

MARC SCHNEIDER

Blickbasierte Beanspruchungsmessung

Entwicklung und Evaluation eines Kalibrierungssystems
zur individuellen Bewertung der mentalen Beanspruchung
in der Mensch-Technik-Interaktion

Marc Schneider

Blickbasierte Beanspruchungsmessung

Entwicklung und Evaluation eines Kalibrierungssystems
zur individuellen Bewertung der mentalen Beanspruchung
in der Mensch-Technik-Interaktion

3 ifab

KARLSRUHER FORSCHUNGSBERICHTE ZUR
ARBEITSWISSENSCHAFT UND BETRIEBSORGANISATION

Blickbasierte Beanspruchungsmessung

Entwicklung und Evaluation eines Kalibrierungssystems
zur individuellen Bewertung der mentalen Beanspruchung
in der Mensch-Technik-Interaktion

von

Marc Schneider

Dissertation, Karlsruher Institut für Technologie
KIT-Fakultät für Maschinenbau

Tag der mündlichen Prüfung: 23. Mai 2017
Hauptreferentin: Prof. Dr.-Ing. Barbara Deml
Korreferent: Prof. Dr. phil. habil. Oliver Sträter

Impressum



Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark
of Karlsruhe Institute of Technology.
Reprint using the book cover is not allowed.

www.ksp.kit.edu



*This document – excluding the cover, pictures and graphs – is licensed
under a Creative Commons Attribution-Share Alike 4.0 International License
(CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>*



*The cover page is licensed under a Creative Commons
Attribution-No Derivatives 4.0 International License (CC BY-ND 4.0):
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2019 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 2363-7315

ISBN 978-3-7315-0809-0

DOI 10.5445/KSP/1000083975

Kurzfassung

In der vorliegenden Arbeit wird ein neues Verfahren zur blickbasierten Erfassung der mentalen Beanspruchung von Menschen in der Interaktion mit technischen Systemen entwickelt. Die zuverlässige Erfassung der mentalen Beanspruchung bildet eine der wichtigsten Voraussetzungen für die menschenzentrierte und nachhaltige Gestaltung sozio-technischer Systeme. Das entwickelte Verfahren berücksichtigt die hohe Individualität der Informationsverarbeitungsprozesse und der physiologischen Reaktionen des Menschen. Durch die Verwendung einer Kalibrierungsaufgabe, wird eine individuelle Beanspruchungsfunktion aus verschiedenen okularen Beanspruchungsindikatoren (z. B. Blinzelrate, Pupillendurchschnittsvariabilität, Fixationsdauer) hergeleitet. Diese individuelle Herangehensweise ermöglicht eine zuverlässige Quantifizierung und Vorhersage der mentalen Beanspruchung zur direkten Ableitung beanspruchungsrelevanter Systemgestaltungsmaßnahmen. Neben den Verhaltensänderungen in Abhängigkeit von der Dauer der Tätigkeit, der Aufgabenart und der Anforderungshöhe, berücksichtigt der Algorithmus ebenfalls bewusste und unbewusste Veränderungen der Verhaltensstrategien von Anwendern. Es werden insgesamt drei experimentelle Studien beschrieben, in denen das Verfahren schrittweise hergeleitet, entwickelt und erprobt wird. Die Aufgaben der Studien, die den Probanden gestellt werden, umfassen sowohl kognitive als auch manuelle Tätigkeiten in kontrollierten Versuchsumgebungen. Es zeigen sich Vorteile des individuellen Ansatzes gegenüber den bisherigen generellen und differenzierten Vorgehensweisen hinsichtlich der Sensitivität und Diagnostizität in der Beschreibung der mentalen Beanspruchung über physiologische Indikatoren. Durch die

kontinuierliche Aufzeichnung und die hohe zeitliche Auflösung des Blickerfassungssystems wird eine genaue Beschreibung des Beanspruchungsverlaufs ermöglicht.

Abstract

In this thesis a new method for ocular mental workload assessment of humans in the interaction with technical systems is developed in order to optimize the human-centered design of socio-technical systems. For the first time, the developed method considers the high individuality of human information processing and the physiological responses. An individual workload function is derived from different ocular workload indicators (e. g. blinks rate, pupil dilations, fixation durations) by the use of a standardized calibration task. This individual approach allows a reliable quantification and prediction of the mental workload for the direct derivation of relevant system design aspects. In addition to the behavioral changes depending on the time on task, the type of task and the level of requirements, the algorithm also considers the distinction of conscious and unconscious changes in behavioral strategies. A total of three experimental studies are described in which the method is gradually derived, developed and tested. The tasks given to the subjects within the studies include both cognitive and manual activities in controlled experimental environments. The individual approach shows advantages regarding the sensitivity and diagnosticity in the description of the mental workload via physiological indicators compared to the common general and differentially procedures. The continuous recording combined with the high temporal resolution of the eye tracking system allows an exact description of the workload profile.

Danksagung

An dieser Stelle möchte ich allen denjenigen Danken, die mich bei der Anfertigung dieser Arbeit unterstützt haben und ohne deren Mithilfe diese Dissertationsschrift niemals zustande gekommen wäre.

Zunächst möchte ich mich bei Frau Prof. Dr. Barbara Deml für die hervorragende Betreuung der Arbeit und das mir entgegengebrachte Vertrauen bedanken. Die mir, während meiner Zeit an Ihrem Institut, übertragenen Aufgaben eigenständig bearbeiten zu dürfen, hat es ermöglicht, immer wieder neue Herangehensweisen zu erforschen und schließlich umzusetzen. Die kurzen Wege und die Möglichkeit zu jeder Zeit ausgiebig oder auch mal auf die Schnelle mit Ihnen fachlich zu diskutieren oder Ihren Rat einholen zu können, hat insbesondere in der Endphase der Dissertation maßgeblich dazu beigetragen, die Arbeit in der geplanten Zeit abzuschließen.

Besonders danken möchte ich Prof. Dr. Oliver Sträter, der mich bereits im ersten Semester meines Studiums für die Arbeitswissenschaft begeisterte und schließlich auch die Ko-Betreuung dieser Arbeit übernahm. Die vielseitigen fachlichen Ratschläge und Hinweise sowie die Möglichkeit meine aktuellen Ergebnisse im Rahmen der Institutskolloquien zu präsentieren und diskutieren zu können, haben die Arbeit intensiv geprägt. In diesem Zuge auch ein herzliches Danke an die Kollegen vom Institut für Arbeits- und Organisationspsychologie der Universität Kassel, die sich bei jedem meiner Besuche intensiv in mein Thema gedacht haben und mir immer wieder neue Impulse geben konnten.

Mein Dank gilt meinen Kollegen vom Institut für Arbeitswissenschaft und Betriebsorganisation, die mir über die drei Jahre hinweg immer zur Seite gestanden haben. Sei es mit fachlichem Input, den Diskussionen über Studiendesigns oder durch die vielen heiteren Momente in den Pausen. Auch in den stressigen Phasen habt ihr dafür gesorgt, dass die Freude und Motivation nie verloren ging. Danke euch allen für die unvergessliche Zeit.

Namentlich hervorzuheben ist Sabrina Bringeland, die sich mit einer unerschütterlichen Geduld jederzeit intensiv in mein Thema und meine Fragen reingedacht hat und mir so stets durch fachliche Diskussionen weitergeholfen hat, auch wenn ich selbst mal den Überblick verloren hatte. Insbesondere in den Bereichen der Mustererkennung und der Physiologie konnte ich unglaublich viel von ihr lernen und in diese Arbeit einbringen.

Besonderer Dank gilt meiner Familie und im speziellen meinen Eltern, die mir meinen Weg bis hier her ermöglicht haben, mich immer unterstützt haben und auch in der Zeit der Promotion mit den erholsamen Tagen zuhause dafür gesorgt haben, dass die Entspannung nicht zu kurz kam.

Danke!

Karlsruhe, im Mai 2017

Marc Schneider

Inhaltsverzeichnis

1	Einleitung	1
1.1	Zielsetzung	3
1.2	Vorgehensweise	4
2	Theoretischer Hintergrund	7
2.1	Kognitive Informationsverarbeitung	7
2.1.1	Modelle der Informationsverarbeitung	8
2.1.2	Von generellen zu differenziellen und individuellen Ansätzen	19
2.2	Mentale Beanspruchung	30
2.2.1	Theorie der mentalen Beanspruchung	31
2.2.2	Individuelle Adaptionstrategien	38
2.3	Ableitung der Fragestellung	45
3	Methoden	49
3.1	Erfassung mentaler Beanspruchung	49
3.1.1	Subjektives Beanspruchungsempfinden	53
3.1.2	Leistungsbasierte Parameter	57
3.1.3	Physiologische Beanspruchungsindikatoren	65
3.2	Verfahren der Datenanalyse	81
3.2.1	Procrustes-Analyse	81
3.2.2	Statistische Verfahren	82

4	Studie 1: Evaluation der physiologischen Parameter	85
4.1	Experimentaldesign der Studie 1	86
4.2	Ergebnisse der Studie 1	89
4.2.1	A1: Flugführung	89
4.2.2	A2: Mentale Rotation	90
4.2.3	Auswertung der physiologischen Daten	93
4.3	Diskussion der Studie 1	102
4.4	Zusammenfassung der Studie 1	104
5	Entwicklung einer Kalibrierungsaufgabe	107
5.1	Konzeption	107
5.2	Informationstechnische Umsetzung	112
5.3	Entwicklung einer individuellen Auswertungslogik	115
5.3.1	Ansatz individueller Regressionsfunktionen	115
5.3.2	Ansatz der individuellen Parameterklassifikation	117
6	Studie 2: Empirische Erprobung der Kalibrierungsaufgabe	125
6.1	Experimentaldesign der Studie 2	126
6.2	Ergebnisse der Studie 2	128
6.2.1	Untersuchung der Operationalisierung der Kalibrierungsaufgabe	128
6.2.2	Auswertung der physiologischen Parameter	135
6.3	Diskussion der Studie 2	142
6.4	Zusammenfassung der Studie 2	144
7	Studie 3: Vorhersage der Beanspruchung in anwendungsnahen Aufgaben	145
7.1	Experimentaldesign der Studie 3	146
7.2	Ergebnisse der Studie 3	152
7.2.1	A1: Qualitätssicherung	152
7.2.2	A2: Transportlogistik	156
7.2.3	Vergleich der beiden Kalibrierungsaufgaben	159
7.2.4	Individuelle Beanspruchungsvorhersage	163

7.3	Diskussion der Studie 3	170
7.4	Zusammenfassung der Studie 3	172
8	Diskussion	175
9	Ausblick	183
	Literaturverzeichnis	189
	Anhang	207

Abbildungsverzeichnis

1.1	Inhalt und Aufbau der Arbeit	6
2.1	Sensorische Verarbeitungsmodelle	9
2.2	Ressourcenbasiertes Verarbeitungsmodell	10
2.3	Rekursive Verarbeitung neuer Stimuli	11
2.4	Menschliche Informationsverarbeitungsschritte	12
2.5	Schematische Informationsverarbeitung	14
2.6	Vereinfachte Darstellung der Kontrollebenen	15
2.7	Der kognitive Verarbeitungszyklus	18
2.8	Klassifizierung der Modelle menschlicher Informationsverarbeitung	20
2.9	Typen der unbewussten Verarbeitung	27
2.10	Task Load und Workload	31
2.11	Das multiple Ressourcenmodell	34
2.12	Beanspruchung und Leistung	37
2.13	Der kontextabhängige Verarbeitungszyklus	40
2.14	Einfache und gemischte Strategien	42
3.1	Trimodaler Ansatz der Beanspruchungsmessung	53
3.2	Geometrien der mentalen Rotation	61
3.3	Wertebereiche des NNI	71
3.4	Muskeln der Pupille	75
3.5	Stufen der Wavelet-Transformation	77
4.1	Studie 1: Aufgaben und Anforderungsniveaus	89

4.2	Flugführung: subjektive und leistungs-basierte Daten	90
4.3	Mentale Rotation: subjektive und leistungs-basierte Daten	91
4.4	Mentale Rotation: subjektive und leistungs-basierte Daten	96
4.5	MDS mit k-Means Clustern	99
5.1	Entwicklung der Kalibrierungsaufgabe	109
5.2	Kugelbewegungen	113
5.3	Vereinfachte Darstellung der sequentiellen Parameterauswahl	117
5.4	Proaktives und reaktives Adaptionsverhalten	120
6.1	Versuchsablauf Studie 2	127
6.2	NASA-TLX Bewertungen Studie 2	129
6.3	Leistung in der Hauptaufgabe Studie 2	131
6.4	Leistungsdaten der Nebenaufgabe Studie 2	133
6.5	NNI über alle Anforderungsniveaus	138
6.6	Blinzelrate über alle Anforderungsniveaus	139
6.7	Mittlere Korrelationskoeffizienten mit der Leistung in der Nebenaufgabe	141
7.1	Versuchsaufbau Aufgabe A1: Qualitätssicherung	148
7.2	Übersicht der Anforderungsniveaus der Aufgabe A1 und A2	151
7.3	Abbildungen/Abbildung ₇₂ .pdf	153
7.4	Abbildungen/Abbildung ₇₃ .pdf	157
7.5	Subjektive und leistungs-basierte Beanspruchungsindikatoren in den Kalibrierungsaufgaben 1 und 2	160
7.6	Ergebnisse des Verfahrens individueller Regressionen	164
7.7	Ergebnisse des Verfahrens zur individuellen Parameterklassifikation	166
7.8	Analyse der Probanden ohne zuverlässige Vorhersage	168

Tabellenverzeichnis

3.1	Übersicht der verwendeten Nebenaufgaben	63
3.2	Berechnung der Grenzwerte für die Wavelet-Analyse	78
3.3	Übersicht der physiologischen Beanspruchungsindikatoren	80
4.1	Leistungsdaten auf Ebene der Geomentrien	93
4.2	Korrelationsanalyse physiologischer Parameter	95
4.3	Korrelationsanalyse auf Gruppenebene	100
5.1	Aufteilung der Signale für die Procrustes-Analyse	119
6.1	Korrelationsanalyse physiologischer Parameter Studie 2	136
7.1	Vergleich von NNI und Sakkadenlängen zwischen Kalibrierung 1 und 2	162

Abkürzungsverzeichnis

AE	Adaptionseffekt
Ari	Arithmetik
BD	Blinzeldauer
BR	Blinzelrate
DPS	Direct Parameter Specification
EEG	Elektroenzephalografie
EKG	Elektrokardiografie
EOG	Elektrookulografie
ERP	Event Related Potentials
FD	Fixationsdauer
HA	Hauptaufgabe
Lkw	Lastkraftwagen
MDS	Multidimensionale Skalierung
n	Stichprobengröße
NA	Nebenaufgabe
NASA	National Aeronautics and Space Administration
NNI	Nearest Neighbor Index
PC	PERCLOS (Percentage of Eye Closure)
Pkw	Personenkraftwagen
PNS	Parasympathisches Nervensystem
POC	Performance Operating Characteristic
PD	Pupillendurchschnittsvariabilität
PFD	Primary Flight Display

SEA	Skala zur Erfassung subjektiv erlebter Anstrengung
SWAT	Subjective Workload Assessment Technique
TLX	Task Load Index
Vis	visueller Reiz
WP	Workload Profile

1 Einleitung

Die zuverlässige Erfassung und Bewertung mentaler Beanspruchung bildet aktuell eine zentrale Forschungsfrage im arbeitswissenschaftlichen Bereich, da sie eine wichtige Voraussetzung für eine menschenzentrierte und nachhaltige Systemgestaltung ist (Packebusch, 2003; Schlick, Bruder & Luczak, 2010; Wickens, Hollands, Banbury & Parasuraman, 2013). Die Gestaltung sozio-technischer Systeme, bestehend aus einer Interaktion von Menschen und Technik, wird aufgrund der steigenden Technisierung und Digitalisierung, einhergehend mit einer kontinuierlich steigenden Menge an verfügbaren Informationen, immer bedeutender (Neudörfer, 2014). Eine beispielhafte Domäne, in der sich die steigenden Anforderungen durch die zunehmende Informationsdichte besonders auswirkt, ist die kommerzielle Luftfahrt. Der hohe Überwachungsaufwand resultierend aus dem hohem Automationsgrad wirkt sich insbesondere in sicherheitskritischen Situationen, die eine Übernahme erfordern, auf die Beanspruchung der Piloten aus (Parasuraman & Rizzo, 2007). In kritischen Flugsituationen müssen die vorhandenen Informationen sehr schnell selektiert werden, um die Störungsquelle zu identifizieren und die richtige Verfahrensweise ableiten zu können. Durch eine kontinuierliche Erfassung der mentalen Beanspruchung können beispielsweise die Menge und die Art der dargestellten Informationen auf den verschiedenen Anzeigeelementen angepasst (adaptiert) werden, um so die Suche nach relevanten Informationen in zeitkritischen Situationen erheblich zu erleichtern. Die konkreten Ziele der Gestaltung sozio-technischer Systeme werden angeführt von der Vermeidung unsicherer Situationen in der Interaktion zwischen Mensch und Technik (Fehlerprävention), was beispielsweise

durch die gestalterische Prävention von mentalen Über- und Unterforderungszuständen der Anwender erreicht werden kann. Zusätzlich wird angestrebt die sozio-technischen Systeme möglichst nachhaltig zu gestalten, sodass Menschen mit individuell verschiedenen Leistungsvoraussetzungen eine optimale und individuell fördernde Interaktion mit technischen Systemen ermöglicht werden kann. Darüber hinaus wird durch die menschenzentrierte Systemgestaltung versucht, die Synergien der Interaktion von Mensch (z. B. Flexibilität, Antizipation) und Technik (z. B. hohe Wiederholungsgenauigkeit, Schnelligkeit) möglichst effizient und effektiv zu nutzen (Schlick et al., 2010). So könnten sich beispielsweise Robotersysteme adaptiv auf den individuellen Anwender anpassen, um eine bedarfsgerechte Unterstützung bei repetitiven Tätigkeiten zu ermöglichen. Um die aufgezeigten Ziele zu erreichen und Gestaltungsmaßnahmen abzuleiten, ist es notwendig die Interaktion direkt in dem jeweiligen System zu untersuchen. Dabei ist es wichtig, die Systeminformationen sowie die individuellen Beanspruchungszustände und die Verhaltensvariabilität der Anwender zu erfassen und zu verstehen. Aus den Bereichen des kognitiven und motorischen Lernens ist bekannt, dass der Mensch überwiegend auf Reize (Informationen) reagiert anstatt zu agieren (Schnotz & Kürschner, 2007). Auch in der Mensch-Technik-Interaktion reagiert der Mensch kontinuierlich mit Handlungen auf Reize, die ihm durch das System dargestellt werden. Die Reaktionen werden dabei sehr stark durch individuelle Erfahrungen, Ziele oder den Trainingsgrad beeinflusst. Diese hohe Individualität der menschlichen Kognition ist ein wesentlicher Grund für die Verhaltensvariabilität, die in der Interaktion verschiedener Menschen mit dem gleichen System beobachtet werden kann (Sträter, 2005). Gleichzeitig wird die zuverlässige Erfassung der mentalen Beanspruchung auf der Ebene einer gesamten Stichprobe erschwert. Die zuverlässige Erfassung der mentalen Beanspruchung ist jedoch eine wichtige Voraussetzung für eine menschenzentrierte Systemgestaltung sozio-technischer Systeme. Basierend auf den individuellen Ansätzen zur Beschreibung der menschlichen Kognition aus dem kognitiven und motorischen Lernen wird im Rahmen dieser Arbeit ein Verfahren zur individuellen und hoch zuverlässigen Beanspruchungsmessung entwickelt, welches die individuellen Unterschiede in der menschlichen

Informationsverarbeitung berücksichtigt. Das Verfahren beruht dabei allein auf den physiologischen Reaktionen des Menschen, da diese, im Gegensatz zu leistungsbasierten und subjektiven Indikatoren, eine kontinuierliche Erfassung während der Anwendung mit hoher zeitlicher Auflösung erlauben (Manzey, 1998).

1.1 Zielsetzung

Die vorliegende Arbeit leistet einen Beitrag zur menschenzentrierten und nachhaltigen Systemgestaltung sozio-technischer Systeme, indem eine Vorgehensweise für die individuelle und zuverlässige Erfassung der mentalen Beanspruchung entwickelt und erprobt wird. Neben der Berücksichtigung der individuellen Unterschiede in der menschlichen Informationsverarbeitung ermöglicht das Verfahren die zuverlässige Nutzung der physiologischen Reaktionen des Menschen, ohne die simultane Erfassung leistungsbasierter oder subjektiver Indikatoren als Referenz. Neben den Vorteilen einer kontinuierlichen und nahezu einschränkungsfreien Erfassung der mentalen Beanspruchung über physiologische (hier: okkulumotorische) Parameter liegen auch Schwierigkeiten in der Interpretation vor. Zum einen unterscheiden sich die physiologischen Reaktionen verschiedener Menschen sehr stark und zum anderen werden die physiologischen Parameter oft nicht allein durch die mentale Beanspruchung, sondern beispielsweise auch durch die Schläfrigkeit beeinflusst (De Waard, 1996). Aus diesem Grund werden physiologische Indikatoren im Rahmen der Korrelationsforschung zumeist nur in Kombination mit leistungsbasierten und subjektiven Maßen eingesetzt, um über den gegenseitigen Abgleich auf die mentale Beanspruchung zu schließen (Ryu & Myung, 2005; Schwalm, 2009; Zhang, Zheng, Duan, Meng & Zhang, 2015). Ein solches Vorgehen führt dazu, dass die genannten Vorteile der physiologischen Beanspruchungsmessung nicht vollständig genutzt werden können. Darüber hinaus würde eine subjektive Erhebung der Beanspruchung den eigentlichen Arbeitsablauf stören und auch die Erfassung der Leistung ist in vielen Arbeitssystemen nicht ohne komplexe Versuchsaufbauten möglich (Wickens et al., 2013). Das Verfahren, welches in

dieser Arbeit entwickelt wird, umfasst eine Kalibrierungsaufgabe, mit der ein individuelles Beanspruchungsmodell basierend auf physiologischen Parametern für den Anwender hergeleitet wird, bevor dieser dann die zu analysierende Tätigkeit ausführt. Die Zielsetzung dieser Arbeit kann wie folgt zusammengefasst werden: Durch die Berücksichtigung bewusster und unbewusster Informationsverarbeitungsprozesse sowie der Individualität der physiologischen Reaktionen des Menschen in der Auswertung von okularen Beanspruchungsparametern wird versucht die Zuverlässigkeit der Vorhersage soweit zu steigern, dass echtzeitfähige Untersuchungen direkt in praktischen Anwendungen ermöglicht werden.

1.2 Vorgehensweise

Zur Erreichung der Zielstellung dieser Arbeit werden in dem folgenden Kapitel zunächst grundlegende Modelle der menschlichen Informationsverarbeitung erläutert. Insbesondere soll dabei auf die Differenzierung von bewussten und unbewussten Wahrnehmungs- und Verarbeitungsprozessen eingegangen werden. Beide Arten von Reaktionen des Menschen auf externe Reize (Informationen) werden stark durch individuelle Erfahrungen, Ziele und die Motivation beeinflusst, weshalb insbesondere Modelle und Ansätze vorgestellt werden, die auf die individuellen Charakteristika der Kognition eingehen. Der Fokus liegt hier auf den Ansätzen von Dehaene und Changeux (2005) sowie Fels, Bauer und Gharabaghi (2015), die unbewusst ablaufende Verarbeitungsprozesse weiter unterteilen und die unterschiedliche Bedeutung für die mentale Beanspruchung aufzeigen. In Kapitel 2.2 wird das Konstrukt der mentalen Beanspruchung anhand der grundlegenden Modelle von De Waard (1996) und Wickens (1984) beschrieben. Dabei wird eine Beziehung zwischen den individuellen bewussten und unbewussten Informationsverarbeitungsprozessen hergestellt.

In Kapitel 3 werden die verschiedenen Methoden zur Erfassung der mentalen Beanspruchung, gegliedert nach subjektiven, leistungsbasierten und physiologischen Indikatoren, beschrieben und verglichen. Kapitel 3.2 umfasst wichtige Methoden der Datenauswertung. Im Zuge der individuellen Datenauswertung wird die Procrustes-Analyse zur Mustererkennung (Gower, 1975) aus dem Bereich der Bewegungsanalyse (Osis, Hettinga, Macdonald & Ferber, 2015) auf die Auswertung physiologischer, okulärer Beanspruchungsparameter übertragen. Im Rahmen der ersten Studie (Kapitel 4) werden die verschiedenen physiologischen Parameter hinsichtlich ihrer Diagnostizität und Sensitivität für die mentale Beanspruchung auf Stichproben-, Gruppen- und individueller Ebene verglichen. Basierend auf den Erkenntnissen aus den vorherigen Kapiteln wird in Kapitel 5 eine individuelle Kalibrierungsaufgabe inklusive zweier möglicher Auswertungsverfahren entwickelt. Dabei stellt der Ansatz der individuellen Regressionsfunktionen eine Übertragung bisheriger Forschungsansätze (Ryu & Myung, 2005; Zhang et al., 2015) auf die individuelle Ebene dar, wohingegen der Ansatz der individuellen Parameterklassifikation eine neue Entwicklung basierend auf den theoretischen Erkenntnissen aus Kapitel 2 darstellt. Die experimentelle Untersuchung der Operationalisierung der Kalibrierungsaufgabe und der Vergleich beider Auswertungsverfahren erfolgt in Kapitel 6. Kapitel 7 beschreibt die letzte Studie, in der die Kalibrierungsaufgabe genutzt wird, um die mentale Beanspruchung in zwei anwendungsnahen Aufgaben anhand von den individuellen Beanspruchungsfunktionen vorherzusagen. Kapitel 8 und 9 bilden mit der Diskussion und dem Ausblick den Abschluss der Arbeit.

Die folgende Abbildung liefert zusammenfassend einen verkürzten Überblick über die Inhalte und die Struktur der Arbeit.



Kap. 2: Theoretischer Hintergrund	● menschliche Informationsverarbeitung ● mentale Beanspruchung
Kap. 3: Methoden	● Beanspruchungsmessung ● Mustererkennung ● statistische Methoden
Kap. 4: Studie 1	● Evaluation blickbasierter Beanspruchungsparameter ● differenzierte und individuelle Auswertung
Kap. 5: Entwicklung der individuellen Kalibrierungsaufgabe	● Aufgabenentwicklung ● informationstechnische Umsetzung ● Auswertungsverfahren
Kap. 6: Studie 2	● Laboruntersuchung der Kalibrierungsaufgabe ● individuelle Auswertung
Kap. 7: Studie 3	● Anwendungserprobung ● Vorhersage der mentalen Beanspruchung
Kap. 8 & 9: Zusammenfassung	● Diskussion der Ergebnisse ● Ausblick und Einsatzpotential

Abb. 1.1: Inhalt und Aufbau der Arbeit (eigene Abbildung)

2 Theoretischer Hintergrund

Die Wahrnehmung und Verarbeitung von Informationen durch den Menschen kann sowohl bewusst als auch unbewusst ablaufen. Dazu existieren verschiedene theoretische Ansätze, die mit unterschiedlichen Herangehensweisen differenzieren, ob ein informationsverarbeitender Prozess bewusst oder unbewusst abläuft. Beide Arten von Informationsverarbeitungsprozessen hängen zum einen von den äußeren Einflüssen (Arbeitsaufgabe, Umgebungsfaktoren) und zum anderen von inneren, individuellen Einflussfaktoren (Motivation, Ziele, Erfahrungen) ab. Die Auswirkungen der äußeren Einflussfaktoren auf die bewusste und unbewusste Wahrnehmung können die Art der Informationsverarbeitung wesentlich verändern, weshalb sie als mentale Beanspruchung bezeichnet werden. Im Folgenden soll zunächst die kognitive Informationsverarbeitung des Menschen beschrieben werden, bevor dann die mentale Beanspruchung eingeführt wird. Der Fokus soll dabei auf den Unterschieden zwischen bewussten und unbewussten Verarbeitungsprozessen, deren Veränderung im zeitlichen Verlauf einer Tätigkeit sowie deren Wechselwirkungen mit der mentalen Beanspruchung liegen. Aus der Vielzahl an Einflussfaktoren auf die mentale Beanspruchung einer individuellen Person werden abschließend die Forschungsfragen dieser Arbeit abgeleitet.

2.1 Kognitive Informationsverarbeitung

Ein umfassendes Verständnis der menschlichen Informationsverarbeitung, insbesondere darüber wie der Mensch mit verschiedenen Arten von Informationen umgeht, bildet eine der wichtigsten Grundlagen für die Erarbeitung von Richt-

linien für die Darstellung von Informationen und die Aufgabengestaltung in technischen Systemen (Parasuraman & Rizzo, 2007). Bereits in den frühen 60er Jahren begann der Einfluss der Kognitionsforschung auf die Ingenieurwissenschaften mit zunehmenden Studien der Technikinteraktion. Es entstand das sogenannte *Cognitive Engineering* womit sich auch die Entwicklungen von quantitativen Modellen der menschlichen Informationsverarbeitung verstärkten (Parasuraman & Rizzo, 2007). Heute existiert ein breites Spektrum an quantitativen und qualitativen Modellen, die sich im Wesentlichen durch die einzelnen Verarbeitungsschritte unterscheiden (Broadbent, 1958; Kahneman, 1973; Rasmussen, 1986; Sträter, 2005). Mittels Studien basierend auf *ERP*'s (Event-related potentials) konnte darüber hinaus ein wesentlicher Einfluss der kognitiven Informationsverarbeitung auf die mentale Beanspruchung und die Leistung festgestellt werden (Kramer, Sirevaag & Braune, 1987; Kramer, Wickens & Donchin, 1983; Sirevaag et al., 1993).

2.1.1 Modelle der Informationsverarbeitung

Es gibt eine Vielzahl an Modellen und Theorien, welche die menschliche Informationsverarbeitung mit verschiedensten Ansätzen beschreiben. Der Umfang der folgenden Modelle zeigt die vielseitigen Facetten der menschlichen Informationsverarbeitung, die insbesondere bei der Beschreibung der mentalen Beanspruchung (Kapitel 2.2) sowie zur Interpretation der empirischen Erkenntnisse immer wieder aufgegriffen werden.

Broadbent (1958) führt mit seiner *Filter-Theorie (Bottleneck-Theory)* eine der bekanntesten und aus heutiger Sicht grundlegendsten Darstellung der menschlichen Informationsverarbeitung ein (Schlick et al., 2010). Das Modell beschreibt die sequentielle Verarbeitung von Informationen, die über verschiedene Informationskanäle aufgenommen werden können. In den ursprünglichen Ausführungen von Broadbent (1958) besteht der sogenannte Filter (*Attention Selector*) aus einer Art Kippschalter, der dafür sorgt, dass nur die Informationen eines Informationskanals (bspw. Augen, Ohren) zeitgleich in den kontrollierten Verarbeitungsprozess gelangen, auf dem aufbauend dann Entscheidungen und

Handlungen ausgeführt werden. Das Prinzip des Kippschalters wird schnell weiter entwickelt, da Forschungen zeigen, dass auch eine simultane Informationsaufnahme über verschiedene sensorische Kanäle (z. B. sehen und hören) möglich ist (Abbildung 2.1).

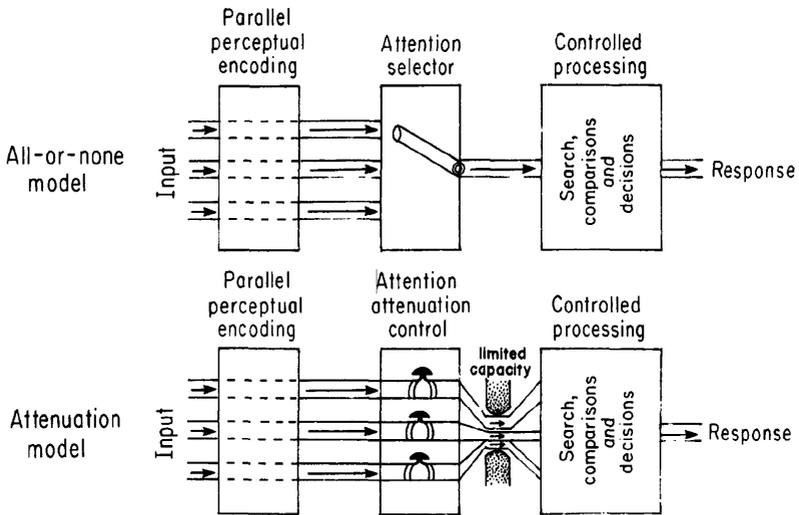


Abb. 2.1: Sensorische Verarbeitungsmodelle (Schneider & Shiffrin, 1977, S. 5)

So stellt Treisman (1960) den Filter als eine Engstelle mit begrenzter Kapazität dar. Die grundlegende Funktionalität des Filters, das Verarbeitungssystem vor Überlastungen zu schützen, bleibt erhalten, wobei jetzt Informationen verschiedener Sinnesorgane zeitgleich, wenn auch begrenzt, verarbeitet werden können. Dabei wird die Menge der eintreffenden Informationen vor dem Hintergrund der zu bearbeitenden Aufgabe priorisiert und selektiert. Dieses eingängige Modell liefert die Grundlage von vielseitigen Weiterentwicklungen der sequentiellen Verarbeitungsmodelle. Unter anderen greift Kahneman (1973) das Filtermodell in seinem sequentiellen Verarbeitungsmodell auf. Er stellt den Punkt heraus, dass der Mensch, trotz vieler eingehender Informationen, nur auf einen Stimulus reagieren (antworten) kann. Die Verortung des Flaschenhalses innerhalb der Informationsverarbeitung kann danach nicht mehr fest definiert werden. Eine

Selektion der Informationen kann so bereits während der sensorischen Wahrnehmung und Speicherung oder erst während des Schritts der Analyse auftreten. Um dieser Dynamik Rechnung zu tragen, führt Kahneman (1973) eines der ersten Ressourcenbasierten Informationsverarbeitungsmodelle ein (Abbildung 2.2).

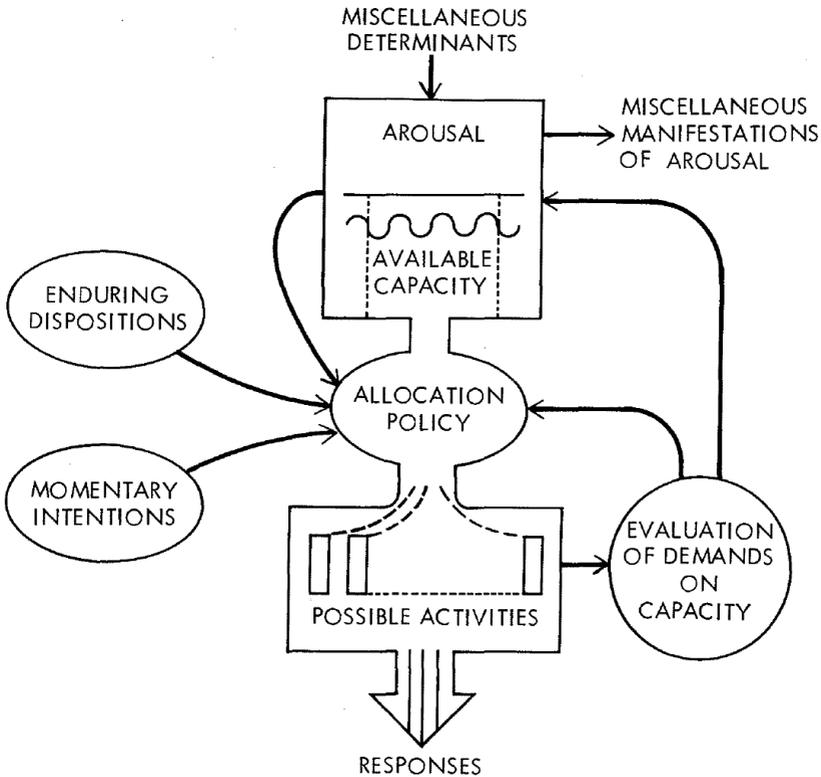


Abb. 2.2: Ressourcenbasiertes Verarbeitungsmodell (Kahneman, 1973, S. 10)

Das Kapazitätsmodell erlaubt eine parallele Informationsverarbeitung, die durch die verfügbaren Ressourcen limitiert ist. Jeder Mensch hat eine individuelle Kapazität an Ressourcen, oder Energie, die in Abhängigkeit der Konstitution, des Trainingsgrades und weiterer individueller Faktoren steht. Diese Ressourcen können für die Verarbeitung von Informationen genutzt werden (Wickens et al.,

2013). Die Menge an genutzten Ressourcen steht, zusammen mit der Stärke des Stimulus, in enger Verbindung mit der *kognitiven Aktivierung*. Die kognitive Aktivierung (*Arousal*) beschreibt den Zustand des Entdeckens, Verarbeitens oder Erlernens von Informationen im Rahmen einer Aufgabe, durch Neugier oder Interesse (Eysenck, 1982). Easterbrook (1959) stellt die Hypothese auf, dass eine steigende Aktivierung zu einer erhöhten Fokussierung der Aufmerksamkeit auf notwendige Informationen führt. Das *Yerkes-Dodson-Gesetz* konkretisiert diese Hypothese indem es das Aktivierungsniveau und die Leistungsfähigkeit in einen invertierten, parabolischen Zusammenhang setzt (Yerkes & Dodson, 1908). Die selektive Wahrnehmung steigt mit zunehmender Aktivierung. Dadurch steigt zunächst die Leistungsfähigkeit an, da irrelevante Informationen herausgefiltert werden. Ab einem gewissen, individuell unterschiedlichen, Aktivierungsniveau beginnt der Mensch jedoch auch relevante Informationen zu ignorieren, was zu einer Abnahme der Leistungsfähigkeit führt. Die Nutzung der Ressourcen ist in diesem Zustand unregelmäßig und unpräzise (Yerkes & Dodson, 1908). Die Art und die Menge der verarbeiteten Informationen sowie die kognitiven Prozesse in dem vorgestellten Modell von Kahneman (1973) hängen demnach von der Aktivität, aber auch von der Aufmerksamkeit und der Anstrengung ab. Weiter findet der Trainingsgrad eines Menschen Berücksichtigung in der Unterscheidung von neuen zu bereits bekannten Stimuli.

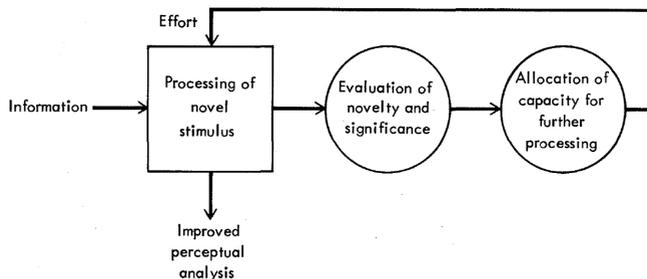


Abb. 2.3: Rekursive Verarbeitung neuer Stimuli (Kahneman, 1973, S. 44)

Kahneman (1973) ergänzt sein Modell durch eine rekursive Funktionsweise im Falle der Verarbeitung von neuen oder unbekanntem Stimuli (Abbildung

2.3). Basierend auf der Theorie des neuronalen Modells von Sokolov (1963) findet zunächst ein Abgleich statt, der prüft, ob ein Stimulus neu ist. In einem rekursiven Zyklus werden dann zusätzliche Ressourcen zur Verarbeitung dieser Informationen bereitgestellt. Sokolov (1963) zeigt in seiner Forschung, dass ein kortikaler Abgleich von neuen Stimuli mit vorhandenen Erfahrungen stattfindet bevor der Entscheidungsprozess eingeleitet wird.

