahrers bei der Bewältigung bestimmter Aufgaben. Eine solche strategische Intervention seitens des Fahrers kann und wird maßgeblich die Höhe der auf den Fahrer wirkenden Beanspruchung beeinflussen. Insbesondere in einer Situation bei der beispielsweise der Fahrer seine ihm zur

Verfügung stehenden Ressourcen auf die Fahraufgabe und die Bearbeitung einer zusätzlichen Nebenaufgabe während des Fahrens aufteilen muss, kann er dies auf sehr unterschiedliche Weise tun. In der Literatur existieren bislang nur wenige umfassende Theorien um ein solches strategisches Verhalten in einer so komplexen Situation, wie der des Autofahrens, genauer beschreiben zu können (s. auch Platten, 2008). Einen Ansatz liefert dazu das kognitiv-energetische Rahmenmodell nach Hockey (1993). Hockey beschreibt hier prinzipiell drei Möglichkeiten um mit einer Erhöhung externaler Anforderungen umzugehen – wie etwa in einer Situation in der der Fahrer zusätzlich zur Fahraufgabe eine Nebenaufgabe bearbeiten muss. Zum einen kann der Fahrer einfach - wie bereits unter Kahnemann (1973) beschrieben - die Höhe der investierten Anstrengung verändern. Er stellt einfach mehr Kapazitäten für die Lösung der Aufgaben bereit. Eine solche Erhöhung der Anstrengung wirkt nach Hockey jedoch prinzipiell aversiv und wird versucht zu vermeiden. Eine weitere Möglichkeit auf eine Erhöhung der situativen bzw. externalen Anforderungen zu reagieren, sieht Hockey in der Vernachlässigung untergeordneter Aufgaben. Gemeint ist damit, dass im Beispiel der Fahrsituation der Fahrer die zu bewältigenden Aufgaben priorisiert. Er reduziert daraufhin die Leistung in den niedriger priorisierten Aufgaben um – unter Gleichhaltung der momentanen Beanspruchung – mehr Kapazitäten für die höher priorisierten Aufgaben zu Verfügung stellen zu können. Eine dritte Möglichkeit der Beanspruchungskompensation nach Hockey besteht darin, bei der Bearbeitung einer Aufgabe in eine effizientere Bearbeitungsstrategie zu wechseln, sofern eine solche verfügbar ist. Gemeint ist damit eine Strategie, die unter Einsatz weniger Ressourcen die gleiche Leistung in einer Aufgabe ermöglicht. Fasst man die soeben genannten Punkte zusammen, so wird deutlich, dass der Fahrer durchaus unterschiedliche Möglichkeiten hat, auf die auf ihn wirkenden Anforderungen mit einer entsprechenden Strategie zu reagieren. So kann es beispielsweise sein, dass der Fahrer zwar die gleiche Leistung in einer Aufgabe zeigt, aber dennoch durch die Bearbeitung einer weiteren Aufgabe eine höhere mentale Beanspruchung zu be-

wältigen hat. Ebenso ist es möglich, dass der Fahrer zwar eine schlechtere Leistung in einer Aufgabe zeigt, dabei aber weniger beansprucht wird.

Welcher dieser Ansätze bei der Entwicklung und Konzeptualisierung neuer Anzeige- und Bedienkonzepte in welcher Situation zu bevorzugen ist, steht nicht im Fokus dieser Arbeit. Dennoch stellen die hier beschriebenen strategischen Effekte der Beanspruchungssteuerung einen wichtigen Baustein für das Verständnis und damit auch der Interpretation gemessener Beanspruchungen in einer solchen Situation dar. Im Verlauf der vorliegenden Arbeit sollen daher die gewonnenen Ergebnisse auch in Hinblick einer solchen strategischen Beanspruchungssteuerung diskutiert und interpretiert werden.

3 Messmethoden mentaler Beanspruchung

Im folgenden Kapitel soll eine Aufstellung, Diskussion und Bewertung der bekannten empirischen Messverfahren für mentale Beanspruchung gegeben werden. Die Diskussion analytischer Verfahren (Lysaght et al. 1989) wie etwa Expertenurteile, mathematische Modelle oder Simulationen stehen hier nicht im Fokus des Interesses. Weitere Informationen dazu können jedoch der Arbeit von Lysaght et al (1989) entnommen werden.

Empirische Messverfahren werden generell in der Literatur in drei Gruppen unterschieden: subjektive Maße, Leistungsmaße (primary und secondary task measures) und physiologische Maße (Yan & Amersing, 2005). Im Folgenden soll jede der drei Gruppen einzeln betrachtet werden. Ein besonderes Augenmerk soll dabei auf der Gruppe der physiologischen Maße liegen, zu der auch die in dieser Arbeit angewandte Methode der Pupillometrie zu zählen ist, die dann im 4. Kapitel ausführlich diskutiert werden wird.

Im Folgenden soll nun zunächst in Anlehnung an O'Donnel und Eggermeier (1986) eine systematische Betrachtung der wichtigsten methodischen Aspekte solcher Verfahren gegeben werden.

Sensitivität

Mit Sensitivität wird die Empfindlichkeit der Methode bezeichnet, mit der sie Unterschiede in der mentalen Beanspruchung detektieren kann. Eng damit verknüpft ist auch die Frage danach, in welchem Bereich der Beanspruchung eine Methode Unterschiede auflösen kann bzw. unter welchen Aspekten die Methode sensitiv auf Beanspruchung reagiert.

Selektivität

Unter Selektivität versteht man die Sensitivität der Methode in Hinblick auf mentale Beanspruchung, d.h. inwieweit die Methode unabhängig von anderen Faktoren, wie beispielsweise physischer Beanspruchung, ist. Je weniger das Ergebnis einer Methode auch von solchen anderen Faktoren abhängig ist, desto valider kann es den tatsächlichen Einfluss mentaler Beanspruchung auf eine Situation wiedergeben.

Diagnostizität

Unter Diagnostizität versteht man das Potenzial der Methode, die unterschiedlichen Anforderungen an einzelne Prozesskapazitäten / Prozesskomponenten zu identifizieren, das heißt in wie weit die Methode Aufschluss über die der mentalen Beanspruchung zugrunde liegenden Faktoren im Sinne einer nach Wickens modalitätsspezifischen Auslastung der Ressourcen geben kann.

In Anlehnung an die zuvor geführte Diskussion der Konzeption von Beanspruchungen in einer anwendungsnahen Situation ist hier jedoch zu hinterfragen, in wie weit ein in diesem Sinn hoch diagnostisches Messinstrument überhaupt sinnvoll ist. Nach Spence & Read ist in einer anwendungsnahen Situation aufgrund auftretender crossmodaler Interaktionen eine Trennung nicht möglich bzw. sinnvoll. Ist der Nachweis einer generellen mentalen Beanspruchung in einer realistischen Anwendungssituation gefordert, so ist Diagnostizität sicherlich nicht das wichtigste Kriterium (De Waard, 1996, S. 28).

Primary-task intrusion

Ein wichtiger Punkt bei der Diskussion von Messmethoden und ihrer Einsetzbarkeit ist der, wie sehr eine Messung den eigentlichen Handlungsablauf in einer Situation stört. Da eine Unterbrechung des Handlungsablaufes durch die Messung eine zusätzliche Beanspruchung für den Probanden bedeutet, sollte es das Ziel jeder Meßmethode sein, diese zusätzliche durch die Messung hervorgerufene Beanspruchung so gering wie möglich zu halten. Secondary-task Methoden, bei der

die Beanspruchung durch die Leistung in einer künstlichen Zweitaufgabe bestimmt wird, nehmen unter Umständen auch Einfluss auf das Verhalten in der eigentlichen Hauptaufgabe (Eggemeier et al. 1991); nicht selten werden hier durch eine solche zusätzliche Aufgabe Störeffekte erzeugt, die den gesuchten Effekt bei weitem übersteigen.

Experimentelle Anforderungen

Mit den experimentellen Anforderungen sind hier praktische Beschränkungen einer Methode gemeint, wie etwa die Anforderungen an ein bestimmtes technisches Equipment oder ein notwendiges technisches Wissen. Bestimmte technisch sehr aufwendige Methoden, wie etwa die klassischen bildgebenden Verfahren (MRT, PET), sind in einem anwendungsnahen Kontext wie etwa Autofahren nicht denkbar. Ein sehr umfangreiches Equipment kann teilweise aber auch den Versuchsablauf erheblich stören oder auf die Versuchsperson ablenkend wirken. Unter den experimentellen Anforderungen einer Messmethode versteht man aber auch beispielsweise den Trainingsaufwand pro Versuchsperson oder ähnliches.

Bei der Diskussion der Charakteristika von Messverfahren für mentale Beanspruchung wird deutlich, dass je nach Kontext und je nach Gegenstand der Untersuchung entschieden werden muss, welches Messverfahren am geeignetsten ist. In einem anwendungsnahen Kontext, wie beispielsweise dem Autofahren, ist aufgrund der in dem vorangegangenen Kapitel geführte Diskussion, eine Methode zu wählen, die einerseits eine hohe Sensitivität als auch eine möglichst geringe Störung der Hauptaufgabe (dem Fahren) verursacht. Gleichzeitig müssen die experimentellen Anforderungen in einem anwendungsnahen Kontext realisierbar sein. Eine geringere Rolle würden die Eigenschaften Selektivität und Diagnostizität spielen, da eine solche Situation zwangsläufig eine Konfundierung der mentalen Beanspruchung mit anderen Faktoren, sowie eine Interaktion verschiedener Prozesskapazitäten bedeutet.

Im Folgenden soll nun eine kurze Aufstellung der wichtigsten in der Literatur diskutierten Messmethoden für mentale Beanspruchung gegeben werden. Dabei werden sowohl Vor- als auch Nachteile der jeweiligen Methoden auf dem Hintergrund der bestehenden empirischen Befundlage kurz beschrieben.

3.1 Subjektive Maße

Subjektive Maße oder auch Maße der Erlebensebene basieren auf der Annahme, dass Personen in der Lage sind, die Schwierigkeit einer Aufgabe und die Wahrnehmung des Ausmaßes mentaler Beanspruchung oder auch die Grenzen ihrer Informationsverarbeitungskapazität zu beurteilen (Rößger, 1996). Subjektive Maße waren und sind immer noch eines der am häufigsten verwendeten Instrumentarien bei der Messung mentaler Beanspruchung. Ein Grund hierfür liegt in der Annahme, dass niemand außer der betreffenden Person selbst Angaben über die subjektiv erlebte Beanspruchung bei einer Aufgabe überhaupt treffen kann (face validity). Manche Autoren sehen somit in den subjektiven Maßen das valideste Kriterium für die Messung von mentaler Beanspruchung „[...] subjective ratings may come closest to tapping the essence of mental workload and provide the most generally valid and sensitiv indicator.“ (Hart & Staveland, 1988). Weitere Vorteile subjektiver Messverfahren liegen darin, dass subjektive Maße einfach, schnell und kostengünstig und in hoher Zahl verfügbar gemacht werden können.

Allerdings gibt es auch die subjektiven Maße betreffende Nachteile, die die Nutzbarkeit der Ergebnisse in bestimmten Situationen in Frage stellen. Zunächst ist es mit subjektiven Maßen nicht möglich, einen momentanen Beanspruchungszustand festzuhalten, da das Durchführen der Befragung den Aufgabenablauf unterbricht und somit selbst wieder Beanspruchung aufwirft und damit eine Messung verfälschen kann. Wie bereits weiter oben diskutiert, basieren subjektive Maße auf der fraglichen Annahme, dass Personen das Ausmaß der Beanspruchung der sie ausgesetzt waren, immer richtig beurteilen können. So können beispielsweise einzelne Beanspruchungsspitzen oder andere Events während der Aufgabenbearbeitung die eigentliche finale Bewertung stark beeinflussen oder verzerren (Vidulich & Tsang, 1986). Darüber hinaus basiert eine subjektive Angabe über das Ausmaß einer mentalen Beanspruchung auf den Informationen, die tatsächlich zum Zeitpunkt der Befragung erinnert werden können (Hart & Staveland, 1988). Vermei-

det man aber eine Unterbrechung des Aufgabenablaufes und macht eine Befragung erst nach Aufgabendurchführung, so sind die einzelnen Phasen der Aufgabendurchführung meist nicht mehr von einander separierbar.

Subjektive Maße laufen somit Gefahr, nicht alle relevanten Informationen zu beinhalten und Verzerrungen durch Situationskonstellationen und Reihenfolgeeffekten zu unterliegen.

Im Folgenden sollen exemplarisch drei in der Literatur beschriebene subjektive Befragungsinventare näher beschrieben werden.

RSME

Der unidimensionale RSME (Rating Scale of Mental Effort) nach Zijlstra (1993) fragt jeweils nach der subjektiv investierten Anstrengung bei der Lösung einer Aufgabe. Das Rating wird auf einer kontinuierlichen Linie durch ein Kreuz gekennzeichnet. Die Linie ist auf einer Länge von insgesamt 150 mm zu jedem 10 mm markiert. Auf der Länge der Linie sind jeweils Aussagen zu der investierten Anstrengung angegeben wie bspw. „fast keine Anstrengung“ oder „extreme Anstrengung“. Der Wert des Tests wird durch die Länge der mit dem Kreuz markierten Strecke in mm ermittelt.

Bei dem RSME handelt es sich um einen unidimensionalen Test der nur über die subjektiv wahrgenommene Höhe der investierten Anstrengung Auskunft gibt. Der Test ist schnell und einfach durchzuführen, bildet allerdings nur einen Aspekt des Beanspruchungskonzeptes ab.

NASA TLX

Der NASA-TLX (Task Load Index) (Hart & Staveland, 1988) wurde ursprünglich als multidimensionaler subjektiver Test zur Messung von mentaler und physischer Beanspruchung für Piloten entwickelt. Der NASA-TLX ist ein multidimensionaler Test, bestehend aus sechs Dimensionen: geistige Aktivität, körperliche Aktivität, zeitliche Anforderung, körperliche und geistige Anstrengung, eine Einschätzung wie gut die Aufgabe generell erfüllt werden konnte und wie wohl sich die

Person während der Aufgabenbearbeitung gefühlt hat. Jede der sechs Dimensionen wird auf einer 18-stufigen Rating Skala von 0 = gar nicht bis 18 = sehr stark pro Aufgabe einzeln erfasst.

Die ursprüngliche Version des NASA-TLX beinhaltet zusätzlich eine subjektive Gewichtung dieser sechs Dimensionen durch paarweise Vergleiche, um einen Gesamtscore aus den sechs Dimensionen bilden zu können. Eine solche Gewichtung trägt somit einer interindividuell unterschiedlichen Konzeption des Beanspruchungsbegriffes Rechnung.

SWAT

Die „Subjective Workload Assesment Technique“ (SWAT) (Reid & Nygren, 1988) wurde insbesondere für die Quantifizierung mentaler Beanspruchung im Kontext von Interaktionen einer Cockpitcrew entwickelt. Die SWAT postuliert ein multidimensionales Modell von Beanspruchung und differenziert drei Dimensionen: 1) Beanspruchung durch zeitliche Faktoren; 2) mentale Anstrengung; 3) psychischer Stress. Auch bei diesem Test werden die Dimensionen ähnlich wie beim NASA TLX von der Person selbst gewichtet und so in einen einzelnen Wert umgerechnet.

3.2 Leistungsmaße

Die zweite Gruppe der empirischen Messverfahren stellen die Leistungsmaße dar. Leistungsmaße schließen aufgrund einer bestimmten Leistung in einer Aufgabe auf die zugrundeliegenden Beanspruchungen. Leistungsmaße können in Leistungsmaße von Primär- und Sekundäraufgaben unterschieden werden.

Leistungsmaße bei Primäraufgaben

Unter Primäraufgabe versteht man in der Literatur die Aufgabe auf die der Hauptaufmerksamkeitsfokus gerichtet ist. Beispielsweise beim Autofahren die visuell-motorische Kontrolle des Fahrzeugs im Verkehr. Dabei genießt die Primäraufgabe meist den höchsten Prioritätsgrad bei der Vergabe von operativen Ressourcen.

Aufgrund der Diversität der verwendeten Primäraufgaben ist die Messung der Leistung hier sehr aufgabenspezifisch (s. a. De Waard, 1996). Typische Leistungsmaße stellen im klassischen Usability Test beispielsweise Reaktionszeiten, der prozentuale Anteil korrekter Antworten, oder die Qualität der Antworten bzw. in einer Fahrsituation die Längs- und Querverführung des Autos auf der Spur, die Geschwindigkeit und der Lenkwinkel dar (Knappe, Keinath, Meinecke, 2006).

Obgleich sich Leistungsmaße in vielen Anwendungsszenarien anbieten um Aussagen über Leistung der Probanden bei einem bestimmten System zu treffen, so muss hier auch berücksichtigt werden, dass solche Leistungsmaße von einer Vielzahl von Komponenten abhängen und somit im strengen Sinne keine direkte Aussage über die eigentliche mentale Beanspruchung eines Nutzers zulassen. So bestimmen zum Beispiel Strategien auf Seiten des Nutzers maßgeblich mit, wie viele der zur Verfügung stehenden Ressourcen für die Lösung einer Aufgabe eingesetzt werden. Es können beispielsweise zwei Nutzer dieselbe Leistung bei der Nutzung eines Systems erbringen, wobei einer der beiden bereits an der Grenze seiner Leistung angelangt ist, während der andere jedoch immer noch eine zusätzliche Aufgabe lösen könnte (De Waard, 1996). Die Interaktion zwischen den Ei-

enschaften eines Systems sowie den Eigenschaften und verwendeten Strategien seitens des Nutzers bestimmt maßgeblich die hier gemessene Leistung.

Daher ist es in vielen Fällen notwendig diese Leistungsmaße mit anderen Messverfahren zu kombinieren, um eine valide Interpretation der quantifizierbaren Leistungsmaße zu erreichen.

Leistungsmaße durch Sekundäraufgaben

Ein anderer Ansatz bei der Erhebung von Leistungsmaßen stellen die Leistungsmaße bei einer zusätzlichen Sekundäraufgabe dar. Dabei wird neben der eigentlichen Primäraufgabe eine zusätzliche Sekundär- bzw. Nebenaufgabe bearbeitet und die Leistung in dieser zusätzlichen Aufgabe ermittelt (O'Donnel & Eggermeier, 1986).

Ein Beispiel für ein Messverfahren mit Sekundäraufgabe im automotiven Kontext ist der Peripheral Detection Task (PDT). Der PDT beruht auf der Wahrnehmung von visuellen Stimuli in der Peripherie des visuellen Wahrnehmungsfeldes. Den Hintergrund dieser Methode bildet die Tatsache, dass sich die Größe des funktionalen Wahrnehmungsfeldes bei steigender Komplexität der Fahraufgabe verringert (Miura, 1986). Dieser sog. Tunneleffekt geht auf die erhöhte kognitive Anforderung bei einer erhöhten Komplexität der Fahraufgabe zurück. Der Tunneleffekt kann somit das Sehfeld in Abhängigkeit der Schwierigkeit der Aufgabe um bis zu 92% verringern (Williams, 1982; 1985). Aus der Detektionsrate der peripher präsentierten Stimuli kann somit auf die mentale Beanspruchung des Fahrers geschlossen werden (Oehme, 2003).

Auch wenn der Ansatz über Leistungsmaße in einer Sekundäraufgabe für verschiedene Anwendungssituationen bei der Messung mentaler Beanspruchung eingesetzt werden kann, so müssen an dieser Stelle auch einige Einschränkungen dieses Ansatzes diskutiert werden. Wie bereits ausgeführt kann die Einführung

einer Sekundäraufgabe auch dazu führen, dass die Leistung in der Primäraufgabe beeinträchtigt wird. Die Situation in der gemessen wird, ist somit immer eine andere als die, über die eine Aussage getroffen werden soll, was die Validität der gewonnenen Ergebnisse beeinflussen kann. Ein weiterer wichtiger Punkt bei der Betrachtung von Situationen in denen zwei oder mehr Aufgaben parallel bearbeitet werden, der bislang noch erstaunlich wenig Beachtung gefunden hat, ist der des Einflusses koordinativer Strategien. Sollen zwei Aufgaben parallel bearbeitet werden, so muss die Person von Situation zu Situation entscheiden, wie sie die ihr zur Verfügung stehenden kapazitiven und zeitlichen Ressourcen auf die vorliegenden Aufgaben verteilt. Es erscheint dabei mehr als fraglich, dass beispielsweise allein die Instruktion, etwa beide Aufgaben gleichermaßen zu bearbeiten, die koordinative Verteilung von Ressourcen bestimmt. Vielmehr ist es anzunehmen, dass eine Interaktion zwischen den Eigenschaften einer Situation und den persönlichen Faktoren der handelnden Person diese Verteilung bedingt. Befindet sich eine Person beispielsweise in einer realistischen Fahrsituation und soll gleichzeitig eine nicht sicherheitsrelevante Nebenaufgabe (z.B. PDT) ausführen, so wird sie bei einer Verknappung der Prozessressourcen, beispielsweise durch eine Erhöhung der Geschwindigkeit, natürlicherweise eher dazu tendieren, die Fahraufgabe auf einem stabilen Leistungsniveau zu halten und die Nebenaufgabe zu vernachlässigen (Schwalm & Zimmer, in Vorbereitung). Das Problem besteht darin, dass Personen die Priorisierung zwischen Haupt- und Nebenaufgabe unterschiedlich betreiben. Bei der Interpretation der gewonnenen Ergebnisse aus den Leistungsmaßen der beiden Aufgaben konfundieren empirisch nicht trennbar die beiden Faktoren der verfügbaren Prozessressourcen und der koordinativen Strategien.

Zusammenfassend bleibt in Bezug auf die Leistungsmaße bei Sekundäraufgaben festzuhalten, dass diese gegenüber den Leistungsmaßen bei Primäraufgaben Vorteile bei der Interpretation der Ergebnisse bieten. Zudem sind sie meist einfach und effizient in einer experimentellen Situation umzusetzen. Dennoch sind auch

sie mit Nachteilen behaftet, die je nach experimenteller Situation und der angestrebten Interpretation der gewonnenen Ergebnisse beachtet werden müssen.

3.3 Physiologische Maße

Die dritte Klasse der zu diskutierenden Messverfahren – die der physiologischen Maße – haben das Ziel, Biosignale und psychophysiologische Parameter für die Beanspruchung bei Informationsverarbeitungsprozessen zu bestimmen. Physiologische Maße sind dabei unterschiedlich sensibel in Bezug auf Sensitivität, Selektivität und Diagnostizität der Messung. Der Hauptvorteil bei der Erhebung von psychophysiologischen Parametern ist, dass sie keine direkte Reaktion seitens der Person verlangen und gleichzeitig die Messung durchgeführt werden kann, ohne dass der Handlungsablauf selbst unterbrochen werden muss. Während subjektive Indikatoren sowie verschiedene Leistungsparameter nur eine zeitlich sehr begrenzte Auflösung der Messung ermöglichen, die meist die Genauigkeit einzelner Aufgabenblöcke oder Aufgabenteile nicht überschreitet, bieten physiologische Maße darüber hinaus den Vorteil, kontinuierlich und mit einer hohen zeitlichen Auflösung Daten erfassen zu können.

Es folgt eine kurze Zusammenstellung der wichtigsten in der Literatur diskutierten Physiologischen Maße für mentale Beanspruchung zusammen mit einer Diskussion aktueller Ergebnisse sowie der Anwendbarkeit dieser Methoden in einem anwendungsorientierten Kontext. Gesondert betrachtete werden soll dabei die Methode der Pupillometrie, aufgrund ihrer zentralen Bedeutung für diese Arbeit.

3.3.1 Elektrokardiogramm (EKG)

Das Elektrokardiogramm (EKG) ist eine Methode, bei der der elektrische Impuls eines Herzschlags gemessen und mittels geeigneter Verfahren dargestellt wird. In Bezug auf die Messung von mentaler Beanspruchung ist dabei insbesondere die Information über a) die Herzrate das heißt die Anzahl der Kontraktionen pro Zeiteinheit sowie b) die Variabilität der Herzrate von Bedeutung.

Herzrate

Die Herzrate gehört wohl zu den am häufigsten verwendeten physiologischen Parametern in Beanspruchungsforschung (Wilson & Eggemeier, 1991). Der Grund hierfür mag in der relativ einfachen Ableitbarkeit des Signals, sowie der augenscheinlich hohen Validität in Bezug auf Stresssituationen liegen, die meist mit einer erhöhten Beanspruchung einhergehen. Der Einfluss mentaler Beanspruchung auf den Herzschlag, meist angegeben in BPM (Beats Per Minute), konnte in vielen unterschiedlicher Labor- und Feldversuchen nachgewiesen werden. So konnte beispielsweise unter Laborbedingungen ein Anstieg in der Herzrate bei zunehmender Beanspruchung durch Überwachungsaufgaben (Fournier, Wilson & Swain, 1999), Kopfrechenaufgaben (Köhler, 1996), bei Flugsimulationen (Veltman & Gaillard, 1996; Backs, Lennemann & Sicard, 1999; Svenson, Angelborg-Thanderez & Wilson, 1999; Backs, 1995) und anderen Aufgabenstellungen gezeigt werden (Ribback, 2003). Auch in Fahrsituationen in einem automotiven Kontext konnten Effekte der Beanspruchung auf die Herzrate berichtet werden (Brookhuis & De Waard, 2001; Richter, Wagner, Heger & Weise, 1998; Backs et al., 2003; Verwey & Veltman; 1996).

Herzratenvariabilität

Neben der Herzrate wird häufig auch die Änderung der Herzfrequenz zur Analyse von EKG Daten in Beanspruchungsuntersuchungen herangezogen. Man unterscheidet Veränderungen der Frequenz in verschiedenen Frequenzbändern (Muller, 1992):

- 0.02 – 0.06 Hz: Unteres Frequenzband, assoziiert mit Veränderungen der Körpertemperatur
- 0.07 – 0.14 Hz: Mittleres Frequenzband, assoziiert mit kurzfristiger Veränderung des Blutdrucks
- 0.15 – 0.40 Hz: Oberes Frequenzband, assoziiert mit respirativen Funktionen

Ein Absinken im mittleren Frequenzband (0.10 Hz Komponente) wird mit einer erhöhten Beanspruchung in Verbindung gebracht (Mulder & Mulder, 1981; Backs & Seljos, 1994). Wie jedoch bei Jorna (1992) diskutiert, sind solche Veränderungen oft nur auf große Veränderungen in der aktuellen Beanspruchung sensibel. So können solche Maße etwa nur zwischen Situationen ohne und mit zusätzlicher Aufgabe unterscheiden, können aber Unterschiede innerhalb einzelner Aufgaben nicht auflösen. Hinzu kommt die Kritik, dass in höheren Bereichen der Beanspruchung solche Maße nicht mehr in einem linearen Zusammenhang mit der Beanspruchung stehen.

Auch wenn einige Ergebnisse zur Herzrate (HR) und zur Herzratenvariabilität (HRV) auf eine Abhängigkeit zum Beanspruchungszustand hinweisen, so gibt es andererseits auch Studien, die keinen Zusammenhang zwischen mentaler Beanspruchung und HR oder HRV zeigen konnten (Wierwille et al. 1985). Einer der Gründe für das Fehlen eines eindeutigen Zusammenhangs ist möglicherweise der, dass kardiologische Maße von einer Vielzahl von Einflussgrößen abhängig sind. So zeigt sich beispielsweise, dass die HRV stark von physischer Anstrengung beeinflussbar ist (Lee & Park, 1990), genauso wie von Müdigkeit (Mascord & Heath, 1992) und insbesondere auch - hervorgerufen durch Veränderungen der Atemcharakteristika - von verbaler Aktivität (Sirevaag et al., 1993). Eine solche Beeinflussbarkeit der kardiologischen Maße macht eine präzise Bestimmung der Komponenten der mentalen Beanspruchung in einer anwendungsnahen Situation – wie beispielsweise einer realistischen Fahrsituation – äußerst schwierig wenn nicht gar unmöglich. Ein weiteres Problem ergibt sich aus der zeitlichen Auflösung, die durch die relativ niedrige durchschnittliche Herzfrequenz von 60 – 70 BPM begrenzt wird. Mulder und Mulder (1981) empfehlen sogar eine Aufzeichnung von mindestens fünf Minuten zur Analyse der 0.1 Hz Komponente. Eine solche Zeitspanne ist insbesondere bei der Betrachtung von Bediensystemen im automotiven Kontext, bei welcher Bedienhandlungen oftmals nur eine Dauer von wenigen Sekunden haben, nicht zielführend. Wohingegen sich ein solch kardiolo-

gisches Verfahren bei der Betrachtung längerfristiger Veränderungen bzw. Adaptationsprozessen durchaus sinnvoll einsetzen ließe.

3.3.2 Elektroencephalogramm (EEG)

Das EEG ist eine Methode, die Spannungsveränderungen an der Schädeloberfläche messbar macht und somit Rückschlüsse auf die neuronale Aktivität des Gehirns ermöglicht.

Die an der Schädeloberfläche abgeleiteten Spannungsveränderungen werden Spontanaktivität genannt. Ein solches Signal kann in unterschiedliche Frequenzbänder unterteilt werden (Tabelle nach Ribbeck, 2003):

Frequenzband	Frequenzbereich (Hz)	Amplitudenbereich (μV)	Region des stärksten Auftretens
Delta	0,5-4	20-200	Variabel
Theta	5-7	5-100	Frontal, temporal
Alpha	8-13	5-100	Okzipital, parietal
Beta	14-30	2-50	Präzentral, frontal
Gamma	30-50	2-10	Präzentral, frontal

Tabelle 1: Frequenzbandeinteilung des Spontan-EEG (nach Ribbeck, 2003)

Veränderungen in der Frequenz und Amplitude in einigen dieser Frequenzbänder werden in der Literatur als Indikatoren für mentale Beanspruchung diskutiert (für eine ausführliche Aufstellung empirischer Befunde sei auf die Arbeit von Rib-

beck, 2003 verwiesen). Allerdings zeigt sich auch bei der Diskussion der Spontanen EEG Befunde, ähnlich wie bei den Befunden zu kardiologischen Kennwerten, eine starke Abhängigkeit der Maße von anderen Störeinflüssen. So kommt Kramer (1991. zitiert in De Waard, 1996) zu dem Schluss, dass die Interpretation der Ergebnisse zum Alpha sowie Theta Band in einem laborspezifischen Kontext durchaus zu aussagekräftigen Ergebnissen führen kann, im Feld jedoch, wegen der großen Artefaktanfälligkeit, nur mit größerer Vorsicht ein zu setzen sind.

Eine weitere in der Literatur diskutierte Methode, die über die Elektroden an der Schädeloberfläche abgeleiteten Spannungsveränderungen zu interpretieren, ist die der ereigniskorrelierten Potentiale (EKP). Unter einem ereigniskorrelierten Potential versteht man die Spannungsveränderungen innerhalb eines definierten Zeitfensters als Reaktion auf exogene oder endogene Reize. Exogene Komponenten entstehen meist innerhalb der ersten 100 ms nach Reizpräsentation und repräsentieren zumeist physikalische Reizeigenschaften (z.B. Lautstärke). Endogene Komponenten stellen die interne Verarbeitung von Reizen dar (z.B. Aufmerksamkeitsprozesse). Als Indikator für mentale Beanspruchung wird die so genannte P300 Komponente diskutiert. Bezeichnet wird damit eine positive Spannungsveränderung nach ca. 300 ms nach Reizpräsentation, die ihre maximale Ausdehnung über dem parietalen Kortex hat (Mecklinger, 1992; Ribbeck, 2003). Mit der P300 Komponente wird sogar eine Diagnostizität im engeren Sinne diskutiert, da diese nicht auf motorische - sondern nur auf perzeptive, sowie kognitive Anforderungen reagiert (Manzey, 1998).

Der Nachteil bei der Methode der ereigniskorrelierten Potentiale ist der, dass die Effekte der EKP im Spontan-EEG nicht direkt sichtbar sind und erst durch Mittelungsprozesse der Antworten auf wiederholte Reize erkennbar werden (Lutzenberger, Elber, Rockstroh & Birnbaumer, 1985). Dieses setzt voraus, dass ein Reiz vielfach wiederholt so dargeboten werden kann, dass er nicht an externer Validität

verliert. Zudem ist, wie bereits beim Spontan-EEG diskutiert, auch bei diesem Verfahren eine hohe Artefaktanfälligkeit zu nennen.

Insgesamt stellt die Methode der Spontan-EEG sowie der ereigniskorrelierten Potentiale aber eine interessante Möglichkeit dar, mentale Beanspruchung über ein physiologisches Verfahren zugänglich zu machen. Dennoch besteht aus den bereits weiter oben diskutierten Punkten weiterer Forschungsbedarf bezüglich der Einsetzbarkeit dieser Methode in einem realen anwendungsorientierten Kontext wie etwa einem automotiven Kontext.

3.3.3 Elektrodermale Aktivität

Die Elektrodermale Aktivität (EDA) beschreibt Veränderungen der elektrischen Leitfähigkeit der menschlichen Haut. Ausgelöst werden diese Veränderungen durch die Aktivität der Schweißdrüsen und der anschließenden Reabsorption der Schweißflüssigkeit (s. Schandry, 1996). Für die Wirksamkeit der Messung mentaler Beanspruchung durch die Messung der elektrodermalen Aktivität liegen in der Literatur zahlreiche Belege vor.

So konnte ein Anstieg der EDA bei steigender Beanspruchung durch Arithmetikaufgaben in einem experimentellen Kontext nachgewiesen werden (Fahrenberg, Foerster, Schneider, Müller & Myrtek, 1984; Köhler, Fricke & Ritz, 1997; Kröner, 1984; Seibt, Scheuch & Boucsein, 1998). Die Wirksamkeit der Elektrodermalen Aktivität als Indikator für mentale Beanspruchung konnte aber auch in verschiedenen anwendungsorientierten Kontexten nachgewiesen werden. So zeigte sich die EDA als Indikator kurzfristiger Steigerungen der Beanspruchung bei Flugsimulationen (Lindholm, Cheatham & Cary, 1983). Aber auch zum Autofahren liegen mehrere experimentelle Evidenzen vor. So konnte ein Zusammenhang zwischen Bremsvorgängen und dem Auftreten von elektrodermalen Aktivität gezeigt werden (Helander, 1978). Zudem konnte ein Anstieg in der elektrodermalen Aktivität bei zusätzlicher Bearbeitung einer kognitiven Nebenaufgabe festgestellt werden (Verwey & Veltman; 1996), sowie der unterschiedliche Beanspruchungs-

grad bei der Verwendung eines Schalt- gegenüber einem Automatikgetriebe quantifiziert werden; eine Erhöhung der EDA bei der Nutzung eines Schaltgetriebes wird hier von den Autoren durch eine erhöhte Konzentration beim Fahrer erklärt.

Allerdings bleibt bei der Interpretation der Ergebnisse bei der Messung der EDA in einem anwendungsnahen Kontext immer auch die Frage nach dem Einfluss andere Einflussfaktoren wie etwa Atmung, Temperatur, Alter, Uhrzeit, als auch von Emotionen (De Waard, 1996).

3.3.4 Weitere psychophysiologischen Maße

Neben der Messung mentaler Beanspruchung durch kardiologische, elektroencephalographische oder elektrodermale Maße wurden in der Vergangenheit eine ganze Reihe weiterer physiologischer Verfahren zur Bestimmung mentaler Beanspruchungen in der Literatur diskutiert. Diese sollen nun aus Vollständigkeitsgründen ebenfalls kurz erwähnt aber im einzelnen nicht genauer diskutiert werden: So gibt es Ansätze mentale Beanspruchung messbar zu machen über Veränderungen in der Atemfrequenz (Veltman & Gaillard, 1998; Wientjes, 1992; Wientjes & Grossman, 1998), des Blutdrucks (Veltman & Gaillard, 1996), über Veränderungen in der Stimme und Sprache (Baber, Mellor, Graham, Noyes & Tunley, 1996; Hagmueller et al. 2005), in der Blinkfrequenz der Augen (Kramer, 1991; Stern, Boyer & Schroeder, 1994), der Muskelspannungen sowie des Gesichtsausdrucks (Boxtel & Jessurun, 1993). Solche Maße unterscheiden sich stark in ihrer tatsächlichen Anwendbarkeit in realistischen Situationen. Manche sind nur mit einem hohen technischen Aufwand realisierbar, andere weisen nur geringe Reliabilitäten der Messung auf oder sind äußerst anfällig auf externe Störeinflüsse und müssen daher vor jeder Anwendung in einer realen Situation sorgfältig auf ihre Aussagekraft geprüft werden.

Im Folgenden soll nun genauer auf ein weiteres physiologisches Verfahren für die Messung mentaler Beanspruchung eingegangen werden, auf dem die experimen-

tellen Untersuchungen der vorliegenden Arbeit beruhen. Es handelt sich dabei um das Verfahren der Pupillometrie, wobei Veränderungen in der Größe der menschlichen Pupille Rückschlüsse auf die aktuelle mentale Beanspruchung erlauben.

4 Pupillometrie

Das physiologische Messverfahren der Pupillometrie zur Bestimmung mentaler Beanspruchung basiert auf der Beobachtung, dass sich die Größe der menschlichen Pupille in Abhängigkeit der mentalen Beanspruchung verändert. Eine solche Veränderung kann gemessen, abgebildet und interpretiert werden. Ein wesentlicher Vorteil der Messung von mentaler Beanspruchung über ein pupillometrisches Verfahren ist - wie für andere physiologische Messverfahren auch - der einer objektiven und störungsfreien Messung (ohne zusätzliche Befragung oder Aufgabe). Die Pupillometrie bietet aber aufgrund der geringen zeitlichen Latenz einer pupillometrischen Reaktion jedoch auch den weiteren entscheidenden Vorteil einer erhöhten zeitlichen Genauigkeit der Messung. In dem nachfolgenden Kapitel sollen zunächst die biologischen Hintergründe einer pupillometrischen Reaktion genauer beschrieben und eine Aufstellung der wichtigsten empirischen Befunde aus der Literatur bzgl. der Messung mentaler Beanspruchung über Reaktionen der Pupille gegeben werden. Anschließend soll hier insbesondere das in der vorliegenden Arbeit verwendete Verfahren des „Index of Cognitive Activity“ (ICA) zur Bestimmung der aktuellen Beanspruchung mittels einer pupillometrischen Messung genauer beschrieben werden.

4.1 Physiologische Grundlagen

Die allgemeine Struktur des Auges kann mit der einer Kamera verglichen werden (Beatty & Wagoner, 2000). Das Auge hat eine Linse, um einfallendes Licht zu brechen und zu bündeln, sowie eine Blende, um die Menge des einfallenden Lichts zu kontrollieren. Diese Funktion übernimmt beim menschlichen Auge die Pupille, eine Öffnung in einer pigmentierten Struktur, der Iris.

Die Iris verfügt über zwei antagonistische Muskelgruppen: den Dilator und den Sphincter Muskel.

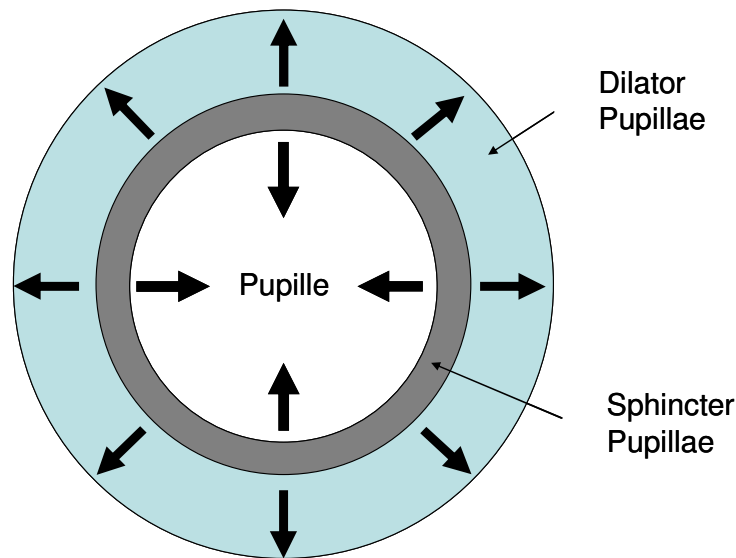


Abbildung 5: Die Muskeln der Iris: Zwei antagonistische Muskeln steuern die Größe der Pupille. Der Sphincter Muskel verkleinert die Pupille, wenn dieser sich zusammenzieht, während der Dilator den Pupillendurchmesser vergrößert.

Jeder dieser Muskeln wird durch eine andere Komponente des Nervensystems innerviert. Der Sphincter Muskel, der die Pupille verkleinert, wenn er sich zusammenzieht, wird parasympathisch und der Dilator, der die Pupille durch Kontraktion vergrößert, wird sympathisch innerviert. Somit ist jede Änderung in der Größe der Pupille ein Ergebnis sowohl sympathischer als auch parasympathischer Aktivität. Ihre letztendliche Größe wird bestimmt durch das Verhältnis der Aktivitäten beider Systeme. Die Größe der menschlichen Pupille variiert dabei je nach anatomischer Gegebenheit zwischen weniger als 1 mm bis zu 9 mm Durchmesser (Beatty & Lucero-Wagoner, 2000).

Einige der Veränderungen der Pupillengröße sind rein optischer Natur. So dient der „Pupillometrische Lichtreflex“ dazu, die Menge des einfallenden Lichtes in das Auge zu regulieren. In hellem Licht ist die Pupille klein und zusammengezogen, wohingegen bei Dämmerlicht die Pupille geöffnet und entspannt ist. Eine solche regulative Reaktion geschieht im menschlichen Auge in einem Zeitraum von ca. 0,2 Sekunden nach Stimulus Onset, wobei die maximale Reaktion ca. 0,5

bis 1 Sekunde später auftritt. Der Pupillen Licht Reflex dient als ein wichtiger diagnostischer Indikator des neurologischen Status. Die „Akkomodations Reaktion“, oder auch „Nähe Reflex“ genannt, dient der Regulation der Krümmung der Pupille, um einen fixierten Gegenstand im Raum durch entsprechende Veränderung der Brechkraft der Linse auf der Retina scharf abzubilden. Wird ein Objekt z. B. in der Nähe fixiert, wird dieses zunächst auf der Retina unscharf abgebildet bis der Akkomodationsreflex die Krümmung der Linse adaptiert hat. Obwohl hier in erster Linie die Brechkraft der Linse verändert wird, kann es dabei auch zu einer Veränderung der Pupillengröße kommen.

Andere Bewegungen der Pupille sind jedoch weder von Unterschieden in der Menge des einfallenden Lichtes noch von der Distanz des Auges zu einem fokussierten Objekt abhängig, sondern werden durch die Verarbeitung sensorischer, mentaler sowie emotionaler Informationen beeinflusst. So zeigt sich, dass sich die Pupille sobald eine mentale Beanspruchung anliegt, relativ zu dem vorangegangenen Zeitintervall in welchem noch keine Beanspruchung anlag, weitet. Das heißt der Pupillendurchmesser vergrößert sich (Beatty & Lucero-Wagoner, 2000). Diese durch den sogenannten psychosensorischen Reflex (Loewenfeld, 1993) hervorgerufenen pupillometrischen Veränderungen sind in einem anwendungsorientierten Kontext von besonderem Interesse, da diese Indikatoren eine Aufmerksamkeitszuwendung sowie eine mentale Aktivität im allgemeinen darstellen (Kahnemann, 1973). Diese kleinen aber reliablen Veränderungen (Beatty & Lucero-Wagoner, 2000) der Pupillengröße überlagern die optisch bedingten Veränderungen der Pupille und müssen von diesen durch geeignete Analyseverfahren separiert werden.

Der psychosensorische Reflex ist dabei nicht nur abhängig von externalen, sensorischen Ereignissen, sondern spiegelt vielmehr eine ganze Bandbreite unterschiedlicher mentaler Prozesse wieder. Darunter auch die Höhe der mentalen Aktivität oder der mentalen Beanspruchung (Beatty & Lucero-Wagoner, 2000). Genau dieser Zusammenhang der Pupillengröße mit höheren kognitiven Funktionen ist der

Grund für ein seit den frühen 60iger Jahren andauerndes Interesse an dem hier beschriebenen psychosensorischen Reflex. Diese Veränderungen der Größe der Pupille dienen als Indikatoren für Vorgänge innerhalb des Gehirns, welche den Aspekten der Dynamik sowie Intensität der menschlichen Kognition unterliegen.

Die Größe der menschlichen Pupille ist also von anderen Reizen und Einflüssen abhängig, insbesondere auch von neuronalen Prozessen, die höhere kognitive Funktionen begleiten. Es ist daher nicht verwunderlich, dass für die Erweiterung und somit Steuerung der Pupille nicht nur ein afferenter Pfad verantwortlich ist, sondern zwei neuronale Pfade eine entscheidende Bedeutung spielen (Steinhauer, Siegle, Condray & Pless, 2004). Zum einen der sympathische Pfad über die posterioren hypothalamischen Kerne zu den Dilator Muskeln, zum anderen der parasympathische Pfad, der eine Hemmung des Edinger-Westphal Komplexes der oculomotorischen Kerne und somit eine Relaxation der Sphincter Muskeln bewirkt (Beatty & Lucero-Wagoner, 2000).

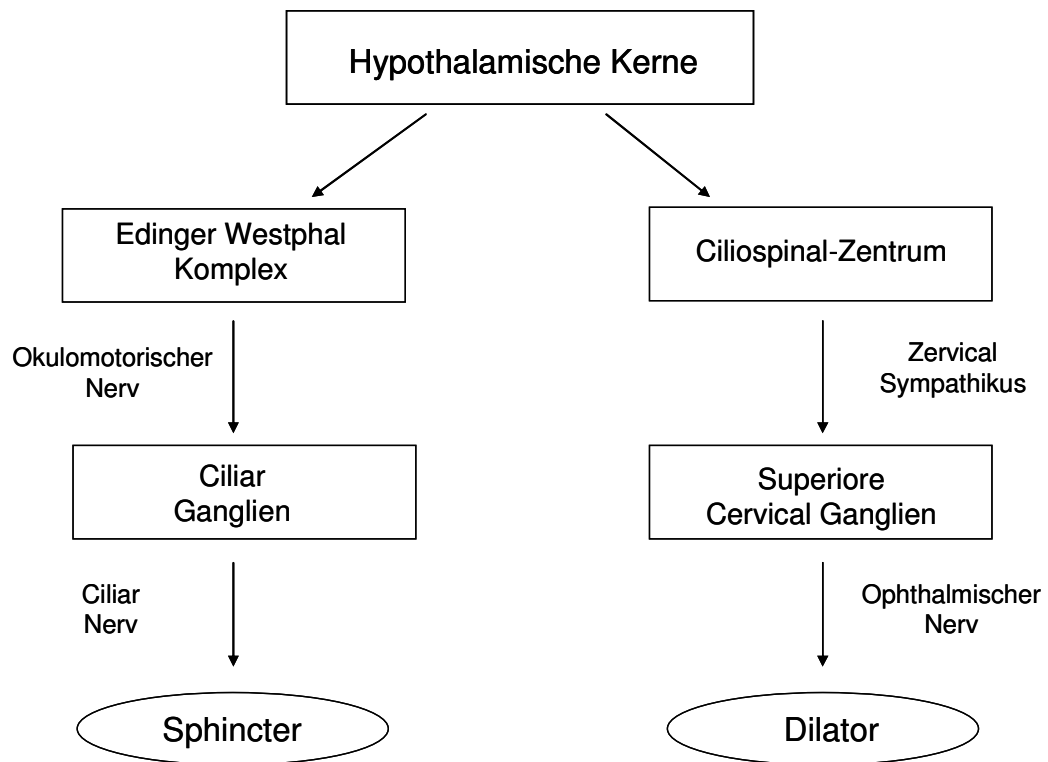


Abbildung 6: Die Abbildung zeigt die neuronalen Strukturen und Pfade über die die Größe der menschlichen Pupille gesteuert wird (nach Beatty und Lucero-Wagoner, 2000).

Wenngleich die neuronalen Pfade, über welche die Größe der menschlichen Pupille gesteuert wird, weitestgehend bekannt sind, so liegen über die Hintergründe für eine Änderung der Pupillengröße unter kognitiver Beanspruchung nur wenige Befunde vor. Anzunehmen ist, dass die Prozesse, welche eine Veränderung der Pupillengröße unter mentaler Beanspruchung hervorrufen, mit Hirnregionen verknüpft sind, welche mit der Steuerung und Aufrechterhaltung von Aufmerksamkeit und der Verarbeitung von Informationen assoziiert sind. So etwa Regionen des frontalen Cortex, welches in Läsions- (Sarter, Givens & Bruto, 2001) sowie bildgebenden Verfahren bestätigt wurde (Siegle et al., 2003). Dennoch findet sich bis heute keine umfassende Erklärung, wie genau sich die Höhe der kognitiven Aktivität in einer solchen sympathischen und parasympathischen Aktivität widerspiegelt, welche wiederum die Pupillengröße unter mentaler Beanspruchung beeinflusst (Steinhauer, Siegle, Condray & Pless, 2004).

Veränderungen der Pupillengröße werden durch die Aktivität der Augenmuskulatur erzeugt, welche der Kontrolle sowohl sympathischer als auch parasympathischer Pfade des autonomen Nervensystems unterliegen. Dennoch muss eine Verknüpfung von Reaktionen der Pupille und kognitiven Prozessen auf einer höheren Ebene des menschlichen Nervensystems stattfinden (Beatty & Lucero Wagoner, 2000).

4.2 Empirische Evidenzen Pupillometrie und mentale Beanspruchung

Bereits seit über hundert Jahren ist bekannt, dass kognitive Prozesse von Reaktionen in der Größe der menschlichen Pupille begleitet werden. So wurde schon von Shiff (1875, Shiff & Foa, 1874, zitiert in Beatty-Lucero-Wagoner, 2000) beschrieben, dass eine ganze Reihe nicht-visueller Stimuli solche Veränderungen verursachen können. Und bereits Ende des 18. Jahrhunderts zeigte Heinrich (1896) in einer experimentellen Studie, dass eine mentale Multiplikationsaufgabe eine Weitung der Pupillengröße hervorrufen kann. Dennoch dauerte es mehr als ein halbes Jahrhundert, bis eine erste systematische, kognitiv-psychologische Betrachtung des hier beschriebenen Effektes erfolgte. Erst 1964 publizierten Hess und Polt (Hess & Polt, 1964) eine erfolgreiche experimentelle Replikation des von Heinrich bereits 1896 beschriebenen Pupillen Dilatationseffektes bei mentaler Arithmetik.

Inzwischen findet sich in der Literatur eine relativ große Anzahl von Belegen für die Möglichkeit der Messung mentaler Beanspruchungen durch Reaktionen der Pupille. Im Folgenden soll eine kurze tabellarische Aufstellung der wichtigsten in der Literatur diskutierten experimentellen Befunde gegeben werden. Die Aufstellung konzentriert sich dabei auf Befunde zu Reaktionen der Pupille auf mentale Beanspruchungen. Die Darstellung unterteilt die Befunde in Hinblick auf verwendete Aufgabenkategorien und erhebt keinen Anspruch auf Vollständigkeit, Autoren sind jeweils nur exemplarisch genannt und teilweise wurden Gebiete zusammengefasst.

Aufgabentyp	Literatur	Beispiele
Mentale Arithmetik	Hess & Polt, 1964; Bradshaw, 1968; Ahern & Beatty, 1979 1981; Beatty & Wagoner, 1978; Rösger, 1996;	Mathematische Rechenaufgaben (Addition, Subtraktion, Multiplikation, Division, Modulo)
Schlussfolgern	Ahern & Beatty, 1979;	Entscheidungsaufgaben
Visuell-räumliches Arbeitsgedächtnis	Van Orden, Limbert & Makeig, 2001;	Visuelle Brooksauflage, Sternberg Paradigma
Visuelle Suche	Backs & Walrath, 1992; Porter, Troscianko & Gilchrist, 2007;	Visuelle Suche und Identifikation von Reizen
Lexikalisches Wissen	Beatty & Wagoner (1978);	Vergleich von Buchstaben
Reizdetektion und Reizdiskrimination	Hakerem & Sutton (1966) Bradshaw, 1979, 1970 Kahnemann & Beatty, 1967 Beatty & Wagoner (1977);	Entdeckung und Unterscheidung von Reizen
Digit span recall	Gardner, Beltramo & Krinsky, 1975; Granholm, Asarnow, Sarkin & Dykes, 1996; Kahneman & Beatty, 1966; Kahnemann, Beatty & Pollack, 1967; Peavler, 1974;	Merken und wiederholen einer Zahlenkette variabler Länge (visuell und auditiv)

Reiz Antizipation	Qiuyuan et al., 1985; Pritchard, 1981;	Erwartung bezüglich des Auftretens eines Stimulus
Grammatikalische Entscheidungen	Ahern & Beatty, 1981;	Entscheidung nach einer spezifischen Satzstruktur (A folgt B oder B folgt A)
Satzverstehen	Metalis, Rhoades, Hess & Petrovich, 1980; Schluroff, 1982; Schluroff et al, 1986; Just & Carpenter, 1993; Kahneman & Wright, 1971; Ben-Nun, 1986;	Semantische, syntaktische und grammatikalische Komplexität von Sätzen
Übersetzen und Interpretieren	Hyönä, Tomola & Alaja (1995);	Übersetzen und Interpretieren von sprachlichen Material
Zeitliche Relationen	Krüger, Nuthmann & Van Der Meer, 2001;	Operationen über zeitliche Ereignisfolgen
Dual Task	Kahnemann & Beatty 1966;	Gleichzeitiges Bearbeiten mehrerer Aufgaben
Überlastung / Leistungsgrenze	Peavler, 1974; Granholm et al. 1997; Granholm, Asarnow, Sarkin & Dykes, 1996;	Bei Leistungsgrenze steigt auch die Pupillengröße nicht weiter an

Tabelle 2: Die vorliegende Tabelle zeigt eine exemplarische Aufstellung der wichtigsten in der Literatur beschriebenen Ergebnisse zum Nachweis mentaler Beanspruchungseffekte durch Reaktionen der menschlichen Pupille.

Während in der Literatur also eine ganze Reihe von Evidenzen für den Zusammenhang zwischen mentaler Beanspruchung und Veränderungen in der Pupillengröße zu finden sind, so hat doch die Methode selbst in einem wirklich anwendungsnahen Kontext bis lang nur vergleichsweise wenig Anwendung gefunden. Dabei bietet die Pupillometrie aufgrund der geringen Latenzzeiten einer pupillometrischen Reaktion auf sich verändernde Beanspruchungen den Vorteil einer

hohen zeitlichen Auflösung gegenüber anderen physiologischer Messverfahren (vgl. Kapitel 3.3). So werden für eine Reaktion der Pupille auf Veränderungen der Beanspruchung eine Latenzzeit von 100 – 200 ms diskutiert (Beatty, 1982). Demgegenüber wirken sich Veränderungen der Beanspruchung im Hautleitwert erst nach 1,4 – 2,5 Sekunden aus. In der Positronen-Emissions-Tomographie sind entsprechende Veränderungen sogar erst nach frühesten 30 sec sichtbar (Kramer, 1991). Zudem ist der technische Aufwand für eine Erfassung der Größenänderung der Pupille im Vergleich z.B. zum EEG relativ gering.

Die Tatsache, dass die Methode der Pupillometrie bisher nur relativ selten in anwendungsnahen Situationen im automotiven Kontext zum Einsatz gekommen ist, ist nicht zuletzt darauf zurückzuführen, dass die Veränderungen der Größe der Pupille, wie eingangs beschrieben, auch von anderen Faktoren, insbesondere von sich verändernden Beleuchtungsbedingungen bestimmt werden. Nimmt man – wie in den bisherigen Studien zur Pupillometrie geschehen – den absoluten Anstieg in der Größe der Pupille als Indikator für einen Anstieg in der mentalen Beanspruchung, so zeigt sich, dass der Einfluss einer mentalen Beanspruchung auf die Pupillengröße gegenüber dem sich verändernder Beleuchtungsbedingungen vergleichsweise gering ist. Dies macht eine saubere Erfassung derartiger Effekte bei sich verändernden Leuchtdichteverhältnissen schwierig oder gar unmöglich. In einer Laborsituation, wie sie in den oben genannten Studien vorlag, können solche Störeinflüsse sowie andere Umwelteinflüsse durch geeignete Maßnahmen auf ein Minimum reduziert werden. In einem realistischen, anwendungsnahen Kontext wie etwa beim Fahren eines Fahrzeugs und/oder bei der Bearbeitung einer Aufgabe in einem Fahrerinformationssystem hingegen ist dies nicht möglich. Hier bedingt die Aufgabe selbst die Veränderung in der Menge des auf das Auge treffenden Lichtes. So etwa durch den Wechsel zwischen Menüs oder bei einem Blickwechsel zwischen Fahrzeuginnenraum und Straße. Dies bedeutet eine entsprechende Konfundierung der Einflussfaktoren auf die gemessene Größe der Pupille, welche im Nachhinein nicht mehr auflösbar ist. Dennoch konnten bereits

einige Studien auch in einem anwendungsnahen Kontext einer Fahraufgabe einen positiven Zusammenhang zwischen einer erhöhten Beanspruchung und pupillometrischen Reaktionen feststellen. So konnten Tsai et al. (2007) in einer simulierten Fahraufgabe mit einer zusätzlichen auditiv präsentierten mathematischen Aufgabe zeigen, dass je höher die Leistung in der Nebenaufgabe ist, desto höher auch die Veränderungen in der gemessenen Pupillengröße ausfallen. Rößger (1996) gelang es, nach einer Korrektur der Pupillengröße auf die sich verändernde Helligkeit, in einer realen Fahrsituation Effekte in der Pupillengröße in Abhängigkeit der Komplexität der Fahrsituation nachzuweisen. So zeigten sich die größten Pupillendurchmesser in einer Abbiegesituation, bei der Vorfahrtsregeln beachtet werden mussten. Die kleinsten Pupillendurchmesser zeigten sich dagegen beim Stehen an einer roten Ampel.