Viele Ansätze der Modelle nach Broadbent (1958) und Kahneman (1973) sind auch in aktuelleren Modellen wieder zu finden. Wickens et al. (2013) beschreibt in seinem Modell zur menschlichen Informationsverarbeitung den Fluss der Informationen durch die einzelnen Verarbeitungsschritte während der Mensch eine Aufgabe ausführt (Abbildung 2.4).

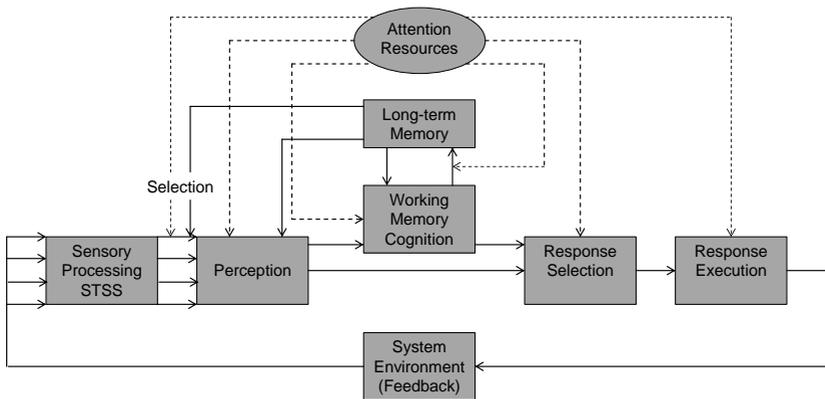


Abb. 2.4: Menschliche Informationsverarbeitungsschritte (Wickens et al., 2013, S. 4)

Informationen aus der Umgebung werden zunächst von den Sinnesorganen erfasst und temporär im Kurzzeitgedächtnis gespeichert, wobei die reine Erfassung nicht zwingend zur Wahrnehmung führt. Lediglich ein geringer Teil der erfassten Informationen wird schließlich wahrgenommen. Wickens et al. (2013) definiert dabei die Wahrnehmung über die Fähigkeit, die Bedeutung von Informationen oder Signalen über gespeicherte Erfahrungen, Fakten oder Bilder

im Langzeitgedächtnis zu bestimmen. Nach der Informationsaufnahme folgt die Information einem von zwei möglichen Wegen. Oftmals löst eine wahrgenommene Situation eine direkte Antwort aus, wobei eine Handlung aus vorhandenen Handlungsmustern ausgewählt wird. Auf der anderen Seite kann der Mensch das Arbeitsgedächtnis zur Berücksichtigung zusätzlicher Informationen nutzen. Von dem Arbeitsgedächtnis aus können Informationsbündel mit gespeicherten Inhalten des Langzeitgedächtnisses verglichen werden, um Handlungsmuster abzuleiten. Darüber hinaus beeinflussen Erfahrungen und Konzepte des Langzeitgedächtnisses auch die Wahrnehmung neuer Informationen. Diese Prozesse nennt Wickens et al. (2013) *Kognition*. Innerhalb des Modells werden zwei Vitalelemente berücksichtigt (*Feedback und Aufmerksamkeit*). Die Handlungen des Menschen beeinflussen und verändern die Umgebung, sodass sich oft neue Umweltzustände und damit auch neue Informationen ergeben, die weitere oder neue Handlungen erfordern. Während einer Aufgabe steht der Mensch in einem ständigen Feedback-Kreislauf mit seiner Umwelt. Die Aufmerksamkeit spielt innerhalb der Informationsverarbeitung zwei zentrale Rollen. Zum einem wird die Ganzheit der erfassten Informationen selektiert und zum anderen weist die Aufmerksamkeit den verschiedenen Verarbeitungsschritten eine bestimmte Menge an Ressourcen oder Energie zu. In Abhängigkeit der Aufgabe brauchen die einzelnen Verarbeitungsschritte mehr oder weniger Energie und das Angebot an Ressourcen ist begrenzt. So kann es vorkommen, dass während der Verarbeitung von einzelnen Informationen nicht genügend Ressourcen für die Bearbeitung paralleler Informationen vorhanden sind (Wickens et al., 2013). Mit dem Anspruch, die verschiedenen Wege der Informationen durch die Verarbeitungsschritte genauer zu differenzieren und zu begründen, entwickelt Rasmussen (1986) ein sequentielles Informationsverarbeitungsmodell (Abbildung 2.5). Das dargestellte Modell ist das Ergebnis einer Studie in der Leitwarte eines Atomkraftwerks. Das zu überwachende System liefert eine Vielzahl an Informationen über den Systemzustand, der sich in Abhängigkeit von bekannten als auch unbekanntem Fehlern aus dem Sollbereich bewegen kann. Die Aufgaben der Probanden umfassen im Wesentlichen das Beobachten und Identifizieren von Abweichungen sowie die Bewertung dieser, gemessen an

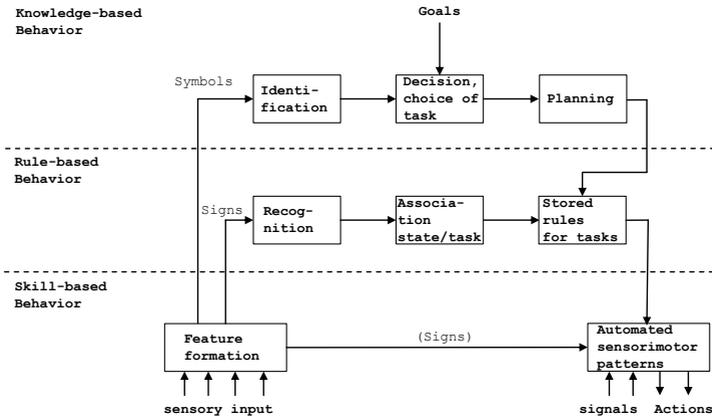


Abb. 2.6: Vereinfachte Darstellung der Kontrollebenen (Rasmussen, 1986, S. 101)

Subroutinen auf, die irgendwann automatisiert werden und genau bei dieser Art des Verhaltens abgerufen werden können. Während einer Aufgabe können mehrere verschiedene Subroutinen abgerufen werden. Die zweite Verhaltensebene beschreibt das *regelbasierte Verhalten*. Gespeicherte Prozeduren aus Erfahrungen, Konversationen oder auch Arbeitsanweisungen in ähnlichen Situationen können bewusst abgerufen und auf die aktuelle Arbeitssituation adaptiert werden. Die Regeln sind zielorientiert und oft implizit in Abhängigkeit von spezifischen Situationen gespeichert. Eine scharfe Trennung von fähigkeitsbasiertem und regelbasiertem Verhalten ist nicht immer möglich, wobei das letztere von Menschen verbalisiert werden kann. Die oberste Ebene beschreibt das *wissensbasierte Verhalten*. In unbekanntem Situationen, für die keine Subroutinen vorhanden sind, ist es für ein zielorientiertes Handeln notwendig, ein konzeptionelles und strukturiertes Modell zu entwickeln. Ziele sind zumeist explizit durch die Umgebung des Menschen definiert. Ein Lösungsvorgehen wird durch das Abwägen von verschiedenen Handlungsweisen gegen die Zielvorgaben entwickelt. Dieses Testen von Vorgehensweisen kann auf der einen Seite durch Probieren oder auf der anderen Seite durch die konzeptuelle Vorher-

sage des Ergebnisses geschehen (Erstellung eines mentalen Modells). Alle drei Ebenen sind eng miteinander verbunden und der Mensch kann während einer Aufgabe die Ebenen, auf denen er handelt, dynamisch wechseln (Rasmussen, 1986).

Alle bisher betrachteten Modelle können durch die Annahme von einzelnen, sequentiell aufeinander aufbauenden, Verarbeitungsschritten zusammengefasst werden (Schlick et al., 2010). Die einzelnen Prozessschritte stehen in einer engen Beziehung miteinander und können nicht getrennt betrachtet werden. In der Begutachtung und Untersuchung von Unfällen oder kritischen Ereignissen, in denen der Mensch mit einem technischen System interagiert, kann die Annahme von sequentiellen Verarbeitungsschritten in den genannten Modellen oft nicht mehr aufrecht erhalten werden. In dieser Art von Modellen fehlt die Berücksichtigung von Absichten, Zielen, Motivation und Einstellungen des Menschen, die seine Informationsaufnahme und -selektion stark beeinflussen (Sträter, 2005). Die Rolle von unbewussten Prozessen geht jedoch weit über die Funktion, Prozesszyklen zu verkürzen (Rasmussen, 1986) hinaus. Menschen verändern ihr Verhalten ohne dabei zwingend die einzelnen Verarbeitungsschritte zu verändern oder den Verarbeitungsprozess zu verkürzen. Der Einfluss von unbewussten Prozessen muss in Modellen der menschlichen Informationsverarbeitung stärker und detaillierter berücksichtigt werden, um Kognition vollständig beschreiben zu können (Simon, 1996; Shiffrin & Schneider, 1977; Schneider & Shiffrin, 1977). Insgesamt bieten die dargestellten Modelle eine eher reaktive Perspektive auf die menschliche Informationsverarbeitung, was sich später auch in einer ebenfalls reaktiven Systemgestaltung niederschlägt (Sträter, 2005). Weiter ist es wichtig die kognitiven Prozesse mit kontextuellen Faktoren zu verknüpfen, um den Prozess der Reflektion eigener Gedanken und die daraus entstehenden neuen Ideen zu beschreiben (Hollnagel, 2004). Die experimentelle Überprüfung der dargestellten Modelle über Such- und Reaktionsaufgaben zeigt, dass die Kognition sehr stark durch die Reflektion, die Erfahrungen und die Ziele beeinflusst wird.

Aufbauend auf den dargestellten Überlegungen zu Modellen menschlicher Informationsverarbeitung entwickelt Sträter (2005) den *kognitiven Informationsverarbeitungszyklus*. Die grundlegenden Annahmen dieses Modells basieren auf dem Prinzip der kognitiven Mühle, welches zwischen einer inneren und einer externen Welt unterscheidet (Neisser, 1976; Katzenberger, 1967). Die externe Welt beschreibt die objektiven Informationen der Umgebung, wohingegen die innere Welt das subjektive, deklarative Wissen basierend auf Erfahrungen, Zielen, Wünschen oder Einstellungen des Menschen beschreibt. Die menschliche Informationsverarbeitung gleicht permanent die externe mit der inneren Welt ab, zeitweise bewusst, aber in der Regel unbewusst. Neben dem Abgleich verändern und erneuern sich die beiden Welten kontinuierlich über Handlungen und Stimuli: der *Wahrnehmungs-Handlungs-Zyklus* (Sträter, 2005). Die kognitive Mühle bildet die Grundlage für adaptives Verhalten von Menschen (Arenius, 2009). In einem kognitiv ausbalancierten Zustand ist keine Adaption des Verhaltens notwendig. In diesem Fall liegt eine Übereinstimmung von der inneren zur äußeren Welt vor. Der Mensch agiert so lange unbewusst, bis es eine Stimulus-bedingte Störung oder Veränderung der externen Welt gibt, sodass es keine passende Entsprechung in der inneren Welt mehr gibt. Dieser Zustand, den der Mensch in kleinem Ausmaß noch fähigkeitsbasiert beheben kann, wird kognitive Dissonanz genannt. Es wird deutlich, dass sich kognitive Prozesse gleichermaßen auf beide Welten beziehen und es nach einer Dissonanz zu einem iterativen Prozess der Wiederherstellung des Gleichgewichtes kommt (Arenius, 2009; Sträter, 2005).

Der kognitive Informationsverarbeitungszyklus beschreibt im Wesentlichen die kognitive Mühle mit zentralen kognitiven Fähigkeiten, um Unterschiede zwischen der externen und der inneren Welt auszugleichen (Abbildung 2.7). Externe Informationen, die in den Zyklus gelangen, werden zu einer Aufgabe und vor dem Hintergrund der vorhandenen inneren Informationen verarbeitet. Liegt keine, oder nur eine sehr geringe Dissonanz zwischen der externen und inneren Welt vor, kann der Mensch fähigkeitsbasiert handeln und reagieren. Kommt es zu einer fehlerhaften Anpassung der Welten infolge einer großen Dissonanz, wird regelbasiert nach ähnlichen inneren Informationen gesucht,

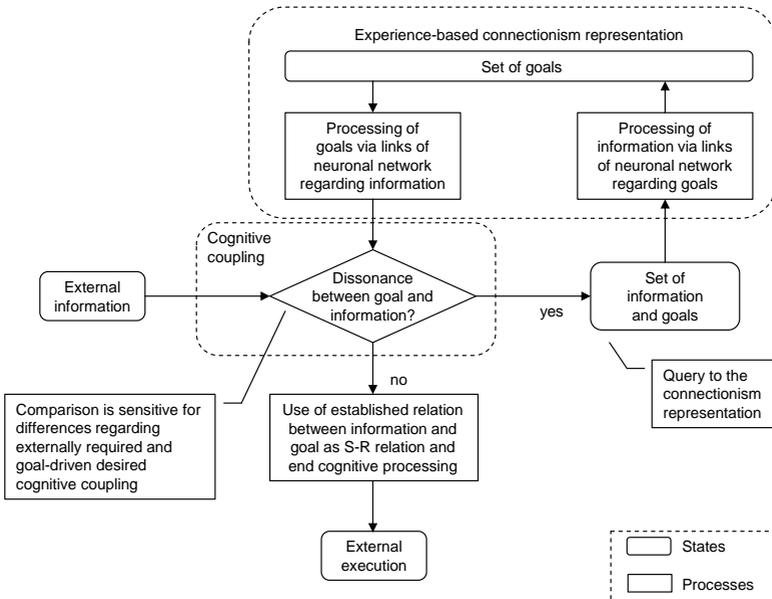


Abb. 2.7: Der kognitive Verarbeitungszyklus (Sträter, 2005, S. 79)

die die entstandene Dissonanz minimieren können. Kann die Dissonanz nicht behoben werden, sind weitere Zyklen notwendig (wissensbasiertes Verhalten). Der Zyklus ermöglicht eine Berücksichtigung der dynamischen Veränderungen von Systemparametern oder anderen Umgebungseinflüssen. Der *zentrale Vergleichler* (verortet im *limbischen System*¹) selektiert die Informationen und kann durch Intentionen und Ziele beeinflusst werden. Dieser beschreibt die Entscheidungsfindung auf verschiedenen kognitiven Ebenen (Sträter 2005). Ein großer Mehrwert dieses Modells liegt in der zuverlässigen Ableitung von Verarbeitungs- und Reaktionszeiten. Insgesamt setzt sich die Reaktionszeit aus vier Prozessen zusammen. Die *kognitive Kopplung* beschreibt die Aufnahme

¹ Das limbische System liegt, ähnlich wie ein Saum, um den Balken und das Zwischenhirn herum und bildet die Verbindung zwischen Neokortex und Hirnstamm. Die wesentlichen Aufgaben umfassen unter anderem die Emotionssteuerung und die Selektion von Informationen (Trepel, 2004, S. 203-210).

der Informationen der externen Welt und deren Verteilung in Abhängigkeit von dem Aktivierungszustand des Menschen. Die *kognitive Bindung* evaluiert wie gut die externen und inneren Informationen zueinander passen. Weiter spiegelt das *kognitive Leveln* im zentralen Vergleich und integriert die Ergebnisse des Abgleichs für weitere Vergleichsprozesse innerhalb des Verarbeitungszyklus bis eine Dissonanz behoben ist. Schließlich nimmt auch die Auslösung von Handlungen über das Kleinhirn noch Zeit in Anspruch, sodass sich bei einmaligem Durchlaufen des Verarbeitungszyklus eine Reaktionszeit von 250 Millisekunden ergibt, die in Abhängigkeit der persönlichen Eigenschaften einer Person variieren kann (Sträter, 2005).

Die dargestellten Modelle zeigen die Vielseitigkeit und die Entwicklung von Ansätzen, die menschliche Informationsverarbeitung zu erklären und zu modellieren. Im Verlauf dieser Arbeit werden die verschiedenen Modelle immer wieder aufgegriffen und mit den empirischen Erkenntnissen abgeglichen. Insbesondere der Informationsverarbeitungszyklus nach Sträter (2005) liefert durch die den kontinuierlichen Abgleich der äußeren mit der inneren Welt viele Ansatzpunkte zur Beschreibung der menschlichen Verhaltensvariabilität. Die starke Individualität der inneren Welt, die sich insbesondere während der unbewussten Informationsverarbeitung auf der Fertigkeitbasierten Ebene (Rasmussen, 1986) auswirkt, motiviert die vertiefte Betrachtung differenzierter und individueller Ansätze im nächsten Kapitel.

2.1.2 Von generellen zu differenziellen und individuellen Ansätzen

Die vorgestellten Modelle der menschlichen Informationsverarbeitung unterscheiden sich sowohl in ihren Ansätzen als auch in den einzelnen Verarbeitungsprozessen. Alle Modelle können jedoch gemeinsam in die Kategorie der generellen Modelle eingeordnet werden. Neben den Vorteilen, wie der leichten Verständlichkeit, gibt es auch erhebliche Nachteile. Die menschliche Informationsverarbeitung wird von subjektiven Faktoren, wie der Motivation, der

Zielsetzung sowie den sensorischen und motorischen Fähigkeiten, beeinflusst, was in diesen Modellen nur in geringstem Maße berücksichtigt werden kann (Driver & Mock, 1975). In dem Modell nach Rasmussen (1986) zeigt sich dies in Ansätzen bereits bei den unbewusst ablaufenden Verarbeitungsprozessen, die durch individuelle Erfahrungen geprägt und entwickelt werden können. Sträter (2005) geht dabei noch einen Schritt weiter, indem er sich mit seinem Informationsverarbeitungszyklus auf das Weltenmodell bezieht. Die innere Welt wird ebenfalls stark durch individuelle Erfahrungen und Konzepte beeinflusst. Aufgrund des allgemeinen Charakters der Modelle ist es nicht möglich, einen größeren Einfluss der individuellen Unterschiede zu berücksichtigen. Neben den generellen Ansätzen lassen sich die differenziellen und die individuellen Ansätze (Abbildung 2.8) unterscheiden (Driver & Mock, 1975). Die *individuellen Ansätze* beschreiben die kognitiven Verarbeitungs- und Entscheidungsprozesse sowie die physiologischen Reaktionen auf Ebene des individuellen Menschen im Sinne einer Einzelfallbetrachtung.

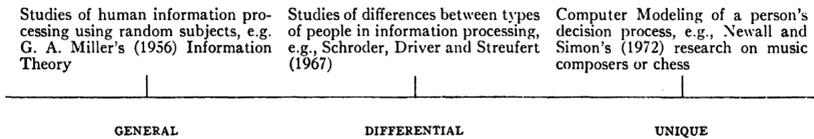


Abb. 2.8: Klassifizierung der Modelle menschlicher Informationsverarbeitung (Driver & Mock, 1975, S. 494)

Die hohe Individualität der menschlichen Informationsverarbeitung durch Erfahrungen, Ziele oder die Motivation führen zu einer deutlichen Variabilität in den Verhaltensweisen unterschiedlicher Menschen, die eine identische Aufgabe bearbeiten (Sträter, 2005). Die *generellen Ansätze* erheben den Anspruch Verarbeitungsprozesse zu beschreiben, die allen Menschen gemein sind, um so allgemein gültige Gesetzmäßigkeiten abzuleiten. Innerhalb der generellen Ansätze (Stichprobenebene), führt die interindividuelle Verhaltensvariabilität häufig zu starken Abweichungen einzelner Probanden von dem Durchschnitt der Stichprobe. Dadurch entstehen, auch bei ausreichenden Stichprobenumfängen, große Streuungen der Daten, die meist nur ansatzweise erklärt werden können.

Sozio-technische Systeme, die basierend auf den resultierenden, generalisierten Erkenntnissen entwickelt und gestaltet werden, bereiten häufig Nutzungsprobleme und hohe Sicherheitsrisiken (Driver & Mock, 1975). Die individuellen Verhaltensweisen werden durch ein solches System nicht voll umfassend unterstützt, weshalb beispielsweise wichtige Informationen in kritischen Situationen häufig von Operateuren nicht wahrgenommen werden. Die Informationen werden nicht dort präsentiert, wo der individuelle Nutzer diese erwartet. Individuelle Modelle können die menschliche Informationsverarbeitung mit sehr hoher Präzision beschreiben, sodass Algorithmen diese beispielsweise replizieren können. Solche Ansätze sind bereits in Schachprogrammen zu finden, die sich nach einer bestimmten Spielzeit auf das individuelle Verhalten des Nutzers einstellen. Ein Nutzer kann so immer optimal durch das Spiel gefordert, aber auch unterstützt, werden. Solche Ansätze können allerdings nicht generalisiert werden (Newell & Simon, 1972). Neben den individuellen Ansätzen gibt es die *differenziellen Ansätze*, die ebenfalls auf die interindividuellen Unterschiede der menschlichen Kognition eingehen. Die Menschen werden dazu in Gruppen eingeteilt, anstatt jede Person individuell zu betrachten, um allgemeingültige Aussagen treffen zu können. Diese Modelle bilden einen Kompromiss zwischen der Einfachheit der allgemeinen und der Präzision der individuellen Modelle. Die Schwierigkeit der differenziellen Ansätze liegt in der Auswahl der Kriterien zur Einteilung der verschiedenen Gruppen. Driver und Mock (1975) schlagen im Wesentlichen zwei Methoden vor, um eine Person zu kategorisieren. Eine wahrnehmungsorientierte Methode, ausgehend von Fragebögen, unterscheidet die Art, in der der Mensch neuen und unbekanntem Informationen gegenüber reagiert. Durch diese Art der Kategorisierung sollen Unterschiede in der Wahrnehmung von Informationen aufgedeckt werden und zu einer höheren Genauigkeit der Wahrnehmungsbeschreibung beitragen. Innerhalb der zweiten vorgeschlagenen entscheidungsorientierten Methode hingegen wird erfragt, wie der Mensch Entscheidungen basierend auf unbekanntem Informationen trifft. Beide Ansätze resultieren jedoch nicht in trennscharfen Gruppen, da Personen häufig zwischen den Kategorien liegen (Driver & Mock, 1975).

Die Ausführungen von Driver und Mock (1975) zeigen, dass es für eine zuverlässige und nachhaltige Systemgestaltung notwendig ist, die individuellen Unterschiede in der menschlichen Informationsverarbeitung zu berücksichtigen, wobei dies einem hohen Aufwand gegenüber steht. Auch die Informationsverarbeitungsmodelle nach Rasmussen (1986) und Sträter (2005) indizieren deutlich, dass es individuelle Unterschiede gibt, die zum einen auf Erfahrungen und Konzepten und zum anderen auch auf dem inneren Weltmodell beruhen können. Um eine individuelle Modellierung weiter zu verfolgen und in einem breiteren Anwendungsfeld nutzbar zu machen, ist es notwendig, die menschliche Kognition noch detaillierter zu betrachten und die Gründe der Unterschiede zu finden. Ein vielversprechender Ansatzpunkt liegt dabei in der Unterscheidung in bewusste und unbewusste Verarbeitungsprozesse (Neumann, 1990), die bereits auf Basis der vorgestellten Modelle mit den individuellen Unterschieden der menschlichen Informationsverarbeitung in Beziehung gebracht werden können (Sträter, 2005; Rasmussen, 1986). Schon in den 70er Jahren weisen Schneider und Shiffrin (1977) mit der Theorie über die zwei Arten der menschlichen Informationsverarbeitung (*Two-process theory*) auf den hohen individuellen Einfluss der unbewussten Prozesse auf die Verarbeitung von Stimuli hin. Sie nutzen Reaktionstests und Suchaufgaben, um eine Unterscheidung von automatischen (unbewusst) und kontrollierten (bewusst) Verarbeitungsprozessen zu zeigen (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Dabei werden die durch zutreffende Reize ausgelösten Aktivierungen oder der Abruf von gelernten Elementen des Langzeitgedächtnisses als automatische Prozesse klassifiziert. Das Langzeitgedächtnis wird als ein großer permanenter Speicher bestehend aus einer Vielzahl an überwiegend passiven Knoten verstanden. Ein automatischer Prozess aktiviert eine Folge von Knoten (Handlungssequenz; Handlungsweise) als Reaktion auf eine spezifische Information, die entweder aus der äußeren oder der inneren Welt stammt, und überträgt diese Sequenz in das Kurzzeitgedächtnis. Dabei werden stets situative und kontextuelle Bedingungen einbezogen. Die Aktivierung eines automatischen Verarbeitungsprozesses bedarf einer gewissen Menge an konsistenten Erfahrungen oder Konzepten, die der aktuellen Situation entsprechen. Erst nachdem beispielsweise ein Pilot einen

Landeanflug mehrmals manuell durchgeführt hat, ist es möglich, den Prozess des Ausgleichens von kleinen Turbulenzen zu automatisieren, sodass nebenbei das Navigations- oder Bordsystem bedient werden kann. Einmal gelernt, sind automatische Prozesse nur sehr schwierig zu unterdrücken oder zu modifizieren. Sträter (2005) beschreibt diesen unbewussten Einfluss von Erfahrungen und Konzepten ebenfalls mit dem Modell der gefestigten Erfahrungen. Ohne die bewusste Kontrolle des Menschen und ohne eine Belastung der Kapazitätsgrenzen der menschlichen Informationsverarbeitung darzustellen, laufen automatisierte Prozesse ohne zwingende Aufmerksamkeit ab. Allerdings können automatische Prozesse die laufende bewusste Verarbeitung von Reizen oder Informationen beeinflussen, indem sie die Informationen mit bestimmten Konzepten verknüpfen. So ist es dem Menschen möglich, einen gezielten Reiz aus einer Menge an Reizen erfolgreich und schnell zu extrahieren (Shiffrin & Schneider, 1977). Die Beschreibungen zeigen den Einfluss von subjektiven automatischen Prozessen auf die bewusste Informationsverarbeitung und damit die Individualität der menschlichen Informationsverarbeitung. Kontrollierte Prozesse hingegen erfordern die Aufmerksamkeit sowie die Kontrolle des Menschen und werden darüber hinaus durch die Kapazitätsgrenzen eingeschränkt. Aus der temporalen Perspektive sind die zuletzt genannten Prozesse sehr hoch aufgelöst und können zu einer sehr schnellen Aktivierung von informationsverarbeitenden Prozessen führen (Schneider & Shiffrin, 1977). Allerdings kann laut Schneider und Shiffrin (1977) simultan nur ein bewusster Prozess ablaufen, sodass sich eine sequentielle Abfolge ergibt, die stark an die Informationsverarbeitungsmodelle nach Broadbent (1958), Kahneman (1973) und Wickens et al. (2013) erinnert. Die Kapazitätsgrenzen von kontrollierten Prozessen werden durch das Kurzzeitgedächtnis determiniert, welches 7 ± 2 Informationseinheiten (oder *Chunks*) verarbeiten kann (Miller, 1956). Auch Experimente mit verdeckten Wörtern bestätigen die vorgestellte Unterscheidung in bewusste und unbewusste Verarbeitungsprozesse (Marcel, 1983b, 1983a). Die fünf zentralen Experimente von Marcel (1983b, 1983a) zeigen, dass beispielsweise ein großer Teil der visuellen Informationsverarbeitung auf unbewusster Ebene abläuft. Die unbewussten Wahrnehmungsprozesse sind notwendig, um auch nicht-aufgabenrelevante In-

formationen aus der Umgebung aufzunehmen. Diese Informationen werden zum einen gebraucht, um die Aufmerksamkeit gezielt auf die relevanten Informationen zu lenken und zum anderen, um die Orientierung sowie das Gleichgewicht zu ermöglichen. Die Experimente von Marcel (1983b, 1983a) zeigen replizierbar die Existenz der unbewussten Verarbeitungsprozesse bei visuellen Stimuli. Auch die Experimente von Coltheart (2000) mit visuell stark eingeschränkten Patienten (Deep Dyslexia) zeigen, dass die Patienten Stimuli zwar bewusst nicht erkennen (Störung des scharfen Sehens), aber ihre Reaktionen trotzdem von der semantischen Bedeutung des Reizes beeinflusst werden. Darüber hinaus zeigen auch die Studien mit Suchaufgaben von Kihlstrom (1987) den Einfluss von unbewussten auf bewusste Verarbeitungsprozesse. Aufgabenrelevante Stimuli werden in Abhängigkeit ihrer Umgebung unterschiedlich gut und schnell erkannt. Die verschiedenen Umgebungen werden dabei von den Probanden nicht bewusst wahrgenommen. Die Erkenntnisse der Studie bestätigen die Aussagen von Marcel (1983b, 1983a), dass eine Abhängigkeit beider Arten von Verarbeitungsprozessen vorliegt. Allerdings weist Kihlstrom (1987) auch darauf hin, dass der Einfluss der unbewussten Prozesse, wie in den Studienergebnissen gezeigt, mit den bislang vorliegenden Beschreibungen der jeweiligen Prozesse nicht vollständig erklärt werden kann. Die genannten Definitionen von automatischen Prozessen basieren auf der Ausprägung durch Erfahrungen und Konzepte und somit auf den ursprünglichen Verarbeitungsebenen nach Rasmussen (1986). Auch eine Studie von Spelke, Hirst und Neisser (1976), in der Probanden einen unbekanntem Text lesen sollten, während sie ein Diktat schreiben, zeigt, dass die Leistung mit der Zeit stark ansteigt. Nach sechs Wochen Übung steigert sich die Leistung der Probanden um 80%, wobei sie sich kaum noch an die Wörter erinnern können, welche sie diktiert niederschreiben. Aufgrund der erfahrungsbasierten Automatisierung von einigen Aufgabenteilen, ist eine bewusste Replikation der Inhalte nicht mehr möglich. Eine Beeinflussung beider Prozessarten ist aber schwer zu erklären, weshalb Kihlstrom (1987) neben der erfahrungsbasierten Ausprägung von automatischen Prozessen, die umgebungsspezifische Ausprägung definierte. Automatische Prozesse werden somit vor den bewussten Prozessen durch die Aufnahme bestimmter Stimuli,

unabhängig von der Intention der Person, ausgelöst. Mit dieser Zweiteilung soll die Abhängigkeit beider Prozesse erklärt werden.

Ein weiterer Ansatz, bewusste und unbewusste Prozesse und ihren Zusammenhang zu beschreiben, liefert das DPS-Modell (*direct parameter specification*) nach Neumann (1990). Er beginnt die Beziehung zwischen der Wahrnehmung eines Stimulus und der Auslösung einer Handlung durch die Gegenüberstellung von mentaler Repräsentation und direkter Parameterextraktion zu diskutieren. Die allgegenwärtige Grundsätze sagen, ein Stimulus muss zuerst verarbeitet werden, bevor eine Handlung oder eine Reaktion abgeleitet werden kann. Experimente, in denen Reaktionszeiten erfasst wurden, indizieren jedoch, dass es zwei Arten der Informationsverarbeitung geben muss (Neumann, 1990). Auf der einen Seite muss ein Stimulus bewusst wahrgenommen werden, um bewusste Entscheidungen und Reaktionen herbeizuführen. Auf der anderen Seite kann ein Stimulus auch direkt eine motorische Reaktion auslösen. Letzteres tritt keinesfalls nur in bekannten und geübten Situationen auf, sondern ist eher die Regel, insbesondere in komplexen Situationen. Ein Bewusstsein ist demnach nicht notwendig, um eine Reiz-Reaktions-Beziehung aufzubauen, dieses ist teilweise der Grund für die großen Differenzen zwischen physiologischen und leistungsbasierten Daten in Experimenten unterschiedlichster Arten. Auf Stichprobenebene müssen nicht zwingend Korrelationen beider Größen vorliegen, da die unbewussten Prozesse hoch individuell sind. Die direkte Parameterspezifikation (DPS) beschreibt den Vorgang, in dem eingehende Informationen direkt Handlungsparameter festlegen können, ohne oder bevor eine mentale Entsprechung (Erfahrungen oder Konzepte) aktiviert wird (Neumann, 1990). Ein wichtiger Indikator für diese Annahme ist der Dissoziationseffekt: ein Stimulus spezifiziert Handlungsmuster, die nicht vollständig mit dem mentalen Modell übereinstimmen (Sträter, 2005). In einer Studie von Kohler (1951) trugen Probanden 10 Tage lang eine prismatische Linse, was das Sehvermögen auch für längere Zeit danach negativ beeinflusst. Es wird deutlich, dass einige Aufgaben besser und andere schlechter bearbeitet werden können. Nach einem Tag können die Probanden wieder etwas greifen und bereits sechs Tage nach

dem Linseneinsatz kann die Mehrheit der Probanden schon wieder Skifahren. Nach dieser Zeit konnten die Probanden zwar Objekte handhaben, diese aber noch nicht vollständig wahrnehmen oder gar beschreiben. Diese Studie bekräftigt die Ansätze des DPS-Modells.

Dehaene, Changeux, Naccache, Sackur und Sergent (2006) greifen die kontroverse Diskussion über die verschiedenen Formen der bewussten Verarbeitungsprozesse erneut auf und berichten von drei möglichen Ursprüngen: entweder dem Thalamus² und dem Hirnstamm³ (Paus, 2000), den ersten visuellen Bereichen (Lamme, 2003) oder den höheren präfrontalen und parietalen Assoziationsbereichen⁴ (Dehaene & Changeux, 2005). Eine der wichtigsten Grundvoraussetzungen für bewusste Verarbeitungsprozesse ist aber die Vigilanz. Nur in einem vollständig wachen Zustand können bewusste Prozesse ablaufen. Darüber hinaus reicht eine rein visuelle Aktivierung noch nicht aus, um den Stimulus bewusst wiederzugeben. Die Forschungen von Tse, Martinez-Conde, Schlegel und Macknik (2005) zeigen einen engen Zusammenhang zwischen dem primären visuellen Cortex⁵ (V1) und auch weiten Teilen des Neocortex (V2 - V5). Diese kortikalen Aktivierungen können mittels ERPs (Event-related potentials) sowohl bis zu 180 ms nach einem Stimulus, als auch bei Personen erfasst werden, die einen Stimulus nicht bewusst wahrgenommen haben. Weiter können bewusste Prozesse durch die Erweiterung der Gehirnaktivitäten auf höher assoziative Bereiche mit langen Verbindungen zu Wahrnehmungsbereichen charakterisiert werden (Dehaene et al., 2006). Bei der Integration der assoziativen Bereiche, resultierend aus der bewussten Verarbeitung von Informationen, können zwei Veränderungen nachgewiesen werden: Die Informationen können

² Der Thalamus besteht aus einzelnen Kerngebieten im Bereich des Zwischenhirns und ist unter anderem für die Integration zerebellärer Informationen zuständig (Trepel, 2004, S. 170-173).

³ Der Hirnstamm umfasst das Mittelhirn, die Brücke (Pons) und das verlängerte Mark (Trepel, 2004, S. 108-122).

⁴ Der parietale Assoziationskortex ist unter anderem für die Sprache und das abstrakte Denken zuständig. Der präfrontale Kortex übernimmt unter anderem Aufgaben der räumlichen Orientierung (Trepel, 2004, S. 213-222).

⁵ Der visuelle Kortex ist im Occipitallappen verortet und als Teil des visuellen Systems verantwortlich für die visuelle Wahrnehmung (Trepel, 2004, S. 224-227).

für lange Zeit, unabhängig von der Dauer des Stimulus, auf Abruf gehalten werden, aber auch schnell an weitere Bereiche des Gehirns weitergereicht werden (Dehaene et al., 2006). Diese neuronale Beschreibung von bewussten Verarbeitungsprozessen liefert eine wichtige Vorarbeit für die Differenzierung von unbewussten Prozessen (Abbildung 2.9).

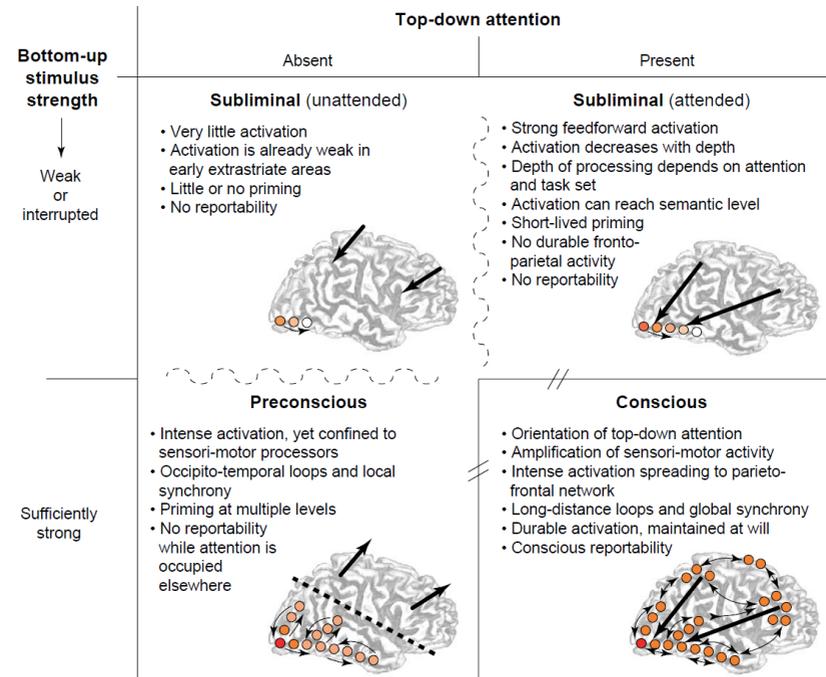


Abb. 2.9: Typen der unbewussten Verarbeitung (Dehaene & Changeux, 2005, S. 206)

Für die Abgrenzung der unbewussten Prozesse bedarf es noch der Klärung des Zusammenhangs zwischen bewussten Prozessen und der Aufmerksamkeit. Bewusste Informationen können wiedergegeben und im Gedächtnis archiviert werden. In Situationen geteilter Aufmerksamkeit zeigt sich, dass auch unverdeckte Wörter nicht immer bewusst erfasst werden. Die Beziehung von Aufmerksamkeit, Stimulus und Bewusstsein ist sehr komplex, da Aufmerksamkeit auch Bottom-Up (stimulus-basiert) ausgelöst werden kann. Dies bedeutet,

dass die Aufmerksamkeit durch einen plötzlichen und unerwarteten Reiz auf diesen gelenkt wird. Durch diese Verschiebung der Fokussierung, z. B. des zentralen Blickfeldes, wird die Wahrscheinlichkeit der bewussten Wahrnehmung der zuvor aufgabenrelevanten Informationen verringert (Dehaene et al., 2006). Die genannten Ausführungen führen zusammen zu einer Unterscheidung in drei Typen von unbewussten Prozessen (Abbildung 2.9). Subliminale Prozesse existieren sowohl in Anwesenheit als auch in Abwesenheit der Aufmerksamkeit und sind somit unabhängig von den Strategien der Person. Eine verbale Wiedergabe ist bei keinem der beiden Prozesse möglich. Diese finden in unbewusst ablaufenden Prozessen der Informationsverarbeitung statt. Bei den vorbewussten Prozessen liegt bereits eine ausreichende stimulus-basierte Aktivierung für einen bewussten Verarbeitungsprozess vor, allerdings sind diese zeitlich vorgelagert und es besteht keine ausreichende Aufmerksamkeit. Vorbewusste Prozesse hängen jedoch nicht von der Stärke eines Stimulus ab, wie auch Studien über das aufmerksamkeitsbasierte Blinzeln zeigen (Dehaene & Changeux, 2005). Im Gegensatz zu früheren Ansätzen (Schneider & Shiffrin, 1977; Kihlstrom, 1987), die unbewusste Prozesse unabhängig von Top-Down Prozessen betrachten, erlaubt das Modell nach Dehaene und Changeux (2005) eine flexible und adaptierbare Analyse unbewusster Verarbeitungsprozesse. Eine klassische Trennung beider Arten der Informationsverarbeitung lässt nur in sehr spezifischen Situationen einen Einfluss von unbewussten auf bewusste Prozesse zu. Ohne diesen Einfluss könnte eine zielgerichtete Handlung oder Reaktion auf Stimuli nur durch das Ignorieren von vielen beeinflussenden Handlungsalternativen ausgeführt werden (Kiefer, Adams & Zovko, 2012). Automatische Prozesse benötigen eine vorgelagerte Konfiguration des kognitiven Systems, abhängig von komplexen Top-Down-Faktoren, wie Aufmerksamkeit, Intention oder der Aufgabe selbst. Auch die Theorie der direkten Parameterspezifikation nach Neumann (1990) zeigt, dass Informationen nur unbewusst verarbeitet werden können, wenn sie den aktuellen Intentionen entsprechen. Aus den aufgeführten Ansätzen leitet Kiefer et al. (2012) das Aufmerksamkeits-Sensibilisierungs-Modell (*attentional sensitization model*) ab. Das Modell suggeriert, dass automatische Prozesse, genau wie kontrollierte Prozesse, abhängig von den zur Verfügung stehenden

kognitiven Ressourcen sind und zusätzlich empfänglich für Top-Down Prozesse (Repräsentationen der aktuellen Aufgabe) sind. Dabei erlaubt das strategische und bewusste Verarbeiten von Informationen eine höhere Adaptionfähigkeit und Flexibilität der Top-Down-Kontrolle, wobei zwischen der präventiven und reaktiven Kontrolle unterschieden werden muss. Lediglich bewusst wahrgenommene Reize können eine reaktive Handlung auf kontinuierliche Stimuli auslösen, da unbewusste Prozesse keine weiteren strategischen Verarbeitungsschritte festlegen können. Eine Top-Down-Kontrolle von unbewussten Prozessen benötigt zwangsläufig die Berücksichtigung von aktuell aktivierten Zielen (Kiefer et al., 2012). Um die Annahmen zu prüfen, verwenden Kiefer et al. (2012) eine wahrnehmungsbasierte Induktionsaufgabe während ERPs und subjektive Fragebögen erhoben werden. Die aufmerksamkeitsbasierte Sensibilisierung (Top-Down) von unbewussten Prozessen belegt die Adaptionfähigkeit des kognitiven Systems. Die Informationsverarbeitung kann hinsichtlich der Verfolgung von beabsichtigten Zielen optimiert werden. Aufgabenrelevante Stimuli werden priorisiert, während irrelevante Reize ignoriert werden können. Beides läuft sowohl auf der bewussten als auch auf der unbewussten Ebene ab. Unbewusste Prozesse können das Risiko des Einflusses von nicht aufgabenrelevanten Informationen auf das Verhalten deutlich verringern (Kiefer et al., 2012).

Zusammenfassend kann festgehalten werden, dass unbewusste Verarbeitungsprozesse nicht nur auf manifestierte Erfahrungen und Konzepte zurück zu führen sind, sondern auch direkt durch unbekannte Reize ausgelöst werden können (Kihlstrom, 1987; Neumann, 1990). Sie beeinflussen in hohem Maße die Bearbeitung von Aufgaben und bestimmte Ausprägungen benötigen, wie auch die bewusste Informationsverarbeitung, kognitive Ressourcen (Dehaene et al., 2006; Kiefer et al., 2012). Somit stellen automatisierte Prozesse einen essentiellen Teil der menschlichen Informationsverarbeitung dar, die in den vorgestellten Modellen der Informationsverarbeitung nur geringe Berücksichtigung finden. Die hohe Individualität der unbewussten Verarbeitungsprozesse steigert somit nochmals die Bedeutung von differenziellen oder individuellen Ansätzen der

Modellierung. Insbesondere vor dem Hintergrund der Beanspruchungsmessung ist eine Berücksichtigung von stark individualisierten unbewussten Prozessen notwendig, um die oft vorhandene Varianz zwischen Leistungsdaten, subjektiven Empfinden und den physiologischen Reaktionen beschreiben zu können (Dehaene & Changeux, 2005).

2.2 Mentale Beanspruchung

Die mentale Beanspruchung beschreibt die kognitive Reaktion des menschlichen Informationsverarbeitungssystems auf äußere Belastungen in Abhängigkeit von seinen persönlichen Voraussetzungen und den individuellen Bewältigungsstrategien (Packebusch, 2003; Niederl, 2007). Bei der Definition von Beanspruchung ist eine Differenzierung von zwei zentralen Begriffen notwendig: Belastung und Beanspruchung (Abbildung 2.10). Die Belastung (Task Load) beschreibt die objektiven Anforderungen durch die Aufgabe oder die Umgebung, die extern auf den Menschen einwirken. Die Belastungen können von dem Menschen selbst nicht beeinflusst werden. Es existiert eine Vielzahl von Belastungsfaktoren, die durch das Arbeitssystem (bspw. Lärm, Beleuchtung, Klima) und die konkrete Arbeitsaufgabe (bspw. Anzahl und Komplexität von Informationen, Darstellungsart, Bedienelemente) auf den Menschen wirken (Schlick et al., 2010).

Die Beanspruchung (Workload) beschreibt die Auswirkungen der Belastungsfaktoren auf den individuellen Menschen. Dabei hängt die subjektiv empfundene Beanspruchung von vielen individuell verschiedenen Faktoren (bspw. Konstitution, Erfahrungen, Strategie, Wissen) ab. Diese Faktoren bestimmen zum einen die Art der Informationsverarbeitung und zum anderen die zur Verfügung stehenden Ressourcen (Hilburn & Jorna, 2001). Ähnlich wie bei der menschlichen Informationsverarbeitung selbst, gibt es ein breites Spektrum an Ansätzen, die versuchen, die mentale Beanspruchung zu beschreiben. Für diese Ausarbeitung werden die Modelle nach Kahneman (1973), Wickens (1984) und De Waard (1996) gewählt. Diese beschreiben die mentale Beanspruchung basierend auf

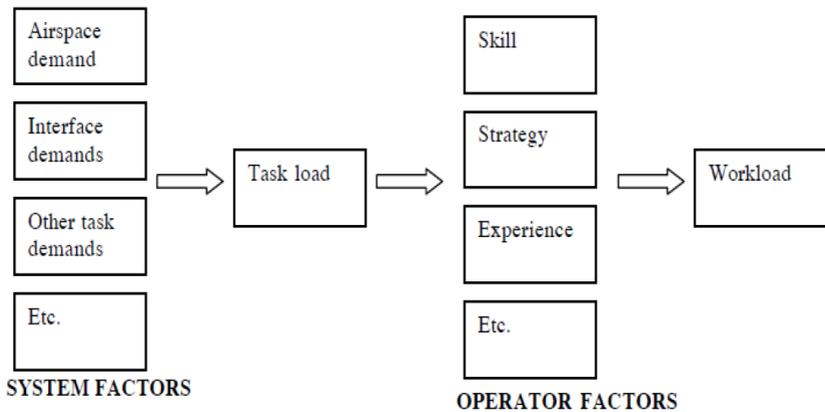


Abb. 2.10: Task Load und Workload (Hilburn & Jorna, 2001, S. 386)

den verfügbaren kognitiven Ressourcen. Die detaillierten Beschreibungen und Abgrenzungen werden nachfolgend dargestellt.

2.2.1 Theorie der mentalen Beanspruchung

Die Kapazität an Ressourcen ist das obere Limit der Informationsverarbeitungsfähigkeit des Menschen. Die Kapazität und die Verfügbarkeit an Ressourcen unterliegen dabei einer gewissen Elastizität in Abhängigkeit der Belastungen (Kahneman, 1973): ein Anstieg der Aufgabenanforderungen kann, bis zu einem gewissen Grad, zu einer Steigerung der Ressourcenmobilisierung führen. Das Verhältnis von bereitgestellten Ressourcen und der Leistung in einer Aufgabe wird, bis alle Ressourcen investiert sind, oft als linear angenommen (Yerkes & Dodson, 1908). Generell lassen sich *ressourcenlimitierte* und *datenlimitierte* Aufgaben unterscheiden. In ressourcenlimitierten Aufgaben kann der Mensch keine zusätzlichen Ressourcen mehr einsetzen und die Leistung bleibt stabil. In datenlimitierten Aufgaben führt ein zusätzliches Investieren von Ressourcen, aufgrund der limitierten Datenqualität, zu keiner Leistungssteigerung (Kahneman, 1973). Die Kapazitätstheorien stehen in einem engen Zusammenhang mit rechnerischen und energetischen Mechanismen der Informationsverarbeitung. Bewusste und unbewusste Verarbeitungsprozesse laufen in mehreren Schritten

ab, von denen jeder unterschiedliche Ressourcen benötigt. Es können drei Arten von energetischen Ressourcen unterschieden werden: *Arousal*, *Activation* und die kompensatorische Ressource *Effort* (Mulder, 1986; Sanders, 1983). Das Arousal beeinflusst die Merkmalsextrahierung und die Activation im Wesentlichen die motorische Organisation. Der Effort (Anstrengung) wird dann aktiviert, wenn der Zustand des Operateurs zu stark von dem erforderlichen Zustand abweicht, das Arbeitsgedächtnis wird benötigt. Der aktuelle Zustand wird an den Zielzustand angeglichen und den Stressoren kann erfolgreich entgegengewirkt werden. Dieser Umstand fördert die Ausbildung von individuellen Strategien. Insgesamt können zwei Arten von Effort unterschieden werden, derjenige für die kontrollierte Informationsverarbeitung und derjenige, der notwendig ist, um den aktuellen energetischen Zustand der Ressourcen zu verändern (Mulder, 1986; Sanders, 1983).

De Waard (1996) beschreibt Beanspruchung als die aufgabenspezifische und individuell erlebte Belastung. Individuelle Fähigkeiten, die Motivation, die Stimmung und die angewendeten Strategien beeinflussen die wahrgenommene Belastung. Beanspruchung kann als die individuelle Spezifikation der Informationsverarbeitungskapazität unter Berücksichtigung der Arbeitsabfolgen und der persönlichen Einschränkungen verstanden werden. Gleiche *Anforderungen* (Demands) führen zu einem individuellen Beanspruchungsniveau. Dabei steht die *Komplexität* einer Aufgabe in direktem Zusammenhang mit den Anforderungen. Sie steigt zusammen mit der Anzahl der für eine Aufgabe benötigten Verarbeitungsschritte an. Beide Konstrukte sind extern durch die Aufgabe gegeben, doch unterliegen sie den individuellen, intrinsischen Zielen des Menschen. Die *Schwierigkeit* (Difficulty) einer Aufgabe bezieht sich auf die erforderliche Anstrengung (Effort) des Individuums (De Waard, 1996). Mentale Beanspruchung ist ein relatives Konzept und hängt von der Relation zwischen den Anforderungen und der Menge an Ressourcen ab, die der Mensch aktivieren kann oder will. Menschen können während einer Aufgabe ihr Verhalten ändern: sie können beispielsweise ihre Strategien und Ziele anpassen, mit schlechteren Leistungen zufrieden sein oder komplett aufgeben. Die Anstrengung kann demnach variieren, während die Leistung gleich bleibt. Diese Verhaltensänderungen sind der

Grund für den häufig großen Unterschied zwischen Leistungsmessungen und subjektiven oder physiologischen Maßen für die mentale Beanspruchung (De Waard, 1996).

Wickens (1984) entwickelt die *multiple Ressourcentheorie*, um die Relation zwischen Belastungen und der Beanspruchung weiter zu konkretisieren. Die Ursprünge seines Modells liegen in den sequentiellen Verarbeitungsmodellen mit Engpässen, wie beispielsweise das vorgestellte Modell nach Broadbent (1958). Auch Kahneman (1973) spricht in seinem Modell von einer limitierten Kapazität, die auf verschiedene Aufgaben aufgeteilt werden kann. Mentale Ressourcen werden für die Bearbeitung von Aufgaben mit einem spezifischen Anforderungsniveau aktiviert. Nicht genutzte Ressourcen können dann für weitere Tätigkeiten eingesetzt werden. Steigt dann jedoch der Ressourcenbedarf einer Aufgabe, so kommt es mehr und mehr zu einer Inferenz zwischen den Aufgaben. Das Kontinuum zwischen Aufgaben, die sehr viele Ressourcen verlangen und Aufgaben, die nahezu keine Ressourcen benötigen, kann ansatzweise mit dem Verhaltenskontinuum (fertigkeitbasiert, regelbasiert und wissensbasiert) nach Rasmussen (1986) beschrieben werden. Eine Vielzahl an Experimenten untersucht den Zusammenhang zwischen der Anforderungshöhe einer Primäraufgabe und der Leistung in einer Nebenaufgabe. Viele Ergebnisse indizieren, dass Interferenzen zwischen den Aufgaben nicht auf die Höhe der Anforderungen in der Primäraufgabe zurückzuführen sind (Wickens, 1984). Das Modell der multiplen Ressourcen setzt genau an diesem Punkt an. Ein großer Teil der Leistungsvarianz in Dual-Task-Situationen lässt sich weder auf die Aufgabenschwierigkeit noch auf die Vorgehensweise der Ressourcenaktivierung (favorisierte Aufgaben) zurückführen. Differenzen in der Effizienz der Zeitaufteilung zwischen zwei Aufgaben können auf Unterschiede in den Anforderungen für verschiedene Verarbeitungsstrukturen zurückgeführt werden (Wickens, 2002). Verschiedene Strukturen der Informationsverarbeitung werden von unabhängigen Ressourcen unterstützt (Wickens, 1984). Die Effizienz der Zeitaufteilung zwischen zwei Aufgaben ist besser, wenn unterschiedliche Wahrnehmungsmodalitäten und Verarbeitungsprozesse rekrutiert werden. Ein Bei-

spiel ist die Informationsaufnahme über die Augen oder die Ohren (Kantowitz & Knight, 1976). Wickens (1984) zeigt mit weitläufigen Analysen variierender Modalitäten in Dual-Task-Situationen, dass sich die unterschiedlichen Paarungen signifikant voneinander unterscheiden. Aus diesen Untersuchungen entsteht das vier-dimensionale multiple Ressourcenmodell (Abbildung 2.11), welches mit vier kategorialen und dichotomen Dimensionen die Varianz bezüglich der Effizienz in der Zeitaufteilung während Dual-Task-Situationen erklärt (Wickens, 2002).

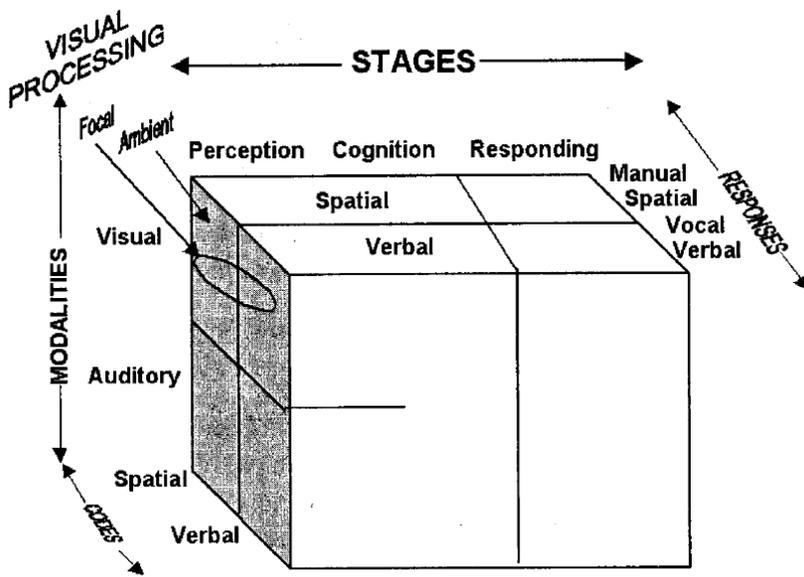


Abb. 2.11: Das multiple Ressourcenmodell (Wickens, 2002, S. 163)

Jede der vier Dimensionen besitzt zwei diskrete Ausprägungen und kann mit verschiedenen physiologischen Mechanismen assoziiert werden. Fordern zwei Aufgaben die gleiche Ausprägung einer gegebenen Dimension, so werden diese zu einem höheren Grad interferieren, als Aufgaben, die unterschiedliche Ausprägungen adressieren. Die erste Dimension (*Stages*) bezieht sich auf die einzelnen Verarbeitungsschritte. Die Ressourcen, die für die Wahrnehmung und Verarbeitung benötigt werden, unterscheiden sich von denen, die für die

Handlungsplanung und -ausführung (Antworten) rekrutiert werden. Variationen bezüglich der Schwierigkeit im Übermitteln von Antworten an ein System haben keinen Einfluss auf die Leistung in einer Wahrnehmungsaufgabe und umgekehrt. Als Beispiel führt Wickens (2002) die Arbeit eines Fluglotsen an, bei welcher die Art der Rückmeldung (manuell oder verbal) an die Flugzeuge keinen Einfluss auf die Wahrnehmung des Luftraums hat. Auch Sträter (2005) entkoppelt in seinem Informationsverarbeitungszyklus die Wahrnehmung und Verarbeitung von der Handlungsauswahl und -durchführung. Bezüglich des Antwortverhaltens zeigt das Gesetz nach Hick (1952) einen logarithmischen Anstieg der Reaktionszeit in Abhängigkeit der Anzahl an Auswahlmöglichkeiten für Eingaben in ein System. Die zweite Dimension unterscheidet die Wahrnehmungsmodalitäten (visuell und auditiv). Eine intermodale Zeitaufteilung ist effizienter als eine intramodale Zeitaufteilung, was weniger auf verschiedenen Ressourcen, als auf peripheren Faktoren beruht. Zwei visuelle Reize, die separiert dargestellt werden, erfordern zusätzliche Blickbewegungen, was nach gewisser Zeit durch adäquate Blickstrategien kompensiert werden kann. Sind die Reize nah beieinander, so können sie sich gegenseitig beeinflussen oder überlagern. Auch auditive Reize können sich leicht gegenseitig überlagern (Wickens, 2002). Die dritte Dimension bezieht sich auf die visuellen Wahrnehmungskanäle. Es wird das foveale vom peripheren Sehen unterschieden (Sträter, 2005). Beide Prozesse werden von verschiedenen Strukturen im Gehirn gesteuert und unterstützen eine effiziente Zeiteinteilung. Der Mensch kann beispielsweise problemlos während des Laufens lesen, was auf unterschiedliche Ressourcen hinweist, die die jeweiligen Prozesse unterstützen (Wickens, 2002). Während der Mensch mit dem fovealen Bereich liest, wird die Umgebung peripher wahrgenommen, sodass ein stabiles Laufen möglich ist. Die letzte Dimension unterscheidet zwischen analogen (räumlichen) und kategorialen (symbolischen) Prozessen. Diese Unterscheidung kann mit den Gehirnhemisphären begründet werden, wobei sich beide Prozesse auf unterschiedliche Ressourcen beziehen. Manuelle Eingaben sind in den meisten Fällen räumlich (Steuern von Fahr- oder Flugzeugen) und unterscheiden sich von verbalen Antworten (sprechen). Während einer Fahraufgabe zeigt das Antworten auf eine Nebenaufgabe mit der

freien Hand höhere Interferenzen mit der Hauptaufgabe (Fahrzeugführung) als das verbale Antworten auf die gleiche Nebenaufgabe (Wickens, 2002).

Das multiple Ressourcenmodell unterscheidet demnach zwischen verschiedenen Ressourcen, die von unterschiedlichen Ausprägungen der vier Dimensionen angesprochen werden, was eine Vorhersage der Leistung und der Interferenzen zwischen Aufgaben in Dual-Task-Situationen ermöglicht. Die einzelnen Ressourcen werden in Abhängigkeit von der Aufgabenschwierigkeit zugewiesen und die mentale Beanspruchung repräsentiert die Menge der zugewiesenen Ressourcen unter Berücksichtigung des individuellen Zustandes und der angewandten Strategie (Wickens, 2002).

Ursprünglich definiert Meister (1976) drei Bereiche A, B und C (Abbildung 2.12), die die Beanspruchung mit der Leistung in Verbindung bringen und für jede Ressource des multiplen Ressourcenmodells angewendet werden können. In Bereich A liegt eine geringe Beanspruchung bei hoher Leistung vor, die unabhängig von Variationen in den Anforderungen sind. Bereich B beschreibt den Abfall der Leistung einhergehend mit einem Anstieg der Beanspruchung infolge weiter zunehmender Anforderungen. Die Erfassung der Leistung in einer Primäraufgabe ist nur in diesem Bereich ein sensitives Werkzeug zur Erfassung der mentalen Beanspruchung.

Auch physiologische und subjektive Verfahren repräsentieren in diesem Bereich die mentale Beanspruchung. Bereich C wird durch ein extremes Anforderungsniveau charakterisiert, welches die Leistung mit zunehmenden Anforderungen auf ein Minimum reduziert. Fragebögen und physiologische Reaktionen zeigen in diesem Bereich eine klare Überbeanspruchung, wohingegen eine Leistungserfassung keinerlei Rückschlüsse auf die Beanspruchung zulässt (Meister, 1976). Basierend auf dem Yerkes-Dodson-Gesetz ist es notwendig, das Modell um eine vierte Region D zu erweitern (Abbildung 2.12). Der Bereich D zeigt die Effekte von monotonen Aufgaben. Geringe Aufgabenanforderungen über eine gewisse Zeit können, aufgrund einer Reduktion der Ressourcenkapazität, zu steigenden Anforderungen und Beanspruchungen führen. Zusätzlich führt

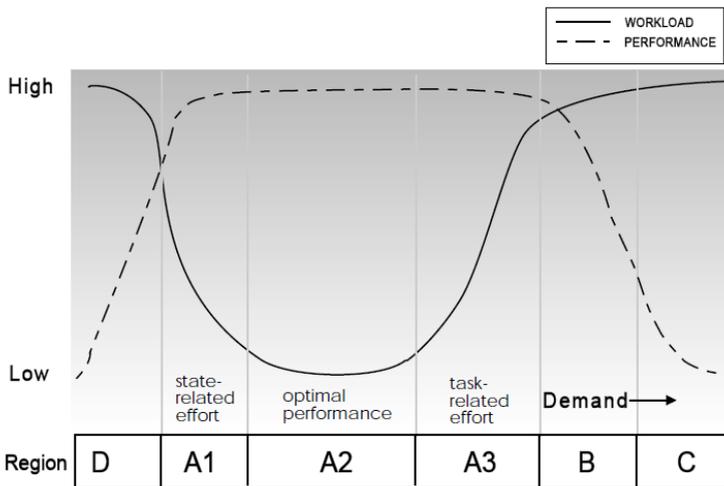


Abb. 2.12: Beanspruchung und Leistung (De Waard, 1996, S. 24)

ein Zustand der Langeweile zu einem erhöhten Ressourcenbedarf bei gleichbleibender Leistung (De Waard, 1996). Auch nach der Beschreibung dieses Modells bleibt die Frage: Ab wann ist Beanspruchung zu hoch? Innerhalb dieses Modells kann eine Höchstgrenze der Beanspruchung eingeführt werden (De Waard, 1996). Mit dem Verweis auf einen merklichen Abfall der Leistung und einen Beanspruchungsanstieg auf subjektiven Skalen, kann diese Grenze zwischen den Bereichen A und B verortet werden. Allerdings wird es nach genauer Betrachtung notwendig, den Bereich A in drei Bereiche A1, A2 und A3 zu teilen (Abbildung 2.12), da der Bereich selbst durch die Leistung definiert wird. Innerhalb von Bereich A2 kann der Operator leicht mit den steigenden Anforderungen umgehen und die Leistung bleibt unverändert hoch. In Bereich A3 zeigt die Leistung noch immer keinen Abfall, jedoch kann der Operator dieses Niveau nur mit einer gesteigerten Anstrengung halten. In dynamischen Aufgaben können kurzzyklische Spitzen der Anforderungen in diesem Bereich nicht immer ohne gesundheitliche Folgen ausgeglichen werden, weshalb die rote Linie der Beanspruchung zwischen den Bereich A2 und A3 gelegt werden sollte. Somit bezieht sich die Höchstgrenze auf die Beanspruchung und nicht

auf den Einbruch der Leistung. Im Bereich A1 beginnen die Auswirkungen der Monotonie und der Operateur muss sich auch hier mehr anstrengen, um keinen Abfall der Leistung zu riskieren. Es wird notwendig, eine zusätzliche Untergrenze der Beanspruchung zwischen den Bereichen A1 und A2 einzuführen (De Waard, 1996). Generell muss bei diesem Modell berücksichtigt werden, dass gleiche Anforderungen (Belastungen) bei jedem Menschen zu unterschiedlichen Beanspruchungen führen können.

2.2.2 Individuelle Adaptionstrategien

Werden die Erkenntnisse aus Kapitel 2.1.2 mit den vorgestellten Modellen der mentalen Beanspruchung kombiniert, so wird deutlich, dass sich die mentale Beanspruchung zum einen individuell, zum anderen aber auch über die Dauer einer Aufgabe hinweg stark unterscheiden kann. Neben bewussten Verarbeitungsprozessen gibt es eine Vielzahl von unbewussten (automatisierten) Prozessen, die ebenfalls auf mentale Ressourcen zurückgreifen (Kiefer et al., 2012; Dehaene & Changeux, 2005). Die Top-Down ablaufenden unbewussten Prozesse sorgen für eine starke Individualisierung der Verarbeitungsprozesse und somit auch der mentalen Beanspruchung. Die von Sträter (2005) gezeigte ständige Wechselwirkung von äußerer und innerer Welt postuliert darüber hinaus eine starke Variabilität des Verhaltens über die Dauer einer Aufgabe hinweg.

Schon Bainbridge (1997) beschreibt es als Notwendigkeit, dass sich Operateure in komplexen Arbeitssituationen an Veränderungen der Umgebung, das eigene Verständnis der Aufgabe und die Planung und Organisation ihres eigenen Verhaltens anpassen. Diese Anpassungsprozesse wirken sich direkt auf die Effektivität der Aufgabenerfüllung aus. Kernmechanismen des menschlichen Verhaltens sind dabei der Aufbau eines situativen Verständnisses und die Flexibilität sich an Veränderungen der Aufgabe und die Entwicklung von Erfahrungen anzupassen (Bainbridge, 1997). Der hohe Grad an Adaptionfähigkeit des menschlichen Verhaltens benötigt situationsabhängige Verarbeitungsmodelle. Die vorgestellten Modelle der Informationsverarbeitung nach Broadbent

(1958), Rasmussen (1986) und Wickens et al. (2013) stellen die Ausgabe (Entscheidung, Handlung) oft nur als einfache Abwandlung der Eingabe (Wahrnehmung) dar, wodurch keine adäquate Beschreibung der adaptiven Varianz möglich ist. Der Informationsverarbeitungszyklus nach Sträter (2005), basierend auf dem Weltenmodell, berücksichtigt zu einem gewissen Anteil schon den Einfluss vorangegangener Entscheidungen und Handlungen auf die aktuelle Handlungsauswahl. Die Aufgabenorganisation, der Umgang mit Interferenzen und die Adaptierbarkeit benötigen einen umfassenden Überblick über die gesamte Arbeitssituation in Relation zu den eigenen Zielen. Erfahrungen und entsprechendes Wissen spielen dabei eine zentrale Rolle. Aus der vorgestellten Adaptionfähigkeit zieht Bainbridge (1997) bestimmte Konsequenzen für das Verhalten und die Verarbeitungsprozesse in komplexen und dynamischen Aufgaben. Während der Aufgabendurchführung erstellen sich Menschen eine Reihe von Verknüpfungen im Kurzzeitgedächtnis (eine Repräsentation des eigenen Verständnisses der Aufgabenstruktur), welche einen Abgleich von zukünftigen Situationen und ihren Plänen ermöglichen. Eine optimale Handlung berücksichtigt den erwarteten Zustand des Systems, wobei ein dauerhafter Abgleich und eine dauerhafte Überarbeitung der Planung stattfindet. Oft ist eine Berücksichtigung der eigenen Handlungen notwendig, um die Evolution von Veränderungen nachvollziehen zu können. Dabei entstehen starke individuelle Unterschiede. Manche Menschen wechseln sequentiell zwischen verschiedenen Aufgabenzielen, wohingegen andere simultan, mit einer adaptiven Gewichtung von Zielen, arbeiten. Operateure nutzen den wesentlichen Teil ihrer Anstrengung zur Bearbeitung der eigentlichen Arbeitsaufgabe und weitere, kleinere Anteile für das Verständnis der Schnittstelle und Nebentätigkeiten. Eine gleiche Aufgabe variiert in der notwendigen Anstrengung in Abhängigkeit der Schnittstellengestaltung, den Erfahrungen und den zur Verfügung stehenden Arbeitsmethoden. Menschen entwickeln sehr schnell eine vorläufige Strategie zur Erreichung von bestimmten Zielen. Jede Person hat eine individuelle Strategie, die sich über die Aufgabe hinweg sehr unterschiedlich entwickeln oder verändern kann. Schließlich führt jede Strategie zu einer spezifischen Beanspruchung, die somit auch über die Aufgabendauer variiert. Komplexe

Situationen können darüber hinaus die neue Entwicklung von Strategien erfordern. Dazu kann der Mensch vorhandene Strategien aus ähnlichen Situationen (Erfahrungen) adaptieren oder eine kognitive Verbindung zwischen dem aktuellen Zustand und dem Zielzustand bilden. Die gewählten Strategien sind dabei stark abhängig von erwarteten Eintrittswahrscheinlichkeiten, Kosten und dem Nutzen. Basierend auf den vorstehenden Überlegungen hat Bainbridge (1997) den kontextabhängigen Verarbeitungszyklus eingeführt (Abbildung 2.13). Die Informationsverarbeitung wird fortlaufend anhand kontextueller Faktoren und den Ergebnissen der eigenen Handlungen aktualisiert. Dabei werden nicht nur die eigenen Handlungen, sondern auch deren mögliche Auswirkungen auf den zukünftigen Status des Systems evaluiert (Bainbridge, 1997).

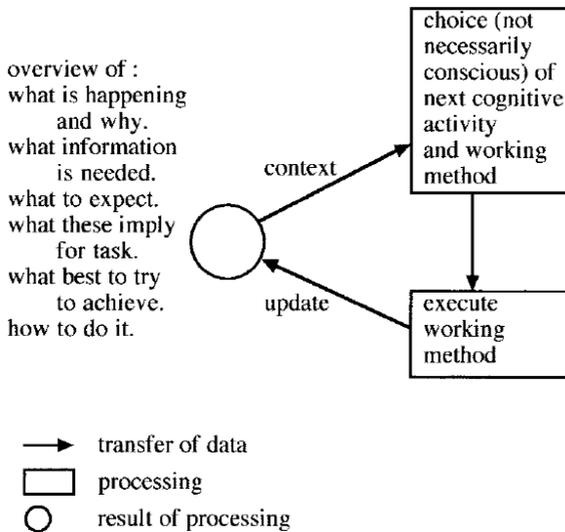


Abb. 2.13: Der kontextabhängige Verarbeitungszyklus (Bainbridge, 1997, S. 354)

An diesem Punkt kann festgehalten werden, dass sich die mentale Beanspruchung nicht nur zwischen Personen, sondern auch über den Verlauf der Aufgabe stark unterscheidet. Diese Veränderungen beeinflussen nicht nur die bewussten Verarbeitungsprozesse, sondern im gleichen Maße auch die unbewusste Informationsverarbeitung.

Sperling und Doshier (1986) beschäftigen sich in ihren Forschungen mit den individuellen Strategien in kognitiven Wahrnehmungsaufgaben. Es ist möglich, einen Teil der Strategien extern zu beobachten (z. B. Blickbewegungen bei der Suche nach wichtigen Informationen). Andere Strategien, wie die Optimierung des Verhaltens unter Berücksichtigung der aktuellen Aufgabe und der individuellen Kapazitäten, können nicht direkt erfasst werden. Beide Arten von Strategien, ob bewusst oder unbewusst, beeinflussen die Entscheidungsfindung, das Verhalten und die Ressourcennutzung (Sperling & Doshier, 1986; Kiefer et al., 2012). Der Grad der Variabilität von Strategien über die Bearbeitungszeit einer Aufgabe hinweg unterscheidet sich individuell sehr stark. Generell können Menschen mit einer einfachen Strategie (*Pure Strategies*) von Menschen mit wechselnden Strategien (*Strategy Mixtures*) unterschieden werden. Eine einfache Strategie kann als die Auswahl definiert werden, nach der der Mensch einen bestimmten Punkt auf einer bestimmten Leistungskurve (*Performance Operating Characteristic - POC*) erreicht. In Situationen, die lediglich eine Auswahl erfordern, wird von einer unidimensionalen Strategie gesprochen, wohingegen mehrere unabhängige Entscheidungen als multidimensionale Strategien bezeichnet werden können. In den meisten gängigen Aufgaben nutzen Menschen jedoch eine Vielzahl an Strategien zwischen denen sie während der Aufgabe wechseln, um so einen optimalen Beanspruchungsgrad (Verhältnis von Ressourcenbedarf und Leistung) zu erreichen. Ein Wechsel der Strategie wirkt sich jedoch zunächst in einem kurzzeitigen Abfall der Leistung während der Wechselphase aus. Stehen zwei unabhängige Strategien S_a und S_b zur Bearbeitung einer Aufgabe zur Verfügung, führen diese zu zwei unterschiedlichen Zielpunkten auf der Leistungskurve (Abbildung 2.14).

Der Mensch nutzt für einen Anteil a der Aufgabe die Strategie S_a und für die übrigen Aufgabenteile b die Strategie S_b (Abbildung 2.14). Ein $a = 0\%$ oder ein $a = 100\%$ stellen dabei einfache Strategien dar, wobei $0\% < a < 100\%$ eine Mischstrategie indiziert. Im Falle eines Strategiewechsels fällt die Leistung entlang der Verbindungslinie von a und b (Punkt x in Abbildung 2.14). Die Gesamtleistung der Aufgabe entspricht demnach dem gewichteten Durchschnitt

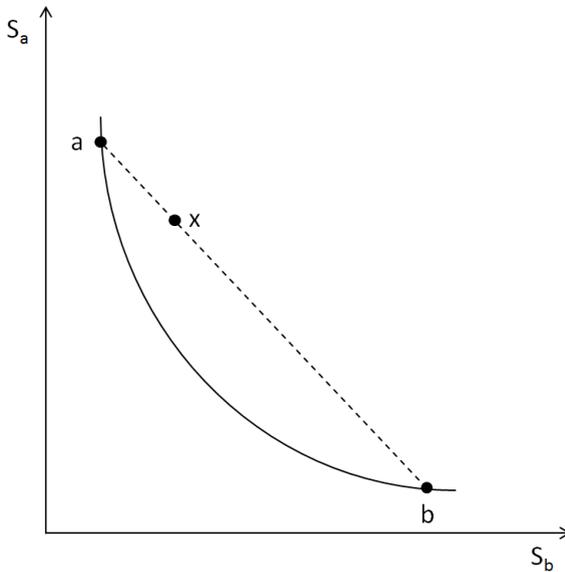


Abb. 2.14: Einfache und gemischte Strategien. Punkt a entspricht dem Anteil der Aufgabenbearbeitung in dem Strategie S_a anwendet wird. Punkt b entspricht dem Anteil der Aufgabenbearbeitung in dem Strategie S_b anwendet wird. Punkt x ist ein Beispiel für einen Strategiewechsel zwischen S_a und S_b (Sperling & Doshier, 1986, S. 48)

der Ergebnisse von S_a und S_b . Beide Arten von Strategien können aus physiologischen und leistungsbezogenen Daten abgeleitet werden (Sperling & Doshier, 1986).

Insbesondere in hoch beanspruchten Zuständen während einer Aufgabe werden Wechsel von Strategien für die vorübergehende Erhaltung der Leistung genutzt (Bereich A3 in Abbildung 2.12). Das aktive Ressourcenmanagement und die Mobilisation von zusätzlicher Anstrengung erlaubt es dem Menschen sich an gegensätzliche Ziele oder veränderte Belastungen anzupassen. Hockey (1997) oder Matthews und Campbell (1998) bezeichnen diese leistungserhaltenden Adaptionen der Strategien als *Coping* und unterteilen diese Regulationen in *aktive* (bewusste) und *passive* (unbewusste) Adaptionen. Aktive Adaptionen können durch den steigenden Regelaufwand des Kurzzeitgedächtnisses mit dem

Ziel, die Anstrengungen an veränderte Aufgabenanforderungen anzupassen, beschrieben werden. Die passiven Adaptionen sind eng mit den unbewussten Verarbeitungsprozessen verbunden. Menschen verändern ihre Strategien unbewusst und bewusst, um eine neue Gewichtung ihrer Ziele zu schaffen. So können sie beispielsweise die Geschwindigkeit oder die Genauigkeit der Aufgabebearbeitung vernachlässigen, um primäre Ziele zu einem höheren Grad zu erfüllen. Hockey (1997) beschreibt die aktiven Adaptionen als unabhängig von den Anforderungen, wohingegen passive Adaptionen erst ab einer bestimmten Anforderungshöhe auftreten. Diese Unterscheidung kann jedoch aufgrund der beschriebenen bewussten und unbewussten Verarbeitungsprozesse (Dehaene & Changeux, 2005; Kiefer et al., 2012) nicht bestätigt werden und auch Matthews und Campbell (1998) treffen diese Unterscheidung nicht. Die beschriebenen Strategiewechsel finden nicht nur zwischen zwei unabhängigen Vorgehensweisen, sondern auch zwischen verschiedenen anstrengenden Ausprägungen der Informationsverarbeitung statt. Ein einfaches Beispiel ist der *speed-accuracy Trade-Off* in Reaktionsaufgaben. In komplexeren Aufgaben wird häufig eine Verringerung der Arbeit des Kurzzeitgedächtnisses oder der Wechsel zu weniger ressourcenintensiven Formen der Kontrolle beobachtet (Hockey, 1997). Sperandio (1978) zeigt diese Art der Strategiewechsel in einer Studie mit Fluglotsen, die eine steigende Anzahl von Flugzeugen in ihrem Luftraum kontrollieren und handhaben müssen. Ab einer individuell unterschiedlichen Anzahl von Flugzeugen beginnen die Lotsen ihre Strategie dahingehend zu verändern, dass die individuellen Routenplanungen aufgegeben werden und eine standardisierte Prozedur für alle Flugzeuge angewendet wird. Das Primärziel der Sicherheit im Luftraum wird sichergestellt, wobei weniger priorisierte Ziele (Staffelung, Passagierkomfort) vernachlässigt werden (Sperandio, 1978). Die leistungserhaltenden Strategiewechsel können ebenso unbewusst stattfinden. Hockey (1997) zeigt diese Kontrollreaktionen des Menschen. Während Sequenzen, in denen ein Leistungsparameter innerhalb von Eingriffsgrenzen liegt, reguliert der Mensch weniger als in den Grenzbereichen, ohne es bewusst verbalisieren zu können.