Bei der Diskussion der hier beschriebenen pupillometrischen Messergebnisse in einem anwendungsnahen Kontext fällt ein Punkt besonders auf. Wenngleich wiederholt in den Studien der Vorteil der Pupillometrie gegenüber anderer Verfahren zur Bestimmung von Beanspruchungen in der hohen zeitlichen Genauigkeit sowie der Kontinuität der Messung diskutiert wird, wird genau dies in den Studien nicht genutzt. Berichtet und diskutiert werden hier lediglich die über die Situationen gemittelten Größen und deren Veränderungen, nicht aber der tatsächliche Verlauf der Beanspruchung. Genau die hohe zeitliche Auflösung der Messung stellt jedoch einen der Hauptvorteile der Methode der Pupillometrie gegenüber anderer Messverfahren dar.

Die hier beschriebenen Ergebnisse zeigen deutlich das Potenzial der Pupillometrie als Methode zur Bestimmung der mentalen Beanspruchung in einer Vielzahl experimenteller Situationen. Auch wenn die neuronalen Ursachen und Hintergründe für eine Veränderung der Pupillengröße unter Beanspruchung bis heute nicht vollständig geklärt sind, so zeigen die hier aufgeführten experimentellen Evidenzen deutlich den korrelativen Zusammenhang zwischen Beanspruchung und Pupillen-

größe. Auch zeigt sich, dass ein pupillometrisches Maß eher einen Indikator der globalen Arbeitsbelastung darstellt (De Waard, 1996), ohne eine explizite diagnostische Aussage in Bezug auf die zugrunde liegenden Ursachen einer solchen erhöhten Beanspruchung zu machen. In einer anwendungsorientierten Situation, wie sie im Fokus der vorliegenden Arbeit steht, ist dies jedoch auch nicht unbedingt notwendig. Wie in Kapitel 2 ausgeführt, handelt es sich bei einer anwendungsnahen Situation wie der des Autofahrens um eine Situation, bei der eine Vielzahl von Faktoren miteinander interagieren und somit die resultierende Beanspruchung bestimmt. Hier wäre es bereits von Vorteil Situationen mit einer erhöhten mentalen Beanspruchung, die potentiell für den Fahrer gefährlich werden könnten, mit einer hohen zeitlichen Genauigkeit identifizieren zu können. Darüber hinaus sollte mithilfe der Pupillometrie eine Vermessung der kontinuierlichen Verläufe von mentalen Beanspruchungen mit einer zeitlichen Genauigkeit möglich sein, wie sie mit keinem anderen klassischen Maß der Usability bis lang erreicht werden kann.

Das bei diesen einmaligen Möglichkeiten, die die Pupillometrie bietet, die Methode nicht schon sehr viel früher zu einer breiten Anwendung bei der Entwicklung und Optimierung von Mensch-Maschine Schnittstellen gefunden hat, liegt hauptsächlich daran, dass die Veränderungen der Größe der Pupille insbesondere auch durch die Lichtverhältnisse mit bestimmt wird. Der Einfluss einer mentalen Beanspruchung auf die Pupillengröße ist demgegenüber vergleichsweise gering. Bislang war es nicht möglich in einer anwendungsnahen Situation, bei der sich der Proband in einer abwechslungsreichen, eventuell sogar dynamisch verändernden Umwelt befunden hat, eine präzise, zeitlich korrelierte Messung der sich verändernden Bedingungen zu erreichen, um diese nachträglich aus dem Signal der Pupillengröße heraus zu rechnen (s. Rößger, 1996). Es ist nahezu unmöglich zu jedem Zeitpunkt zu bestimmen, wie viel Helligkeit tatsächlich das Auge erreicht. Dieses führt auch dazu, dass bestehenden Studien die hohe zeitliche Genauigkeit der Messung pupillometrischer Reaktionen bisher nicht ausschöpfen konnten.

In der vorliegenden Arbeit soll daher ein neues pupillometrisches Verfahren diskutiert werden, bei dem nicht etwa die einzelnen Einflussfaktoren extern gemessen und daraufhin aus dem Signal herausgenommen werden, und auch nicht ein absoluter Anstieg in der Pupillengröße als Indikator herangezogen wird, sondern bei dem stattdessen nach ganz bestimmten charakteristischen zeitlichen Veränderungen in dem pupillometrischen Signal gesucht wird, aus denen daraufhin ein korrelatives Maß für die Höhe der mentalen Beanspruchung abgeleitet werden kann.

4.3 Index of Cognitive Activity (ICA)

Im Folgenden soll das in der vorliegenden Arbeit verwendete Verfahren des „Index of Cognitive Activity“ (ICA) (Marshall, Davis & Knust, 2004; Marshall, 2000, 2005, 2007) beschrieben werden. Das Verfahren des ICA beschreibt dabei eine neue Methode der Messung von mentalen Beanspruchungen über Reaktionen der Pupille, bei dem weder, wie in den bisherigen Studien zur Pupillometrie, über einen absoluten Anstieg in der Größe der Pupille noch über eine externe Erfassung von Störeinflüssen argumentiert wird. Stattdessen sucht das hier beschriebene Verfahren nach ganz spezifischen Charakteristika in der Veränderung der Größe der Pupille welche mit dem Auftreten mentaler Beanspruchungen assoziiert sind.

Die wohl am häufigsten verwendete Technik (Beatty & Lucero-Wagoner, 2000) bei der Bestimmung pupillometrischer Veränderungen in Bezug auf mentale Reaktionen ist die der Aufgaben-induzierten Pupillenreaktion oder „Task-evoked Pupillary Response“ (TEPR). Ein Verfahren entwickelt durch Jackson Beatty und Kollegen Ende der 80iger Jahre (Richer, Silverman & Beatty, 1983). Dieses Verfahren ist analog zu dem der Aufgaben evozierten Potentiale in der EKP Forschung. Hier werden in einer Studie verschiedene Stimuli wiederholt für eine kurze Zeitspanne dargeboten. Die so gewonnenen und von Artefakten (Blinks etc.) befreiten Daten werden daraufhin über die Durchgänge hinweg gemittelt. Eine Baseline Messung, meist am Trial Onset gemessen (das heißt eine Messung vor der eigentlichen Aufgabenbearbeitung), wird daraufhin von der Peak-Dilatation bzw. von der mittleren Pupillenerweiterung bei Verarbeitung/Bearbeitung des Stimulus abgezogen. Das Ergebnis ist eine Abschätzung der durchschnittlichen Veränderung in der Größe der Pupille bei der Bearbeitung einer bestimmten Aufgabe.

Eine solche Methode der TEPRs ist jedoch nur auf eine sehr begrenzte Menge an Forschungsfragen zu applizieren. Nicht jede Fragestellung ermöglicht bestimmte Stimuli zu einer exakt bestimmten Geschwindigkeit und unter identischen Bedingungen wiederholt zu präsentieren. Insbesondere bei der Betrachtung anwendungsorientierter Fragestellungen treten zumeist Situationen auf, bei denen die Probanden selbst die Geschwindigkeit und den Ablauf einer Aufgabe steuern müssen. Oder es handelt sich um Situationen, bei der aus einer ganzen Reihe aufeinander folgender mentaler Operationen eine spezifische besonders interessiert. Hier bietet ein solches Maß keine saubere Möglichkeit die Baseline zu einem bestimmten Zeitpunkt zu bestimmen.

Wie bereits in dem vorangegangenen Abschnitt beschrieben besteht eine weitere Einschränkung eines solchen Verfahrens in der Abhängigkeit der Pupillengröße von Veränderungen in der Helligkeit des betrachteten Stimulus. Werden die Bedingungen nicht exakt gleich gehalten, so spiegelt eine Veränderung in der Pupille nicht nur die auftretende mentale Beanspruchung wieder sondern auch die Veränderungen in der Helligkeit. Ein solches Artefakt verhindert den sinnvollen Einsatz einer solchen Methode in einem Anwendungsszenario bei dem neben dem Stimulus selbst auch bestimmte Umgebungsvariablen veränderbar sind, wie beispielsweise Farbgebung und Entfernungen zu Objekten.

Zusätzlich ergibt sich bei der Methode der TEPRs aufgrund der absoluten anstatt einer relativen Betrachtung der Pupillengröße ein zusätzliches Problem. Während bereits Veränderungen in der Größe der Pupille um 0.1 bis 0.5 mm als signifikant erachtet werden (Beatty & Lucero-Wagoner, 2000), variiert die Größe der menschlichen Pupille selbst bereits zwischen 4 und 9 mm Durchmesser. Dies bedeutet, dass eine Veränderung um 0.4 mm bei einer Pupillengröße von 4 mm eine Veränderung von 10%, bei einer Größe von 8 mm jedoch nur eine Veränderung um 8% bedeutet. Das hier beschriebene Problem könnte mit einem relativen Maß für die Veränderung in der Größe der Pupille gelöst werden.

Der Index of Cognitive Activity (ICA) ist nun eine Methode bei der anders als bei TEPRs nicht etwa nach gesamthaften Veränderungen der absoluten Pupillengröße gesucht wird. Stattdessen sucht das Verfahren des ICA nach ganz bestimmten schnellen und kurzfristigen Veränderungen in dem Signal der Pupillengröße. Diese kurzfristigen und schnellen Änderungsanteile sind ebenfalls relativ zu kognitiven Prozessen und begleiten genauso wie eine absolute Weitung der Pupille, wie bereits unter dem psychosensorischen Reflex beschrieben, einen Anstieg in der mentalen Beanspruchung. Diese kleinen, kurzfristigen und kognitiv relatierten Änderungsanteile im Pupillensignal haben allerdings bislang in der pupillometrischen Forschung keine Beachtung gefunden. Bei dieser Methode werden zudem nur relative Veränderungen anstatt absolute Veränderungen in der Pupillengröße betrachtet. Da in diesem Verfahren nur ganz bestimmte Charakteristika von Veränderungen in der Pupillengröße betrachtet werden, können diese von Veränderungen, welche durch andere Faktoren wie Helligkeitsunterschiede hervorgerufen werden, unterschieden werden. Schließlich ermöglicht ein solches Maß neben der Betrachtung diskreter Aufgaben auch den Zugang und die Erfassung der Pupillenveränderung bei kontinuierlichen Aufgaben und zeitlichen Verläufen mit einer hohen zeitlichen Präzision (Marshall, Davis & Knust, 2004).

4.3.1 Beschreibung des Verfahrens

Der Index of Cognitive Activity (ICA) wird nun im Gegensatz zu den bisher diskutierten Verfahren der Pupillometrischen Forschung aus einem bestimmten Anteil des Signals, den hochfrequenten Komponenten des Größensignals der Pupille berechnet. Diese kurzfristigen und schnellen Änderungen begleiten neben einer gesamthaften Vergrößerung der Pupille ebenfalls eine mentale Beanspruchung und können somit als Indikatoren herangezogen werden. Sie sind daher ein Maß für die relative Veränderung der Größe der Pupille und beschreiben dabei die Anzahl der kurzfristigen, unregelmäßigen und schnellen Anstiege in der Amplitude des Pupillensignals. Das Verfahren selbst wurde von Marshall und Kollegen ent-

wickelt und unter dem US Patent No. 6,090,051 patentiert (Marshall, 2000, 2005, 2007; Marshall, Davis & Knust, 2004).

Um den Einfluss des bereits zuvor diskutierten Lichtreflexes gegenüber dem Dilatationsreflex bei mentalen Prozessen unterscheiden zu können, nutzt der ICA folgende Gegebenheiten (nach Marshall 2000): Während des Lichtreflexes (Licht-an) fungiert der Sphincter Muskel als Agonist und wird angeregt sich zusammen zu ziehen, während der Dilator als Antagonist gehemmt wird und so daran gehindert wird sich ebenfalls zusammen zu ziehen. Wenn der Lichteinfall reduziert wird (Licht-off) fällt die Innervation des Sphincters weg, die Kontraktion lässt nach und die Pupille kehrt in die Ruhelage zurück. Wenn nun eine Person eine Aufgabe mit einer signifikanten mentalen Beanspruchung zu leisten hat, so zeigt sich in der Pupille eine Reaktion, die sich von der bei einem Lichtreflex unterscheidet. Hier ist der Prozess der reziproken Innervation umgekehrt. Während der Dilator Muskel angeregt wird sich zusammen zu ziehen, was eine Erweiterung der Pupille bewirkt, wird gleichzeitig auch der Sphincter Muskel inhibiert, was ebenfalls eine Erweiterung der Pupille bewirkt. Das Ergebnis ist hier eine kurzfristige Weitung der Pupille, welche schneller und größer ist als dass sie eine Muskelgruppe alleine hätte produzieren können und somit auch schneller und stärker ausfällt als eine Reaktion nach einem Lichtreflex bei der die Pupille wieder in die Ursprungslage zurückkehrt.

Solche kurzfristigen, unregelmäßigen und schnellen Änderungen wurden auch bereits früher schon in der Literatur unter dem Begriff „Pupillary-Hippus“ beschrieben (Beatty & Lucero-Wagoner, 2000). Der Begriff leitet sich aus dem griechischen Wort für Pferd ab und beschreibt kleine, schnelle und unregelmäßige Veränderungen der Größe der menschlichen Pupille. Diese Veränderungen wurden bereits zuvor als kognitiv relatiert und visuell insignifikante dynamische Veränderungen beschrieben (Beatty & Lucero-Wagoner, 2000). Diese rhythmischen aber unregelmäßigen Veränderungen der Pupille sind unabhängig von Augenbe-

wegungen oder Veränderungen der Leuchtdichte (McLaren, Erie & Brubaker, 1992).

Ähnlich wie der Lichtreflex ist der Dilatationsreflex ein transitives Ereignis. In beiden Fällen beobachtet man ein Pulsieren des Durchmessers, wobei beim Dilatationsreflex die Pulse unregelmäßig und scharf abgegrenzt sind. Hier folgt einer kurzfristigen und schnellen Weitung der Pupille auch eine Verkleinerung (Marshall, 2000). Das hier beschriebene Verfahren des ICA sucht nun mit entsprechenden Verfahren nach diesen unregelmäßigen, kurzfristigen, starken und schnellen Weitungen in der Pupillengröße als Korrelat für eine erhöhte mentale Aktivität. Je mehr solcher kurzfristigen Änderungen pro Sekunde in dem Signal auftreten, desto höher ist die vorliegende Beanspruchung.

4.3.2 Datenanalyse

Am Ausgangspunkt des hier beschriebenen Verfahrens steht die Information über die Größe der Pupille. Die Datengrundlage bietet dabei - wie auch im gesamten Verlauf der vorliegenden Arbeit verwendet - die über einen aktiven Infrarot Eyetracker aufgezeichnete Pupillengröße getrennt für das linke und das rechte Auge. Die Datenaufzeichnungsrate beträgt 250 Hz mit einer Genauigkeit der Messung der Pupillengröße von 0,01 % (für eine detaillierte Beschreibung des in der vorliegenden Arbeit verwendeten Eyetrackers s. Kapitel 5.2).

Um die kurzfristigen und schnellen Änderungsanteile in der gemessenen Pupillengröße identifizieren zu können, wird die Information zunächst das Rohsignal der aufgezeichneten Pupillengröße nach Blinzlern und Teil-Blinzlern, also ein kurzfristiges komplettes Verschließen oder ein teilweises Verschließen des Augenlids bereinigt. Ein solcher Blinzler dauert normalerweise circa 70-100 ms. In dieser Zeit können die Kameras die Pupille nicht erfassen und an den entsprechenden Stellen im Signal treten Pupillengrößen von null auf. Um dennoch eine reibungslose Auswertung der Daten zu gewährleisten, wird hier ein Verfahren

angewandt, bei dem diese Stellen im Signal detektiert und eliminiert werden. Zusätzlich kann es bei solchen Blinzlern auch kurz vor oder nach dem Blinzeln durch ein teilweises Verdecken der Pupille zu unrealistisch kleinen Werten kommen. Für eine weitere Auswertung werden daher bei dem hier beschriebenen Verfahren die Pupillendaten während eines Blinzler oder Teil-Blinzler durch eine lineare Interpolation der Daten ersetzt. Dabei werden auch die einem Blinzler vorangegangenen und die darauffolgenden 6 Messzeitpunkte auf Extremwerte überprüft und gegebenenfalls ebenso durch einen entsprechenden interpolierten Wert ersetzt. Insgesamt treten Blinzler in ungefähr 3-4 % der beobachteten Zeit auf, Teil-Blinzler in weiteren 1 % der Zeit.

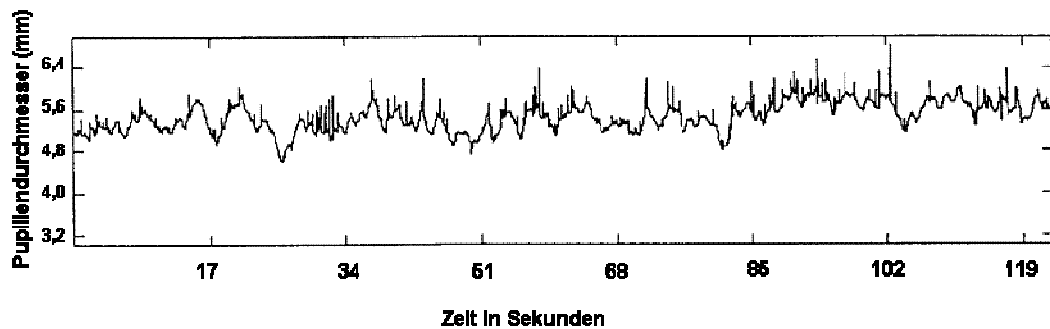


Abbildung 7: Die Abbildung zeigt exemplarisch den Verlauf der gemessenen Pupillengröße für eine Person über einen Zeitabschnitt. Ein solches Signal ist Ausgangspunkt der Berechnung des ICA (nach Marshall, 2000).

Um die gesuchten Änderungen in der Größe der Pupille identifizieren zu können, kommt bei dem Verfahren des „Index of Cognitive Activity“ eine entsprechende Wavelet Analyse zum Einsatz. Eine Wavelet Analyse besteht aus der wiederholten orthogonalen Transformation eines Signals. Das Ziel dabei besteht darin, das ursprüngliche Signal in eine Vielzahl unabhängige Komponenten zu zerlegen, wobei jede dieser Komponenten für sich analysiert und interpretiert werden kann. Der Vorteil einer solchen Wavelet Transformation gegenüber anderer Frequenzanalysen, wie etwa der Fourier Transformation, ist der, dass entsprechend der jeweils betrachteten Frequenz eines Signal der betrachtete Zeitabschnitt angepasst werden kann. Die Basis der Transformation stellt hier nicht etwa eine unbegrenzte

Sinus- und Kosinus-Funktion, sondern eine lokale begrenzte Funktion, dar. Somit eignet sich die Wavelet Analyse für den Einsatz von Signalen mit kurzen und sich abrupt ändernder Charakteristiken. Mittels einer Wavelet Analyse ist eine präzise Lokalisation solcher Diskontinuitäten im Signalverlauf möglich.

Die Basis für eine Wavelet Analyse stellt dabei eine sogenannte „Mutter-Wavelet“ Funktion dar. Diese ist eine kleine Funktion, die oszillierend sowohl in positive wie auch in negative Richtung schnell zu null abfällt.

In der Literatur findet sich eine ganze Reihe unterschiedlicher Familien solcher Mutter-Wavelet Funktionen. Bei der Berechnung des hier beschriebenen Index of Cognitive Activity wird die Familie der sogenannten Daubechies Wavelets (nach Ingrid Daubechies, 1988) verwendet. Diese entsprechen mit ihrer oszillierenden Form natürlichen physiologischen Signalen, wie sie in den Signalverläufen für die Pupillengröße zu finden sind. Solche Daubechies Mutter-Wavelets können unterschiedliche Größen sowie Formen annehmen, welche wiederum die Anzahl der in der Analyse auftretenden Koeffizienten bestimmen.

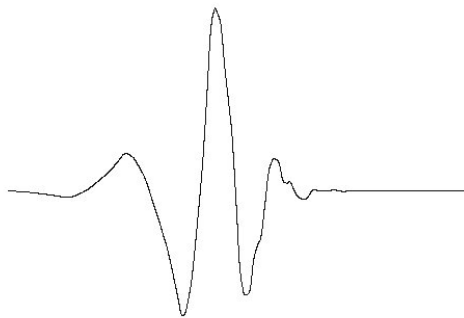


Abbildung 8: Die Abbildung zeigt exemplarisch ein Daubechies-Wavelet achten Grades.

Bei der Bestimmung des ICA wird nun ein Daubechies Mutter-Wavelet des achten Grades (siehe Abbildung 8) mit insgesamt 16 Koeffizienten verwendet (für eine genaue Aufstellung der Koeffizienten s. Marschall, Davis & Knust, 2004).

Die Koeffizienten beschreiben zusammen mit dem Grad die Form und Größe des entsprechenden Mutter-Wavelets.

Bei der Wavelet Analyse wird nun iterativ eine Transformation des ursprünglichen Signals in einzelne Wavelet-Koeffizienten zusammen mit einer Skalierungsfunktion durchgeführt. Dabei beschreiben die einzelnen Koeffizienten den Anteil des Signals welcher jeweils durch eine bestimmte Kombinationen von Ausdehnung und Verschiebung des Mutter-Wavelets (sog. Baby-Wavelets) innerhalb eines Zeitintervalls beschrieben werden kann. Die so gebildeten Koeffizienten sind orthogonal zueinander und bilden den Satz der sog. Wavelet Koeffizienten.

Bei der Bestimmung des ICA wird nun mit dem zuvor spezifizierten Wavelet die erste Iterationsstufe der Transformation durchgeführt. Diese erste Transformation des Signals erfasst die größten und schnellsten Änderungen bzw. Diskontinuitäten im ursprünglichen Signal und bildet diese mittels entsprechender Koeffizienten ab. In einem nächsten Schritt wird nun all die Information ermittelt, welche durch diesen ersten Transformationsschritt nicht erfasst wurde. Das Ergebnis ist eine geglättete Version des ursprünglichen Signals, ohne die in der Transformation erfassten Diskontinuitäten. Eine Differenz zwischen der geglätteten und der ursprünglichen Version des Signals ergibt daraufhin das sogenannte „Detail-Set“ (s. Abbildung 9). Dieses ist Ausgangspunkt der weiteren Bestimmung des ICA.

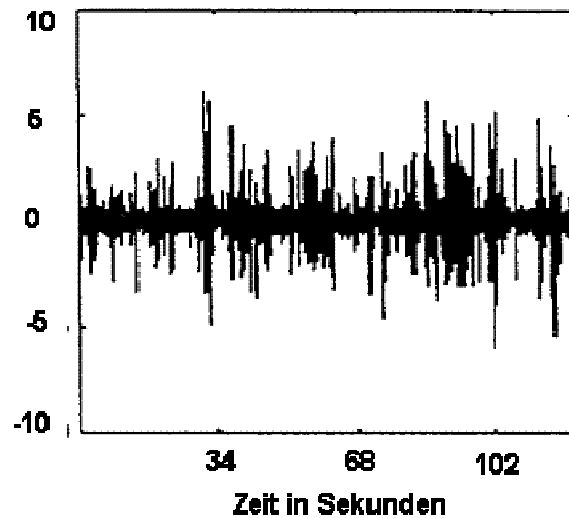


Abbildung 9: Die Abbildung zeigt die ermittelten Differenzen zwischen dem Original Signal und der für die Hoch-Frequenten Änderungsanteile bereinigten Signal. (nach Marshall, 2000).

Da nun, wie bereits weiter oben beschrieben, für die Berechnung des ICA nur die kurzfristigen und schnellen signifikanten Vergrößerungen der Pupille betrachtet werden sollen, werden in einem zusätzlichen Schritt der Analyse nur die Datenpunkte bestimmt, welche zeitlich in dem Originalsignal auch mit einer Erweiterung der Pupillengröße assoziiert sind. Somit wird in einem weiteren Schritt der Datenanalyse der Verlauf das gewonnene Detail-Set an dem ursprünglichen Pupillensignal referenziert. Es werden durch einen entsprechenden Filter nur die Stellen identifiziert, bei denen es in dem ursprünglichen Verlauf der Pupillengröße auch zu einer Vergrößerung der Pupille gekommen ist.

In einem weiteren Schritt der Datenanalyse wird daraufhin ein entsprechendes Threshold Verfahren oder auch de-noising Technik (engl. ent-rauschen) auf die so ermittelten Daten angewandt. Das Ziel besteht darin, aus den ermittelten Daten die als „unrelevant“ definierten Änderungsanteile als sogenanntes Rauschen aus dem Signal heraus zu nehmen. Erreicht werden kann somit eine Konzentration der Daten auf die kurzfristigen und signifikanten Änderungsanteile des gemessenen Signals. Für ein solches Threshold Verfahren werden in der Literatur verschiedene Ansätze diskutiert (Ogden & Parzen, 1996). In dem vorliegenden Fall kommt das

Verfahren der „minimax threshold estimation“ (Donoho & Johnstone, 1994) zum Einsatz. Ein solches Verfahren entfernt alle Wavelet Koeffizienten, welche unterhalb eines kritischen, das heißt signifikanten Niveaus liegen. Dieses wird über eine angenommene Normalverteilung des Fehler bzw. des Rauschens über den gemessenen Werten bestimmt. Verändert wird das Kriterium hier über die Form, das heißt Breite der zu Grunde liegenden Verteilung. Für eine genaue Beschreibung sowie Spezifikation der für die Bestimmung des ICA verwendeten Threshold Parameter sei an dieser Stelle auf die Arbeiten nach Marshall et al. verwiesen (Marshall, Davis & Knust, 2004; Marshall, 2000).

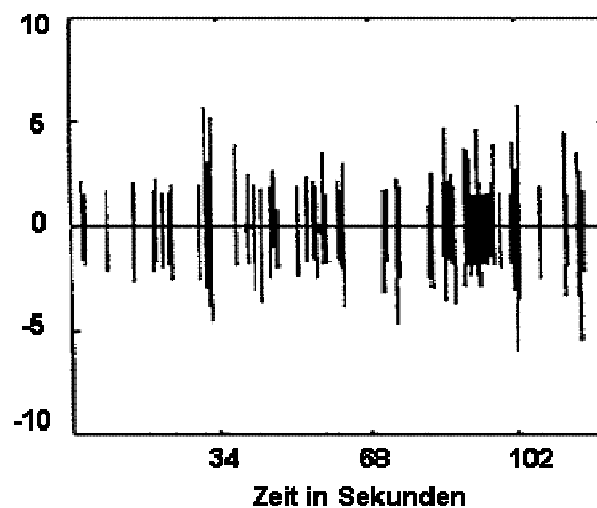


Abbildung 10: Die Abbildung zeigt die ermittelten Signaldetails nach Applikation der entsprechenden Filter (nach Marshall, 2000).

Der in der vorliegenden Arbeit diskutierte ICA Wert bezieht sich nun auf die Anzahl der so ermittelten kurzfristigen und entsprechend signifikanten Änderungen in der Pupillengröße innerhalb eines bestimmten Zeitraums. Während prinzipiell zu jedem betrachteten Messzeitpunkt ein solcher signifikanter Wert auftreten kann, bedeutet dies, dass bei einer Datenaufzeichnungsrate von 250 Hz und einem betrachteten Zeitraum von 1 Sekunde der ICA Index prinzipiell einen Wert zwischen 0 und 250 annehmen kann. Der hier gewählte Zeitraum von 1 Sekunde ist

prinzipiell variabel und kann auch kleiner gewählt werden. Er wird lediglich durch die zeitliche Auflösung der Messung begrenzt.

4.3.3 Experimentelle Evidenz

Die Wirksamkeit der Methode des „Index of Cognitive Activity“ (ICA) bei der Bestimmung von mentaler Beanspruchung konnte bereits in einer Reihe unterschiedlicher Studien nachgewiesen werden. Im Folgenden sollen nun die bereits bestehenden experimentellen Evidenzen genauer beschrieben und diskutiert werden.

Beleuchtungseinflüsse

Wie bereits bei der Konzeption des Verfahrens diskutiert, liegt einer der Hauptvorteile des Verfahrens des ICA in der Unabhängigkeit des Maßes von Beleuchtungseinflüssen. Bereits in mehreren Studien konnte die Unabhängigkeit des hier beschriebenen Verfahrens des ICA von Unterschieden in der Menge des in die Pupille einfallenden Lichtes und somit dem Pupillenlichtreflexes experimentell bestätigt werden (Marshall, Davis, Knust, 2004; Marshall, 2000, Marshall, 2007). Hier wurden der ICA in unterschiedlich beanspruchenden Situationen und unter unterschiedlichen Beleuchtungsbedingungen erhoben. Die Studien zeigten ohne zusätzliche Beanspruchung keine Unterschiede in den gemessenen ICA Werten, trotz sich verändernder Beleuchtungsbedingungen. Bei zusätzlicher Beanspruchung durch die Bearbeitung unterschiedlicher kognitiver Aufgaben zeigte sich jeweils nur ein Einfluss der unterschiedlich schweren Aufgaben, nicht aber der unterschiedlichen Beleuchtungsbedingungen. Die vorliegenden Studien zeigen, dass der hier beschriebene Index unter nahezu allen normalen Beleuchtungsbedingungen einsetzbar ist.

Validierung anhand TEPRs

Eine Studie von Marshall et al. (Marshall, Davis & Knust, 2004) zeigte die Validierung des ICA anhand der bereits zuvor beschriebenen und bei pupillometrischen Untersuchungen am häufigsten eingesetzten Methode der „Task-evoked Pupillary Responses“ (TEPR) (Kahnemann & Beatty, 1966; Beatty & Lucero-Wagoner, 2000). Verwendet wurde hier eine visuelle digit-span Aufgabe, bei der unterschiedlich lange Sequenzen (2 bis 7 Items) von einstelligen Zahlen (1-9) in der Reihenfolge memoriert werden mussten. Einflüsse von Beleuchtungsfaktoren auf die Pupillengröße wurden experimentell ausgeschlossen.

Ermittelt wurde hier pro Sequenzgröße der mittlere ICA, der mittlere Anstieg (mean dilation) in der Pupillengröße gemessen an der vor jeder Bedingung ermittelten Baseline sowie die mittlere maximale Pupillendilatation.

Die Ergebnisse zeigten für alle drei ermittelten Maße einen klaren signifikanten linearen Anstieg über die Sequenzlänge und zeigten somit die Sensibilität der drei Maße für die experimentelle Manipulation der Beanspruchung. Eine Korrelationsanalyse zeigte darüber hinaus hoch signifikante Zusammenhänge zwischen den drei ermittelten pupillometrischen Maßen.

Kognitive Aufgaben

Insgesamt zeigte sich der ICA als ein sensibles Maß für die Erfassung auftretender mentaler Beanspruchungen für unterschiedliche grundlegende kognitive Aufgaben.

Der ICA zeigte sich auch sensibel für die Bearbeitung von einfachen mathematischen Aufgaben (Marshall, Davis & Knust, 2004). Hier wurden einfache Additionsaufgaben vom Typ $14 + 18$ verwendet. Die Probanden antworteten verbal. Es zeigten sich hier klare Unterschiede in den ermittelten ICA Werten zwischen den Situationen mit und ohne Bearbeitung von Aufgaben. Bei einer weiteren Studie (Marshall, 2007) zeigte sich ebenfalls ein solcher Unterschied zwischen Situatio-

nen mit Bearbeitung einer mathematischen Aufgabe und ohne. Hier bestanden die Aufgaben aus Additionen, Subtraktionen sowie Multiplikationen.

Eine weitere Studie zeigte die Sensibilität des ICA für auftretende Beanspruchung bei der Memorierung unterschiedlicher Information und Stimuli. So konnten Marshall, Davis & Knust (2007) in der bereits zuvor erwähnten Studie auch in einer Sequenzaufgabe die Sensibilität des ICA für unterschiedlich lange Sequenzen und der somit auftretenden Beanspruchung bei der Memorierung von Informationen nachweisen. Getestet wurden Sequenzen von 2 bis 7 Items mit einstelligen Zahlen, Buchstaben, Farben und einfachen geometrischen Formen. Die Ergebnisse zeigten sowohl für die Zahlen, Buchstaben, Farben als auch die geometrischen Formen einen klaren Anstieg in den ermittelten ICA Werten über die Zunahme in der Sequenzlänge. Somit zeigte sich der ICA hier als sensibles Maß für auftretende Beanspruchungen bei der Memorierung von einer Vielzahl von Stimuli.

Evidenzen aus anwendungsnahen Fragestellungen

Neben dem Nachweis der Wirksamkeit der Messung mentaler Beanspruchung in einem kontrollierten experimentellen Kontext wie etwa der Memorierung unterschiedlicher Stimuli oder der Bearbeitung unterschiedlicher Mathematik Aufgaben, zeigte sich der ICA auch sensibel für die Detektion auftretender mentaler Beanspruchungen in sehr viel komplexeren sowie anwendungsnahen Situationen.

So zeigte eine Studie (Marshall, 2007) auftretende Beanspruchungen mittels des ICA bei einer taktischen militärischen Entscheidungsfindung im Team. Verwendet wurde hier die „Distributed Dynamic Decision Making Simulation“ nach Kleinman, Young und Higgins (1996). Dabei löst ein Team von Entscheidungsträgern eine bestimmte Mission mit zuvor definierten Zielen unter Berücksichtigung bestimmter zur Verfügung stehender Informationen. Bei der hier beschriebenen Studie wurden zusätzlich die Teams in Ihrer Organisation entweder passend

zur jeweiligen Aufgabe trainiert oder aber unpassend. Eine der Missionssituation unangepasste Vorbereitung zeigte sich daraufhin in einem deutlich höheren ICA Wert bei den jeweiligen Beteiligten in den entsprechenden Entscheidungssituationen, begleitet von einem erheblich veränderten Kommunikationsverhalten. Beides wird hier als eine Kompensationsreaktion für eine unpassende Vorbereitung gesehen.

Eine weitere Studie von Marshall et al. (Marshall, Pleydell-Pearce & Dickson, 2002) konnte zeigen, dass mittels des ICA strategische Aufmerksamkeitseffekte in einer anwendungsnahen Regelaufgabe nachgewiesen werden können. Verwendet wurde hier ein Szenario bei dem die Probanden auf einer Anzeigetafel insgesamt fünf Anzeigeelemente kontinuierlich überwachen und bei Bedarf regelnd eingreifen mussten. Die Schwierigkeit der Aufgabe variiert dabei mit der Geschwindigkeit der Veränderung in den jeweiligen Anzeigen und der damit verbundenen Notwendigkeit der Kontrolle und Korrektur. Die Anzeigetafel war dabei so aufgebaut, dass es sowohl eine numerische Anzeige der jeweiligen Elemente gab als auch eine entsprechende Warnleuchte für jedes der fünf Elemente, die die Notwendigkeit einer manuellen Korrektur in eine bestimmte Richtung (kodiert über die Farbe) durch den Operator anzeigte. Die Aufgabe der Probanden bestand darin die jeweiligen Anzeigen in einem bestimmten Bereich zu halten. Bei zunehmender Übung durch Wiederholungen und zunehmender Schwierigkeit der Aufgabe zeigte sich bei den Probanden eine Veränderung in der Strategie der Überwachung der einzelnen Anzeigen. Während zu Beginn hauptsächlich die numerischen Werte betrachtet wurden um zu entscheiden ob eine Korrektur notwendig war oder nicht zeigte sich bei zunehmender Übung eine Konzentration auf die entsprechenden Warnleuchten. Somit erreichten die Probanden eine Reduzierung der zu verarbeitenden Information, um eine Entscheidung über ein Eingreifen zu treffen. Während die Beobachtung der numerischen Zahlen sowohl eine Verarbeitung der Information der Zahl sowie eine Einschätzung und eine Entscheidung über das

Eingreifen verlangt, zeigt die entsprechende Warnleuchte direkt die Notwendigkeit eines Eingreifens sowie die entsprechende Richtung an.

Die Ergebnisse der Studie zeigten einen deutlichen Effekt des hier beschriebenen strategischen Wechsels bei der Bearbeitung der Aufgabe sowohl in dem beobachteten Blickverhalten der Probanden als auch den ermittelten ICA Werten. Wechselten die Probanden zu einer Strategie bei der sie nur noch die Warnleuchten betrachteten, so reduzierte sich gleichzeitig der beobachtete ICA Wert der Probanden. Die Studie zeigte somit, dass mittels des ICA hier nicht nur auftretende Beanspruchungen in einem Aufgabenverlauf abgebildet werden können, sondern dass darüber hinaus auch komplexe Strategien in Bezug auf die Beanspruchungssteuerung der Probanden dargestellt werden können.

Zusammenfassend lässt sich festhalten, dass die Wirksamkeit der Methode des Index of Cognitive Activity bei der Bestimmung von mentaler Beanspruchung bereits in einer Reihe von Studien belegt und als Methode validiert werden konnte. Die hier beschriebenen Studien zeigen, dass der ICA anders als bislang existierende Verfahren der Pupillometrie unabhängig von sich verändernden Beleuchtungseinflüssen ist, aber dennoch die gleichen Ergebnisse wie die bereits existierenden und über viele Experimente bestätigten Verfahren der Pupillometrie liefert. Durch diese Unabhängigkeit des Verfahrens von externen Beleuchtungseinflüssen konnte das Verfahren bereits für verschiedene komplexe sowie anwendungsnahe Szenarien erfolgreich eingesetzt werden.

4.3.4 Zusammenfassung Index of Cognitive Activity

Insgesamt zeigt sich der hier beschriebene Index of Cognitive Activity (ICA) als ein äußerst vielversprechendes pupillometrisches Verfahren zur Bestimmung der mentalen Beanspruchung. Insbesondere durch die Unabhängigkeit des Verfahrens von externen Beleuchtungseinflüssen, der hohen zeitlichen Auflösung sowie der Kontinuität der Messung erscheint das hier beschriebene Verfahren geeignet um

insbesondere in einem anwendungsorientierten Kontext, wie dem einer realistischen Fahrsituation, die auftretenden Beanspruchungen für den Fahrer abbilden zu können.

Ein Aspekt welcher bei bisherigen Studien zum ICA dabei noch nicht genauer betrachtet wurde, jedoch auch nicht im Fokus dieser Arbeit steht, aber an dieser Stelle erwähnt werden sollte, ist der Einfluss eines zuvor beschriebenen Nah- Fern Adaptationsprozesses auf die Pupillengröße. Für den Nachweis der prinzipiellen Anwendbarkeit des Verfahrens wurde in der vorliegenden Arbeit ein experimentelles Setting verwendet, bei dem alle relevanten Stimuli auf einer Ebene mit einer konstanten Entfernung zum Betrachter abgebildet wurden. Die Überprüfung des Einflusses eines Nah- Fern Adaptationsprozesses auf den ICA steht dabei nicht im Fokus der vorliegenden Arbeit, sollte jedoch in weiteren Studien zum ICA untersucht werden.

Der ICA verspricht - anders als herkömmliche Beanspruchungsmaße - eine direkte störungsfreie Messung der Beanspruchung des Fahrers. Mit der durch das Verfahren zur Verfügung stehenden zeitlichen Genauigkeit ergibt sich darüber hinaus die Möglichkeit, Bedienabläufe bzw. Handlungsabläufe in einer zeitlichen Genauigkeit hinsichtlich der für den Fahrer auftretenden Beanspruchungen mit bisher unerreichter zeitlicher Genauigkeit zu analysieren und Optimierungspotential aufzuzeigen.

Die Zielsetzung der vorliegenden Arbeit ist es daher den hier beschriebenen vielversprechenden Ansatz des Index of Cognitive Activity in die Anwendung bei der Entwicklung neuer Anzeige- sowie Bedienkonzepte im automotiven Kontext zu bringen. Die Methode bietet dabei das Potential für eine Optimierung solcher Konzepte in einer Genauigkeit und Granularität wie sie mit bisherigen Methoden nicht erreicht werden konnte.

Teil 2

5 Empirischer Teil

5.1 Allgemeines

Im empirischen Teil dieser Arbeit sollen vier Experimente zur Anwendung des ICA bei der Bestimmung von auftretender mentaler Beanspruchung beschrieben werden. Bei den ersten zwei Experimenten handelt es sich um eine Validierung des ICA Verfahrens anhand von Mathematikaufgaben unterschiedlicher Schwierigkeit und Darbietungsmodalität. Experiment 2 beschreibt die erste Anwendung des ICA in einer Studie mit einer einfachen Fahrsimulation mit und ohne eine zusätzliche Aufgabe sowie unterschiedlichen Instruktionsbedingungen. In Experiment 3 wird die Methode des ICA genutzt, um unterschiedliche Beanspruchungsfaktoren in einer simulierten Fahraufgabe zu identifizieren. Zum Einsatz kommen hier unterschiedliche Nebenaufgaben zusätzlich zur Fahraufgabe in unterschiedlichen Schwierigkeitsstufen. In diesem Experiment sollte zusätzlich auch der Nachweis von strategischen Aufmerksamkeitseffekten mittels des ICA überprüft werden. Experiment 4 beschreibt schließlich die Anwendung des ICA bei der Identifikation von Beanspruchungsprofilen bei der Durchführung von realen Aufgaben in unterschiedlichen Systemen. Zum Einsatz kamen hier zwei unterschiedliche Systeme sowie Eingabemodalitäten für alphanumerische Zeichen. Experiment 5 dient somit dem Nachweis der Wirksamkeit der Pupillometrie und insbesondere des Index of Cognitive Activity als Methode zur Messung mentaler Beanspruchung bei der konkreten Gestaltung und Optimierung von Anzeige- und Bedienkonzepten im automotiven Kontext.

Sämtliche in dieser Arbeit beschriebenen Experimente wurden ausschließlich im Usability Labor der BMW AG in München durchgeführt.

5.2 Verwendeter Eyetracker

Bei den in der vorliegenden Arbeit diskutierten Experimenten zur Bestimmung des Index of Cognitive Activity kam ausschließlich das Eyetracking Gerät Eye-Link 2 von Eye-Tracking Systems (SR Research, Ltd.; Mississauga, Ontario, Canada) zum Einsatz. Bei dem Gerät handelt es sich um einen binokularen Eyetracker mit zwei Videokameras, welche an einem leichtgewichtigen Kopfband befestigt sind. Die Kameras sind ausgestattet mit Infrarot Beleuchtungseinheiten, welche mittels einer digitalen Steuerung eine gleichmäßige Ausleuchtung und somit einen gleichbleibenden Kontrast der aufgezeichneten Bilder auch bei wechselnden Bedingungen ermöglichen. Das Gerät liefert eine zeitliche Auflösung der Messung bis zu 500 Hz, verwendet wurden in den hier beschriebenen Studien zur Bestimmung des Index of Cognitive Activity jedoch eine zeitliche Auflösung von 250 Hz. Das Gerät liefert eine Blickerfassung mit einer Genauigkeit von weniger als $0,01^\circ$ und eine Aufzeichnung des Pupillendurchmessers mit einer Genauigkeit von 0,01 %.



Abbildung 11: Die Abbildung zeigt den für die Bestimmung des ICA verwendeten Eyetracker Eyelink 2

Die Erfassung der Pupillengröße geschieht bei dem Gerät EyeLink 2 über eine Bilderkennung für den Farbkontrast der Pupille. In Abbildung 12 dargestellt durch die blaue Fläche.

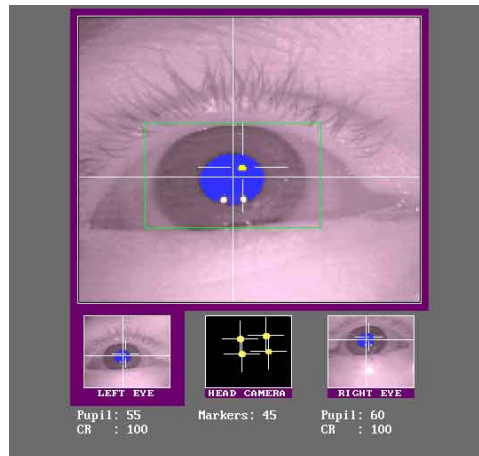


Abbildung 12: Die Abbildung zeigt die mittels des Geräts EyeLink 2 aufgezeichneten Infrarot Bild des Auges. Die Erkennung und Vermessung der Pupille geschieht mittels einer Farbkontrastbestimmung (blaue Fläche).

Der Vorteil des hier verwendeten kopfbasierten Eyetrackers liegt in einer erhöhten Präzision der Vermessung der Pupille gegenüber herkömmlichen remote Eyetrackern, welche eine Erfassung der Pupille durch Kamerasysteme vornehmen, welche auf dem Tisch vor der Versuchsperson installiert sind und über Teleobjektive die Pupille erfassen. Darüber hinaus bietet ein solcher Eyetracker auch in Situationen bei denen die Probanden natürlicherweise neben der Bewegung der Augen auch eine Bewegung des Kopfes durchführen immer noch ein exaktes Bild der Pupille.

Die Erfassung des Blickpunktes geschah mittels einer zusätzlichen, auf dem Kopfband befestigten Kamera. Diese ermöglicht die Erfassung der Blickrichtung aus der Perspektive der Versuchsperson in Echtzeit zusammen mit den aus den beiden Infrarot Kameras gewonnen Information über den exakten Blickpunkt. Die Kopfposition wurde mittels entsprechender Infrarot Markern auf der Ebene des zu betrachtenden Monitors ermittelt. Die Kalibrierung des Eyetrackers geschah vor

jeder Versuchsdurchführung anhand einer 9-Punkt-Methode auf dem zu betrachtenden Bildschirm.

5.3 Datenaufbereitung

Die Aufbereitung der pupillometrischen Signale sowie der Eyetracking Informationen geschah mittel drei Standard PCs und der entsprechend Software.

Ein PC ausgestattet mit einer spezialisierten Software diente der Steuerung und Aufzeichnung des Eyetrackers sowie der Erfassung und Artefaktbereinigung der pupillometrischen Daten.

Die pupillometrischen Daten wurden online an einen weiteren PC via LAN übermittelt. Dieser ermittelte nach dem in Kapitel 4.3 beschriebenen Verfahren den aktuellen ICA Index basierend auf den gemessenen pupillometrischen Daten getrennt für beide Augen. Für die Berechnung des ICA kam dabei ein auf dem Softwarepaket MATLAB Wavelet Toolbox von MathWorks, Inc. basierendes Programm nach Marshall (2000) zum Einsatz.

Bei den in der vorliegenden Arbeit diskutierten ICA Werten handelt es sich um die für das linke und das rechte Auge ermittelten und pro Sekunde aufsummierten ICA Werte. Bei einem Teil der hier berichteten Experimente wurde zusätzlich eine interindividuelle Standardisierung der ermittelten ICA Werte durchgeführt, um so Differenzen in dem absoluten Niveau der gemessenen Beanspruchungen zwischen Personen herauszunehmen.

Der dritte in dem vorliegenden experimentellen Setting verwendete PC wurde zur zeitlichen Synchronisation der Aufzeichnung der Daten sowie der Aufzeichnung von Videos verwendet.

6 Experiment 1: Untersuchung zur grundlegenden Validierung des experimentellen Settings und des ICA

Das Ziel des ersten in dieser Arbeit berichteten Experimentes war es, zunächst eine grundlegende Validierung des bereits weiter oben beschriebenen experimentellen Laboraufbaus zur Bestimmung des Index of Cognitive Activity zu erreichen. Dafür wurden in dem vorliegenden Experiment unterschiedliche Arten von Kopfrechenaufgaben in unterschiedlichen Schwierigkeitsstufen (einfach, mittel, schwer) untersucht. Zusätzlich wurde in dem vorliegenden Experiment die Sensibilität der Methode gegenüber der Darbietungsmodalität (auditiv, visuell) der Aufgaben überprüft.

6.1 Hintergrund

Bei den in diesem Experiment verwendeten Mathematik Aufgaben lag die Annahme zugrunde, dass eine „einfache“ Rechenaufgabe weniger mentale Beanspruchung erzeugt als eine „schwere“ Aufgabe. Es wird ferner angenommen, dass die Lösung der in dem vorliegenden Experiment verwendeten Aufgaben hauptsächlich eine Beanspruchung einer zentral-kognitiven Prozessressource darstellt. Es sollten somit Störeffekte aufgrund modalitätsspezifischer Stimuluseigenschaften weitestgehend minimiert werden.

Hauptziel der ersten hier berichteten experimentellen Studie war es zu zeigen, dass mittels des ICA modalitätsunspezifisch mentale Beanspruchung nachzuweisen ist und die Methode somit Situationen unterschiedlich hoher Beanspruchungen mit einem sehr kontrollierten Stimulusmaterial identifizieren kann. Dafür wurde zusätzlich zur Schwierigkeit auch die Darbietungsmodalität variiert. Es wird davon ausgegangen, dass in einer auditiven Instruktionsmodalität, bei der die

Probanden die entsprechende Aufgabe lediglich auditiv über einen Kopfhörer präsentiert bekommen, visuelle Einflüsse auf die Pupille weitestgehend ausgeschlossen werden können.

6.2 Fragestellung

Die hier verwendete Manipulation der Schwierigkeit der Aufgaben lag der Annahme zugrunde, dass eine „einfache“ Rechenaufgabe für den Probanden eine geringere mentale Beanspruchung erzeugt als eine „schwere“ Aufgabe, wenn die zur Verfügung stehende Zeit eine Lösung zu erarbeiten konstant gehalten wird.

Mit der Methode des ICA sollte es möglich sein, unterschiedliche Schwierigkeitsstufen bei Kopfrechenaufgaben und die damit einhergehende unterschiedlich hohe mentale Beanspruchung für den Probanden zu unterscheiden.

Aufgrund der Annahme, dass die Pupillometrie und somit auch der ICA als Methode, wie bereits in Kapitel 4 diskutiert, ein Maß für eine global-zentrale Beanspruchung darstellt und somit unabhängig von der verwendeten Stimulusmodalität ist, ergibt sich eine weitere Forderung für die hier diskutierte experimentelle Manipulation.

Die Methode des ICA sollte modalitätsunspezifisch unterschiedlich hohe zentral-kognitive mentale Beanspruchungen für den Probanden abbilden können.

6.3 Methode

6.3.1 Verwendetes Material

Bei dem vorliegenden Experiment wurden Kopfrechenaufgaben in Form von mathematischen Gleichungen als Stimulusmaterial verwendet. Hier mussten Probanden mittels Tastendruck angeben, ob das dargestellte Ergebnis der Gleichung richtig oder falsch war. Ein solches Unterscheidungsparadigma wurde gewählt, um die konfundierenden Einflüsse durch eine motorische Handlung oder eine verbale Reaktion bei der Angabe der Lösung durch den Probanden auf die pupillometrische Messung so gering wie möglich zu halten.

Die drei unterschiedlichen Schwierigkeitsstufen wurden durch Additions-, Multiplikations- sowie Moduloaufgaben erzeugt. Ähnliche Aufgaben wurden bereits in anderen Studien zur Pupillometrie verwendet und validiert (Rösger, 1996) und erschienen daher an dieser Stelle als geeignet.

Schwierigkeit	Aufgabentyp	Beispiel
Einfach	Additions- und Subtraktionsaufgaben	$15 + 20 = 35$
Mittel	Multiplikation- und Divisionsaufgaben	$12 * 4 = 48$
Schwer	Moduloaufgaben	$37 \bmod 3 = 1$

Tabelle 3: Die Tabelle zeigt die in Experiment 1 verwendeten Aufgabentypen in drei Schwierigkeitsstufen mit jeweiligen Beispielen

Die Darbietung der Aufgaben erfolgte in jeder der drei Schwierigkeitsstufen entweder visuell auf einem Bildschirm vor dem Probanden in geschriebener Form oder auditiv verbal in gesprochener Form über Kopfhörer.

6.3.2 Ablauf

Zunächst wurde jede Versuchsperson über das Ziel sowie dem Ablauf der aktuellen Untersuchung informiert. Nachdem die Probanden vor dem Hauptmonitor Platz genommen hatten, wurde der Eyetracker vom Versuchsleiter aufgesetzt und entsprechend kalibriert. Es wurde sichergestellt, dass die Informationen über die Größe der Pupille bei beiden Augen verfügbar waren.

Die Darbietung der Aufgaben erfolgte in jeder der drei Schwierigkeitsstufen sowohl visuell als auch auditiv. Dabei wurde jede der Aufgaben pro Proband nur einmal bearbeitet und die Zuordnung der Aufgaben zu einer Darbietungsmodalität erfolgte randomisiert. Die Präsentation eines Aufgabentyps in einer Darbietungsmodalität erfolgte geblockt. In jeder Bedingung wurden insgesamt 20 Aufgaben eines Typs gelöst, wobei die Darbietungsgeschwindigkeit innerhalb der jeweiligen Modalitäten konstant gehalten wurde. Jedoch unterschied sich die Darbietungszeit der visuell präsentierten Aufgaben zu der der auditiv präsentierten Aufgaben. Die Reihenfolge der Darbietungsmodalität wurde zwischen den Probanden systematisch variiert. Insgesamt absolvierte somit jeder Proband 120 Aufgaben in unterschiedlichen Modalitäten und Schwierigkeiten.

Die visuelle Darbietung der Aufgaben geschah mittels einer schwarzen Schrift auf weißem Hintergrund auf dem vor dem Probanden platzierten Hauptbildschirm. Die auditive Darbietung geschah über Kopfhörer und zuvor aufgezeichneten Audiofiles. Sowohl in der visuellen als auch der auditiven Darbietungsmodalität wurden die Probanden instruiert während der gesamten Versuchsdurchführung entspannt auf den Bildschirm vor Ihnen zu schauen. Sowohl in der visuellen als auch der auditiven Bedingung wurde eine neue Aufgabe durch ein Fixationskreuz auf dem Bildschirm angezeigt. In der auditiven Bedingung wurden während der Präsentation der Aufgabe keinerlei Informationen auf dem Bildschirm dargeboten.

Dennoch waren die Probanden instruiert auch hier während der Bearbeitung der Aufgabe an die Stelle des vorher präsentierten Fixationskreuzes zu schauen.

Bei der Bearbeitung der Aufgaben und der Abgaben einer Antwort standen den Probanden in der visuellen Bedingung 5 Sekunden ab Beginn der Aufgabenpräsentation zur Verfügung. In der auditiven Bedingung wurde dieses Zeitfenster aufgrund der längeren Dauer der auditiven Aufzeichnung auf 9 Sekunden verlängert wobei die Dauer der eigentlichen auditiven Darbietung immer 5 Sekunden betrug. In der visuellen Bedingung blieb die Aufgabe auch nach Abgabe der Antwort über den Zeitraum der 5 Sekunden auf dem Bildschirm sichtbar.

Zusätzlich zu der Bearbeitung der Aufgaben wurde am Ende jeder experimentellen Bedingung die subjektiv erlebte Beanspruchung mittels des standardisierten und ins deutsche übersetzten Fragebogeninventars NASA TLX (Task Load Index) ohne Gewichtung der Dimensionen erhoben. Der NASA TLX beinhaltet sechs Dimensionen: geistige Aktivität, körperliche Aktivität, zeitliche Anforderung, körperliche und geistige Anstrengung, wie gut die Aufgabe generell erfüllt werden konnte und wie wohl sich die Person während der Aufgabenbearbeitung gefühlt hat. Die hier verwendete Version des Fragebogens findet sich im Anhang der vorliegenden Arbeit.

6.3.3 Stichprobe

Insgesamt nahmen an dieser Studie 17 Probanden teil, wovon 4 Datensätze aufgrund fehlender Daten oder anderer Probleme bei der Erfassung der Pupillengröße von den weiteren Analysen ausgeschlossen werden mussten. Insgesamt bezieht sich somit die weitere Auswertung auf insgesamt 13 gültige Datensätze von insgesamt 2 weiblichen und 11 männlichen Probanden. Das Durchschnittsalter lag bei 30,35 in einem Bereich von 23 bis 58 Jahren.

6.4 Ergebnisse

Im Folgenden sollen zunächst die Ergebnisse der pro Aufgabe ermittelten ICA Werte diskutiert werden. Hierbei ist anzumerken, dass die Berechnung des mittleren ICA Wertes pro Versuchsperson innerhalb des vollständigen Zeitintervalls der Aufgabenpräsentation erfolgte. Somit auch die Zeit nach Abgabe der Antwort bis zum Erscheinen des nächsten Fixationskreuzes beinhaltet. Die entsprechenden Abschnitte wurden nach Versuchsdurchführung mittels Videoanalyse und einem zeitlich synchronisierten Videoaufzeichnung des Bildschirminhaltes identifiziert und errechnet. Bei den hier berichteten ICA Werten handelt es sich um die innerhalb der Versuchsperson standardisierten ICA Werte für die Summe aus den für das linke als auch das rechte Auge errechneten ICA Werten.

Die Auswertung erfolgt über die Faktoren Richtigkeit der Antwort (richtige vs. falsche Antworten), Schwierigkeit der Aufgabe (einfach, mittel, schwer) sowie der Darbietungsmodalität (auditiv, visuell).

6.4.1 ICA

Insgesamt ergab sich für eine ANOVA mit Messwiederholung für die Faktoren Richtigkeit der Antwort, Schwierigkeit sowie Darbietungsmodalität kein signifikanter Haupteffekt für die Unterscheidung visuell zu auditiv ($F(1,12) = 3.9$; $p = 0.07$). Ein Post Hoc Test zeigte hier jedoch einen signifikanten Unterschied zwischen visuell und auditiv bei richtig gelösten Aufgaben ($p = 0.04$; LSD Post Hoc Test) wobei die ICA Daten für die auditiven Aufgaben über denen der visuellen Aufgaben lagen. Es zeigte sich kein signifikanter Unterschied zwischen visuell und auditiv bei den falsch gelösten Aufgaben ($p = 0.08$; LSD Post Hoc Test). Es zeigte sich ein signifikanter Haupteffekt für den Faktor Richtigkeit der Antwort ($F(1,12) = 5.4$; $p = 0.03$). Wobei die ICA Werte bei falschen Antworten insgesamt höher lagen als bei richtigen Antworten.