Iqbal, Adamczyk, Zheng und Bailey (2005) untersuchen den Verlauf men-

taler Beanspruchung über die Dauer von interaktiven, hierarchischen Aufgaben (Wegplanung, Bearbeitung von Dokumenten), um Zeitpunkte zu definieren, an denen Zusatzinformationen den Nutzer am wenigsten von der eigentlichen Aufgabe ablenken. Die hierarchischen Aufgaben bestehen aus mehreren Unteraufgaben, die jeweils eine andere Priorisierung haben und eine unterschiedliche Vorgehensweise (Strategie) erfordern. Zunächst kann bestätigt werden, dass verschiedene Unteraufgaben zu unterschiedlichen mentalen Beanspruchungen führen, was die Annahmen über die Auswirkungen bewusster und unbewusster Strategiewechsel auf die mentale Beanspruchung bestätigt. Weiter sinkt die mentale Beanspruchung an den Grenzen oder Übergängen von einer Unteraufgabe zu der nachfolgenden. Diese Erkenntnis steht vordergründig im Widerspruch zu den Erkenntnissen von Sperling und Doshier (1986), die einen Abfall der Leistung bei Strategiewechseln postulieren. Der Operateur tätigt in der Zeit des Wechsels von Unteraufgaben und Strategien keine Eingaben, bearbeitet also keine Aufgaben. Dieser Zustand kann sich schon vor dem Beenden der bearbeiteten Unteraufgabe einstellen und die aktiven Eingaben können unter einer Latenzwirkung erst nach Beginn des nächsten Aufgabenteils wieder einsetzen. Demnach sinkt sowohl die Leistung, als auch die mentale Beanspruchung in diesem Bereich. Aus diesem Grund finden auch die Erkenntnisse von Sperling und Doshier (1986) in diesen Ergebnissen Bestätigung. Darüber hinaus berichten Iqbal et al. (2005) von stärkeren Abfällen bezüglich der mentalen Beanspruchung bei wichtigeren Unteraufgaben. Dies kann mit der höheren Priorität der Ziele begründet werden, die eine deutlichere Abgrenzung von vorhergehenden Aufgaben erfordert. Auch auf der gleichen Hierarchieebene unterscheiden sich die Veränderungen der mentalen Beanspruchung an den Grenzen der Unteraufgaben stark, was die Individualität der Beanspruchung erneut indiziert. Genauer betrachtet bestehen interaktive Aufgaben aus hierarchischen und rekursiven Schleifen der Zielformulierung, Ausführung und Evaluation. So werden, auch unabhängig von induzierten Veränderungen auf Ebene der Unteraufgaben, kontinuierlich neue Grenzen von Zielen geschaffen, die sich ähnlich verhalten. Die mentale Beanspruchung sinkt hier direkt bevor eine neue Zielformulierung beginnt (Iqbal et al., 2005). Die Erkenntnisse

decken sich mit dem Informationsverarbeitungszyklus nach Sträter (2005) und dem kontextbasierten Modell nach Bainbridge (1997).

Gagnon et al. (2011) forschen auf dem Gebiet der adaptiven Automatisierung und zeigen mit den Ergebnissen einer Dual-Task-Studie die Abhängigkeit der mentalen Beanspruchung zum einen von der Aufgabe und zum anderen von dem Individuum (Erfahrung, verfügbare mentale Ressourcen, Bearbeitungszeit). Die Ergebnisse unterstreichen den individuellen Ansatz und zeigen die Notwendigkeit einer individuellen Kalibrierung für die Ableitung von funktionalen Zusammenhängen. Eine weitere aktuelle Studie von Fels et al. (2015) zeigt die Möglichkeit, Beanspruchungsprofile über einen differenziellen Ansatz zu extrahieren. In einer Feedback-Aufgabe mit drei Schwierigkeitsstufen wird die mentale Beanspruchung mit Hilfe von EEG-Signalen und dem NASA-TLX erfasst. Die Stichprobe kann anhand der NASA-TLX-Angaben in Gruppen eingeteilt werden, die sich signifikant voneinander unterscheiden. Jedes dieser Cluster kann mit einem anderen Beanspruchungsprofil beschrieben werden und auch die EEG-Signale (Thetaband) unterschieden sich signifikant zwischen den Gruppen. Gleiche Aufgaben führen demnach zu unterschiedlich wahrgenommenen Beanspruchungen, dieses bekräftigt die zuvor genannten Ausführungen erneut.

2.3 Ableitung der Fragestellung

Die menschliche Informationsverarbeitung unterliegt, unabhängig von äußeren Anforderungen, starken inter- und intraindividuellen Unterschieden, die auf Erfahrungen und Konzepten, aber auch auf Handlungsweisen und Zielformulierungen beruhen. Bereits frühe Modelle der Informationsverarbeitung weisen auf die individuell verschiedenen mentalen Ressourcen hin, die für die Bearbeitung einer oder mehrerer Aufgaben zur Verfügung stehen (Schneider & Shiffrin, 1977; Kahneman, 1973). Der Informationsverarbeitungszyklus nach Sträter (2005) spricht darüber hinaus von einer kontinuierlichen gegenseitigen Beeinflussung der äußeren und der inneren Welt in Abhängigkeit von individuellen

Merkmale. Die Unterscheidung in bewusste und unbewusste Verarbeitungsprozesse macht diese individuellen Unterschiede noch deutlicher, da unbewusste Prozesse im Wesentlichen auf Erfahrungen und Konzepten der Menschen beruhen (Schneider & Shiffrin, 1977; Marcel, 1983b; Kihlstrom, 1987). Jüngere Forschungen von Dehaene und Changeux (2005) unterteilen die unbewussten Prozesse und zeigen, dass bestimmte Ausprägungen ebenfalls kognitive Ressourcen verbrauchen und somit auch für die mentale Beanspruchung relevant sind. Im Bereich der Forschungen um die mentale Beanspruchung haben die individuellen Strategien und Vorgehensweisen einen großen Einfluss (Bainbridge, 1997; Hockey, 1997; Matthews & Campbell, 1998). Menschen adaptieren ihre Strategien über die Bearbeitungszeit einer Aufgabe hinweg unterschiedlich stark mit individuellen Zielen. Auch diese Strategiewechsel geschehen sowohl bewusst, als auch unbewusst (Hockey, 1997). Der Grad der Adaptionen von Strategien kann nicht nur auf äußere Anforderungen zurückgeführt werden, sondern wird genauso von individuell variablen Zielen und Motiven determiniert (Iqbal et al., 2005). Um die mentale Beanspruchung von Menschen in der Interaktion mit technischen Systemen hoch zuverlässig vorhersagen zu können, bedarf es des Schrittes von einem generellen zu einem individuellen Ansatz (Fels et al., 2015; Driver & Mock, 1975; Galy, Cariou & Mélan, 2012). Bereits Gagnon et al. (2011) beschreiben den Bedarf einer individuellen Kalibrierung, um Beanspruchung auf einem Niveau der Zuverlässigkeit zu bestimmen, welches für die Systemgestaltung notwendig ist.

Basierend auf den genannten theoretischen Vorarbeiten stellt sich die zentrale Forschungsfrage dieser Arbeit: kann die mentale Beanspruchung mit einem individuellen Ansatz hoch zuverlässig über blickbasierte, physiologische Reaktionen bestimmt werden? Diese Fragestellung untergliedert sich in die drei zentralen Fragen, die adressiert werden sollen:

- Kann eine individuelle Kalibrierung entwickelt werden, die es ermöglicht, einen individuellen Ansatz der Beanspruchungsmessung unabhän-

gig von der Arbeitsaufgabe anzuwenden? Insbesondere steht hier die Gestaltung und Entwicklung einer solchen Aufgabe im Vordergrund.

- Sind individuelle Adaptionsstrategien in den physiologischen Reaktionen erkennbar? Verschiedene zeitliche Auflösungen von physiologischen Signalen sollen verwendet werden, um die Verläufe der mentalen Beanspruchung über die Aufgabendauer hinweg abzugrenzen.
- Können individuelle Beanspruchungsmuster oder -profile auf Basis der Kalibrierungsaufgabe erstellt werden? Es soll geprüft werden, inwieweit diese Modelle aufgabenunabhängig angewendet werden können.

3 Methoden

Nachfolgend werden die in dieser Arbeit verwendeten Methoden dargestellt. Dabei werden zunächst die verschiedenen Messinstrumente für die mentale Beanspruchung beschrieben. Diese teilen sich in subjektive, leistungsorientierte und physiologische Indikatoren, die zusammen den trimodalen Ansatz der Beanspruchungsmessung bilden. Die physiologischen Indikatoren werden darüber hinaus in bewusste und unbewusste Parameter unterteilt, die eine Entsprechung der jeweiligen Verarbeitungsprozesse darstellen. Diese Differenzierung wird sowohl theoretisch als auch empirisch hergeleitet und geprüft. In Kapitel 3.2 wird dann auf die verwendeten Methoden der Signalverarbeitung und der Statistik eingegangen. Im Fokus der Signalverarbeitung steht die Procrustes-Analyse, die ein wesentlicher Bestandteil des Auswertungsverfahrens ist, welches in Kapitel 5.3.2 hergeleitet wird.

3.1 Erfassung mentaler Beanspruchung

Die theoretischen Ausführungen haben gezeigt, dass die mentale Beanspruchung und damit auch deren Erfassung von einer Vielzahl an Faktoren (z. B. Anforderungshöhe, Aufgabenart, individuelle Erfahrungen und Konzepte) sowohl intra- als auch interindividuell beeinflusst werden. Grundsätzlich lassen sich drei Kategorien von Erhebungsmethoden unterscheiden: *subjektive*, *leistungsorientierte* und *physiologische* Messinstrumente (De Waard, 1996; Wickens et al., 2013; Manzey, 1998). Alle drei Kategorien haben speziell in Abhängigkeit des Versuchsdesigns und des Untersuchungsgegenstandes verschiedene Vor- und Nachteile, die zusätzlich durch die starken individuellen Unterschiede in

der menschlichen Informationsverarbeitung beeinflusst werden. Die einzelnen Parameter aller drei Kategorien können grundsätzlich anhand von fünf Kriterien für eine zielgerichtete Anwendung bewertet werden. Die fünf Kriterien umfassen die *Sensitivität*, *Diagnostizität*, *Beeinflussung der Hauptaufgabe*, *Anwendungsaufwand* und die *Nutzerakzeptanz* (De Waard, 1996). Die Sensitivität beschreibt die Fähigkeit von Parametern, Veränderungen in der Beanspruchung (z. B. durch Anforderungsniveaus) zu indizieren. Diese Fähigkeit unterscheidet sich beispielsweise sehr stark in Abhängigkeit von den Anforderungen einer Aufgabe. Um das Modell nach De Waard (1996) nochmals aufzugreifen (Abbildung 2.12), kann beispielsweise die Messung der Leistung in einer Hauptaufgabe nicht in den Anforderungsbereichen A und C eingesetzt werden, da die Leistung dort, im Gegensatz zu den Bereichen B und D, konstant bleibt (Abbildung 2.12). In Abhängigkeit der einzelnen Anforderungsbereiche können auch die verschiedenen Kategorien von Messinstrumenten eine unterschiedliche Sensitivität aufweisen. Das hohe Beanspruchungsniveau in Bereich C kann von Probanden subjektiv meist sehr gut artikuliert werden (Manzey, 1998). Die Diagnostizität beschreibt die Fähigkeit eines Parameters, die Art der Beanspruchung zu differenzieren: also die Beanspruchung einem spezifischen Belastungsfaktor der Arbeitsaufgabe (z. B. Aufgabenart) zuordnen zu können (De Waard, 1996). Basierend auf dem Ressourcenmodell nach Wickens (1984) besitzt ein Beanspruchungsparameter eine hohe Diagnostizität, wenn er nur sensitiv für die Veränderung einer spezifischen Ressource ist. Parameter hoher Diagnostizität beschreiben ausschließlich sehr spezifische Anforderungen, wohingegen Parameter mit geringer Diagnostizität eher die allgemeinen Anforderungen einer Aufgabe widerspiegeln (De Waard, 1996). Insbesondere die Unterscheidung von bewussten und unbewussten Informationsverarbeitungsprozessen sowie deren Abhängigkeit von individuellen Erfahrungen und Konzepten kann die Bewertung der Diagnostizität erheblich erschweren. Aufgrund der zeitvariablen Veränderung der inneren Welt, unter anderem durch anforderungsunabhängige Veränderungen der Erwartungen und Ziele (Dehaene & Changeux, 2005), ist die eindeutige Definition der Diagnostizität für einen Parameter nur unter stark kontrollierten Bedingungen möglich. Die Beeinflussung der Hauptaufgabe bewertet,

inwieweit der Aufbau oder der Einsatz des Messinstrumentes die Bearbeitung der eigentlichen Arbeitsaufgabe beeinflusst oder stört. Die Verwendung von Nebenaufgaben hat mitunter den größten Einfluss auf die Hauptaufgabe, da Interaktionseffekte zwischen Haupt- und Nebenaufgabe benötigt werden, um Rückschlüsse auf die verwendeten kognitiven Ressourcen ziehen zu können (Wickens et al., 2013). Dementsprechend ist von großer Bedeutung, dass die Nebenaufgabe gut in den Versuchsablauf oder den Handlungsstrang der Hauptaufgabe eingebunden wird (Schneider, Bringeland & Deml, in Review). Auch die Befragung von Probanden während der Aufgabenbearbeitung bildet eine Störvariable für den flüssigen Arbeitsablauf, sodass Fragebögen ausschließlich nach Bearbeitung der Aufgabe gereicht werden sollten (De Waard, 1996). Der Anwendungsaufwand eines Messinstrumentes umfasst beispielsweise die Anzahl der benötigten Apparaturen und den Aufwand diese dem Probanden anzulegen (De Waard, 1996). Die Nutzerakzeptanz als letztes Kriterium beeinflusst stark die Motivation eines Probanden und insbesondere im Falle von subjektiven Messinstrumenten wird weiter die Richtigkeit und Präzision der Angaben erheblich beeinflusst (De Waard, 1996). Die Messinstrumente sollten so gewählt werden, dass sie den Probanden möglichst wenig stören. Die kopfbasierte Blickerfassung gilt dabei als ein beispielhaftes physiologisches Messinstrument mit sehr hoher Nutzerakzeptanz (Sträter, 2005; Manzey, 1998). Die fünf grundlegenden Kriterien zur Bewertung von Messinstrumenten können um die Selektivität und die Reliabilität erweitert werden (De Waard, 1996). Die Selektivität beschreibt die Validität eines Messinstrumentes. Eine hohe Selektivität bedeutet, dass ein Parameter lediglich für ein spezielles Konstrukt, wie die mentale Beanspruchung, sensitiv ist. Der NASA Task Load Index (*NASA-TLX*) nach Hart und Staveland (1988) beispielsweise erfragt neben den geistigen auch die körperlichen Anforderungen, was dazu führt, dass er nicht nur sensitiv für eine Art der Beanspruchung ist. Die Reliabilität beschreibt die Stabilität eines Messinstrumentes über verschiedene Versuche hinweg. Nicht alle Parameter, die unter Laborbedingungen konstante Ergebnisse liefern, weisen auch in Feldstudien eine vergleichbare Zuverlässigkeit auf (Rubio, Díaz, Martín & Puente, 2004). Ein großer Teil der Varianz zwischen unterschiedlichen Versuchen, kann

auch hier wieder auf das Anforderungsniveau zurückgeführt werden. Manche Messinstrumente zeigen nur Unterschiede in hohen Anforderungsbereichen, andere lediglich in niedrigeren Bereichen (De Waard, 1996). Insgesamt können die genannten Kriterien nicht unabhängig voneinander betrachtet werden, da sie sich häufig gegenseitig bedingen. Eine hohe Diagnostizität beispielsweise schränkt die Sensitivität eines Parameters stark ein. Häufig werden in der Literatur die Sensitivität, Reliabilität und die Beeinflussung der Hauptaufgabe als die wichtigsten Kriterien für Messinstrumente herausgestellt (Wickens et al., 2013; De Waard, 1996; Manzey, 1998).

Die drei Kategorien von Messinstrumenten für die mentale Beanspruchung unterscheiden sich hinsichtlich der vorgestellten Kriterien für die gleiche Aufgabenstellung in Abhängigkeit mehrerer Einflussgrößen. So kann beispielsweise ein Abfall der Leistung mit einem abnehmenden, subjektiven Beanspruchungsempfinden einhergehen (Yeh & Wickens, 1988; Vidulich & Wickens, 1986), was unter anderem auf eine Veränderung in der Anstrengung des Probanden oder auf einen Strategiewechsel beziehungsweise einen Adaptionsprozess zurückgeführt werden könnte (Sperling & Doshier, 1986; Fels et al., 2015). Aufgrund dieser Unterschiede bietet sich unter Laborbedingungen ein trimodaler Ansatz der Beanspruchungsmessung an, in dem Messinstrumente aus allen drei Kategorien gemeinsam eingesetzt werden (Abbildung 3.1). Subjektive, leistungsbasierte und physiologische Messinstrumente bilden unterschiedliche Facetten der Beanspruchung ab. Durch eine Kombination von Parametern aus allen drei Kategorien kann, basierend auf einer Kreuzvalidierung, eine umfassende Beschreibung der Beanspruchung geschaffen werden. Die physiologischen Parameter sind dabei am schwierigsten zu interpretieren, da sie die kognitiven Prozesse des Menschen anhand von physiologischen Reaktionen messen. Die Vielzahl an Einflussfaktoren (z. B. bewusste und unbewusste Adaptionsprozesse, Veränderungen der Zielsetzung), die aus dem theoretischen Hintergrund hergeleitet wurden, wirken sich hier am deutlichsten aus (Fels et al., 2015; Sträter, 2005; Manzey, 1998).

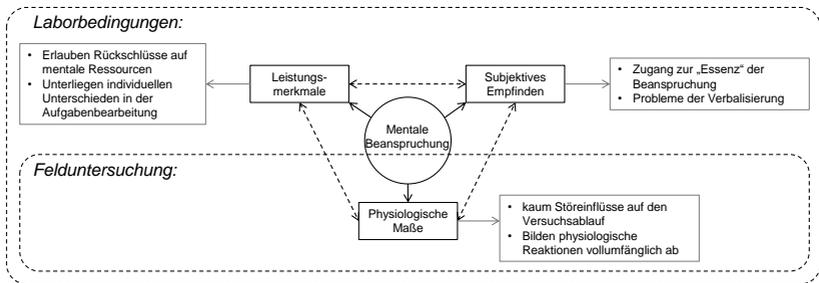


Abb. 3.1: Trimodaler Ansatz der Beanspruchungsmessung (eigene Abbildung)

Allerdings können die physiologischen Parameter sehr gut in Feldstudien eingesetzt werden, in denen subjektive und leistungsbasierte Messungen nur mit einer erheblichen Beeinflussung des Probanden oder gar nicht möglich sind. Das subjektive Beanspruchungsempfinden kann oft nur nach Arbeitstätigkeiten abgefragt werden, um den betrieblichen Ablauf und vor allem die zu analysierende Arbeitstätigkeit nicht zu unterbrechen. Dementsprechend ist es oft schwer, die subjektive Bewertung auf spezifische Situationen zu beziehen (Vidulich & Tsang, 1986), um konkrete Handlungsempfehlungen ableiten zu können. Die Leistung eines Probanden darf in einem praktischen Umfeld oft aus rechtlichen oder betriebsinternen Gründen gar nicht erhoben werden (Vidulich & Tsang, 1986; Rubio et al., 2004), da diese weitreichendere Rückschlüsse zulassen würde. Diese Situation motiviert erneut die Fragestellung dieser Arbeit. Nur wenn die Zuverlässigkeit der physiologischen Beanspruchungsmessung aufgabenunabhängig erhöht wird, ist eine Erfassung der mentalen Beanspruchung direkt in Arbeitssystemen möglich. Der direkte Bezug zu dem zu gestaltenden Arbeitssystem ist eine notwendige Voraussetzung für die menschenzentrierte und nachhaltige Arbeitsgestaltung komplexer Systeme.

3.1.1 Subjektives Beanspruchungsempfinden

Das subjektive Beanspruchungsempfinden, erfasst über Fragebögen, wird häufig als die „Essenz“ der mentalen Beanspruchung bezeichnet. Nur so kann erfasst

werden, wie der Proband selbst die Anforderungen einer Aufgabe empfunden hat (De Waard, 1996; Manzey, 1998). Moray (1979) berichtet, dass eine Person angestrengt und beansprucht ist, wenn sie dies in einem Fragebogen oder Interview angibt, unabhängig davon, was leistungsorientierte oder physiologische Messinstrumente zeigen. Das subjektive Empfinden ist auch heute noch oft der entscheidende Faktor in der Bewertung von der Effektivität beanspruchungsoptimierender Systemgestaltung (Wickens et al., 2013). Allerdings sind subjektive Messinstrumente nicht immer zuverlässige Indikatoren für die mentale Beanspruchung (Vidulich & Tsang, 1986). Es ist oft schwierig für den Anwender, seine Beanspruchung korrekt und zutreffend zu verbalisieren. So haben die Erfahrungen mit solchen Fragebögen, als auch die Zeitpunkte der Abfrage einen erheblichen Einfluss auf die abgegebenen Bewertungen (De Waard, 1996). Weiter entstehen Probleme bei der Abbildung von zeitlich variablen Anforderungen über die Dauer einer Arbeitsaufgabe hinweg. Thornton (1985) vergleicht in einer Studie Aufgaben mit gleichbleibenden Anforderungen mit Aufgaben, die jeweils eine kurzzeitige Anforderungsspitze (hoher Anforderungsanstieg) beinhalten. Die Ergebnisse von Thornton (1985) zeigen, dass eine kurze Beanspruchungsspitze in einer längeren Aufgabe zu einer deutlichen Verzerrung der subjektiven Bewertung hin zu dem Maximalwert führt. Dabei hängt die Effektstärke von der Höhe der Beanspruchungsspitze ab. In den Studien von Yeh und Wickens (1988) hingegen befanden sich die subjektiven Beanspruchungsbewertungen auf einem mittleren Anforderungsniveau der verschiedenen Teilaufgaben. Allerdings waren die Veränderungen in der Anforderungshöhe weitaus geringer als in den Studien von Thornton (1985). Die Befunde aus beiden Studien können Vorteile hinsichtlich der subjektiven Beanspruchungsmessung bedeuten. Vor dem Hintergrund einer engpassorientierten Systembetrachtung ist es sehr wichtig, Beanspruchungsspitzen während der Arbeitstätigkeit zu vermeiden, weshalb deren Identifikation einen großen Mehrwert liefern würde (Wickens et al., 2013). Andernfalls haben die Ergebnisse nach Yeh und Wickens (1988) dahingehend einen Vorteil, dass das Gesamtbeanspruchungsniveau für Arbeitstätigkeiten definiert werden kann. Estes (2015) bezieht das subjektive Beanspruchungsempfinden ausschließlich auf

bewusste kognitive Verarbeitungsprozesse. Die bewussten Verarbeitungsprozesse attribuiert Estes (2015) vorwiegend auf der Ebene des Kurzzeitgedächtnisses, weshalb das subjektive Empfinden im Wesentlichen durch die Anforderungen an das Kurzzeitgedächtnis determiniert wird. Faktoren wie die verfügbaren Ressourcen, die Aktivierung oder die Aufmerksamkeitsfokussierung beeinflussen die Beanspruchung des Kurzzeitgedächtnisses ebenso wie die Präsentationsrate von Reizen und deren Verarbeitungsgeschwindigkeit (Yeh & Wickens, 1988). Aktuelle Konzepte der menschlichen Informationsverarbeitung nach Dehaene und Changeux (2005) oder Kiefer et al. (2012) zeigen, dass auch bestimmte unbewusste Verarbeitungsprozesse auf kognitive Ressourcen zurückgreifen (Kapitel 2.1.2). Ausgehend von dem theoretischen Hintergrund sowie dem Zusammenhang zwischen dem subjektiven Beanspruchungsempfinden und dem Verbrauch kognitiver Ressourcen (Yeh & Wickens, 1988; Estes, 2015) kann davon ausgegangen werden, dass subjektive Messinstrumente auch sensitiv für unbewusste Verarbeitungsprozesse sind.

Aufgrund der diversen Vor- und Nachteile subjektiver Verfahren, wurde bis heute eine Vielzahl an Fragebögen entwickelt. Zu den bekanntesten zählen der *NASA-Task-Load-Index (NASA-TLX)* nach Hart und Staveland (1988), die *Cooper-Harper Skala* (Cooper & Harper, 1969), das *Workload-Profile (WP)* nach Tsang und Velazquez (1996), der *SWAT Fragebogen* (Reid & Nygren, 1988) oder die *SEA Skala* (Eilers, Nachreiner & Hänecke, 1986). Generell lassen sich zwei Arten von subjektiven Erfassungsmethoden unterscheiden: *eindimensionale* und *mehrdimensionale Verfahren*. Eindimensionale Verfahren (z. B. SEA) erfragen die mentale Beanspruchung auf einer Skala ohne eine weitere Differenzierung in einzelne Facetten, wie es bei den mehrdimensionalen Methoden (z. B. NASA-TLX) der Fall ist, wodurch häufig eine höhere Sensitivität für Veränderungen der Anforderungen geschaffen wird (De Waard, 1996). In einer empirischen Studie von Vidulich und Tsang (1986) wird der NASA-TLX mit dem SWAT Fragebogen in einem Laborversuch hinsichtlich ihrer Validität verglichen. Die Konstruktvalidität wird anhand der Sensitivität für Veränderungen der Anforderungen geprüft. Die Werte beider Fragebögen

steigen zusammen mit den Anforderungen in der Hauptaufgabe (Verfolgungsaufgabe), wobei lediglich der NASA-TLX eine signifikante Differenzierung der Anforderungsniveaus ermöglicht. Beide Erfassungsmethoden zeigen signifikante Zusammenhänge mit der Leistung, weshalb die Kriteriumsvalidität für beide angenommen werden kann (Vidulich & Tsang, 1986). Beide Messinstrumente zeigen eine hohe Sensitivität für die Wahrnehmungsmodalitäten der Aufgabe, aber nicht für die Eingabemodalitäten. Auch Wickens et al. (2013) bestätigen die hohe Sensitivität subjektiver Erfassungsmethoden für die Art der Informationsdarstellung, was einen großen Vorteil für die Bewertung von Informationssystemen bietet. Vidulich und Tsang (1986) vermuten zwei mögliche Gründe für die fehlende Sensitivität gegenüber der Eingabemodalität. Es wäre möglich, dass die verwendeten Modalitäten nicht unterschieden werden können oder, dass generell keine Eingabemodalitäten unterschieden werden können. Dies würde insbesondere für Mensch-Roboter-Interaktion eine erhebliche Restriktion darstellen, da so keine Steuerungskonzepte subjektiv bewertet werden könnten. Eine empirische Studie von Schneider und Deml (2016) zum Vergleich verschiedener Eingabemodalitäten in der Mensch-Roboter-Interaktion unter Verwendung des NASA-TLX zeigt jedoch signifikante Unterschiede zwischen den Steuerungsarten ($F(2, 36) = 7,08, p < 0,01$). Motiviert durch diese Befunde, wurde der NASA-TLX auch für die Studien im Rahmen dieser Arbeit gewählt. Nach Rubio et al. (2004), die den NASA-TLX mit dem SWAT und dem WP vergleichen, zeichnet sich der NASA-TLX besonders durch seine hohe Validität, Reliabilität und Nutzerakzeptanz aus. Die genannte Studie von Schneider und Deml (2016) zeigt, dass der NASA-TLX auch in dynamischen Umgebungen in der Lage ist, Anforderungsniveaus signifikant zu differenzieren und eine hohe Diagnostizität hinsichtlich leistungsbasierter Parameter aufzuweisen. Der NASA-TLX besteht aus sechs Items (*geistige Anforderungen, körperliche Anforderungen, zeitliche Anforderungen, Leistung, Anstrengung, Frustration*), die jeweils auf einer 20-stufigen Skala von niedrig bis hoch abgefragt werden. In der ursprünglichen Version des Fragebogens folgt auf die Bewertung der sechs Items ein paarweiser Vergleich der gesamten Items, sodass sich eine Gewichtung zur Berechnung des gewichteten Mittelwertes ergibt (Hart & Staveland,

1988). Auf diesen paarweisen Vergleich wurde hier verzichtet und der einfache Mittelwert der Items für eine Gesamtbewertung gebildet (Vidulich & Wickens, 1986). Der NASA-TLX kann auch bezüglich der Item-Struktur angepasst werden (Zhang, Yin & Wang, 2015). Zhang et al. (2015) beispielsweise erweiterten den Fragebogen um zwei weitere Items (Stress und Müdigkeit). Die Aufgaben in den Studien dieser Arbeit werden ausschließlich im Sitzen ausgeführt. In der ersten Studie wurden die körperlichen Anforderungen als vernachlässigbar gering eingeschätzt ($M = 0$, $SD = 0$), sodass der NASA-TLX Fragebogen mit nur fünf Items, ohne „körperliche Anforderungen“, verwendet wird.

3.1.2 Leistungsbasierte Parameter

Das Ressourcenmodell nach Wickens (1984) bildet eine der wichtigsten Grundlagen für die leistungsbasierte Beanspruchungsmessung. Eine Veränderung der kognitiven Ressourcenallokation ist eine der Hauptursachen für die Leistungsvariabilität, weshalb die Leistung direkt mit der mentalen Beanspruchung verbunden ist (Eggemeier & Wilson, 1991). Grundsätzlich unterscheidet sich die leistungsbasierte Beanspruchungsmessung in die *Messung der Hauptaufgabe* und in *Nebenaufgabenmessungen* (Wickens et al., 2013; De Waard, 1996; Manzey, 1998; Eggemeier & Wilson, 1991). Die Leistung der Hauptaufgabe kann im Wesentlichen durch die Präzision, die Anzahl an Fehlern, die Bearbeitungsgeschwindigkeit oder die Reaktionszeiten erfasst und quantifiziert werden (De Waard, 1996). Bezogen auf die Anforderungsbereiche nach De Waard (1996) ist die Leistung in einer Hauptaufgabe sensitiv für die mentale Beanspruchung in den Bereichen B und D (Abbildung 2.12), wobei die Grenzen zwischen den Bereichen nicht als statisch angesehen werden dürfen. Insbesondere die automatischen Informationsverarbeitungsprozesse, ausgehend von Erfahrungen und Konzepten (Sträter, 2005), oder das erfolgreiche Anwenden einer neuen Bearbeitungsstrategie (Sperling & Doshier, 1986) können eine effektive Erweiterung des optimalen Anforderungsbereichs A2 ermöglichen. Darüber hinaus können auch Motivatoren (z. B. Zielsetzung) zu einer Erweiterung oder Einschränkung des Bereiches A2 führen (Eggemeier & Wilson, 1991). Durch die

genannten Einflussfaktoren ist die Leistung in der Hauptaufgabe nicht immer sensitiv für die mentale Beanspruchung, insbesondere in den Anforderungsbereichen A1 und A3. In Laborumgebungen können die Einflussfaktoren relativ gut kontrolliert werden, wohingegen in Feldstudien große Schwankungen auftreten können. Die Gesamteffektivität einer Mensch-Maschine-Schnittstelle lässt sich zwar aufgrund der Leistung in der Hauptaufgabe quantifizieren, für eine genaue Betrachtung der Beanspruchung müssen aber oft weitere Messinstrumente herangezogen werden. Zwei Probanden mit der gleichen Leistung können beispielsweise ganz unterschiedlich beansprucht sein. Die Leistung in einer Nebenaufgabe gilt dahingegen als sensitiver bezüglich der mentalen Beanspruchung (Eggemeier & Wilson, 1991). Für die Erfassung der Leistung (z. B. Fehlerrate, Reaktionszeiten) in einer Nebenaufgabe werden zwei Paradigmen unterschieden: das „*Loading Task Paradigm*“ und das „*Subsidiary Task Paradigm*“ (De Waard, 1996). Innerhalb des Loading Task Paradigm werden Probanden dahingehend instruiert zu versuchen, die Leistung in der Nebenaufgabe konstant zu halten, auch wenn dadurch Einbußen in der Hauptaufgabenleistung zu verzeichnen sind. In diesem Fall findet ein Übergang von Anforderungsbereich A in den Bereich B statt, mit dem sich der Anstieg in der Beanspruchung beschreiben lässt (De Waard, 1996). Das Subsidiary Task Paradigm gilt als der weiter verbreitete Ansatz (Eggemeier & Wilson, 1991). Hier wird der Proband instruiert, die Leistung in der Hauptaufgabe möglichst konstant zu halten und die Nebenaufgabe mit einer untergeordneten Priorität zu bearbeiten. Die Nebenaufgabenleistung indiziert in diesem Fall die freien kognitiven Kapazitäten, die bei der isolierten Bearbeitung der Hauptaufgabe für weitere Tätigkeiten genutzt werden könnten (De Waard, 1996). Nach Wickens (1984) steigt die Sensitivität bezüglich der mentalen Beanspruchung mit steigenden Interaktionseffekten bezüglich der geforderten Ressourcen zwischen Haupt- und Nebenaufgabe. Allerdings hängt die Sensitivität stark von der Anforderungshöhe der Hauptaufgabe und dem Verhalten bezüglich der Ressourcenallokation der individuellen Person ab (temporär präferierte Aufgabe). So kann beispielsweise bei Nebenaufgaben, die eine verbale Rückmeldung des Probanden erfordern, der Effekt der sozialen Erwünschtheit dazu führen,

dass die Nebenaufgabe temporär die höchste Priorität besitzt (Wickens et al., 2013). Weiter hat die Nutzerakzeptanz gegenüber zusätzlichen Aufgaben, insbesondere bei alltäglichen Hauptaufgaben, einen erheblichen Einfluss auf die Sensitivität. Es ist folglich sehr wichtig, welche Nebenaufgabe gewählt wird. Grundlegend muss die verwendete Nebenaufgabe kontinuierlich messbar sein, mit der Hauptaufgabe interferieren und ein geeignetes Anforderungsniveau aufweisen. Eggemeier und Wilson (1991) haben die genutzten Nebenaufgaben in verschiedenen Studien hinsichtlich der Sensitivität und der Anwendbarkeit in komplexen Umgebungen verglichen, wobei im Wesentlichen Reaktions-, Zeitschätzungs-, Gedächtnis- und Kopfrechenaufgaben berücksichtigt wurden. Für eine detaillierte Übersicht soll an dieser Stelle an Eggemeier und Wilson (1991) verwiesen werden. Die verschiedenen Studien haben allerdings auch potentielle Nachteile bei der Verwendung von Nebenaufgaben mit einem Interaktionseffekt zu der analysierenden Hauptaufgabe verdeutlicht. Mögliche Nachteile basieren dabei im Wesentlichen auf den benötigten zusätzlichen Versuchsaufbauten, den einzugehenden Kompromissen in der Systemsicherheit (bei Feldstudien) und der Akzeptanz der Probanden. Eine Lösung dieser potentiellen Probleme ist das Paradigma der „*Embedded Secondary Task*“ (Eggemeier & Wilson, 1991). Eine *Embedded Secondary Task* ist eine normale Systemfunktion oder Teiltätigkeit im Arbeitssystem der zu untersuchenden Hauptaufgabe, die aber distinkt von der Hauptaufgabe bearbeitet wird (De Waard, 1996). Diese Art von Aufgaben hat eine geringere Priorität und repräsentiert alltägliche Systemfunktionen, wodurch das Untersuchungssystem repräsentativer wird und die Nutzerakzeptanz hoch ist. Als beispielhafte Anwendung verweisen Eggemeier und Wilson (1991) auf Shingledecker und Crabtree (1982), die in mehreren Studien von Fluglotsentätigkeiten die Sensitivität von Funkaufgaben als *Embedded Secondary Tasks* analysiert haben. Die Ergebnisse stimmen mit den Annahmen der Ressourcentheorie überein und zeigen, dass vier von acht verschiedenen Aufgaben erfolgreich die Anforderungsniveaus der Hauptaufgabe differenzieren können (Shingledecker & Crabtree, 1982). In den Studien dieser Arbeit werden verschiedene Nebenaufgaben nach dem *Subsidiary Task Paradigm* verwendet, die im Folgenden beschrieben werden.

Visuelle n-back Aufgabe

Die *n-back Aufgabe* gehört zu den *visuellen Gedächtnisaufgaben* (Wickens, 2002). Die Probanden bekommen in definierten Zeitabständen (hier: vier Sekunden) eine Zahl im Zahlenraum von eins bis neun für eine Sekunde visuell auf einem Bildschirm präsentiert. Sobald eine neue Zahl erscheint, soll der Proband verbal die vorherige Zahl wiedergeben. Die verbalen Antworten werden registriert und anschließend anhand der Fehlerrate ausgewertet. Die *n-back* (hier: *1-back*) Aufgabe wurde bereits in mehreren Studien auf die Sensitivität hinsichtlich der mentalen Beanspruchung geprüft (Trowbridge & Collins McLaughlin, 2013; Anand, Terken & Hogema, 2013). Innerhalb des *Subsidiary Task Paradigm* hat diese Aufgabe allerdings Nachteile bezüglich kumulativer Fehler, was die Sensitivität in gewissem Maße einschränkt. Falls sich ein Proband die aktuelle Zahl merken kann, aber zum Zeitpunkt der nächsten Zahl mit der Hauptaufgabe beschäftigt ist, entstehen häufig zwei Fehler aufgrund von einer versäumten Zahl. Zum einen kann die gemerkte Zahl nicht wiedergeben werden und zum anderen wird die nächste Zahl verpasst, sodass zum darauf folgenden Zeitpunkt ein Folgefehler entsteht.

Mentale Rotation

Die *mentale Rotation* zählt zu den *Auswahl- und Reaktionsaufgaben* (Eggemeier & Wilson, 1991). Es existieren verschiedene Varianten dieser Aufgabe mit unterschiedlichen Anforderungsniveaus (Xu & Franconeri, 2015). In der Version, die im Rahmen dieser Arbeit verwendet wurde, sollen zwei zweidimensionale Geometrien im mittleren Anforderungsbereich (acht bis zwölf Konturpunkte) miteinander verglichen werden (Bertrau-Pavy, Raber & Piper, 2011). Es werden in Abständen von fünf Sekunden jeweils zwei Geometrien nebeneinander für eine Dauer von drei Sekunden präsentiert.

Der Proband soll innerhalb der drei Sekunden verbal oder über einen Tastendruck bewerten, ob die beiden Geometrien gleich oder gespiegelt sind. Die

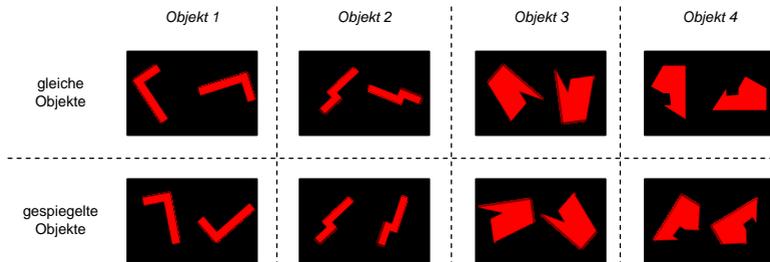


Abb. 3.2: Geometrien der mentalen Rotation (eigene Abbildung)

zufällig ausgewählten Geometrien erscheinen dabei in einer randomisierten Orientierung im Raum. Abbildung 3.2 zeigt Beispiele der vier Geometrien mit acht Konturpunkten, die in randomisierter Reihenfolge präsentiert werden. Oben sind jeweils gleiche Paare dargestellt, wohingegen die unteren Paare gespiegelt sind. Die Antworten der Probanden werden erfasst und anschließend hinsichtlich der Fehlerquote ausgewertet (Bertrau-Pavy et al., 2011).

Visueller Reiz

Ein *visueller Reiz* im peripheren Sichtfeld wird als *Reaktionsaufgabe* klassifiziert, die überwiegend in Studien zur Erforschung der Beanspruchung bei Autofahrern eingesetzt wird (Engström, Åberg, Johansson & Hammarbäck, 2005; Harbluk, Burns, Hernandez, Tam & Glazduri, 2013; Sonnleitner, Simon, Kincses, Buchner & Schrauf, 2012). Ein kleines Quadrat (5mm x 5mm) wird dabei in randomisierten zeitlichen Abständen in zufälligen Ecken des Bildschirms, auf dem die Hauptaufgabe bearbeitet wird, für 0,3 Sekunden angezeigt. Sobald der Proband einen solchen Reiz wahrnimmt, soll er dies so schnell wie möglich über einen Tastendruck bestätigen. Die Taste wird dabei so gewählt, dass diese ohne zeitlichen Aufwand während der Hauptaufgabe betätigt werden kann. Zur Auswertung der Leistung werden die Reaktionszeiten und die Anzahl der Fehler (nicht-Reagieren auf einen Reiz) erhoben (Engström et al., 2005). Tastendrucke

werden dabei erst ab dem Zeitpunkt des Erscheinen eines Reizes registriert, sodass keine zufälligen Tastendrucke (Reaktionen ohne vorangegangenen Reiz) die Auswertung verzerren können.

Mentale Arithmetik

Weiterhin wird im Rahmen dieser Arbeit die *mentale Arithmetik* verwendet, die eine eigene Kategorie von Nebenaufgaben bildet (Eggemeier & Wilson, 1991). Dem Probanden werden in fünf Sekunden Abständen verschiedene mathematische Aufgaben auditiv präsentiert. Innerhalb der fünf Sekunden nach einer Aufgabe, soll der Proband mit dem möglichst richtigen Ergebnis verbal antworten, wobei die Fehlerquote zur Quantifizierung der Leistung erfasst wird. Die mentale Arithmetik wurde als Nebenaufgabe bereits in einer Vielzahl an Studien hinsichtlich der Differenzierbarkeit von Anforderungen verschiedener Aufgabentypen und der Sensitivität hinsichtlich der mentalen Beanspruchung untersucht (Verner, Herrmann, Troche, Roebers & Rammsayer, 2013; Hitch, 1978; Ketelsen & Welsh, 2010; Lee & Kang, 2002; Ashcraft, 1992). Die mentale Arithmetik erfordert alle grundlegenden kognitiven Funktionen: Informationsaufnahme und -verarbeitung, die Handlungsausführung und den kontinuierlichen Abgleich von Informationen (Verner et al., 2013). Innerhalb der mentalen Arithmetik unterschieden sich Aufgaben nach ihren Anforderungsniveaus. Additionsaufgaben im Zahlenraum von 1 - 100 bilden dabei das leichte Anforderungsniveau. Es besteht jedoch ein signifikanter Unterschied darin, welche Zahl bei Aufgaben mit einer einstelligen und einer zweistelligen Zahl zuerst genannt wird: „23 + 9“ oder „9 + 23“. Der Unterschied begründet sich in der Anzahl der Informationsverarbeitungsschritte, die notwendig sind, um die einzelnen Zahlen vor ihrer Addition zu zerlegen. So unterscheidet sich das genannte Beispiel um einen Verarbeitungsschritt: „23 + 9“ oder „9 + 3 + 20“. Für detailliertere Untersuchungen der Aufgabenarten soll an dieser Stelle auf Hitch (1978) und Verner et al. (2013) verwiesen werden. Für die zweite Studie im Rahmen dieser Arbeit (Kapitel 6) werden leichte Additionsaufgaben als Nebenaufgabe verwendet. Aufgrund der zu geringen Fehlerquoten wird

in der dritten Studie (Kapitel 7) eine Kombination aus zwei Aufgabentypen im mittleren Anforderungsbereich gewählt; die Subtraktion einer einstelligen Zahl von einer zweistelligen Zahl und die Subtraktion einer zweistelligen Zahl von einer zweistelligen Zahl im Zahlenraum von 1 - 100 (Ketelsen & Welsh, 2010). Die zwei Aufgabentypen werden gleichverteilt und hinsichtlich des Minuend, des Subtrahend und der Reihenfolge randomisiert. Tabelle 3.1 fasst die verwendeten Nebenaufgaben in den jeweiligen Studien mit den erhobenen Leistungsindikatoren zusammen.

Tab. 3.1: Übersicht der verwendeten Nebenaufgaben (eigene Tabelle)

Aufgabe	Aufgabentyp	Darstellungsmodalität	Eingabemodalität	Leistungsmaß	Verwendet in
<i>n-back Aufgabe</i>	visuelle Gedächtnisaufgabe	visuell	verbal	Fehlerquote	Studie 1
<i>mentale Rotation</i>	Auswahl- und Reaktionsaufgabe	visuell	verbal / manuell	Fehlerquote	Studie 1, Studie 2
<i>visueller Reiz</i>	Reaktionsaufgabe	visuell	manuell	Fehlerquote, Reaktionszeit	Studie 2, Studie 3
<i>mentale Arithmetik</i>	mentale Arithmetik	auditiv	verbal	Fehlerquote	Studie 2, Studie 3

Bezüglich der Leistung in den verschiedenen Nebenaufgaben gibt es häufig einen trade-off zwischen den verschiedenen Leistungsparametern, sodass die einzelnen Leistungsindikatoren nicht immer hochpräzise die mentale Beanspruchung beschreiben (Galy et al., 2012). Paas und VanMerriënboer (1993) kombinierten die Leistung in der Nebenaufgabe mit der subjektiv empfundenen Anstrengung zu dem Maß der mentalen Effizienz E (Formel 3.1).

$$E = \frac{|R - P|}{\sqrt{2}} \quad (3.1)$$

Zur Berechnung der mentalen Effizienz (Formel 3.1) wird der z-Wert der Leistung in der Nebenaufgabe (mentale Beanspruchung durch Aufgabe oder Umwelt) von dem z-Wert der subjektiv empfundenen Anstrengung (aktivierte kognitive Ressourcen) subtrahiert und anschließend durch die Wurzel von zwei dividiert (Paas & VanMerriënboer, 1993). Der Parameter bietet ein umfassendes Bild der mentalen Beanspruchung, da die Anstrengung sowie die Leistung zwei zentrale, messbare Dimensionen der Beanspruchung darstellen (De Waard, 1996). Personen bearbeiten eine Aufgabe demnach effizienter, wenn der Grad der Anstrengung im Verhältnis geringer ist als die resultierende Leistung. Eine hohe Leistung bei niedriger Anstrengung wäre folglich die effizienteste Kombination mit der geringsten Beanspruchung (Paas & VanMerriënboer, 1993). Eine weitere Alternative ist die Kombination aus Genauigkeit und Geschwindigkeit der Antworten (Fournier, Wilson & Swain, 1999). Auf diese Weise kann den genannten methodologischen Schwierigkeiten bei der Auswertung der Leistung in einer Nebenaufgabe entgegengewirkt werden. Fournier et al. (1999) nennen den Parameter die globale Leistung und dividieren dazu den z-Wert der mittleren Reaktionszeiten in der Nebenaufgabe durch den z-Wert der Quote der richtigen Antworten. Die globale Leistung zeigt sich als hoch sensitiv für die mentale Beanspruchung. Fournier et al. (1999) können unter Verwendung der globalen Leistung die Trainingseffekte sehr detailliert beschreiben. Galy et al. (2012) verwenden beide Parameter in einer Studie zur Unterscheidung von verschiedenen kognitiven Beanspruchungsarten, wobei die Ergebnisse die hohe Sensitivität beider Parameter (mentale Effizienz, globale Leistung) bestätigen. Auch weitere Studien bestätigen diese Ergebnisse (Galy & Mélan, 2015; Jong, 2010). Die mentale Effizienz nach Paas und VanMerriënboer (1993) wird für die gesamten Nebenaufgaben dieser Arbeit berechnet. Aus datenschutzrechtlichen und technischen Gründen können verbale Antworten in den Studien nicht aufgezeichnet und hinsichtlich der Reaktionszeiten ausgewertet werden, weshalb die globale Leistung nach Fournier et al. (1999) in den Studien dieser Arbeit nicht berücksichtigt werden kann.

3.1.3 Physiologische Beanspruchungsindikatoren

Physiologische Messinstrumente erfassen die mentale Beanspruchung ganzheitlich anhand der physiologischen Reaktionen des menschlichen Körpers, wobei eine kontinuierliche Messung mit einer hohen zeitlichen Auflösung realisiert wird (Wickens et al., 2013; De Waard, 1996; Manzey, 1998). Die physiologischen Reaktionen des Körpers (z. B. Pupillenveränderungen, Herzratenveränderungen) werden durch das periphere Nervensystem autonom (unbewusst) aktiviert (De Waard, 1996). Das periphere Nervensystem lässt sich in das parasympathische Nervensystem (PNS), welches für Entspannungsreaktionen verantwortlich ist, und das sympathische Nervensystem (SNS), welches Stress- und Beanspruchungsreaktionen auslöst, unterteilen. Nicht alle physiologischen Reaktionen entstehen jedoch in den beiden genannten Strukturen, weiter gibt es Reaktionen (z. B. Blickbewegungen) die in anderen Bereichen des peripheren Nervensystems entstehen. Für eine detaillierte Übersicht über die Verortung physiologischer Reaktionen im Nervensystem wird auf O'Donnell und Eggemeier (1986) verwiesen. Das periphere Nervensystem aktiviert die messbaren physiologischen Reaktionen des Körpers im Wesentlichen basierend auf den kognitiven Prozessen innerhalb des zentralen Nervensystems (u. a. Gehirn, Gehirnstamm, Spinalzellen). Die bewussten und unbewussten Verarbeitungs- und Adaptionsprozesse, welche kognitive Ressourcen verbrauchen und somit den mentalen Beanspruchungszustand definieren, lösen verschiedene physiologische Reaktionen aus (O'Donnell & Eggemeier, 1986). Es existiert eine Vielzahl an physiologischen Messinstrumenten zur Erfassung der mentalen Beanspruchung (u. A. Elektroenzephalografie, Elektrokardiografie, Elektrookulografie). Neben den Vorteilen von physiologischen Parametern, gibt es auch Nachteile, die durch die Vielzahl an Einflussfaktoren (z. B. bewusste und unbewusste Adaptionsprozesse, Veränderungen der Zielsetzung) entstehen. Viele aktuelle Forschungsansätze beschäftigen sich mit der Zuverlässigkeitssteigerung in der Interpretation physiologischer Parameter. Ryu und Myung (2005) kombinieren Messungen der Alpha-Wellen (EEG) mit der Herzratenvariabilität und der Blinzelrate über ein Regressionsmodell, um so mentale

Beanspruchung in einer Verfolgungsaufgabe zuverlässig bewerten zu können. Das Modell kann erfolgreich zwischen verschiedenen Anforderungsniveaus unterscheiden und korreliert signifikant ($p < 0,01$) mit der Leistung. Allerdings liegt eine sehr hohe Multikolarität der einzelnen Parameter vor, was die Interpretation erschwert (Ryu & Myung, 2005). Zhang et al. (2015) nutzen ebenfalls eine kombinierte Messung von der Elektroenzephalografie (EEG), Elektrokardiografie (EKG) und Elektrookulografie (EOG), um mit Hilfe einer adaptiven Support-Vector-Machine die Beanspruchung zu klassifizieren. Das Modell korreliert dabei signifikant ($p < 0,01$) mit dem subjektiven Beanspruchungsempfinden. Allerdings sind die verschiedenen Anforderungsniveaus der Studie nicht randomisiert, weshalb ein Lerneffekt die Ergebnisse beeinflusst. Weiter ist für das Vorgehen ein sehr komplexer Versuchsaufbau mit drei physiologischen Messsystemen notwendig (Zhang et al., 2015). Weitere Ansätze beschäftigen sich mit der Echtzeiterfassung der mentalen Beanspruchung durch die Kombination von Parametern (Hoover, Singh, Fishel-Brown & Muth, 2012).

Für die Entwicklung der individuellen Beanspruchungsmessung wurden physiologische Reaktionen mittels eines binokularen, kopfbasierenden Blickerfassungssystems erfasst. Die Blick- und Augenerfassung bietet essenzielle Vorteile bei der Verwendung in Feldstudien und praktischen Anwendungen, die das Zielsystem für die Entwicklungen dieser Arbeit darstellen. Das verwendete kopfbasierte Blickerfassungssystem der Firma Ergoneers GmbH (Dikalbis Professional 60Hz) ähnelt einem herkömmlichen Brillengestell, wobei zwei Infrarotkameras die Positions- und Größeninformationen der beiden Pupillen erfassen und eine Feldkamera das Blickfeld der Probanden aufzeichnet. Das System nutzt die Purkinje Lichtreflexionen 1. Ordnung¹, um die Position und die Größe der Pupille zu bestimmen. Nach einer vier-Punkte Kalibrierung²

¹ Lichtreflexe, die durch einfallendes Licht auf der Hornhaut entstehen. Als Lichtquelle werden Infrarotdioden nahe der Augenkameras verwendet, die nicht in deren optischen Achsen liegen (Dark Pupil Tracking), sodass Blendungseffekte verhindert werden können (Manzey, 1998)

² Der Proband schaut vier definierte Punkte an, um die Pupillenpositionen dem Blickpunkt zuzuordnen und so die Feldkamera mit den Augenkameras zu synchronisieren

erlaubt das System präzise Aufzeichnungen mit Abweichungen von bis zu $0,3^\circ$. Der Proband selbst ist in den Aufzeichnungen nicht zu sehen und kann auch durch die Pupilleninformationen nicht identifiziert werden, was zusammen mit dem vergleichsweise hohen Tragekomfort und dem leichten Anlegen zu einer sehr hohen Nutzakzeptanz führt. Darüber hinaus haben die physiologischen Reaktionen des Auges einen weiteren Vorteil bezüglich der kurzen Latenzzeiten durch die anatomische Nähe zu den informationsverarbeitenden, kognitiven Systemen, was wiederum die Sensitivität positiv beeinflusst (Sträter, 2005). Das menschliche Auge hat darüber hinaus eine sehr hohe sensorische Relevanz, da ungefähr 80% der Reize über die Augen wahrgenommen werden (Rötting, 2001). Im Rahmen dieser Arbeit werden sieben Blick- und Augenparameter zur Erfassung der mentalen Beanspruchung, basierend auf den Daten des Blickerfassungssystems, ausgewertet und für die Entwicklung der individuellen Beanspruchungsfunktionen berücksichtigt. So kann die Vielzahl an individuellen, kognitiven Einflussfaktoren (Kapitel 2) auf verschiedene Parameter erfasst und beschrieben werden. Die sieben Parameter werden zur Auswertung, nach Kenntnis des Autors, erstmals in bewusst und nicht-bewusst beeinflussbare Parameter differenziert, um die Unterschiede, die durch bewusste und unbewusste Verarbeitungsprozesse entstehen, abbilden zu können. Einige Studien können bereits mittels Priming-Experimenten oder Versuchen zur visuellen Trennung des rechten und linken Auges zeigen, dass Lidschlussparameter (z. B. Blinzelfrate, Blinzeldauer) (Shapiro, Arnell & Raymond, 1997; Ichikawa & Ohira, 2004) und die Pupillendurchschnittsvariabilität (Tamietto et al., 2010; Laeng, Sirois & Gredebäck, 2012) sensitiv gegenüber unbewussten Verarbeitungsprozessen sind. Veränderungen in diesen Parametern während der Bearbeitung einer Aufgabe können von den Probanden nicht verbalisiert werden. In einer Teilstudie zur Erprobung der, im Rahmen dieser Arbeit entwickelten, Kalibrierungsaufgabe (Kapitel 5) werden die Probanden ($N = 23$) nach der Bearbeitung der verschiedenen Anforderungsniveaus hinsichtlich durchgeführter Strategiewechsel befragt. Die subjektiven Angaben der Probanden werden dann mit dem Verhalten der okularen Beanspruchungsindikatoren (Kapitel 3.1.3.1 und 3.1.3.2) verglichen. Die Ergebnisse zeigen, dass es, unabhängig von dem Probanden,

der Aufgabenart und dem Anforderungsniveau, signifikante Unterschiede in der Blickstrategie (*NNI*) zwischen Versuchsteilen mit keinem, einem oder mehreren berichteten Strategiewechseln der Probanden gibt ($F(2, 44) = 5,88, p < 0,01$). Die Fixationsdauern ($F(2, 44) = 3,73, p < 0,05$) und die Sakkadenlängen ($F(2, 44) = 3,67, p < 0,05$) zeigen die gleichen Effekte. Weiter zeigen sich signifikante Unterschiede in den Varianzen des *NNI* ($t(19,36) = -2,19, p < 0,05$) und der Sakkadenlängen ($t(14,62) = -2,14, p < 0,05$) zwischen Versuchsteilen mit verbalisierten Strategiewechseln und denen ohne. Für die unbewussten Parameter (*Blinzelrate*, *Blinzeldauer*, *PERCLOS* und *Pupillendurchschnittsvariabilität*) zeigen sich keine Effekte. Basierend auf den Ergebnissen und den Befunden aus der Literatur kann die Unterteilung in bewusste und unbewusste Parameter motiviert und begründet werden. Im Folgenden sollen die einzelnen Parameter und deren Berechnung beschrieben werden.

3.1.3.1 Bewusst beeinflussbare Parameter

Die bewusst beeinflussbaren Parameter bilden hauptsächlich die bewusst ablaufenden Verarbeitungsprozesse ab und deren Veränderungen können von den Probanden nach einer Aufgabe verbalisiert werden. Es werden drei bewusste, okulare Parameter erfasst: Der *Nearest Neighbor Index (NNI)*, die *Fixationsdauern (FD)* und die *Sakkadenlängen (SK)*. Die Augenbewegungen werden durch das periphere Nervensystem, allerdings nicht durch das autonome Nervensystem (ANS) gesteuert, was ein wesentlicher Grund für das Bewusstsein über diese physiologischen Reaktionen ist (O'Donnell & Eggemeier, 1986). Grundsätzlich werden Augen- von Blickbewegungen unterschieden. Augenbewegungen beschreiben die Bewegungen des Auges, die lediglich durch die Erfassung des Auges analysiert werden können. Blickbewegungen sind Augenbewegungen unter Berücksichtigung der betrachteten Objekte (Rötting, 2001). Die drei bewussten Parameter, die im Folgenden beschrieben werden, können als Augenbewegungen klassifiziert werden, da kein konkreter Objektbezug besteht oder hergestellt werden soll. Das menschliche Blickfeld umfasst einen Kegel mit Radius $r > 100^\circ$. Der foveale Bereich des scharfen Sehens ist deutlich kleiner, weshalb Augenbewegungen notwendig sind, um Informationen von verschiede-

nen Objekten oder Orten wahrzunehmen (Rötting, 2001). Sakkaden beschreiben die sprunghaften Bewegungen, um den Blick auf eine neue Informationsquelle zu richten. Sakkaden werden sowohl unwillkürlich von Bewegungsänderungen im peripheren Sichtfeld, als auch von gezielten Verhaltensplänen bei der Wahrnehmung und Interpretation von Informationen ausgelöst. Es existiert eine Vielzahl an kognitions- und neurowissenschaftlichen Ansätzen zur Beschreibung der Entstehung von Sakkaden. Sakkaden haben eine Dauer von etwa 30 ms bis 40 ms wobei Geschwindigkeiten bis 1000°/s erreicht werden können. Bis zu 120 ms vor und nach dem Beginn einer Sakkade ist das Wahrnehmungsvermögen deutlich reduziert (bei kurzen Sakkaden also auch während der folgenden Fixation). Einige theoretische und empirische Arbeiten können die Sensitivität der *Sakkadenlänge* für Such-, Informationsaufnahme- und Verarbeitungsprozesse abbilden und erklären (Rötting, 2001; Cardona & Quevedo, 2014; Bodala, Ke, Mir, Thakor & Al-Nashash, 2014; Matthews, Reinerman-Jones, Barber & Abich, 2015; May, Kennedy, Williams, Dunlap & Brannan, 1990), wonach in dieser Arbeit die durchschnittliche Sakkadenlänge in 30 s Intervallen als bewusster Indikator für die mentale Beanspruchung berechnet wird. Befindet sich das Auge in relativer Ruhe zu einem Suchobjekt, um Informationen aufzunehmen, wird von einer Fixation gesprochen (Rötting, 2001). Die minimale Fixationsdauer, um Informationen aufnehmen zu können, liegt bei 175 ms, wobei die Dauer in Abhängigkeit von vielen Einflussfaktoren (z. B. Alter, Umweltbedingungen, Informationsdarstellung) schwankt (Sträter, 2005). Studien zur Erfassung der mentalen Beanspruchung zeigen, dass die Fixationsdauer in einem direkten Zusammenhang mit der benötigten Zeit zur Informationsaufnahme oder der Zeit zum Verständnis einer wahrgenommenen Information steht und so Rückschlüsse auf die mentale Beanspruchung zulässt (Di Stasi, Antolí & Canas, 2013; Matthews et al., 2015). In den Studien dieser Arbeit wird die durchschnittliche Fixationsdauer in je 30 s Intervallen ausgewertet. Weiter wird in dieser Arbeit die Blickstrategie, operationalisiert über den *Nearest Neighbor Index (NNI)*, zur Bewertung der mentalen Beanspruchung herangezogen. Die Berechnung des NNI basiert auf dem Zusammenhang zwischen der Entropie (Zustand der Unordnung) des Blickverhaltens der mentalen Beanspruchung, der bereits von

Harris, Tole, Stephens und Ephrat (1982), Harris, Glover und Spady (1986) und Hilburn, Jorna, Byrne und Parasuraman (1996) erforscht wurde. In den Studien wurden a priori bestimmte statische Areas of Interest (AOIs) definiert, um Blickwechsel zwischen den AOIs in Übergangsmatrizen hinsichtlich der Entropie zu untersuchen. Die Ergebnisse zeigen signifikante Zusammenhänge zwischen der Entropie und der mentalen Beanspruchung. Di Nocera, Terenzi und Camilli (2006) entwickelten, basierend auf diesen Erkenntnissen, einen neuen Ansatz zur Untersuchung der Entropie des Blickverhaltens. Durch die Verwendung des NNI, der ursprünglich von Clark und Evans (1954) zur Untersuchung des räumlichen Verteilungsverhaltens von Pflanzenpopulationen entwickelt wurde, ist es nicht länger notwendig, a priori bestimmte AOIs zu definieren. Der NNI (Formel 3.2) beschreibt das Verhältnis zwischen dem durchschnittlichen Mindestabstand von Punkten in einer vorliegenden Verteilung und der mittleren Distanz einer komplett zufälligen und regellosen Verteilung von der gleichen Anzahl an Punkten (Di Nocera et al., 2006).

$$NNI = \frac{d(NNI)}{d(ran)} \quad (3.2)$$

$$d(NNI) = \sum_{i=1}^N \left[\min \frac{(d_{ij})}{N} \right], 1 \leq j \leq N, i \neq j \quad (3.3)$$

$$d(ran) = 0,5 \times \sqrt{\frac{A}{N}} \quad (3.4)$$

Dabei wird der durchschnittliche Mindestabstand der aktuellen Verteilung von Fixationen, $d(NNI)$, über Summe an Distanzen von einer Fixation i zu der am nächsten liegenden Fixation j (wobei $i \neq j$), dividiert durch die Anzahl N der Fixationen, berechnet (Formel 3.3). Die mittlere Distanz einer zufälligen Verteilung, $d(ran)$, wird aus der Relation der Fläche A , in der die Fixationen liegen, und der Anzahl an Fixationen N berechnet (Formel 3.4). Für die inhaltliche Interpretation des NNI wird der Wertebereich $NNI \in [0; 2,1491]$ in drei Bereiche differenziert (Di Nocera, Camilli & Terenzi, 2006). Hat der NNI einen Wert von 1, entspricht die aktuelle Verteilung genau der rein zufälligen Verteilung

von Fixationen. Für den $NNI \in [0; 1)$ bilden sich Informationsgruppen und für den $NNI \in (1; 2,1491]$ liegt eine regelmäßige Verteilung vor, wobei nach Clark und Evans (1954) für $NNI = 2,1491$ eine exakt hexagonale Verteilung vorliegt (Abbildung 3.3). Entgegen den ursprünglichen Erkenntnissen von Harris et al. (1982), Harris et al. (1986) und Hilburn et al. (1996) zeigen die Ergebnisse der Studien von Di Nocera et al. (2006), Di Nocera et al. (2006) und Di Nocera, Camilli und Terenzi (2007), dass der NNI hoch sensitiv für die mentale Beanspruchung ist, sich die Beanspruchung aber proportional zu einer zunehmenden Entropie verhält. Konvergiert der NNI gegen 1, so wird eine maximale Beanspruchung postuliert. Diese Ergebnisse decken sich mit den Erkenntnissen aus den theoretischen Ansätzen zur menschlichen Informationsverarbeitung und mentalen Beanspruchung.

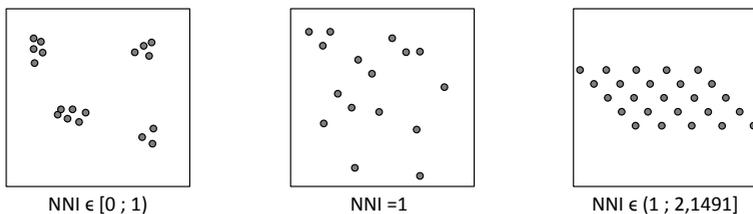


Abb. 3.3: Verteilung von Fixationen im Wertebereich des NNI (eigene Abbildung)

Je höher die Beanspruchung, desto weniger folgt das Blickverhalten einer klaren Strategie. Es wird mehr und mehr von nicht-zielgerichtetem Suchverhalten nach Informationen bestimmt (Sträter, 2005). Eine klare Gruppierung von Fixationen ($NNI < 1$) entspricht einer eindeutigen Strategie, bei der der Operateur genau weiß, wo er welche benötigten Informationen zu welchem Zeitpunkt findet. Ein reguläres Blickverhalten ($NNI > 1$) impliziert schon Ansätze des Suchverhaltens beziehungsweise Strategiewechsel, die auf ein regelbasiertes Verhalten und somit auf vorliegende kognitive Dissonanzen schließen lassen (Sträter, 2005). Ein wichtiger Aspekt bei der Berechnung des NNI ist die Definition der Bezugsfläche A . Das ursprüngliche Verfahren zur Berechnung des NNI nach Di

Nocera et al. (2006) verwendet ein Rechteck basierend auf der maximalen Verteilungsbreite in X- und Y-Richtung. Liegt eine Fixation aufgrund einer Interferenz oder Bildstörung der Aufzeichnung außerhalb des aufgabenrelevanten Blickfeldes, wirkt sich dies sofort artifiziell auf den Wert des NNI aus. Die Bezugsfläche ist dann überproportional größer, als die tatsächliche Fläche der Verteilung basierend auf einer konvexen Hülle. Aufgrund dessen wird im Rahmen dieser Arbeit der NNI neben der ursprünglichen Bezugsfläche auch mit einer dynamischen Ellipsenfläche berechnet. Die Berechnungsmethode der Ellipse wird ursprünglich im Bereich der Bewegungsanalyse, genauer zur Schwerpunktberechnung des menschlichen Körpers, angewendet. Die Ellipse wird bezüglich der Orientierung und der Größe dynamisch an die Verteilung der Fixationen angepasst, sodass sich die Fläche optimal an die vorliegende Verteilung anschmiegt (Schubert & Kirchner, 2014). Weiter wird der NNI in Zeitintervallen von 30 s berechnet, um sicherzustellen, dass eine ausreichende Anzahl an Fixationen vorliegt (Di Nocera et al., 2007).

3.1.3.2 Nicht-bewusst beeinflussbare Parameter

Die nicht-bewusst beeinflussbaren Parameter, umfassen die *Blinzelrate (BR)*, die *Blinzeldauer (BD)*, den *PERCLOS (percentage of eye closure, PC)* und die *Pupillendurchschnittsvariabilität (PD)*. Diese werden durch das autonome Nervensystem (ANS) ausgelöst und können von Probanden nicht verbalisiert werden. Zahlreiche Studien haben gezeigt, dass diese Parameter sensitiv für unbewusst ablaufende Verarbeitungsprozesse sind (Tamietto et al., 2010; Laeng et al., 2012; Ichikawa & Ohira, 2004; Lin & He, 2009; Shapiro et al., 1997; Dovidio, Kawakami & Johnson, 1997). Viele Modelle der menschlichen Informationsverarbeitung beschreiben die bewusst wahrgenommenen Informationen als allein verantwortlich für Entscheidungen und Handlungen (van Gaal & Lamme, 2012). Aktuelle theoretische und empirische Forschungen zeigen, dass unbewusste Verarbeitungsprozesse, ohne bewusste oder freiwillige Kontrolle, einen erheblichen Teil des alltäglichen Verhaltens bestimmen (Dehaene et al., 2006; Kiefer et al., 2012; van Gaal & Lamme, 2012). Bestimmte Arten von unbewussten Prozessen, die sich zwar von bewussten Prozessen unterschei-

den, aber trotzdem mentale Ressourcen verbrauchen, machen eine getrennte Betrachtung der unbewussten Indikatoren für eine zuverlässige Beanspruchungsmessung notwendig.

Die *Blinzelrate* beschreibt die Anzahl der Lidschlüsse in einem definierten Zeitintervall. Ein Lidschluss dauert zwischen 70 - 100 ms (Marshall, 18.07.2000). In der Regel werden Betrachtungsintervalle von einer Minute gewählt, welche aber für die Auswertung von sehr dynamischen oder kurzzyklischen Tätigkeiten auch auf 30 s verkürzt werden können (Wilson, 2009). Zahlreiche Veröffentlichungen untersuchten bereits den Zusammenhang zwischen der Blinzelrate und der mentalen Beanspruchung. Die Studien von Veltman und Gaillard (1998) sowie von Wilson (2009) und Jessee (2010) zeigen die Abnahme der Frequenz spontaner, nicht willkürlicher Lidschlüsse mit steigender mentaler, insbesondere visueller Beanspruchung. Weiter zeigten Goldstein, Walrath, Stern und Strock (1985) neben der abnehmenden Frequenz der Lidschläge zusätzlich eine kürzere Dauer spontaner Lidschlüsse. Die genannten Studien verwenden ausschließlich visuelle Diskriminationsaufgaben. Ergebnisse aktuellerer Studien in Dual-Task-Situationen zeigen eine Zunahme der Blinzelrate mit steigenden Aufgabenanforderungen (Faure, Lobjois & Benguigui, 2016; Chen & Epps, 2014; Recarte, Perez, Conchillo & Nunes, 2008). Teilweise kann die Diskrepanz zwischen den verschiedenen Studien auf die Abhängigkeit der Blinzelrate von unbewussten Verarbeitungsprozessen zurückgeführt werden. Unbewusste Prozesse unterscheiden sich mitunter sehr stark, basierend auf den individuellen Erfahrungen und Konzepten sowie der aktuellen Zielsetzung, gespiegelt an den Umgebungsfaktoren (Sträter, 2005; Kiefer et al., 2012; Dehaene & Changeux, 2005). Die Blinzelrate verspricht demnach insbesondere auf differenzierter und individueller Auswertungsebene eine zuverlässige Beanspruchungsvorhersage. Die *Blinzeldauer* kann zwischen 70 - 500 ms liegen und zeigt in einer Studie von Benedetto et al. (2011) eine zuverlässigere Beanspruchungsbewertung als die Blinzelrate. In der verwendeten Lane-Change-Task in Kombination mit einer visuellen Nebenaufgabe zeigt sich eine Abnahme der durchschnittlichen Blinzeldauer mit zunehmender Beanspruchung (Benedetto et al., 2011).

Brookhuis und De Waard (2010) berechnen neben der durchschnittlichen Blinzeldauer auch die kumulierte Blinzeldauer, den *PERCLOS* (Martins & Carvalho, 2015), in einer simulierten Fahraufgabe. Es zeigen sich signifikante Zusammenhänge mit der Leistung der Probanden. Aufgrund dieser vielversprechenden Ergebnisse wird der *PERCLOS*, neben der Blinzeldauer und der Blinzelrate, in den folgenden Auswertungen berücksichtigt. Als letzter physiologischer Beanspruchungsindikator wird die *Pupillendurchschnittsvariabilität* betrachtet. Die menschliche Pupille ist eine Öffnung in der Iris und verhält sich wie eine Linse, die die Menge des einfallenden Lichts reguliert, dieses bricht und bündelt (Beatty & Lucero-Wagoner, 2007). Die Iris verfügt über zwei antagonistische Muskeln, welche die Pupille umgeben: der *Musculus Sphincter Pupillae* (verkleinert die Pupille) und der *Musculus Dilator Pupillae* (vergrößert die Pupille) (Abbildung 3.4). Beide Muskeln werden durch verschiedene Komponenten des Nervensystems innerviert: Der Sphincter wird durch den Parasymphikus und der Dilator durch den Sympathikus aktiviert. Jede Größenänderung der Pupille ist somit das Resultat sowohl sympathischer als auch parasymphischer Aktivität, wobei sich die letztendliche Größe der Pupille aus dem Verhältnis der Aktivitäten beider Systeme bestimmt (Schwalm, 2009).

Einige Veränderungen der Pupillengröße sind rein optische Prozesse. Der pupillometrische Lichtreflex reguliert die Menge des einfallenden Lichtes und adaptiert die Pupille so an die vorliegenden Beleuchtungsbedingungen. In hellen Umgebungen ist die Pupille klein und zusammengezogen und in dunklen Umgebungen geöffnet und entspannt. Während des *Lichtreflexes* von dunkel nach hell (Licht an) wird der Sphincter als Agonist angeregt sich zusammen zu ziehen. Der Dilator wird als Antagonist gehemmt und so daran gehindert sich ebenfalls zusammen zu ziehen – die Pupille verkleinert sich. Bei dem Lichtreflex von hell nach dunkel fällt die Innervation des Sphincters weg, die Kontraktion lässt nach und die Pupille entspannt sich (Beatty & Lucero-Wagoner, 2007; Schwalm, 2009). Ein weiterer Reflex innerhalb der Pupille ist der *Akkomodationsreflex*. Dieser verändert die Krümmung der Pupille, um so Objekte in unterschiedlichen Entfernungen scharf sehen zu können, was sich ebenfalls auf die Pupillengröße auswirkt. Aufgrund der Innervation der Pupillenmuskeln durch parasymph-

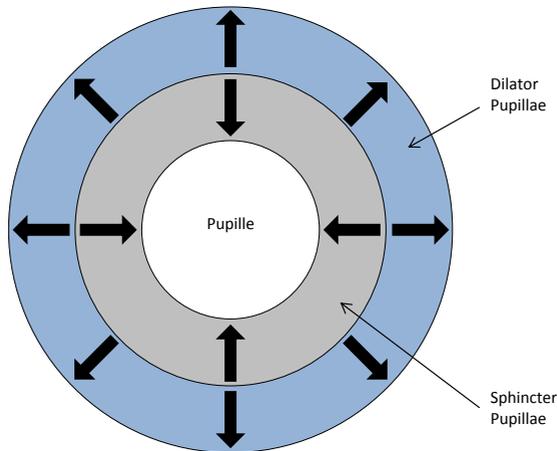


Abb. 3.4: Muskeln der Pupille. Die Pfeile entsprechen der Funktionsrichtung der Muskeln (Schwalm, 2009, S. 49)

sche und insbesondere sympathische Aktivität, verändert sich die Pupillengröße, neben den rein optischen Adaptionen, auch aufgrund der mentalen Aktivität und Beanspruchung (Kahneman, 1973). Der *psychosensorische Reflex* (Loewenfeld, 1993) gilt als Indikator für Vorgänge innerhalb des zentralen Nervensystems, welche der Dynamik sowie Intensität der menschlichen Informationsverarbeitung unterliegen und somit mentale Aktivität und Beanspruchung indizieren (Beatty & Lucero-Wagoner, 2007; Steinhauer, Siegle, Condray & Pless, 2004; Loewenfeld, 1993). Während Aufgaben mit signifikanter mentaler Beanspruchung wird der Prozess der reziproken Innervation umgekehrt. Der Dilator wird sympathisch angeregt sich zusammen zu ziehen, was eine Erweiterung der Pupille bewirkt. Zeitgleich wird der Sphincter inhibiert, was ebenfalls eine Erweiterung der Pupille bewirkt. Diese Kombination führt zu unregelmäßigen, sehr abrupten Fluktuationen der Pupille, die häufig durch extrem starke Anstiege gefolgt von schnellen Abfällen in der Pupillengröße charakterisiert sind (Marshall, 18.07.2000). Eine Vielzahl an Studien untersuchte bereits die Sensitivität des psychosensorischen Reflexes hinsichtlich der mentalen Beanspruchung. Für eine tabellarische Übersicht soll an dieser Stelle auf Schwalm (2009, S. 55-56)

verwiesen werden. Ein essentielles Problem bei der Erfassung der mentalen Beanspruchung liegt in der Überlagerung der einzelnen Reflexe, die durch geeignete Verfahren getrennt werden müssen. Bisherige Vorgehensweisen basieren auf der relativen Größenänderung im Vergleich zu einer Baseline Bedingung oder auf dem gleitenden Mittelwert der Pupillenfläche (Rößger, 1997; Richer, Silverman & Beatty, 1983). Diese Ansätze sind jedoch nur sehr begrenzt anwendbar, da sie gleichbleibende Beleuchtungsverhältnisse in der Baseline und der gesamten Versuchszeit voraussetzen. Darüber hinaus werden individuelle Unterschiede in der Pupillengröße nur unzureichend berücksichtigt (Schwalm, 2009). Ein neuer Ansatz nach Marshall (18.07.2000), der *Index of Cognitive Activity (ICA)*, distanziert sich von der Betrachtung der absoluten Pupillenfläche und fokussiert die Unterscheidung von der kurzfristigen und unregelmäßigen Charakterisierung des psychosensorischen Reflexes im Vergleich zu langsameren und kontinuierlicheren Licht- und Akkomodationsreflexen. Es gibt bereits vielseitige empirische Evidenzen für die Sensitivität des ICA hinsichtlich der mentalen Beanspruchung unabhängig von den Beleuchtungsverhältnissen (Marshall, 2007, 2002; Marshall, Pleydell-Pearce & Dickson, 2003; Schwalm, 2009; Demberg, 2013). Marshall (18.07.2000) nutzt eine *Wavelet-Analyse*, um die psychosensorischen Reflexe erfolgreich zu extrahieren. Die Wavelet-Analyse ist ein weit verbreitetes Verfahren in der Signalverarbeitung und -dekomposition, wobei ein Signal orthogonal transformiert wird, sodass die Frequenz über die Zeit aufgetragen werden kann (Daubechies, 2006). Während der Transformation wird das Signal in seine approximativen Komponenten (niederfrequent) und die Detailkomponenten (hochfrequent) zerlegt, was eine schnelle Extrahierung von hochfrequenten, unregelmäßigen Veränderungen der Pupillenfläche ohne einen großen Verlust der zeitlichen Auflösung erlaubt (Wilkinson & Cox, 1996). Die nachstehenden Transformationsformeln (Formel 3.5 und 3.6) beschreiben die Komponenten (cA = approximativ und cD = Detail) der Daubechies db4 Transformation (Wilkinson & Cox, 1996), die laut Marshall (18.07.2000) besonders für eine Aufnahme Frequenz von 60 Hz des hier verwendeten Blickerfassungssystems geeignet ist. Die Daubechies db4 Wavelet-Transformation benötigt lediglich vier Koeffizienten Φ_i . Bei einer Aufnahme mit 60 Hz entspricht ein

Frame einer Dauer von 16,7 ms, sodass nur $4 \times 16,7 \text{ ms} = 66,8 \text{ ms}$ des Ursprungssignals Φ zu einem orthogonal transformierten Punkt in cA_i und cD_i zusammengefasst werden (Wilkinson & Cox, 1996).

$$cA_i = \frac{1}{4\sqrt{2}} (\Phi_{2i}(1+\sqrt{3}) + \Phi_{2i+1}(3+\sqrt{3}) + \Phi_{2i+2}(3-\sqrt{3}) + \Phi_{2i+3}(1-\sqrt{3})) \quad (3.5)$$

$$cD_i = \frac{1}{4\sqrt{2}} (\Phi_{2i}(1-\sqrt{3}) + \Phi_{2i+1}(\sqrt{3}-3) + \Phi_{2i+2}(3+\sqrt{3}) + \Phi_{2i+3}(-1-\sqrt{3})) \quad (3.6)$$

Das Verfahren der Wavelet-Transformation ist iterativ. Die vorstehenden Formeln beschreiben eine Level 1 - Transformation ausgehend von dem Ursprungssignal. Der approximative Teil des transformierten Signals kann jetzt erneut in approximative Komponenten und Detailkomponenten zerlegt werden (Level 2). Dieses Vorgehen kann wiederholt werden, bis eine weitere Zerlegung aufgrund der Signallänge nicht mehr möglich ist (Daubechies, 2006). Abbildung 3.5 stellt schematisch dar, dass die Auflösung des Frequenzbereichs auf jeder Stufe ansteigt, wobei dies zu Lasten der zeitlichen Auflösung geht.