Es zeigte sich ein signifikanter Haupteffekt für den Faktor Schwierigkeit der Aufgabe ($F(2,24) = 20.30$, $p < .001$). Für die richtig gelösten Aufgaben ergab sich in der auditiven Darbietungsmodalität ein signifikanter Anstieg in den ICA Daten von den einfachen zu den schwierigen Aufgaben ($p < .05$; LSD Post Hoc Test) aber kein signifikanter Anstieg zwischen den einfachen und den mittleren sowie zwischen den mittleren zu den schwierigen Aufgaben ($p > .05$; LSD Post Hoc Test). Unter der visuellen Darbietungsmodalität ergab sich ebenfalls ein signifikanter Anstieg in den ICA Werten zwischen den einfachen und den schwierigen Aufgaben ($p < .05$; LSD Post Hoc Test) und kein signifikanter Anstieg zwischen den einfachen und den mittleren sowie den mittleren zu den schweren Aufgaben ($p < .05$; LSD Post Hoc Test).

Für die falsch gelösten Aufgaben ergab sich über die auditive und visuelle Darbietungsmodalität hinweg ein signifikanter Anstieg in den ICA Daten zwischen den einfachen und den mittel schweren Aufgaben ($p < .01$; LSD Post Hoc Test) und kein signifikanter Anstieg in den ICA Daten zwischen den mittelschweren und den schweren Aufgaben.

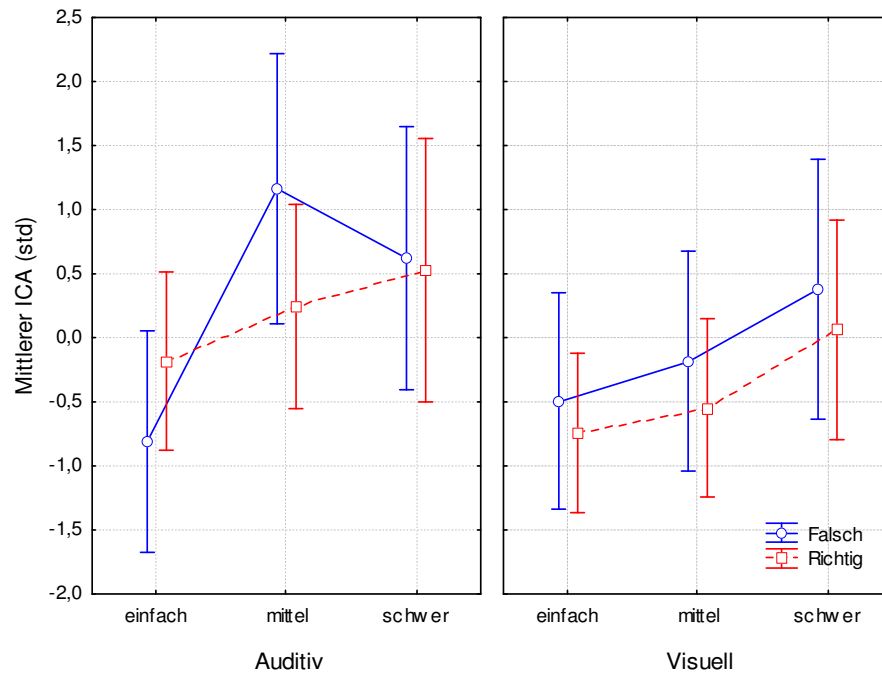


Abbildung 13: Die Abbildung zeigt die die standardisierten ICA Werte getrennt nach Modalität (visuell und auditiv), Schwierigkeit der Aufgabe sowie Richtigkeit der Antworten (richtig vs. falsche Antworten) [vertikale Balken zeigen +/- 1 Standardabweichung].

6.4.2 Leistung

Für die Leistung bei den Aufgaben an sich ergab sich durchgehend ein klares Befundmuster wie in Abbildung 14 dargestellt. Insgesamt zeigte hier eine Varianzanalyse mit Messwiederholung eine Verringerung der Anzahl richtig gelöster Aufgaben mit zunehmender Schwierigkeit der Aufgaben ($F(2, 24) = 44.60$; $p < .001$). Zwischen den Bedingungen auditiv und visuell gab es jedoch keinen signifikanten Unterschied bei der Leistung ($F(1, 24) = 3.92$; $p = .07$).

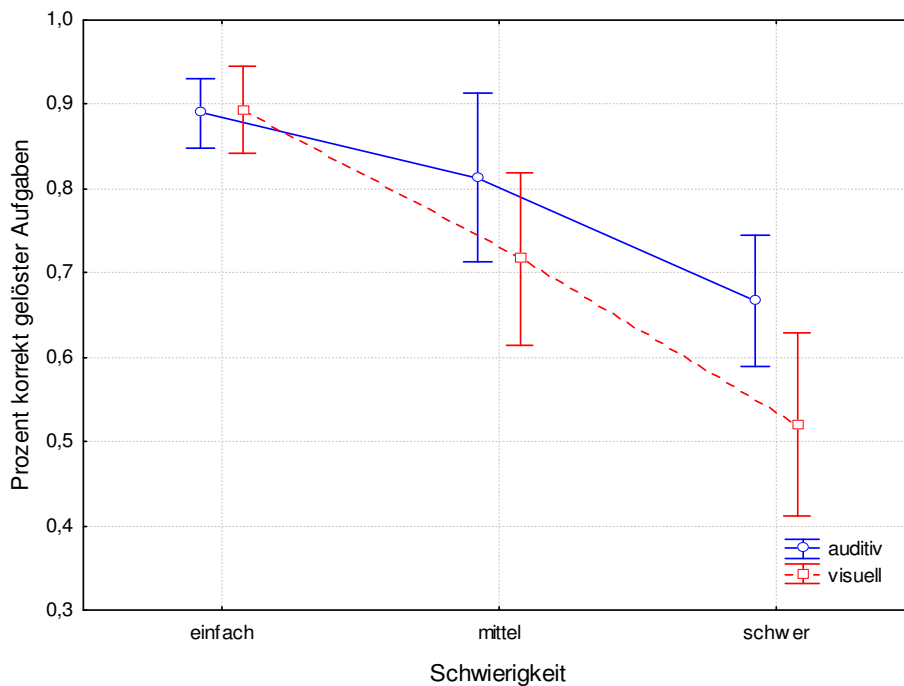


Abbildung 14: Die Abbildung zeigt den prozentualen Anteil richtig gelöster Aufgaben für die Darbietungsmodalitäten Auditiv und Visuell [vertikale Balken zeigen +/- 1 Standardabweichung].

6.4.3 Reaktionszeiten

Für die mittleren Reaktionszeiten der richtig gelösten Aufgaben zeigte sich, wie in Abbildung 15 dargestellt, insgesamt auf Aufgabenebene ein klarer Anstieg über die drei Schwierigkeitsstufen, $F(2, 24) = 45.53$, $p < .001$ sowie aufgrund der unterschiedlichen Stimuluspräsentationsdauern für visuelle und auditive Stimuli auch ein signifikanter Unterschied für die Darbietungsmodalität, ($p < .001$). Eine Post Hoc Analyse zeigte signifikante Unterschiede zwischen jeder der drei Schwierigkeitsstufen einfach, mittel und schwer sowohl für auditiv als auch visuell präsentierte Stimuli.

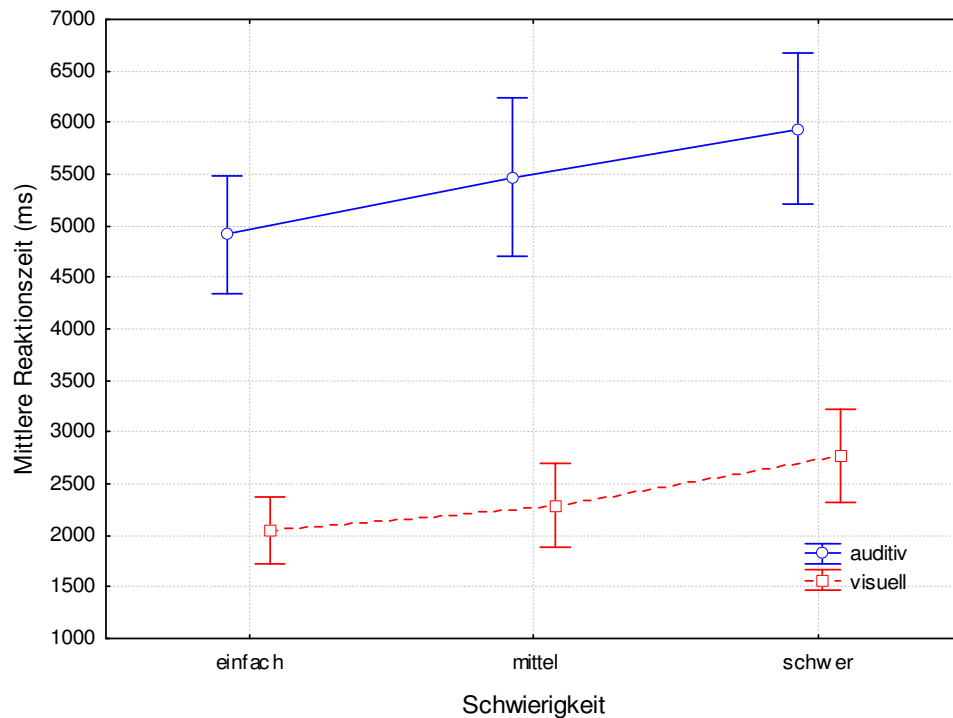


Abbildung 15: Die Abbildung zeigt die mittleren Reaktionszeiten für auditiv sowie visuell präsentierten Aufgaben getrennt nach den drei Schwierigkeiten einfach, mittel und schwer [vertikale Balken zeigen +/- 1 Standardabweichung].

6.4.4 Subjektive Daten

Bei der Analyse der im Anschluss an jede experimentelle Bedingung mittels des NASA TLX Fragebogens erhobenen subjektiven Daten ergab sich eine signifikante Separierung ($p < .05$) der drei Schwierigkeitsstufen von Aufgaben auf den Dimensionen: geistige Aktivität, körperliche Aktivität, zeitliche Anforderung, körperliche und geistige Anstrengung, sowie der Frage wie gut die Aufgabe generell erfüllt werden konnte. Eine signifikante Unterscheidung zwischen der visuellen und der auditiven Darbietungsmodalität ergab sich dabei nur für die Dimensionen der zeitliche Anforderung ($F(1, 14) = 8.43, p = 0.01$). Auf der Dimension wie wohl sich die Versuchsperson während der Versuchsdurchführung gefühlt hat, ergaben sich keine Unterschiede.

Eine Korrelationsanalyse der ermittelten ICA Werte und der zwei für die vorliegende Arbeit wichtigsten Dimensionen des NASA TLX ergaben folgende signifi-

kanten ($p < .05$) Zusammenhänge zusammen mit den ebenfalls angegebenen Beta-Gewichten aus einer linearen Regression: geistige Aktivität ($R = .32$; $b = .32$ $p < .01$); generelle Anstrengung ($R = .27$; $b = .27$ $p = .01$).

6.5 Zusammenhang subjektive Schwierigkeit und ICA

Da während der Aufgabenbearbeitung in dem hier vorgestellten experimentellen Setting keine individuelle Einschätzung der Schwierigkeit jeder einzelnen Aufgabe durchgeführt werden konnte, wurde stattdessen im Anschluss an den Versuch bei insgesamt 6 der Probanden mittels eines Fragebogens ein subjektives Rating für die empfundene Schwierigkeit einer Aufgabe auf einer Skala von 1 bis 10 für jede der 120 bearbeiteten Aufgaben durchgeführt. Für dieses so erhaltene Rating der Schwierigkeit auf Itemebene wurden zusätzlich die Korrelationen zu den ermittelten ICA Werten pro Aufgabe ermittelt. Hier zeigte sich zwischen den subjektiven Angaben und den ermittelten ICA pro Aufgabe für die richtig gelösten Aufgaben eine signifikante Korrelation von $R = .52$. Für die falsch gelösten Aufgaben hingegen zeigte sich erwartungsgemäß kein signifikanter Zusammenhang, $R = .05$.

6.6 Diskussion

Die Ergebnisse des ersten hier beschriebenen Experiments zeigen deutlich, dass es möglich ist mithilfe des ICA Situationen mit unterschiedlich hohen mentalen Beanspruchungen zu identifizieren. Darüber hinaus zeigen die Ergebnisse, dass es mit Hilfe des ICA möglich ist diese unterschiedlich hohen Beanspruchungen unabhängig von der jeweiligen Darbietungsmodalität abzubilden.

Die hier berichteten ICA Daten zeigen einen klaren Anstieg der auftretenden Beanspruchungen sowohl in der auditiven als auch der visuellen Modalität bei der Bearbeitung von einfachen Additions- und Subtraktionsaufgaben bis hin zu der Bearbeitung von Moduloaufgaben. Die Ergebnisse zeigen einen klaren Unter-

schied der auftretenden Beanspruchungen bei auditiv dargebotenen zu visuell dargebotenen Aufgaben. Dieses ist dadurch zu erklären, dass bei der Bearbeitung der auditiven Mathematik Aufgaben zusätzlich zu der mathematischen Transformation auch die Inhalte mental bis zur Erarbeitung der Lösung einer Aufgabe gehalten werden mussten. Durch das Halten von Informationen wurde eine zusätzliche Beanspruchung generiert.

Es soll an dieser Stelle noch einmal darauf hingewiesen werden, dass unter der auditiven Instruktionsmodalität keinerlei visuelle Information auf dem Bildschirm zu sehen war. Durchgänge bei denen ein Proband während der Aufgabenbearbeitung nicht auf den Bildschirm geschaut hatte wurden aus der Analyse ausgeschlossen. Die hier berichteten pupillometrischen Ergebnisse lassen sich somit nur durch eine unterschiedlich hohe kognitive Beanspruchung in der jeweiligen Situation erklären.

Unterstützt wird eine solche Aussage ebenfalls durch die hier ebenfalls beschriebenen hohen korrelativen Zusammenhänge zwischen den ermittelten ICA Daten und den post hoc ermittelten subjektiven Einschätzungen auf Itemebene. Je schwieriger eine Aufgabe empfunden wird, desto höher ist auch die mittels des ICA ermittelte Beanspruchung.

Zusammenfassend zeigen die hier diskutierten Ergebnisse eine klare Bestätigung, dass mittels des ICA modalitätsunspezifisch Situationen mit unterschiedlich hohen Beanspruchungen identifiziert werden können.

7 Experiment 2: Untersuchung zum grundlegenden Nachweis von Beanspruchungseffekten mittels des ICA bei einer simulierten Fahraufgabe

Das Ziel des hier beschriebenen zweiten Versuchs der vorliegenden Arbeit war es zu zeigen, dass mittels der Methode des ICA Situationen mit einer unterschiedlich hohen Beanspruchung für den Fahrer während einer simulierten Fahraufgabe in einer hohen zeitlichen Präzision kontinuierlich über die Zeit abgebildet werden können.

Dafür wurde in dem vorliegenden experimentellen Setting eine simulierte Fahraufgabe verwendet, bei der die Probanden auf unterschiedliche Instruktionen mit entsprechenden Fahrmanövern, das heißt Spurwechsel reagieren mussten. Variiert wurde in dem vorliegenden Experiment die Instruktionsmodalität (visuell vs. auditiv) sowie die Beanspruchung durch eine zusätzliche Nebenaufgabe zusätzlich zur Fahraufgabe (mit vs. ohne Nebenaufgabe).

Eine Veröffentlichung des hier beschriebenen Versuchs findet sich in Schwalm, Keinath & Zimmer (2008).

7.1 Hintergrund

In dem vorliegenden Experiment sollte überprüft werden, in wie weit es mittels des ICA möglich ist in einer vereinfachten, aber dennoch realistischen Fahraufgabe Situationen mit einer erhöhten Beanspruchung für den Fahrer kontinuierlich über die Zeit abzubilden. Dabei lag das Hauptaugenmerk insbesondere auf der Identifikation von Beanspruchungsspitzen im Verlauf einer Fahrt. Da solche Situationen, mit einer punktuell besonders hohen Beanspruchung, für den Fahrer ein erhöhtes Risiko für Fahrfehler bedeuten (De Waard, 1996), besteht an dieser Stelle ein besonderes Interesse solche Situationen mit einer hohen zeitlichen Genauigkeit identifizieren zu können. Erst wenn Situationskonstellationen, bei denen es zu einer erhöhten Beanspruchung für den Fahrer kommt, eindeutig und zeitlich exakt beschrieben werden können, können entsprechende Warn- und Assistenz-

systeme an die daraus abgeleiteten Anforderungen angepasst und entwickelt werden.

Für das vorliegende Experiment wurde eine vereinfachte simulierte Fahraufgabe verwendet, bei der die Probanden als Fahrer entsprechend visuell gegebener Instruktionen (über Schilder in der Nähe der Straße) die Fahrspur wechseln sollten. Eine solche Instruktion über einen Spurwechsel bedeutet für den Fahrer, dass hier zunächst die entsprechende Information auf dem Schild wahrgenommen und semantisch verarbeitet werden muss. Die so gewonnene Information muss daraufhin motorisch korrekt in einen Handlungsplan umgesetzt werden. Dabei muss das Fahrzeug so schnell, aber auch so exakt wie möglich auf die andere Fahrspur gebracht werden. Das wiederum erfordert eine visuell-motorische Kontrolle der Bewegung entsprechend der zuvor verarbeiteten Richtungsinformation. Insgesamt sollte somit der gesamte Vorgang eine deutliche Erhöhung der Beanspruchung für den Fahrer bedeuten, gegenüber der Situation bei der das Fahrzeug nur auf der Spur gehalten werden muss.

Um zu überprüfen, dass eine solche Erhöhung der Beanspruchung nicht allein durch perzeptuelle Prozesse bei der Wahrnehmung, der auf den Schilder gegebenen Information beruht, wurde zusätzlich in dem vorliegenden Experiment auch eine auditive Instruktionsmodalität verwendet, bei der keine Information auf den Schildern zu sehen war.

Zusätzlich zu der Identifikation von Phasen einer erhöhten Beanspruchung, ausgelöst durch Gegebenheiten in der Fahrsituation selbst, sollte in der vorliegenden Studie auch gezeigt werden, dass mittels des ICA auch Fahrsituationen mit einer erhöhten Beanspruchung, ausgelöst durch die Bearbeitung einer zusätzlichen Aufgabe während der Fahrt, abgebildet werden können. Insbesondere wenn ein Fahrer seine Aufmerksamkeit zwischen der Fahraufgabe und einer zusätzlichen Aufgabe, wie etwa der Verwendung eines Fahrerinformationssystems, aufteilen muss, steigt die gesamthafte Beanspruchung für den Fahrer (De Waard, 1996). Auch für diese Situation gilt: Steigt die resultierende Beanspruchung für den Fahrer, so erhöht

sich das Risiko von Fahrfehlern. Daher besteht die Anforderung bei der Entwicklung neuer Fahrerinformations- und Kommunikationssysteme darin, diese Systeme in einer Art und Weise zu gestalten, dass durch diese die zusätzliche Beanspruchung für den Fahrer auf ein Minimum reduziert wird. Daher soll in dem vorliegenden Experiment gezeigt werden, dass mittels der Pupillometrie und des in dieser Arbeit diskutierten ICA diese Beanspruchungen zeitlich präzise abgebildet werden können.

Dafür wurde in dem vorliegenden Experiment eine Nebenaufgabe verwendet, welche zusätzlich zur Fahraufgabe bearbeitet werden musste.

7.2 Fragestellung

Das Ziel des vorliegenden Experiments war es die Methode des ICA auf eine simulierte Fahraufgabe zu übertragen. Sie sollte verwendet werden um Situationen mit erhöhten Beanspruchungen während des Fahrens zu identifizieren. Es wird erwartet, dass mittels des ICA Situationen mit einer erhöhten Beanspruchung, ausgelöst durch Eigenschaften der Fahraufgabe selbst, nachgewiesen werden können.

Mittels des ICA sollte es möglich sein Situationen mit einer erhöhten Beanspruchung für den Fahrer, ausgelöst durch visuelle Instruktionen die Spur zu wechseln, abzubilden.

Da die so gemessene Beanspruchung ein Resultat der Verarbeitung und Umsetzung der semantischen Instruktion darstellt und somit nicht die visuell-perzeptuellen Prozesse, ergibt sich auch die folgende Hypothese.

Mittels des ICA sollte es auch möglich sein erhöhte Beanspruchungen für den Fahrer, ausgelöst durch auditiv gegebene Instruktionen die Spur zu wechseln, abzubilden.

Zusätzlich zu den erhöhten Beanspruchungen durch Situationen in der Fahraufgabe selbst, sollte es mittels des ICA auch möglich sein Situationen mit erhöhten Beanspruchungen durch eine zusätzliche Aufgabe nachzuweisen.

Mittels des ICA sollte es möglich sein erhöhte Beanspruchungen für den Fahrer, ausgelöst durch die Bearbeitung einer zusätzlichen Aufgabe zur Fahraufgabe, abzubilden.

7.3 Methode

7.3.1 Verwendete Aufgaben

Fahraufgabe

Für die Fahraufgabe wurde die Software mit dem Namen Lane Change Task (LCT) (Mattes 2003) verwendet. Der LCT wurde als eine PC-basierte Fahrsimulationssoftware für die Messung von Fahrerablenkung entwickelt. Hier müssen Fahrer wiederholt Spurwechsel durchführen, während sie auf einer dreispurigen Strecke mit einer Gesamtlänge von etwa 3000 m fahren. Die Fahrer werden instruiert während der gesamten Aufgabenbearbeitung mit Vollgas zu fahren, wobei die Maximalgeschwindigkeit auf 60 km/h begrenzt ist und somit die Gesamtfahrtdauer für eine Strecke 180 Sekunden beträgt. In der ursprünglichen Version des LCT werden die Fahrer über die Spurwechsel visuell über entsprechende Schilder, welche links und rechts der Straße aufgestellt sind, instruiert. Jedes Schild zeigt jeweils die korrekte Spur an (s. Abbildung 16). Insgesamt gibt es 18 Schilder verteilt über die gesamte Strecke. Somit erscheint eine Instruktion die Spur zu wechseln circa alle 150 m oder alle 9 Sekunden. Um Lerneffekte zu vermeiden variiert der Abstand zwischen den Schildern um wenige Sekunden. In der vorliegenden Studie wurde der LCT zusätzlich auch in einer leicht abgewandelten Form

verwendet, bei der die Fahrer nicht visuell sondern auditiv über die Spurwechsel instruiert wurden. Die Instruktionen wurden über verbal-auditive Instruktionen über Kopfhörer, anstatt über die Schilder, gegeben, während die Schilder keinerlei Information zeigten. Während die Fahrer instruiert wurden die Spurwechsel zu schnell aber auch so präzise wie möglich durchzuführen, sollte die entsprechende Spur zwischen den Wechseln so exakt wie möglich gehalten werden.

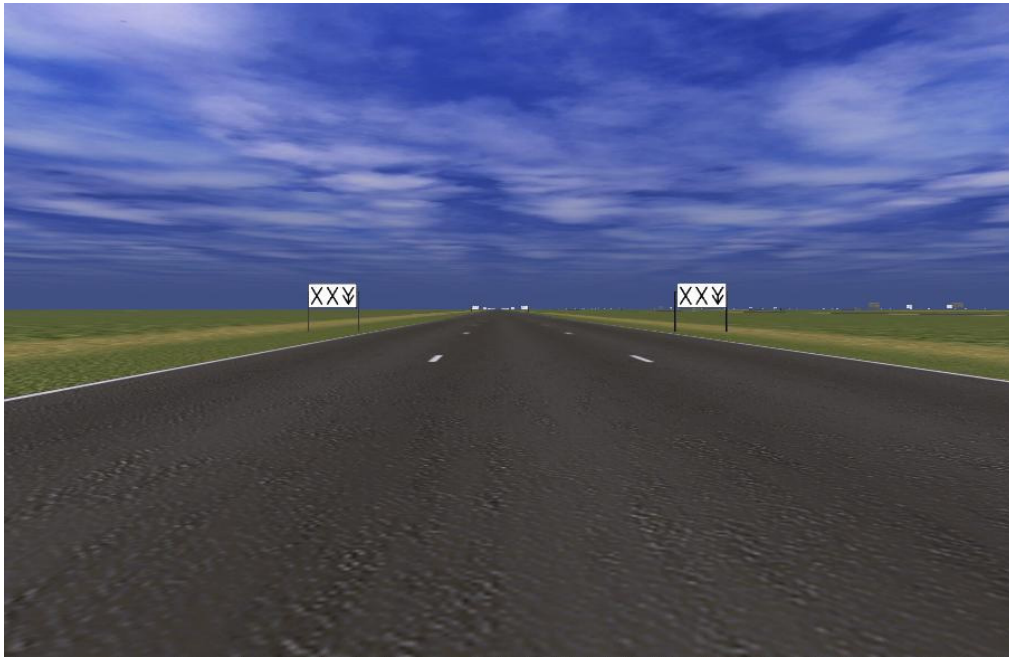


Abbildung 16: Die Abbildung zeigt die Ansicht für den Fahrer auf die dreispurige Straße des Lane Change Task (LCT) mit den Schildern, welche die Instruktion zum Spurwechsel geben.

In dem hier verwendeten LCT wurde die Fahrleistung über die Abweichung zwischen einer normativen, das heißt optimalen Fahrspur basierend auf den jeweiligen Spurwechseln, und der tatsächlich gefahrenen Spur errechnet. Die Mittelung erfolgt über die gesamte, während einer Aufgabenbearbeitung gefahrene Strecke. Frühere Studien konnten zeigen, dass der LCT sowohl für visuelle als auch für kognitive Ablenkungen während des Fahrens sensibel ist (Mattes, 2003, 2006; ISO Draft TC22/SC13/WG8).

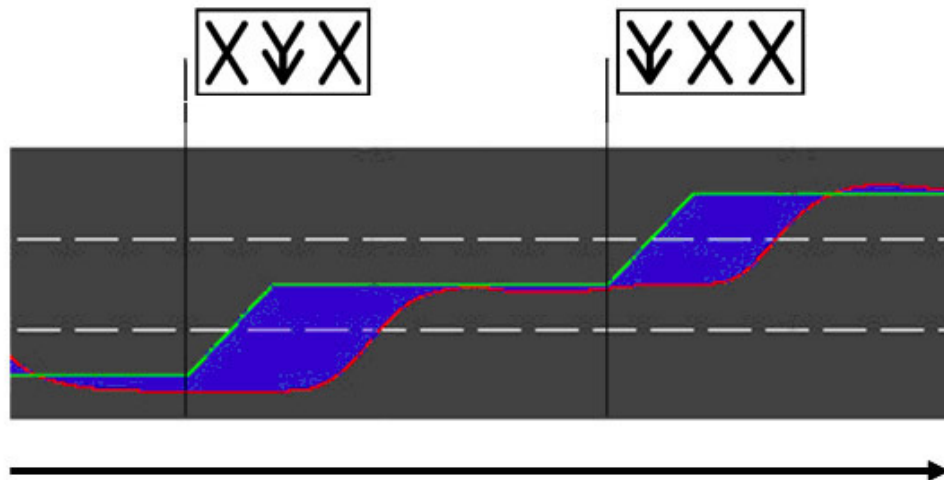


Abbildung 17: Die Abbildung zeigt exemplarisch die Erfassung der Fahrleistung über den Abstand (blau) zwischen der optimalen Fahrspur (grün) und der tatsächlich gefahrenen Spur (rot).

Insgesamt bleibt bei der Betrachtung der LCT als Fahraufgabe festzuhalten und heraus zu stellen, dass es sich hierbei um eine stark vereinfachte und nur in einem abstrakten Sinne vergleichbare Aufgabe zu einer realen, anwendungsnahen Fahrsituation handelt. Die durch die LCT Aufgabe gewonnenen Ergebnisse, insbesondere zu den konkret ermittelten Spurabweichungen auf der Straße, können nicht 1:1 auf die Realität übertragen werden. Dennoch bietet die LCT die Möglichkeit einer realistischen Fahraufgabe sehr ähnliche kognitive und motorisch beanspruchende Effekte zu erzeugen und abzubilden.

Nebenaufgabe

Um die Beanspruchungen für den Fahrer bei der Durchführung des Lane Change Task zu verändern, wird der LCT meist zusammen mit unterschiedlichen Nebenaufgaben verwendet. In der vorliegenden Studie wurde der so genannte Surrogate Reference Task (SURT) (Mattes, 2006) verwendet. Der Surrogate Reference Task besteht aus einer visuellen Suchaufgabe, bei der die Probanden die Position eines Targets (ein weißer Ring) unter 50 Distraktoren (weiße Ringe mit geringerem Durchmesser als das Target) identifizieren müssen. Die Nebenaufgabe wird dabei

auf einem kleinen Monitor neben der Versuchsperson gezeigt. Sobald die Versuchsperson das Target identifiziert hat, soll sie dieses mittels eines grauen Balkens, welcher mittels zweier Tasten auf dem Bildschirm verschoben werden kann, auf dem Display markieren. Die Aufgabenbearbeitung kann während der Durchführung beliebig unterbrochen und wieder aufgenommen werden. Sobald eine Aufgabe bearbeitet wurde, erscheint direkt darauf eine Folgende.

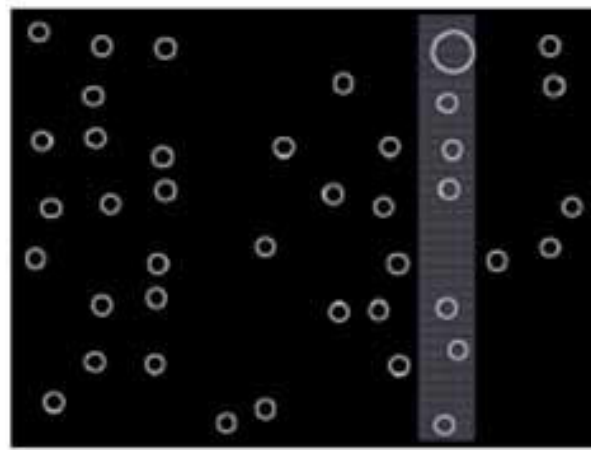


Abbildung 18: Surrogate Reference Task (SURT) mit Target und Distraktoren.

7.3.2 Versuchsaufbau

Bei dem hier beschriebenen Experiment kam ein Laboraufbau, wie in Abbildung 19 dargestellt, zum Einsatz. Die Abbildung zeigt einen Probanden mit dem Eyetracker vor dem Hauptbildschirm mit einer simulierten Fahraufgabe. Der kleine Bildschirm rechts daneben dient der Präsentation der entsprechenden Nebenaufgaben. Der Monitor für die Fahraufgabe sowie der kleinere Monitor für die Nebenaufgabe befanden sich dabei auf einer Ebene mit einer konstanten Entfernung zum Probanden.



Abbildung 19: Die Abbildung zeigt den allgemeinen experimentellen Aufbau. Zu sehen ist die Versuchsperson mit Eyetracker und den beiden Monitoren für die Haupt- sowie Nebenaufgabe.

7.3.3 Ablauf

Zunächst wurde jede Versuchsperson kurz über das Ziel der vorliegenden Studie aufgeklärt. Zu Beginn hatte jede Versuchsperson die Möglichkeit den Lane Change Task zu üben, wobei sie instruiert wurde den jeweiligen Spurwechsel so schnell, aber auch so effizient wie möglich durchzuführen. Nachdem die Probanden sicher im Umgang mit der Aufgabe waren, wurde eine Strecke so lange durchfahren, bis das Kriterium einer Abweichung < 1 erfüllt werden konnte (entsprechend ISO Draft TC22/SC13/WG8).

Nach der Kalibrierung des Eyetrackers führte jede Versuchsperson jede der drei experimentellen Bedingungen für die Dauer der Länge einer im Lane Change Task gefahrenen Strecke durch: LCT visuell, LCT visuell plus SURT, LCT auditiv. Die Abfolge der einzelnen Bedingungen wurde experimentell systematisch variiert. Bei der Durchführung der Bedingung LCT plus SURT wurden die Ver-

suchspersonen instruiert keine der beiden Aufgaben zu priorisieren, sondern beide Aufgaben so gut und so effizient wie möglich zu lösen.

Direkt im Anschluss an jede experimentelle Bedingung gaben die Probanden eine subjektive Einschätzung der auftretenden Beanspruchung bei der Bearbeitung der soeben durchgeführten Aufgabe mittels einer deutschen Version des NASA TLX ohne Gewichtung.

7.3.4 Stichprobe

An dem vorliegenden Experiment nahmen insgesamt 20 Probanden teil (10 männlich und 10 weiblich). Die Teilnehmer waren im Durchschnitt 26,6 Jahre, und zwischen 23 bis 33 Jahre alt.

7.4 Ergebnisse

Fahrleistung

Für die Erfassung der Fahrleistung wurde die mittlere Abweichung auf der Strecke zwischen der Optimallinie und der tatsächlich gefahrenen Spur ermittelt. Ein solches Maß ist ein geeigneter Indikator um die Fahrerablenkung zu erfassen, da es empfindlich ist gegenüber Unterschieden in der Schnelligkeit der Lenkreaktion (Spurwechsel vollzieht sich schneller oder langsamer) sowie der Güte der lateralen Kontrolle, das heißt wie gut der Fahrer die Spur halten kann (ISO Draft TC22/SC13/WG8). Die Ergebnisse zur Fahrleistung sind in Abbildung 20 dargestellt.

Eine ANOVA mit Messwiederholung zeigte einen signifikanten Effekt für den Faktor der experimentellen Bedingung, $F(2, 28)=55.9$, $p<.001$. Ein geplanter Kontrast zeigte keinen signifikanten Unterschied zwischen den beiden Bedingungen LCT visuell und LCT auditiv ohne zusätzliche Nebenaufgabe, $F(1, 14)=0.03$, $p=.86$. Ein signifikanter Anstieg in der mittleren Abweichung zeigte sich hingegen zwischen der Bedingung LCT plus SURT und den beiden Bedingungen LCT visuell und LCT auditiv, $F(1, 14)=70.85$, $p<.001$.

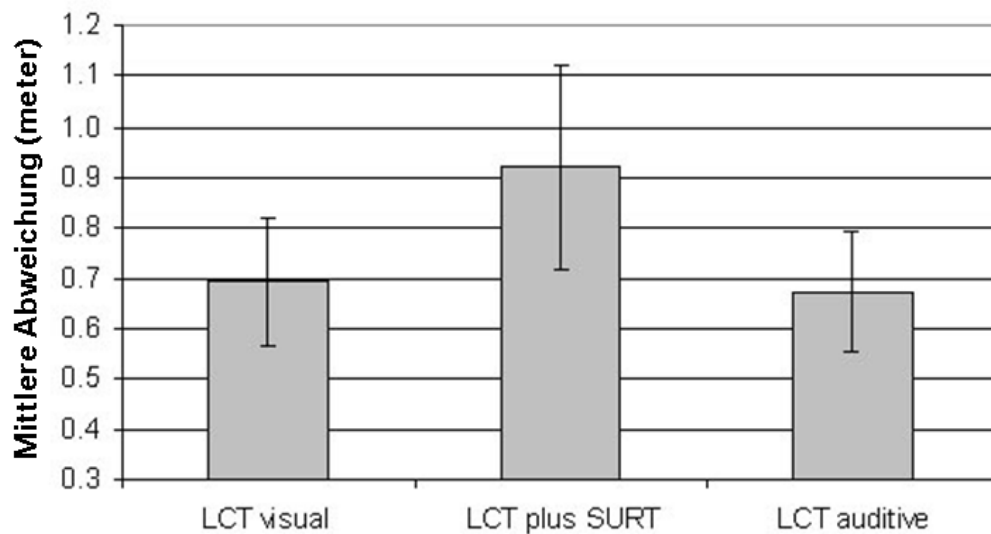


Abbildung 20: Die Abbildung zeigt die mittleren Abweichungen (in Metern) in der Fahraufgabe für die drei experimentellen Bedingungen [vertikale Balken zeigen +/- 1 Standardabweichung].

Subjektive Angaben

Die subjektiven Ratings bezüglich der wahrgenommenen Beanspruchung wie sie durch den NASA TLX Fragebogen am Ende jeder experimentellen Bedingung erhoben wurden, zeigten entsprechend der Fahrleistungsdaten den gleichen Befund. Eine ANOVA mit Messwiederholung zeigte einen signifikanten Unterschied zwischen den drei experimentellen Bedingungen über jede der sechs Dimensionen des NASA TLX, $F(2, 28) = 17.59$, $p < .001$. Ein geplanter Vergleich zeigte hier einen signifikanten Unterschied zwischen den Bedingungen ohne zusätzlicher Nebenaufgabe und der Bedingung LCT plus SURT, $F(1, 14) = 30.96$, $p < .001$. Zusätzlich zeigte ein weiterer geplanter Vergleich keinen Unterschied zwischen den beiden Bedingungen LCT visuell und LCT auditiv über die sechs Dimensionen des NASA TLX, $F(1, 14) = 0.89$, $p = .36$.

Insgesamt zeigte sich, dass die Bedingung mit zusätzlicher Nebenaufgabe als deutlich beanspruchender eingeschätzt wurde als die beiden Bedingungen ohne zusätzliche Aufgabe.

Index of Cognitive Activity

Der ermittelte ICA wurde zunächst für jede Versuchsperson jeweils für beide Augen aufsummiert und über die gesamte Dauer einer jeden experimentellen Bedingung erfasst (s. Abbildung 21). Eine ANOVA mit Messwiederholung zeigte einen signifikanten Unterschied zwischen den Bedingungen in den ICA Werten, $F(2, 28) = 217.3$, $p < .001$. Ein geplanter Vergleich zeigte einen signifikanten Anstieg in den gemessenen ICA Werten zwischen den beiden Bedingungen ohne zusätzliche Nebenaufgabe und der Bedingung LCT plus SURT, $F(1, 14) = 2259.17$, $p < .001$. Zwischen den beiden Bedingungen LCT visuell und LCT auditiv zeigte ein weiterer geplanter Vergleich keinen Unterschied zwischen den beiden Bedingungen in Bezug auf die gemessenen Beanspruchungen, $F(1, 14) = 0.25$, $p = .62$.

Die gemessenen Beanspruchungsdaten unterstützen somit die bereits für die Fahrleistung sowie die subjektiven Daten beschriebenen Befunde.

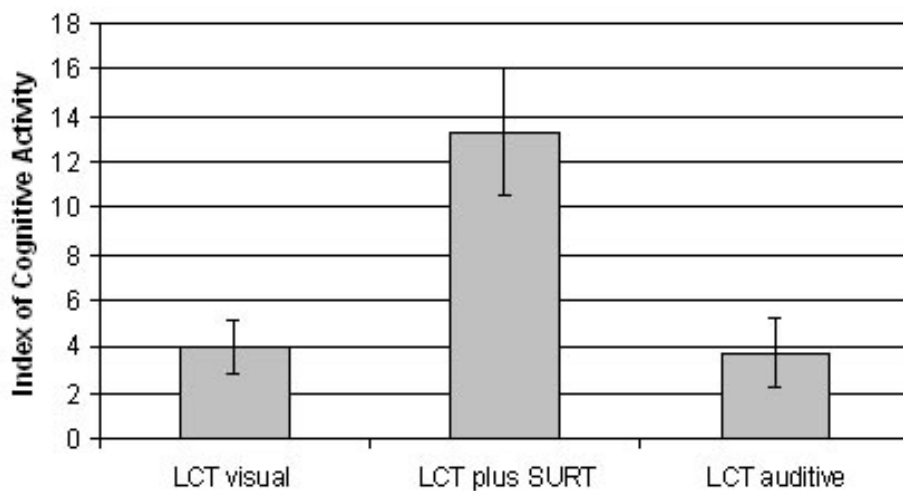


Abbildung 21: Die Abbildung zeigt den mittleren Index of Cognitive Activity (ICA) für die Bedingungen LCT visuell, LCT auditiv sowie LCT plus SURT [vertikale Balken zeigen +/- 1 Standardabweichung].

Da der Vorteil des ICA jedoch in der Tatsache liegt, dass dieser eine kontinuierliche Messung der auftretenden Beanspruchungen mit einer hohen zeitlichen Auflö-

sung ermöglicht, wurde in der vorliegenden Studie eine genauere Analyse der Beanspruchungsverläufe während der Fahrt durchgeführt. Dafür wurden die individuellen Verläufe der ICA Daten aller 20 Versuchspersonen auf den Startpunkt der Strecke zeitlich synchronisiert und die ICA Werte für jeden Zeitpunkt gemittelt. Abbildung 22 und Abbildung 23 zeigen die mittleren ICA Werte für die Bedingungen LCT visuell und LCT auditiv. Die grauen Balken zeigen das Zeitintervall in denen die Versuchspersonen jeweils die Instruktion (visuell oder auditiv) verarbeiten und entsprechend einen Spurwechsel durchführen mussten. Der zeitliche Verlauf des ICA Signals zeigt für jede der beiden Bedingungen ein klares zeitliches Muster, wonach die Amplituden im ICA Signal dem Auftreten der einzelnen Schilder folgen. Der ICA steigt jeweils zu einem lokalen Maximum bei jedem der 18 Spurwechsel an. Das lokale Maximum in den ICA Werten (ca. 10) ist hier etwa zweimal so hoch wie in den Situationen ohne Fahrmanöver (ca. 4). Unter Berücksichtigung der Standardabweichungen auf der Ebene der einzelnen experimentellen Bedingungen sowie der absoluten Regelmäßigkeit des Effektes, kann dieser Effekt als reliabel betrachtet werden.

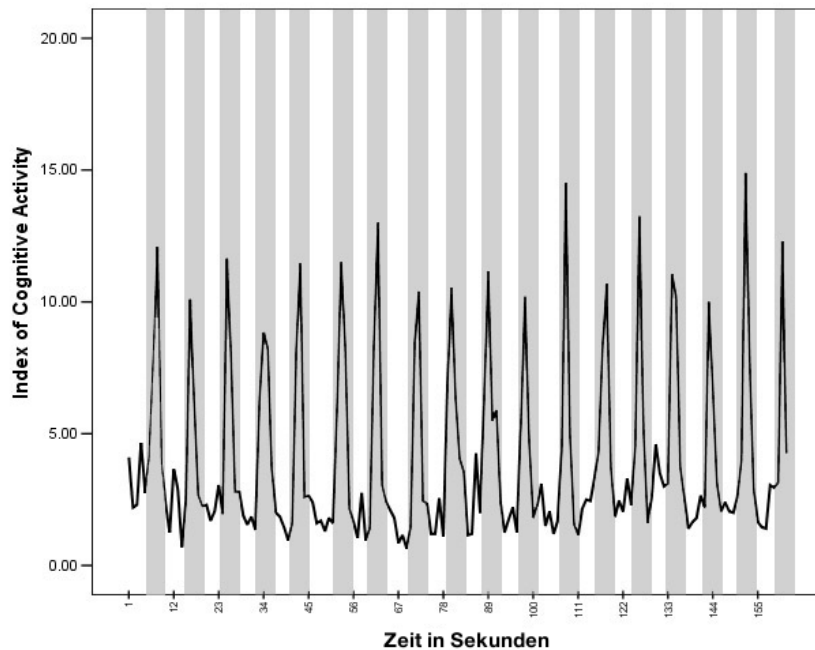


Abbildung 22: Die Abbildung zeigt die gemittelten ICA Verläufe in der Bedingung LCT visuell über die Dauer einer gefahrenen Strecke. Die grauen Balken zeigen die 18 Spurwechsel auf der Strecke an.

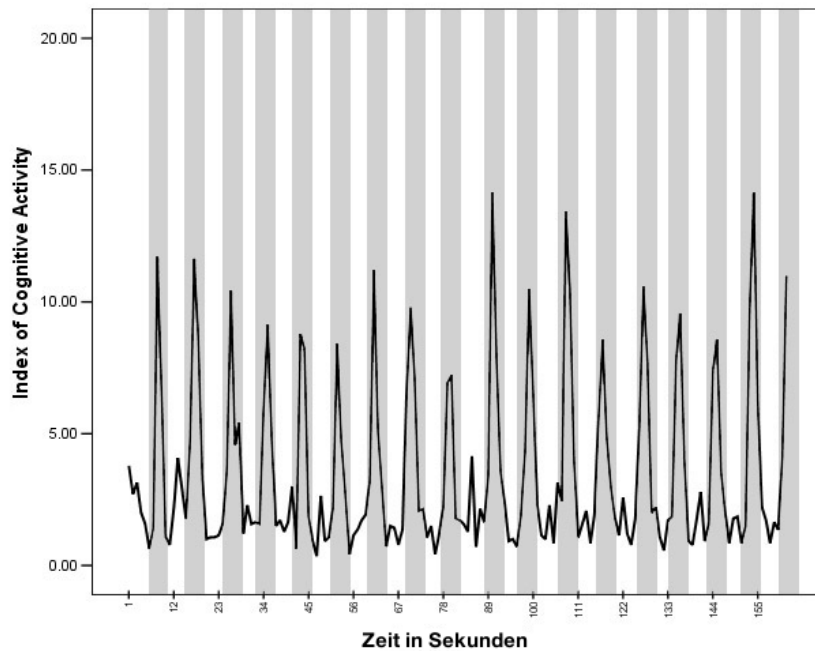


Abbildung 23: Die Abbildung zeigt die gemittelten ICA Verläufe in der Bedingung LCT auditiv über die Dauer einer Strecke. Die grauen Balken zeigen die 18 Spurwechsel auf der Strecke.

Zusätzlich zu den Bedingungen ohne eine zusätzliche Nebenaufgabe wurde auf die gleiche Weise der Verlauf der ICA Werte unter der Bedingung LCT plus SURT analysiert. Hier wird zunächst deutlich, dass sich die ICA Daten über die gesamte Zeit der Aufgabenbearbeitung auf einem deutlich höheren Niveau befinden, als unter den Bedingungen LCT visuell und LCT auditiv.

Darüber hinaus zeigte sich in den Daten ein zunächst unerwarteter Effekt. In der Bedingung LCT plus SURT zeigte sich bei den insgesamt erhöhten ICA Werten bei jedem der 18 Schilder ein deutlicher Abfall in den gemessenen Werten. Eine ANOVA zeigte hier einen signifikanten Unterschied zwischen den Situationen in denen ein Spurwechselmanöver durchgeführt werden musste (Mittelwert = 11,73) im Vergleich zu den Situationen in denen lediglich die Spur gehalten werden musste (Mittelwert = 14,47), $F(1, 155) = 92.65$, $p < .001$.

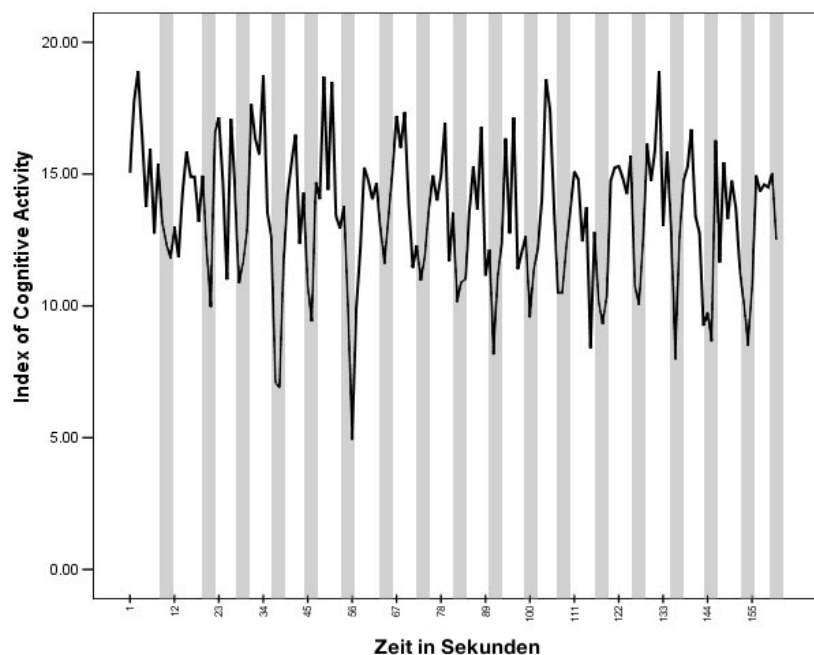


Abbildung 24: Die Abbildung zeigt die gemittelten ICA Verläufe in der Bedingung LCT plus SURT. Die grauen Balken zeigen die 18 Spurwechsel auf der Strecke.

Leistung Nebenaufgabe

In Bezug auf die Leistung in der Nebenaufgabe wurden in dem vorliegenden Experiment die Anzahl der richtig gelösten Aufgaben betrachtet. Hier ergab sich ein prozentualer Anteil richtig gelöster Aufgaben mit einem Mittelwert von 98% und einer Standardabweichung von 1.2.

Zusätzlich ergab sich eine signifikante Korrelation von $r=.55$ ($p<.05$) zwischen der Anzahl richtig gelöster Aufgaben und den mittleren ICA Werten gemittelt über die gesamte Strecke. Dies bedeutet, je mehr Aufgaben während der Fahrt bearbeitet wurden, desto höher lag auch die durchschnittlich gemessene Beanspruchung.

7.5 Diskussion

Der Lane Change Task wurde entwickelt, um die mentalen Beanspruchungen in einer Fahrsituation experimentell zu erzeugen sowie über ein entsprechendes Leistungskriterium abbilden zu können. Entsprechend konnte in dem vorliegenden Experiment gezeigt werden, dass sich die mittlere Abweichung zur Optimallinie bei der Bearbeitung einer zusätzlichen Aufgabe zusätzlich zur Fahraufgabe deutlich erhöht. Die mit der Bearbeitung einer zusätzlichen Aufgabe einhergehende erhöhte Beanspruchung für den Fahrer konnte auch über die subjektiven Maße sowie den gemessenen ICA Wert bestätigt werden. Insgesamt zeigten sich keine Unterschiede zwischen den beiden Instruktionsmodalitäten visuell und auditiv. Somit ist anzunehmen, dass hier die perzeptuelle Verarbeitung der Schilder nur eine minimale zusätzliche Beanspruchung bedeutet und somit die Fahraufgabe nicht durch modalitätsspezifische Interferenzen beeinflusst wird.

Neben der Bestätigung der mittels der vorhandenen Performanz Daten gefundenen Ergebnisse in Bezug auf die erhöhte Beanspruchung durch eine zusätzliche Aufgabe, zeigt das vorliegende Experiment jedoch auch den großen Vorteil der Me-

thode des ICA. Die Ergebnisse zeigen, dass mittels des ICA neben der absoluten Höhe der Beanspruchung in einer bestimmten Bedingung, es auch möglich ist den ganz speziellen zeitlichen Verlauf der Beanspruchung bei der Bearbeitung einer Aufgabe abzubilden. So zeigt sich hier bei der Bearbeitung der Fahraufgabe ein signifikanter Anstieg in den Beanspruchungswerten jedes Mal, wenn der Fahrer eine Instruktion wahrnehmen, verstehen und motorisch umsetzen muss, indem er die Spuren so schnell aber auch so präzise wie möglich wechselt. Somit wird deutlich, dass es mittels des ICA möglich ist eine sehr viel genauere zeitliche Auflösung bei der Messung von Beanspruchungen zu erreichen als mit herkömmlichen Performanz Maßen, wie etwa der Spurabweichung oder subjektiven Daten. Dadurch dass der hier beobachtete Effekt – wie bereits zuvor diskutiert - unabhängig von der verwendeten Instruktionsmodalität ist, ist davon auszugehen, dass die hier berichteten zeitlich sehr begrenzten Anstiege in der gemessenen Beanspruchung auf eine zusätzliche Beanspruchung in höheren kognitiven Prozessen zurückzuführen sind und nicht etwa auf rein perzeptuelle Prozesse.

Der unerwartete Abfall in den beobachteten ICA Daten bei zusätzlicher Nebenaufgabe zeigt das Potential des ICA für das Gebiet der Aufgabenanalyse sowie der Erfassung der aktuellen Beanspruchungen. Wenn zusätzlich zu der Fahraufgabe eine Nebenaufgabe bearbeitet wird, reduziert sich der ICA deutlich zu jedem Spurwechsel. Solch ein Abfall in den Beanspruchungswerten, besonders um den Zeitpunkt eines kritischen Manövers in der Fahraufgabe, ist zunächst konterintuitiv.

Allerdings wird bei näherer Betrachtung des eigentlichen Verhaltens des Fahrers rund um ein solches kritisches Event deutlich, warum sich die Beanspruchung in diesem Fall so verhalten muss. Der Fahrer hört auf an der Nebenaufgabe (SURT) zu arbeiten, wenn er die Instruktion bekommt die Spur zu wechseln und konzentriert sich stattdessen auf die Fahraufgabe. In anderen Worten wechselt der Fahrer an dieser Stelle von einer klaren dual-task Situation (also eine Situation bei der er wirklich zwei Aufgaben simultan bearbeiten) zu einer single-task Situation (nur

Fahren). Dieses ermöglicht insbesondere die Struktur der hier gewählten Aufgabe des SURT. Die Aufgabenbearbeitung kann jederzeit unterbrochen und ohne Nachteile wieder begonnen werden, wenn es denn die Situation zulässt. Was somit in diesem Fall passiert, ist folgendes: Nähert sich der Fahrer einem kritischen Event in der Fahraufgabe wird er diese als primäre / sicherheitskritische Aufgabe identifizieren und die Bearbeitung der Nebenaufgabe unterbrechen. Erst wenn er die kritische Situation in der Fahraufgabe (Spurwechsel) gemeistert hat, wird er sich wieder der Nebenaufgabe zuwenden.

Somit zeigt das vorliegende Ergebnis deutlich, dass es wirklich einer detaillierten zeitlich hochauflösenden Analyse der Situation bedarf, um zu verstehen was in einer solchen Situation tatsächlich passiert. Eine Analyse auf Aufgabenebene, bei der lediglich die Aufgabe als ganzes betrachtet werden kann, ist nicht ausreichend. Man muss davon ausgehen, dass der Fahrer eine solche strategische, serielle Planung der beiden vorliegenden Aufgaben vornimmt, sobald es die diskrete Charakteristik der Aufgabe (oder der Hauptaufgabe) erlaubt. Angewandt auf die Fahrsituation bedeutet dies, dass der Fahrer hier aktiv und strategisch seine Kapazitäten auf die einzelnen Beanspruchungen verteilt und zu deren Lösung einsetzt. Der Fahrer handelt somit hier als eine Art aktiver Manager seiner ihm zu Verfügung stehenden Ressourcen.

Folgt man der oben gegebenen Interpretation, so ist anzunehmen, dass es mittels des ICA möglich ist strategische Wechsel bei der Verteilung der Aufmerksamkeit, wie etwa bei einem Wechsel zwischen einer dual-task und einer single-task Situation, abzubilden. Der ICA ermöglicht somit nicht nur die Erfassung der momentanen Beanspruchung insgesamt, sondern vielmehr einen Einblick in die internalen Strategien eines Fahrers und seiner aktiven Strategien der Beanspruchungsbewältigung. Dass der ICA genutzt werden kann um strategische Veränderungen in der Aufmerksamkeit nachzuweisen, konnte bereits in früheren Studien nachgewiesen werden (Marshall, Pleydell-Pearce, Dickson, 2002, Marshall, 2002, 2007). Dies

ist jedoch das erste Mal, dass dies in einem solch komplexen Fahrkontext gezeigt werden konnte.

Zusammenfassend zeigt das vorliegende Experiment, dass es mittels des ICA möglich ist auftretende Beanspruchungen während des Fahrens sowie strategische Veränderungen bei der Aufmerksamkeitszuwendung zu erfassen.

8 Experiment 3: Untersuchung zur weiteren Validierung der Methode im Fahrkontext sowie der Nachweis strategischer Beanspruchungseffekte in einer Fahraufgabe

In dem nun vorliegenden dritten Experiment dieser Arbeit soll eine weitere genauere Validierung der Methode des Index of Cognitive Activity (ICA) in Bezug auf den Nachweis von Beanspruchungseffekten in einer simulierten Fahraufgabe erreicht werden. Ausgehend von den in Experiment 2 gefundenen Ergebnissen soll zusätzlich in dem vorliegenden Experiment eine genauere Analyse des Potentials der Methode in Bezug auf den Nachweis von impliziten strategischen Aufmerksamkeitseffekten erreicht werden.

Verwendet wurden in der vorliegenden Untersuchung verschiedene Nebenaufgaben mit unterschiedlichen Schwierigkeitsstufen. Zusätzlich wurde die Art der Instruktion in der Fahrsituation experimentell variiert. Die vorliegende Untersuchung erlaubt einen Abgleich der über die Methode des ICA gefundenen Ergebnisse mit denen klassischer Performanz Maße wie etwa Fahrleistungsdaten sowie Leistungen in der Nebenaufgabe.

8.1 Hintergrund

In Experiment 2 zeigte sich, dass der ICA bei der Verwendung des Lane Change Task (LCT) als simulierte Fahraufgabe ausgesprochen sensibel auf eine erhöhte Beanspruchung rund um jeden der 18 Spurwechsel reagiert. In dem vorliegenden Experiment sollte nun geklärt werden, welche Faktoren genau bei der Instruktion und Durchführung der einzelnen Spurwechsel eine erhöhte Beanspruchung für den Fahrer erzeugen. Dafür wurde in der vorliegenden Untersuchung die Situation rund um die jeweiligen Spurwechsel mit einer deutlich erhöhten zeitlichen Genauigkeit untersucht. Mit der Replikation der gefundenen Ergebnisse aus Experiment 2 sollte mit dem vorliegenden Experiment einen weiteren Nachweis der Validität der Methode erbracht werden.