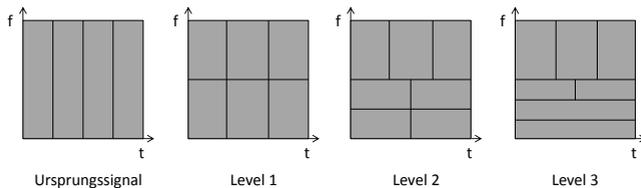


Abb. 3.5: Stufen der Wavelet-Transformation (eigene Abbildung)

Durch die Charakterisierung des psychosensorischen Reflexes als hochfrequent und unregelmäßig (Marshall, 18.07.2000), reicht eine Level 1 - Transformation aus, da so noch eine gute zeitliche Verortung vorgenommen werden kann. Im nächsten Schritt der Analyse müssen die auffälligen, unregelmäßigen Charakteristika der Detailkomponenten, die den psychosensorischen Reflex repräsentieren, extrahiert werden. Dazu etablierte sich ein Verfahren nach Ogden und Parzen (1996) und Ogden (1997), die nach statistisch signifikanten Sprüngen in dem Signal suchen und diese über einen Grenzwert extrahieren. Bei dem

Verfahren des „*thresholding*“ bleiben nur die Komponenten des Signals erhalten, die oberhalb des festgelegten Grenzwertes liegen, das restliche Signal wird gleich Null gesetzt. So unterscheiden sich nur diejenigen, besonders auffälligen, Charakteristika des Signals von Null, welche laut Marshall (18.07.2000) den psychosensorischen Reflexen entsprechen. Es gibt viele Verfahren, um die Grenzwerte festzulegen, deren Auswahl im Wesentlichen von der Qualität und dem Rauschen (hier: Licht- und Akkomodationsreflexe) des Signals abhängt. Marshall (18.07.2000) empfiehlt einen *Minimax-Grenzwert*, bei dem eine harte und eine weiche Grenze definiert wird und der Zwischenbereich noch einmal geglättet und anschließend zugeordnet wird. Durch die hohe Individualität der physiologischen Reaktionen, werden in dieser Arbeit vier verschiedene Grenzwerte verwendet (*hart, weich, universal, standardisiert*), die in Tabelle 3.2 zusammengefasst sind.

Tab. 3.2: Berechnung der Grenzwerte für die Wavelet-Analyse (eigene Tabelle)

Grenzwert	Berechnung	
hart	$\sqrt{2 \times \log(n)}$	(3.7)
weich	$\sqrt{2 \times \log(n) \times \log(n)}$	(3.8)
universal	$\sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n}} \times \sqrt{\frac{2 \times \log(n)}{n}}$	(3.9)
standardisiert	$\sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n}} \times \sqrt{2 \times \log(n)}$	(3.10)

Für eine genaue Herleitung und Beschreibung der Berechnungsverfahren soll an dieser Stelle auf Donoho und Johnstone (1994), Donoho (1995), Donoho, Johnstone, Kerkyacharian und Picard (1996), Donoho und Johnstone (1998) und Chang, Yu und Vetterli (2000) verwiesen werden. Im letzten Schritt der Analyse wird die Anzahl der psychosensorischen Reflexe pro Sekunde durch die maximale Anzahl der physiologisch möglichen Dilationen pro Sekunde (30) dividiert, um einen einheitlichen, vergleichbaren und skalierbaren Parameter zu erhalten (Demberg, 2013). Aus der technischen Umsetzung von kopfbasierten

Blickerfassungssystemen mit einer festen Position der Augenkameras, leicht unterhalb der Augen, ergibt sich allerdings ein weiterer Störeinfluss auf die Erfassung des psychosensorischen Reflexes, der in den bisherigen Forschungen noch nicht berücksichtigt wurde. Bei Augenbewegungen nach oben und unten verändert sich die, durch die Augenkamera erfassbare, Fläche der Pupille erheblich. Diese Bewegungsartefakte können leicht als unregelmäßige, abrupte Veränderungen und somit als psychosensorische Reflexe klassifiziert werden. In der Berechnung der Pupillendurchschnittsvariabilität werden, im Rahmen dieser Arbeit, auch die Bewegungen um die vertikale y-Achse als Kontrollvariable berücksichtigt. Psychosensorische Reflexe, die zu Zeitpunkten einer Pupillenbewegung nach oben oder unten detektiert werden, werden ausgeschlossen. Die Studie 1 (Kapitel 4) hat gezeigt, dass mit diesem Vorgehen aufgabenabhängig etwa ein bis zwei Reflexe weniger erfasst werden. Die genaue Sichtprüfung der Videoaufzeichnungen bestätigt dann die Augenbewegungen als Auslöser für die herausgefilterten Reflexe.

Abschließend zur Beschreibung der physiologischen Beanspruchungsparameter, fasst die folgende Tabelle nochmals alle, in der Auswertung verwendeten, Parameter mit einer Auswahl der wichtigsten Referenzen zusammen.

Tab. 3.3: Übersicht der in dieser Arbeit verwendeten physiologischen Beanspruchungsindikatoren (eigene Tabelle)

Parameter	Kurzbeschreibung	Verweise
<i>bewusste Parameter</i>		
∅-Fixationsdauer	Durchschnittliche Dauer von Fixationen in 30 s Intervallen	Di Stasi et al. (2013),- Matthews et al. (2015)
∅-Sakkadenlänge	Mittlere Länge der Sakkaden in 30 s Intervallen	Cardona und Quevedo (2014), Rötting (2001)
NNI	Statistisches Verteilungsmaß für Fixationen im Blickfeld	Di Nocera et al. (2006), Di Nocera et al. (2007)
<i>unbewusste Parameter</i>		
Blinzelrate	Anzahl von Lidschlüssen in 30 s Intervallen	Faure et al. (2016), Chen und Epps (2014)
Blinzeldauer	Mittlere Dauer von Lidschlüssen in 30 s Intervallen	Goldstein et al. (1985),- Benedetto et al. (2011)
PERCLOS	Kumulierte Lidschlussdauer über 30 s Intervalle	Brookhuis und De Waard (2010), Martins und Carvalho (2015)
Pupillendurchschnittsvariabilität	Mittlere Anzahl an Fluktuationen der Pupille in 5 s Intervallen	Marshall et al. (2003),- Schwalm (2009)

3.2 Verfahren der Datenanalyse

Im folgenden Abschnitt werden die verwendeten Verfahren zur Datenanalyse beschrieben. Zunächst wird auf die Procrustes-Analyse eingegangen, welche später für die Mustererkennung der individuellen Beanspruchungsmodellbildung verwendet wird. Im Anschluss werden die statistischen Verfahren erläutert, die für die Auswertung der Daten gewählt wurden. Der Beanspruchungswerte aus den leistungs-basierten und subjektiven Erhebungsverfahren werden vor der Auswertung normiert, indem die jeweiligen Werte durch den Skalenumfang dividiert werden. Die physiologischen Parameter bleiben unverändert, um eine direkte Verarbeitung der Daten zu ermöglichen.

3.2.1 Procrustes-Analyse

Entwickelt von Gower (1975) ist die *Procrustes-Analyse* heute ein weit verbreitetes Verfahren der Mustererkennung, welches häufig in der Bewegungsanalyse und Bildverarbeitung Anwendung findet (Osis et al., 2015). Ziel der Analyse ist es, die Ähnlichkeit zweier n -dimensionaler Signale zu prüfen und zu quantifizieren. Im Rahmen dieser Arbeit werden die Signale der physiologischen Parameter über die Zeit aufgetragen und mit einer Referenzgröße (z. B. Leistung, subjektives Beanspruchungsempfinden) verglichen. Beide Signale liegen in $m \times 2$ Matrizen vor (m = Anzahl Messpunkte). Die Procrustes-Analyse besteht aus drei Schritten, die sequentiell durchgeführt werden (Handl, 2010). Zunächst werden beide Signale translatiert, sodass sie den gleichen Mittelpunkt (Ursprung) besitzen. Dazu werden die Mittelwerte der x -Koordinaten und der y -Koordinaten von den einzelnen Koordinatenpunkten k der Matrizen subtrahiert. Im zweiten Schritt werden die Signale skaliert, sodass sie eine einheitliche Größe haben. Osis et al. (2015) verwenden für diesen Schritt einen Skalierungsfaktor s , durch den jeder Koordinatenpunkt k der Matrix dividiert wird (Formel 3.11). In anderen Varianten der Procrustes-Analyse wird nur ein Signal durch ein Minimierungsverfahren skaliert, sodass der durchschnittliche Abstand der einzelnen Punkte minimiert wird (Handl, 2010). Da die Signale der hier betrachteten Daten aufgrund der gleichen x -Koordinaten (Messzeitpunkte)

bereits eine hohe Ähnlichkeit bezüglich der Skalierung aufweisen, erweist sich das Verfahren nach Osis et al. (2015) als effizienter und effektiver.

$$s = \sqrt{\frac{\sum_{n=1}^k (x_n^2 + y_n^2)}{k}} \quad (3.11)$$

Nach der Skalierung wird das Signal des physiologischen Parameters um den Winkel $\alpha = 1^\circ$, $\alpha \in [1, 360]$ rotiert, indem die Matrix mit der Rotationsmatrix multipliziert wird (Handl, 2010). In Folge der Rotation wird die mittlere quadratische Abweichung d zwischen den Punkten der Parametermatrix und den Punkten der Referenzmatrix berechnet. Dieser Vorgang wird für alle Winkel α des Wertebereichs wiederholt. Schließlich wird der Winkel α ausgewählt, für den der mittlere Abstand d zwischen den Signalen minimal ist (Handl, 2010). Die Werte von α und d sind das Ergebnis der Procrustes-Analyse (Osis et al., 2015) und geben Aufschluss darüber, wie ähnlich sich die zwei Signale sind (d) und wie stark sie dazu verändert werden mussten (α). In Kapitel 5.2.3 wird das Verfahren verwendet, um die Ähnlichkeit einzelner physiologischer Indikatoren zu einer Referenzgröße zu ermitteln.

3.2.2 Statistische Verfahren

Hinsichtlich der statistischen Auswertung der Daten werden die Stichprobengrößen so gewählt, dass sie nach dem *zentralen Grenzwerttheorem* ($N > 30$) der Normalverteilung entsprechen (Bortz & Schuster, 2010). Zudem wird die Normalverteilung vor jeder Auswertung unter Verwendung des *Shapiro-Wilk-Tests* geprüft (Field, 2013). Im Falle einer vorliegenden Normalverteilung werden parametrische Testverfahren für die Untersuchung der Daten verwendet. Zur Prüfung von linearen Zusammenhängen wird der Korrelationskoeffizient r nach Bravais-Pearson verwendet, wobei die Varianzaufklärung r^2 als wichtiges Gütekriterium herangezogen wird (Field, 2013). Diese Analyse wird in den folgenden Studien im Wesentlichen für die Prüfung der Diagnostizität von physiologischen Beanspruchungsparametern (Kapitel 3.1) verwendet. Für die Untersuchung von Unterschieden zwischen verschiedenen Anforderungsnive-

aus bei zwei Schwierigkeitsstufen wird der messwiederholte *t-Test* verwendet. Für die Differenzierung von drei oder mehr Schwierigkeitsstufen wird auf eine messwiederholte *ANOVA* mit Bonferroni-korrigierten *t-Tests* als Post-Hoc-Test zurückgegriffen (Field, 2013). Mit Hilfe dieser Analysen wird innerhalb der Auswertungen die Sensitivität der verschiedenen Beanspruchungsindikatoren hinsichtlich der Anforderungsniveaus untersucht. Im weiteren Verlauf der Auswertungen wird eine einfaktorische, *lineare Regression* verwendet, um die verschiedenen Beanspruchungsindikatoren in einen funktionalen Zusammenhang zu bringen (Bortz & Schuster, 2010). Insofern nicht-normalverteilte Daten vorliegen, werden nicht-parametrische Testverfahren verwendet. Die Korrelationsanalysen werden nach Spearman durchgeführt (Field, 2013). Für die Differenzierung der Anforderungsniveaus wird der *Mann-Whitney-U-Test* (2 Bedingungen) oder der *Friedman-Test* mit dem *Wilcoxon-Vorzeichen-Rang-Test* als Post-Hoc-Test (3 oder mehr Bedingungen) verwendet.

Für die Extrahierung von Gruppen zur Auswertung auf der differenzierten (Gruppen-) Ebene wird die *multidimensionale Skalierung (MDS)* in Kombination mit der *k-Means* Analyse verwendet. Die MDS gruppiert Punkte anhand ihrer Ähnlichkeit, gemessen über die Euklidischen Distanzen, im zweidimensionalen Raum und gibt so einen optischen Eindruck über zusammengehörige Gruppierungen. Die *k-Means* Analyse ist ein standardisiertes Instrument der Statistik und unterteilt die Cluster durch die Bildung und Verschiebung von *k* Clusterzentren (Mittelpunkte), von denen ausgehend die euklidischen Distanzen zu allen umliegenden Punkten berechnet werden. Zwei Cluster unterscheiden sich signifikant, wenn die durchschnittliche Distanz der Punkte aus Cluster A zu dem Clusterzentrum A signifikant geringer ist, als zu Clusterzentrum B des Clusters B und umgekehrt. Die Anzahl $k \in [1, n]$ der Cluster ist dabei eine Eingabegröße in die Analyse und wird iterativ von 1 bis *n* erhöht. Es wird die Anzahl *k* und Anordnung der Cluster gewählt, die die stärkste Heterogenität (Signifikanzniveau der Varianzanalyse zwischen den Distanzen) zwischen den Clustern und die stärkste Homogenität (kleiner Clusterumfang) innerhalb der Cluster ausweist. Innerhalb der Auswertungen werden die Stichproben anhand

von definierten Kriterien in kleinere Gruppen unterteilt, um so eine differenzierte Betrachtung der mentalen Beanspruchung zu ermöglichen.

4 Studie 1: Evaluation der physiologischen Parameter

Ziel der ersten Studie ist die Untersuchung der physiologischen Beanspruchungsparameter hinsichtlich der Diagnostizität und Sensitivität auf Stichproben- (genereller Ansatz), Gruppen- (differenzieller Ansatz) und Individualebene unter kontrollierten Laborbedingungen. Die Studie dient zur Replikation vorhandener theoretischer und empirischer Arbeiten hinsichtlich der drei Betrachtungsebenen und schafft somit die Grundlage für die Entwicklung einer individuellen Auswertungslogik. Die Differenzierung von bewussten und hoch individuellen unbewussten Informationsverarbeitungsprozessen, die nicht nur auf manifestierte Erfahrungen und Konzepte zurückzuführen ist, beeinflusst in hohem Maße die Bearbeitung von Aufgaben. Die Forschungsarbeiten von Dehaene und Changeux (2005) sowie Kiefer et al. (2012) zeigen, dass auch Ausprägungen unbewusster Prozesse kognitive Ressourcen verbrauchen, was zu einer steigenden Bedeutung von differenziellen und individuellen Modellen für die zuverlässige Beanspruchungsmessung führt. Zusätzlich verändert der Mensch während der Bearbeitung einer Aufgabe sowohl bewusst als auch auf unbewusster Ebene seine Strategien. Diese Änderungen des Verhaltens erschweren die Betrachtung der mentalen Beanspruchung auf Stichprobenebene (Matthews & Campbell, 1998). Diese Adaptionprozesse finden nicht nur aufgrund von Veränderungen der Aufgaben oder des Anforderungsniveaus, sondern auch durch Veränderungen der individuellen Zielstellungen (Sträter, 2005) oder zur Optimierung der Anstrengung sowie der Leistung statt (Sperling & Doshier, 1986). Zusammenfassend zeigt der theoretische Hintergrund in Kapitel 2 die

starken inter- und intraindividuellen Unterschiede in der menschlichen Informationsverarbeitung und somit auch in der mentalen Beanspruchung, was die Beschreibung der kognitiven Prozesse auf differenzieller und individueller Ebene motiviert (Driver & Mock, 1975). Insbesondere die Studie von Fels et al. (2015) zeigt, dass die Zuverlässigkeit der physiologischen Beanspruchungsmessung bei der Betrachtung auf der differenzierten Ebene im Vergleich zu der generellen Ebene erheblich gesteigert werden kann (Kapitel 2.2.2). Gagnon et al. (2011) postulieren, basierend auf ihren empirischen Befunden, dass die Möglichkeit einer zuverlässigen Beanspruchungsmessung über physiologische Parameter ausschließlich auf der individuellen Ebene funktionieren kann. Zusammenfassend ist der Bestandteil dieser Studie zunächst die Untersuchung der Diagnostizität und Sensitivität physiologischer Parameter auf Stichprobenebene. Nachfolgend werden die Ergebnisse mit der differenziellen und individuellen Ebene verglichen. Dementsprechend wird eine mögliche Steigerung der Zuverlässigkeit von physiologischen Indikatoren aufgrund einer individuellen Auswertung geprüft.

4.1 Experimentaldesign der Studie 1

Die Probanden bearbeiten in dieser Studie zwei Aufgaben (Flugführung, mentale Rotation) mit jeweils drei fest definierten Anforderungsniveaus. Sowohl die Reihenfolge der Aufgaben als auch der Anforderungsniveaus wird randomisiert, um Reihenfolgeeffekte zu vermeiden. Die mentale Beanspruchung wird entsprechend dem trimodalen Ansatz (Kapitel 3.1) über subjektive, leistungsbezogene und physiologische Parameter erfasst. Das subjektive Beanspruchungsempfinden wird über den NASA-TLX erfasst (Kapitel 3.1.1), welcher den Probanden nach jedem Anforderungsniveau (insgesamt 6-mal) gereicht wird. Für die Erfassung der physiologischen, okularen Beanspruchungsparameter (Kapitel 3.1.3) tragen die Probanden während des gesamten Versuches das kopfbasierte Blickerfassungssystem Dikablis Professional (binokular, 60Hz) der Firma Ergoneers GmbH. Die leistungsbezogenen Messungen unterscheiden sich zwischen den Aufgaben und werden folgend im Rahmen der detaillierten Aufgabenbeschrei-

bungen näher erläutert. Die Studie dauert je Proband, inklusive der Instruktionen und der Bearbeitung der Aufgaben, zirka 60 min. Bevor die Probanden mit der Bearbeitung der Aufgaben beginnen, erhalten sie eine umfassende Einführung in die Messinstrumente und die Erklärung des jeweils direkt folgenden Versuchsteils durch den Versuchsleiter. Der Versuchsleiter führt die Instruktionen nach einem zuvor ausgearbeiteten Skript durch, um Versuchsleitereffekte als Störvariable weitestgehend zu minimieren. Das Blickerfassungssystem wird vor jedem der sechs Versuchsteile neu kalibriert, um Messartefakte, die durch das Verrutschen des Systems entstehen könnten, zu verhindern.

Aufgabe A1: Flugführung

Aufgabe A1 besteht aus einer Flugführungsaufgabe in einem Flugsimulator (Microsoft Flight Simulator X). Flugführungsaufgaben werden in empirischen Forschungsarbeiten zur Beanspruchungsmessung sehr häufig verwendet, da sie gut skalierbare mentale Anforderungsniveaus umfassen, die ein breites Spektrum an praktischen Anwendungsszenarien repräsentieren (Wickens et al., 2013; Cardona & Quevedo, 2014; Di Nocera et al., 2007). Der Autopilot wird hier als defekt simuliert, sodass die Probanden das Flugzeug (Boeing 737) manuell für eine Dauer von je vier Minuten auf der aktuellen Flughöhe (5000ft) und den aktuellen Steuerkurs (360°) halten müssen. Der Flug findet unter standardisiert eingestellten, mittleren Turbulenzen statt, sodass die Probanden die genannten Flugparameter kontinuierlich überwachen und über den Steuerknüppel korrigieren müssen. Das Primary Flight Display (PFD) zur Überwachung von Flughöhe und Steuerkurs wird auf dem Bildschirm vergrößert dargestellt (Abbildung 4.1), um die Informationssuche der überwiegend studentischen Teilnehmer mit keiner bis wenig Flugerfahrung nicht zusätzlich zu erschweren. Die drei Anforderungsniveaus (Versuchsteile) werden in dieser Aufgabe über eine Nebenaufgabe operationalisiert (Abbildung 4.1). Während die Probanden in dem niedrigen Anforderungsniveau ohne Nebenaufgabe fliegen, sollen sie in den beiden anderen Versuchsteilen eine n-back Nebenaufgabe (Kapitel 3.1.2) nach dem Subsidiary Task Paradigm bearbeiten, die einmal visuell und einmal auditiv

präsentiert wird. Dabei wird den Probanden in Abständen von drei Sekunden eine Zahl zwischen Null und Zehn präsentiert. Die Probanden werden so instruiert, dass zum Zeitpunkt der Präsentation einer neuen Zahl die vorhergehende Zahl verbal wiedergeben werden soll. Als Leistungsmaß werden die Fehler in der Nebenaufgabe ausgewertet. Die Unterscheidung in eine visuelle und eine auditive Nebenaufgabenbedingung basiert auf dem Ressourcenmodell nach Wickens (1984). Auch wenn die Nebenaufgaben hinsichtlich des Inhaltes und der Anforderungen identisch sind, sollten sie sich aufgrund der unterschiedlich starken Interaktionseffekte mit der Hauptaufgabe anders auf die Beanspruchung des Probanden auswirken. Die visuelle Präsentation soll demnach eine höhere Beanspruchung induzieren (Wickens et al., 2013).

Aufgabe A2: Mentale Rotation

Aufgabe A2 besteht aus einer mentalen Rotation (Kapitel 3.1.2), welche auch in den Studien von Eggeheimer und Wilson (1991) sowie Xu und Franconeri (2015) zur Induktion mentaler Beanspruchung verwendet wurde. Innerhalb dieser Studie wird die mentale Rotation als Hauptaufgabe, ohne zusätzliche Nebenaufgabe, verwendet. Die Probanden überprüfen in dieser Aufgabe je Anforderungsniveau 128 Paare von Geometrien auf Gleichheit. Die zweidimensionalen Geometrien werden zufällig in der Ebene rotiert und sind entweder gleich oder gespiegelt. Diese Aufgabe wird ohne eine Nebenaufgabe ausgeführt, da die drei Anforderungsniveaus durch die Anzahl der Konturpunkte der Geometrien definiert werden (Abbildung 4.1). Die Probanden haben pro Vergleich drei Sekunden Zeit die Antwort über einen Tastendruck einzugeben. Falls keine Eingabe in dieser Zeit erfolgt, springt der Test weiter zum nächsten Vergleich. Nach jedem Vergleich erhalten die Probanden Rückmeldung, ob ihre Eingabe richtig oder falsch war, um die Motivation über die Aufgabendauer hinweg zu erhalten. Als Leistungsmaß werden hier die Fehler und die Reaktionszeiten der mentalen Rotation ausgewertet.

A1: Flugführung			A2: Mentale Rotation		
Anforderungsniveaus					
Keine Nebenaufgabe	Akustische n-back Aufgabe	Visuelle n-back Aufgabe	6 - 8 Konturpunkte	12 - 16 Konturpunkte	20 - 24 Konturpunkte
			<p>Leicht</p>  <p>Press 'hold' for same and 'release' for different</p>  <p>Press 'hold' for same and 'release' for different</p>	<p>Mittel</p>  <p>Press 'hold' for same and 'release' for different</p>  <p>Press 'hold' for same and 'release' for different</p>	<p>Schwer</p>  <p>Press 'hold' for same and 'release' for different</p>  <p>Press 'hold' for same and 'release' for different</p>

Abb. 4.1: Studie 1: Aufgaben und Anforderungsniveaus (eigene Abbildung)

4.2 Ergebnisse der Studie 1

Die Stichprobe umfasst $N = 48$ (32 männliche, 16 weibliche) Probanden, die im Durchschnitt 22,44 Jahre alt sind ($SD = 3,26$). Alle Probanden haben ein uneingeschränktes oder korrigiert-uneingeschränktes Sehvermögen (13 Brillenträger, 6 Kontaktlinsenträger). Die Ergebnisse der Studie werden in drei Abschnitten berichtet: Zuerst werden die subjektiven und leistungsbezogenen Daten für beide Aufgaben A1 und A2 vorgestellt. Im Weiteren erfolgt die Auswertung der physiologischen Daten auf genereller sowie auf differenzieller und individueller Ebene. Aufgrund des Stichprobenumfangs werden zur Auswertung dieser Studie, gemäß dem zentralen Grenzwerttheorems, parametrische Tests verwendet (Kapitel 3.2.2).

4.2.1 A1: Flugführung

Für Aufgabe A1: Flugführung zeigen sich signifikante Unterschiede hinsichtlich des subjektiven Beanspruchungsempfindens zwischen den drei Anforderungsniveaus ($F(2, 94) = 64,77, p < 0,01$). Die paarweisen, Bonferroni-korrigierten t -Tests zeigen jedoch, dass die Unterschiede lediglich zwischen der Bedingung ohne Nebenaufgabe und auditiven ($t(47) = -10,50, p < 0,01$) beziehungsweise der visuellen ($t(47) = -10,79, p < 0,01$) Nebenaufgabenbedingung liegen (Abbildung 4.1). Die beiden Anforderungsniveaus mit Nebenaufgabe unterscheiden

sich nicht signifikant ($t(47) = 0,66, n. s.$). Diese Ergebnisse werden durch die Fehlerquoten in den Nebenaufgaben unterstützt, welche ebenfalls keine signifikanten Unterschiede zwischen visueller und auditiver Bedingung aufweisen ($t(47) = 0,65, n. s.$).

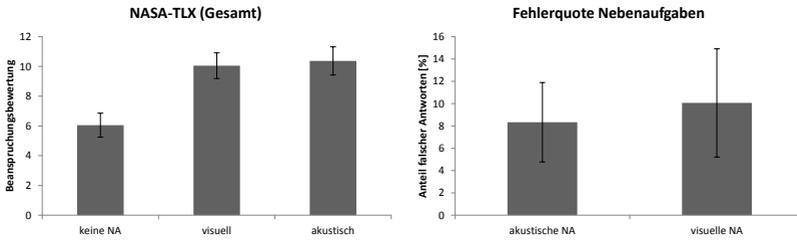


Abb. 4.2: Flugführung: subjektive und leistungs-basierte Daten. NA = Nebenaufgabe (eigene Abbildung)

Aufgrund der Ergebnisse der subjektiven und leistungs-basierten Daten können die Interaktionseffekte zwischen Haupt- und Nebenaufgabe nicht, wie im Ressourcenmodell nach Wickens (1984) postuliert, hinsichtlich der Modalität differenziert werden. Die Korrelationsanalyse zwischen beiden Erfassungsmethoden bekräftigt dieses Ergebnis ($r = 0,42, p < 0,01$). Für die Aufgabe der Flugführung kann basierend auf dem NASA-TLX und der Leistung in der Nebenaufgabe nur zwischen der Bedingung ohne Nebenaufgabe und beiden Bedingungen mit Nebenaufgabe unterschieden werden.

4.2.2 A2: Mentale Rotation

Auch für die mentale Rotation liegen signifikante Unterschiede zwischen den drei Anforderungsniveaus vor ($F(2, 94) = 25,38, p < 0,01$).

Die schwere Bedingung zeigt signifikant höher empfundene Beanspruchungen als die leichte ($t(47) = -4,57, p < 0,01$) und die mittlere Bedingung ($t(47) = -7,21, p < 0,01$). Die NASA-TLX Werte zwischen der leichten und der mittleren Bedingung unterscheiden sich nicht signifikant ($t(47) = 1,75, n. s.$). Das

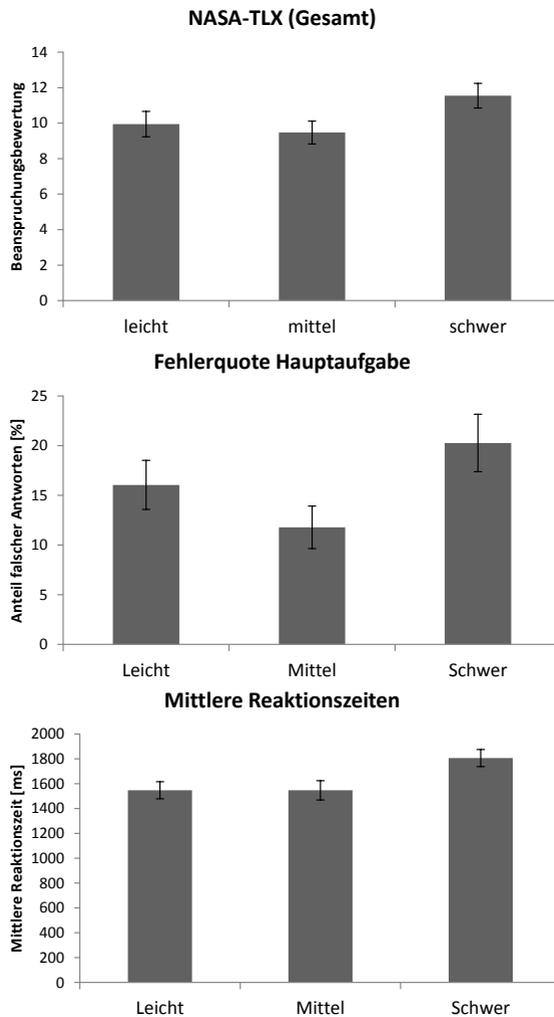


Abb. 4.3: Mentale Rotation: subjektive und leistungsorientierte Daten. NA = Nebenaufgabe (eigene Abbildung)

Diagramm der Mittelwerte (Abbildung 4.3) zeigt darüber hinaus, dass der Mittelwert des NASA-TLX im mittleren Anforderungsniveau ($M = 9,47$, $SD = 2,26$) sogar etwas unterhalb des Mittelwertes der leichten Bedingung ($M =$

9,95, $SD = 2,55$) liegt. Dies ist, laut Aussagen der Probanden, auf die z-förmige Geometrie der leichten Bedingung zurückzuführen, welche, trotz der wenigen Konturpunkte, nur schwer zu bewerten war. Auch die Fehlerquote ($F(2, 94) = 33,65, p < 0,01$) und die mittleren Reaktionszeiten ($F(2, 94) = 48,98, p < 0,01$) zeigen signifikante Unterschiede zwischen den Anforderungsniveaus (Abbildung 4.3). Die Reaktionszeiten ($r = 0,998, p < 0,01$) und die Fehlerquote ($r = 0,64, p < 0,01$) bestätigen das subjektive Empfinden der Probanden. Die mittleren Reaktionszeiten der schweren Bedingung sind signifikant höher als die der leichten ($t(47) = -10,53, p < 0,01$) und mittleren ($t(47) = -7,63, p < 0,01$) Bedingung. Weiter liegen die Reaktionszeiten in der mittleren Bedingung, wenn auch nicht signifikant ($t(47) = 0,04, n. s.$), unterhalb der leichten Bedingung. Bezüglich der Fehlerquote wird dieser Effekt signifikant ($t(47) = 5,19, p < 0,01$). Darüber hinaus unterscheidet sich auch hier die schwere Bedingung signifikant von dem leichten ($t(47) = -3,75, p < 0,01$) und mittleren ($t(47) = -7,53, p < 0,01$) Anforderungsniveau (Abbildung 4.3). Für die mentale Rotationsaufgabe kann anhand des subjektiven Empfindens und der mittleren Reaktionszeiten nur zwischen dem schweren und den beiden leichteren Anforderungsniveaus differenziert werden. Die Fehlerquote ist in der mittleren Bedingung signifikant geringer als in der leichten Bedingung, was den subjektiven Eindruck der Probanden bekräftigt. Die nachfolgende Tabelle 4.1 zeigt die Auswertung der Leistungsdaten auf der Ebene der einzelnen Geometrien. Die Auswertung zeigt signifikante Unterschiede zwischen den einzelnen Geometrien sowohl hinsichtlich der mittleren Reaktionszeiten ($F(5, 235) = 140,18, p < 0,01$) als auch der Fehlerquote ($F(5, 235) = 84,24, p < 0,01$).

Die Reaktionszeiten unterscheiden sich jeweils signifikant zwischen den leichten ($t(47) = 12,22, p < 0,01$), den mittleren ($t(47) = -3,79, p < 0,01$) und den schweren Geometrien ($t(47) = 22,52, p < 0,01$). Weiter hat das erste Objekt, welches von den Probanden als sehr schwer empfunden wurde, signifikant höhere Reaktionszeiten als alle anderen Objekte, mit Ausnahme des ersten Objekts des schweren Anforderungsniveaus. Die Fehlerquote bestätigt diese Ergebnisse: die leichten ($t(47) = 10,98, p < 0,01$), mittleren ($t(47) = -2,35, p < 0,05$) und schweren Geometrien ($t(47) = 12,48, p < 0,01$) unterscheiden sich jeweils

Tab. 4.1: Leistungsdaten auf Ebene der Geomentrien (eigene Tabelle)

Leicht		Mittel		Schwer	
					
<i>mittlere Reaktionszeiten [ms]</i>					
$M =$ 1687,35	$M =$ 1406,99	$M =$ 1501,85	$M =$ 1593,55	$M =$ 2184,16	$M =$ 1429,27
$SD =$ 253,97	$SD =$ 253,68	$SD =$ 289,73	$SD =$ 286,47	$SD =$ 236,25	$SD =$ 302,49
<i>Fehlerquote</i>					
$M = 0,24$ $SD = 0,11$	$M = 0,11$ $SD = 0,09$	$M = 0,12$ $SD = 0,09$	$M = 0,15$ $SD = 0,08$	$M = 0,33$ $SD = 0,15$	$M = 0,1$ $SD = 0,09$

signifikant voneinander. Auch hier weist das erste Objekt signifikant höhere Fehlerquoten als alle anderen Objekte, bis auf das erste Objekt des schwersten Anforderungsniveaus, auf. Basierend auf den Ergebnissen können ein leichtes (Geometrie 2 und 6), ein mittleres (Geometrie 3 und 4) und ein schweres (Geometrie 1 und 5) Anforderungsniveau definiert werden. Diese unterscheiden sich signifikant, widersprechen aber der Definition von Anforderungsniveaus über die Anzahl der Konturpunkte. Für die mentale Rotationsaufgabe kann festgehalten werden, dass die Anzahl der Konturpunkte allein, insbesondere im Bereich der leichten und mittleren Bedingung, nicht ausreichend für die Definition von Anforderungsniveaus ist. Die Formgebung der Geomentrien hat einen signifikanten Einfluss auf die mittleren Reaktionszeiten und die Fehlerquoten der Probanden, was bei der Interpretation der Ergebnisse für die mentale Rotation berücksichtigt werden muss.

4.2.3 Auswertung der physiologischen Daten

Die Ergebnisse der subjektiven und leistungsbasierten Beanspruchungsindikatoren liefern bereits einen guten Überblick über die Anforderungsniveaus beider Aufgaben und bilden somit die Referenz für die Auswertung der physio-

logischen (blickbasierten) Parameter. Für die Flugaufgabe zeigen sowohl die subjektiven als auch die leistungsbasierten Daten, dass lediglich eine Unterscheidung zwischen vorhandener und nicht vorhandener Nebenaufgabe getroffen werden kann. Auch für die mentale Rotation lässt sich über alle bisher betrachteten Beanspruchungsindikatoren hinweg nur zwischen der schweren und den beiden leichteren Bedingungen differenzieren. Die Fehlerquote indiziert allerdings darüber hinaus noch Unterschiede zwischen dem leichten und mittleren Anforderungsniveau. Die physiologischen Parameter werden demnach fokussiert auf deren Sensitivität hinsichtlich der genannten, signifikanten Unterschiede zwischen den Anforderungsniveaus untersucht. In dieser Studie wurde der PERCLOS als Beanspruchungsindikator noch nicht berücksichtigt, weshalb nur sechs okuläre Beanspruchungsparameter betrachtet werden. Die folgende Auswertung teilt sich in drei Abschnitte: zunächst soll die generelle Stichprobenebene betrachtet werden, bevor dann eine differenzierte und schließlich eine individuelle Auswertung vorgenommen wird.

Generelle Ebene

Zunächst wird die Diagnostizität der einzelnen Parameter bezüglich der subjektiven und leistungsbasierten Parameter bewertet. Die Tabelle 4.2 fasst die Ergebnisse der Korrelationsanalysen beider Aufgaben zusammen.

Für die Flugaufgabe zeigen lediglich die Blinzelrate und die durchschnittliche Sakkadenlänge anforderungsunabhängige, signifikante Zusammenhänge zu subjektiven und leistungsbasierten Parametern. Die Pupillendurchschnittsvariabilität korreliert ausschließlich mit dem NASA-TLX. Alle weiteren Parameter zeigen, trotz der Vielzahl an empirischen Evidenzen (Kapitel 3.1.3), keine Effekte auf Stichprobenebene. Die Diagnostizität hinsichtlich der mentalen Beanspruchung von Blinzelrate und Sakkadenlängen kann aber, trotz der signifikanten Korrelationen, in dieser Studie lediglich als gering bewertet werden. Sowohl die Blinzelrate (NASA-TLX: $r^2 = 0,06$, Leistung: $r^2 = 0,03$) als auch die Sakkadenlängen (NASA-TLX: $r^2 = 0,02$, Leistung: $r^2 = 0,02$) haben nur eine sehr geringe Varianzaufklärung bezüglich des subjektiven Empfindens und

Tab. 4.2: Korrelationsanalyse physiologischer Parameter (eigene Tabelle)

	Flugaufgabe		mentale Rotation		
	<i>NASA-TLX</i>	<i>Fehlerquote Nebenaufgabe</i>	<i>NASA-TLX</i>	<i>Fehlerquote Hauptaufgabe</i>	<i>mittlere Reaktionszeiten</i>
Blinzelrate	$r = 0,25,$ $p < 0,01$	$r = 0,17,$ $p < 0,05$	$r = 0,1,$ <i>n. s.</i>	$r = 0,15,$ $p < 0,05$	$r = 0,07,$ <i>n. s.</i>
Blinzel-dauer	$r = 0,02,$ <i>n. s.</i>	$r = 0,02,$ <i>n. s.</i>	$r = 0,02,$ <i>n. s.</i>	$r = -0,11,$ <i>n. s.</i>	$r = -0,19,$ $p < 0,05$
Pupillen-durchschnitts-variabilität	$r = 0,19,$ $p < 0,05$	$r = 0,03,$ <i>n. s.</i>	$r = 0,04,$ <i>n. s.</i>	$r = -0,19,$ $p < 0,05$	$r = -0,03,$ <i>n. s.</i>
NNI	$r = -0,00,$ <i>n. s.</i>	$r = 0,02,$ <i>n. s.</i>	$r = -0,09,$ <i>n. s.</i>	$r = -0,07,$ <i>n. s.</i>	$r = -0,08,$ <i>n. s.</i>
Fixations-dauer	$r = -0,12,$ <i>n. s.</i>	$r = -0,08,$ <i>n. s.</i>	$r = 0,01,$ <i>n. s.</i>	$r = -0,20,$ $p < 0,01$	$r = -0,07,$ <i>n. s.</i>
Sakkaden-länge	$r = 0,15,$ $p < 0,05$	$r = 0,15,$ $p < 0,05$	$r = -0,08,$ <i>n. s.</i>	$r = 0,12,$ <i>n. s.</i>	$r = 0,23,$ $p < 0,05$

der Leistungsparameter, sodass keine zuverlässigen Aussagen hinsichtlich der mentalen Beanspruchung getroffen werden können. Für die mentale Rotationsaufgabe zeigen sich ähnliche Ergebnisse. Es liegen nur vereinzelt signifikante Zusammenhänge vor, die allerdings nur einen sehr geringen Teil der Varianz in subjektiven und leistungsbasierten Messungen erklären. Bei der Betrachtung der Fixationsdauer und der Sakkadenlänge fällt auf, dass sich die Vorzeichen der Korrelationen zwischen NASA-TLX und den Leistungsparametern umkehren, obwohl diese untereinander sehr stark positiv korrelieren (Kapitel 4.2.2). Die Analyse der Sensitivität der physiologischen Parameter liefert deutlichere Ergebnisse. Für die Flugaufgabe zeigen sich signifikante Unterschiede zwischen den Anforderungsniveaus in allen Parametern, ausgenommen dem NNI (Abbildung 4.4).

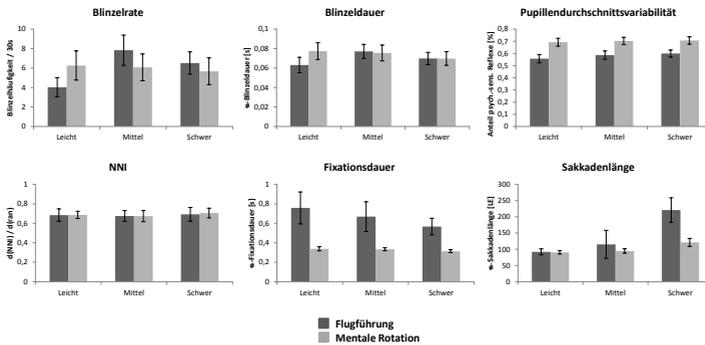


Abb. 4.4: Mentale Rotation: subjektive und leistungs-basierte Daten. NA = Nebenaufgabe (eigene Abbildung)

Die unbewussten Parameter bestätigen dabei die Ergebnisse der subjektiven und leistungs-basierten Parameter. Die Blinzelrate beispielsweise zeigt signifikante Unterschiede zwischen der Bedingung ohne Nebenaufgabe und sowohl der auditiven ($t(47) = -6,89, p < 0,01$) als auch der visuellen ($t(47) = -5,06, p < 0,01$) Nebenaufgabenbedingung. Die Modalität der Nebenaufgabe kann dabei nicht unterschieden werden ($t(47) = 2,32, n. s.$). Die Pupillendurchschnittsvariabilität zeigt ein sehr ähnliches Muster. Auch dieser Parameter kann zwischen keiner Nebenaufgabe und vorhandener auditiver ($t(47) = -3,88, p < 0,01$) oder visueller Nebenaufgabe ($t(47) = -4,41, p < 0,01$) unterscheiden. Die bewussten Parameter, insbesondere die Sakkadenlängen, erlauben eine Differenzierung von der vorhandenen visuellen Nebenaufgabe zu keiner ($t(47) = -4,46, p < 0,01$) oder einer auditiven Nebenaufgabe ($t(47) = -6,69, p < 0,01$). Dies kann in erster Linie auf das veränderte Blickverhalten durch eine neue visuelle Informationsquelle zurückgeführt werden. Die Sakkaden zur Nebenaufgabe und wieder zurück zum PFD sind deutlich länger, als die Sakkaden zwischen den einzelnen Anzeigeelementen der Hauptaufgabe, was sich dann in den signifikant höheren Sakkadenlängen niederschlägt. Für die mentale Rotationsaufgabe ist eine Differenzierung der Anforderungsniveaus lediglich über die Blinzeldauer ($F(2, 94) = 4,17, p < 0,05$), die Fixationsdauer ($F(2, 94) = 23,02, p < 0,01$) und die Sakkadenlänge ($F(2, 94) = 4,67, p < 0,05$) möglich. Dabei entsprechen die

paarweisen Vergleiche aller drei Parameter den Ergebnissen der subjektiven und leistungsbasierten Beanspruchungsindikatoren: Die schwere Bedingung unterscheidet sich von leicht und mittel, wobei die beiden letztgenannten sich untereinander nicht unterscheiden. Die eingeschränkte Sensitivität in der mentalen Rotationsaufgabe, kann dabei nicht zuletzt auch auf die Unterschiede zwischen den einzelnen Geometrien eines Anforderungsniveaus zurückgeführt werden (Kapitel 4.2.2).

Die sehr gute Sensitivität vereinzelter physiologischer Parameter, insbesondere in der Flugführungsaufgabe, bestätigt das hohe Potential der ausgewählten okularen Parameter zur Erfassung der mentalen Beanspruchung. Die Ergebnisse vorhandener empirischer Forschungsarbeiten können somit bestätigt werden (Wickens et al., 2013; Cardona & Quevedo, 2014; Di Nocera et al., 2007). Die Ergebnisse bezüglich der Diagnostizität bestätigen jedoch die theoretischen Überlegungen bezüglich der starken intra- und interindividuellen Unterschiede der mentalen Beanspruchung (Kapitel 2.2.2), sowohl für die Flugaufgabe, als auch für die mentale Rotation. Die Vielzahl an nicht signifikanten Korrelationskoeffizienten und auch die geringe Varianzaufklärung der signifikanten Zusammenhänge indizieren die hohe Individualität der mentalen Beanspruchung. Einzelne Parameter verhalten sich nicht über alle Probanden hinweg gleich, was sich zum Teil auch schon in den kontroversen Ergebnissen hinsichtlich der Lidschlussparameter vorhandener Studien zeigt (Faure et al., 2016; Wilson, 2009). Keiner der analysierten Parameter kann jeden Probanden der Stichprobe in gleichem Maße hinsichtlich der mentalen Beanspruchung beschreiben und bezüglich der Anforderungsniveaus differenzieren. Darüber hinaus verhalten sich Parameter teilweise für unterschiedliche Probanden genau gegenläufig: proportional oder antiproportional zu den steigenden Anforderungen. Die Auswertung auf Stichprobenebene wird durch diese Unregelmäßigkeiten deutlich erschwert und eine differenzierte Auswertung motiviert.

Differenzielle Ebene

Die Auswertungen auf der generellen Ebene haben das Potential der okularen Beanspruchungsindikatoren bezüglich ihrer Sensitivität gezeigt. Insbesondere aber bezüglich der Diagnostizität zeigt sich der Bedarf einer differenzierten oder individuellen Auswertung zum Ziel der zuverlässigen Beanspruchungsbewertung. Für eine differenzierte Auswertung wird die Stichprobe in Gruppen (Cluster) eingeteilt, in denen die Probanden ein ähnliches Verhalten während der Bearbeitung der Aufgaben zeigen. Diese Einteilung wird unter Verwendung der MDS und der k-Means-Analyse (Kapitel 3.2.2) basierend auf den einzelnen bewussten und unbewussten Parametern vorgenommen. Als Eingabe für beide Analysen dienen die euklidischen Distanzen (d) zwischen den einzelnen Probanden, die auf dem invertierten Betrag der Korrelationskoeffizienten (r) zwischen den Probanden basieren (Formel 4.1).

$$d = 1 - |r| \quad (4.1)$$

Für die Berechnung der Korrelationskoeffizienten werden alle physiologischen Parameter eines Probanden über alle Aufgabenteile hinweg berücksichtigt, um aufgabenunabhängig die Varianz in bewussten und unbewussten Parametern abbilden zu können. Probanden, die sich hinsichtlich jedes Parameters über alle Aufgabenteile hinweg gleich verhalten, werden in der Nähe zueinander angeordnet, wohingegen Probanden, die sich sehr heterogen verhalten, weit auseinander angeordnet werden. Abbildung 4.5 zeigt das Ergebnis der MDS, wobei acht Probanden aufgrund von teilweise artifiziellen Parameterdaten ausgeschlossen werden mussten.

Die eingezeichneten Cluster A - H sind das Ergebnis der k-Means-Analyse. Die Abstände der einzelnen Probanden (Punkte) zu den jeweiligen, zugeordneten, Clusterzentren sind für die gesamte Stichprobe signifikant geringer als zu den anderen Clusterzentren, wodurch sich auch die einzelnen Cluster signifikant voneinander unterscheiden. Die entsprechenden F -Werte der ANOVAs sind im Anhang (S. 213) zu finden. Die Analyse der Diagnostizität hinsichtlich des sub-

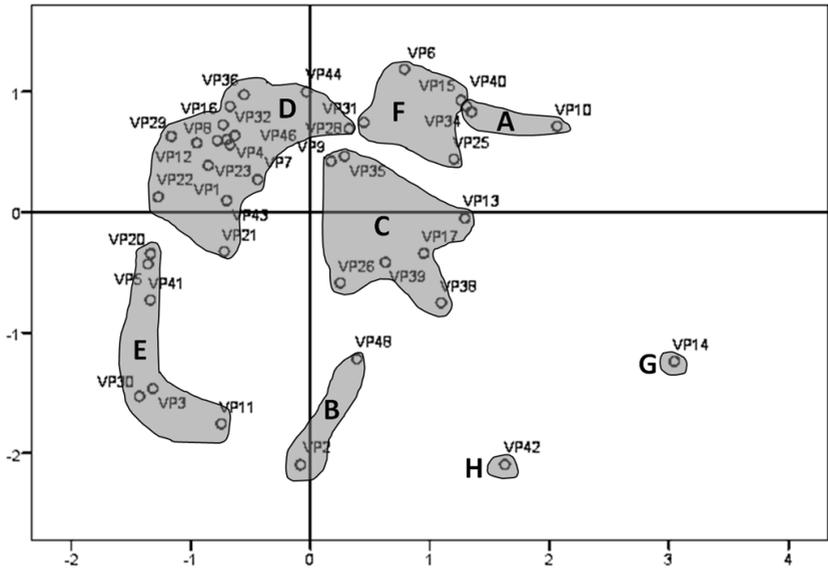


Abb. 4.5: MDS mit k-Means Clustern (eigene Abbildung)

jektiven Empfindens und der leistungsbasierten Daten zeigt, dass jedes Cluster durch einen spezifischen physiologischen Parameter zuverlässig beschrieben werden kann (Tabelle 4.3).

Um die Ergebnisse auf Gruppenebene mit denen auf genereller Ebene vergleichen zu können, werden die Korrelationskoeffizienten des höchst korrelierenden Parameters je Cluster hinsichtlich des NASA-TLX Gesamtwertes und der Leistungswerte transformiert (Fischers' z-Transformation¹), anschließend mit denen der anderen Cluster gemittelt, bevor der Mittelwert wieder rücktransformiert wird. So ergeben sich über alle Gruppen hinweg signifikante mittlere Korrelationen bezüglich des subjektiven Empfindens ($r_Z = 0,55$, $\sigma_Z = 0,16$, $p < 0,01$) und der Fehlerquote ($r_Z = 0,63$, $\sigma_Z = 0,16$, $p < 0,01$). Verglichen mit den Werten der Korrelationsanalyse auf genereller Ebene (Tabelle 4.2) kann zunächst eine

¹ Da Korrelationskoeffizienten nicht intervallskaliert sind, müssen diese zur Berechnung von durchschnittlichen Korrelationen transformiert werden. Fisher führt dazu eine asymptotische Normalisierung durch: $Z = \frac{1}{2} \times \ln\left(\frac{1+r}{1-r}\right)$ (Field, 2013)

Tab. 4.3: Korrelationsanalyse auf Gruppenebene (eigene Tabelle)

Gruppe	Korrelationen NASA-TLX		Korrelationen Fehlerquote		N (Gruppe)
	Parameter	Korrelation	Parameter	Korrelation	
A	PD	$r = 0,52,$ $p < 0,01$	FD	$r = -0,46,$ $p < 0,05$	3
B	PD	$r = 0,63,$ $p < 0,01$	FD	$r = 0,38,$ <i>n. s.</i>	2
C	FD	$r = -0,39,$ $p < 0,01$	FD	$r = -0,56,$ $p < 0,01$	7
D	PD	$r = 0,26,$ $p < 0,01$	FD	$r = -0,32,$ $p < 0,01$	14
E	FD	$r = -0,49,$ $p < 0,01$	FD	$r = -0,38,$ $p < 0,01$	6
F	FD	$r = -0,50,$ $p < 0,01$	BD	$r = 0,28,$ $p < 0,05$	6
G	PD	$r = 0,89,$ $p < 0,01$	BD	$r = 0,92,$ $p < 0,01$	1
H	FD	$r = -0,88,$ $p < 0,01$	NNI	$r = -0,71,$ <i>n. s.</i>	1

deutliche Steigerung der Diagnostizität festgestellt werden. Weiter fällt auf, dass die Pupillendurchschnittsvariabilität (PD) und die Fixationsdauer (FD) sehr häufig die Parameter mit der höchsten Diagnostizität bezüglich des NASA-TLX als auch bezüglich der Fehlerquote für das jeweilige Cluster sind (Tabelle 4.3). Auf der generellen Ebene zeigen beide Parameter mitunter die niedrigsten Korrelationswerte (Tabelle 4.3). Dieses Ergebnis bestätigt die starke Differenz in den physiologischen Reaktionen zwischen verschiedenen Personen. Einzelne Parameter verfügen über eine individuell verschiedene Diagnostizität und Sensitivität hinsichtlich der mentalen Beanspruchung. So können sich individuell hoch diagnostische und sensitive physiologische Reaktionen zweier Personen bei einem annähernd gleichen Beanspruchungszustand sehr stark unterscheiden. Beispielsweise zeigen sich längere Fixationsdauern mit steigender Fehlerquo-

te in Cluster B, wohingegen die Personen in Cluster A eine Verkürzung der durchschnittlichen Fixationsdauer mit steigender Fehlerquote aufweisen (Tabelle 4.3). Abschließend sollen die Cluster G und H näher betrachtet werden. Beide umfassen jeweils nur eine Person und weisen mit Abstand die höchste Diagnostizität hinsichtlich des NASA-TLX und der Fehlerquote auf. Gruppe D dagegen umfasst 14 Personen und hat insgesamt die niedrigste Diagnostizität. Diese Beobachtung zeigt, dass selbst nach der Unterteilung der Stichprobe in Cluster eine hohe Variabilität hinsichtlich der physiologischen Reaktionen auf einen Beanspruchungszustand innerhalb der einzelnen Gruppen besteht. Die theoretischen Überlegungen und die Ableitung des Bedarfs einer individuellen Auswertung mit dem Ziel mentale Beanspruchung über physiologische Parameter zuverlässig erfassen zu können, wird dadurch erneut bestätigt und bekräftigt. Nachfolgend werden daher die physiologischen Reaktionen auf individueller Ebene ausgewertet.

Individuelle Ebene

Basierend auf den Erkenntnissen der generellen und differenziellen Auswertung, werden folgend die einzelnen Probanden auf individueller Ebene betrachtet. Dazu werden für jeden Probanden die einzelnen Parameter über alle Aufgabenteile hinweg mit dem subjektiven Empfinden und der Leistung korreliert und der Parameter mit der höchsten Diagnostizität ausgewählt. Transformiert (Fischers' z-Transformation), über alle Probanden gemittelt und schließlich rücktransformiert, ergeben sich über alle Probanden hinweg hoch signifikante mittlere Korrelationen im Hinblick auf das subjektive Beanspruchungsempfinden ($r_Z = 0,83$, $\sigma_Z = 0,13$, $p < 0,01$) und die Fehlerquote ($r_Z = 0,87$, $\sigma_Z = 0,11$, $p < 0,01$). Verglichen mit der differenziellen Ebene kann die Diagnostizität der physiologischen Parameter, den Erwartungen entsprechend, durch eine individuelle Auswertung erneut deutlich gesteigert werden. Die Ergebnisse zeigen, dass es für jeden Probanden einen spezifischen physiologischen Parameter gibt, der zuverlässig und hoch signifikant die mentale Beanspruchung indiziert. Schon auf der differenziellen, aber insbesondere auf der individuellen

Ebene, fällt auf, dass für einen Probanden unterschiedliche physiologische Parameter die jeweils höchste Diagnostizität hinsichtlich NASA-TLX und der Leistung aufweisen. Lediglich bei $N = 9$ (22,5%) Probanden wird der gleiche Parameter für beide Vorhersagen ausgewählt. Dieses Ergebnis bestätigt die im Theorieteil hergeleitete Differenz zwischen subjektiven und leistungsbasierten Indikatoren, basierend auf den bewussten und unbewussten Adaptionsstrategien in Folge von Veränderungen der individuellen Zielstellungen, zur Optimierung der Anstrengung oder der Leistung (Kapitel 2.2).

4.3 Diskussion der Studie 1

Im Rahmen der ersten Studie haben die Probanden zwei Aufgaben mit jeweils drei Anforderungsniveaus unter kontrollierten Laborbedingungen bearbeitet, wobei die mentale Beanspruchung über den trimodalen Ansatz erfasst wurde (Kapitel 3.1). Bei der Interpretation und Diskussion der Ergebnisse müssen jedoch einige Aspekte bezüglich der Operationalisierung der Anforderungsniveaus berücksichtigt werden. Bei der Flugaufgabe soll das Problem der Folgefehler in der n-back Nebenaufgabe (Kapitel 3.1.2) und dessen Auswirkungen auf die Datenauswertung noch einmal aufgegriffen werden. Durch den Umstand, dass ein Proband eine Zahl nicht wiedergeben kann, obwohl er zu dem Zeitpunkt die notwendigen freien Kapazitäten hat, entsteht eine Abnahme der Diagnostizität des Leistungsparameters. Auch die Sensitivität zur Differenzierung der beiden Ausgabemodalitäten nimmt ab, da das Beanspruchungsniveau in diesen Zeitenintervallen abnimmt und sich so die Interferenzeffekte zwischen Haupt- und Nebenaufgabe weniger stark auswirken (Wickens et al., 2013). Dies würde wiederum zu einem gewissen Anteil die nicht signifikanten Unterschiede zwischen den Nebenaufgabenbedingungen erklären. Bei der mentalen Rotationsaufgabe wurden die Anforderungsniveaus anhand der Anzahl an Konturpunkten der Geometrien definiert (Bertrau-Pavy et al., 2011). Anhand der detaillierten Auswertung der Leistungsdaten, auf der Ebene der einzelnen Geometrien, kann dieses Kriterium zur Definition der Anforderungsniveaus als nicht ausreichend klassifiziert werden. Die deutlichen Unterschiede in den einzelnen Anforderungsniveaus

rungsniveaus zeigen, dass weitere Einflussfaktoren, wie z. B. die Formgebung, berücksichtigt werden müssen, was einen guten Anknüpfungspunkt für weitere Forschungen zur mentalen Rotation oder auch zur visuellen Komplexität bildet. Die genannten Aspekte hinsichtlich der Operationalisierung von Anforderungsniveaus müssen in der folgenden Diskussion der physiologischen Parameter berücksichtigt werden, da sie sowohl die Bewertung der Diagnostizität als auch der Sensitivität beeinflussen können.

Die Auswertung der physiologischen Beanspruchungsparameter auf der generellen Ebene hat das Potential okularer Parameter durch eine gute Sensitivität gezeigt. Allerdings erschwert die fehlende Diagnostizität eine Interpretation der Parameter und besonders die Ableitung von systemgestalterischen Maßnahmen basierend auf physiologischen Messungen erheblich. Es ist möglich, anhand der physiologischen Reaktionen auf genereller Ebene zu differenzieren, aber nicht zu quantifizieren, was die Auswertung auf der differenziellen Ebene motiviert. Die Auswertung auf einer differenziellen Gruppenebene zeigt eine deutliche Steigerung der Diagnostizität im Vergleich zu einer generellen Herangehensweise. Allerdings liegt auch die Varianzaufklärung hinsichtlich des subjektiven Empfindens und der leistungsbasierten Parameter noch weit unter 50%, was bei einer alleinigen Erfassung der physiologischen Parameter noch keine zuverlässige Systemgestaltung erlauben würde. Die Auswirkungen der individuellen Unterschiede in den physiologischen Reaktionen sind auch auf Gruppenebene noch zu groß. Dies wird durch die Unterschiede in der Varianzaufklärung zwischen großen Clustern und den beiden Clustern mit nur jeweils einem Probanden wiederholt verdeutlicht. Weiter ist der Ansatz über die Gruppierung der Stichprobe nicht ausreichend stabil. Die statistischen Verfahren zur Gruppierung einer Stichprobe unterliegen bestimmten Schwankungen. Die MDS beispielsweise verändert sich bei wiederholtem Ausführen mit der gleichen Dateneingabe immer wieder leicht, was eine sichere und zuverlässige Interpretation deutlich erschwert. Darüber hinaus konnten keine konstanten Merkmale hinsichtlich der Zuordnung eines Probanden zu einem Cluster extrahiert werden, weshalb eine trimodale Erfassung der Beanspruchung, wie

auf genereller Ebene, notwendig bleibt. Abschließend wurde die individuelle Ebene betrachtet, welche großes Potential hinsichtlich der Diagnostizität zeigt. Es gibt für jeden Probanden eine physiologische Reaktion, die zuverlässige Rückschlüsse auf den mentalen Beanspruchungszustand zulässt. Die individuelle Auswertung, wie sie hier vorgenommen wurde, ist jedoch durch einen erheblichen Aufwand geprägt und erfordert weiterhin den Abgleich physiologischer Daten mit subjektiven und leistungs-basierten Parametern. Die Ergebnisse der ersten Studie motivieren die Entwicklung einer Kalibrierungsaufgabe zur Erstellung eines individuellen Beanspruchungsmodells, da die individuelle Auswertung indiziert, dass ein solches Modell für jede Person existieren kann. Eine mögliche Kalibrierungsaufgabe muss jedoch generalisierbar sein und keinen spezifischen Aufgabenbereich adressieren, sodass ein individuelles Modell aus der Kalibrierungsaufgabe auf eine tatsächliche Anwendung übertragen werden kann.

4.4 Zusammenfassung der Studie 1

Die erste Studie hat gezeigt, dass es für jeden der untersuchten Probanden einen okularen Beanspruchungsparameter gibt, der individuell die mentale Beanspruchung zuverlässig und aufgabenübergreifend beschreiben kann. Die theoretischen Zusammenhänge zwischen der hoch individuellen Informationsverarbeitung und der mentalen Beanspruchung können damit bestätigt werden. Die Ergebnisse verdeutlichen zum einen das Potential okulärer Beanspruchungsindikatoren und zum anderen den Bedarf einer Kalibrierungsaufgabe zur individuellen Modellierung der physiologischen Reaktionen. Physiologische Reaktionen werden nicht nur durch die mentale Beanspruchung ausgelöst, sondern bilden eine Vielzahl an Konstrukten (z. B. Müdigkeit, Schläfrigkeit) ab. Die Auswirkung verschiedener mentaler Prozesse auf einzelne physiologische Indikatoren ist individuell stark verschieden, sodass für jede individuelle Person ein anderer Parameter, in anderem Umfang und mit individuellem Wirkungszusammenhang die mentale Beanspruchung am zuverlässigsten indiziert. Um die mentale Beanspruchung in praktischen Anwendungen (ohne parallele Erfassung

des subjektiven Empfindens und der Leistung) zuverlässig über okulare Parameter bestimmen zu können, bedarf es einer vorgelagerten Kalibrierungsaufgabe. Anhand dieser Kalibrierung wird eine individuelle Beanspruchungsfunktion, basierend auf den individuell zuverlässigsten physiologischen Indikatoren, gebildet. Darüber hinaus konnten aus der ersten Studie wichtige Erkenntnisse zur Operationalisierung von Anforderungsniveaus in mentalen Rotationsaufgaben gewonnen werden. Eine Definition allein über Konturpunkte ist nicht ausreichend und muss um die Formgebung der Geometrien erweitert werden. Die Studie bietet dazu einen guten Ausgangspunkt für weitere Forschungsarbeiten im Bereich der Aufgabengestaltung.

5 Entwicklung einer Kalibrierungsaufgabe

Ausgehend von den Ergebnissen der ersten Studie wird im Folgenden eine Kalibrierungsaufgabe entwickelt, mit der ein individuelles Beanspruchungsmodell, basierend auf physiologischen Parametern, abgeleitet werden kann. Die Kalibrierungsaufgabe soll von jedem Probanden vor der eigentlich zu analysierenden Tätigkeit ausgeführt werden. Definierte Anforderungsniveaus und die trimodale Beanspruchungsmessung während der Aufgabe bilden die Grundlage für die Entwicklung einer individuellen Beanspruchungsfunktion. Zusammenfassend soll die mobile Kalibrierungsaufgabe kontrollierte Laborbedingungen schaffen, um ein physiologisches Messsystem auf einen individuellen Probanden zu kalibrieren. Um dieses Ziel zu erreichen, werden im Rahmen der Konzeption Anforderungen definiert, die in einem konkreten Aufgabendesign münden, welches später informationstechnisch umgesetzt wird.

5.1 Konzeption

Die Konzeption der Kalibrierungsaufgabe beginnt mit der Erarbeitung von Anforderungen, die anschließend in ein konkretes Aufgabendesign überführt werden können. Bezogen auf die spätere Anwendung in Feldstudien ergeben sich zunächst drei grundlegende Anforderungen. Der Versuchsaufbau für die Kalibrierungsaufgabe muss mobil sein. Idealerweise wird die Aufgabe auf einem Laptop oder Tablet-PC realisiert, ohne dabei weitere Versuchsaufbauten zu benötigen. Weiter sollte die Aufgabe die zu analysierende Tätigkeit so gering

wie möglich beeinflussen. Eine Vorer müdung der Probanden durch eine lange Kalibrierungsdauer sollte in jedem Fall vermieden werden. Angestrebt wird daher eine Gesamtbearbeitungsdauer von zirka fünf Minuten. Diese Zeit wird in der Regel auch für diverse Vorbereitungen der eigentlichen Untersuchung benötigt, sodass im Idealfall Kalibrierung und Vorbereitungen simultan ablaufen könnten. Schließlich soll die Kalibrierungsaufgabe unabhängig von dem verwendeten physiologischen Messsystem sein, um so perspektivisch noch weitere physiologische Indikatoren in die Betrachtung einschließen zu können. Beispielsweise darf die Aufgabe keine großen Bewegungen benötigen, um später auch die Verwendung von bewegungssensitiven Messsystemen (z. B. EEG) zu ermöglichen. Bezüglich der Aufgabengestaltung ergeben sich aus den genannten Zielen weitere Anforderungen. Die Kalibrierungsaufgabe und deren Steuerung müssen intuitiv sein. Die Aufgabe muss soweit abstrahiert werden, dass diese ohne einen Übungsdurchgang und unabhängig von individuellen Leistungsvoraussetzungen und Vorkenntnissen durchführbar ist. Es werden verschiedene Anforderungsniveaus mit einer sehr guten Differenzierbarkeit benötigt, um die Sensitivität der Beanspruchungsmessung bewerten zu können. Die Unterschiede zwischen den Aufgabenteilen müssen mit bestimmten Aufgabencharakteristika klar attribuiert werden können, um so Unterschiede nicht nur zu erzeugen, sondern diese auch erklären zu können. Idealerweise umfasst die Kalibrierungsaufgabe verschiedene Aufgabenarten, um so die Aufgabenabhängigkeit von den physiologischen Parametern indizieren zu können. Des Weiteren muss die Aufgabe generalisierbar sein. Die Kalibrierungsaufgabe müsste die wichtigsten Aspekte der Mensch-Maschine-Interaktion in sozio-technischen Systemen umfassen (z. B. Auge-Hand-Koordination, Überwachungsanteile, Verfolgungstätigkeiten, Mehrfachbelastung), ohne dabei eine spezifische Anwendung allein zu adressieren. So kann gewährleistet werden, dass das individuelle Beanspruchungsmodell aus der Kalibrierungsaufgabe für die Analyse aller Arten von sozio-technischen Systemen verwendet werden kann. Abschließend muss die Aufgabe die Erfassung des subjektiven Beanspruchungsempfindens und der Leistung, als Referenzgrößen für das Beanspruchungsmodell, ermöglichen.

Ausgehend von den definierten Anforderungen folgt die Ableitung eines konkreten Aufgabendesigns. Dieses beginnt mit der Recherche nach einer Mensch-Maschine-Schnittstelle, die sowohl die grundlegenden Interaktionsaspekte soziotechnischer Systeme im Allgemeinen umfasst, als auch möglichst gut zu abstrahieren ist. Aus dem Bereich der Flugführung, welcher bereits aus der ersten Studie in dieser Arbeit bekannt ist (Kapitel 4), kann ein solches Design abgeleitet werden. Mehr als 500 wissenschaftliche Publikationen bezüglich Flugführungsaufgaben in Verbindung mit mentaler Beanspruchung, die allein im Jahr 2016 veröffentlicht wurden, unterstreichen die hohe Relevanz dieses Bereiches in der Beanspruchungsforschung. Darüber hinaus bietet die manuelle Flugführung alle genannten Interaktionsaspekte, durch die sozio-technische Systeme beschrieben werden (Kapitel 1). Die Anforderungen der Steuerungs- beziehungsweise Verfolgungsaufgabe (z. B. Flughöhe, Steuerkurs, Geschwindigkeit halten) sind durch die Möglichkeit der Teilautomatisierung (Autopilot) sehr gut skalierbar.

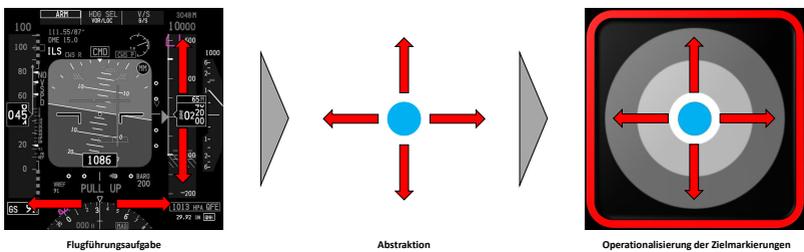


Abb. 5.1: Entwicklung der Kalibrierungsaufgabe (eigene Abbildung)

Für die Kalibrierungsaufgabe wurde eine Verfolgungsaufgabe der klassischen, manuellen Flugführung ausgewählt. Diese umfasst das Halten des Steuerkurses und der Flughöhe anhand der Vorgaben der Flugsicherung. Die vorgegebene Flughöhe und der zu haltende Steuerkurs werden, neben den aktuellen Flugparametern, auf dem Primary Flight Display (PFD) über magentafarbene Markierungen auf dem Kurskreisel (unten) und dem Höhenband (rechts) angezeigt (Abbildung 5.1 links). Das Ziel des Operators ist es, die aktuelle Flughöhe und den aktuellen Steuerkurs über die Steuerung mittels Steuerknü-

pel (Flughöhe: drücken / ziehen; Steuerkurs: links / rechts) in Deckung mit den magentafarbenen Markierungen zu bringen. Ständige Bewegungen der Luft und auftretende Turbulenzen erfordern kontinuierliche Korrekturen beziehungsweise Steuerungseingaben durch den Piloten. Dies stellt eine kontinuierliche Anforderung an den Piloten dar und benötigt seine dauerhafte Aufmerksamkeit. Das PFD ist allerdings, genau wie die typische Steuerung eines Flugzeuges, wenig intuitiv und generalisierbar, was zwei essentielle Anforderungen an die Kalibrierungsaufgabe sind. Die zweidimensionalen Bewegungen, die das PFD anzeigt, werden für die Kalibrierungsaufgabe auf eine Kugel übertragen (Abbildung 5.1 Mitte). Dieser Abstrahierungsschritt erhält alle Funktionen des PFD hinsichtlich Flughöhe und Steuerkurs, macht diese aber zugänglich für jede Person, unabhängig ihrer Vorkenntnisse. Die magentafarbenen Zielmarkierungen werden im nächsten Schritt über eine Art Zielscheibe operationalisiert (Abbildung 5.1 rechts). Die Kugel bewegt sich während der Aufgabe unterschiedlich stark aus der Mitte der Zielscheibe in zufällige Richtungen und muss durch den Probanden über eine Steuerungseinheit (Tablet) oder Tasteneingaben (Laptop) in der Mitte gehalten werden. Sollte die Kugel die Mitte verlassen, soll diese so schnell und präzise wie möglich dorthin zurückgesteuert werden. Die Leistung in dieser Aufgabe kann über die mittlere Kugelposition quantifiziert werden. Diese abstrahierte Flugführungsaufgabe ist generalisierbar, da sie die wichtigsten Aspekte von Mensch-Maschine-Schnittstellen in sozio-technischen Systemen umfasst (Auge-Hand-Koordination, Überwachungsanteile, Verfolgungstätigkeiten) und dabei keine Vorkenntnisse über das Führen eines Luftfahrzeuges erfordert. Zur Operationalisierung verschiedener Anforderungsniveaus werden zwei Kugeln auf zwei Zielscheiben nebeneinander angezeigt. In einer leichteren Bedingung soll nur eine Kugel in der Mitte gehalten werden, wobei in einer schwereren Bedingung beide Kugeln simultan gesteuert werden sollen. Somit wird auch der Aspekt der Mehrfachbelastung in der Interaktion mit sozio-technischen Systemen berücksichtigt und differenzierbare Anforderungsniveaus geschaffen. Zur Prüfung der Aufgabenunabhängigkeit von physiologischen Parametern soll die Kalibrierungsaufgabe zwei unterschiedliche Aufgabenarten beinhalten, was über die Verwendung von zwei unterschiedlichen Nebenaufgaben reali-

siert wird. Die beiden genannten Anforderungsniveaus werden im Rahmen der Kalibrierung zweimal ausgeführt, einmal mit einer auditiven und einmal mit einer visuellen Nebenaufgabe, sodass sich insgesamt vier Anforderungsniveaus verteilt auf zwei Nebenaufgabenbedingungen ergeben. Ausgehend von der Ressourcentheorie nach Wickens (1984) haben die beiden Nebenaufgaben, aufgrund der Modalitäten, unterschiedlich starke Interaktionseffekte mit der Hauptaufgabe, sodass sich die Anforderungen für den Probanden erheblich unterscheiden. Die auditive Nebenaufgabe besteht aus der mentalen Arithmetik (Kapitel 3.1.2). Diese beinhaltet optional (durch Versuchsleiter einstellbar) leichte (Addition) oder mittelschwere (Subtraktion) Aufgaben, die dem Probanden in einem Vier-Sekunden-Takt präsentiert werden. Der Proband soll dann verbal so schnell wie möglich mit dem richtigen Ergebnis antworten. Die visuelle Nebenaufgabe besteht aus einem visuellen Reiz (5 mm x 5 mm weißer Punkt), der in unterschiedlichen zeitlichen Abständen und in zufälligen Positionen des peripheren Sichtfeldes für 0,3 Sekunden erscheint (Kapitel 3.1.2). Sobald der Proband den Reiz wahrnimmt, soll er dies über einen Tastendruck so schnell als möglich bestätigen. Beide Nebenaufgabenbedingungen werden gemäß dem Subsidiary-Task-Paradigm instruiert (Kapitel 3.1.2). Dieses besagt, dass sich die Probanden komplett auf die Hauptaufgabe fokussieren und die Nebenaufgabe nur dann bearbeiten, wenn freie Kapazitäten zur Verfügung stehen. Die Reize beider Nebenaufgaben werden durch verschiedene Informationskanäle aufgenommen, weshalb sich die Aufgaben wesentlich voneinander unterscheiden. Durch die unterschiedliche Reizwahrnehmung lassen sich die Aufgabenarten anhand der auditiven und visuellen Nebenaufgabenbedingung klar differenzieren. Jedes der vier Anforderungsniveaus wird für 60 Sekunden bearbeitet. Die Zeit von 60 Sekunden ist nicht zu lang, reicht aber gerade aus, um auch bei physiologischen Parametern mit einem minimalen Betrachtungsintervall von 30 Sekunden (z. B. NNI) zwei Messpunkte zu erhalten. Nach jedem Anforderungsniveau bekommt der Proband einen Fragebogen zum subjektiven Beanspruchungsempfinden eingeblendet (frei wählbar, hier: NASA-TLX), sodass sich eine gesamte Kalibrierungsdauer von zirka fünf Minuten ergibt. Die Reihenfolge der Nebenaufgabenbedingungen wird randomisiert, innerhalb der Bedingungen werden

immer zunächst eine, dann zwei Kugeln gesteuert. Ob für die Bedingung mit einer Kugel die rechte oder linke Kugel verwendet wird, kann frei von dem Versuchsleiter gewählt oder randomisiert werden. In den folgenden Studien wird aufgrund der dominanten Augenführung (Porac & Coren, 1976) die linke Kugel in dieser Bedingung verwendet.