Zunächst soll eine genauere Betrachtung der Fahraufgabe im LCT und seine für den Fahrer beanspruchenden Komponenten erfolgen. Bei der Aufgabe des LCT muss der Fahrer auf ein Schild oder eine auditive Instruktion mit einem Spurwechsel reagieren. Dabei können die folgenden für den Fahrer mental beanspruchenden Faktoren identifiziert werden:

- 1) Zunächst muss der Fahrer die Instruktion perzeptuell verarbeiten. Das heißt das Schild visuell wahrnehmen bzw. die auditive Instruktion hören.
- 2) Der Fahrer muss daraufhin den semantischen Gehalt der Instruktion verstehen und mit einer entsprechenden Handlungsanweisung verknüpfen.
- 3) Der Fahrer muss die Instruktion in eine willentliche Handlung überführen, indem er eine Handlungsplanung durchführt und die Handlung initiiert.

Während in dem vorangegangenen Experiment gezeigt werden konnte, dass der ICA insgesamt bei den jeweiligen Spurwechseln einen klaren Anstieg in den Werten zeigte, sollten in dem vorliegenden Experiment die bereits weiter oben beschriebenen Schritte separiert und auf ihren Beanspruchungsgehalt für den Fahrer untersucht werden.

Zusätzlich sollte in dem vorliegenden Experiment untersucht werden, in wie weit eine visuelle Veränderung an sich, hervorgerufen beispielsweise durch die Veränderung des visuellen Flusses bei dem Wechsel der Spur, einen Einfluss auf die gemessenen ICA Werte hat.

Dabei sollte die Analyse der jeweiligen Fahrsituation unter ihrem zeitlichen Verlauf noch detaillierter erfolgen als in Experiment 2. Dafür wurden die Abschnitte der 18 unter einer Bedingung durchfahrenen Schilder jeweils gemittelt, um so eine erhöhte Validität der Ergebnisse zu erzielen.

- 1) Um den Einfluss der visuellen Wahrnehmung der Instruktionen in Form von Schildern zu identifizieren, wurde der LCT in der vorliegenden Studie einmal unter einer normalen visuellen Instruktionsbedingung durchgeführt (Schilder zeigen visuell die Richtung des Spurwechsels an) und einmal - wie bereits in Experiment 2 geschehen - in einer auditiven Instruktionsmodalität (Schilder wurden gelöscht und die Instruktion die Spur zu wechseln auditiv dargeboten).
- 2) Um eine Trennung zwischen dem Einfluss einer semantischen Verarbeitung der Instruktionen und die intentionale Umsetzung und Initiierung einer Handlung zu erreichen, wurde in dem vorliegenden Experiment die Bedingung auditiv-imperativ eingeführt. Hier bekamen die Probanden 4 Sekunden vor dem eigentlichen Spurwechsel die semantische Information auf welche Spur gewechselt werden sollte in Form eines auditiven Kommandos („links“, „rechts“, „mitte“). Das Besondere an dieser Bedingung war, dass die Probanden instruiert waren zu diesem Zeitpunkt lediglich die Instruktion zu verstehen aber noch nicht in eine Fahrhandlung umzusetzen. Erst 2 Sekunden vor dem Schild (exakt der Zeitpunkt an dem normalerweise die visuelle Instruktion auf den Schildern sichtbar wird) bekamen die Fahrer mittels eines imperativen akustischen Signals die Instruktion das soeben gehörte Kommando in eine entsprechende Fahrhandlung umzusetzen. Somit konnte zeitlich eine semantische Verarbeitung von der tatsächlichen willentlichen Umsetzung einer Handlung unterschieden werden.
- 3) Um den Einfluss der Veränderung des visuellen Eindrucks bzw. des visuellen Flusses bei der Durchführung eines Spurwechsels auf den ICA zu ermitteln, wurde in der vorliegenden Studie eine zusätzliche Bedingung eingeführt bei der die Fahrer völlig frei eine Strecke des LCT abgefahren sind und dabei nach Belieben die Spuren wechseln konnten ohne dabei auf

Schilder oder andere auditive Instruktionen in irgendeiner Art reagieren zu müssen.

Neben der Validierung der Ergebnisse des ICA als Methode zur Messung mentaler Beanspruchung in einer simulierten Fahraufgabe, sollte in dem vorliegenden Experiment auch das Thema des Nachweises von impliziten Strategien der Aufmerksamkeitsverteilung mittels des ICA bei realen Fahraufgaben, wie bereits in Experiment 2 beschrieben, genauer beleuchtet werden.

In Experiment 2 konnte bereits gezeigt werden, dass es bei der Bearbeitung einer zusätzlichen Aufgabe neben der Fahraufgabe selbst rund um die kritischen Situationen in der Fahraufgabe anstatt zu einer Erhöhung der gemessenen Beanspruchung zu einer kurzfristigen Verringerung der gemessenen Beanspruchung gekommen ist. Dieser Effekt kann damit erklärt werden, dass in diesen kritischen Situationen der Fahrer von einer wirklichen dual-task Situation, also einer Situation bei der wirklich zwei Aufgaben parallel bearbeitet werden, zu einer single-task Situation wechselt, bei der nur noch die Fahraufgabe bearbeitet wird. Das Ergebnis ist ein entsprechender kompensativer, strategischer Effekt bei dem der Fahrer proaktiv seine momentane Beanspruchung in Hinblick auf eine kritische Situation in der Fahraufgabe anpasst. So gelingt es ihm, in der kritischen Situation genügend Kapazitäten freustellen zu können, um die Situation sicher und effizient zu lösen.

In dem vorliegenden Experiment sollte nun dieser Effekt zunächst noch einmal mit einer deutlich erhöhten Genauigkeit untersucht werden, um so ein genaueres Verständnis dafür zu generieren, welche Strategien der Aufgabenbearbeitung von den Fahrern eingesetzt werden und in wie weit es möglich ist mittels des ICA in diese Prozesse Einsicht zu erlangen.

In dem vorliegenden Experiment wurden daher zusätzlich zu den Fahrleistungsdaten auch die Aktivitätsdaten aus den Nebenaufgaben erhoben und über ihren zeitlichen Verlauf abgebildet. Für eine genauere Analyse der auftretenden Effekte

wurden hier ebenfalls die Ereignisse gemittelt über die jeweils 18 Spurwechsel in einer Bedingung dargestellt.

Folgt man der Diskussion einer impliziten Strategie der Aufmerksamkeitsverteilung bei der der Fahrer seine Kapazitäten situationsangepasst aktiv auf die einzelnen Aufgaben verteilt, stellt sich auch die Frage in wie weit ein solches Verhalten durch bestimmte situative Faktoren beeinflusst werden kann. So ist zu vermuten, dass bereits die Charakteristika einer Nebenaufgabe ein solches Verhalten beeinflussen können.

Um dieses zu überprüfen wurden in dem vorliegenden Experiment zwei unterschiedliche artifizielle Nebenaufgaben in jeweils drei unterschiedlichen Schwierigkeitsstufen verwendet.

Der SURT (Surrogate Reference Task) (Mattes, 2006) als eine diskrete Aufgabe. Hier kann der Fahrer die Aufgabe zu jeder Zeit unterbrechen und fortsetzen ohne Einbußen in der Leistung der Nebenaufgabe befürchten zu müssen.

Der CTT (Critical Tracking Task) als eine kontinuierliche Aufgabe. Hier ist eine kontinuierliche Zuwendung des Fahrers zur Aufgabe notwendig. Wird die Bearbeitung der Aufgabe über einen bestimmten Zeitraum unterbrochen, so hat dies Leistungseinbußen in der Aufgabe zur Folge.

Folgt man der Diskussion über eine aktive kompensatorische Reaktion bei der Verteilung der Aufmerksamkeitskapazitäten, so ist zu erwarten, dass insbesondere für den diskreten Aufgabentyp eine solche kompensative Strategie unterstützend wirkt und mittels des ICA abgebildet werden kann. Wohingegen unter einer kontinuierlichen Aufgabe die Nutzung solcher Strategien erschwert wird.

Um zu überprüfen in wie weit die Verwendung einer solchen Strategie von der Schwierigkeit der verwendeten Aufgabe abhängt, wurden hier zusätzlich die bei-

den Nebenaufgaben in jeweils drei unterschiedlichen Schwierigkeitsstufen (einfach, mittel, schwer) dargeboten.

Zusätzlich zu der Verwendung der bereits weiter oben beschriebenen künstlichen Nebenaufgaben sollte in dem vorliegenden Experiment auch direkt ein Vergleich der auftretenden Beanspruchungen sowie Bearbeitungsstrategien mit einer Aufgabe in einem realen, anwendungsnahen System erfolgen. Für einen solchen Vergleich wurde in dem vorliegenden Experiment von den Probanden zusätzlich eine Navigationszieleingabe in einem BMW iDrive Fahrerinformationssystem aus dem Vorgängersystem der aktuellen 5er Modellreihe durchgeführt. Die Bedienung erfolgte hier über einen externen iDrive Controller.

Da die Struktur des BMW iDrive Systems explizit darauf ausgelegt ist, dass die Bearbeitung einer Aufgabe zu jedem Zeitpunkt nach Belieben unterbrochen und wieder aufgenommen werden kann, ohne dass hier für den Benutzer Nachteile bei der Benutzung entstehen, sind hier Bearbeitungsstrategien ähnlich denen bei der Bearbeitung der SURT Aufgabe zu erwarten. Es ist zu erwarten, dass hier die Struktur eines solchen Systems die Verwendung kompensativer Strategien unterstützen kann und diese mittels der zeitlichen Verläufe über den ICA abzubilden sind.

8.2 Fragestellung

Das vorliegende Experiment soll zunächst eine klare Replikation der in Experiment 2 gefundenen Ergebnisse zum Nachweis der Messung auftretender Beanspruchungen in einer simulierten Fahraufgabe mittels des ICA bieten. Dabei soll mit dem vorliegenden Experiment auch ein genaueres Verständnis darüber erreicht werden, welche beanspruchenden Komponenten der Situation mittels des ICA erfasst werden können.

Mittels des ICA sollte es möglich sein Beanspruchungen durch Spurwechsel in einer Fahrsituation unabhängig von der jeweiligen Instruktionsmodalität (visuell und auditiv) nachzuweisen.

In Bezug auf den Nachweis einzelnen Komponenten der Beanspruchung mittels des ICA sollte hier eine Trennung der auftretenden Beanspruchung, ausgelöst durch eine semantische Verarbeitung und der Umsetzung einer Instruktion in eine Handlung, möglich sein.

Mittels des ICA sollte es möglich sein auftretende Beanspruchungen, ausgelöst durch semantische Verarbeitung sowie die willentliche, motorische Umsetzung in eine Handlung, abzubilden und getrennt messbar machen zu können.

Da davon ausgegangen wird, dass die Unterschiede in den ICA Daten auf Unterschiede in der ganzheitlichen Beanspruchung des Fahrers zurückzuführen sind und nicht etwa durch eine Veränderung des visuellen Eindrucks oder etwa einer motorischen Reaktion an sich, ergibt sich eine weitere Forderung.

Die gefundenen Ergebnisse einer Erhöhung der ICA Werte ausgelöst durch Spurwechsel sind nicht durch die Veränderung des visuellen Eindrucks oder einer motorischen Reaktion allein begründet.

In Bezug auf die Diskussion kompensativer Strategien bei der Bearbeitung einer zusätzlichen Aufgabe während des Fahrens sowie der Überprüfung der in Experiment 2 geführten Argumentation in Bezug auf die Bearbeitungsstrategie bei künstlichen Nebenaufgaben, ergibt sich eine weitere Hypothese für das vorliegende Experiment.

Mittels geeigneter Aktivitätsmaße für die Nebenaufgabe sowie der gemessenen Beanspruchung über den ICA sollte es möglich sein in Abhängigkeit der Charakteristika sowie der Schwierigkeit einer künstlichen Nebenaufgabe zusätzlich zur Fahraufgabe unterschiedliche kompensative Strategien der Aufgabenbearbeitung in einer Fahrsituation nachzuweisen.

Aufgrund der theoretischen Ableitung auf Basis der postulierten Bearbeitungsstrategien bei künstlichen Nebenaufgaben sollten die hier beschriebenen Effekte in Bezug auf die Beanspruchungsmaße auch auf Situationen mit einem realen Fahrerinformationssystem übertragbar sein.

Mittels der gemessenen Beanspruchung über den ICA sollte es möglich sein auch für die Verwendung eines realen Fahrerinformationssystems auftretende Beanspruchungen sowie eine kompensative Bearbeitungsstrategie zu identifizieren.

8.3 Methode

In Experiment 3 wurde derselbe experimentelle Aufbau und Eyetracker sowie dieselbe Fahraufgabe (LCT), wie bereits unter Experiment 2 beschrieben, verwendet.

Zur Überprüfung der oben angeführten Hypothesen wurden jedoch unter der Fahraufgabe einige zusätzliche Bedingungen eingeführt.

8.3.1 Verwendete Aufgaben

Fahraufgabe

Für die Fahraufgabe wurde erneut die Software Lane Change Task (LCT) (Mattes 2003), wie bereits unter Experiment 2 beschrieben, verwendet. Zur Überprüfung der für das Experiment 3 formulierten Hypothesen wurde der LCT unter folgenden Instruktionsbedingungen durchgeführt: 1) LCT visuell (Visuell Baseline). Hier wurden die Probanden, wie bereits unter Experiment 2 beschrieben, visuell über die entsprechenden Richtungen der Spurwechsel mittels Schilder, welche rechts und links der Straße aufgestellt waren, instruiert. 2) LCT auditiv (auditiv normal). Hier wurde keine Information auf den Schildern präsentiert. Stattdessen wurden die Instruktionen auditiv über Kopfhörer an der entsprechenden Position auf der Strecke gegeben, an der auch das Schild sichtbar geworden wäre. 3) LCT auditiv imperativ. Hier wurden die Fahrer ebenfalls auditiv über die Richtung des Spurwechsels informiert (Schilder zeigten keine Information). Allerdings wurde hier die semantische Information über die Richtung des Spurwechsels bereits 3 Sekunden vor dem eigentlichen Zeitpunkt gegeben, an dem tatsächlich die Spur gewechselt werden sollte. Das heißt 5 Sekunden vor dem Passieren des Schildes. Dieses wurde dann (zum gleichen Zeitpunkt, an dem auch die visuelle Information auf den Schildern sichtbar geworden wäre) durch ein kurzes akustisches Signal angezeigt. Die Fahrer konnten hier somit die semantische Information bereits vor der Ausführung der eigentlichen Handlung verarbeiten und die Handlung entspre-

chend vorbereiten. 4) LCT freies Fahren. Hier wurde die Versuchsperson zu Beginn der Fahrt instruiert über die Dauer einer gesamten Strecke im LCT nach freiem Ermessen die Spuren zu wechseln. Die Schilder zeigten hier keinerlei Informationen. Der Zeitpunkt des Spurwechsels sowie die Spur konnten hier frei von der Versuchsperson gewählt werden.

Nebenaufgaben

Sollte zusätzlich zu der Fahraufgabe eine Nebenaufgabe bearbeitet werden, wurde in dem vorliegenden Experiment die visuelle Instruktionsmodalität des LCT verwendet.

Surrogate Reference Task (SURT)

In der vorliegenden Studie wurde, wie bereits in Experiment 2, der so genannte Surrogate Reference Task (SURT) (Mattes, 2006) verwendet. Der Surrogate Reference Task besteht aus einer visuellen Suchaufgabe, bei der die Probanden die Position eines Targets (ein weißer Ring) unter 50 Distraktoren (weiße Ringe mit geringerem Durchmesser als das Target) identifizieren mussten. In dem vorliegenden Experiment wurde die Schwierigkeit das Target zu identifizieren durch eine Veränderung der Unterschiede in der Größe zwischen Target und Distraktoren in drei Schritten manipuliert. In dem vorliegenden Experiment wurde die Größe der Distraktoren in 3 Stufen verändert, um somit eine einfach, mittlere und schwierige Aufgabe zu erzeugen. Die entsprechenden Verhältnisse in der Größe zwischen Target und Distraktor wurden dabei in den drei Stufen festgesetzt: einfach 150/80, mittel 150/100 und schwer 150/110.

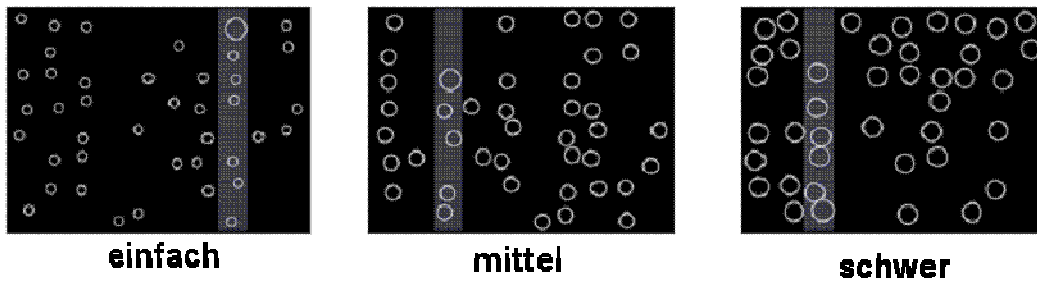


Abbildung 25: Die Abbildung zeigt die drei Schwierigkeitsstufen des Surrogate Reference Tasks (SURT) einfach, mittel und schwer.

Sobald die Versuchsperson das Target identifiziert hat, soll sie dieses mittels eines grauen Balkens, welcher mittels zweier Tasten auf dem Bildschirm verschoben werden kann, auf dem Display markieren. Die Aufgabenbearbeitung kann dabei während der Durchführung beliebig unterbrochen und wieder aufgenommen werden. Sobald eine Aufgabe bearbeitet wurde, erscheint direkt darauf die Folgende.

Critical Tracking Task (CTT)

Bei der Nebenaufgabe des Critical Tracking Task (CTT) handelt es sich um eine kontinuierliche, visuell motorische Trackingaufgabe. Hier muss ein sich kontinuierlich bewegendes Balken mittels zweier Tasten möglichst gut in der Mitte des Bildschirms gehalten werden. Die Bewegungen des Balkens verlaufen dabei per Zufall zum oberen oder unteren Bildschirmrand. Die Bewegungsgeschwindigkeit nimmt logarithmisch mit zunehmender Entfernung von der Mitte zu. Die Bewegungen des Balkens sind somit in der Mitte des Bildschirms geringer als in den Außenbereichen. Damit wird erreicht, dass je länger die Versuchsperson keine Korrekturen in der Aufgabe vornimmt die Abweichung vom Sollzustand größer wird.

Die Stärke der Bewegung des Balkens wurde von $\Lambda = 0.5$ (einfach) zu 1 (mittel) zu 1.5 (schwer) variiert. Somit braucht eine stärkere Bewegung des Balkens ein höheres Maß an visueller und motorischer Kontrolle, da mit einem höheren Λ die zufälligen Bewegungen des Balkens stark zunehmen und somit eine häufigere Kontrolle notwendig machen.

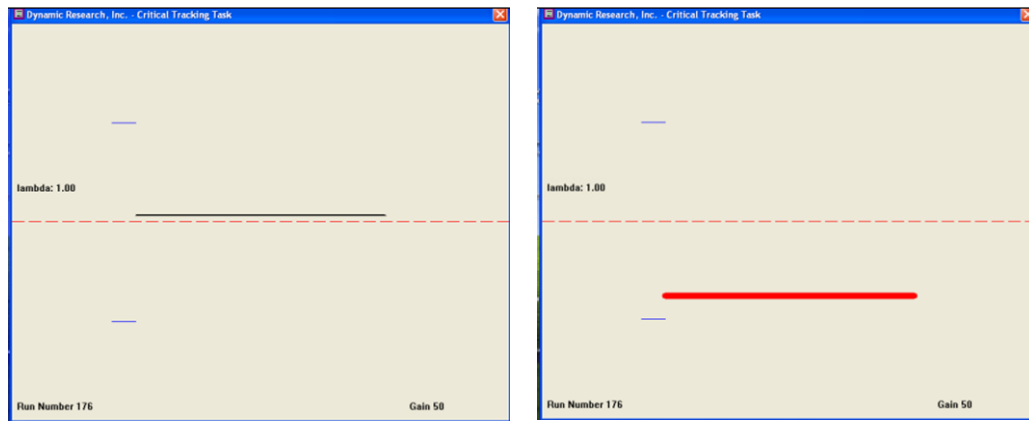


Abbildung 26: Critical Tracking Task (CTT). Der Balken muss so gut wie möglich in der Mitte des Bildes gehalten werden (rote gestrichelte Linie). Erhöht sich die Distanz zur Mitte wird die Bewegung des Balkens schneller und dieser färbt sich rot (rechtes Bild).

Navigationszieleingabe IDrive System

Bei der in dem vorliegenden Versuch verwendeten realen Nebenaufgabe wurde eine Navigationszieleingabe in einem BMW IDrive System verwendet. Es handelte sich dabei um das Vorgängermodell des aktuell in der 5er Modellreihe eingesetzten Bordcomputersystems. Bei dem vorliegenden Versuch wurde eine PC basierte Simulation des Systems sowie ein Tischaufbau mit dem rechts neben der Versuchsperson angeordneten Mittelarmlehne mit IDrive Controller verwendet, welcher blind von der Person bedient werden konnte.



Abbildung 27: Die Abbildung zeigt das Hauptmenü des IDrive Systems (BMW 5er Modellreihe).

Die einzelnen Menüpunkte sowie Untermenüs sind mittels Drehen und Drücken des IDrive Controllers zu erreichen. Bei dem vorliegenden Versuch sollten die Probanden ein Navigationsziel eingeben und die Zielführung starten. Der Usecase startet im Hauptmenü und endet mit dem Starten der Zielführung.

Die Probanden wurden instruiert das Navigationssystem zu nutzen und die Zieladresse ‚Paulusgasse in Hannover‘ zu verwenden.

Die Auswahl der einzelnen Buchstaben erfolgt dabei über die Auswahl der horizontal angeordneten Buchstaben mittels des IDrive Controllers.

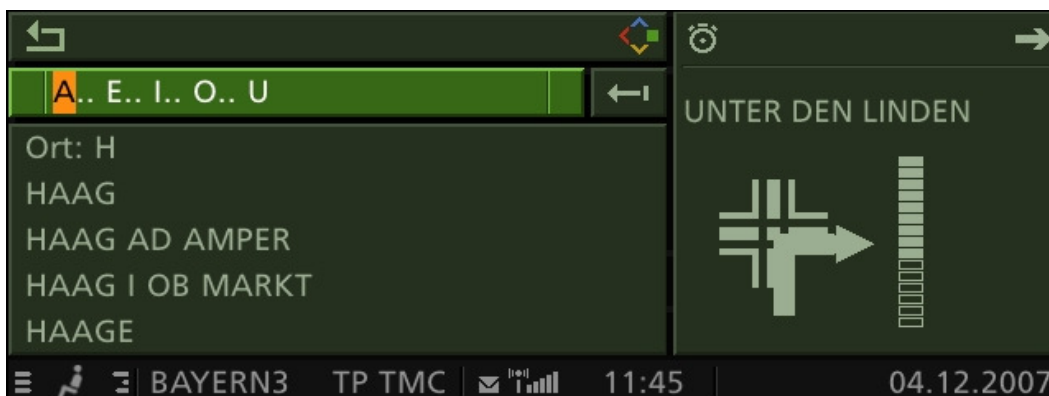


Abbildung 28: Die Abbildung zeigt die Eingabe von alphanumerischen Zeichen über das IDrive System (Eingabe Ort).

Die Eingabe und Auswahl eines Zieles kann dabei während der Bearbeitung der Aufgabe beliebig unterbrochen und fortgesetzt werden.

8.3.2 Ablauf

Auch in diesem Experiment wurde zunächst jede Versuchsperson kurz über das Ziel der vorliegenden Studie aufgeklärt und hatte zu Beginn die Möglichkeit den Lane Change Task zu üben. Die Probanden wurden dabei instruiert den jeweiligen Spurwechsel so schnell, aber auch so exakt wie möglich durchzuführen. Es wurde sichergestellt, dass die Probanden sicher im Umgang mit der Aufgabe waren und

eine Abweichung in der Fahraufgabe < 1 erreicht werden konnte (entsprechend ISO Draft TC22/SC13/WG8).

Nach der Kalibrierung des Eyetrackers führte jede Versuchsperson jede der experimentellen Bedingungen für die Dauer der Länge einer im Lane Change Task gefahrenen Strecke (ca. 3 min) durch: LCT auditiv, LCT auditiv imperativ, LCT visuell (Visuell Baseline), LCT visuell plus SURT (einfach, mittel, schwer), LCT visuell plus CTT (einfach, mittel, schwer) sowie LCT plus Zieleingabe und LCT freies Fahren. Zusätzlich wurden die Nebenaufgaben SURT (einfach, mittel, schwer), CTT (einfach, mittel, schwer) sowie die Navigationszieleingabe jeweils auch ohne Fahraufgabe erhoben. Die Nebenaufgaben wurden über einen Zeitraum von 3 min und die Navigationszieleingabe einmal durchgeführt. Die Abfolge der einzelnen Bedingungen wurde experimentell systematisch variiert.

Insgesamt bearbeitete so jede Versuchsperson 11 Aufgaben in Zusammenhang mit der Fahraufgabe und weitere 7 Aufgaben ohne zusätzliche Fahraufgabe. Bei der Durchführung der Bedingungen unter LCT plus zusätzliche Nebenaufgaben wurden die Versuchspersonen instruiert keine der beiden Aufgaben zu priorisieren, sondern beide Aufgaben so gut und so effizient wie möglich zu bearbeiten.

Direkt im Anschluss an die Bedingungen mit zusätzlicher Fahraufgabe SURT einfach und schwer sowie CTT einfach und schwer gaben die Probanden eine subjektive Einschätzung der auftretenden Beanspruchung bei der Bearbeitung der soeben durchgeführten Aufgabe mittels einer deutschen Version des NASA TLX ohne Gewichtung. Zu Ende des Experiments wurde zusätzlich noch ein demographischer Fragebogen erhoben.

Die Gesamtdauer des hier beschriebenen Experiments betrug ca. 2 Stunden.

8.3.3 Stichprobe

An dem vorliegenden Experiment nahmen insgesamt 18 Probanden teil (14 männlich und 4 weiblich). Die Teilnehmer hatten ein Durchschnittsalter von 28,9 Jahren und waren zwischen 24 bis 47 Jahre alt. Alle Probanden waren zum Zeitpunkt des Experiments im Besitz einer gültigen Fahrerlaubnis und bezeichneten sich selbst als erfahrene Fahrer. Zehn Probanden gaben an durchschnittlich 5.000 bis 10.000 Kilometer pro Jahr zu fahren, acht Probanden gaben an durchschnittlich 10.000 bis 20.000 Kilometer zu fahren.

8.4 Ergebnisse

8.4.1 Fahrleistung

Für die Erfassung der Fahrleistung wurde, wie bereits unter Experiment 2 beschrieben, die mittlere Abweichung auf der Strecke zwischen der Optimallinie und der tatsächlich gefahrenen Spur ermittelt. Es bleibt dabei, wie bereits unter Experiment 2 diskutiert, anzumerken, dass die so ermittelten Abweichungen auf der Straße, aufgrund der Eigenschaften der gewählten Aufgabe und der Bestimmung der Spurabweichung nicht 1:1 auf eine reale Fahrsituation übertragen werden können, sondern vielmehr exemplarisch auftretende beanspruchende Effekte abbilden.

Bei der Analyse der Fahrleistung mussten 2 Probanden aufgrund technischer Probleme bei der Datenerfassung aus der Analyse ausgeschlossen werden. Abbildung 29 zeigt die gemessenen mittleren Spurabweichungen getrennt für alle experimentellen Bedingungen.

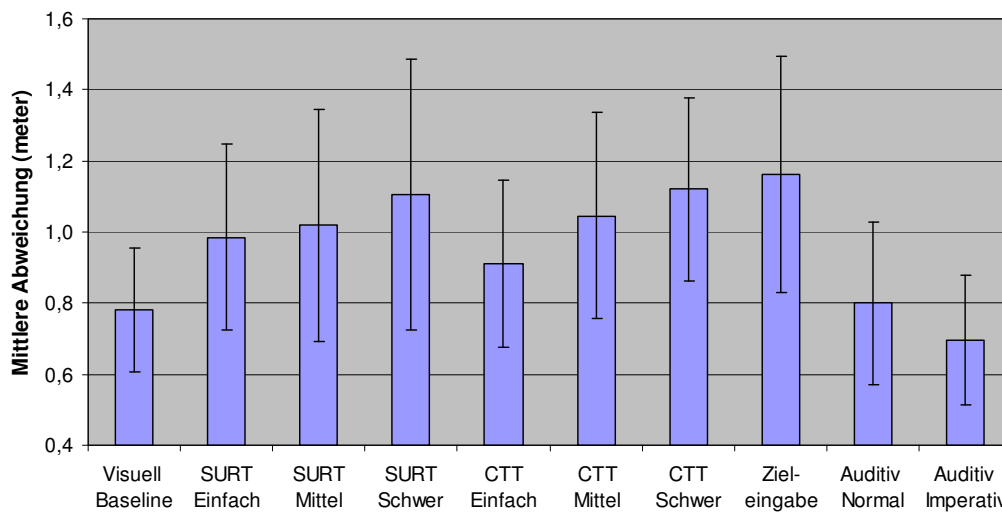


Abbildung 29: Die Abbildung zeigt die Abweichungen in der Fahraufgabe von der Ideallinie über alle experimentellen Bedingungen [Die vertikalen Balken zeigen +/- 1 Standardabweichung].

Eine ANOVA mit Messwiederholung über alle 10 experimentellen Bedingungen zeigte einen signifikanten Effekt für den Faktor der experimentellen Bedingung, $F(9, 135)=23.17$, $p<.001$. Ein geplanter Vergleich zeigte einen signifikanten Unterschied zwischen der visuellen Baseline Bedingung und den drei SURT Bedingungen (einfach, mittel, schwer), $F(1, 15)=60.75$, $p<.001$. Ein weitere geplanter Vergleich zeigte auch einen signifikanten Unterschied zwischen der visuellen Baseline Bedingung und den drei CTT Bedingungen (einfach, mittel, schwer), $F(1, 15)=99.79$, $p<.001$.

Surrogate Reference Task

In einer ANOVA mit Messwiederholung für die drei Stufen des Surrogate Reference Tasks zeigen sich keine signifikanten Unterschiede in der Fahrleistung für die drei Schwierigkeitsstufen der Aufgabe, $F(2, 30)=2.84$; $p=.07$.

Critical Tracking Task

Im Gegensatz zu den Ergebnissen bei der Bearbeitung des SURT zeigte eine ANOVA mit Messwiederholung einen signifikanten Unterschied zwischen den

drei Schwierigkeitsstufen des CTT, $F(2, 30)=38.85$; $p < .01$. Ein LSD PostHoc Test zeigte einen signifikanten Anstieg in der mittleren Abweichung von der Optimallinie zwischen jedem der drei Schwierigkeitsstufen ($p < 0.01$).

Auditive Bedingungen

In einer ANOVA mit Messwiederholung für die Fahrleistungen in den Bedingungen Baseline (visuell), auditiv unspezifisch und auditiv imperativ zeigte sich ein signifikanter Haupteffekt, $F(2,30)=4.07$; $p=.027$. Ein LSD PostHoc Test zeigte hier keinen signifikanten Unterschied zwischen den Bedingungen visuell Baseline und Auditiv normal ($p=.63$). Dafür zeigte sich ein signifikanter Unterschied zwischen der auditiv imperativen Bedingung und den Bedingungen visuell Baseline ($p=.03$) sowie auditiv normal ($p=.01$). Dies bedeutet, dass unter der auditiv imperativen Bedingung die beste Fahrleistung gezeigt wurde.

8.4.2 Index of Cognitive Activity

Im Folgenden sollen die Ergebnisse zu den ermittelten ICA Daten berichtet werden.

Bei den hier berichteten ICA Daten handelt es sich um die für beide Augen aufsummierten und interindividuell über die gesamte Länge des Versuches standardisierte Daten. Somit konnten interindividuellen Differenzen bei auftretenden Beanspruchungen Rechnung getragen werden und aus der experimentellen Betrachtung herausgenommen werden. Bei der Analyse der Daten mussten 3 der insgesamt 18 Versuchspersonen aufgrund technischer Probleme während der Aufzeichnung oder anderer Störeinflüsse während der Versuchsdurchführung aus der Analyse ausgeschlossen werden. Die hier berichteten Analysen beziehen sich somit auf insgesamt 15 gültige Datensätze.

Um in dem vorliegenden Experiment eine noch genauere Analyse der kritischen Situation rund um den Zeitpunkt des Passierens des Schildes (Spurwechsels) zu

erreichen, wurden hier während der Durchführung der Fahrsimulation an jeweils genau den Stellen, an denen die Probanden die Schilder passierten, dieses mittels eines Markers in den ICA Daten vermerkt.

Eine solche Markierung ermöglicht es die jeweils kritischen Situationen über alle 18 Spurwechsel einer im LCT gefahrenen Strecke hinweg zu mitteln und somit die Validität der hier gefundenen Ergebnisse deutlich zu erhöhen.

8.4.2.1 Auswertung Bedingungen mit Fahraufgabe

Aufbereitung der Daten

Für die Aufbereitung der Daten in den Bedingungen mit Fahraufgabe wurden, wie bereits oben beschrieben, jeweils die individuell standardisierten Daten für 7 Sekunden vor dem Passieren des Schildes sowie 2 Sekunden danach bestimmt. Die so gewonnenen Daten wurden dann pro Versuchsperson für jede Sekunde einzeln über alle 18 Spurwechsel gemittelt.

Das Passieren des Schildes geschieht, wie in den folgenden Abbildungen dargestellt, immer zwischen Sekunde 9 und Sekunde 1. Die Sichtbarkeit des Schildes bzw. die auditive Instruktion (oder imperative Signal) wurden auf 40 Meter vor dem Schild festgelegt. Dies bedeutet bei einer konstanten Geschwindigkeit von 60 km/h, dass die Instruktion die Spur zu wechseln jeweils in Sekunde 8 gegeben wurde.

Visuell Baseline und Freies Fahren LCT

Für die Auswertung der Bedingung Freies Fahren wurden während der Durchführung der Fahrt jeweils die einzelnen Spurwechsel der Versuchsperson zeitlich markiert. Um eine Konfundierung der Daten zu vermeiden wurden für die weitere Analysen der ICA Daten daraufhin nur die Fälle mit in die Analyse aufgenommen, bei denen sowohl 7 Sekunden vor dem Spurwechsel als auch 2 Sekunden danach keine weiteren Spurwechsel aufgetreten waren.

In der Bedingung der visuellen Baseline Bedingung sowie des Freien Fahrens ergab sich für die Situation rund um das Passieren des Schildes, wie in Abbildung 30 dargestellt, ein sehr klares Muster in den ermittelten ICA Werten.

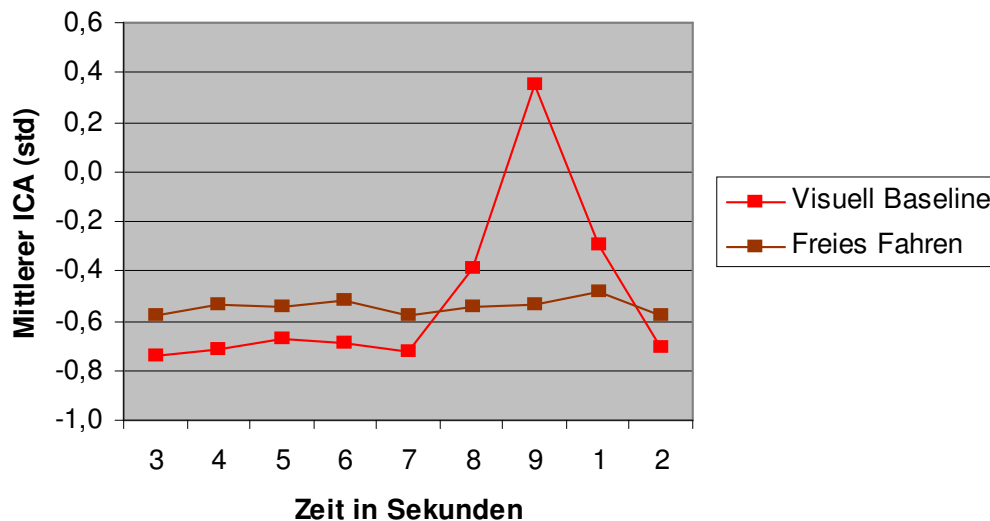


Abbildung 30: Die Abbildung zeigt die mittleren interindividuell standardisierten ICA Werte für die Bedingung Visuell Baseline und Freies Fahren. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

Hier zeigt sich deutlich der Anstieg in den ICA Werten vor dem Schild in der visuellen Baseline Bedingung. Kein Anstieg rund um den Spurwechsel zeigt sich hingegen unter der freien Fahrt Bedingung.

Bei einer ANOVA mit Messwiederholung für die standardisierten ICA Werte der Probanden pro Sekunde, zeigte sich in der Visuellen Baseline Bedingung in einem geplanten Vergleich ein signifikanter Anstieg der ICA Daten in den Sekunden 8, 9 und 1 rund um das Auftreten des Schildes, $F(1, 14)=108.37$; $p < .001$. Im Gegensatz dazu zeigte sich in der Freien Fahrt Bedingung in den Sekunden 8, 9 und 1 rund um das Auftreten des Schildes in einem geplanten Vergleich kein signifikanter Anstieg der ICA Werte, $F(1, 14)= 2.32$; $p = .15$. Eine LSD PostHoc Analyse

zeigte zudem einen signifikanten ($p < .01$) Unterschied zwischen den Bedingungen Visuell Baseline und Freies Fahren in den Sekunden 8, 9 und 1. Kein signifikanter Unterschied zeigte sich für die Sekunden 5 und 6. Der mittlere ICA Wert lag dabei für die Sekunden 3 – 7 sowie 2 für die Visuelle Baseline Bedingung mit -0.709 unter dem für die freie Fahrt Bedingung -0.55.

Auditiv Baseline und Auditiv Imperativ LCT

Für die auditiv-imperative Bedingung wurde den Probanden in Sekunde 5 die semantische Information über den auszuführenden Spurwechsel gegeben. Das imperative Signal in Form eines kurzen auditiven Signaltons, wurde dann in Sekunde 8 präsentiert und leitete die Durchführung des eigentlichen Spurwechsels ein.

Für die beiden Bedingungen mit auditiven Instruktionen, bei denen die Schilder des Lane Change Tasks gelöscht wurden und die Instruktionen stattdessen auditiv via Kopfhörer präsentiert wurden, ergab sich ebenfalls ein klares Bild in den ermittelten ICA Daten, wie in Abbildung 31 dargestellt.

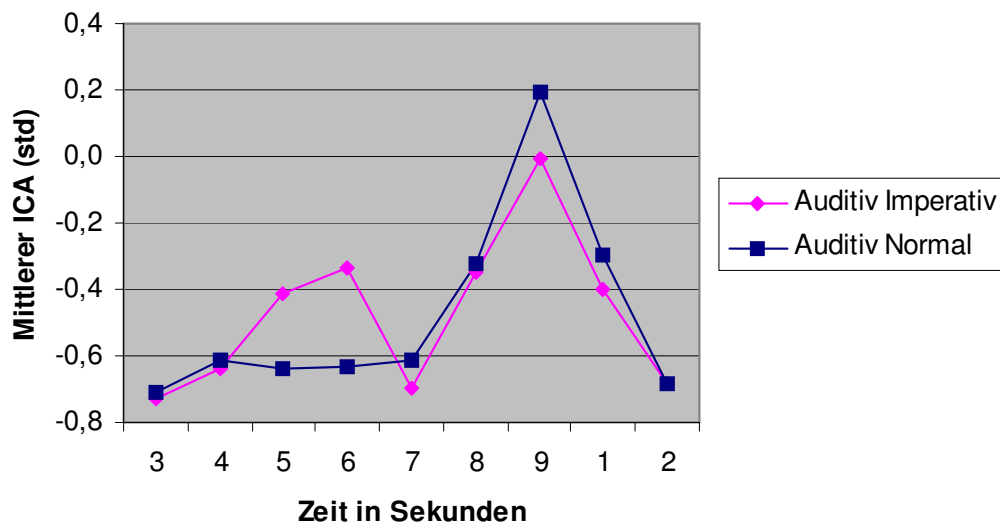


Abbildung 31: Die Abbildung zeigt die mittleren interindividuell standardisierten ICA Werte für die Bedingung Auditiv Normal und Auditiv Imperativ. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

Deutlich erkennbar ist der signifikante Anstieg in der auditiv-imperativen Bedingung durch die Verarbeitung der semantischen Instruktion um die Sekunde 5 und 6 gegenüber der auditiv-normalen Bedingung. Anzumerken bleibt hier, dass zu diesem Zeitpunkt für die Probanden keine weitere visuelle Information in der Fahrsituation gegeben wurde. Die beiden Bedingungen unterscheiden sich somit an diesem Zeitpunkt nur durch die auditive Instruktion. Der visuelle Eindruck für die Probanden hingegen war der gleiche.

In einer ANOVA mit Messwiederholung für die wieder über alle Spurwechsel und pro Versuchsperson gemittelten ICA Werte ergab sich ein signifikanter geplanter Kontrast für die Sekunden 5 und 6 zwischen den Bedingungen auditiv-normal und auditiv-imperativ, $F(1, 14)=21.32$; $p<.001$). Hier lag das ICA Niveau in der auditiv-imperativen Bedingung mit einem mittleren ICA von -0,375 deutlich über dem der auditiv-normalen Bedingung mit -0.635.

Auch ergab sich für die Sekunden 8, 9 und 1 ein signifikanter Unterschied zwischen auditiv-imperativ (Mittelwert = -0.253) und auditiv-normal (Mittelwert = -0.143), $F(1,14)=5.581$; $p=.033$. Hier lagen die ICA Werte unter der Bedingung auditiv-imperativ signifikant unterhalb der Werte in der Bedingung auditiv-normal. Eine solch reduzierte Beanspruchung in der Bedingung auditiv-imperativ lässt sich dadurch erklären, dass unter dieser Bedingung zum Zeitpunkt der Initialisierung der Handlung die semantische Verarbeitung bereits erfolgt war und somit die resultierende Gesamtbeanspruchung reduziert werden konnte.

Für die Sekunden 3, 4, 7 und 2 ergab sich entsprechend der Erwartung kein signifikanter Unterschied in den ICA Werten zwischen den beiden auditiven Bedingungen ($F(1, 14)=2.414$; $p = .14$).

Visuell Baseline und Auditiv Normal

Bei einem Vergleich der beiden Instruktionsmodalitäten Visuell Baseline und Auditiv Normal (s. Abbildung 32) zeigte sich kein signifikanter Unterschied in den jeweiligen Beanspruchungsverläufen der ICA Daten.

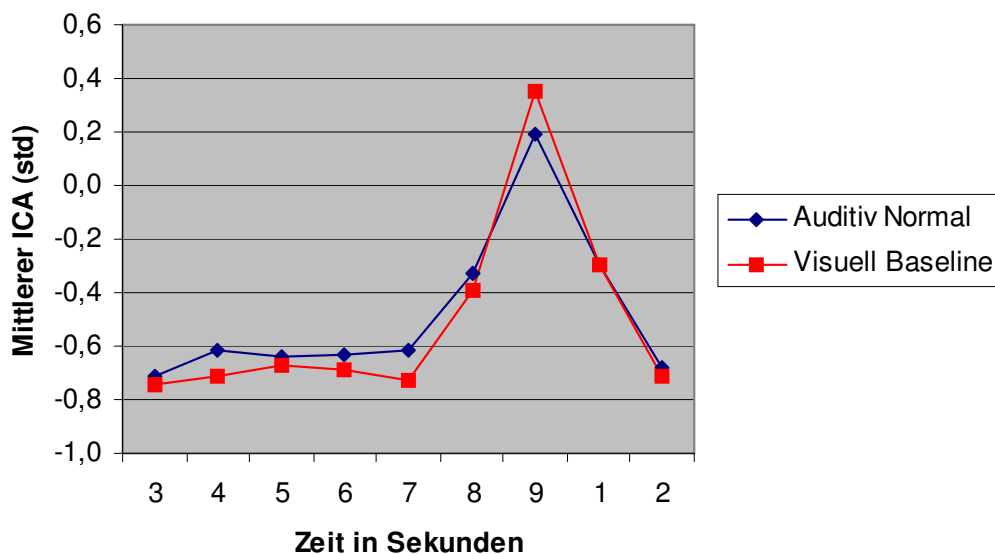


Abbildung 32: Die Abbildung zeigt die mittleren interindividuell standardisierten ICA Werte für die Bedingung Visuell Baseline und Auditiv Normal. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

In einer ANOVA mit Messwiederholung für die jeweils über alle Spurwechsel und pro Versuchsperson gemittelten ICA Werte ergab sich insgesamt über alle Sekunden hinweg kein signifikanter Unterschied zwischen den beiden Bedingungen Visuell Baseline und Auditiv Normal, $F(1, 14)=1.28$, $p=.27$.

Das vorliegende Ergebnis zeigt, dass hier die Modalität der Instruktion und somit die perzeptuelle Wahrnehmung alleine bzw. eine Orientierungsreaktion hin zum Schild einen solchen Effekt in den ICA Daten nicht beeinflussen kann.

LCT plus SURT einfach, mittel und schwer

Eine ANOVA mit Messwiederholung über die drei Bedingungen des LCT plus SURT unter den Schwierigkeitsstufen einfach, mittel und schwer zeigte keinen Haupteffekt in den ICA Werten für die Schwierigkeit der Aufgabe über alle Sekunden hinweg, $F(2, 26)=0.043$, $p = .95$. Dafür zeigte ein geplanter Vergleich entsprechend der Erwartung aus den Ergebnissen des vorangegangenen Experiments einen signifikanten Abfall in den ICA Werten während des Passierens des Schildes über die drei Schwierigkeitsstufen hinweg für die Sekunden 3 bis 9 gegen Sekunde 1 und 2 (s. Abbildung 33).

Darüber hinaus zeigte sich über die drei Schwierigkeitsstufen hinweg ein signifikanter Unterschied zwischen den Sekunden 3 bis 7 gegenüber Sekunde 8. Zur Anmerkung: Bei Sekunde 8 wird das Schild im LCT sichtbar. Somit scheint sich hier eine Art kompensativer strategischer Effekt für die Beanspruchung zu zeigen, welcher im Folgenden auch über die jeweiligen Aktivitätsmuster in der SURT zu diskutieren sein wird.

Ein LSD PostHoc Test zeigte zudem einen signifikanten Anstieg ($p=.01$) von Sekunde 8 zu Sekunde 9 für die schwierige Bedingung des SURT; in der mittleren sowie der einfachen Bedingung konnte von Sekunde 8 zu Sekunde 9 kein statistisch signifikanter Anstieg festgestellt werden.

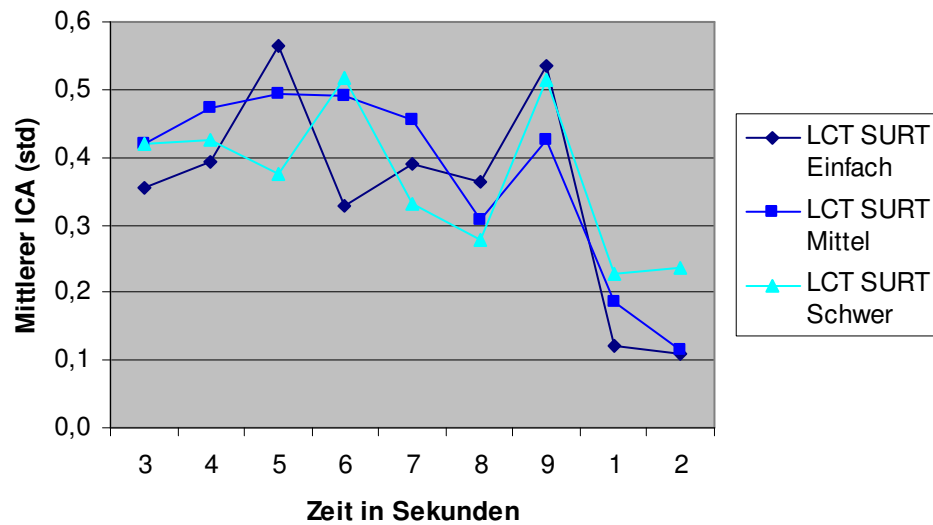


Abbildung 33: Die Abbildung zeigt die ICA Werte für die drei Schwierigkeitsstufen der SURT über den zeitlichen Verlauf hinweg. Deutlich wird dabei auch der Abfall in den ICA Werten um die Sekunden 8 und 9 direkt nach dem Spurwechsel.

LCT plus SURTeinfach, mittel und schwer und Navigationszieleingabe

Für die Bedingung LCT mit zusätzlicher Bearbeitung der Navigationszieleingabe, wie in Abbildung 34 dargestellt, zeigte sich in einer ANOVA mit Messwiederholung bei einem geplanten Vergleich kein signifikanter Abfall in den ICA Werten für die Sekunden 1 bis 7 gegenüber Sekunde 8 und 9, $F(1, 10) < 1$, $p = .99$. Dafür zeigte sich ein signifikanter Abfall in den ICA Werten für die Sekunde 8 gegenüber der Sekunden 3 bis 7 sowie 9 bis 2, $F(1, 10) = 5.519$; $p = .04$.

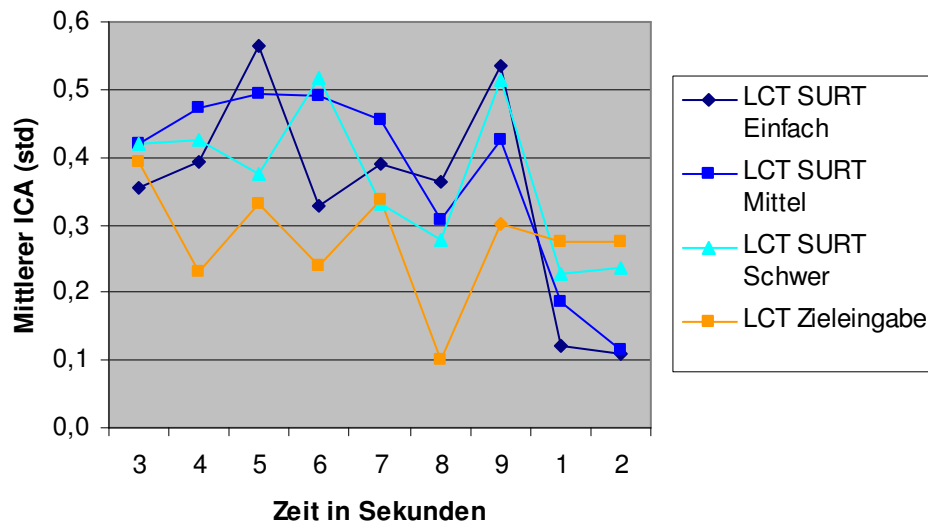


Abbildung 34: Die Abbildung zeigt die ICA Werte bei der Bearbeitung der Zieleingabe sowie der drei Schwierigkeitsstufen der SURT über die Zeit. Es zeigt sich bei der Bearbeitung der Navigationszieleingabe ein deutlicher Abfall der ICA Werte für die Sekunde 8.

Vergleicht man nun die zeitlichen Muster der Bedingungen LCT plus Navigationszieleingabe und LCT plus SURT (einfach, mittel, schwer), so können folgende Befunde festgehalten werden. Kommt es bei der Bearbeitung des SURT zu einem Abfall der ICA Daten von Sekunde 9 zu Sekunde 1, so bleiben hier die Werte unter der Bearbeitung einer Navigationszieleingabe auf einem stabilen Niveau. Sowohl unter den drei Schwierigkeitsstufen des SURT als auch der Navigationszieleingabe kommt es zu Sekunde 8 zu einem signifikanten Abfall in den ICA Werten. Ein solcher Abfall ist unter der Bearbeitung einer Navigationszieleingabe deutlich stärker ausgeprägt, als unter den drei SURT Bedingungen.

Da zu Sekunde 8 das Schild und somit auch die Instruktion die Spur zu wechseln für die Versuchsperson sichtbar wird, ist ein solcher Effekt mit einer Art kompensatorischer Reaktion zu erklären. Insbesondere bei der Bearbeitung einer realistischen Aufgabe zusätzlich zur Fahraufgabe versucht der Fahrer hier für die anstehende kritische Fahrsituation, bei der ein Spurwechsel durchgeführt werden muss, möglichst viele Kapazitäten frei zu stellen, um sicher und effizient in der Fahrsituation reagieren zu können.

Vergleich LCT SURT (einfach, mittel, schwer) sowie Zieleingabe gegenüber Visuell Baseline

In einer weiteren ANOVA mit Messwiederholung wurde ein Vergleich zwischen den drei Schwierigkeitsstufen des SURT sowie der Zieleingabe bei gleichzeitiger Bearbeitung der Fahraufgabe gegenüber der visuellen Baseline Bedingung durchgeführt (s. Abbildung 35). Ein geplanter Vergleich zwischen der Bedingung visuell Baseline und den drei Schwierigkeitsstufen des SURT sowie der Navigationszieleingabe über alle 9 Sekunden zeigte zunächst signifikant niedrigere ICA Werte unter der visuellen Baseline Bedingung als unter den anderen Bedingungen, $F(1, 10)=212.83$, $p<.001$. Eine genauere Analyse der Daten mittels eines LSD Post Hoc Tests zeigte allerdings einen erstaunlichen Befund, wonach sich die drei Bedingungen mit zusätzlicher SURT sowie die Navigationszieleingabe zu der visuellen Baseline nur in den Sekunden 1 bis 8 signifikant unterschieden ($p<.001$). In Sekunde 9 hingegen zeigte sich kein Unterschied in den ermittelten ICA Daten zwischen den Bedingungen visuell Baseline und den übrigen Bedingungen ($p>.05$). Somit zeigt sich nur zum Zeitpunkt der kritischen Fahrsituation kein signifikanter Unterschied zwischen den Bedingungen mit und ohne zusätzliche SURT Aufgabe oder bei der Bearbeitung einer Navigationszieleingabe.

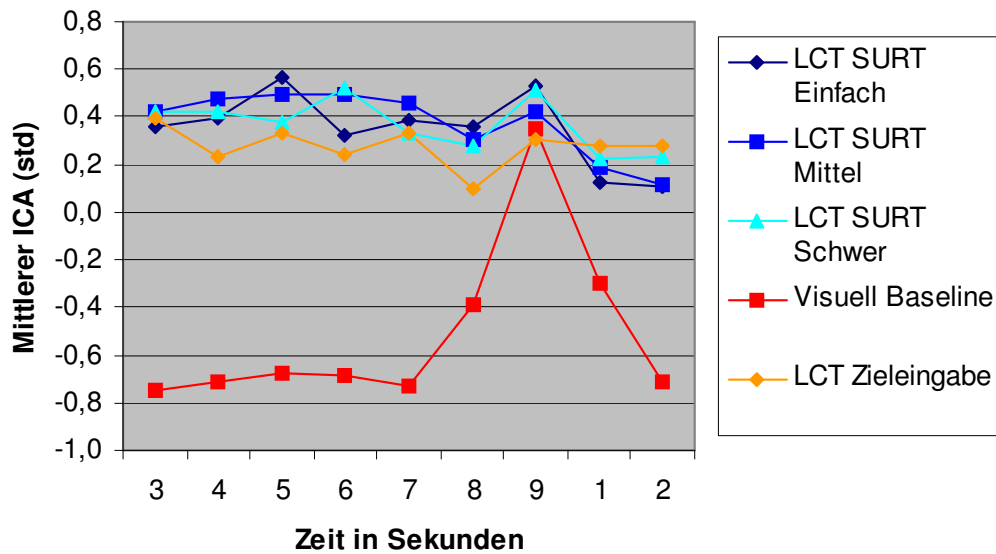


Abbildung 35: Die Abbildung zeigt die ICA Werte über den zeitlichen Verlauf für die drei Schwierigkeitsstufen der SURT, der Navigationszieleingabe sowie der visuellen Baseline Bedingung. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

Critical Tracking Task (CTT) einfach, mittel, schwer mit LCT

Für die Bedingung LCT mit zusätzlicher CTT Aufgabe zeigte sich, wie in Abbildung 36 dargestellt, ein deutlich anderes Befundmuster als für die SURT oder die Navigationszieleingabe.

Eine ANOVA mit Messwiederholung auf den drei Schwierigkeitsstufen des CTT zeigte in einem geplanten Vergleich einen signifikanten Anstieg in den ICA Werten für die Sekunden 8, 9 und 1 gegenüber den Sekunden 2, und 3 bis 7 über die drei Schwierigkeitsstufen hinweg, $F(1, 12)=17.29$, $p=.001$. Zudem zeigten sich in einem LSD PostHoc Test ein signifikanter Unterschied in der Sekunde 9 und 1 zwischen der einfachen und der schweren Bedingung ($p<.001$) sowie der einfachen und der mittleren Bedingung des CTT ($p<.001$). Insgesamt zeigte sich eine schnellere Normalisierung der ICA Werte nach dem Peak um Sekunde 1 für die einfache Bedingung gegenüber der mittleren und schweren Bedingung.

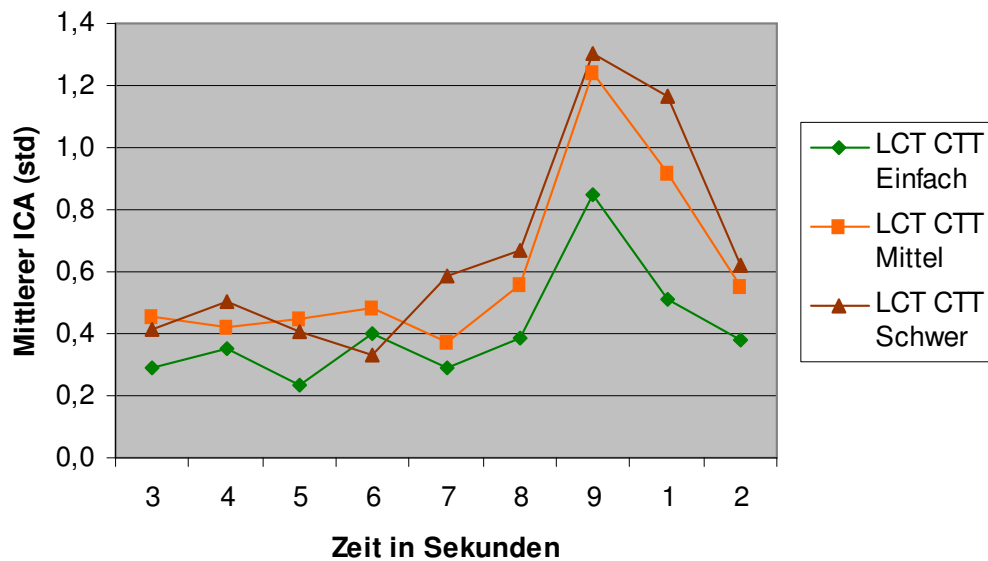


Abbildung 36: Die Abbildung zeigt den Verlauf der ICA Werte über die Zeit bei der Bearbeitung der CTT zusammen mit der Fahraufgabe. Deutlich wird der Anstieg in den ICA Werten in den Sekunden 6, 7 und 8 während des Spurwechsels.

Vergleich LCT CTT (einfach, mittel, schwer) gegenüber Visuell Baseline

In einer weiteren ANOVA mit Messwiederholung für die drei Schwierigkeitsstufen des CTT sowie der visuellen Baseline Bedingung zeigte ein geplanter Vergleich zwischen der Bedingung visuell Baseline und den drei Schwierigkeitsstufen des CTT über alle 9 Sekunden signifikant niedrigere ICA Werte unter der visuellen Baseline Bedingung als unter den anderen Bedingungen, $F(1, 12)=155.87$, $p<.001$. Eine genauere Analyse der Daten mittels eines LSD Post Hoc zeigte hier einen signifikanten Unterschied ($p<.001$) zwischen der visuellen Baseline Bedingung und den drei Bedingungen des CTT zu jeder der 9 Sekunden. Dies bedeutet, dass hier, anders als unter der Bearbeitung der SURT Aufgabe, auch zu den Zeitpunkten der kritischen Fahrsituationen ein deutlicher Anstieg in den ICA Werten zu verzeichnen war.

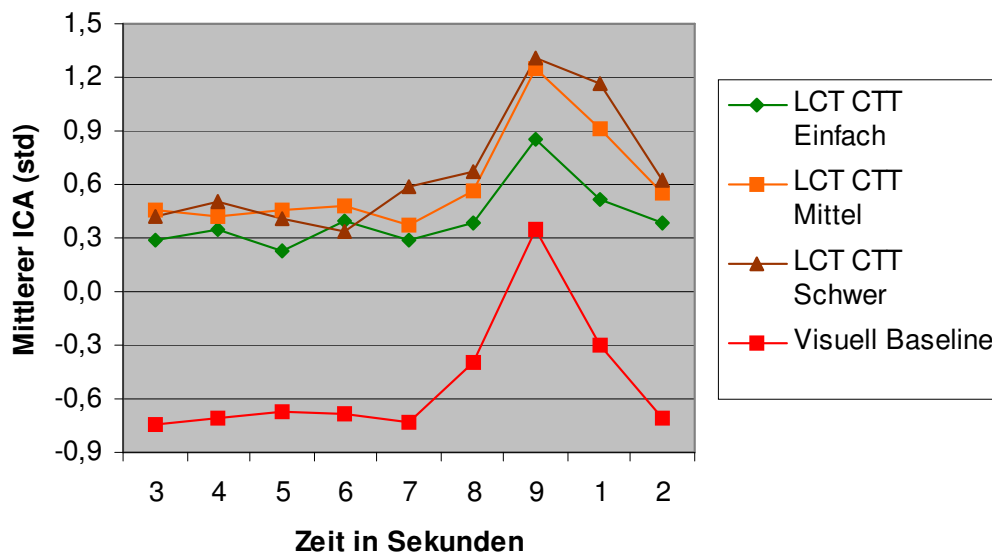


Abbildung 37: Die Abbildung zeigt die ICA Werte über den zeitlichen Verlauf für die drei Schwierigkeitsstufen der CTT sowie der visuellen Baseline Bedingung. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

Vergleich aller Nebenaufgaben bei zusätzlicher Fahraufgabe

Vergleicht man nun zusammenfassend die Ergebnisse aus den ICA Daten für die Bedingungen LCT mit SURT, CTT und Navigationszieleingabe, so wird deutlich, dass die Probanden für die SURT sowie die CTT und Navigationszieleingabe, bezogen auf die Situationen rund um die Spurwechsel, jeweils unterschiedliche aufgabenspezifische Dynamiken bzw. Strategien entwickeln (s. Abbildung 38). Bei einer vergleichenden ANOVA mit Messwiederholung über alle Bedingungen im LCT bei zusätzlicher Nebenaufgabe (CTT einfach, mittel, schwer, SURT einfach, mittel, schwer und Navigationszieleingabe) zeigte sich insgesamt kein signifikanter Unterschied in den ICA Werten für die Bedingungen SURT (einfach, mittel, schwer) und der Navigationszieleingabe, $F(1, 10)=2.254$, $p=.16$. Ein Vergleich der ICA Werte zwischen den drei Schwierigkeitsstufen des CTT und der Navigationszieleingabe zeigt hingegen einen deutlichen Unterschied in den Beanspruchungsdaten, $F(1, 10)=25.09$, $p<.001$.

Insgesamt liegen die ICA Werte in den CTT Bedingungen in den Sekunden 8, 9, 1 und 2 deutlich über den Werten in der SURT oder der Navigationszieleingabe,

$F(1, 10)=66.18$, $p<.001$. Wohingegen sich für die Sekunden 3 – 7 kein signifikanter Unterschied für die ICA Werte in der SURT und Navigationszieleingabe gegenüber den Werten in der CTT Aufgabe zeigt, $F(1, 10)=0.11$, $p=.79$.

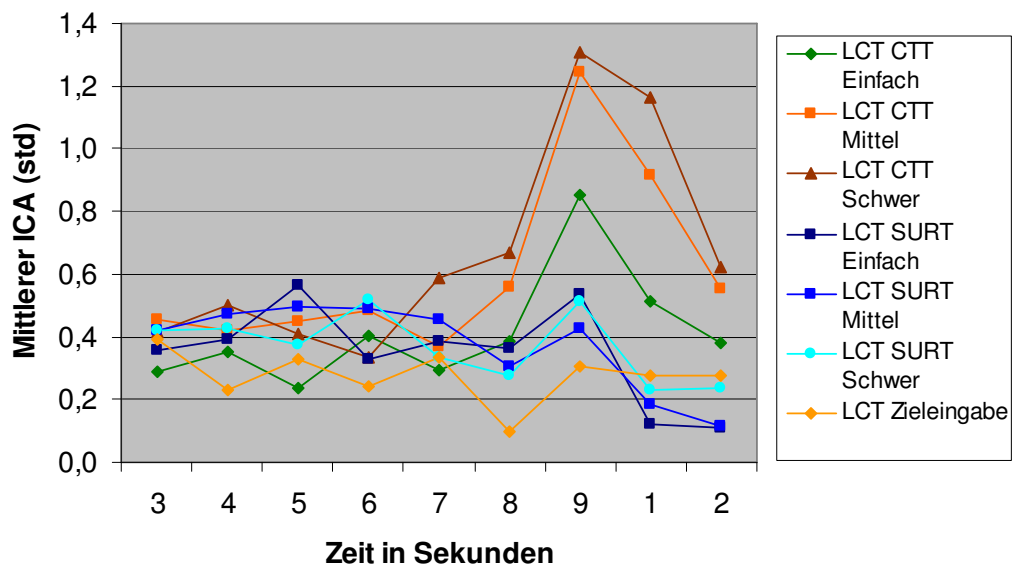


Abbildung 38: Die Abbildung zeigt den Verlauf der ICA Werte für jede der 7 Nebenaufgaben. Deutlich wird der unterschiedliche Verlauf bei der CTT gegenüber der SURT und der Navigationszieleingabe. Das Passieren des Schildes erfolgte zwischen Sekunde 9 und 1.

Die hier aufgezeigten Unterschiede machen deutlich, dass sich insbesondere rund um die kritischen Situationen der Spurwechsel die Beanspruchungsverläufe unterscheiden. Insgesamt ist festzustellen, dass bei der Bearbeitung der CTT Aufgabe, welche eine kontinuierliche Zuwendung zur Aufgabe erfordert, im Gegensatz zur SURT oder der Bedienung des IDrive Systems, bei denen eine Unterbrechung des Aufgabenverlaufes zu jeder Zeit möglich ist, ein deutlicher Anstieg in den Beanspruchungsverläufen zu beobachten ist. Dies legt die Vermutung nahe, dass hier ein kontinuierliches Aufgabencharakteristikum eine Art kompensative Reaktion, wie sie bei einer diskreten Aufgabencharakteristik auftritt, verhindert. Die Folge ist ein deutlicher Anstieg in der Beanspruchung für den Fahrer.

Fahraufgabe mit und ohne Nebenaufgabe

Bei einem Vergleich der ICA Werte für die Visuelle Baseline Bedingungen gegen die Bedingungen mit zusätzlicher Fahraufgabe zeigte sich in einer ANOVA mit Messwiederholung mit den Bedingungen Visuell Baseline, CTT (einfach, mittel schwer) sowie SURT (einfach, mittel, schwer) und Navigationszieleingabe für einen geplanten Vergleich visuell Baseline gegen Bedingungen mit Nebenaufgabe insgesamt ein signifikanter Unterschied, $F(1, 10)=198.99$, $p<.001$.

Eine weitere PostHoc Analyse zeigte für Sekunde 9 zwischen den visuellen Baseline Bedingung und der CTT für einfach, mittel und schwer einen signifikanten ($p<.01$) Unterschied. Hier ergaben sich höhere ICA Werte für die CTT Aufgabe. Für die SURT einfach, mittel und schwer sowie die Navigationszieleingabe ergaben sich in Sekunde 9 keine signifikanten ($p>.05$) Unterschiede zu der visuellen Baseline in den ICA Werten.

Die gemessene Beanspruchung lag somit in den kritischen Situationen, zum Zeitpunkt des Spurwechsels, nur bei einer kontinuierlichen Nebenaufgabe (CTT) deutlich über der Beanspruchung der visuellen Baselinebedingung. Zu dem Zeitpunkt der kritischen Situationen gab es keine signifikanten Unterschiede in der gemessenen Beanspruchung zwischen der visuellen Baselinebedingung und den diskreten Aufgaben SURT sowie Navigationszieleingabe.

8.4.2.2 Bedingungen ohne Fahraufgabe

Insgesamt bearbeiteten die Probanden die drei Nebenaufgaben auch ohne zusätzliche Fahraufgabe insgesamt jeweils für etwa eine Minute. Auch hier kamen für die Auswertung der ICA Daten wie in Abbildung 39 dargestellt die über beide Augen aufsummierten und individuell standardisierten Daten zum Einsatz.

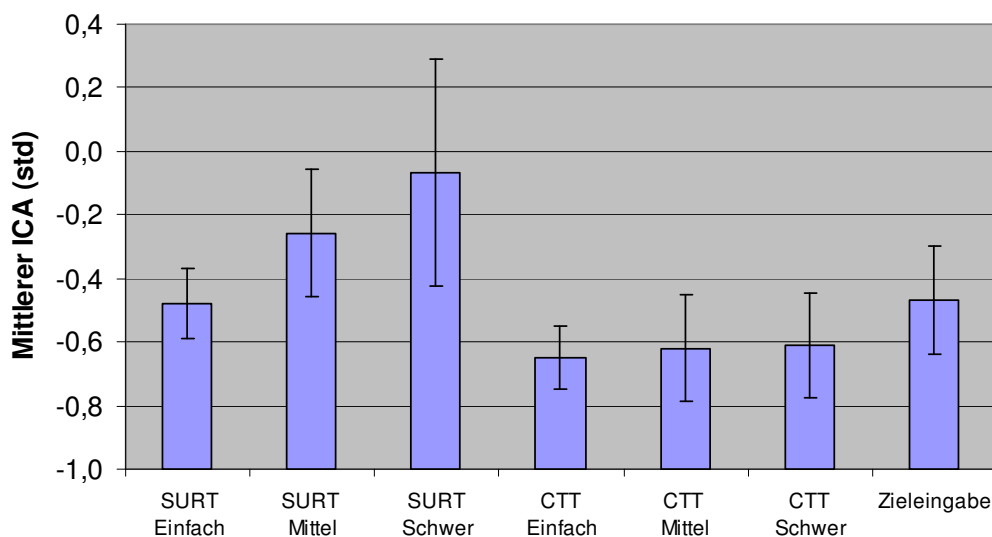


Abbildung 39: Die Abbildung zeigt die mittleren ICA Werte für die drei Schwierigkeitsstufen des SURT sowie des CTT und der Navigationszieleingabe [vertikale Balken zeigen +/- 1 Standardabweichung].

Eine ANOVA mit Messwiederholung zeigte einen signifikanten Anstieg in jeder der drei Schwierigkeitsstufen des SURT, $F(2, 26) = 12.97$, $p < .001$. Wohingegen in einer ANOVA mit Messwiederholung kein Anstieg in den ICA Werten für die drei Schwierigkeitsstufen der CTT Aufgaben zu beobachten war, $F(2, 26) = 0.13$, $p = .87$. Die Navigationszieleingabe verhält sich hier in Bezug auf die ICA Werte ähnlich der einfachen SURT Aufgabe.

Vergleich ICA Werte mit und ohne Fahraufgabe

Für einen Vergleich der ICA Werte in den Situationen mit und ohne Nebenaufgabe wurden die individuellen Mittelwerte für die jeweiligen Bedingungen ermittelt und in einer ANOVA mit Messwiederholung analysiert. Es zeigte sich eine klare Unterscheidung der Aufgaben mit vs. ohne Fahraufgabe, $F(1, 9)=244.93$, $p<.001$. Ein LSD PostHoc Test zeigte einen signifikanten Anstieg für jede einzelne Bedingung für den ICA bei zusätzlicher Fahraufgabe ($p<.001$)

8.4.3 Ergebnisse Nebenaufgaben

Im Folgenden sollen nun die Ergebnisse aus den Nebenaufgaben SURT und CTT berichtet werden. Dabei sollen zunächst die Eingabeaktivitäten und anschließend die Leistung unter beiden Aufgaben diskutiert werden. Für die Navigationszeileingabe konnten aufgrund der Struktur der Aufgabe keine kontinuierlichen Eingabeaktivitäten über die Zeit bestimmt werden.

Bei der Analyse der Eingabeaktivitäten wurde, wie bereits bei der Diskussion der ICA Daten, ein Verfahren verwendet, bei dem jede Eingabe seitens des Probanden an einem externen Gerät zur Steuerung der Aufgabe mit einem synchronisierten Zeitstempel versehen wurde. So konnten für die Auswertung Handlungen in der Nebenaufgabe und Situationen in der Fahraufgabe miteinander in Beziehung gesetzt werden.

Eine Person musste bei der Analyse der Daten wegen technischer Probleme ausgeschlossen werden.

8.4.3.1 Surrogate Reference Task

Aktivitätsverlauf

Für die Bearbeitung der SURT wurden zunächst die Aktivitätsverläufe bei der Bearbeitung der Aufgabe zusätzlich zur Fahraufgabe bestimmt. Dazu wurde analog zu der Analyse der ICA Daten, zu jeder Sekunde rund um den Zeitpunkt des Spurwechsels in der Fahraufgabe, die Wahrscheinlichkeit einer Eingabe zu einer bestimmten Sekunde in der Nebenaufgabe bestimmt (in Prozent). Die so gewonnenen Daten wurden daraufhin über alle 18 Spurwechsel gemittelt und sind in Abbildung 40 dargestellt.

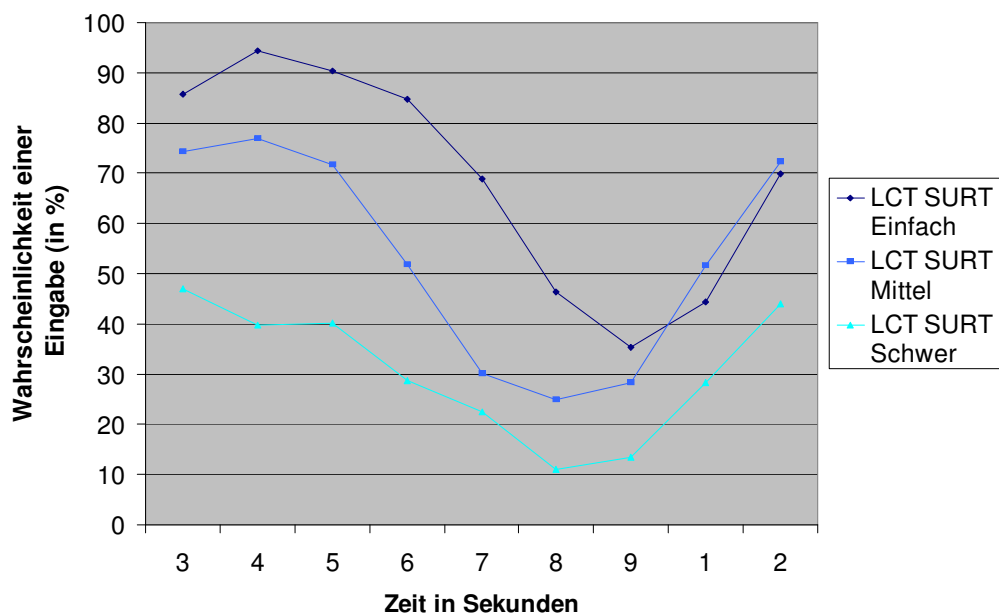


Abbildung 40: Die Abbildung zeigt die prozentuale Wahrscheinlichkeit einer Eingabe unter den drei Schwierigkeiten des SURT rund um den Zeitpunkt eines Spurwechsels (Zwischen Sekunde 9 und 1).

Bei der Betrachtung der Aktivitätsverläufe bei der Bearbeitung der SURT zeigte sich insbesondere ein auffälliger Abfall in der Wahrscheinlichkeit einer Eingabe

kurz vor dem Passieren des Schildes sowie ein erneutes Ansteigen der Werte kurz nach Passieren des Schildes (zwischen Sekunde 9 und 1).

Eine ANOVA mit Messwiederholung über die pro Person ermittelten Wahrscheinlichkeiten einer Eingabe in den Sekunden 1 bis 9 sowie der drei unterschiedlichen Schwierigkeitsstufen zeigte einen signifikanten Unterschied zwischen den drei Schwierigkeitsstufen, $F(2, 32)=48.54$, $p<.001$. Zudem zeigte sich ein signifikanter Haupteffekt für den Zeitpunkt, $F(8, 128)=91.81$, $p<.001$.

Ein geplanter Vergleich zwischen den Sekunden rund um den Spurwechsel (Sekunden 8,9 und 1) und zu den Zeitpunkten an denen keine Spurwechsel erforderlich waren (Sekunden 2 – 7), zeigte eine deutliche Reduktion der Eingaben über alle drei Schwierigkeitsstufen, $F(1,16)=8.70$; $p<.001$. Weitere Vergleiche zeigten hier einen signifikanten Abfall wovon? getrennt für jede der drei Schwierigkeitsstufen zum Zeitpunkt des Passierens des Schildes.

Analysiert man die Aktivitätsverläufe unter den drei Schwierigkeitsstufen genauer, so zeigte ein entsprechender LSD Post Hoc Test stabile Aktivitätsniveaus zwischen den Sekunden 3 bis 5 (mittel und schwer) bzw. 6 (einfach). Dies bedeutet, dass zu diesem Zeitpunkt die Wahrscheinlichkeit einer Eingabe auf einem etwa gleichbleibend hohen Niveau gehalten wurde. Ein signifikanter Abfall in der Aktivität ist unter einer einfachen Bedingung ab Sekunde 6 zu beobachten, unter der mittleren und schweren Bedingung bereits ab Sekunde 5. Dies ist insofern erstaunlich, dass das eigentliche Schild, welches den entsprechenden Spurwechsel instruiert, erst ab Sekunde 8 überhaupt sichtbar wird.

Vergleicht man zusätzlich zu den Aktivitätsverläufen während der Fahrt auch die Aktivitätsniveaus in den Bedingungen ohne zusätzliche Fahraufgabe, wie in Abbildung 41 dargestellt, so zeigte sich, dass die Aktivitätsniveaus in den Sekunden 3 bis 5 (mittel und schwer) bzw. 3 bis 6 (einfach) sehr nahe an denen ohne zusätzliche Fahraufgabe lagen.

In den Aktivitätsmustern ohne zusätzliche Fahraufgabe zeigten sich keinerlei strategischen Effekte. Die drei Schwierigkeitsstufen unterscheiden sich in den Bedingungen ohne zusätzliche Nebenaufgabe signifikant auf jeder Stufe ($F(2, 30)=82.17, p<.001$) und sind daher in Abbildung 41 als konstante Werte dargestellt. Insgesamt werden unter der einfachen Bedingung deutlich mehr Eingaben getätigt als unter der mittleren und unter dieser wiederum deutlich mehr als unter der schweren Bedingung.

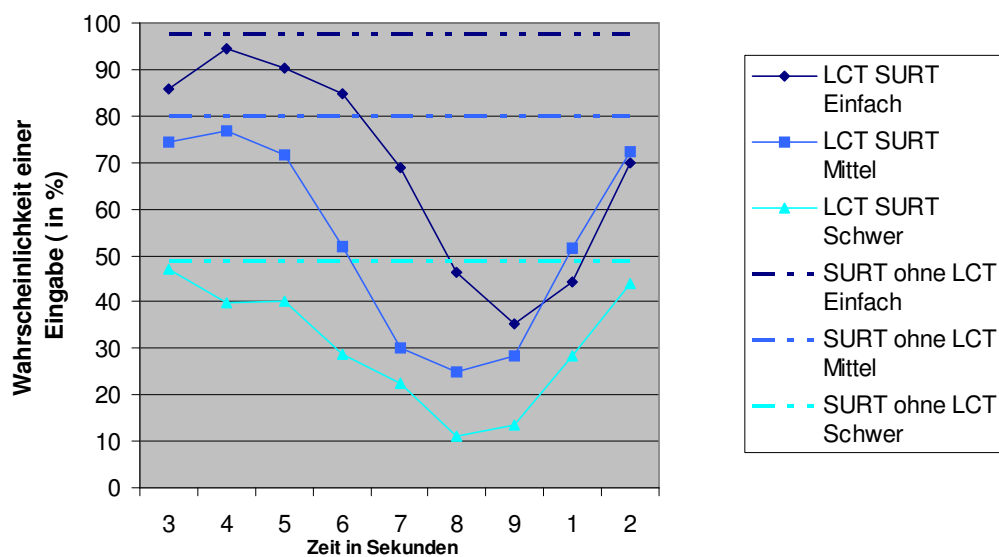


Abbildung 41: Die Abbildung zeigt die Aktivitätsverläufe in den Bedingungen mit Fahraufgabe zusammen mit den mittleren Aktivitätsniveaus ohne zusätzliche Fahraufgabe (horizontale gestrichelte Linien).

Anzahl an bearbeiteten Aufgaben und Fehlerrate

Im Folgenden soll die Anzahl der richtig und der falsch gelösten Aufgaben pro Bedingung sowohl in der Bedingung mit als auch ohne zusätzliche Fahraufgabe berichtet werden.

Da die Gesamtbearbeitungsdauer der Aufgabe ohne zusätzliche Fahraufgabe um ca. 100 Sekunden kürzer war als in der Bedingung mit Fahraufgabe (hier wurde die Bearbeitung während des Durchfahrens einer gesamten Teilstrecke des LCT durchgeführt ~ 3min) wurde die Anzahl der Aufgaben an der jeweiligen Gesamt-

bearbeitungszeit relativiert. Die hier berichteten Leistungsdaten beziehen sich somit auf die errechnete Leistung pro Zeiteinheit (1 Sekunde); der Mittelwert mit LCT beträgt 163 Sekunden, ohne LCT 60 Sekunden.

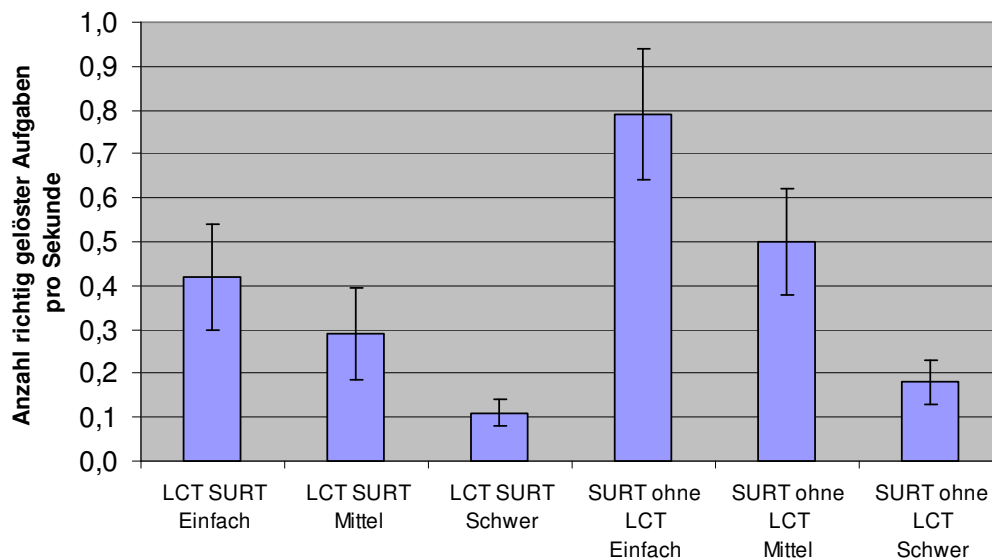


Abbildung 42: Die Abbildung zeigt die Anzahl der richtig gelösten Aufgaben pro Sekunde, errechnet an der jeweiligen Gesamtbearbeitungszeit pro Bedingung [vertikale Balken zeigen +/- 1 Standardabweichung].

In einer ANOVA mit Messwiederholung über die Anzahl richtig gelöster Aufgaben mit und ohne Fahraufgabe ergab sich ein signifikanter Unterschied für den Faktor Fahraufgabe, $F(1, 14)=250.58$, $p<.001$. Es wurden deutlich mehr Aufgaben in der Bedingung ohne zusätzliche Fahraufgabe gelöst.

Zusätzlich zeigte sich ein signifikanter Unterschied für die drei Schwierigkeitsstufen der Aufgabe, $F(2, 28)=102.01$ $p<.001$. Je schwieriger die Bedingung, desto weniger Aufgaben werden gelöst. Auch zeigte sich eine signifikante Interaktion zwischen beiden Faktoren, $F(2, 28)= 53.86$; $p<.001$.

Eine zusätzliche Analyse des prozentualen Fehleranteils der bearbeiteten Aufgaben zeigte keinen signifikanten Unterschied zwischen den Bedingungen mit und ohne Fahraufgabe, $F(1, 14)=0.05$ $p=.81$. Die Genauigkeit der Aufgabenbearbei-

tung zeigte sich somit als unabhängig von der Bearbeitung einer parallelen Fahraufgabe.

8.4.3.2 CTT

Bei der Auswertung der Daten für die Leistung in der Critical Tracking Task (CTT) wurden sowohl die Daten für die Aktivität als auch die Daten für die mittlere absolute Abweichung von der Optimallinie betrachtet. Eine Person musste aus technischen Gründen bei der Datenaufzeichnung aus der Analyse ausgeschlossen werden.

Aktivität

Für die Bearbeitungsaktivität wurde die ursprüngliche Aufzeichnungsrate von 50 Hz auf 1 Hz reduziert. Dazu wurde analog zu der Analyse der SURT Daten, zu jeder Sekunde rund um den Zeitpunkt des Spurwechsels in der Fahraufgabe, die Wahrscheinlichkeit einer Eingabe zu einer bestimmten Sekunde in der Nebenaufgabe bestimmt (in Prozent). Die so gewonnenen Daten wurden daraufhin über alle 18 Spurwechsel gemittelt. Für die Eingabewahrscheinlichkeiten ergab sich in der CTT Aufgabe bei gleichzeitiger Fahraufgabe das in Abbildung 43 dargestellte Bild.

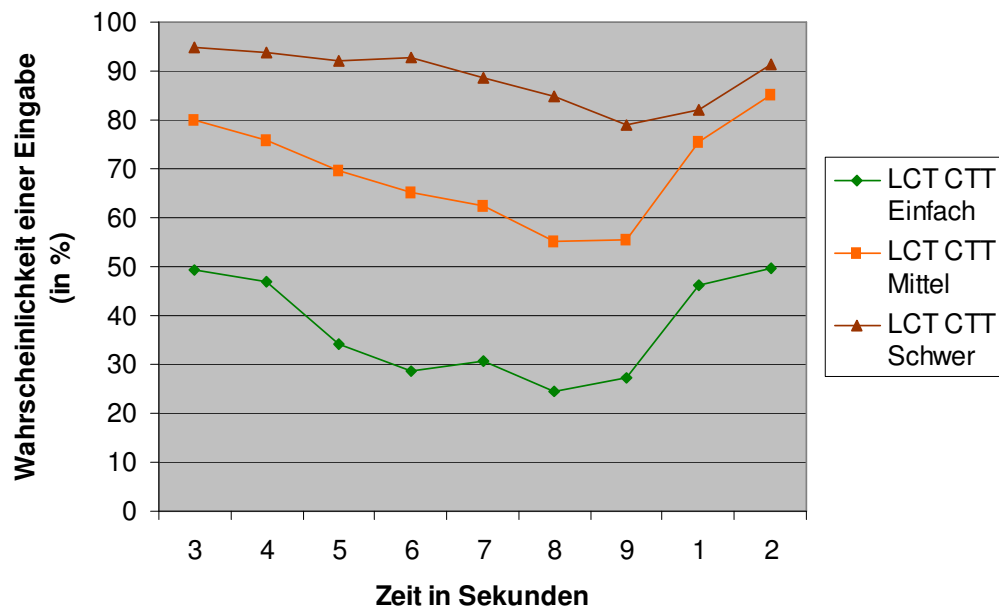


Abbildung 43: Die Abbildung zeigt die prozentuale Wahrscheinlichkeit einer Eingabe unter den drei Schwierigkeiten des CTT rund um den Zeitpunkt eines Spurwechsels bei zusätzlicher Fahraufgabe (Spurwechsel zwischen Sekunde 9 und 1).

Deutlich zu sehen sind hier die Unterschiede zwischen den drei Bedingungen einfach, mittel und schwer sowie der Abfall in der Aktivität vor dem Spurwechsel (Sekunde 9).

Für die Aktivität in der CTT Aufgabe bei zusätzlicher Fahraufgabe zeigte sich in einer ANOVA mit Messwiederholung für die Faktoren Aufgabenschwierigkeit und Zeitpunkt ein signifikanter Unterschied zwischen den drei Schwierigkeitsstufen einfach, mittel und schwer ($F(2, 30)=454.39, p<.001$). Zudem zeigte sich ein signifikanter Haupteffekt für den Zeitpunkt, $F(8, 120)=16.63, p<.001$.

Ein LSD Post Hoc Test zeigte für die einfache Bedingung einen signifikanten Anstieg ($p<.001$) in der Aktivität von Sekunde 9 zu Sekunde 1. Sowie ein stabiles Aktivitätsniveau für die Sekunden 1 bis 4 und einen signifikanten Abfall ($p<.001$) von Sekunde 4 bis Sekunde 7.

Für die mittlere Bedingung zeigte der Post Hoc Test ebenfalls einen signifikanten Anstieg ($p < .001$) von Sekunde 9 zu Sekunde 1 in der Aktivität. Hier zeigte sich auch noch ein signifikanter Anstieg in der Aktivität bis Sekunde 2. Nach dem Maxima der Aktivität in Sekunde 2 folgt in der mittleren Bedingung dann ein stufenweise signifikanter Abfall der Aktivität bis Sekunde 8 und dort eine Stagnation bis Sekunde 9.

Für die schwere Bedingung zeigte der PostHoc Test keinen signifikanten Anstieg in der Aktivität von Sekunde 9 zu Sekunde 1. Von Sekunde 2 bis Sekunde 8 bleibt das Aktivitätsniveau dann auf einem stabilen Niveau von ca. 90 %.

Aktivität CTT ohne zusätzliche Fahraufgabe

Für die Aktivität in der CTT ohne zusätzliche Fahraufgabe zeigte sich für die drei Schwierigkeitsstufen jeweils ein stabiles Niveau für die Wahrscheinlichkeit einer Eingabe gemittelt über Gesamtdauer der Aktivität. Hier ergab sich ein signifikanter Effekt für die Schwierigkeit der Aufgabe, $F(2, 30) = 53.523$, $p < .001$ (einfach 44,04 %, mittel 63,37 % und schwer 83,40 %).

Vergleich Aktivität CTT mit und ohne zusätzliche Fahraufgabe

In einer zusätzlichen ANOVA mit Messwiederholung über die interindividuellen Aktivitätsmittelwerte über die jeweiligen experimentellen Bedingungen konnte kein signifikanter Effekt für die Unterscheidung mit vs. ohne zusätzliche Fahraufgabe gefunden werden, $F(1, 15) = 0.12$, $p = .72$.

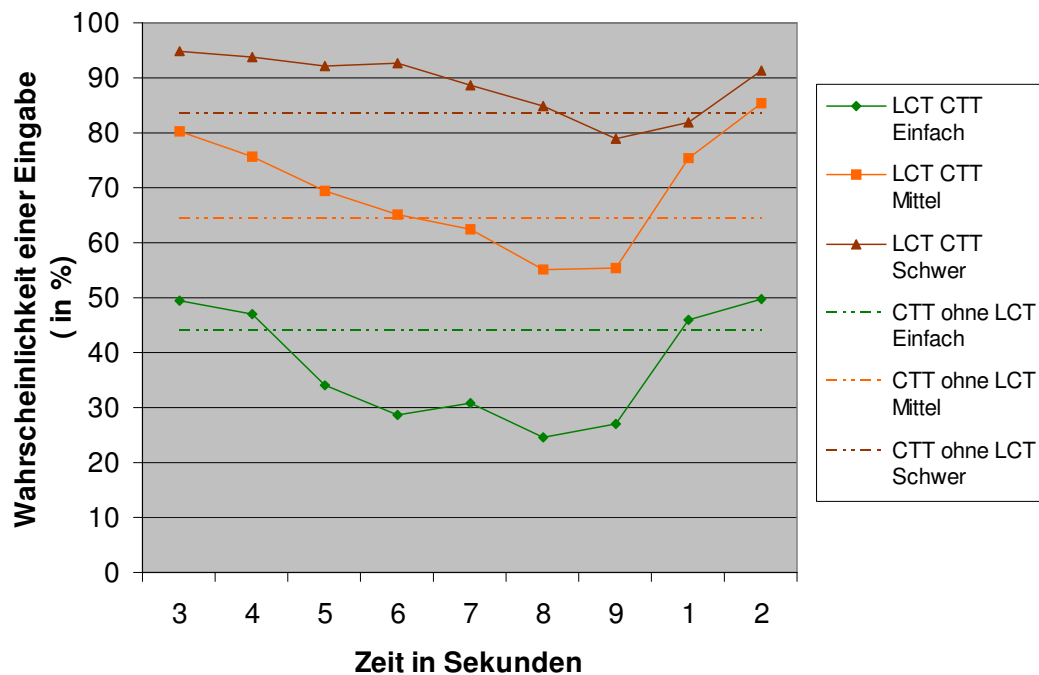


Abbildung 44: Die Abbildung zeigt die Aktivitätsverläufe mit und ohne zusätzliche Fahraufgabe unter den drei Schwierigkeitsstufen der CTT.

Abweichung

Auch für die Auswertung der Abweichungen in der CTT wurde die ursprüngliche Aufzeichnungsrate von 50 Hz auf 1 Hz reduziert. Daraufhin wurde die Abweichung des Balkens von der Bildschirmmitte im CTT zu jeder Sekunde rund um den Zeitpunkt des Spurwechsels bestimmt. Die so gewonnenen Daten wurden daraufhin über alle 18 Spurwechsel gemittelt.

Abbildung 45 zeigt den Verlauf der Abweichungen unter den drei Schwierigkeitsstufen des CTT bei denen gleichzeitig eine Fahraufgabe bearbeitet werden musste. Die Abweichungswerte können dabei von 0 bis 100 variieren.

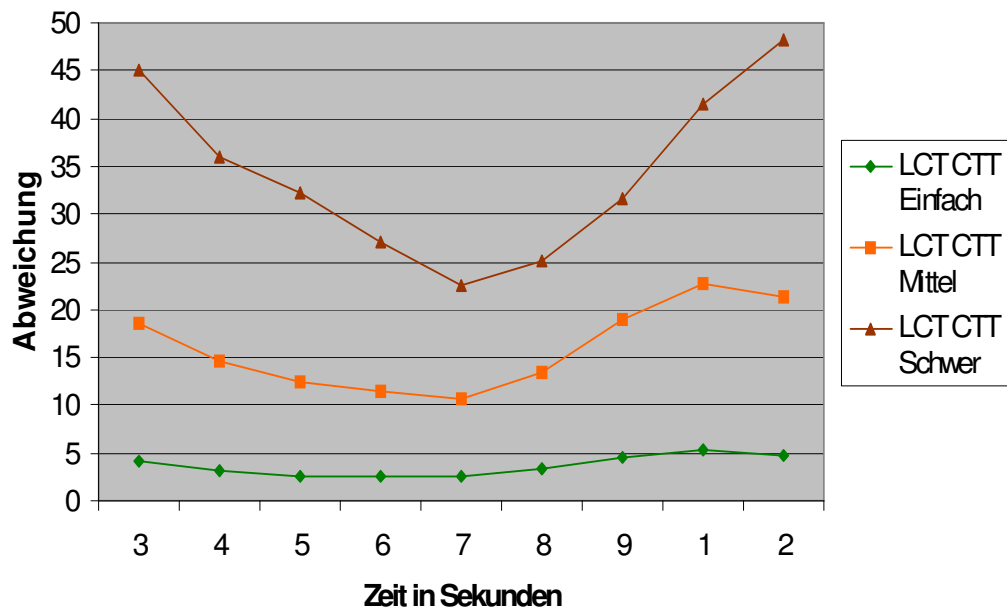


Abbildung 45: CTT. Die Abbildung zeigt die Abweichung der Ziellinie von der Mitte des Bildschirms getrennt für die drei Schwierigkeitsstufen über den Zeitraum von 9 Sekunden (das Schild wird zwischen Sekunde 9 und 1 passiert).

Für die Abweichung in der CTT Aufgabe bei zusätzlicher Fahraufgabe zeigte sich in einer ANOVA mit Messwiederholung für den Faktor Aufgabenschwierigkeit und Zeitpunkt ein signifikanter Unterschied zwischen den drei Schwierigkeitsstufen einfach, mittel und schwer ($F(2, 30)=45.09$, $p<.001$). Zudem zeigte sich ein signifikanter Haupteffekt für den Zeitpunkt, $F(8, 120)=44.44$, $p<.001$.

Ein LSD Post Hoc Test zeigte für die schwierige Bedingung des CTT einen signifikanten ($p<.01$) Anstieg in der Abweichung zwischen Sekunde 8 und 9 sowie zwischen Sekunde 9 und 1 sowie 1 und 2. Dabei zeigte sich aber auch ein stabiles Niveau in der Abweichung für die Sekunden 6 bis 8.

Für die mittlere Bedingung des CTT zeigte sich ein signifikanter Anstieg ($p<.001$) in der Abweichung zwischen Sekunde 8 und 9 sowie zwischen Sekunde 9 und 1. Ein stabiles Niveau zeigt sich hier für die Sekunden 4 bis 8.

Für die einfache Bedingung ergaben sich keinerlei signifikante Unterschiede über den zeitlichen Verlauf in der Abweichung.

Für die Abweichung in der CTT ohne zusätzliche Fahraufgabe zeigte sich ein Anstieg für die drei Schwierigkeitsstufen gemittelt über Gesamtdauer der Aktivität. Hier ergab sich ein signifikanter Haupteffekt für die Schwierigkeit der Aufgabe, $F(2, 30)=8.89$, $p<.001$ (einfach 2.18, mittel 3.94 und schwer 10.70).

Eine ANOVA mit Messwiederholung über die individuellen Abweichungsmittelwerte über die jeweiligen experimentellen Bedingungen sowie den Bedingungen mit vs. ohne zusätzliche Fahraufgabe zeigte einen signifikanten Anstieg in der Abweichung bei zusätzlicher Fahraufgabe über die drei Schwierigkeitsstufen hinweg, $F(1, 15)=79.83$, $p<.001$. Zusätzlich zeigte sich eine signifikante Wechselwirkung zwischen den beiden Bedingungen mit vs. ohne zusätzliche Fahraufgabe und den drei Schwierigkeitsstufen, $F(2, 30)=31.59$, $p<.001$.

Ein LSD Post Hoc Test zeigte einen signifikanten Anstieg in der Abweichung zwischen den Bedingungen mit und ohne Fahraufgabe nur für die Bedingungen mittel und schwer. Die mittleren Abweichungen in der einfachen Bedingung des CTT unterschieden sich dabei nicht zwischen den Bedingungen mit und ohne Nebenaufgabe ($p=.45$).

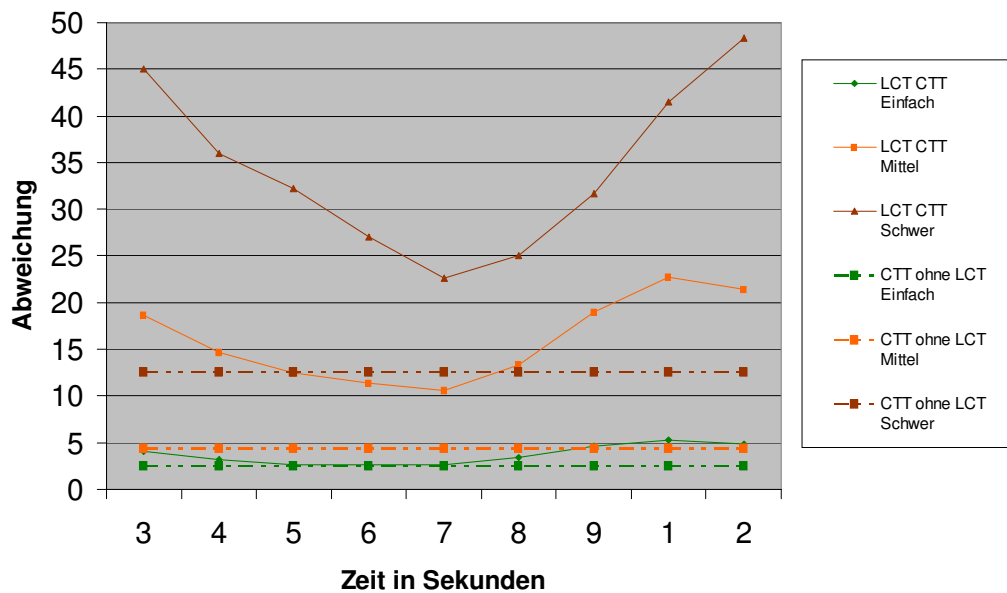


Abbildung 46: Die Abbildung zeigt die Abweichung der Ziellinie von der Mitte des Bildschirms getrennt für die drei Schwierigkeitsstufen mit und ohne LCT (das Schild wird zwischen Sekunde 9 und 1 passiert).

8.4.4 Subjektive Daten

Für die Erfassung der subjektiv erlebten Beanspruchung wurde in dem vorliegenden Experiment zusätzlich der NASA TLX verwendet. Die Daten wurden jedoch aus Zeitgründen nur für die Bedingungen SURT einfach und schwer und CTT einfach und schwer jeweils unter zusätzlicher Fahraufgabe erhoben. Die hier berichteten subjektiven Ratings wurden interindividuell standardisiert.

Getrennte Analysen für die Dimensionen körperliche Anstrengung, Zeitdruck sowie der Frage nach dem Wohlbefinden konnten keine signifikanten Unterschiede zwischen den beiden Aufgaben SURT und CTT zeigen. Jedoch zeigte sich auf jeder der drei Dimensionen ein signifikanter Effekt für die Schwierigkeit der Aufgaben (einfach vs. schwer). Mit einer erhöhten Schwierigkeit der Aufgabe gingen somit eine erhöhte körperliche Anstrengung, empfundener Zeitdruck und eine Verringerung des Wohlbefindens einher.

Für die Dimension der Einschätzung wie gut die Aufgabe erfüllt werden konnte, zeigte sich ein signifikanter Effekt für die Schwierigkeit der Aufgabe, $F(1, 17)=25.52$, $p<.001$, sowie eine signifikante Interaktion zwischen Schwierigkeit der Aufgabe und Aufgabentyp, $F(1, 17)=25.52$, $p<.001$. Unter der leichten Bedingung wurde die subjektiv empfundene Leistung im CTT als besser eingeschätzt als im SURT. Im Gegensatz dazu wurde unter der schwierigen Bedingung die Aufgabenerfüllung unter SURT als besser eingeschätzt als unter CTT.

Auf der Dimension der empfundenen geistigen Aktivität bei der Bearbeitung der Aufgaben zeigte sich ein signifikanter Effekt für den Faktor Aufgabe, $F(1, 17)=11.21$, $p<.01$, sowie für den Faktor Schwierigkeit, $F(1, 17)=37.39$, $p<.001$. Hier wurde die geistige Aktivität unter der Bearbeitung der SURT als höher eingeschätzt als unter der Bearbeitung der CTT. Eine schwierige Aufgabe erforderte in beiden Aufgaben mehr geistige Aktivität als eine einfache Aufgabe.

Für die Dimension der erlebten Anstrengung allgemein zeigte sich ein signifikanter Effekt für den Faktor Aufgabe, $F(1, 17)=8.95$, $p<.01$, sowie für den Faktor Schwierigkeit, $F(1, 17)=34.88$, $p<.001$. Auch unter dieser Dimension wurde die Bearbeitung des SURT als allgemein anstrengender empfunden als die des CTT. Sowie eine schwierige Aufgabe als anstrengender als eine einfache Aufgabe.

8.5 Diskussion

Bereits in Experiment 2 konnte gezeigt werden, dass es mittels des ICA möglich ist neben der absoluten Höhe der Beanspruchung in einer Fahrsituation auch deren zeitlichen Verlauf während der Bearbeitung einer Aufgabe abzubilden. So zeigte sich in Experiment 2 bei der Bearbeitung einer Fahraufgabe, bei der entsprechend gegebener Instruktionen die Fahrspur so schnell aber auch so präzise wie möglich gewechselt werden musste, ein signifikanter Anstieg zu den jeweiligen Spurwechseln während der Fahrt.

In dem vorliegenden Experiment 3 sollte nun zunächst eine klare Replikation dieser in Experiment 2 gefundenen Ergebnisse erreicht werden. Darüber hinaus sollte genauer ermittelt werden, welche Komponenten der Situation und somit der auf den Fahrer wirkenden Beanspruchung sich in den ICA Daten widerspiegeln. Dafür sollte zusätzlich zu einer geeigneten experimentellen Manipulation der Situation eine Trennung der unterschiedlichen Einflussvariablen auf die Beanspruchung des Fahrers während des Spurwechsels erreicht werden.

Zudem sollte in dem vorliegenden Experiment auch überprüft werden, inwieweit es möglich ist mittels des ICA sowie weiteren geeigneten Aktivitätsmaßen die bereits in Experiment 2 postulierten Strategien der Beanspruchungssteuerung bei der Bearbeitung von Nebenaufgabe zusätzlich zur Fahraufgabe abzubilden und zu identifizieren. Entsprechend der Diskussion aus Experiment 2 sollte an dieser Stelle auch überprüft werden, inwieweit unterschiedliche Aufgabencharakteristika ein solches strategisches Beanspruchungsverhalten beeinflussen können.

Eine detaillierte zeitliche Analyse der Situationen rund um die kritischen Fahrmanöver sollte in dem vorliegenden Experiment zusätzlich auch eine erhöhte zeitliche Genauigkeit bei der Interpretation der gefundenen Ergebnisse ermöglichen.

Komponenten der Beanspruchung

Zunächst zeigte die Analyse der auftretenden ICA Daten unter der visuellen Standardinstruktionsbedingung (visuell Baseline) einen zeitlich begrenzten Anstieg in den ICA Werten rund um die kritischen Situationen des Spurwechsels. Dies ist konform zu den bereits in Experiment 2 berichteten Befunden und zeigt, dass es möglich ist, mittels des ICA solche beanspruchenden Situationen im Verlauf einer Fahraufgabe zuverlässig zu identifizieren.

Um den Einfluss der reinen visuellen Wahrnehmung von Instruktionen in Form von Schildern auf den ICA zu identifizieren, wurde der LCT in der vorliegenden Studie nicht nur unter einer visuellen Instruktionsbedingung durchgeführt (Schilder geben die Richtung des Spurwechsels an), sondern wie bereits in Experiment 2 geschehen, auch unter einer auditiven Instruktionsmodalität (Schilder wurden gelöscht und die Instruktion die Spur zu wechseln auditiv dargeboten). Eine genauere zeitliche Analyse der Situation vor, während und nach Passieren des Schildes zeigte hier keinerlei Unterschiede in den Verläufen zwischen einer visuellen und einer auditiven Instruktionsmodalität. Unter beiden Instruktionsmodalitäten zeigte sich hier ein klarer Anstieg rund um die Situation des Spurwechsels. Die vorliegenden Ergebnisse unterstützen somit die bereits in Experiment 2 geführte Argumentation, dass die hier berichteten zeitlich sehr begrenzten Anstiege in der gemessenen Beanspruchung rund um die kritischen Spurwechsel nicht allein auf modalitätsspezifisch perzeptuelle Prozesse zurückzuführen sind.

Da davon ausgegangen werden soll, dass die während der Spurwechsel gemessenen Veränderungen in den ICA Daten auf Unterschiede in der ganzheitlichen Beanspruchung des Fahrers zurückzuführen sind und nicht etwa durch eine Veränderung des visuellen Eindrucks, oder etwa einer motorischen Bewegung (Lenkbewegung) an sich, wurde in dem vorliegenden Experiment eine weitere Bedingung mit einer freien Fahrt durchgeführt. Hier sollten die Probanden nach freiem Ermessen die Spur wechseln, ohne dass dies durch eine Instruktion angezeigt worden war (Schilder wurden gelöscht). Hier zeigte sich in den ICA Daten

keinerlei Effekt zum Zeitpunkt des Spurwechsels im Vergleich zu den Zeitpunkten davor und danach. Ein solches Ergebnis zeigt deutlich, dass die unter den weiteren Bedingungen beobachteten kurzen und präzisen Anstiege in den ICA Daten nicht allein durch die Perzeption eines sich verändernden visuellen Eindrucks oder allein der Durchführung einer motorischen Handlung hervorgerufen werden können. Vielmehr deuten die Ergebnisse darauf hin, dass die gemessenen Anstiege in den ICA Daten rund um den jeweiligen Spurwechsel eine willentliche Umsetzung einer expliziten Instruktion in eine Handlung widerspiegeln.

Zusätzlich wurde in dem vorliegenden Experiment untersucht, inwieweit eine semantische Verarbeitung einer Instruktion Einfluss auf den Verlauf der gemessenen ICA Daten hat. Um eine Trennung zwischen dem Einfluss einer semantischen Verarbeitung der Instruktionen und die intentionale Umsetzung und Initiierung einer Handlung zu erreichen, wurde in dem vorliegenden Experiment die Bedingung auditiv-imperativ eingeführt. Hier bekamen die Probanden bereits 4 Sekunden vor dem eigentlichen Spurwechsel die semantische Information auf welche Spur gewechselt werden sollte in Form eines auditiven Kommandos. Das Besondere an dieser Bedingung jedoch war, dass die Probanden instruiert waren zu diesem Zeitpunkt lediglich die Instruktion zu verstehen, aber noch nicht in eine Fahrhandlung umzusetzen. Erst 2 Sekunden vor dem Schild bekamen die Fahrer mittels eines imperativen akustischen Signals die Instruktion das soeben gehörte Kommando in eine entsprechende Fahrhandlung umzusetzen. Die Ergebnisse zeigten auch hier einen klaren Anstieg in den ICA Daten zum Zeitpunkt der Verarbeitung der semantischen Instruktion. Zusätzlich zeigte sich zum Zeitpunkt des eigentlichen Spurwechsels ein leicht verringertes Niveau in den ICA Daten unter der Bedingung auditiv-imperativ im Vergleich zu einer normalen auditiven Instruktionsbedingung. Dies ist damit zu erklären, dass hier die Gesamtbeanspruchung durch das Fehlen einer semantischen Verarbeitung und Verständnis der Instruktion reduziert wird.

Zusammenfassend lassen sich an dieser Stelle in Bezug auf den Nachweis von Beanspruchungseffekten in einer simulierten Fahraufgabe mittels des ICA folgende Ergebnisse festhalten: Mittels des ICA können zuverlässig Situationen mit einer erhöhten Beanspruchung für den Fahrer in einer simulierten Fahrsituation zeitlich präzise identifiziert und analysiert werden. In dem vorliegenden Experiment konnte gezeigt werden, dass die hier beschriebenen Effekte weder allein aufgrund modalitätsspezifischer perzeptueller Prozesse bei der Instruktion, noch aufgrund visueller Veränderungen in der Fahraufgabe oder allein einer motorischen Handlung zustande gekommen sein können. Ferner konnte gezeigt werden, dass die so abgebildeten Effekte in den ICA Daten direkt mit semantischen Verarbeitungsprozessen assoziiert werden können und somit einer willentlichen Umsetzung einer expliziten Instruktion zugeschrieben werden können.

Strategische Effekte

Eine zusätzliche Fragestellung, die in dem vorliegenden Experiment adressiert werden sollte, war die des Nachweises auftretender strategischer Beanspruchungseffekte bei der Durchführung unterschiedlicher Nebenaufgaben zusätzlich zur Fahraufgabe.

Mit dem Begriff der kompensativen Strategien sollte dabei in dem vorliegenden Experiment insbesondere der bereits unter Experiment 2 diskutierte Befund adressiert werden, dass es insbesondere unter einer Fahraufgabe mit zusätzlicher Bearbeitung einer Nebenaufgabe rund um die kritischen Situationen zu einer kurzfristigen Verringerung oder Konstanthaltung anstatt zu einer Erhöhung der gemessenen Beanspruchung kommen kann. Dieser Effekt kann damit erklärt werden, dass in diesen kritischen Situationen der Fahrer von einer dual-task Situation, also einer Situation bei der zwei Aufgaben parallel bearbeitet werden, zu einer single-task Situation wechselt, bei der nur noch die Fahraufgabe durchgeführt und die Nebenaufgabe ignoriert wird. Das Ergebnis ist ein entsprechender kompensativer, strategischer Effekt, bei dem der Fahrer proaktiv seine momentane Bean-

spruchung in Hinblick auf eine kritische Situation in der Fahraufgabe anpasst, um so in der kritischen Situation genügend Kapazitäten freizustellen und so die Situation sicher und effizient zu lösen. Dieses wird möglich, indem der Fahrer die vor ihm liegende Situation - und somit die ihm für die Bearbeitung einer Aufgabe zu Verfügung stehende Zeit einschätzen und für die Bearbeitung einer zusätzlichen Aufgabe nutzen kann. Dies ist insbesondere aufgrund der Eigenschaften der in dem vorliegenden Experiment verwendeten Fahraufgabe möglich.

In dem vorliegenden Experiment sollte nun dieser Effekt zunächst noch einmal mit einer deutlich erhöhten Genauigkeit untersucht werden, um so ein genaueres Verständnis dafür zu generieren, welche Strategien der Aufgabenbearbeitung von den Fahrern eingesetzt werden und inwieweit es möglich ist mittels des ICA in diese Prozesse Einsicht zu erreichen.

Variiert wurden in dem vorliegenden Experiment sowohl die Schwierigkeit als auch die Charakteristik der Nebenaufgabe. Der Surrogate Reference Task (SURT) ermöglicht aufgrund der Struktur der Aufgabe zu jedem beliebigen Zeitpunkt während der Bearbeitung eine Unterbrechung und eine Fortsetzung der Aufgabe, ohne dass dies Leistungseinbußen zur Folge hat oder Informationen verloren gehen. Ein solches Bearbeitungscharakteristikum sollte auch bei einer realistischen Aufgabe wie der Eingabe eines Navigationsziels mittels des IDrive Controllers auftreten. Auch hier kann die Bearbeitung der Aufgabe beliebig unterbrochen und wieder fortgesetzt werden, ohne dass hier Informationen für den Fahrer verloren gehen. Der Critical Tracking Task (CTT) hingegen ist eine Aufgabe, die eine wirklich kontinuierliche Zuwendung bei der Bearbeitung erforderlich macht. Eine Unterbrechung der Bearbeitung der Aufgabe hat hier einen deutlichen Leistungsabfall und Einbußen in der Performance zur Folge. Folgt man der Diskussion einer bewussten strategischen Bearbeitung von Aufgaben, so ist zu erwarten, dass insbesondere unter der Bearbeitung der SURT kompensative Bearbeitungsstrategien gezeigt werden können, wohingegen diese unter der CTT deutlich geringer ausfallen müssten.

Die Ergebnisse des vorliegenden Experiments zeigten entsprechend der Erwartung einen deutlichen Abfall in der Bearbeitungswahrscheinlichkeit des SURT kurz, vor- und während dem Passieren eines Schildes. Die Wahrscheinlichkeit eine Eingabe zu tätigen erreicht hier ein Minimum kurz vor dem Spurwechsel, unabhängig von der Schwierigkeit der Aufgabe. Zusätzlich ist zu erwähnen, dass hier bereits vor dem eigentlichen Erscheinen der Information auf den Schildern und somit der Verfügbarkeit der Instruktion, eine Reduktion der Aktivität in der Nebenaufgabe zu beobachten ist. Dies ist damit zu begründen, dass der Fahrer hier aktiv eine heuristische Einschätzung der verbleibenden Zeit bis zum nächsten Spurwechsel vornimmt und mit dieser Abschätzung bereits proaktiv sein Verhalten beeinflusst.

Für die Bearbeitung des CTT zeigt sich ein solches Muster in der Bearbeitungswahrscheinlichkeit in einer deutlich reduzierten Form. Auch hier kommt es vor dem Passieren der Schilder zu einer leichten Reduktion der Eingabeaktivität. Allerdings in einer deutlich geringeren Ausprägung. Unter der schwierigen Bedingung ist keine kompensative Reaktion in Hinblick auf die Vorbereitung des Spurwechsels noch vor dem Erscheinen der Instruktion zu beobachten.

In Bezug auf die gemessenen ICA Daten ergab sich ebenfalls entsprechend der Daten der Nebenaufgabe ein erwartungskonformes Bild in den Ergebnissen. Eine detaillierte Analyse der Verläufe der ICA Werte auf Ebene der Spurwechsel zeigte einen deutlichen Abfall zum Zeitpunkt des Spurwechsels unter Bearbeitung der SURT. Ebenso zeigte sich bei der Bearbeitung der Navigationszieleingabe ein solcher Abfall in den ICA Werten kurz vor dem Passieren des Schildes. Insgesamt zeigte sich, dass unter der Bearbeitung dieser beiden Aufgabentypen das Niveau der gemessenen Beanspruchung zum Zeitpunkt des Spurwechsels auf demselben Niveau lag wie unter der Bedingung in der keine Nebenaufgabe zusätzlich zur Fahraufgabe bearbeitet wurde. Ein solches Befundmuster steht im Einklang mit den bereits zuvor diskutierten Ergebnissen aus den Aktivitätsverläufen. Aufgrund der Charakteristika der SURT Aufgabe verändern die Probanden ihr Verhalten

von einer dual-task zu einer single-task Situation und wirken somit einer erneuten Steigerung der Beanspruchung durch das anstehende Fahrmanöver entgegen. Die gefundenen Ergebnisse unter der Bearbeitung einer realistischen Nebenaufgabe der Navigationszieleingabe, decken sich mit den für die SURT gefundenen Ergebnissen.

Bei der Bearbeitung der CTT hingegen zeigte sich zum Zeitpunkt des Spurwechsels ein deutlicher Anstieg in den gemessenen ICA Daten. Hier ist ein deutlicher Unterschied im Vergleich zu der Bedingung ohne zusätzliche Nebenaufgabe zum Zeitpunkt des Spurwechsels zu beobachten. Wie bereits für die Aktivitätsdaten diskutiert, erschwert hier die Eigenschaft der Aufgabe, insbesondere in der schwierigen Bedingung des CTT, ein strategisches Vorgehen wie unter der SURT Aufgabe.

Die Ergebnisse in Bezug auf die subjektiven Angaben der Probanden in diesem Versuch werfen für die Interpretation der Ergebnisse jedoch einige interessante Fragen auf. Für die Einschätzung der für die Bearbeitung der Aufgaben notwendigen körperlichen Anstrengung, der entstehende Zeitdruck sowie der Frage nach dem Wohlbefinden (Frustration bei der Bearbeitung) zeigten sich überraschenderweise keine Unterschiede zwischen den beiden Aufgabentypen. Wohingegen die subjektiv wahrgenommene geistige Aktivität sowie die Anstrengung insgesamt bei der Bearbeitung der Aufgaben für die SURT als höher eingeschätzt wurde als unter der Bearbeitung der CTT Aufgabe. Gleichzeitig jedoch ist die subjektive Einschätzung wie gut eine Aufgabe erfüllt werden konnte, insbesondere unter den schwierigen Aufgaben des SURT, höher als unter der schwierigen Aufgabe des CTT. An dieser Stelle kann vermutet werden, dass hier die Anwendung einer aktiven, kompensativen Strategie, wie durch die Aktivitäts- sowie Beanspruchungsdaten gezeigt, durch das aktive Zutun des Probanden zwar insgesamt zu einer erhöhten subjektiv erlebten Anstrengung führt, die tatsächlich resultierende Beanspruchung jedoch geringer ist.

9 Experiment 4: Einsatz der Methode für die Bewertung und Vergleich zweier realer Bediensysteme sowie Interaktionsmodalitäten

In dem vorliegenden vierten Experiment dieser Arbeit sollte nun nach der vorangegangenen grundlegenden methodischen Validierung der Methode des ICA der Schritt hin zu einer tatsächlichen Anwendbarkeit der Methode bei der Bewertung und Entwicklung neuer Anzeige- und Bediensysteme im Kontext einer Usability Untersuchung im automotiven Kontext vollzogen werden.

Zum Einsatz kamen hier zwei Fahrerinformationssysteme mit unterschiedlichen Ansätzen in Bezug auf die Informationsstrukturierung sowie -Darstellung. Zudem wurde die für den Benutzer zu Verfügung stehende Eingabemodalität variiert.

Bei den beiden Systemen handelte es sich einmal um das Vorgängersystem des in der aktuellen BMW 5er Serie eingesetzten Fahrerinformationssystems sowie einer konzeptuellen Neuentwicklung für ein Fahrerinformationssystem im Prototypenstatus mit einer neuartigen Informationsdarstellung und Strukturierung. Die beiden zu Verfügung stehenden Eingabemodalitäten für alphanumerische Zeichen bestanden bei der hier beschriebenen Studie in beiden Systemen einmal in der Eingabe durch die visuelle Auswahl von Zeichen und durch eine Handschrifterkennung über ein Touchpad. Beide hier betrachteten Systeme stellen unterschiedliche Anforderungen an den Benutzer und sind unterschiedlich gut an die jeweiligen Eingabemodalitäten angepasst. Dadurch ergaben sich für einen solchen experimentellen Vergleich unterschiedliche Erwartungen hinsichtlich der auftretenden Beanspruchungen bei der Verwendung der Systeme. Um die Einsetzbarkeit der beiden Systeme im fahrrelevanten Kontext zu überprüfen, wurde auch in dem vorliegenden Experiment der Lane Change Task (LCT) als zusätzliche simulierte Fahraufgabe verwendet.

9.1 Hintergrund

Während die vorangegangenen in dieser Arbeit diskutierten Experimente in erster Linie einer eher grundsätzlichen Validierung der Methode des „Index of Cognitive Activity“ (ICA) gedient haben, sollte mit dem vorliegenden Experiment nun der Schritt hin zu einer tatsächlichen Anwendung bei dem Vergleich und der Bewertung unterschiedlicher Fahrerinformationssysteme sowie Eingabemodalitäten erreicht werden. In dem vorliegenden Experiment kamen dabei zwei unterschiedliche Fahrerinformationssysteme zum Einsatz. Zusätzlich wurden hier zwei Eingabemodalitäten für alphanumerische Zeichen variiert. Beide Systeme sowie Eingabemodalitäten wurden zudem mit als auch ohne zusätzliche Fahraufgabe (LCT) untersucht.

Bei den beiden Systemen handelte es sich um das Vorgängersystem des in der aktuellen BMW 5er Serie verbauten Fahrerinformationssystem (IDrive) sowie um eine Neuentwicklung für ein Fahrerinformationssystem im Prototypenstatus mit einer neuartigen Informationsdarstellung und Strukturierung, dem System „Hybride Mensch Maschine Interaktion“ (HMMI) (Graf et al., 2007). Die beiden in dem vorliegenden Experiment realisierten Eingabemodalitäten waren einmal die Eingabe über ein Spellersystem bei der Zahlen visuell ausgewählt werden mussten und einmal die Eingabe von alphanumerischen Zeichen über eine Handschrifterkennung via Touchpad.

Für die Bewertung der beiden Systeme sowie Eingabemodalitäten sollte in der vorliegenden Studie überprüft werden, inwieweit es mit der Methode des ICA als physiologisches Maß für die Beanspruchung möglich ist, hier ein – für ein bestimmtes System sowie Aufgabe und Eingabemodalität - charakteristisches Beanspruchungsprofil kontinuierlich über die Zeit hinweg abzubilden. Die Bestimmung eines solchen Beanspruchungsprofils über die Zeit bietet bei der Entwicklung und Konzeptualisierung neuer Anzeige- und Bedienkonzepte die einzigartige

Möglichkeit der Identifikation einzelner kritischer Bedienhandlungen und deren Optimierung. Bedienschritte mit einer erhöhten Beanspruchung für den Fahrer können identifiziert und optimiert werden und dies mit einer Genauigkeit, welche mit herkömmlichen Usability Maßen, wie etwa Bedienzeiten oder subjektive Einschätzungen nicht zu erreichen wäre. Denn solche Maße bieten jeweils nur einen Überblick über ganze Abschnitte der Aufgabenbearbeitung, lassen jedoch keinen Einblick in deren Verlauf zu.

In dem vorliegenden Experiment sollte daher für die Erstellung eines solchen Beanspruchungsprofils über die Zeit ein entsprechendes methodisches Verfahren entwickelt und realisiert werden. Die Herausforderung besteht an dieser Stelle darin, den interindividuellen Varianzen bei der Dauer der Bearbeitung von Aufgaben Rechnung zu tragen, um so ein zeitlich einheitliches und somit vergleichbares Profil erzeugen zu können.

Neben der Betrachtung der über die Pupillometrie ermittelten Beanspruchungsdaten sollten in der vorliegenden Untersuchung auch klassische Usability Maße wie Fahrleistungsdaten, Bedienzeiten, Blickverhalten sowie subjektive Bewertungen berichtet werden. Allerdings soll an dieser Stelle auf eine ausführliche Diskussion dieser Maße aus Platzgründen und im Hinblick der Fokussierung der vorliegenden Arbeit auf die Methode der Pupillometrie verzichtet werden.

9.2 Fragestellung

Basierend auf den vorangegangenen Studien zur grundlegenden Validierung des ICA zum Nachweis von Beanspruchungseffekten sollte in dem vorliegenden Experiment gezeigt werden, dass der ICA auch im Kontext eines realen experimentellen Vergleiches zwischen verschiedenen Systemen sowie Usecases hinsichtlich der Bewertung von auftretenden Beanspruchungsverläufen einsetzbar ist.

Mittels des ICA sollte es möglich sein für ein bestimmtes System sowie Usecase ein stabiles und charakteristisches Beanspruchungsprofil über die Zeit abzubilden und vergleichen zu können.

Über ein solches Beanspruchungsprofil sollte es möglich sein systemcharakteristische Besonderheiten zu identifizieren und auf ihren Beanspruchungsgehalt hin zu überprüfen.

Ein so erzeugtes Beanspruchungsprofil ermöglicht es systemcharakteristische Besonderheiten hinsichtlich ihrer Anforderungen an den Benutzer zu identifizieren.

Darüber hinaus sollte es möglich sein, Interaktionen zwischen bestimmten Systemcharakteristika sowie weiteren Eigenschaften der Bediensituation aufdecken zu können. So zum Beispiel zwischen den Eigenschaften eines Systems und den verwendeten Eingabemodalitäten.

Mittels eines solchen Beanspruchungsprofils sollte es auch möglich sein, auftretenden Interaktionen zwischen Systemcharakteristika und verwendeter Eingabemodalitäten abzubilden.

9.3 Methode

In Experiment 4 wurde die gleiche Fahraufgabe (LCT) sowie das gleiche Eyetracking Gerät, wie bereits unter Experiment 2 und 3 beschrieben, verwendet. Auch dieser Versuch wurde an einem Laboraufbau im Usability Labor der BMW Group durchgeführt.

9.3.1 Versuchsaufbau

Der Versuchsaufbau ähnelt dem bereits in Experiment 2 und Experiment 3 beschriebenen Setting mit einem Standard LCT Aufbau und einem Logitech Lenk-
rad für die simulierte Fahraufgabe. Rechts neben dem Hauptmonitor mit der Fahr-
simulation befand sich ein zweiter Monitor in gleicher Entfernung zur
Versuchsperson auf dem die unterschiedlichen Systeme dargestellt wurden. Der
Controller zur Bedienung der beiden Systeme befand sich rechts zur Versuchsperson. Abbildung 47 zeigt den verwendeten Versuchsaufbau sowie den zur Erfas-
sung des ICA verwendeten Eyetracker EyeLink 2 von SR Research, Ltd.

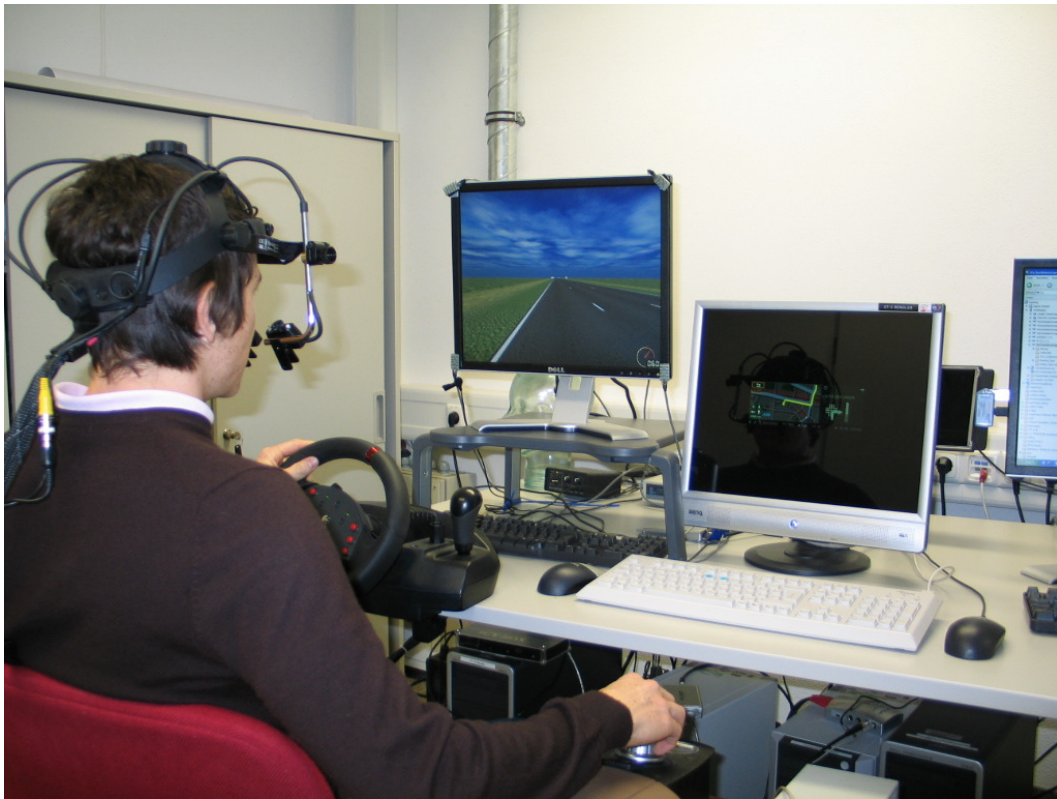


Abbildung 47: Die Abbildung zeigt einen Probanden bei der Bearbeitung einer Aufgabe während des Fahrens.

9.3.2 Verwendete Aufgaben und Systeme

9.3.2.1 Fahraufgabe

Für die Fahraufgabe wurde erneut die Software Lane Change Task (LCT) (Mattes 2003), wie bereits unter Experiment 2 und 3 beschrieben verwendet.

9.3.2.2 IDrive System

Bei dem hier verwendeten IDrive System handelte es sich um eine PC basierte Simulation des Vorgängersystems des in der aktuellen BMW 5er Reihe eingesetzten Fahrerinformationssystems, wie bereits unter Experiment 3 beschrieben und verwendet. Das Hauptmenü ist in die vier Bereiche Klima, Kommunikation, Na-

vigation und Entertainment unterteilt. die einzelnen Untermenüs sind durch das Schieben des Controllers in die entsprechende Richtung zu erreichen.



Abbildung 48: Die Abbildung zeigt das Hauptmenü des BMW iDrive Systems

Die einzelnen Untermenüs sind hierarchisch horizontal aufgebaut. Die Oberkategorien sind in der obersten Leiste aufgeführt. Die Kategorien darunter bieten entsprechende Unterfunktionen. Die Auswahl der einzelnen Untermenüs erfolgt durch Drehen, Kippen und Drücken. Die Eingabe alphanumerischer Zeichen erfolgt standardmäßig über eine horizontale Spellerleiste in der mittels Drehung des Controllers zwischen den einzelnen Buchstaben gewechselt werden kann.



Abbildung 49: Die Abbildung zeigt das Untermenü Kommunikation des BMW iDrive Systems.

Befindet sich der Benutzer in einem bestimmten Untermenü des IDrive Systems, so kann er auch nur auf die Funktionen und Inhalte dieses Untermenüs zugreifen. Um an Informationen anderer Untermenüs zu kommen muss er diese zuerst auswählen, um dann mit der Informationssuche beginnen zu können.

9.3.2.3 System HMMI

Das Konzept „Hybride Mensch Maschine Interaktion“ (HMMI) wurde bei der BMW Forschung und Technik GmbH als ein alternativer Ansatz für eine menübasierte Systemkonzeption entwickelt. Während das zuvor beschriebene IDrive System eine klassische Menüstruktur aufweist, welche Informationen hierarchisch innerhalb verschiedener Untermenüs strukturiert, arbeitet das Konzept HMMI mit einer Informationsstrukturierung, welche weitestgehend auf eine solche hierarchische Struktur der Information verzichtet. Die Idee hier ist es Quellen- und Themenabhängig Informationen ganzheitlich darzustellen und dem Nutzer mittels einer einheitlichen Suchfunktion zu Verfügung zu stellen. Der Nutzer benötigt keinerlei Vorwissen unter welcher Untermenüstruktur welche Information zu finden ist. Für eine detaillierter theoretische Diskussion eines solchen Suche-basierten Ansatzes in der Systemkonzeption und des hier beschriebenen Systems sei an dieser Stelle auf die Arbeiten nach Graf et al. (2008) verwiesen.

Das System HMMI verfügt über 2 Hauptbereiche. Im linken Teil des Displays befindet sich der Direktzugriffsbereich. Im Direktzugriffsbereich findet der Nutzer die am häufigsten verwendeten Funktionen und Einträge – diese Liste verändert sich also in Abhängigkeit des Nutzerverhaltens. Da diese Funktionalität in der vorliegenden Studie nicht im Fokus des Interesses stand, soll diese hier jedoch auch nicht weiter beschrieben werden.

Der zweite Hauptbereich des Systems stellt den allgemeinen Suchbereich dar (s. Abbildung 50).



Abbildung 50: Die Abbildung zeigt den Suchbereich des HMMI Systems.

Hier ist es möglich die Suche durch Eingabe von Buchstaben einzuschränken. Das System HMMI ist dabei von der grundlegenden Konzeption auf die Verwendung über eine alphanumerische Eingabe über Touch ausgelegt. Eine Visualisierung einer Buchstabenleiste wie unter dem System IDrive fehlt. Generell ist es hier möglich sowohl in der gesamten Datenbasis zu suchen als auch einen Filter für einzelne Kategorien von Inhalten zu setzen. Nach der Eingabe von Buchstaben zum Starten der Suche öffnet sich eine Liste mit entsprechenden Einträgen, welche über die Bedienung eines Controllers ausgewählt werden können.

Bei Auswahl eines Listeneintrags öffnet sich eine perspektivische Würfeldarstellung (s. Abbildung 51), die alle Optionen rund um den Eintrag enthält.



Abbildung 51: Perspektivische Würfel-Darstellung HMMI System.

In dem in Abbildung 51 dargestellten Beispiel enthält der Würfel alle Daten zum ausgewählten Kontakt. Der Eintrag kann hier bearbeitet, der Teilnehmer angerufen oder die Zielführung zu ihm gestartet werden. Im Würfel kann durch Drehen und Drücken des Controllers die entsprechende Funktion aufgerufen werden. Durch das Schieben des Controllers kann der Würfel rotiert werden, um weitere Optionen aufzurufen (bspw. ‚SMS schreiben‘ oder zu ‚Favoriten hinzufügen‘).

Die Besonderheiten des HMMI Konzepts lassen sich wie folgt zusammenfassen:

- Weitgehender Verzicht auf eine Menü-Hierarchie
- Die Suche nach Inhalten erstreckt sich standardmäßig auf die gesamte Datenbasis
- Daten- statt Geräteorientierung. Das heißt es werden inhaltlich Daten zu Verfügung gestellt unabhängig der dahinterstehenden technischen Systeme.
- Die Darstellung einzelner semantischer Datenobjekte sowie Relationen in visuell-räumlicher Form.

Die Bedienung des HMMI Konzeptes geschieht über den gleichen Controller (Dreh- Drücksteller) wie für das IDrive System. Dabei ist das HMMI System, anders als das System IDrive, grundsätzlich für die Verwendung über die Eingabe über Touch bei alphanumerischen Zeichen konzeptuell ausgelegt worden. So fehlt hier unter dieser Eingabemodalität eine Spellerleiste, wie unter dem System IDrive beschrieben, vollständig. Eine Option eine alphanumerische Eingabe mittels Speller zu erzeugen wurde erst für den vorliegenden Versuch realisiert.

9.3.3 Alphanumerische Eingabe

9.3.3.1 Verwendete Bedienelemente

Wie bereits zu Beginn erwähnt, wurden in der vorliegenden Studie die beiden Systeme IDrive sowie HMMI mit einer alphanumerischen Eingabe sowohl über ein Touchpad mit Handschrifterkennung als auch mit der Eingabe über Speller verglichen. Dafür wurde ein speziell für diese Untersuchung angefertigtes Bedienelement bzw. Controller verwendet, dem ein entsprechendes kapazitives Touchpad vorgelagert war. Bei dem hier dargestellten Bedienelement handelt es sich um die Standard Controller Einheit des IDrive Systems.

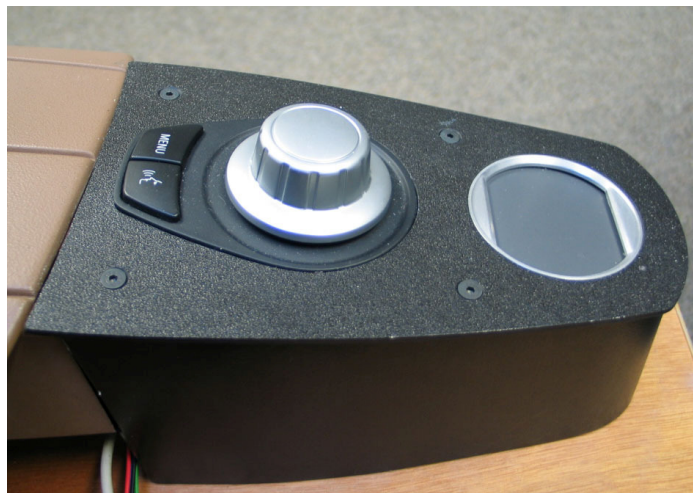


Abbildung 52: Die Abbildung zeigt den verwendeten Controller mit vorgelagertem kapazitiven Touchpad (graue Fläche).

9.3.3.2 Eingabe über Speller

Bei der alphanumerischen Eingabe über Speller geschieht die Eingabe von Ziffern und Buchstaben über die Auswahl in einer Liste von Zeichen. Der Nutzer bewegt den Fokus durch Drehen des Controllers auf den gewünschten Buchstaben und wählt diesen mittels drücken des Controllers aus. Die beiden in der vorliegenden Studie verwendeten Systeme unterscheiden sich dabei in der Darstellung der zu wählenden Zeichen nicht. Bei beiden Systemen ist die Anordnung der Zeichen

horizontal nebeneinander, wobei nicht immer alle Zeichen gleichzeitig sichtbar sind. Die jeweils nicht mehr verfügbaren Buchstaben (d.h. es ist kein entsprechender Eintrag in der Datenbank vorhanden) werden in allen drei Systemen ausgeblendet. Sobald ein Buchstabe ausgewählt wurde, erscheint dieser in der Statuszeile der Eingabe.

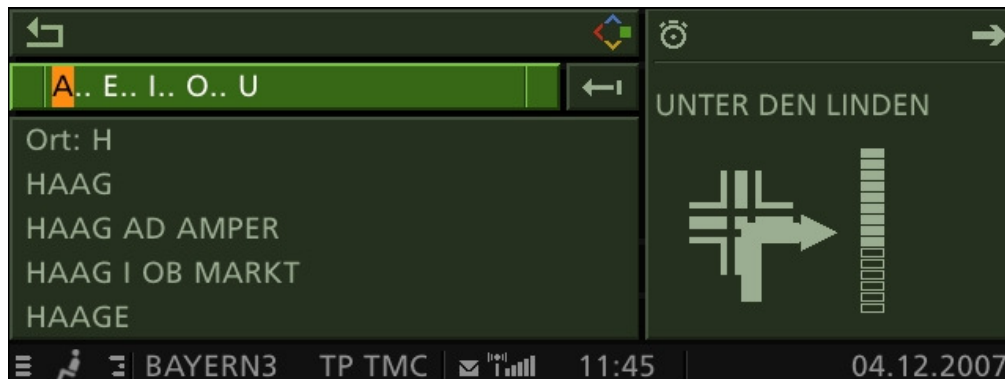


Abbildung 53: Die Abbildung zeigt das IDrive System bei der Verwendung einer alphanumerischen Eingabe über Speller. Der jeweils ausgewählte Buchstabe wird rot hervorgehoben.

Die Eingabe über Speller geschieht für das HMMI System ähnlich wie für das IDrive. Hier jedoch werden alle Buchstaben in der Spellerleiste auf einmal sichtbar.



Abbildung 54: Die Abbildung zeigt die Eingabe über Speller für das System HMMI.

9.3.3.3 Eingabe mittels Handschrifterkennung

Bei der alphanumerischen Eingabe über das Touchpad mussten die Probanden über Schreibschrift die jeweiligen Zahlen oder Buchstaben auf dem über dem Controller angebrachten kapazitiven Touchpad eingeben. Die jeweiligen Eingaben wurden mittels eines geeigneten Handschrifterkenners für Einzelbuchstaben identifiziert und wie bei der Spellerversion auch jeweils als Feedback in geschriebener Form in der Statuszeile dargestellt. Auf ein verbales Feedback seitens des Systems wurde hier gänzlich verzichtet. Kam es bei der Erkennung über das Touchpad zu einem Fehler so wurde dies dem Probanden durch eine kurze akustische Rückmeldung in Form eines kurzen Tonsignals signalisiert.

Bei der alphanumerischen Eingabe über Touch wurde für das IDrive System aus technischen Gründen die gleiche Eingabemaske wie unter der Spellerversion verwendet. Die Spellerleiste zusammen mit der Hervorhebung der einzelnen ausgewählten Buchstaben blieb hier auch bei der Eingabe über Handschrifterkennung die ganze Zeit sichtbar.

Bei der Eingabe über das Touchpad für das System HMMI fehlte die Spellerleiste vollständig. Das Feedback, welcher Buchstabe erkannt wurde, wurde über die Statuszeile gegeben.



Abbildung 55: Die Abbildung zeigt die Eingabe über Touch für das System HMMI.

9.3.4 Beschreibung der Bedienvorgänge

Da es sich bei der vorliegenden Untersuchung um einen vollständigen Systemvergleich zweier Systemen sowie zweier Eingabemodalitäten mit und ohne zusätzliche Fahraufgabe handelte, musste aufgrund einer zeitlichen Beschränkung eine Auswahl der von den Probanden durchgeführten Bedienvorgängen (Usecases) getroffen werden. Darüber hinaus sollten Bedienvorgänge zuvor geübt werden, um auftretende Varianz aufgrund von Lernvorgängen zu minimieren.

Nachfolgend sollen die in der vorliegenden Untersuchung verwendeten Aufgaben beschrieben werden. Die Auswahl der Usecases geschah dabei mit Blick auf die Vergleichbarkeit der beiden Systeme, dass heißt die Aufgaben mussten in den beiden Systemen gleichermaßen durchführbar sein. Die Aufgaben sollten dabei auch exemplarisch sowohl die Eingabe von Ziffern als auch von Buchstaben enthalten.

Navigationszieleingabe (Navi):

Hier sollten die Probanden ein Navigationsziel eingeben und die Zielführung starten. Der Usecase startet im Hauptmenü und endet mit dem Starten der Zielführung.

„Nutzen Sie das Navigationssystem und lassen Sie sich zur Paulusgasse in Hannover führen“

Telefonnummerneingabe (Tele):

Hier sollten die Probanden eine 6-stellige Telefonnummer wählen. Der Usecase startet im Hauptmenü und endet mit dem Wahlvorgang der vollständig eingegebenen Rufnummer.

„Bitte wählen Sie die Rufnummer 283576.“

9.3.5 Ablauf

Nach einer kurzen Einführung in den Gegenstand der aktuellen Untersuchung hatte jeder Proband zunächst einmal die Möglichkeit sich mit der Fahraufgabe vertraut zu machen. Anschließend wurde eine Baselinefahrt erhoben, wobei das Fahren mit dem Lane Change Task (LCT) so lange wiederholt wurde bis das Kriterium einer Abweichung in der Spur von < 1 erfüllt werden konnte.

Nach der Erhebung der Baseline Performance in der Fahraufgabe wurde das Eyetracking System mittels einer 9 Punkt Methode auf dem Hauptbildschirm kalibriert.

Anschließend wurden den Probanden vom Versuchsleiter die relevanten Bedienvorgänge in den jeweiligen Systemen sowie Eingabemodalitäten erläutert und die Probanden wurden aufgefordert diese ohne zusätzliche Fahraufgabe mindestens dreimal zu wiederholen bis ein sicherer Umgang mit den Systemen sowie Bedienvorgängen erreicht werden konnte.

Nach der Übungs- und Eingewöhnungsphase bearbeitete jeder Proband jeden der zwei Usecases sowohl mit als auch ohne zusätzliche Fahraufgabe, mit Speller sowie Touch als auch in jedem der beiden in diesem Versuch zum Einsatz kommenden Systeme. Jeder Proband führte die Aufgaben dabei jeweils zweimal durch. Den Usecase Zieleingabe im IDrive System zu Validierungszwecken sogar vier Mal. Vor der Durchführung jedes Usecases wurde erneut sichergestellt, dass der Proband sicher in dem Umgang mit dem jeweiligen System sowie der Aufgabenstellung war. Die Aufgabenabfolge wurde dabei interindividuell randomisiert dargeboten, wobei die Darbietung geblockt pro System erfolgte. Die Reihenfolge der Systeme wurde ebenfalls experimentell manipuliert.

Nach der Durchführung jeder experimentellen Bedingung wurden die Probanden befragt wie anstrengend sie die gerade bearbeitete Aufgabe (auf einer Skala von 1

bis 10) einschätzen würden. Am Ende des Versuches wurde zudem ein demographischer Fragebogen erhoben.

Die Gesamtdauer der Untersuchung betrug pro Versuchsperson ca. 120 min. bisher immer in Stunden angegeben, würde daher auch jetzt 2h schreiben.

9.3.6 Stichprobe

An dem vorliegenden Experiment nahmen insgesamt 18 Probanden teil (16 männlich und 2 weiblich). Die Teilnehmer variierten im Alter zwischen 24 und 30 Jahren mit einem Altersdurchschnitt von 26.2 Jahren. 17 der Probanden gaben an rechtshändig - und 1 Proband linkshändig zu sein. Alle Probanden waren zum Zeitpunkt des Experiments im Besitz einer gültigen Fahrerlaubnis und bezeichneten sich selbst als erfahrene Fahrer.

9.3.7 Erhobene Daten

Im Folgenden sollen nun die in dem vorliegenden Versuch gemessenen und im weiteren Verlauf der Experimentalbeschreibung diskutierten Variablen detailliert dargestellt werden. In der vorliegenden Studie kamen dabei neben der Bestimmung der auftretenden Beanspruchungen mittels des ICA auch die klassischen Usability Maße wie Bearbeitungsdauer, Fahrleistung, Blickdaten sowie subjektive Maße zum Einsatz.

Bearbeitungsdauer

Zur Bestimmung der Bearbeitungsdauern wurden jeweils die Zeitabschnitte von Anfang der Bearbeitung der Aufgabe bis zum Beenden der Aufgabe festgehalten.

Fahrleistung

Um die Ablenkungswirkung durch die Bedienung und die Beanspruchung des Fahrers in einer simulierten Fahrsituation bewerten zu können, wurde die Leistung bei der Spurhaltung im LCT gemessen. Es handelt sich dabei um die Abweichung der tatsächlichen gefahrenen Spur zur Optimallinie (für eine Beschreibung des Verfahrens s. Experiment 2).

Blickdaten

Mittels des verwendeten Eyetracking Systems wurde zudem das Blickverhalten der Probanden während der Experimentalfahrten aufgezeichnet. Diese Vorgehensweise erlaubt es sowohl die Gesamtblickdauer für die Bearbeitung der einzelnen Bedienvorgänge als auch die Einzelblickdauern sowie die Blickabwendung von der Straße zu berechnen.

Subjektive Daten

Neben der physiologischen Messung der mentalen Beanspruchung, wurde zusätzlich die subjektiv empfundene Beanspruchung für den Benutzer erhoben. Dafür wurden die Probanden nach der Durchführung eines Usecases um eine verbale Einschätzung der erlebten Beanspruchung von 1 (wenig beanspruchend) bis 10 (stark beanspruchend) gebeten.

9.4 Ergebnisse

9.4.1 Index of Cognitive Activity

Im Folgenden sollen zunächst die Ergebnisse aus den pupillometrischen Messungen dem „Index of Cognitive Activity“ (ICA) berichtet werden.

Bei den hier berichteten Daten handelt es sich um die individuell, über die gesamte Länge des Versuchs standardisierten Daten. Bei den hier berichteten Daten handelt es sich zudem um die kumulierten und individuell über die gesamte Länge des Versuchs standardisierten ICA Werte für das rechte sowie das linke Auge.

Zeitliche Normierung

Das Ziel einer zeitlichen Normierung der Daten war es, für ein System sowie Aufgabe und Eingabemodalität charakteristisches Beanspruchungsprofil über die Zeit und über die gewonnenen ICA Daten aller Versuchspersonen zu generieren. Dafür war es notwendig der interindividuellen aber auch intraindividuellen Variabilität (Wiederholungen des gleichen Usecases) bei den Bearbeitungsdauern Rechnung zu tragen. Dies wurde in der vorliegenden Studie dadurch gelöst, dass vom Versuchsleiter jeweils per Hand der Anfang und das Ende sowie ausgewählte Unterschritte bei der Bearbeitung der Usecases in Abhängigkeit des Systems per elektronischem Marker zeitlich in den Daten kodiert wurde. Zusätzlich wurde die Tatsache angenommen, dass auch wenn sich die individuell gemessenen Bedienzeiten für die Gesamtbearbeitungsdauer sowie einzelne Abschnitte einer Aufgabe unterscheiden, sich die einzelnen Arbeitsschritte die der Benutzer innerhalb dieser Abschnitte durchführt, zeitlich relational gleich verhalten. Dies bedeutet, auch wenn eine Versuchsperson zum Beispiel für die Eingabe einer Stadt in ein Navigationssystem länger benötigt als eine andere, sie die Eingabe der einzelnen Buchstaben relational zu der Gesamtzeit zu den selben Zeitpunkten durchführt. Für die zeitliche Normierung wurden die einzelnen Abschnitte auf jeweils 250 Datenpunkte gestaucht bzw. gesteckt und so auf eine einheitliche Länge gebracht. Um eine Konfundierung mit der zeitlichen Beeinflussung durch Bedienfehler ausschließen zu können, wurden nur solche Durchgänge in die Analyse mit aufgenommen, bei denen keine Bedienfehler aufgetreten waren. Aufgrund des mehrmaligen Übens vor der eigentlichen experimentellen Durchführung konnten diese auf ein Minimum reduziert werden.

Die Unterschritte wurden für die einzelnen Systeme und Usecases wie folgt festgelegt:

- IDrive Navigationszieleingabe: Hier wurde jeweils der erste Abschnitt vom Start im Hauptmenü bis zum Start der Eingabe der Stadt, der zweite von dem Start Eingabe Stadt bis Start Eingabe Straße, der dritte von Start Eingabe Stadt bis zum Starten der Zielführung kodiert.
- HMMI Navigationszieleingabe: Hier wurden aufgrund der fehlenden Menüstruktur jeweils nur zwei Abschnitte kodiert: Der erste Abschnitt vom Start bis zum Ende der Eingabe der Straße, der zweite Abschnitt vom Start der Eingabe der Stadt bis zum Starten der Zielführung.
- Telefonnummerneingabe IDrive und HMMI: Für den Usecase der Telefonnummernwahl wurde aufgrund der sehr kurzen Bedienzeiten in den beiden Systemen jeweils nur ein Abschnitt kodiert.

Im Folgenden sollen nun die Ergebnisse dieser so zeitlich normierten Daten für jedes System sowie Anwendungsszenario (Usecase) einzeln diskutiert werden. Es soll dabei jeweils überprüft werden, wie gut die normierten ICA Verläufe für die einzelnen Wiederholungen der Usecases aufeinander abbildbar sind, um so einen Anhaltspunkt für die Reliabilität der gewonnenen Messung zu geben. Bei der Betrachtung der Verläufe soll auch diskutiert werden in wie weit eine zusätzliche Beanspruchung wie durch die zusätzliche Fahraufgabe sowie die Nutzung einer alternativen Eingabemodalität (Speller vs. Touch) den Verlauf der ICA Werte beeinflussen kann.

Unter einigen ausgewählten Bedingungen wurden zusätzlich zu den ermittelten ICA Daten exemplarisch an einer Stichprobe der Probanden (N=6) eine nachträgliche Videoanalyse der Aufgaben durchgeführt. Hier wurde die Eingabe von einzelnen Zeichen oder der Durchführung einzelner Bedienschritte auch innerhalb der einzelnen Abschnitte kodiert. Die so gewonnenen Videokodierungen wurden

daraufhin über das gleiche Verfahren, wie bereits für die ICA Daten beschrieben, zeitlich synchronisiert und konnten somit auf die jeweiligen ICA Verläufe referenziert werden.

Die Darstellung der Ergebnisse folgt der Struktur, dass zunächst die Ergebnisse für die Durchgänge ohne zusätzliche Fahraufgabe sowohl für Speller als auch für Touch einzeln für die beiden Systeme diskutiert werden sollen. Danach folgt die Diskussion der Ergebnisse unter zusätzlicher Fahraufgabe.

9.4.1.1 Ergebnisse ohne Fahraufgabe

IDrive

Navigationszieleingabe

Für das System IDrive System wurden in der Bedingung ohne zusätzliche Fahraufgabe mit Speller unter dem Usecase Navigationszieleingabe zu Validierungszwecken insgesamt 4 Durchgänge erhoben (s. Abbildung 56), um hier die pupillometrischen Messungen auf Reliabilität überprüfen zu können. In den übrigen Bedingungen wurden jeweils nur zwei Wiederholungen durchgeführt.

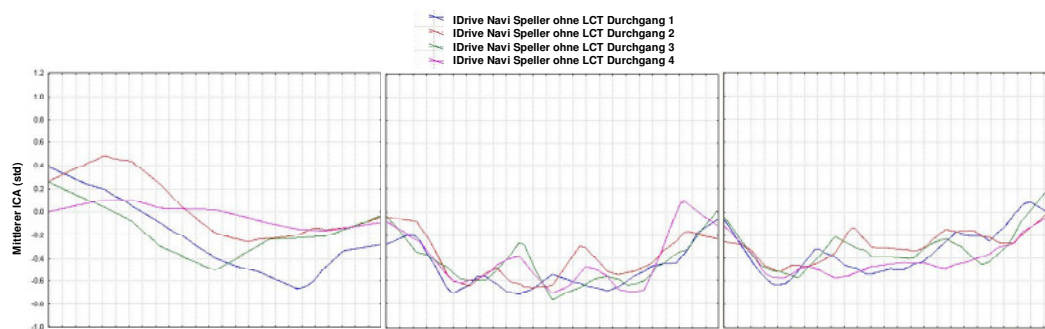


Abbildung 56: Die Abbildung zeigt die zeitlich normierten und interindividuell standardisierten ICA Verläufe über die drei Abschnitte für die Zieleingabe mittels Speller im IDrive.

Bei der Betrachtung der Daten bleibt anzumerken, dass es sich bei der Länge der betrachteten Abschnitte – trotz der gleichen Darstellung – wie bereits weiter oben beschrieben, um unterschiedliche Bearbeitungsdauern handelt. So liegt die mittlere Bearbeitungsdauer in dem ersten hier gezeigten Abschnitt (von links nach rechts) bei 4,65 s. (Auswählen des Navigationsmenüs), in dem zweiten Abschnitt bei 8,15 s (Eingabe und Auswahl der Zielstadt) und in dem dritten Abschnitt bei 8,59 s (Eingabe und Auswahl der Zielstraße und Starten der Zielführung). Zudem unterschieden sich auch die einzelnen Durchgänge 1 bis 4 in Bezug auf die Bearbeitungsdauern.

Trotz der unterschiedlichen Bearbeitungsdauern zeigte sich dennoch bereits rein deskriptiv ein äußerst homogenes Bild in den einzelnen zeitlich normalisierten ICA Verläufen. Eine statistische Betrachtung der Reliabilitäten zwischen den Durchgängen über die Bestimmung eines des Cronbach's Alpha zeigte hier einen Koeffizient von .92.

Dies bedeutet, dass die hier über den ICA gemessenen Beanspruchungsverläufe für getrennt ermittelte experimentelle Durchgänge auch statistisch ein äußerst homogenes Bild erzeugen konnten.

Aufgrund der hohen signifikanten Korrelationen zwischen den vier Durchgängen, soll in der folgenden Diskussion auf eine Trennung der vier Durchgänge zugunsten einer verbesserten Übersicht verzichtet werden.

Bei den hier beschriebenen ICA Verläufen zeigt sich jeweils ein klarer Anstieg in den betrachteten ICA Werten bei einem Abschnittswechsel. An diesen Stellen mussten die Probanden jeweils die Stadt bzw. die Straße des einzugebenden Zieles erinnern, sowie eine entsprechende Änderung in dem jeweiligen Menü vornehmen.

Für die Bedingung der Navigationszieleingabe über Touch (Abbildung 57) ergab sich für das System IDrive für die beiden experimentellen Wiederholungen ebenfalls eine hohe signifikante Korrelation von $r=.83$. Aufgrund dieser hohen Korrelation zwischen den Durchgängen ist auch unter dieser Bedingung eine hohe Reliabilität der Messung über die Wiederholungen hinweg anzunehmen.

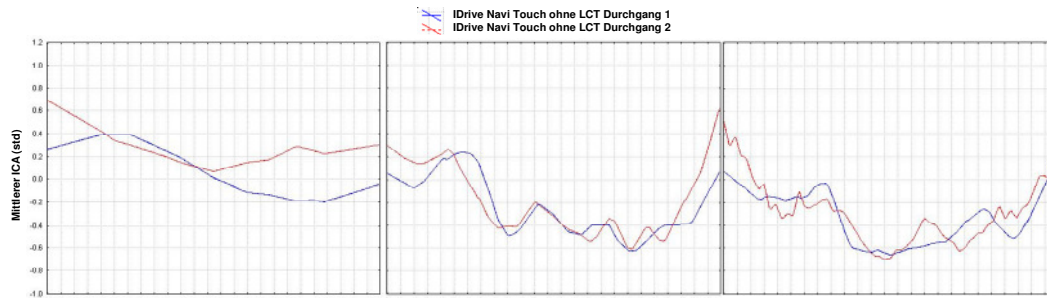


Abbildung 57: Die Abbildung zeigt die zeitlich normierten und interindividuell standardisierten ICA Verläufe über die drei Abschnitte für die Navigationszieleingabe (Navi) mittels Touch im IDrive.

Bei einer inhaltlichen Betrachtung der Ergebnisse des Beanspruchungsverlaufes wird auch bei der Betrachtung der Bedingung mit alphanumerischer Eingabe über Touch ein Anstieg bei den jeweiligen Abschnittswechseln deutlich.

Bei einem Vergleich der beiden Eingabemodalitäten Speller und Touch für den Usecase Navigationszieleingabe unter IDrive, gemittelt über die 4 bzw. 2 Wiederholungen, zeigt sich das folgende Bild in den Daten:

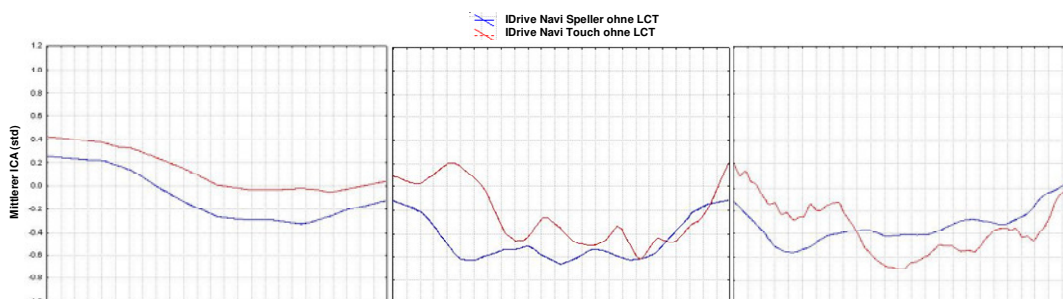


Abbildung 58: Die Abbildung zeigt den Vergleich für die Bedingungen Zieleingabe (Navi) unter IDrive für die Bedingungen Speller und Touch ohne LCT gemittelt über jeweils zwei Wiederholungen.

Telefonnummerneingabe

Für den Usecase Telefonnummerneingabe unter IDrive ergab sich für die Eingabe über Speller eine signifikante Korrelation von $r=.60$ und für die Eingabe über Touch eine ebenfalls signifikante Korrelation von $r=.88$ jeweils für die beiden experimentellen Wiederholungen (s. Abbildung 59). In der nachfolgenden Abbil-

derung sind die gemittelten Verläufe über die beiden Wiederholungen sowohl für Speller als auch für Touch abgebildet. Zusätzlich werden in Abbildung 59 auch die mittels einer nachträglich gewonnenen Videoanalyse an einer zufälligen Auswahl von Probanden (N= 6) identifizierten und zeitlich normalisierten Eingaben der einzelnen Ziffern (Zahlen) unter beiden Bedingungen dargestellt. Die Zeitpunkte für die Eingabe einzelner Ziffern wurden hier farblich markiert (Zahlen geben die einzugebende Ziffer an).

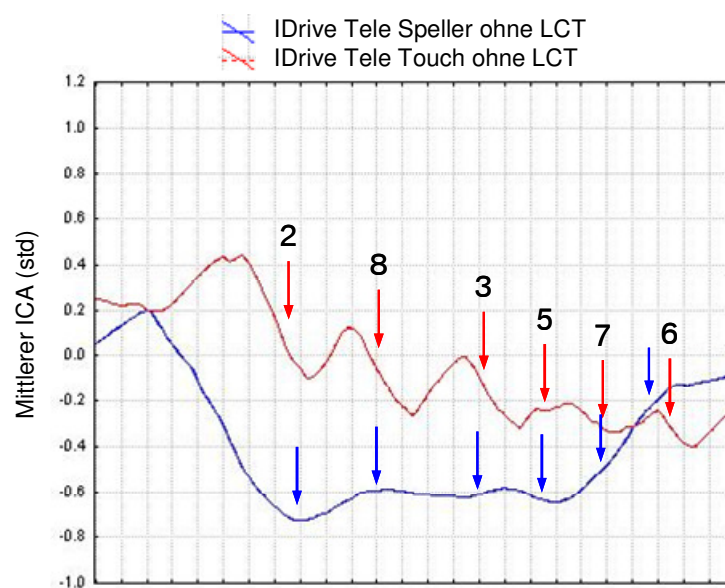


Abbildung 59: Die Abbildung zeigt den Vergleich für die Aufgabe der Telefonnummerneingabe (Tele) unter IDrive für die Bedingungen Speller und Touch ohne LCT gemittelt über jeweils zwei Wiederholungen.

Bei einem Vergleich der beiden Eingabemodalitäten für den Usecase der Telefonnummerneingabe ohne LCT zeigt sich zu Beginn sowie zum Ende des Usecases ein sehr ähnliches Bild in den Verläufen. Hier unterschieden sich die beiden Bedingungen nicht, da hier hauptsächlich MenüEinstellungen vorzunehmen sind, die sich zwischen den beiden Modalitäten nicht unterscheiden. In der Mitte des Usecases hingegen - bei der Eingabe der Ziffern - unterschieden sich die beiden Eingabemodalitäten signifikant voneinander $F(1, 17)=38.67$, $p<.001$ [Gerechnet über eine ANOVA mit Messwiederholung auf Personenbasis über die von Messzeit-

punkt 30 bis 230 gemittelten Werte]. In diesem Intervall zeigten sich signifikant höhere Werte für die Bedingung Touch gegenüber der Bedingung Speller.

Ein auffälliges in den Daten zu beobachtetes Muster für die Eingabe über Touch ist ein recht regelmäßiges phasisches Ansteigen und Abfallen in den ICA Daten während der Eingabe der einzelnen Ziffern, welches zum Ende hin geringer wird. Es zeigt sich, dass hier jeweils das Maxima der Beanspruchung jeweils kurz vor der Eingabe der jeweiligen Ziffer auftritt. Unter der Bedingung Speller zeigte sich kein solches Muster.

HMMI

Navigationszieleingabe

Für das System HMMI ergaben sich in der Bedingung Zieleingabe ohne LCT mit Speller eine signifikante Korrelation von .55 und für die Bedingung Touch eine Korrelation von .70 zwischen den ICA Verläufen der beiden experimentellen Wiederholungen.

Bei dem Vergleich der ICA Verläufe für die Bedingungen Speller und Touch zeigt sich, anders als unter IDrive, ein sehr ähnliches Bild für beide Eingabemodalitäten, wie in Abbildung 60 dargestellt.

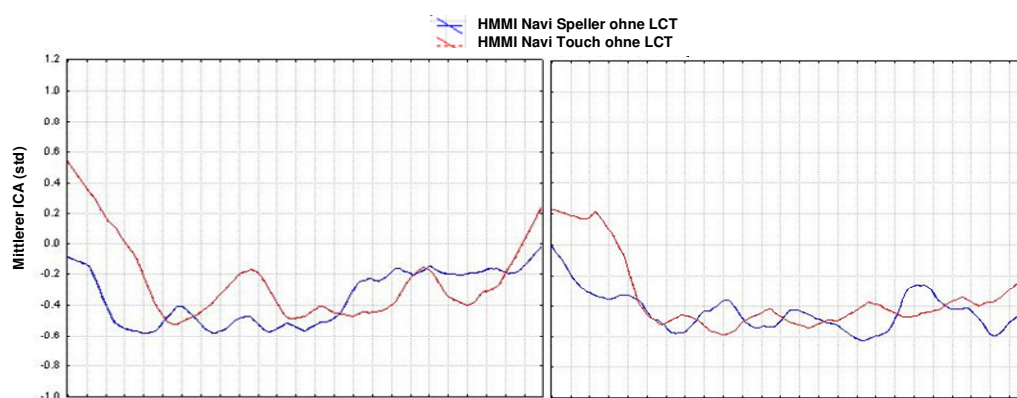


Abbildung 60: Die Abbildung zeigt den Vergleich für die Bedingungen Zieleingabe (Navi) unter HMMI für die Bedingungen Speller und Touch ohne LCT gemittelt über jeweils zwei Wiederholungen.

Zu beobachten ist bei dem Verlauf der ICA Werte unter dem System HMMI, dass es zu einem Anstieg in der Beanspruchung rund um den Abschnittswechsel kommt. An dieser Stelle beinhaltet das Konzept HMMI sowohl in der Bedingung Touch wie auch der Bedingung Speller gleichermaßen eine räumliche Darstellung eines Würfels welcher beim Abschnittswechsel gedreht wird. Eine solche räumliche mentale Rotation, welche für eine Orientierung im System notwendig ist, könnte an dieser Stelle den erhöhten ICA Wert erklären.

Telefonnummerneingabe

Für den Usecase Telefon ergab sich in der Bedingung Speller für die beiden Durchgänge eine signifikante Korrelation von $r=.76$ und unter der Bedingung Touch ebenfalls eine signifikante Korrelation von $r=.89$.

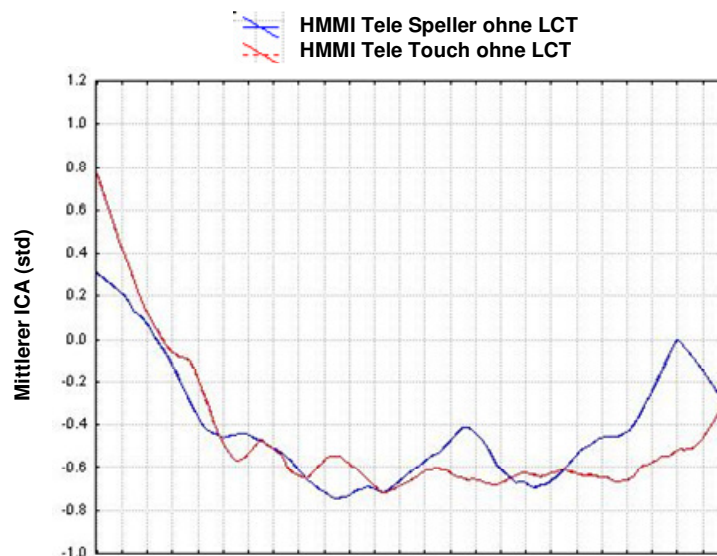


Abbildung 61: Die Abbildung zeigt den Vergleich für die Bedingungen Telefonnummerneingabe (Tele) unter HMMI für die Bedingungen Speller und Touch ohne LCT gemittelt über jeweils zwei Wiederholungen.

Auch unter dem Usecase HMMI fällt bei der Betrachtung der Ergebnisse, wie in Abbildung 61 dargestellt, auf, dass die Eingabemodalität Touch anders als für das System IDrive diskutiert hier nicht zu einer signifikanten Erhöhung der Beanspruchung gegenüber der Bedingung mit Speller geführt hat, $F(1, 17)=0.49$, $p=.49$.

Vergleich der Systeme

Ein Vergleich der beiden Systeme soll anhand des Usecases der Telefonnummerneingabe erfolgen. Hier ist die Abfolge und Struktur der Aufgabe aufgrund der Gegebenheiten der hier zu betrachtenden Systeme identisch.

Für die Eingabemodalität Speller (s. Abbildung 62) ergab eine ANOVA mit Messwiederholung für die individuell über die Wiederholungen gemittelten ICA Werte von Messpunkt 30 bis 230 keinen signifikanten Unterschied zwischen den beiden Systemen, $F(1, 17)=.002$, $p=.96$.

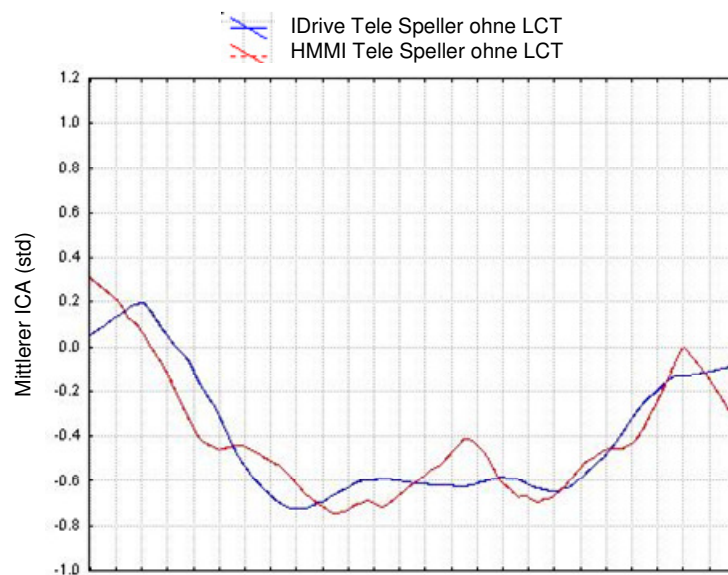


Abbildung 62: Die Abbildung zeigt den Vergleich der ICA Verläufe für die beiden Systeme IDrive sowie HMMI für den Usecase Telefonnummerneingabe für die Eingabemodalität Speller.

Für die Eingabemodalität Touch ergab eine ANOVA mit Messwiederholung für die individuell gemittelten ICA Werte von Messpunkt 30 bis 230 einen signifikanten Unterschied zwischen den beiden Systemen, $F(1, 16)=16.86$, $p<.001$. Wobei die ICA Werte bei der Telefonnummerneingabe über Touch für das IDrive höher waren als für das System HMMI.

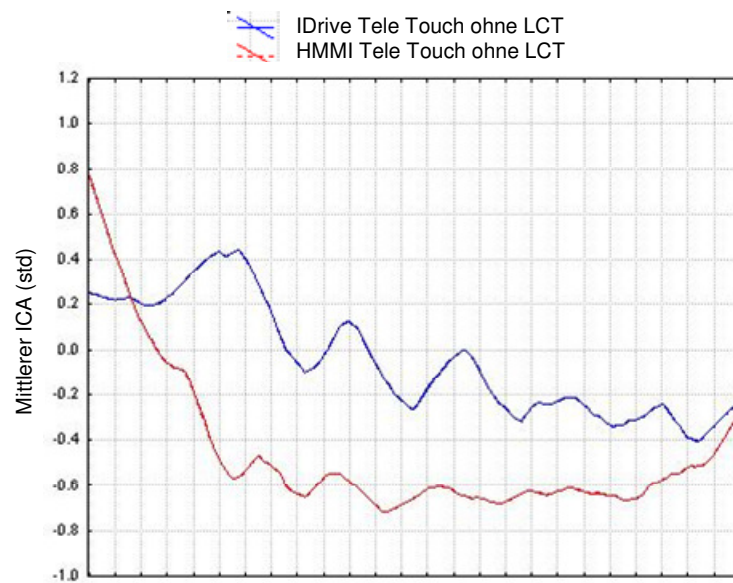


Abbildung 63: Die Abbildung zeigt den Vergleich der ICA Verläufe für die beiden Systeme IDrive sowie HMMI für den Usecase Telefonnummerneingabe für die Eingabemodalität Touch.

9.4.1.2 Ergebnisse mit Fahraufgabe

Bei der Analyse der ICA Daten für die Bedingungen mit zusätzlicher Fahraufgabe soll hier insbesondere der Anstieg der Beanspruchung durch die Fahraufgabe allgemein diskutiert werden. Eine detaillierte Analyse der einzelnen Beanspruchungsverläufe erscheint an dieser Stelle nicht zielführend, da für die Bearbeitung der Aufgaben je nach Bearbeitungsdauer jeweils unterschiedlich lange Strecken gefahren wurden und wie aus Experiment 2 und 3 bekannt hier die Schilder einen maßgeblichen Einfluss auf die Beanspruchungsmessung haben. Daher konfundieren hier zwei Einflussgrößen: Die Beanspruchungen durch die systemspezifischen Charakteristika der Aufgabe und die Beanspruchungen durch die Gegebenheiten der Fahrsituation.

Im Folgenden sollen die über die beiden Wiederholungen pro Bedingung gemittelten ICA Daten beschrieben werden. Es folgt eine Darstellung der ICA Daten für die beiden Eingabemodalitäten mit und ohne Fahraufgabe für die Navigationszeileingabe sowie die Telefonnummernwahl zunächst getrennt pro System.

IDrive

Für das System IDrive und den Usecase der Navigationszieleingabe (s. Abbildung 64) zeigte sich in einer ANOVA mit Messwiederholung auf Versuchspersonenebene ein signifikant höherer ICA Wert in den Bedingungen mit LCT gegenüber den Bedingungen ohne LCT, $F(1, 16) = 44.45$, $p < .001$.

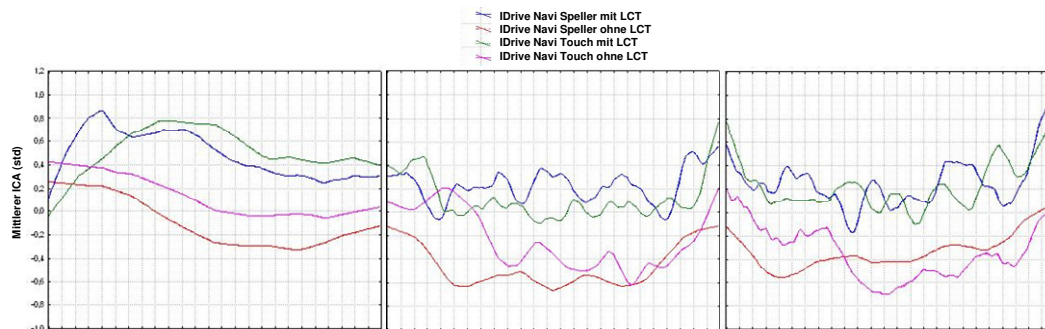


Abbildung 64: Die Abbildung zeigt die ICA Verläufe für den Usecase Navigationszieleingabe (Navi) unter IDrive unter Touch sowie Speller mit als auch ohne LCT gemittelt über die beiden ersten Wiederholungen pro Bedingung.

Ein Post Hoc Test zeigte keinen signifikanten Unterschied zwischen den beiden Eingabemodalitäten Speller und Touch unter der Bedingung mit zusätzlicher Fahraufgabe.

Für die Usecases der Telefonnummerneingabe für das System IDrive zeigte sich das in Abbildung 65 dargestellte Bild in den ICA Daten:

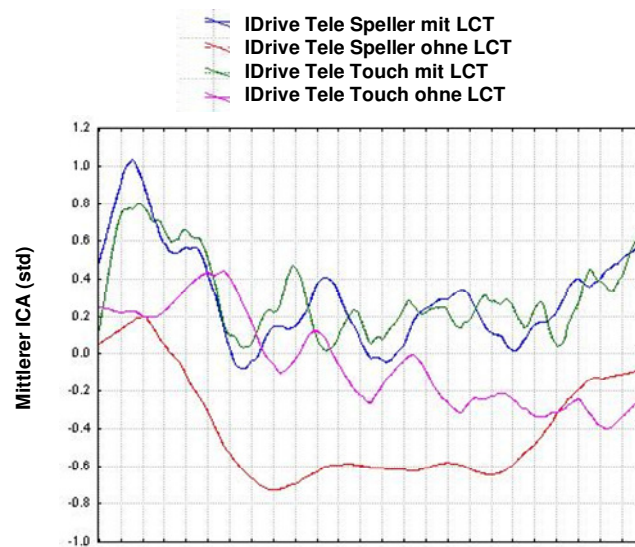


Abbildung 65: Die Abbildung zeigt die ICA Verläufe für den Usecase Telefonnummerneingabe (Tele) unter IDrive unter Touch sowie Speller mit als auch ohne LCT gemittelt über beide Wiederholungen pro Bedingung.

Auch für den Usecase der Telefonnummerneingabe zeigte sich in den ICA Daten ein signifikanter Anstieg für die Bedingungen mit zusätzlicher Fahraufgabe, $F(1, 16)=37.35$, $p<.001$. Ein Post Hoc Test zeigte keine Unterschiede zwischen den beiden Eingabemodalitäten Speller und Touch bei zusätzlicher Fahraufgabe.

HMMI

Auch für das System ID4 zeigte sich für den Usecase der Navigationszieleingabe (s. Abbildung 66) ein signifikant höherer ICA Wert in den Bedingungen mit LCT gegenüber ohne LCT, $F(1, 15)=75.13$, $p<.001$. Auch hier zeigte ein Post Hoc Test keinen signifikanten Unterschied zwischen den beiden Eingabemodalitäten Speller und Touch unter der Bedingung mit zusätzlicher Fahraufgabe.

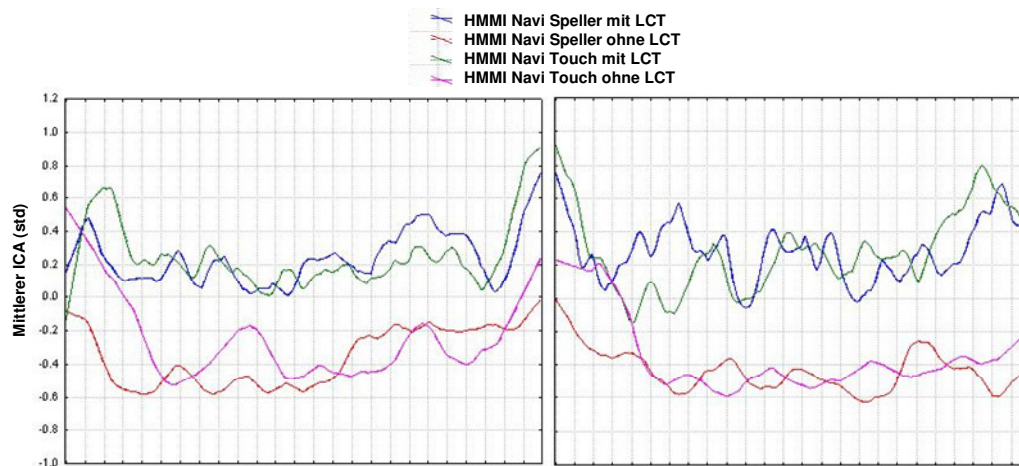


Abbildung 66: Die Abbildung zeigt die ICA Verläufe für den Usecase Navigationszieleingabe (Navi) unter HMMI unter Touch sowie Speller mit als auch ohne LCT gemittelt über beide Wiederholungen pro Bedingung.

Auffällig ist hier bei der Betrachtung der Beanspruchungsverläufe für das System HMMI mit zusätzlicher Fahraufgabe der starke Anstieg zum Abschnittswechsel. Hier mussten in dem System die Operationen in einem animierten 3D Würfel vorgenommen werden. Eine solche räumliche Transformation erfordert einen teilweise erheblichen kognitiven Aufwand, welcher sich insbesondere unter einer zusätzlichen Fahraufgabe deutlich bemerkbar macht.

Für den Usecase der Telefonnummerneingabe unter HMMI zeigte sich das in Abbildung 67 dargestellte Bild in den ICA Daten:

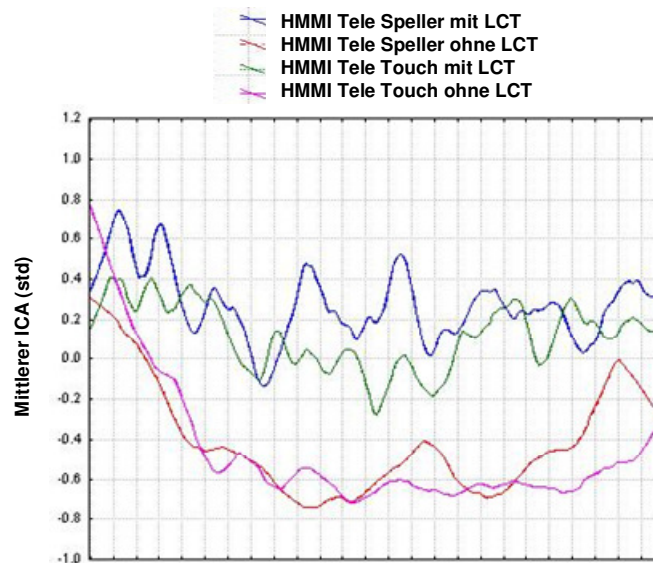


Abbildung 67: Die Abbildung zeigt die ICA Verläufe für den Usecase Telefonnummerneingabe (Tele) unter HMMI unter Touch sowie Speller mit als auch ohne LCT gemittelt über beide Wiederholungen pro Bedingung.

Für den Usecase der Telefonnummerneingabe zeigte sich in den ICA Daten ein signifikanter Anstieg für die Bedingungen mit zusätzlicher Fahraufgabe, $F(1, 16)=96.3$, $p<.001$. Auch hier zeigten sich in einem Post Hoc Test keine signifikanten Unterschiede zwischen den beiden Eingabemodalitäten bei zusätzlicher Fahraufgabe.

Vergleich zwischen den Systemen unter zusätzlicher Fahraufgabe

Bei dem Vergleich der Systeme unter LCT Bedingung stand insbesondere die Frage im Vordergrund, ob die unterschiedlichen Systeme insgesamt eine unterschiedlich hohe Beanspruchung unter einer Fahraufgabe für den Fahrer induzieren.

Die folgenden Analysen basieren auf dem interindividuell über die Zeit und Wiederholungen gemittelten Datensatz pro experimenteller Bedingung.

Für den Usecase Navi zeigte eine ANOVA mit Messwiederholung über die Bedingungen mit zusätzlicher Fahraufgabe keinen signifikanten Unterschied zwi-

schen den Systemen $F(1, 15)=0.92$, $p=.35$ sowie keine Wechselwirkung für die Eingabemodalität mit dem Faktor System, $F(1, 15)=0.19$, $p=.66$.

Für den Usecase Telefonnummerneingabe zeigte sich in einer ANOVA mit Messwiederholung über die Bedingungen mit zusätzlicher Fahraufgabe ebenfalls kein signifikanter Unterschied zwischen den Systemen, $F(1, 16)=3.73$, $p=.07$ sowie keine Wechselwirkung zwischen System und Eingabemodalität, $F(1, 16)=2.57$, $p=.12$.

9.4.1.3 Zusammenfassung Ergebnisse Pupillometrie

Das wohl wichtigste Ergebnis des hier beschriebenen Experimentes in Bezug auf die pupillometrischen Daten ist, dass es mittels des ICA und der hier vorgeschlagenen Methode der zeitlichen Normierung möglich ist, reliable, über die Wiederholungen hinweg konstante Beanspruchungsprofile über die Zeit der Bearbeitung einer Aufgabe abzubilden. Die statistischen Kennwerte zeigen hier unter allen experimentellen Bedingungen eine hohe korrelative Beziehung zwischen den einzelnen Durchgängen. Mittels der Beanspruchungsprofile war es hier möglich Unterschiede zwischen den Systemen sowie Eingabemodalitäten abzubilden und statistisch zu belegen und dies in einer zeitlichen Genauigkeit wie es klassische Usability Maße nicht ermöglichen. Es können somit einzelne Abschnitte von wenigen Sekunden Dauer gegeneinander verglichen werden. In den vorliegenden Ergebnissen ist es beispielsweise möglich zu beschreiben, wo genau im Bedienablauf eine Eingabemodalität über Touch eine erhöhte Beanspruchung für den Nutzer bedeutet und an welchen Stellen sie zu keiner Erhöhung der Beanspruchung führt. Mittels der hier berichteten ICA Daten konnten Unterschiede in der für den Nutzer auftretenden Beanspruchung zwischen den einzelnen Aufgaben sowie experimentellen Situationen (mit und ohne Fahraufgabe) nachgewiesen werden.

9.4.2 Bearbeitungsdauern

Bei der Diskussion der Bearbeitungsdauern sollen zunächst die Gesamtbedienzeiten für die einzelnen Usecases sowie Systeme und Eingabemodalitäten gemittelt über die jeweils zwei Wiederholungen getrennt nach den Bedingungen mit und ohne zusätzliche Fahraufgabe diskutiert werden.

Bearbeitungsdauern ohne Fahraufgabe

Für die Bearbeitungsdauern ohne zusätzliche Fahraufgabe zeigten sich für die alphanumerische Eingabe mit Speller sowie Touch für die drei Systeme die folgenden in den beiden unteren Abbildungen dargestellten mittleren Bearbeitungsdauern. In Abbildung 68 sind die Bearbeitungsdauern gemittelt über alle Probanden dargestellt.

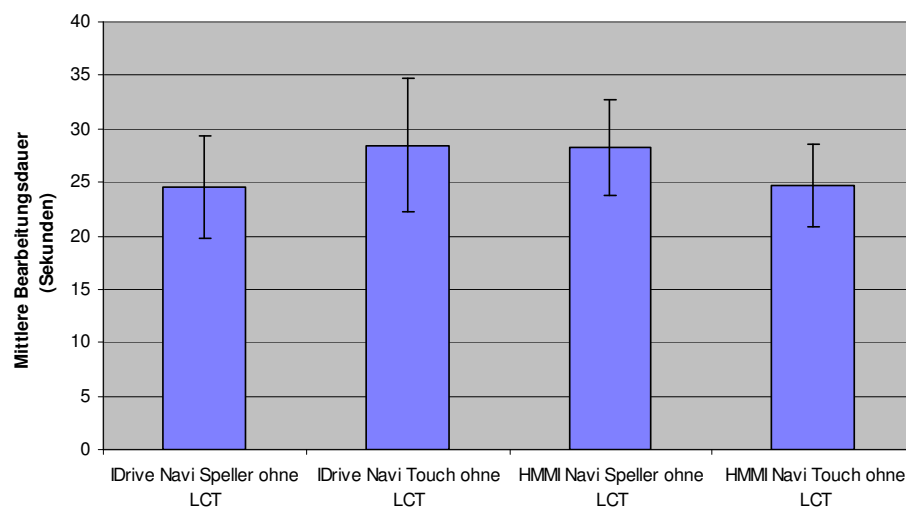


Abbildung 68: Die Abbildung zeigt die mittleren Bearbeitungsdauern für den Usecase Navigationszieleingabe (Navi) unter Speller sowie Touch ohne zusätzliche Fahraufgabe.

Eine Varianzanalyse mit Messwiederholung für den Usecase Navigationszieleingabe ohne zusätzliche Fahraufgabe zeigte keinen signifikanten Effekt zwischen

den beiden Systemen $F(1, 13)=1.087$, $p=.31$ sowie für den Faktor Eingabemodalität, $F(1, 13)=1.44$, $p=.25$.

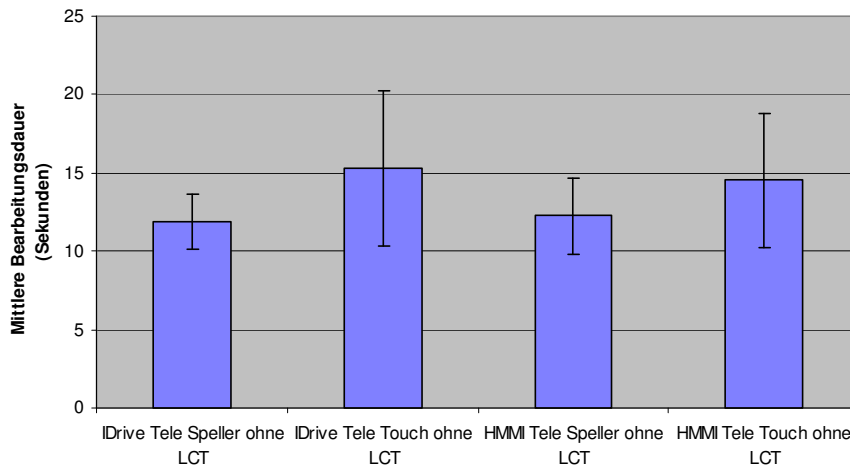


Abbildung 69: Die Abbildung zeigt die mittleren Bearbeitungsdauern für den Usecase Telefonnummerneingabe (Tele) mit Speller und Touch ohne zusätzliche Fahraufgabe.

Für den Usecase Telefonnummerneingabe ohne zusätzliche Fahraufgabe zeigte sich in einer weiteren ANOVA mit Messwiederholung ebenfalls kein signifikanter Unterschied zwischen den beiden Systemen, $F(1, 16)=0.03$, $p=.85$. Allerdings zeigte sich für den Usecase der Telefonnummerneingabe ein signifikanter Unterschied zwischen der Eingabemodalität Speller vs. Touch, $F(1, 16)=13.15$, $p<.01$ wobei die Bedienzeiten unter Touch höher lagen als für Speller.

Insgesamt zeigte sich aufgrund der Struktur der verwendeten Aufgaben ein deutlicher Unterschied zwischen den beiden Aufgaben Navigationszieleingabe und Telefonnummerneingabe. Wobei die Bearbeitungsdauern für die Navigationszieleingabe deutlich über denen der Telefonnummerneingabe lagen. In Bezug auf die Eingabemodalität ergaben sich in der Bedingung ohne zusätzliche Fahraufgabe nur bei der Telefonnummerneingabe niedrigere Bedienzeiten für die Eingabe über Speller als unter Touch.

Bearbeitungsdauern mit Fahraufgabe

Bei der Betrachtung der Bearbeitungsdauern unter zusätzlicher Fahraufgabe ergaben sich für die alphanumerische Eingabe über Speller sowie Touch für die folgenden in den beiden unteren Abbildungen dargestellten Bedienzeiten. Abbildung 70 und Abbildung 71 stellen die Bearbeitungsdauern bei gleichzeitiger Fahraufgabe gemittelt über alle Probanden dar.

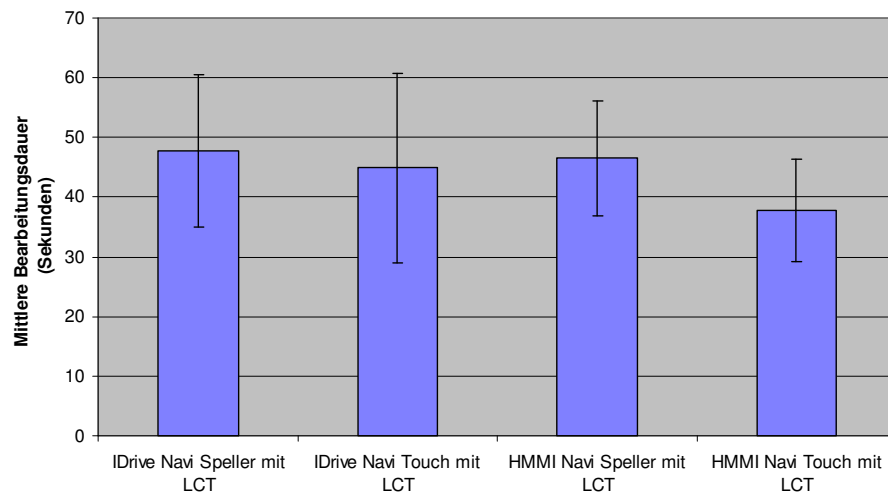


Abbildung 70: Die Abbildung zeigt die mittleren Bearbeitungsdauern für den Usecase Navigationszieleingabe (Navi) unter Speller sowie Touch unter zusätzlicher Fahraufgabe.

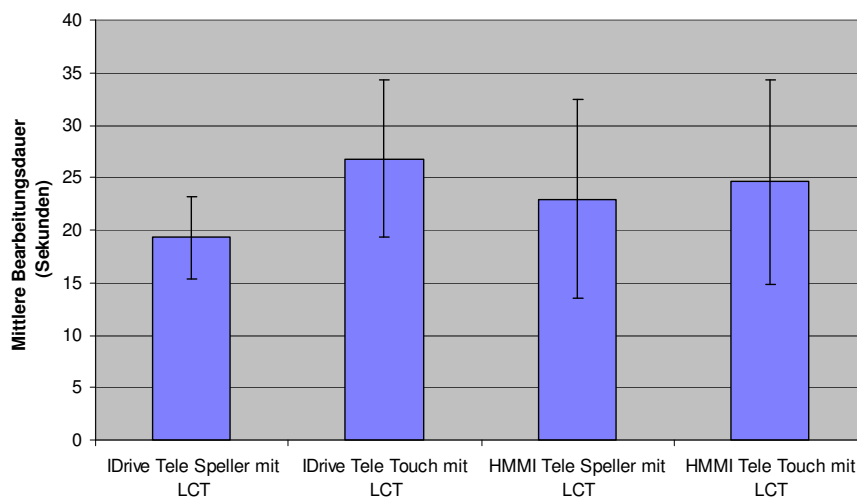


Abbildung 71: Die Abbildung zeigt die mittleren Bearbeitungsdauern für den Usecase Telefonnummerneingabe (Tele) mit Speller und Touch unter zusätzlicher Fahraufgabe.

Eine Varianzanalyse mit Messwiederholung für den Usecase der Navigationszieleingabe mit zusätzlicher Fahraufgabe zeigte keinen signifikanten Unterschied zwischen den beiden Systemen $F(1, 16)=3.95$, $p=.06$. Allerdings zeigte sich ein signifikanter Effekt auf 5% Niveau für den Faktor Eingabemodalität, $F(1, 16)=4.82$, $p=.04$. Wobei hier die Eingabemodalität Touch schneller war als unter Speller. Der Effekt der Wechselwirkung zwischen dem Faktor System und Eingabemodalität wurde nicht signifikant, $F(1, 16)=2.10$, $p=.16$.

Für den Usecase Telefonnummerneingabe mit Fahraufgabe zeigte sich in einer weiteren ANOVA mit Messwiederholung kein signifikanter Unterschied zwischen den beiden Systemen, $F(1, 15)=0.07$, $p=.78$. Allerdings zeigte sich für den Usecase Telefonnummerneingabe unter zusätzlicher Fahraufgabe ein signifikanter Unterschied zwischen der Eingabemodalität Speller vs. Touch, $F(1, 15)=4.78$, $p=.04$. Wobei hier die Bedienzeiten – anders als für den Usecase Navi - unter Touch höhere lagen als für den Speller.

Insgesamt zeigte sich aufgrund der Struktur der verwendeten Aufgaben auch hier wieder ein deutlicher Unterschied zwischen den beiden Aufgaben Navigationszieleingabe und Telefonnummerneingabe. Wobei die Bearbeitungszeiten für Navigationszieleingabe deutlich über denen von Telefonnummerneingabe lagen.

Zusammenfassung Ergebnisse Bearbeitungsdauern

Für die in dem vorliegenden experimentellen Setting erhobenen Bediendauern zeigten sich generell keine Unterschiede in den Bedienzeiten zwischen den beiden Systemen IDrive und HMMI. Es ergaben sich aufgrund der Struktur und Stellung der Aufgaben deutlich höhere Bedienzeiten für den Usecase Navigationszieleingabe gegenüber der Telefonnummerneingabe. Das wohl interessanteste Ergebnis resultierte aus der Art der Eingabemodalität. Hier zeigte sich zunächst ohne zusätzliche Fahraufgabe ein deutlicher Nachteil für die Eingabe über das Touchpad, wohingegen unter einer zusätzlichen Fahraufgabe eine Reduzierung der Bedienzeiten zu beobachten war.

9.4.3 Fahrleistung

Die nachfolgend beschriebenen Daten zur Fahrleistung beziehen sich nur jeweils auf die Abschnitte der LCT Fahraufgabe in denen die Bearbeitung der Nebenaufgabe erfolgt ist. Diskutiert wird hier die mittlere Abweichung auf der Spur zur Optimallinie wie sie im ISO Draft Standard für den Lane Change Task beschrieben wird. Zudem sollen hier die Ergebnisse der Baseline Fahrt, das heißt die Durchgänge ohne zusätzliche Bearbeitung einer Zweitaufgabe, zu Vergleichszwecken jeweils zusätzlich abgebildet werden ohne dieses statistisch zu diskutieren.

Es bleibt, wie bereits zu den vorangegangenen Experimenten diskutiert, anzumerken, dass aufgrund der Konstruktion und der Erfassung der Fahrleistungen im LCT eine 1:1 Übertragung der hier berichteten Fahrleistungsmaße auf reale Fahrleistungsmaße nicht möglich ist.

Im Folgenden soll zunächst eine Diskussion der Fahrleistungsdaten getrennt pro Usecase erfolgen. Es ist zu beachten, dass es sich bei den hier diskutierten Ergebnissen um Daten von insgesamt 15 der 18 Versuchspersonen handelt. Der Datensatz von drei Versuchspersonen musste aufgrund technischer Probleme bei der Datenerfassung aus der Analyse ausgeschlossen werden

Für den Usecase Navigationszieleingabe (s. Abbildung 72) ergab eine ANOVA mit Messwiederholung für die mittlere Abweichung auf der Spur einen signifikanten Effekt für den Faktor System, $F(1, 14)=13.76$, $p<.001$. Hier zeigte sich eine signifikant geringere Abweichung für das System HMMI gegenüber IDrive. Ein Post Hoc Test zeigte eine signifikant geringere Abweichung unter der Eingabemodalität Touch für das System HMMI gegenüber IDrive und keinen signifikanten Unterschied zwischen den beiden Systemen unter der Eingabemodalität Speller. Die varianzanalytische Betrachtung zeigt jedoch insgesamt keine signifikante Unterscheidung für die mittlere Abweichung auf der Spur zwischen den beiden

Eingabemodalitäten $F(1, 14)=0.85$, $p=.37$. Auch die Wechselwirkung zwischen System und Eingabemodalität wurde nicht signifikant, $F(1, 14)=2.91$, $p=.10$.

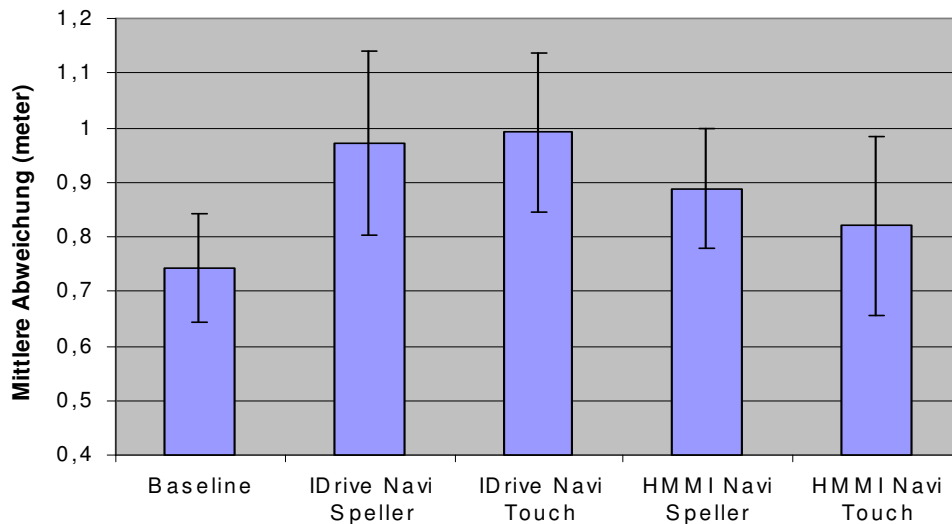


Abbildung 72: Die Abbildung zeigt die mittlere Abweichung auf der Spur für die zwei Systeme und die Eingabemodalitäten Speller sowie Touch für den Usecase Navigationszeileingabe.

Für den Usecase Telefonnummerneingabe ergab eine ANOVA mit Messwiederholung für die mittlere Abweichung der Spur keinen signifikanten Unterschied zwischen den Systemen, $F(1, 14)=0.07$, $p=.79$. Darüber hinaus zeigte sich ebenfalls kein signifikanter Effekt für die Eingabemodalität $F(1, 14)<1$, $p=.90$ oder die Wechselwirkung zwischen Eingabemodalität und System $F(1, 14)<1$, $p=.87$.

Bei der Interpretation der hier berichteten Ergebnisse aus dem Usecase der Telefonnummerneingabe muss allerdings bedacht werden, dass es sich hier bei den Bedienzeiten um – für den LCT – extrem kurze Abschnitte handelt. Dies bedeutet, da die Bearbeitungsdauer durchschnittlich bei ca. 20 Sekunden liegt, hat der Proband in dieser Zeit maximal 2 Schilder passiert. Dieses erklärt auch die hohen Varianzen in den gemessenen Spurabweichungen.

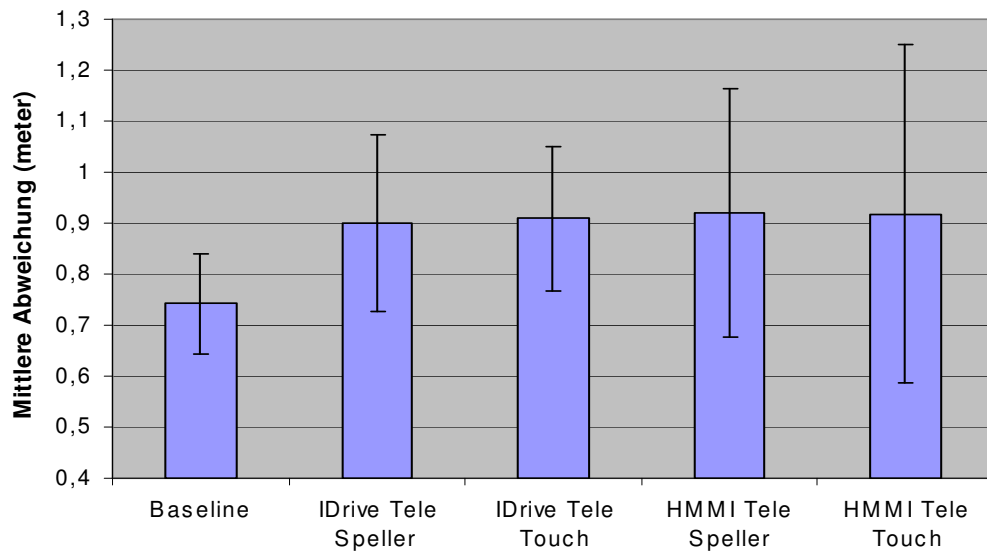


Abbildung 73: Die Abbildung zeigt die mittlere Abweichung auf der Spur für die zwei Systeme und die Eingabemodalitäten Speller sowie Touch für den Usecase Telefonnummerneingabe.

9.4.4 Blickdaten

Bei der Diskussion der Blickdaten sollen im Folgenden die Ergebnisse in Hinblick auf die mittlere Anzahl der Blicke, die mittleren Blickdauern sowie die prozentuale Blickzuwendungszeiten auf die Nebenaufgabe bei zusätzlicher Fahraufgabe beschrieben und diskutiert werden.

Datenaufbereitung

Die mittels des verwendeten Eyetrackers aufgezeichneten Blickdaten wurden mittels der während der Versuchsdurchführung gesetzten zeitlichen Markierungen entsprechend der einzelnen Abschnitte, in denen eine bestimmte Aufgabe bearbeitet wurde, aufbereitet. Die Daten wurden nach Blinzlern sowie anderen Artefaktquellen bereinigt. Daraufhin wurden entsprechend dem verwendeten Weltmodell die Areas of Interest (AOI) nach Fahraufgabe und Nebenaufgabe definiert und die Blicke aus den Eyetracking Daten bestimmt. Die hier berichteten Daten beziehen

sich nur auf die Durchgänge in denen zusätzlich zur Fahraufgabe eine Nebenaufgabe bearbeitet wurde.

Anzahl der Blicke

In Bezug auf die Auswertung der Anzahl der Blicke während der Fahrsituation wurden pro Versuchsperson sowie pro Bedingung die Anzahl der Blicke auf das Display der Nebenaufgabe gemittelt über die beiden ersten Wiederholungen bestimmt (s. Abbildung 74). Hier ergab sich für den Usecase Navigationszieleingabe in einer ANOVA mit Messwiederholung ein klarer Effekt für den Faktor System, $F(1, 17)=22.06$, $p<.001$. Hier zeigte sich eine geringere Anzahl Blicke unter HMMI gegenüber dem IDrive. Die Analyse zeigte keinen Effekt für den Faktor Eingabemodalität, $F(1, 17)=4.14$, $p=.06$. Eine Interaktion zwischen den Faktoren System und Eingabemodalität wurde nicht signifikant, $F(1, 17)<1$, $p=.96$.

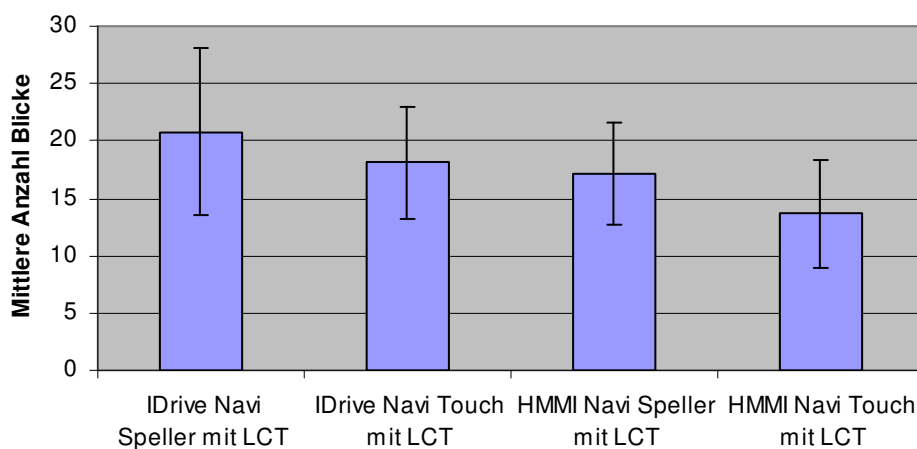


Abbildung 74: Die Abbildung zeigt die mittlere Anzahl der Blicke auf die Nebenaufgabe für den Usecase Navigationszieleingabe (Navi) mit Speller sowie Touch.

Für den Usecase Telefonnummerneingabe (s. Abbildung 75) zeigte sich in der mittleren Anzahl der Blicke in einer ANOVA mit Messwiederholung ein signifikanter Effekt für den Faktor System, $F(1, 17)=16.07$, $p<.001$. Darüber hinaus zeigte sich kein signifikanter Effekt für den Faktor der Eingabemodalität,

$F(1, 17)=3.46$, $p=.07$. Allerdings zeigte sich eine signifikante Wechselwirkung mit $F(1, 17)=9.49$, $p<.01$. Ein LSD Post Hoc Test zeigte nur einen signifikanten Anstieg in der Anzahl der Blicke unter IDrive und alphanumerischen Eingabe über Touch.

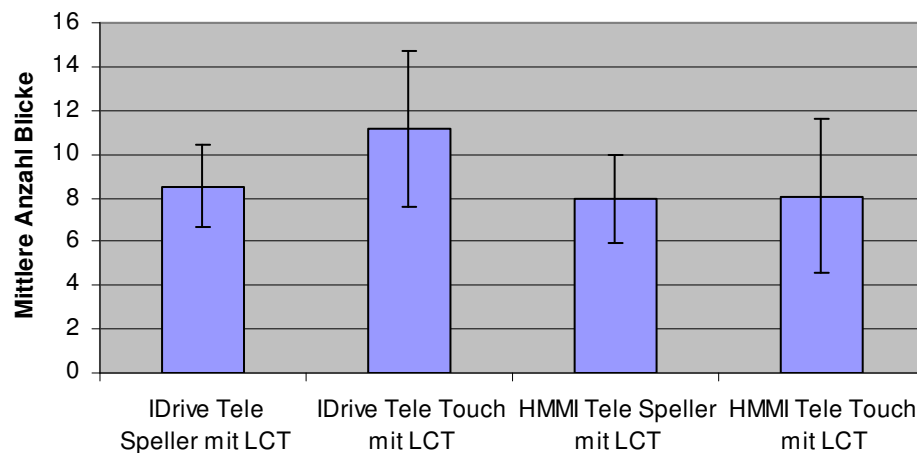


Abbildung 75: Die Abbildung zeigt die mittlere Anzahl der Blicke auf die Nebenaufgabe für den Usecase Telefonnummerneingabe (Tele) mit Speller sowie Touch.

Mittlere Blickdauer

Für die Erfassung der mittleren Blickdauern auf die Nebenaufgabe wurden wiederum die Daten aus den beiden ersten experimentellen Wiederholungen zusammengefasst betrachtet. Für den Usecase Navigationszieleingabe zeigte eine ANOVA mit Messwiederholung ein äußerst klares Bild in den Ergebnissen in Bezug auf den Faktor Eingabemodalität, $F(1, 17)=32.83$, $p<.001$. Hier zeigten sich deutlich niedrigere mittlere Blickdauern auf die Nebenaufgabe unter der Eingabe über Touch gegenüber einer Eingabe mittels Speller. Insgesamt zeigte sich kein signifikanter Unterschied zwischen den beiden betrachteten Systemen, $F(1, 17)<1$, $p=.69$.

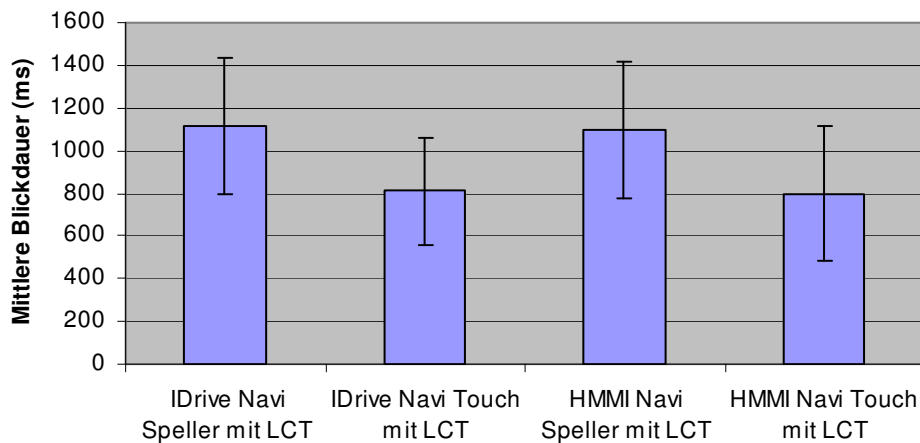


Abbildung 76: Die Abbildung zeigt die mittlere Blickdauer auf die Nebenaufgabe für den Usecase Navigationszieleingabe (Navi) mit Speller sowie Touch.

Für den Usecase der Telefonnummerneingabe zeigte sich in den Daten für die mittleren Blickdauern ein ähnliches Bild in den Ergebnissen. Eine ANOVA mit Messwiederholung zeigte für den Faktor Eingabemodalität einen signifikanten Effekt für den Faktor Eingabemodalität, $F(1, 17)=29.31$, $p<.001$. Ein Post hoc Test zeigte eine signifikant geringere Blickdauer für Touch gegenüber Speller unter beiden Systemen. Insgesamt zeigten sich jedoch kein signifikanter Unterschied zwischen den beiden Systemen, $F(1, 17)=3.27$, $p=.08$.

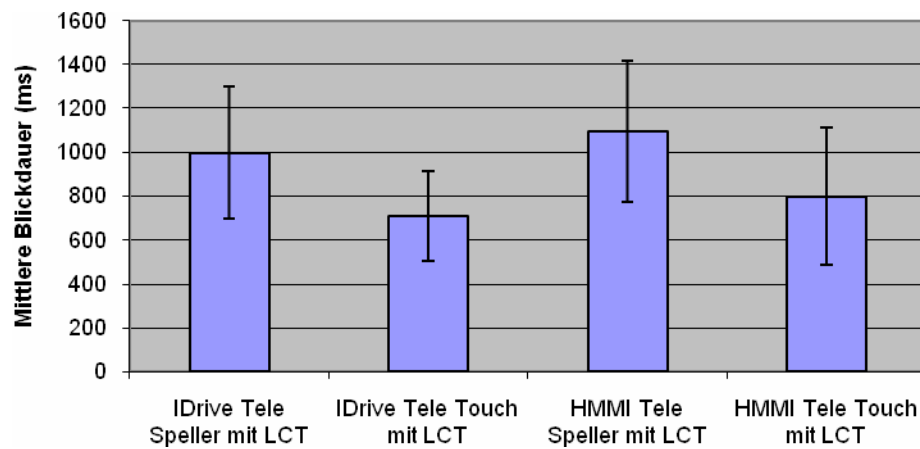


Abbildung 77: Die Abbildung zeigt die mittlere Blickdauer auf die Nebenaufgabe für den Usecase Telefonnummerneingabe (Tele) mit Speller sowie Touch.

Prozentuale Blickzuwendung auf Nebenaufgabe

Für eine Abschätzung bezüglich des Potentials einer alphanumerischen Eingabe über Touch wurde zusätzlich zu der Anzahl der Blicke sowie der mittleren Blickzuwendungsdauer auch die prozentuale Blickzuwendungsdauer auf die Nebenaufgabe bestimmt. Diese errechnet sich aus dem Quotient der mittleren Blickdauer auf die Nebenaufgabe zu der Gesamtblickdauer bzw. Gesamtbearbeitungsdauer.

Für den Usecase Navigationszieleingabe (s. Abbildung 78) zeigte sich in einer ANOVA mit Messwiederholung für die prozentuale Blickzuwendungsdauer ein hoch signifikanter Effekt für den Faktor Eingabemodalität, $F(1, 17)=60.85$, $p<.001$. Ein Post Hoc Test zeigte hier eine signifikant niedrigere prozentuale Blickzuwendungsdauer für die Eingabemodalität Touch unter beiden Systemen. Konkret ergaben sich für den Usecase Navigationszieleingabe unter IDrive für die Eingabe unter Touch gegenüber Speller eine prozentuale Verringerung der Blickzuwendungszeiten auf die Nebenaufgabe um 16 % und bei HMMI um 9 %. Auch zeigte sich ein signifikanter Unterschied zwischen den Systemen,

$F(1, 17)=8.98$, $p<.01$. Unter dem System HMMI kam es zu einer geringeren Blickzuwendungszeit auf die Nebenaufgabe als für das System IDrive.

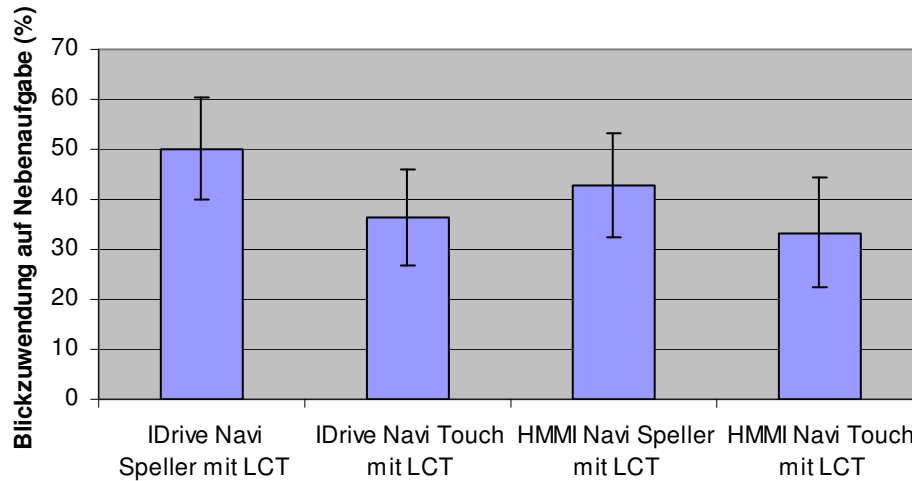


Abbildung 78: Die Abbildung zeigt den prozentualen Anteil der Blickzuwendung auf die Nebenaufgabe für den Usecase Navigationszieleingabe (Navi) mit Speller sowie Touch.

Für den Usecase Telefonnummerneingabe (s. Abbildung 79) zeigte eine ANOVA mit Messwiederholung für den Faktor Eingabemodalität ebenfalls einen hoch signifikanten Effekt, $F(1, 17)=32.24$, $p<.001$. Eine Post hoc Analyse zeigte einen signifikanten Unterschied zwischen Touch und Speller für beide Systeme. Konkret ergab sich hier eine Verringerung der prozentualen Blickzuwendungszeiten auf die Nebenaufgabe von 15% für das System IDrive für Touch gegenüber Speller und für HMMI um 16%. Für den Usecase Telefonnummerneingabe zeigte sich insgesamt kein signifikanter Unterschied zwischen den beiden Systemen, $F(1, 17)=4.10$, $p=.05$.

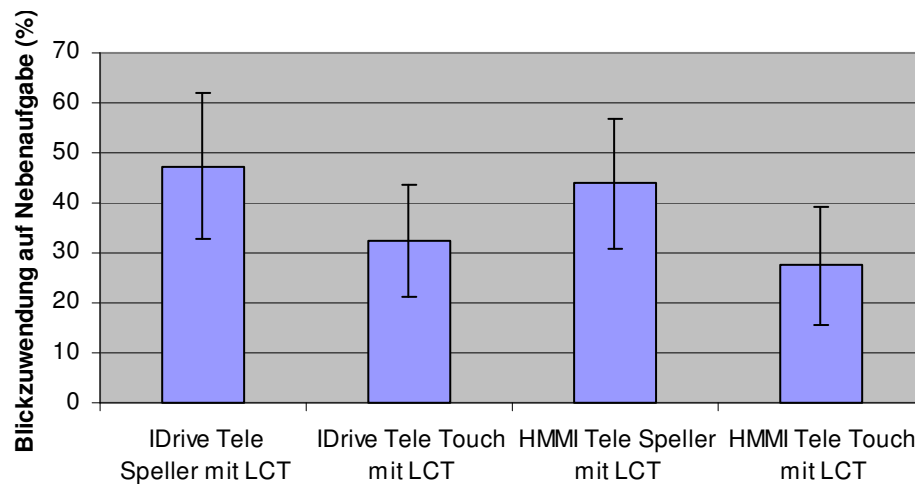


Abbildung 79: Die Abbildung zeigt den prozentualen Anteil der Blickzuwendung auf die Nebenaufgabe für den Usecase Navigationszieleingabe (Navi) mit Speller sowie Touch.

Zusammenfassung Ergebnisse Blickdaten

Die hier berichteten Ergebnisse in Bezug auf die ermittelten Blickdaten zeigen deutlich das Potential einer alphanumerischen Eingabe über Touch gegenüber einer Eingabe über Speller. Insbesondere bei der Dauer der einzelnen Blicke sowie der prozentualen Blickzuwendungsdauern konnte in dem vorliegenden experimentellen Setting eindeutig Potential für die Eingabe über Touch gegenüber einer Eingabe über Speller festgestellt werden, wonach hier die Blickabwendungszeiten von der Straße deutlich reduziert werden konnten.

9.4.5 Subjektive Daten

Im Anschluss an die Bearbeitung einer jeweiligen Aufgabe wurde von den Probanden jeweils mündlich eine Einschätzung der soeben während der Bearbeitung der Aufgabe erlebten Anstrengung auf einer Skala von 1 (sehr wenig) bis 10 (sehr stark) erhoben. Die Ergebnisse dieser subjektiven Angaben sollen nachfolgend diskutiert werden.

Zunächst werden die Ergebnisse der Durchgänge ohne zusätzliche Fahraufgabe für die beiden Systeme und die beiden Eingabemodalitäten berichtet.

Eine ANOVA mit Messwiederholung zeigte für die Bedingungen ohne zusätzliche Fahraufgabe für die Analyse über alle Systeme, Usecases sowie Eingabemodalitäten einen signifikanten Unterschied für die subjektiven Ratings der Beanspruchung zwischen den beiden Systemen, $F(1, 17)=5.73$, $p=.02$. Hier wurde insgesamt die Beanspruchung unter dem System HMMI als geringer empfunden als unter IDrive. Zusätzlich zeigte sich ein signifikanter Unterschied zwischen den beiden Usecases, $F(1, 17)=42$, $p<.001$. Hier wurde die erlebte Beanspruchung für die Navigationszieleingabe als deutlich höher erlebt als unter der Telefonnummereingabe. Für den Faktor der Eingabemodalität zeigte sich hier hingegen kein signifikanter Effekt in den subjektiven Angaben, $F(1, 17)<1$, $p=.55$.

Insgesamt ist bei den Beanspruchungsratings ohne zusätzliche Nebenaufgabe, wie in Abbildung 80 dargestellt, anzumerken, dass sich hier die subjektiv empfundenen Beanspruchungen generell auf einem sehr niedrigen Niveau bewegen. Dieses lässt darauf schließen, dass sich hier für die Probanden generell keine Schwierigkeiten oder Probleme bei der Bearbeitung der Aufgaben ergeben haben.

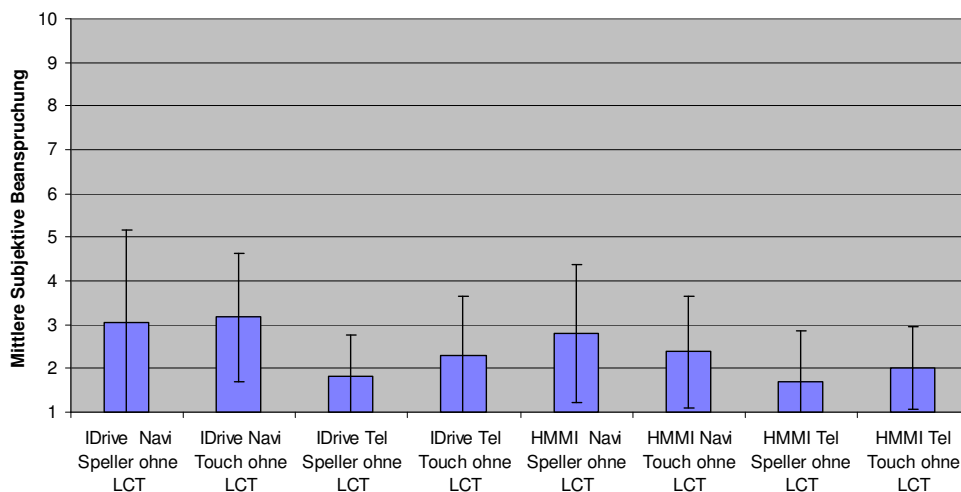


Abbildung 80: Die Abbildung zeigt die mittleren Beanspruchungsratings in den Bedingungen ohne zusätzliche Fahraufgabe (1 = sehr wenig, 10 = sehr stark).

Eine ANOVA mit Messwiederholung zeigte für die Bedingungen unter zusätzlicher Fahraufgabe, wie in Abbildung 81 dargestellt, keinen signifikanten Unterschied für die subjektiv erlebte Beanspruchung zwischen den Systemen, $F(1, 17)=2.34$, $p=.14$. Darüber hinaus zeigte sich auch unter der Bedingung mit zusätzlicher Fahraufgabe kein signifikanter Unterschied zwischen den beiden Eingabemodalitäten Speller und Toch, $F(1, 17)=1.34$, $p=.26$. Wohingegen sich erneut ein signifikanter Unterschied für die beiden Usecases zeigte, $F(1, 17)=19.8$, $p<.001$. Auch unter einer zusätzlichen Fahraufgabe wurde die Zieleingabe als deutlich beanspruchender erlebt als die Telefonnummerneingabe.

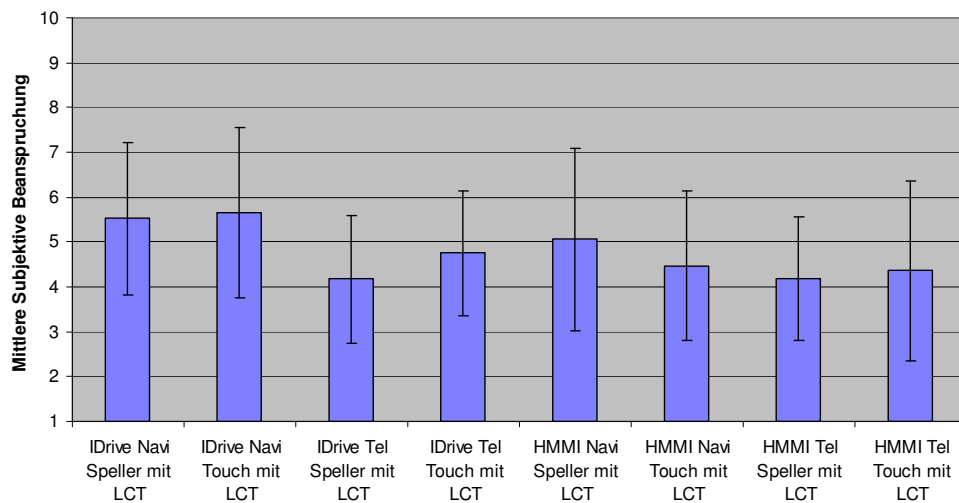


Abbildung 81: Die Abbildung zeigt die Beanspruchungsratings für die Bedingungen mit zusätzlicher Fahraufgabe.

9.5 Diskussion

Das vorliegende vierte Experiment dieser Arbeit sollte nun nach einer methodischen Validierung des ICA in den vorangegangenen Experimenten den Schritt hin zur Anwendbarkeit der Methode im Kontext der Bewertung und Optimierung von Anzeige- und Bedienkonzepten vollziehen. In dem vorliegenden experimentellen Setting wurden daher zwei reale Systeme, wie sie im Kontext einer Mensch-Maschine Interaktion im Fahrzeug eingesetzt werden, untersucht. Zusätzlich sollten in dem vorliegenden Experiment unterschiedliche, zu Verfügung stehende Eingabemodalitäten verglichen werden.

Ein Vorteil des physiologischen Maßes der Pupillometrie und der hier beschriebenen Methode des ICA liegt in der hohen zeitlichen Auflösung der Beanspruchungsmessung. Hier kann kontinuierlich über die Zeit hinweg die jeweilige Beanspruchung für den Benutzer ermittelt werden. Eine der wichtigsten Fragen bei der Gestaltung und Optimierung von Mensch-Maschine Interaktionssystemen liegt darin zu identifizieren wo und an welcher Stelle der Interaktion des Benutzers mit dem System Schwierigkeiten bzw. Schwachstellen auftreten. Wie bereits eingangs erwähnt, lag eines der Ziele der vorliegenden Studie darin zu zeigen, dass es mittels des ICA möglich ist für ein bestimmtes System und eine bestimmte Aufgabe ein charakteristisches Beanspruchungsprofil über die Zeit hinweg aufzuzeigen. Aufgrund der hohen zeitlichen Auflösung der Messung gelingt es somit – anders als bei herkömmlichen Leistungsmaßen wie etwa Bearbeitungsdauern – genau zu bestimmen wo im Ablauf der Bearbeitung einer Aufgabe Phasen mit erhöhter kognitiver Beanspruchung zu finden sind. Mittels eines solchen Beanspruchungsprofils wird somit ein genaueres Verständnis dessen möglich, was aufgrund der Gegebenheiten des vorliegenden Systems und Aufgabe tatsächlich beim Benutzer bei der Nutzung des Systems passiert.

Bei den beiden in dem vorliegenden Experiment eingesetzten Systemen handelte es sich einmal um das bereits in der BMW 5er Serie verbauten Fahrerinformationssystem IDrive sowie einem implementierten Konzept für ein Fahrerinformationssystem das HMMI System. Die beiden Systeme unterschieden sich hinsichtlich verschiedener Aspekte der Informationsstrukturierung und der Informationsdarstellung. Während das IDrive System eine klassische Menüstruktur aufweist, welche Informationen hierarchisch innerhalb verschiedener Untermenüs strukturiert, arbeitet das Konzept HMMI mit einer Informationsstrukturierung und Darstellung, welche weitestgehend auf eine solche hierarchische Struktur der Information verzichtet. Die Idee hier ist es quellen- und themenunabhängig Informationen ganzheitlich darzustellen und dem Nutzer mittels einer einheitlichen Suchfunktion zu Verfügung zu stellen. Der Nutzer benötigt keinerlei Vorwissen unter welcher Untermenüstruktur welche Information zu finden ist. Zusätzlich wurden in dem vorliegenden Experiment die alphanumerische Eingabe über ein Spellersystem sowie die Eingabe über Handschrifterkennung mittels Touchtechnologie realisiert. Während die Eingabe über Speller auf einer rein visuellen Suche und Auswahl basiert, müssen bei der Handschrifteingabe über Touch Buchstaben in Schreibschrift auf dem Touchpad umgesetzt werden.

Für die Überprüfung der vorliegenden Fragestellungen hinsichtlich der Erfassung und Beschreibung eines Beanspruchungsprofils über den ICA sowie einer Vergleichbarkeit der beiden Systeme und Eingabemodalitäten wurden in dem vorliegenden Experiment zwei Usecases realisiert. Die Eingabe eines vollständigen Navigationsziels und die Eingabe und Auswahl einer Telefonnummer.

Ein charakteristisches Beanspruchungsprofil

In Bezug auf die Erwartungen der vorliegenden Untersuchung zu der Erstellung eines für ein bestimmtes System sowie Usecase charakteristisches Beanspruchungsprofils, zeigten sich in den hier ermittelten Daten sehr eindeutig das Potential für die ermittelten ICA Daten. Durch eine entsprechende zeitliche Normierung

der Daten konnten hier sehr hohe Übereinstimmungen für die Verläufe zwischen den einzelnen Wiederholungen erreicht werden. Für den Usecase der Navigationszieleingabe ohne zusätzliche Fahraufgabe wurde unter dem System IDrive für die vier gänzlich unabhängigen Wiederholungen desselben Usecases ein Cronbachs Alpha von .92 ermittelt. Und das während sich innerhalb des Verlaufs der Bearbeitung der Aufgabe das Niveau der Beanspruchung durchaus verändert hat. Das heißt obwohl sich innerhalb der Bearbeitung der Aufgabe die Beanspruchung durchaus zwischen verschiedenen Bearbeitungsschritten unterschied, konnten diese Unterschiede über die Zeit unter allen experimentellen Wiederholungen des Usecases repliziert werden. Ähnlich hohe korrelative Beziehungen zwischen den einzelnen Wiederholungen konnten für die weiteren experimentellen Bedingungen identifiziert werden.

Es kann somit angenommen werden, dass die in dieser Studie dargestellten Beanspruchungsprofile über die Zeit durch eine systematische System- sowie Aufgabenspezifische Charakteristik ausgelöst werden.

Vergleicht man beispielsweise die Beanspruchungsverläufe zwischen den Eingabemodalitäten Touch und Speller für den Usecase der Telefonnummereingabe für das System IDrive, so wird hier deutlich, dass es unter Verwendung der Eingabemodalität Touch jeweils vor der Eingabe einer jeden Ziffer zu einem deutlichen Anstieg in den gemessenen Beanspruchungen kommt. Während hier insgesamt das Niveau deutlich höher liegt, zeigen sich zum Anfang und Ende der Durchführung des Usecases sehr ähnliche Beanspruchungswerte. Dieses ist anzunehmen, da hier die Menüführung und Menüstruktur, welche hier zu Beginn und zum Ende hin eine Rolle spielen, unter beiden Eingabemodalitäten die gleichen waren. Für die ebenfalls ermittelten subjektiven Angaben zu der empfundenen Beanspruchung konnte hier jedoch allenfalls ein deskriptiver Unterschied zwischen den beiden Eingabemodalitäten festgestellt werden. Dies mag daran gelegen haben, dass hier die Beanspruchung insgesamt als sehr niedrig empfunden wurde

und daher die Unterschiede an dieser Stelle subjektiv nicht zu entscheidenden Veränderungen geführt haben.

Für den Usecase der Telefonnummerneingabe in dem System HMMI zeigte sich anders als für das System IDrive kein signifikanter Unterschied in den auftretenden Beanspruchungsverläufen unter Touch sowie Speller. Wobei sich dennoch unter der Eingabe über Touch hier ganz ähnliche phasische Verläufe in den ICA Daten wie unter dem System IDrive zeigten. Ein solchermaßen reduziertes Niveau in den gemessenen Beanspruchungen lässt sich dadurch erklären, dass hier das System HMMI konzeptuell bereits auf die Verwendung über Touch ausgelegt war, wohingegen das System IDrive auf die Verwendung über Speller konzipiert wurde. Unter der Eingabe über Touch fehlte bei HMMI eine entsprechende Spellerleiste vollständig, welche bei IDrive aufgrund systemspezifischer Gegebenheiten immer sichtbar war. Zusammen mit einer entsprechenden Erwartungshaltung einer Eingabe über Speller im System IDrive – alle Versuchspersonen waren im Umgang mit dem System IDrive vertraut – ließe sich ein solches Ergebnis interpretieren.

Auch für den Usecase der Navigationszieleingabe zeigten sich bei der Verwendung des HMMI Systems interessante Befunde hinsichtlich der ermittelten Beanspruchungsprofile. Hier zeigte sich beispielsweise für den Usecase der Navigationszieleingabe zum Zeitpunkt des Wechsels zwischen der Eingabe der Stadt und der Eingabe der Straße ein deutlicher Anstieg in den gemessenen Beanspruchungsverläufen. Zu diesem Zeitpunkt musste ein in 3D dargestellter Würfel auf die richtige Seite gekippt werden um hier die entsprechenden Informationen auszuwählen. Eine solche dreidimensionale Rotation jedoch bedeutet einen nicht unerheblichen kognitiven Aufwand, welcher sich in den gemessenen ICA Daten zeigt. Insbesondere unter einer zusätzlichen Fahraufgabe können hier visuell-räumliche Anforderungen aus der Fahraufgabe und aus der Bearbeitung der Nebenaufgabe konkurrieren.

Für die Betrachtung der Bedingungen unter zusätzlicher Fahraufgabe zeigte sich ein deutlicher Anstieg in den Beanspruchungsdaten insgesamt. Jedoch beschränkt sich die Interpretierbarkeit der einzelnen Profile hier auf die Diskussion des absoluten Niveaus der Beanspruchung, da wie bereits aus den Experimenten 2 und 3 bekannt, hier die einzelnen Spurwechsel alleine bereits einen deutlichen Effekt in der gemessenen Beanspruchung erzeugen. Da jedoch die Bearbeitungsdauern und somit die Anzahl und Zeitpunkte der passierten Schilder zwischen den Wiederholungen nicht identisch waren, können hier die einzelnen Profile nur schwer interpretiert werden.

Dennoch bleibt hier anzumerken, dass das hier gemessene Beanspruchungsniveau über die einzelnen Bedingungen mit zusätzlicher Fahraufgabe hinweg auf einem nahezu stabilen Niveau gehalten wurde. Dieses geht konform mit der bereits unter Experiment 3 diskutierten Annahme, dass hier eine Art subjektiv sicheres Beanspruchungsniveau formuliert wird, welches es ermöglicht die Fahrleistungsdaten auf einem annähernd stabilen Niveau zu halten. Die Bedienung in beiden der betrachteten Systeme konnte zu jedem beliebigen Zeitpunkt unterbrochen und wieder aufgenommen werden.

10 Allgemeine Zusammenfassung und Diskussion

Ziel der vorliegenden Arbeit war es, die Methode der Pupillometrie bei der Gestaltung und Optimierung von neuen Anzeige- und Bedienkonzepten hinsichtlich der bei der Bedienung auftretenden Beanspruchungen für den Fahrer im automobilen Bereich einsetzbar zu machen. Die Erfassung auftretender Beanspruchungen für den Fahrer ist von entscheidender Bedeutung, da durch eine solche Optimierung der Bedienkomfort sowie die Sicherheit des Fahrers maßgeblich verbessert und Unfälle vermieden werden können (De Waard, 1996).

In der Literatur wurde bereits eine Vielzahl von Studien beschrieben, in welchen die Relation zwischen Größenveränderungen der menschlichen Pupille in Abhängigkeit der auftretenden Beanspruchungen beschrieben wurden. Allerdings wiesen diese Verfahren bislang immer einen entscheidenden Nachteil auf. Dieser ergibt sich dadurch, dass der Einfluss von Faktoren wie etwa die Menge des einfallenden Lichtes auf die Pupille einen sehr viel größeren Einfluss auf die Größe der Pupille haben, als etwa die Höhe der momentanen Beanspruchung. Dieses war bislang ein entscheidender Nachteil bei der Anwendung der Methode der Pupillometrie in realistischen Situationen. Insbesondere in Situationen, in denen eine Vielzahl unterschiedlicher Stimuli auf eine Person einwirken, waren bisherige pupillometrische Verfahren nur sehr begrenzt in der Lage explizit nur solche Veränderungen der Pupille abzubilden und zu identifizieren, welche auch mit einer Veränderung der mentalen Beanspruchung assoziiert sind.

In der vorliegenden Arbeit wurde nun ein neues pupillometrisches Verfahren der „Index of Cognitive Activity“ (ICA) diskutiert und verwendet. Der ICA stellt dabei ein Verfahren dar, welches als relatives Maß für die Veränderungen der Pupillengröße den Anspruch erhebt, auch in solchen komplexen Situationen wie sie während einer realen Mensch-Maschine-Interaktion im Fahrzeugkontext auftreten, mentale Beanspruchungen abbilden zu können. Der entscheidende Unterschied bei dem hier verwendeten Verfahren liegt darin, dass hier nicht etwa durch die

parallele Erfassung auftretender Veränderungen der Leuchtdichte in der Umgebung versucht wird diese im Nachhinein aus dem Pupillensignal heraus zu rechnen - wie in verschiedenen anderen Studien im anwendungsnahen Kontext geschehen (Rösger, 1996, Kerkau, 2005). Ein solches Herauspartialisieren von externen Einflüssen gestaltete sich immer dann schwierig, wenn die jeweiligen Einflussgrößen extern nicht exakt bestimmt werden können. So zum Beispiel in einer Fahrsituation bei der ein Blick in wenigen Sekunden von einem dunklen Fahrzeuginnenraum zu hellen Anzeigeelementen und zum hellen Himmel außerhalb des Fahrzeugs wandern kann. Hier ist es nahezu unmöglich zu jedem Zeitpunkt zu bestimmen wie viel Licht tatsächlich die Pupille erreicht. Dieses Verfahren erscheint somit in einer solchen Situation nur bedingt geeignet um auftretende Beanspruchungen über Veränderungen der Pupillengröße zu bestimmen.

In dem in dieser Arbeit diskutierten Verfahren wird stattdessen nach einem charakteristischen Veränderungsprofil in dem Pupillensignal gesucht, welches mit der ressourcengebundenen Auslastung kognitiver Funktionen im korrelativen Zusammenhang steht. Das Verfahren basiert auf der Tatsache, dass die durch eine Erhöhung der mentalen Beanspruchung hervorgerufenen Veränderungen in der Pupillengröße ein anderes charakteristisches Profil aufweisen, als solche, welche durch andere Faktoren, wie etwa Leuchtdichte, ausgelöst werden. Somit macht es ein solches Verfahren möglich, die Höhe der gemessenen Beanspruchung auch ohne die Kenntnis der übrigen Einflussgrößen zu bestimmen.

Die vorliegende Studie hatte nun zum Ziel dieses hier beschriebene Verfahren der Pupillometrie für das Anwendungsgebiet der Gestaltung und Optimierung von neuen Anzeige- und Bedienkonzepten im automotiven Bereich einsetzbar zu machen. Dafür sollte in der vorliegenden Arbeit zunächst überprüft werden, inwieweit die Methode bei der Erfassung auftretender Beanspruchungen allgemein und bei Fahrsituationen sowie bei der Bewertung verschiedener Fahrerinformationssysteme im Speziellen einsetzbar ist.

In der vorliegenden Studie wurden insgesamt 4 Experimente beschrieben, deren jeweiliger Fokus und Ergebnisse im Folgenden noch einmal kurz zusammenfassend dargestellt und diskutiert werden sollen.

Das erste in der vorliegenden Arbeit beschriebene Experiment diente einer grundlegenden Validierung der Methode des „Index of Cognitive Activity“ (ICA). Zusätzlich diente dieses erste Experiment auch der Validierung des in dieser Arbeit verwendeten experimentellen Aufbaus anhand unterschiedlich schwerer Kopfrechenaufgaben. Die Ergebnisse zeigten deutlich, dass es mittels der Methode des ICA möglich ist hier Aufgaben unterschiedlicher Schwierigkeiten voneinander zu differenzieren, unabhängig von der verwendeten Darbietungsmodalität der Aufgaben. Hier war es möglich auch unter Verwendung einer auditiven Modalität, bei der für den Probanden keinerlei visuelle Information auf dem Bildschirm zu sehen war, Unterschiede in den Schwierigkeiten der dargebotenen Aufgaben abzubilden. Daher ist anzunehmen, dass die mittels der Methode erfassten Effekte nicht durch perzeptuelle Prozesse, sondern vielmehr durch höhere mentale Beanspruchungen erklärt werden können.

Nach einer sehr grundlegenden Validierung der Methode in Experiment 1 konnte in Experiment 2 der Nachweis erbracht werden, dass mittels der Methode des ICA hier innerhalb einer simulierten Fahraufgabe Situationen mit einer erhöhten Beanspruchung für den Fahrer mit einer hohen zeitlichen Genauigkeit identifiziert und abgebildet werden können. Es zeigte sich, dass eine Instruktion die Fahrspur zu wechseln zu einem signifikanten Anstieg in den gemessenen Beanspruchungsdaten geführt hat, unabhängig von der jeweils verwendeten Modalität der Instruktion. Mittels der Methode des ICA war es hier möglich einzelne Spurwechsel im zeitlichen Verlauf sicher und präzise zu bestimmen. Ein weiteres wichtiges Ergebnis zeigte sich hier bei der Betrachtung der Ergebnisse für die Bedingung einer Fahraufgabe plus zusätzliche Nebenaufgabe. Hier zeigte sich neben einer generellen Erhöhung der gemessenen Beanspruchung eine Art strategischer Kompensati-

onseffekt bei der Durchführung von Spurwechseln unter gleichzeitiger Bearbeitung einer Nebenaufgabe. Hier verringerte sich kurzfristig vor dem Wechseln der Fahrspur die gemessene Beanspruchung. Ein solcher strategischer Kompensationseffekt wurde hier dadurch erklärt, dass aufgrund der Struktur der gewählten Fahraufgabe, bei der es möglich ist das Auftreten eines kritischen Spurwechsels recht präzise vorherzusagen, es hier für den Fahrer möglich war die Aktivität in der Nebenaufgabe und daher auch die auftretende Beanspruchung kompensativ zu regulieren. Der Fahrer bereitet sich auf das kritische Fahrmanöver vor, indem er vorausschauend Kapazitäten dafür freistellt und die Bearbeitung der Nebenaufgabe unterbricht.

In Experiment 3 sollte nun, in Anlehnung an das zweite Experiment dieser Arbeit, die Validierung der Methode im simulierten Fahrkontext erweitert werden. Dafür wurden in Experiment 3 weitere experimentelle Manipulationen der Fahrsituation eingeführt, um zu identifizieren welche Faktoren der Situation hier die über den ICA gemessene Beanspruchung beeinflussen. Es zeigte sich, dass weder die Veränderung des visuellen Eindrucks, noch die motorische Reaktion die gemessenen Veränderungen in den Beanspruchungswerten beeinflussen. Vielmehr konnte gezeigt werden, dass auch eine semantische Verarbeitung von Instruktionen ohne visuelle Veränderung oder eine motorische Reaktion über die Methode des ICA präzise über die Zeit abgebildet werden können. Insgesamt zeigte sich erneut kein Unterschied zwischen den gemessenen Beanspruchungen unter einer auditiven und einer visuellen Instruktionsmodalität.

Zusätzlich sollte in Experiment 3 in Anlehnung an den in Experiment 2 diskutierten strategisch-kompensativen Effekt bei der Steuerung auftretender Beanspruchungen für den Fahrer überprüft werden. Dafür kamen in Experiment 3 zwei verschiedene Aufgabentypen mit unterschiedlichen Schwierigkeitsstufen zum Einsatz. Der Unterschied zwischen den Aufgaben bestand darin, dass diese aufgrund ihrer Struktur entweder die Nutzung einer kompensativen Reaktion durch eine Unterbrechung der Aufgabenbearbeitung in Hinblick auf die kritischen Fahr-

situationen unterstützten oder aber diese eher verhinderten. Die Ergebnisse zeigten eine klare Bestätigung der bereits unter Experiment 2 beschriebenen Effekte. Es zeigten sich sowohl für die Performance Daten der Nebenaufgabe als auch für die gemessenen Beanspruchungen über den ICA klare kompensativ-strategische Effekte bei der Bearbeitung einer Aufgabe zusätzlich zur Durchführung der Fahraufgabe. In einer kritischen Situation steuert der Fahrer aktiv durch eine entsprechend strategische Ausrichtung seiner Handlungen die auf ihn wirkenden Beanspruchungen. Eine solch strategische Ausrichtung kann darüber hinaus durch die Charakteristika der zusätzlich bearbeiteten Aufgabe maßgeblich beeinflusst werden.

Nach der Validierung der Methode des ICA in einem simulierten Fahrkontext sollte in Experiment 4 der vorliegenden Arbeit überprüft werden, in wie weit es möglich ist, mittels dieser Methode der Pupillometrie, ein aufgabencharakteristisches Beanspruchungsprofil für ein im automotiven Kontext eingesetztes Fahrerinformationssystem über die Zeit abzubilden. Ein solches Beanspruchungsprofil ermöglicht es Optimierungspotentiale auf der Basis einzelner Bedienhandlungen im Kontext eines ganzheitlichen Bedienablaufes aufzuzeigen und hinsichtlich der für den Benutzer auftretenden Beanspruchungen zu optimieren. Dafür wurden in Experiment 4 zwei konzeptuell unterschiedliche Systeme unter zwei verschiedenen Eingabemodalitäten für alphanumerische Zeichen untersucht. Eine zeitliche Normierung auf der Basis einzelner Bedienbögen wurde eingesetzt, um individuellen Differenzen in der Dauer der Aufgabenbearbeitung Rechnung tragen zu können.

Es zeigten sich für eine bestimmte Aufgabe, System und Eingabemodalität charakteristische Verläufe der gemessenen Beanspruchung mit einer äußerst hohen Übereinstimmung zwischen experimentellen Wiederholungen. Das bedeutet, dass obwohl die Verläufe insgesamt für bestimmte Aufgaben und Eingabemodalitäten charakteristische Varianzen aufwiesen, diese über die durchgeführten experimentellen Wiederholungen hinweg äußerst konstant abgebildet werden konnten. Hier

konnten mittels der Methode des ICA innerhalb der Verläufe einzelner Bedienungsvorgänge Bedienschritte mit einer erhöhten Beanspruchung für den Benutzer identifiziert werden. Zudem zeigten sich charakteristische Unterschiede zwischen den jeweiligen Systemen, Usecases und Eingabemodalitäten.

Insgesamt folgte der argumentative Aufbau der experimentellen Betrachtung der Methode des ICA in der vorliegenden Arbeit einer klaren Struktur, angefangen bei einer streng kontrollierten Laborsituation bis hin zu einer Anwendungssituation mit realen Systemen.

Die Methode des ICA zeigte sich dabei erwartungsgemäß als ein allgemeines Maß der Beanspruchung unabhängig von modalitätsspezifischen Einflüssen.

Dies entspricht den Aussagen wie sie bereits Kahnemann (1973) in seiner Theorie zu einer limitierten Prozessressource formuliert hat. Er sah die Pupillometrie als einen Indikator für eine Art generelle Auslastung der dem Menschen zu Verfügung stehenden limitierten kognitiven Ressourcen. Auch Jackson Beatty (1982) sah aufgrund der Sensibilität einer pupillometrischen Reaktion sowohl für perzeptuelle-, kognitive- als auch reaktionsbezogene Beanspruchungen pupillometrische Reaktionen weniger als einen diagnostischen als vielmehr einen globalen Indikator mentaler Beanspruchung.

Entsprechend des unter Kapitel 2.3 diskutierten Modells der Fahrerbeanspruchung wird auch deutlich, dass es sich bei einer komplexen Anwendungssituation, wie der des Autofahrens, um eine Aufgabe handelt, bei der eine Vielzahl von Informationen auf einer höheren kognitiv-koordinativen Ebene integriert und in eine effiziente Handlungsplanung sowie -Umsetzung überführt werden müssen. In einer solchen Situation fließen alle relevanten System- und Benutzerfaktoren sowie auftretende Interaktionen zwischen diesen beiden zusammen und bestimmen somit gemeinsam die resultierende Beanspruchung. Ähnliches – wenn auch in seiner Komplexität reduziert – kann auch für die Bearbeitung einer Aufgabe in einem realistischen Fahrerinformationssystem erwartet werden. Auch hier interagieren

verschiedene Informationen mit system- sowie nutzerspezifischen Eigenschaften. Es erscheint somit sinnvoll, ja sogar notwendig, dass auftretende Beanspruchungen hier zunächst in ihrer Gesamtheit abbilden zu können. Nur so können Schwachstellen der miteinander interagierender Komponenten in einem solchen ganzheitlichen System identifiziert werden. Die Methode der Pupillometrie und der in dieser Arbeit diskutierte ICA Index als Indikator einer ganzheitlichen, modalitätsunspezifischen Beanspruchung erscheint daher geeignet um genau diesen Anforderungen methodisch zu begegnen.

Auch in Hinblick auf die in Experiment 2 und 3 geführte Diskussion eines Fahrers als aktiver Verwalter seiner mentalen Ressourcen, der durch den Einsatz angepasster Strategien die auf ihn einwirkenden Beanspruchungsfaktoren beeinflussen kann, erscheint eine hoch diagnostische Methode mit einem sehr selektiven Fokus der Messung einzelner Komponenten der Beanspruchung nicht unbedingt zielführend. So konnte gezeigt werden, dass ein Fahrer in bestimmten Situationen seine Kapazitäten von einer Aufgabe auf die andere verteilt, um so beispielsweise kompensativ auf Veränderungen des Fahrkontextes reagieren zu können. Mit einem solchen Wechsel jedoch können sich auch die Zusammensetzung der Komponenten der Fahrerbeanspruchung verändern. Eine Methode die lediglich eine ganz bestimmte solche Komponente detektieren und nachweisen kann, würde einen solchen Wechsel eventuell anders oder sogar falsch interpretieren. Es wäre dabei möglich, dass in einer bestimmten Komponente der Beanspruchung eine Veränderung detektiert wird, obwohl sich die Beanspruchung insgesamt für den Fahrer nicht geändert hat. Es ist daher zur Diskussion zu stellen, wie sinnvoll der Einsatz einer hoch selektiven Methode zur Abbildung einzelner Komponenten der Beanspruchung in einer Situation wie der einer realen Fahraufgabe ist, solange nicht ein umfassendes Modell des Fahrerverhaltens mitsamt der von ihm eingesetzten Strategien und zugrundeliegenden aktiven Prozesse der Beanspruchungssteuerung vorliegt.

In Bezug auf die Einordnung nach den Charakteristika von Messverfahren (s. Kapitel 3) zeigt sich die Methode des ICA in den Experimenten der vorliegenden Arbeit als hinreichend sensibel um Unterschiede hinsichtlich bestehender Beanspruchungen in einer Reihe unterschiedlicher Situationen und Anwendungsfällen mit einer hohen zeitlichen Präzision zu bestimmen. Zusätzlich zeigen die Ergebnisse eine hohe Stabilität der gemessenen Beanspruchungen. In Hinblick auf die Selektivität und Diagnostizität der Messung muss der ICA wie alle pupillometrische Verfahren eher als ein globaler Indikator mentaler Beanspruchung gesehen werden. Dabei bietet die Methode den großen Vorteil, dass sie den eigentlichen Handlungsablauf in einer Situation nicht stört und somit den ursprünglichen und natürlichen Beanspruchungsverlauf einer Situation abbilden kann. Eine störungsfreie Messung der Beanspruchung bedeutet hier aber auch, dass die Erfassung auftretender Beanspruchungen nicht etwa die Einführung weiterer Aufgaben notwendig macht. Die relevante Aufgabe kann als solche direkt betrachtet werden. Wie in Experiment 3 gezeigt, führen unterschiedliche Strategien bei der Bearbeitung zweier paralleler Aufgaben in einer Fahrsituation nicht immer auch zu einer Veränderung der Fahrleistung, wohl aber zu unterschiedlich hohen Beanspruchungen während der Bearbeitung der Aufgaben. Allein die Betrachtung der Fahrleistungen würde hier eine fehlerhafte oder sogar falsche Interpretation der Ergebnisse erzeugen.

In Bezug auf die experimentellen Anforderungen der Methode ergeben sich nur insofern Beeinträchtigungen, als dass für die Erfassung der relevanten Pupillendaten ein entsprechender kopfbasierter Eyetracker (s. Kapitel 5.2) verwendet werden muss. Dieser wurde jedoch in den vorliegenden Studien von den Probanden insgesamt nicht als störend empfunden.

Die hohe zeitliche Genauigkeit der Messung von Beanspruchung durch die Methode der Pupillometrie stellt hier einen weiteren wichtigen Vorteil der Methode dar. Anders als bisherige indirekte Maße der Beanspruchung, wie etwa Bearbeitungsdauern, lässt die hier beschriebenen Methode, wie in Experiment 4 dieser

Arbeit gezeigt, erstmals eine genaue Betrachtung eines charakteristischen Beanspruchungsprofils einer Aufgabe zu. Die Ergebnisse zeigen hier, dass über die Wiederholungen hinweg für eine bestimmte Aufgabe ein ganz charakteristisches Beanspruchungsprofil auf Basis einzelner Bedienschritte abgebildet werden kann. Es wird hier deutlich, dass bei der Diskussion solcher Ergebnisse nicht nur ein Blick auf das absolute Niveau der Beanspruchung ausreicht, sondern dass hier tatsächlich eine Diskussion der Beanspruchungsverläufe erfolgen muss, um entsprechende Schwachstellen in einem Anzeige- und Bedienkonzept zu identifizieren. Unter einer Berücksichtigung solcher Faktoren erscheint es sinnvoll, ja sogar notwendig, die bisherige Praxis, bei welcher die absolute Bearbeitungsdauer als Maß für die Einsetzbarkeit eines Systems betrachtet wird, zu überdenken. Eventuell kann es für den Anwender durchaus zielführender sein, die Bearbeitungsdauer zu verlängern, dafür jedoch Beanspruchungsspitzen – welche die gesamte Kapazität des Benutzers erfordern – zu vermeiden.

Der größte Vorteil insgesamt und somit auch die Attraktivität des hier beschriebene Verfahren der Pupillometrie für die Anwendung in der Praxis im automobilen Bereich liegt zum einen in einer störungsfreien Messung der Beanspruchung, einer direkten und objektiven Erfassung der Beanspruchung sowie einer sehr hohen zeitlichen Genauigkeit der Messung. Darüber hinaus kann mittels der Methode der Pupillometrie erstmals auch eine kontinuierliche Messung von Beanspruchung in seinem Verlauf über die gesamte Dauer der Bearbeitung einer realen Aufgabe erfasst und abgebildet werden.

11 Kritische Diskussion

Wenngleich die Ergebnisse der vorliegenden Untersuchung zu der Methode des Index of Cognitive Activity (ICA) das Potential der Methode der Pupillometrie in einem anwendungsnahen Kontext aufzeigen konnten, so sollten an dieser Stelle dennoch einige interessante kritische Diskussionspunkte aufgegriffen werden.

Ein Aspekt, welcher in der vorliegenden Studie nicht weiter überprüft wurde, jedoch in weiterführenden Studien zum ICA genauer diskutiert werden muss, ist der Einfluss eines Nah-Fern-Adaptationsprozesses auf die Pupillengröße. Für den Nachweis der prinzipiellen Anwendbarkeit des Verfahrens wurde in der vorliegenden Arbeit durchgehend ein experimentelles Setting verwendet, bei dem alle relevanten Stimuli auf einer Ebene mit einer konstanten Entfernung zum Betrachter abgebildet wurden. Die Überprüfung des Einflusses eines Nah-Fern-Adaptationsprozesses auf den ICA stand dabei nicht im Fokus der vorliegenden Arbeit, muss jedoch in nachfolgenden Studien zum ICA detailliert experimentell betrachtet werden.

Wie bereits in dem vorangegangenen Abschnitt diskutiert, ist die Methode der Pupillometrie eher als ein globaler Indikator mentaler Beanspruchung zu interpretieren, welcher es ermöglicht eine Art ganzheitliche Beanspruchung des kognitiven Apparates zu erfassen.

Wenngleich die neurokognitiven Mechanismen zur Steuerung der Pupillengröße weitestgehend bekannt sind, fehlen bis heute noch immer entsprechende Theorien, die vollständig erklären können, warum es überhaupt zu einer Veränderung der Pupillengröße unter Beanspruchung kommt. Bestätigt wurde in zahlreichen Studien bislang jedoch der korrelative Zusammenhang zwischen Veränderungen in der menschlichen Pupille und einer kognitiven Beanspruchung. Ein solcher korrelativer Zusammenhang jedoch kann auch kritisiert werden, da dieser Zusammen-

hang bis zu einem genauen Verständnis der zugrundeliegenden hirnpfysiologischen Prozesse indirekt und nicht zwingend kausal ist.

Diese Diskussion trifft natürlich auch auf die Erfassung und Interpretation des in der vorliegenden Studie verwendeten ICA zu. Auch hier basieren die zugrunde gelegten Verfahren und Auswertungsalgorithmen auf empirische Untersuchungen. Allerdings stellt sich an dieser Stelle die Frage, ob es sich hierbei tatsächlich um ein auf die Methode der Pupillometrie begrenztes Problem handelt oder aber ein generelles in der empirischen Humanwissenschaft verbreitetes Vorgehen. Das wohl bekannteste Beispiel einer gemessenen körperlichen Reaktion stellt hier die Erfassung von Reaktionszeit dar. Erhöht man die Schwierigkeit einer Aufgabe, so erhöht sich die Reaktionszeit. Dies beschreibt jedoch im engeren Sinne - wenn auch unter der Kontrolle weiterer Einflussgrößen - auch nur einen korrelativen Zusammenhang, wobei die zugrundeliegenden kausalen kognitiven Vorgänge ebenfalls nicht vollständig erfasst werden können bzw. bekannt sind.

Bis zu einem ganzheitlichen Verständnis bzw. einem Modell des menschlichen kognitiven Apparates, unter welchem die einzelnen kausalen Vorgänge, die zu einer Reaktion der Pupille oder einer anderen körperlichen Reaktion führen, vollständig verstanden sind, bleiben solche empirischen Betrachtungen im engeren Sinn generell korrelativer Natur. Hier stellt die in der vorliegenden Studie beschriebene Methode keine Ausnahme dar.

Einen weiteren interessanten Diskussionspunkt in Bezug auf die in der vorliegenden Arbeit vorgestellte Methode stellt die Frage nach einer Art allgemeingültigen Skala der Beanspruchung dar. Fassen wir mentale Beanspruchung als ein Konstrukt auf, welches im Kahnemannschen Sinne dem Füllstand eines großen Behälters gleicht, welcher als Ressourcenpool kognitiver Kapazitäten anstehenden Aufgaben zu Verfügung steht, so stellt sich hier die Frage nach einer Art Skala der Auslastung eines solchen Ressourcenpools. Kann hier wirklich ein Minimum und ein Maximum der Auslastung definiert und empirisch festgelegt werden? Eine solche bildliche Darstellung von Prozessressourcen mag zwar auf den ersten Blick

passend und greifbar erscheinen, muss jedoch recht schnell in Anbetracht empirischer Evidenzen erweitert und verändert werden. So müssen modalitätsspezifische Ressourcen betrachtet werden welche wiederum auf einer höheren kognitiven Ebene modalitätsunspezifisch verbunden sein müssen (s. Kapitel 2.1). In Bezug auf die vorliegende Studie zeigen sich zudem bei der Betrachtung angewandter Regeln und Strategien bei einer Beanspruchungssteuerung, dass ein solches Modell zu einfach ist. Die Verteilung mentaler Ressourcen folgt bestimmten Regeln und Strategien, welche bislang nur ansatzweise bekannt sind. Eng damit verknüpft ist somit auch die Frage nach einem Minimum sowie Maximum von Beanspruchung und einer Art Nullpunkt der Skala. Um beispielsweise ein Maximum bestimmen zu können, müssten wir dieses tatsächlich empirisch erreichen können. Es erscheint jedoch aufgrund auch der in der vorliegenden Arbeit diskutierten Ergebnisse eher unwahrscheinlich, dass ein solches Maximum aufgrund auftretender regulierender Strategien überhaupt je erreicht werden kann. Somit kann auch die Frage nach einem „kritischen“ ICA Niveau, bei dem es beispielsweise in einer Fahrsituation tatsächlich zu schwerwiegenden Fahrfehlern kommt, nicht ohne weiteres beantwortet werden. Zu stark hängen hier situative Faktoren, verwendete Strategien und auftretenden Beanspruchungen voneinander ab.

Bis ein vollständiges Bild der zugrundeliegenden Strukturen, Regeln sowie Strategien bei der Verteilung solcher Ressourcen vorliegt, bleibt eine solche Skala mit fest definierten Minimum und kritischen Maximum unerreicht. Was jedoch mit einer Methode wie dem ICA hier erreicht werden kann, ist ein weitere Schritt hin zu einem genaueren Verständnis relativer Zusammenhänge von gemessenen Beanspruchungen sowie der zugrundeliegenden regulierenden Prozesse, Regeln und Strategien.

12 **Ausblick**

In der vorliegende Studie konnte klar das Potential der hier beschriebenen pupillometrischen Methode des „Index of Cognitive Activity“ (ICA) zur Bestimmung mentaler Beanspruchungen im Kontext einer anwendungsnahen automotiven Entwicklung aufgezeigt werden. In der Folge von insgesamt vier Experimenten konnte generell die Wirksamkeit der Methode zum Nachweis mentaler Beanspruchungseffekte von sehr grundlegenden experimentellen Situationen bis hin zu anwendungsorientierten Fragestellungen erbracht werden.

Der größte Vorteil und somit auch die Attraktivität des hier beschriebene Verfahren der Pupillometrie für die Anwendung in der Praxis im automobilen Bereich liegt zum einen in einer störungsfreien Messung der Beanspruchung, einer direkten und objektiven Erfassung der Beanspruchung sowie einer sehr hohen zeitlichen Genauigkeit der Messung. Mit Hilfe der Methode der Pupillometrie und insbesondere des ICA können nun erstmals auch Prozesse der Steuerung von Beanspruchungen in einem zeitlichen Verlauf präzise und genau erfasst werden ohne die eigentliche Aufgabe zu beeinflussen. Dies ist bislang mit herkömmlichen Maßen, wie etwa Performanz und Fahrerleistungsmaßen, nicht möglich gewesen.

Die vorliegenden experimentellen Evidenzen zu auftretenden strategischen Effekten bei einer aktiven Steuerung der Beanspruchung seitens des Fahrers lassen vermuten, dass eine solch zeitlich präzise Messung auftretender Beanspruchungen den Schlüssel für ein sehr viel grundlegenderes Verständnis von Fahrleistungsmaßen sowie der Steuerung von Beanspruchung insgesamt bieten kann.

Ein solches detailliertes Verständnis der dynamischen und strategischen Prozesse sowie zeitlichen Verläufe hinter dem Fahrverhalten könnten hier gänzlich neue Wege der Optimierung einer Mensch-Maschine Schnittstelle aufzeigen. Natürli-

che Strategien der Beanspruchungssteuerung des Fahrers können aufgegriffen und unterstützt werden. Gleiches gilt auch für die Betrachtung der in der vorliegenden Arbeit beschriebenen Beanspruchungsverläufe bei der Bearbeitung einer Aufgabe in einem spezifischen System. Bei der Bearbeitung einer solchen Aufgabe handelt es sich um einen zeitlich verlaufenden Prozess. Herkömmliche Usability Kriterien für die Bewertung solcher Systeme jedoch lassen einen solchen zeitlichen Prozess zumeist gänzlich unbeachtet. Berichtet werden zumeist lediglich Daten, wie etwa die Bearbeitungszeit insgesamt, Anzahl Fehler, Blickzuwendungszeiten, subjektive Einschätzungen etc. Auch hier zeigt sich das Potential der in der vorliegenden Arbeit diskutierten Methode der Pupillometrie.

Insgesamt konnte die vorliegende Arbeit das Potential der Methode des ICA bei der Applikation anwendungsnaher Fragestellungen prinzipiell aufzeigen. Dennoch muss nun in weiterer Forschungsarbeit eine Anwendung der Methode auf eine Vielzahl unterschiedlicher Fahr- und Interaktionssituationen erfolgen. Es muss darüber hinaus empirisch geprüft werden, in welchen Bereichen von Situationen die Methode sinnvoll und gewinnbringend einsetzbar ist. Dabei ist auch zu prüfen, in welchen Situationen die Applikation der hier beschriebenen Methode des ICA die Interpretation herkömmlicher Leistungsmaße unterstützen und erweitern kann. Dazu gehört auch eine umfassende und systematische Betrachtung der in der vorliegenden Arbeit nur exemplarisch aufgezeigten strategischen Effekte der menschlichen Beanspruchungssteuerung.

Die Ergebnisse der vorliegenden Arbeit zur Methode des ICA zeigen eine neuartige Möglichkeit der Optimierung zukünftiger Mensch-Maschine Interaktionssysteme mit einer bis lang unerreichten Genauigkeit. Die Anwendbarkeit der Methode bleibt dabei keineswegs nur auf den Bereich einer automotiven Forschung beschränkt, sondern gilt für das Gebiet der Mensch-Maschine Interaktion allgemein.

13 Literatur

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science*, 205(4412), 1289-1292.

Ahern, S., & Beatty, J. (1981). Physiological evidence that demand of processing capacity varies with intelligence. In M. Friedman, J. Das & O. C. N. (Eds.), *Intelligence and Learning* (pp. 121-128). New York: Plenum.

Baber, C., Mellor, B., Graham, R., Noyes, J. M., & Tunley, C. (1996). Workload and the use of automatic speech recognition: The effects of time and resource demands. *Speech Communication*, 20(1), 37-53.

Backs, R. W. (1995). Going beyond heart rate: Autonomic space and cardiovascular assessment of mental workload. *International Journal of Aviation Psychology*, 5(1), 25-48.

Backs, R. W., Lenneman, J. K., & Sicard, J. L. (1999). The use of autonomic components to improve cardiovascular assessment of mental workload in flight simulation. *International Journal of Aviation Psychology*, 9(1), 33-47.

Backs, R. W., Lenneman, J. K., Wetzel, J. M., & Green, P. (2003). Cardiac Measures of Driver Workload during Simulated Driving with and without Visual Occlusion. *Human Factors*, 45(4), 525-538.

Backs, R. W., & Seljos, K. A. (1994). Metabolic and cardiorespiratory measures of mental effort: The effects of level of difficulty in a working memory task. *International Journal of Psychophysiology*, 16(1), 57-68.

- Backs, R. W., & Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics*, 23(4), 243-254.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2), 276-292.
- Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, 5(10), 371-372.
- Beatty, J., & Lucero-Wagoner, B. (1975). *Pupillary measurement of sensory and decision processes in a signal-detection task*. Paper presented at the Meeting of Psychonomic Society Denver.
- Beatty, J., Lucero-Wagoner, B., Cacioppo, J. T., Tassinari, L. G., & Berntson, G. G. (2000). *The pupillary system*. New York, NY, US: Cambridge University Press.
- Beatty, J., & Wagoner, B. L. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science*, 199(4334), 1216-1218.
- Ben-Nun, Y. (1986). The use of pupillometry in the study of on-line verbal processing: Evidence for depths of processing. *Brain and Language*, 28(1), 1-11.
- Bradshaw, J. L. (1968). Load and pupillary changes in continuous processing tasks. *British Journal of Psychology*, 59(3), 265-271.
- Bradshaw, J. L. (1968). Pupil Size and Problem solving. *The Quarterly Journal of Experimental Psychology*, 20(2), 116-122.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.

Brookhuis, K. A., De Waard, D., & Fairclough, S. H. (2003). Criteria for driver impairment. *Ergonomics*, 46(5), 433-445.

Brookhuis, K. A., Waard, D. d., Hancock, P. A., & Desmond, P. A. (2001). Assessment of drivers' workload: Performance and subjective and physiological indexes. In P.A. Desmond & P.A. Hancock (Eds.), *Stress, workload, and fatigue*. (pp. 321-333). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Calvert, G., Spence, C., & Stein, B. E. (2004). *The Handbook of multisensory processing*. Cambridge: MIT Press.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications in Pure and Applied Mathematics*, 41, 909-996.

Donoho, D. L., & Johnston, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90, 1200-1224.

Driver, J., & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology*, 8(2), 245-253.

Driver, J., & Spence, C. (2004). Crossmodal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), *Crossmodal Space and cross-modal attention*. Oxford: Oxford University Press.

Eggemeier, F. T., Wilson, G. F., Kramer, A. F., & Damos, D. L. (1991). Workload assessment in multi-task environments. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 207-216). London: Taylor & Francis.

Fahrenberg, J., Schneider, H., Foerster, F., Myrtek, M., & Müller, W. (1985). The quantification of cardiovascular reactivity in longitudinal studies. In H. Steptoe, H. Rüddel & H. Neus (Eds.), *Clinical and methodological issues in cardiovascular psychophysiology* (pp. 107-102). Berlin: Springer.

Fournier, L. R., Wilson, G. F., & Swain, C. R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: Manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31(2), 129-145.

Gardner, R. M., Beltramo, J. S., & Krinsky, R. (1975). Pupillary changes during encoding, storage, and retrieval of information. *Perceptual and Motor Skills*, 41(3), 951-955.

Graf, S., Keinath, A., Broy, V., & Rigoll, G. (2007). H-MMI - Interaktionskonzept für variable Daten und Funktionen. In Mensch und Computer Workshop-band. (Hrsg.) Thilo Paul-Stueve, Verlag der Bauhausuniversität Weimar, Weimar, 2007, 31-36 ISBN 978-3-86068-319-4

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33(4), 457-461.

Granholm, E., Morris, S. K., Sarkin, A. J., Asarnow, R. F., & Jeste, D. V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *Journal of Abnormal Psychology*, 106(3), 458-467.

Hagmueller, M., Rank, E., & Kubin, G. (2005). *Evaluation of the Human Voice for Indications of Workload Induced Stress in the Aviation Environment*: Eurocontrol.

Haigney, D., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics*, 44(2), 132-143.

Hakerem, G. A. D., & Sutton, S. (1966). Pupillary response at visual threshold. *Nature*, 212(5061), 485-486.

Hart, S. G., Staveland, L. E., Hancock, P. A., & Meshkati, N. (1988). *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research*. Oxford, England: North-Holland.

Heinrich, W. (1896). Die Aufmerksamkeit und die Funktion der Sinnesorgane. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 9, 342-388.

Helander, M. (1978). Applicability of drivers' electrodermal response to the design of the traffic environment. *Journal of Applied Psychology*, 63(4), 481-488.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143(3611), 1190-1192.

Hilburn, B., & Jorna, P. G. (2001). Workload and air traffic control. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue*. Mahwah: L. Erlbaum.

Hockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention, Selection, Awareness and Control: A Tribute to Donald Broadbent* (pp. 328-345). Oxford: OUP.

Hyönä, J., Tommola, J., & Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 48(3), 598-612.

Jex, H. R. (1988). Measuring Mental Workload: Problems, Progress and Promises. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland.

Jorna, P. G. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, 34(2), 237-257.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47(2), 310-339.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, 4(1), 56-88.

Kahneman, D. (1973). *Attention and effort*. New Jersey, USA: Prentice Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science (New York, N.Y.)*, 154(756), 1583-1585.

Kahneman, D., & Beatty, J. (1967). Pupillary Responses in a pitch-discrimination Task. *Perception & Psychophysics*, 2(3), 101-105.

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during mental task. *Science (New York, N.Y.)*, 157, 218-219.

Kahneman, D., & Wright, P. (1971). Changes of pupil size and rehearsal strategies in a short-term memory task. *The Quarterly Journal of Experimental Psychology*, 23(2), 187-196.

Kerkau, F. (2005). *Biosignale der Pupille zur Steuerung intelligenter User Interfaces*. Unveröffentlichte Dissertationsschrift, Berlin.

Kleinman, D., Young, P., & Higgins, G. (1996). *The DDD-III: A Tool for Empirical Research in Adaptive Organisations*: CCRTS.

Knappe, G., Keinath, A., & C., M. (2006). Empfehlungen für die Bestimmung der Spurhalteegüte im Kontext der Fahrsimulation. *MMI-Interaktiv*, 11, 3-13.

Köhler, T. (1996). Cardiovascular, electrodermal and catecholamine responses to stress in borderline hypertensives. *Stress Medicine*, 66(5), 261-267.

Kohler, T., Fricke, M., Ritz, T., & Scherbaum, N. (1997). Psychophysiological reactivity of borderline hypertensives and their recovery after mental stress. *Psychotherapy and Psychosomatics*, 66(5), 261-267.

Kramer, A. F. (1991). Physiological metrics of mental workload: a review of recent progress. In D. L. Damos (Ed.), *Multiple-Task Performance* (pp. 279-328). London: Taylor & Francis.

Kröner, B. (1984). Psychophysiologische Korrelate chronischer Kopfschmerzen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 31(4), 610-639.

Krüger, F., Nuthmann, A., & van der Meer, E. (2001). Pupillometric indices of temporal order representation in semantic memory. *Zeitschrift für angewandte Psychologie*, 209(4), 402-415.

Lee, D. H., & Park, K. S. (1990). Multivariate analysis of mental and physical load components in sinus arrhythmia scores. *Ergonomics*, 33(1), 35-47.

Lindholm, E., & Cheatham, C. M. (1983). Autonomic activity and workload during learning of a simulated aircraft carrier landing task. *Aviation, Space, and Environmental Medicine*, 54(5), 435-439.

Loewenfeld, I. E. (1993). *The Pupil: Anatomy, Physiology, and clinical Applications* (Vol. 1, 2). Ames: Iowa State University Press.

Lutzenberger, W., Elbert, T., Rockstroh, B., & Birbaumer, N. (1985). *Das EEG*. Berlin Heidelberg: Springer-Verlag.

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., Linton, P. M., Wierville, W. W., et al. (1989). *Operator Workload: Comprehensive Review and Evaluation of Operator Workload Methodologies* Alexandria, Virginia: U.S. Army Research Institute for Behavioral and Social Sciences.

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rösler (Ed.), *Ergebnisse und Anwendungen der Psychophysiologie (Enzyklopädie der Psychologie, C, Serie 1, Bd. 5)* (pp. 799 - 864). Göttingen: Hogrefe.

Marshall. (2002). *The Index of cognitive Activity: Measuring Cognitive Workload*. Paper presented at the 7th IEEE Conference on Human Factors and Power Plants, New York.

Marshall. (2005). *Assessing cognitive engagement and cognitive state from eye metrics*. Paper presented at the 1st International Conference on Augmented Cognition, Las Vegas.

Marshall, S. (2007). Measures of Attention and Cognitive Effort in Tactical Decision Making. In C. Malcom, J. M. Noyes & Y. Masakowski (Eds.), *Decision Making in Complex Environments* (pp. 321-332). Ashgate Publishing.

Marshall, S. P. (2000). U.S. Patent No. 6,090,051. Washington, DC: U.S. Patent & Trademark Office.

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space, And Environmental Medicine*, 78 (5 Suppl), B165.

Marshall, S. P., Davis, C., & Knust, S. (2004). *The Index of cognitive Activity: estimating cognitive effort from pupil dilation*. San Diego, CA: Eyetracking Inc. Technical Report ETI-0401.

Marshall, S. P., Pleydell-Pearce, C., & Dickson, B. (2002). *Integrating Psychophysiological Measures of Cognitive Workload and Eye Movements to Detect Strategy Shifts*. Paper presented at the 36th Hawaii International conference on System Sciences.

Mascord, D. J., & Heath, R. A. (1992). Behavioral and physiological indices of fatigue in a visual tracking task. *Journal of Safety Research*, 23(1), 19-25.

Mattes, S. (2003). *The lane-change-task as a tool for driver distraction evaluation*. Paper presented at the Annual Spring Conference of the GfA/17th Annual Conference of International-Society-for-Occupational-Ergonomics-and-Safety (ISOES).

- Mattes, S. (2006). *Messung der Fahrerablenkung in der Fahrzeugentwicklung*. Paper presented at the Vehicle Interaction Summit 3, Fraunhofer IAO, Stuttgart.
- McLaren, J., Erie, J., & Brubaker, R. (1992). Computerised analysis of pupillograms in studies of alertness. *Investigative Ophthalmology and Visual Science*, 33, 671-676.
- Mecklinger, A. (1992). *Gedächtnissuchprozesse. Eine Analyse ereigniskorrelierter Potentiale und der EEG-Spontanaktivität*. Weinheim: Psychologie Verlags Union.
- Metalis, S. A., Rhoades, B. K., Hess, E. H., & Petrovich, S. B. (1980). Pupilometric assessment of reading using materials in normal and reversed orientations. *Journal of Applied Psychology*, 65(3), 359-363.
- Michon, J. A. (1985). A critical view of driver behaviour models: what do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), *Human Behaviour & Traffic Safety* (pp. 485-524). New York: Plenum Press.
- Miura, T. (1986). Coping with situational demands: A study of eyemovements and peripheral vision performance. In A. Gale, I. Brown, C. Haselgrave, P. Smith & S. Taylor (Eds.), *Vision in Vehicles-II*. Amsterdam: Elsevier.
- Mulder, L. J. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, 34(2), 205-236.
- Mulder, L. J. M., & Mulder, G. (1981). Information Processing and Cardiovascular control. *Psychophysiology*, 18, 392-402.

- O'Donnell, R. D., Eggemeier, F. T., Boff, K. R., Kaufman, L., & Thomas, J. P. (1986). *Workload assessment methodology*. Oxford, England: John Wiley & Sons.
- Oehme, A. (2003). The Peripheral Detection Task - A sensitive method for measuring drivers' workload. Technische Universität Chemnitz.
- Ogden, T., & Parzen, E. (1996). Data dependent wavelet thresholding in non-parametric regression with change-point applications. *Computational Statistics and Data Analysis*, 22, 53-70.
- Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of mental workload. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue*. Mahwah: L. Erlbaum.
- Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, 11(5), 559-566.
- Platten, F. (2008). Analyse von Bearbeitungsstrategien in Fahrsituationen mit Hilfe der Pupillometrie zur verbesserten Bewertung von Anzeige- und Bedienkonzepten. Unveröffentlichte Diplomarbeit: Universität Trier.
- Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology*, 60(2), 211-229.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale: Erlbaum.
- Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, 89(3), 506-540.

Qiuyuan, J., Richter, F., Wagoner, B. L., & Beatty, J. (1985). The pupil and stimulus probability. *Psychophysiology*, 22, 204-207.

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*. Amsterdam: North Holland.

Ribback, S. (2003). *Psychophysiologische Untersuchung mentaler Beanspruchung in simulierten Mensch-Maschine-Interaktionen*. Dissertationsschrift. Universität Potsdam: Potsdam.

Richer, F., Silverman, C., & Beatty, J. (1983). Response selection and initiation in speeded reactions: A pupillometric analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 360-370.

Richter, P., Wagner, T., Heger, R., & Weise, G. n. (1998). Psychophysiological analysis of mental load during driving on rural roads: A quasi-experimental field study. *Ergonomics*, 41(5), 593-609.

Rößger, P. (1997). *Die Entwicklung der Pupillometrie zu einer Methode der Messung mentaler Beanspruchung in der Arbeitswissenschaft*. Dissertationsschrift. Sinzheim: Pro Universitate Verlag.

Rouse, W. B., Edwards, S. L., & Hammer, J. M. (1993). Modelling the dynamics of mental workload and human performance in complex systems. *IEE transactions on systems, man and cybernetics*, 23, 1662-1671.

Sarter, M., Givens, B., & Bruno, J. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain Research, Rev.* 35, 146-160.

Schandry, R. (1996). *Lehrbuch der Psychophysiologie*. München: Verlags Union.

Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain And Language*, 17(1), 133-148.

Schluroff, M., Zimmermann, T. E., Freeman, R. B., Jr., Hofmeister, K., Lorscheid, T., & Weber, A. (1986). Pupillary responses to syntactic ambiguity of sentences. *Brain And Language*, 27(2), 322-344.

Schönplflug, W. (1987). Beanspruchung und Belastung bei der Arbeit - Konzepte und Theorien. In U. Kleinbeck & J. Rutenfranz (Eds.), *Enzyklopädie der Psychologie. Arbeitspsychologie* (Vol. Bd. 1, pp. 130-184). Göttingen: Hogrefe.

Schwalm, M., & Zimmer, H., (in Vorbereitung). Utilizing spatially congruent driving instructions to enhance performance in assisted navigation.

Schwalm, M., Keinath, A., & Zimmer, H., (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. In D. de Waard, F.O. Flemisch, B. Lorenz, H. Oberheid, & K.A. Brookhuis (Eds.), *Human Factors for assistance and automation* (pp. 1 - 13). Maastricht, the Netherlands: Shaker Publishing.

Seibt, R., Scheuch, K., Boucsein, W., & Grass, A. (1998). Cardiovascular reactivity of different mental stress models in normotensives, borderline hypertensives and hypertensives. *Stress Medicine*, 14(3), 183-193.

Shiff, J. M. (1875). *La pupille considérée comme esthésiomètre* (übersetzt durch R. G. de Choisy). Paris: Bailliére.

Shiff, J. M., & Foa, P. (1874). La pupille considerée comme esthésiomètre (übersetzt durch R. G. de Choisy). *Marseille Médical*, 2, 736-741.

Siegle, G., Steinhauer, S., Stenger, V., Konecky, R., & Carter, C. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage*, 20, 114-124.

Sirevaag, E. J., Kramer, A. F., Wickens, C. D., & Reisweber, M. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. *Ergonomics*, 36(9), 1121-1140.

Spence, C., & Read, L. (2003). Speech shadowing while driving: on the difficulty of splitting attention between eye and ear. *Psychological science : a journal of the American Psychological Society*, 2003 May; Vol 14 (3), pp 251-256.

Spence, C., & Walton, M. (2005). On the inability to ignore touch when responding to vision in the crossmodal congruency task. *Acta psychologica*, 2005 Jan-Feb; Vol 118 (1-2), pp 47-70 Electronic Publication.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception & Performance*, 20(3), 555-574.

Staal, M. (2004). *Stress, Cognition, and human Performance: A Literature Review and Conceptual Framework*: NASA.

Steinhauer, S., Siegle, G., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, 52, 77-86.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors*, 36(2), 285-297.

Stokes, A. F., & Kite, K. (1994). On grasping a nettle and becomming emotional. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue*. Mahwah: L. Erlbaum.

Svenson, E., Angelborg-Thanderz, M., & Wilson, G. F. (Writer) (1999). Models of pilot performance for systems and mission evaluation - psychological and psychophysiological aspects., *USAF AMRL Technical Report (Wright-Patterson)*. August, 1-51.

Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., & Jung, T.-P. (2007). Task Performance and Eye Activity: Predicting Behavior Relating to Cognitive Workload. *Aviation, Space, and Environmental Medicine*, 78, B176-B185.

Van Boxtel, A., & Jessurun, M. (1993). Amplitude and bilateral coherency of facial and jaw-elevator EMG activity as an index of effort during a two-choice serial reaction task. *Psychophysiology*, 30(6), 589-604.

Van der Molen, H. h., & Bötticher, A. (1988). A hierachical risk model for traffic participants. *Ergonomics*, 31, 537-555.

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T.-P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors*, 43(1), 111-121.

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, 42(3), 323-342.

- Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5), 656-669.
- Verwey, W. B., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied*, 2(3), 270-285.
- Vidulich, M. A., & Tsang, P. S. (1986). Techniques of subjective workload assessment: A comparison of SWAT and the NASA-bipolar methods. *Ergonomics*, 29, 1385-1398.
- Vincente, K. J., Thornton, D. C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: a measure of mental effort. *Human Factors*(29), 171-182.
- Waard, D. d. (1996). *The measurement of drivers' mental workload*. Groningen, The Traffic Research Center.
- Wickens, D. (1984). Processing Resources in attention. In Parasuraman & D. R. Davis (Eds.), *Varieties of attention* (pp. 63-102). London: Academic Press.
- Wickens, D., & Hollands, G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). New Jersey: Prentice Hall.
- Wientjes, C. J. (1992). Respiration in psychophysiology: Methods and applications. *Biological Psychology*, 34(2), 179-203.
- Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human Factors*, 27(5), 489-502.

Williams, L. (1982). Cognitive load and the functional field of view. *Human Factors*, 24(6), 683-692.

Williams, L. (1985). Tunnel vision induced by foveal load manipulation. *Human Factors*, 24(2), 221-227.

Wilson, G. F., & Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 329-360). London: Taylor & Francis.

Yan, Z., & Amersing, L. (2005). Subjective mental workload measures. *Ergonomics*, 27(3), 199-206.

Zijlstra, F. R. (1993). *Efficiency in work behaviour. A design approach for modern tools*. Delft: Delft University of Technology.

14 Anhang

NASA Task-Load-Index (TLX)

VP:

Datum:

Einschätzung der gestellten Anforderungen:

Versuchen Sie bitte, die durch den Versuch gestellten Anforderungen zusammenfassend mit den folgenden Skalen zu bewerten (jeweils ein Feld ankreuzen):

Wieviel **geistige Aktivität** war erforderlich?

[illegible]

wenig viel

Wieviel **körperliche Aktivität** war erforderlich?

wenig viel



wenig viel

Wie waren die **zeitlichen Anforderungen** (Zeitdruck)?

wenig viel

Kategorie	Anteil
1	1
2	1
3	2
4	3
5	10
6	3
7	2
8	1
9	1
10	1

wenig viel

Wie groß war die aufzubringende (geistige und körperliche) **Anstrengung**?

[illegible]

wenig viel


Wie gut konnten Sie die gestellten **Aufgaben** erfüllen?

wenig viel

wenig viel

Wie **wohl** fühlten Sie sich während des Versuchs?

wenig viel



wenig viel

Anhang 1: Deutsche Version des NASA TLX ohne Gewichtung

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt zu haben.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.