5.2 Informationstechnische Umsetzung

Die Kalibrierungsaufgabe wurde im Rahmen dieser Arbeit zunächst für die Nutzung auf Laptops und Desktop-PCs umgesetzt. Durch die Verwendung der vielseitigen Programmiersprache C# in der Programmierumgebung Visual Studio 2015 ist eine spätere Übertragung der Applikation auf ein mobiles Endgerät (z. B. Android, Windows Mobile, Apple iOS) ohne größeren Aufwand möglich. Der erste Schritt in der Programmierung umfasst die Bewegungen der Kugeln. Abbildung 5.2 zeigt das Layout der linken Kugel mit den Steuerungstasten und einem beispielhaften Vektorensystem für die Kugelbewegungen. Der Bewegungsvektor der Kugel \vec{b} wird automatisch erzeugt, wobei die x - und y -Werte zufällig gewählt werden, sodass $|\vec{b}| \leq 9 \frac{cm}{s}$ gilt. Die Länge von neun Zentimetern entspricht dabei dem Abstand vom Mittelpunkt der Zielscheibe bis zu dem äußeren Rand des dunkelgrauen Rings, der in Abbildung 5.2 mit 3 codiert ist. Beide grauen Ringe sind jeweils drei Zentimeter breit. Aus der dargestellten Codierung ergibt sich zugleich die Bewertungseinheit für die mittlere Kugelposition. Bei einem Wert von Eins für die mittlere Kugelposition befand sich die Kugel folglich während der gesamten Aufgabe perfekt in der Mitte.

Die zufälligen Bewegungen der Kugeln werden so programmiert und ausbalanciert, dass diese ohne Eingreifen des Probanden eine mittlere Abweichung von $d = 3$ erzeugen. Durch das Betätigen der Steuerungstasten wird ein Steuerungsvektor \vec{s} erzeugt. Für den Vektor \vec{s} gilt $|\vec{s}| \geq 1,5 \times |\vec{b}|$, wobei der Faktor 1,5 mit zunehmendem Betrag des Vektors \vec{b} linear skaliert. Dieser Skalierungsfaktor ist notwendig, um die Steuerbarkeit der Kugel zu jedem Zeitpunkt zu gewährleisten und ein schnelles Reagieren des Probanden in eine schnelle Kontrolle über die Kugel zu überführen. Der Mittelpunkt der Kugel M_K ist der Ausgangspunkt der

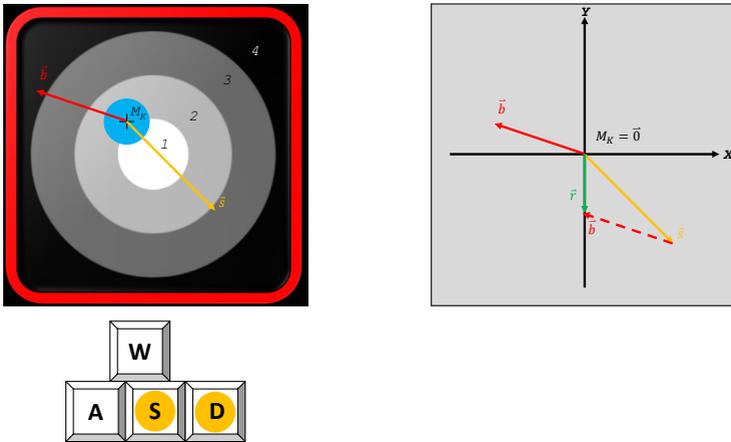


Abb. 5.2: Kugelbewegungen. Der rote Bewegungsvektor wird automatisiert und randomisiert erzeugt. Der gelbe Steuerungsvektor wird manuell über die entsprechenden Tasten erzeugt. Das rechte Koordinatensystem zeigt die Zusammensetzung der resultierenden Kugelbewegung (eigene Abbildung)

beiden Vektoren, für den in dem betrachteten Koordinatensystem (Abbildung 5.2 rechts) gilt: $M_K = \vec{0}$. Der Mittelpunkt der Kugel bildet demnach zu jeder Zeit den Koordinatenursprung, unabhängig von der aktuellen Position der Kugel, was die Berechnung und Ausbalancierung der automatischen Bewegungen erheblich erleichtert. Betätigt der Proband zwei Steuerungstasten, wie in Abbildung 5.2 dargestellt, liegt der Winkel des Steuerungsvektors genau im 45° Winkel zwischen den gewählten Bewegungsrichtungen, was einer herkömmlichen Computerspielsteuerung entspricht. Aufgrund der Bedingung $M_K = \vec{0}$ ergibt sich der resultierende Bewegungsvektor \vec{r} (Abbildung 5.2 rechts) aus der einfachen Vektoraddition gemäß Formel 5.1.

$$\vec{r} = \vec{b} + \vec{s} = \begin{pmatrix} b_x \\ b_y \end{pmatrix} + \begin{pmatrix} s_x \\ s_y \end{pmatrix} \quad (5.1)$$

Der resultierende Vektor \vec{r} ist im Wesentlichen von dem Skalierungsfaktor innerhalb des Steuerungsvektors abhängig, der im Zuge der Entwicklungen und Erprobungen der Aufgabe auf 1,5 festgelegt wurde. Dieser Faktor bewirkt

eine für das menschliche Auge flüssige Bewegung der Kugeln. Der Faktor und die Geschwindigkeit der automatischen Kugelbewegungen kann bei Bedarf im Nachhinein angepasst werden. Die Leistung in der Hauptaufgabe wird, wie oben beschrieben, durch die mittlere Kugelposition bestimmt. Die Leistung in der visuellen Nebenaufgabe (mittlere Reaktionszeiten und Fehler) wird ebenfalls automatisiert erfasst und ausgewertet. Um eine Vergleichbarkeit mit anderen Nebenaufgaben zu schaffen, werden beide Leistungsparameter normiert und gleichgewichtet miteinander kombiniert (Formel 5.2).

$$Leistung = 0,5 \times \frac{Fehler}{AnzahlReize} + 0,5 \times \frac{mittlere\ Reaktionszeit - 0,3}{1,1} \quad (5.2)$$

Die mittlere Reaktionszeit wird normiert, indem diese um die minimal mögliche Reaktionszeit des Menschen (0,3 Sekunden) reduziert (Sträter, 2005) und anschließend durch das Zeitfenster dividiert wird, welches dem Probanden für einen Tastendruck als Bestätigung zur Verfügung steht. Dieses Zeitfenster entspricht dem minimal möglichen Abstand von 1,1 (+ 0,3) Sekunden zwischen zwei aufeinander folgenden Reizen. Die minimale Reaktionszeit des Menschen wird subtrahiert, um zu vermeiden, dass verspätete Reaktionen auf den vorangegangenen Reiz als richtige Antwort auf den neuen Reiz anstatt eines Fehlers gewertet werden. Die Fehlerquote wird in Bezug zu der Gesamtanzahl an Reizen gebildet. Für die auditive, arithmetische, Nebenaufgabe wird nur die Fehlerquote ausgewertet. Eine Berücksichtigung der Reaktionszeiten war aufgrund des manuellen Fehlerprotokolls zum Zeitpunkt der folgenden Studien nicht möglich. Perspektivisch soll die Kalibrierungsaufgabe noch um eine automatisierte Spracherkennung erweitert werden, sodass auch die Leistung in der arithmetischen Nebenaufgabe, inklusive der Reaktionszeiten, automatisiert ausgewertet werden kann. Die Leistung der Nebenaufgabe wird, wie auch die physiologischen Parameter, in 30-Sekunden-Intervallen ausgewertet, sodass sich für die gesamte Kalibrierungsaufgabe acht Messzeitpunkte ergeben. Zwischen den einzelnen Aufgabenteilen wird der NASA-TLX Fragebogen eingeblendet, der von dem Probanden durch das Anklicken der entsprechenden Felder ausgefüllt werden soll. Der Proband kann dabei seine Auswahl durch erneutes Klicken korrigieren

und erst die Bearbeitung des nächsten Anforderungsniveaus fortsetzen, nachdem alle Items beantwortet wurden. Die Ergebnisse des Fragebogens werden im Anschluss der Kalibrierung ebenfalls automatisiert ausgewertet.

5.3 Entwicklung einer individuellen Auswertungslogik

Nach der Kalibrierungsaufgabe werden zunächst die sieben okularen Beanspruchungsparameter (Kapitel 3.1.3) in verschiedenen Varianten für jedes Anforderungsniveau automatisiert berechnet. Alle Parameter werden, wie auch die Leistungsdaten, aus Haupt- und Nebenaufgabe, in 30-Sekunden-Intervallen ausgegeben, sodass auch hier insgesamt acht Messpunkte betrachtet werden. Neben den verschiedenen Bezugsflächen für den NNI (Kapitel 3.1.3.1) und den unterschiedlichen Grenzwerten für die Extrahierung des psychosensorischen Pupillenreflexes (Kapitel 3.1.3.2) werden verschiedene Kenngrößen (Maximum, Minimum, Median, Mittelwert, erster Wert) der einzelnen Parametervarianten berechnet. Insgesamt ergeben sich so $k = 61$ physiologische, okulare Beanspruchungsparameter, die in die Erstellung des individuellen Beanspruchungsmodells eingehen können. Die Ergebnisse aus der individuellen Auswertung der ersten Studie (Kapitel 4.2.3) haben gezeigt, dass für jede untersuchte Person ein individueller Parameter existiert, der die mentale Beanspruchung zuverlässig vorhersagen kann. Dementsprechend wird versucht, ausgehend von der Kalibrierungsaufgabe, ein oder mehrere individuell zuverlässige Beanspruchungsindikatoren aus den physiologischen Parametervarianten zu extrahieren, mit denen anschließend eine Anwendungsaufgabe, ohne die zusätzliche Verwendung von Fragebögen oder Leistungsmaßen, analysiert werden kann. Folgend sollen dazu zwei mögliche Vorgehensweisen vorgestellt und diskutiert werden, bevor diese dann in den Studien zwei und drei erprobt werden.

5.3.1 Ansatz individueller Regressionsfunktionen

Die aktuellen Forschungsarbeiten von Jo, Myung und Yoon (2012) und Ryu und Myung (2005) zeigen eine deutliche Steigerung der Varianzaufklärung in der

physiologischen Vorhersage mentaler Beanspruchung durch die Kombination mehrerer Parameter auf Stichprobenebene. Dieser Ansatz einer Regressionsfunktion wird aufgegriffen und auf individueller Ebene angewendet. Zunächst wird die gewünschte Referenzgröße (Leistung Nebenaufgabe, subjektives Empfinden, mentale Effizienz) ausgewählt, anhand derer dann die individuelle Beanspruchungsfunktion entwickelt wird. Die Berechnung der Regressionsfunktion erfolgt mit möglichst wenigen unabhängigen Variablen (physiologischen Parametern), um die Multikollinearität gering zu halten. Durch die theoretisch hergeleitete und empirisch belegte Unterscheidung in bewusste und unbewusste Parameter liegen bereits zwei inhaltlich unabhängige Gruppen vor. Für die Bildung der Regressionsfunktion findet eine Auswahl von einem bewussten und einem unbewussten Parameter statt, die entsprechend der Formel 5.3 in einen funktionalen linearen Zusammenhang mit der standardisierten Referenzgröße gebracht werden.

$$f(p) = b_0 + b_1 \times P_{Bewusst} + b_2 \times P_{Unbewusst} \quad (5.3)$$

Die Faktoren b_i entsprechen den Regressionskoeffizienten einer linearen Regression. Durch die Standardisierung der Referenzgrößen gilt: $f(p) \in [0; 1]$, wodurch eine Vergleichbarkeit zwischen den individuellen Funktionen jeder Person ermöglicht wird. Im Zuge der Parameterauswahl wird die Korrelation zwischen den einzelnen Parametervarianten und der Referenzgröße gebildet. Anschließend werden der bewusste und der unbewusste Parameter mit der jeweils höchsten Varianzaufklärung ausgewählt. Diese individuelle Funktion bildet die Prognosefunktion für die Vorhersage der mentalen Beanspruchung in einer Anwendungsaufgabe. Allerdings weisen die theoretischen Vorüberlegungen aus Kapitel 2 darauf hin, dass diese Vorgehensweise der Parameterauswahl nicht ausreichend sein kann, um die Beanspruchungsfunktion aufgabenunabhängig einsetzen zu können. Die Regressionsfunktion, basierend auf den höchst korrelierenden Parametern, berücksichtigt zwar die interindividuellen, aber nicht die intraindividuellen Unterschiede der Informationsverarbeitung. Jeder Mensch verändert während der Bearbeitung einer Aufgabe sowohl bewusst als auch auf

unbewusster Ebene seine Strategien mit unterschiedlich starker Ausprägung (Matthews & Campbell, 1998). Diese Adaptionprozesse finden nicht nur aufgrund von Veränderungen der Aufgaben oder des Anforderungsniveaus, sondern auch durch Veränderungen der individuellen Zielstellungen (Sträter, 2005) oder zur Optimierung der Anstrengung sowie der Leistung statt (Sperling & Doshier, 1986). Aus diesem Grund kann es bei der Verwendung der Varianzaufklärung als einziges Auswahlkriterium der Parameter zu Problemen bei der Übertragung auf andere Aufgaben kommen. Nachfolgend soll daher ein Verfahren zur Parameterklassifikation entwickelt werden, welches diese intraindividuellen Unterschiede adressiert.

5.3.2 Ansatz der individuellen Parameterklassifikation

Der Ansatz zur individuellen Parameterklassifikation kombiniert mehrere methodische Vorgehensweisen, um die Aufgaben-, Anforderungs- und Zeitabhängigkeit sowie das Adaptionsverhalten der einzelnen physiologischen Indikatoren zu beschreiben und zu klassifizieren. Die Parameterauswahl findet sequentiell über verschiedene Schritte statt. Die Anzahl der Parameter wird bei jedem Schritt anhand verschiedener Kriterien reduziert (Abbildung 5.3).

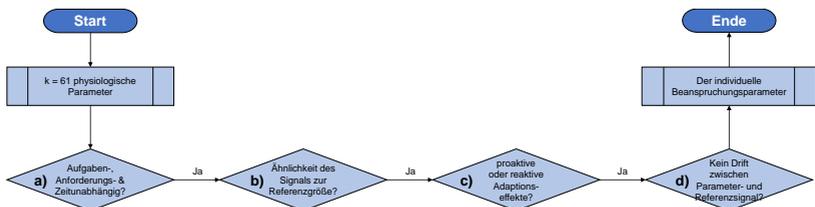


Abb. 5.3: Vereinfachte Darstellung der sequentiellen Parameterauswahl (eigene Abbildung)

Nach der Auswahl der Referenzgröße wird im ersten Schritt der Modellierung eine Procrustes-Analyse durchgeführt (Kapitel 3.2.1). Sowohl das Referenzsignal als auch die Signale der $k = 61$ physiologischen Parameter werden für die Analyse in $m \times 2$ Matrizen ($m =$ Anzahl der Messpunkte) aufbereitet, bevor die Signale translatiert, skaliert und rotiert werden. Die mittlere quadratische Abweichung d und der Rotationswinkel α werden für jeden Parameter gespeichert.

Sie beschreiben, wie weit (d) das Parametersignal unter welcher Veränderung (α) an das Referenzsignal angenähert werden kann. Das bedeutet, dass berechnet wird, wie ähnlich sich beide Signale sind. Die Procrustes-Analyse hat dabei entscheidende Vorteile gegenüber einer Korrelationsanalyse, da diese keine linearen Zusammenhänge voraussetzt. Darüber hinaus können auch Verschiebungen und Verdrehungen des Signals berücksichtigt werden. Diese Analyse wird für jeden Parameter k in verschiedenen Varianten durchgeführt, um die essentiellen Einflussfaktoren (Aufgabenart, Anforderungshöhe, Aufgabendauer) auf physiologische Parameter erstmals beschreiben und quantifizieren zu können (Tabelle 5.1). Neben dem Signal über die gesamte Kalibrierung (acht Messpunkte), werden auch jeweils die einzelnen Parametersignale nach Anforderungsniveau, Aufgabenart und Aufgabendauer (jeweils vier Messpunkte) zerlegt betrachtet. Über eine Procrustes-Analyse werden die einzelnen Parametersignale (Tabelle 5.1) mit den entsprechenden Referenzsignalen verglichen. Neben dem Signal über die gesamte Kalibrierung (acht Messpunkte), werden auch jeweils die einzelnen Parametersignale nach Anforderungsniveau, Aufgabenart und Aufgabendauer (jeweils vier Messpunkte) zerlegt betrachtet. Über eine Procrustes-Analyse werden die einzelnen Parametersignale (Tabelle 5.1) mit den entsprechenden Referenzsignalen verglichen.

Für jeden der $k = 61$ physiologischen Parameter werden demnach sieben Procrustes-Analysen durchgeführt. Durch den Vergleich der mittleren quadratischen Abweichung d der zerlegten Signale mit dem Signal der gesamten Kalibrierung, lassen sich die genannten Einflüsse auf den jeweiligen Parameter quantifizieren. Sind die quadratischen Abweichungen der zerlegten Signale (bspw. nach Anforderungsniveau) im Mittel kleiner als die des Gesamtsignals, liegt eine Abhängigkeit vor, da die beiden einzelnen Abschnitte jeweils besser vorhergesagt werden können. Durch eine Trennung des Signals würde die Vorhersagekraft des Parameters steigen. Dieser Vergleich der mittleren quadratischen Abweichungen wird für jedes Zerlegungskriterium angewendet und es werden weiter nur Parameter berücksichtigt, welche die folgenden Kriterien erfüllen:

Tab. 5.1: Aufteilung der Signale für die Procrustes-Analyse (eigene Tabelle)

Eingangssignal	Matrix	Verwendung
<i>Gesamte Kalibrierung</i> (8 Datenpunkte)	8 x 2	Quantifizierung der Ähnlichkeit des Gesamtsignals
<i>Leicht</i> (Hauptaufgabe 1; 4 Datenpunkte), <i>Schwer</i> (Hauptaufgabe 2; 4 Datenpunkte)	4 x 2 4 x 2	Aufteilung des Signals nach Anforderungsniveau (leicht / schwer)
<i>Arithmetik</i> (Nebenaufgabe arithmetisch; 4 Datenpunkte), <i>Visuell</i> (Nebenaufgabe visuell; 4 Datenpunkte)	4 x 2 4 x 2	Aufteilung des Signals nach Aufgabenart (arithmetisch / visuell)
<i>Aufgabenbeginn</i> (0 - 30 s; 4 Datenpunkte), <i>Aufgabende</i> (30 - 60 s; 4 Datenpunkte)	4 x 2 4 x 2	Aufteilung des Signals nach der Aufgabendauer (0 - 30 s / 30 - 60 s)

$$\begin{aligned}
 d_{Gesamt} &< \bar{x}(d_{Anforderungsniveau}) \ \&\& \\
 d_{Gesamt} &< \bar{x}(d_{Aufgabenart}) \ \&\& \\
 d_{Gesamt} &< \bar{x}(d_{Aufgabendauer})
 \end{aligned}
 \tag{5.4}$$

Nur Parameter, die von dem Anforderungsniveau, der Aufgabenart und der Aufgabendauer unabhängig sind, eignen sich für eine zuverlässige Vorhersage einer Anwendungsaufgabe (a in Abbildung 5.3). Für die verbleibenden Parameter werden jeweils drei Korrelationen über alle Messzeitpunkte hinweg mit der Referenzgröße gebildet. Neben der normalen Korrelation werden auch zwei weitere Korrelationen, basierend auf einer Verschiebung des Parametersignals, berechnet. Dazu wird das Parametersignal einmal um einen Messzeitpunkt nach vorne (r_{Plus}) und einmal nach hinten (r_{Minus}) verschoben. So kann festgestellt werden, ob sich ein Parameter schon vor einer Änderung der Referenzgröße verändert (*proaktive Adaption*) oder erst danach (*reaktive Adaption*). Beide Adaptionarten können sowohl auf bewusster als auch auf unbewusster Ebene

stattfinden (Matthews & Campbell, 1998). Die proaktiven Adaptionsprozesse (Abbildung 5.4 links) können vorwiegend Veränderungen in der individuellen Zielstellung oder der Optimierung der Anstrengung zugeordnet werden (Sträter, 2005). Die reaktiven Adaptionen (Abbildung 5.4 rechts) hingegen werden zu- meist durch Optimierungsprozesse hinsichtlich der Leistung, aber auch teilweise durch eine Veränderung des Anforderungsniveaus bestimmt (Sperling & Doshier, 1986).

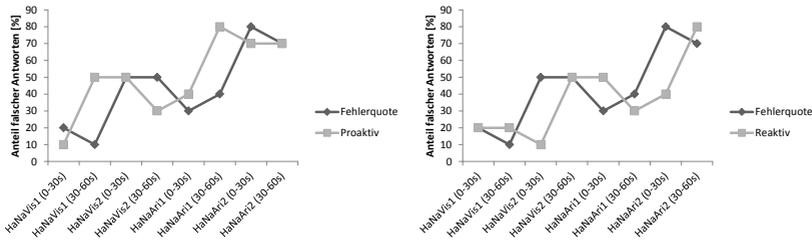


Abb. 5.4: Proaktives und reaktives Adaptionsverhalten: *HAI* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

Die in Abbildung 5.4 dargestellten Beispiele verdeutlichen die Notwendigkeit, das individuelle Adaptionsverhalten in der Beanspruchungsmessung zu berücksichtigen (b in Abbildung 5.3). Sowohl die Signale der rechten, als auch die Signale der linken Abbildung korrelieren mit $r = 0,53$. Nach der Verschiebung des proaktiven Signals nach hinten und des reaktiven Signals nach vorne ergibt sich für beide Fälle eine Korrelationskoeffizient von $r = 1$. Nicht nur für die Systemgestaltung, sondern auch für die Parameterauswahl ergeben sich durch diese Betrachtung erhebliche Vorteile. Zeitpunkte von Beanspruchungsveränderungen können noch genauer zugeordnet und in systemischen Gestaltungsmaßnahmen berücksichtigt werden. Physiologische Indikatoren, die auf den ersten Blick ungeeignet erscheinen, zeigen ein deutlich höheres Potential, insbesondere zur proaktiven Fehlerprävention in Echtzeitsystemen. Im nächsten Schritt der Parameterklassifikation und -auswahl wird zunächst geprüft, ob die normalen Korrelationen zwischen den Parametern und der Referenzgröße besonders hoch

und der Winkel α aus der Procrustes-Analyse des Gesamtsignals besonders klein ist, da die grundlegende Ähnlichkeit der Signale eine weitere Anforderung der Parameterauswahl darstellt (c in Abbildung 5.3). Dazu werden die 95 % Konfidenzintervalle der Korrelationskoeffizienten r_{Normal} und der Winkel α über die verbleibenden bewussten und unbewussten Parameter berechnet. Es werden lediglich Parameter weiter berücksichtigt, für die gilt:

$$r_{Normal} > \left(\bar{r} + z_{\left(1-\frac{0,05}{2}\right)} \frac{\sigma}{\sqrt{n}} \right) \&\& \quad (5.5)$$

$$\alpha_{Gesamt} < \left(\bar{\alpha} - z_{\left(1-\frac{0,05}{2}\right)} \frac{\sigma}{\sqrt{n}} \right) \quad (5.6)$$

Für die verbleibenden $k = 61 - x$ physiologischen Parameter werden die mittleren Adaptionseffekte AE der bewussten und unbewussten Parameter verglichen. Der Adaptionseffekt jedes verbleibenden Parameters berechnet sich dabei anhand der Steigerung in der Varianzaufklärung, die durch eine Verschiebung des Signals bewirkt wird:

$$\begin{aligned} r_{Max} &= r_{Normal}, \text{ dann } AE = 0 \\ r_{Max} &= r_{Plus}, \text{ dann } AE = r_{Plus} - r_{Normal} \\ r_{Max} &= r_{Minus}, \text{ dann } AE = r_{Minus} - r_{Normal} \end{aligned} \quad (5.7)$$

Es werden nachfolgend die bewussten oder unbewussten Parameter zum Verbleib in der weiteren Parameterauswahl anhand des mittleren Adaptionseffektes und der Homogenität des Adaptionsverhaltens ausgewählt. Ein einheitliches Verhalten hinsichtlich proaktiver oder reaktiver Adaptionen im Zusammenhang mit großen Adaptionseffekten innerhalb einer Parameterkategorie weist auf eine höhere Stabilität des Verhaltens, auch bei der Übertragung auf andere Aufgaben, hin. Zur Klassifikation der Homogenität wird die Varianz zwischen den kodierten maximalen Korrelationskoeffizienten herangezogen. Nach der Festlegung auf die verbleibenden bewussten oder unbewussten Parameter wird als letztes der Steigungswinkel φ der Gesamtsignale betrachtet. Ausgangspunkt für diese

Betrachtung ist die partielle Ausnutzung des Wertebereichs der Referenzgröße durch die vier Anforderungsniveaus der Kalibrierungsaufgabe. In einer späteren Anwendungsaufgabe können beispielsweise Fehlerquoten auftreten, die weit über oder unter den erfassten Werten aus der Kalibrierungsaufgabe liegen, auf denen aufbauend das individuelle Beanspruchungsmodell erstellt wird. Um die Übertragbarkeit der Parameter auf den gesamten Wertebereich der Referenzgröße bewerten zu können, wird der Steigungswinkel φ der Parametersignale mit dem des Referenzsignals verglichen, um einen möglichen Drift der beiden Kurven ausschließen zu können (d in Abbildung 5.3). Nur, wenn eine der folgenden Kausalitätsketten erfüllt ist, kann sichergestellt werden, dass die Parameter sich über den gesamten Wertebereich der Referenzgröße einheitlich zu den betrachteten Messpunkten der Kalibrierung verhalten:

$$\begin{aligned} \varphi_{Parameter} > \varphi_{Referenz} \ \&\& \ Min_{Parameter}(y) < Min_{Referenz}(y), \text{ dann kein Drift} \\ \varphi_{Parameter} < \varphi_{Referenz} \ \&\& \ Min_{Parameter}(y) > Min_{Referenz}(y), \text{ dann kein Drift} \\ &&& \text{sonst, Drift} \end{aligned} \tag{5.8}$$

Nach Anwendung dieses letzten Kriteriums zur Parameterauswahl, können drei Möglichkeiten bezüglich der verbleibenden Anzahl k der Parameter mit den nachstehenden Schlussfolgerungen eintreten:

$$\begin{aligned} k = 0, \text{ keine Vorhersage dieser Person möglich} \\ k = 1, \text{ Vorhersage mit diesem Parameter möglich} \\ k > 1, \text{ Auswahl anhand des größten Adaptionseffektes} \end{aligned} \tag{5.9}$$

Sollte nur noch $k = 1$ Parameter nach den Auswahlkriterien übrigbleiben, so wird dieser für die Vorhersage verwendet. Sollten mehr als ein Parameter die Anforderungen erfüllen, so wird der Parameter mit dem größten Adaptionseffekt ausgewählt (Formel 5.7), da dieser die deutlichste Vorraussage und somit

das höchste Potential der Übertragbarkeit auf andere Aufgaben bietet. Für den Fall, dass kein Parameter die Kriterien erfüllt ($k = 0$), so ist eine zuverlässige Vorhersage des Probanden anhand der definierten Anforderungen nicht möglich. Die Auswertungssoftware bietet dem Versuchsleiter jedoch die Option in diesem Fall bestimmte Kriterien (z. B. Aufgaben- oder Zeitabhängigkeit) auszuschließen. Durch diesen Schritt wird die Zuverlässigkeit der Vorhersage deutlich reduziert, was in praktischen Anwendungen nicht zu empfehlen ist. In Laboruntersuchungen zur mentalen Beanspruchung und deren Verlauf, kann die Variation oder Definition neuer Kriterien großes Potential bieten. Um abschließend die Vergleichbarkeit zwischen den Probanden zu ermöglichen, wird ähnlich dem Ansatz der individuellen Regressionsfunktionen ein funktioneller Zusammenhang zwischen dem Parameter und der gewählten standardisierten Referenzgröße gebildet.

$$f(p) = b_0 + b_1 \times P \quad (5.10)$$

Perspektivisch kann die dargestellte Vorgehensweise auf die Verwendung eines bewussten und eines unbewussten Parameters in einer Regressionsfunktion erweitert werden. Diese Art der Parameterauswahl bietet viele Anknüpfungspunkte für weitere Forschungsarbeiten zur mentalen Beanspruchung. In der nachfolgenden Studie 2 (Kapitel 6) soll zunächst die Kalibrierungsaufgabe und insbesondere die Operationalisierung der Schwierigkeitsstufen empirisch untersucht werden, bevor in Studie 3 (Kapitel 7) die beiden vorgestellten Auswertungsverfahren verglichen werden sollen. Dabei soll der Fokus auf der Möglichkeit der Vorhersage einer zweiten anwendungsnahen Aufgabe liegen.

6 Studie 2: Empirische Erprobung der Kalibrierungsaufgabe

Die erste Studie dieser Arbeit (Kapitel 4) umfasste die Untersuchung von physiologischen okularen Beanspruchungsindikatoren hinsichtlich ihrer Sensitivität und Diagnostizität auf der generellen, der differenziellen sowie der individuellen Ebene. Die Ergebnisse bestätigten die Steigerung der Zuverlässigkeit der physiologischen Beanspruchungsmessung auf der differenziellen Ebene, wie auch Fels et al. (2015) herausgestellt haben. Allerdings war die Auswertung auf der Gruppenebene nicht stabil, was einen Einsatz zur Vorhersage der Beanspruchung deutlich erschwert. Die Betrachtung der Daten auf individueller Ebene konnte die Varianzaufklärung der physiologischen Parameter an dem subjektiven Empfinden sowie der Leistung nochmals deutlich steigern, was die Vermutungen von Gagnon et al. (2011) hinsichtlich des großen Potentials einer individuellen Auswertung bestätigt. Um eine individuelle, aber vergleichbare Untersuchung der mentalen Beanspruchung über physiologische Parameter mit besonderem Blick auf den Aufwand praktikabel zu machen, wurde in Kapitel 5 eine Kalibrierungsaufgabe entwickelt. Diese Aufgabe soll von den Probanden vor der eigentlich zu untersuchenden Tätigkeit ausgeführt werden, um eine individuelle Beanspruchungsfunktion anhand einer wählbaren Referenz (subjektives Empfinden, Leistung, mentale Effizienz) abzuleiten. Mit dieser Funktion wird versucht die mentale Beanspruchung in der Hauptaufgabe vorhersagen zu können. In dieser Studie wird die entwickelte Kalibrierungsaufgabe empirisch untersucht, insbesondere steht dabei die Operationalisierung der einzelnen Anforderungsniveaus im Fokus. Bei der Verwendung von Nebenaufgaben ist

die Wechselwirkung mit der Hauptaufgabe ein kritischer Punkt (Kapitel 3.1.2). Aus diesem Grund werden sowohl die Haupt- als auch die Nebenaufgaben einzeln und in Kombination durchgeführt. So lassen sich die Anforderungen der einzelnen Aufgaben sowie die Interaktionseffekte zwischen den Aufgaben beschreiben. Weiter soll die Kalibrierungsaufgabe hier nach dem Subsidiary Task Paradigm ausgeführt werden, weshalb die Leistung der Hauptaufgabe unter Hinzunahme der Nebenaufgabe nahezu konstant bleiben sollte, während die Leistung in der Nebenaufgabe aber deutlich abfallen müsste. Darüber hinaus müssen die Anforderungen so abgestimmt werden, dass weder eine Über- noch eine Unterforderung auftritt. Folgend dem Modell nach De Waard (1996) sollten sich die Anforderungen der Aufgabenkombinationen idealerweise innerhalb der Bereiche A1 bis A3 bewegen, um die mentale Beanspruchung nach dem trimodalen Ansatz optimal erfassen zu können (Kapitel 2.2.1).

6.1 Experimentaldesign der Studie 2

Im Rahmen dieser Studie wird eine Stichprobe von $N = 39$ Probanden untersucht. Die Probanden führen zunächst nur die Hauptaufgaben (eine Kugel, zwei Kugeln) der Kalibrierungsaufgabe aus, bevor dann die beiden Nebenaufgaben alleinstehend bearbeitet werden. Abschließend werden die vier Anforderungsniveaus der Kalibrierungsaufgabe als Kombination aus Haupt- und Nebenaufgaben ausgeführt. Die Reihenfolge der Aufgabenteile wird randomisiert, um Reihenfolgeeffekte zu vermeiden. Die mentale Beanspruchung wird nach dem trimodalen Ansatz (Kapitel 3.1) über subjektive, leistungsbasierte und physiologische Parameter erfasst. Zur Erfassung der physiologischen, okularen Beanspruchungsparameter wird, wie auch in Studie 1, das binokulare 60Hz Blickerfassungssystem Dikablis Professional (Ergoneers GmbH) verwendet. Das subjektive Beanspruchungsempfinden wird mittels des NASA-TLX Fragebogens nach jedem Aufgabenteil erfasst. Die Leistung wird sowohl für die Hauptaufgaben, als auch für die Nebenaufgabe, entsprechend den Beschreibungen der Leistungsauswertung in Kapitel 5.2, teilautomatisiert erhoben. Der Versuch dauert insgesamt zirka 60 Minuten. Zu Beginn des Versuchs wird der

Proband über die verwendete Messtechnik und den Versuchsablauf informiert. Danach wird das Blickerfassungssystem angelegt und kalibriert. Die Instruktionen des Versuchsleiters folgen einem, vor dem Versuch ausgearbeiteten, Protokoll, um Störeinflüsse durch abweichende oder unvollständige Instruktionen zu vermeiden. Das Blickerfassungssystem wird während des gesamten Versuchs getragen, allerdings vor jedem Versuchsteil neu kalibriert, um Messartefakte durch das mögliche Verrutschen des Systems zu vermeiden.

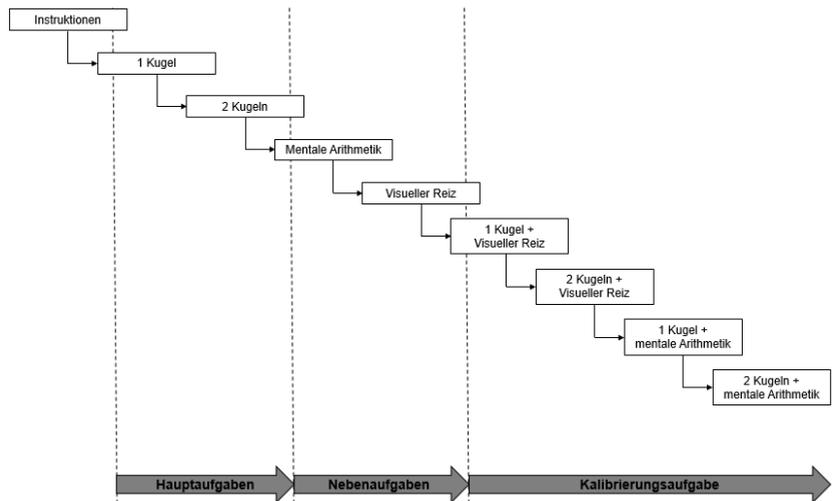


Abb. 6.1: Versuchsablauf Studie 2. Die Nebenaufgabenbedingungen in dem zweiten und dritten Abschnitt wurden randomisiert (eigene Abbildung)

Nach den Instruktionen beginnt der Proband mit der Bearbeitung der Hauptaufgabe ohne Nebenaufgabe. Zunächst soll der Proband nur die linke Kugel steuern, bevor sich dann im zweiten Teil beide Kugeln bewegen. Im Anschluss werden die arithmetische (NA-Ari) und die visuelle Nebenaufgabe (NA-Vis) in randomisierter Reihenfolge ohne Hauptaufgabe ausgeführt. Für die arithmetische Nebenaufgabe werden in dieser Studie Additionsaufgaben der Kategorie „Leicht“ verwendet (Kapitel 3.1.2). Während dieser beiden Aufgabenteile werden die Probanden gebeten auf ein weißes Kreuz in der Mitte des schwarzen Bildschirms zu sehen, um zum einen ein Verschließen der Augen während der

mentalen Arithmetik zu vermeiden und zum anderen zu gewährleisten, dass sich die visuellen Reize im peripheren Sichtfeld befinden. Die letzten vier Versuchsteile umfassen die definierten Anforderungsniveaus der Kalibrierungsaufgabe, bestehend aus der Kombination von Haupt- und Nebenaufgaben. Dabei wird die Reihenfolge der Nebenaufgabenbedingungen ebenfalls randomisiert (Abbildung 6.1).

6.2 Ergebnisse der Studie 2

Die Auswertung der zweiten Studie teilt sich in zwei Abschnitte auf. Zunächst werden die Operationalisierung der Haupt- und Nebenaufgaben sowie deren Interaktionseffekte in der Kalibrierungsaufgabe untersucht. Dabei werden im ersten Abschnitt (Kapitel 6.2.1) die subjektiven und leistungsbasierten Beanspruchungsindikatoren im Fokus stehen, bevor dann in Abschnitt 6.2.2 die individuelle Auswertung der physiologischen Parameter folgt. Die Stichprobe umfasst $N = 39$ (25 männliche, 14 weibliche) Probanden, im Wesentlichen Studenten, die im Mittel 24,31 Jahre alt sind ($SD = 3,26$). Nahezu alle Probanden haben ein uneingeschränktes oder korrigiert-uneingeschränktes Sehvermögen (11 Brillenträger), ein Proband hat eine leichte Rot-Grün-Schwäche. Aufgrund des Stichprobenumfangs werden zur Auswertung dieser Studie, gemäß dem zentralen Grenzwerttheorem, parametrische Tests verwendet (Kapitel 3.2.2).

6.2.1 Unersuchung der Operationalisierung der Kalibrierungsaufgabe

Um die Operationalisierung der Haupt- und Nebenaufgaben sowie deren Interaktionseffekte zu untersuchen, sollen zunächst die subjektiven und leistungs-basierten Daten herangezogen werden. Das subjektive Beanspruchungsempfinden (Abbildung 6.2) unterscheidet sich signifikant zwischen den einzelnen Aufgabenteilen ($F(7, 266) = 34,81, p < 0,01$). In den ersten Versuchsteilen, in denen nur die Nebenaufgaben bearbeitet wurden (NA-Ari, NA-Vis), ist die subjektiv empfundene Beanspruchung bei der mentalen Arithmetik signifikant höher als bei dem visuellen Reiz ($t(38) = 2,96, p < 0,01$). Dieser Unterschied begrün-

det sich zunächst darin, dass sich die Anforderungen des visuellen Reizes, im Vergleich zur mentalen Arithmetik, erst in Anwesenheit einer Hauptaufgabe vollständig auswirken. Weiter könnte der Effekt der sozialen Erwünschtheit bei der mentalen Arithmetik einen höheren Einfluss nehmen und so auch auf die empfundene Beanspruchung wirken (Bortz & Schuster, 2010).

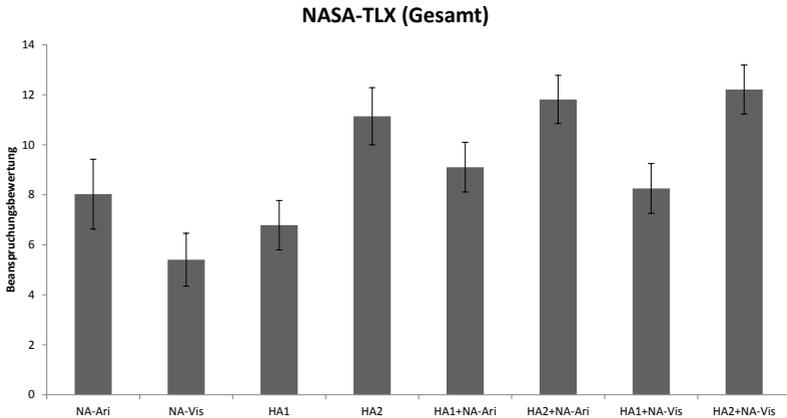


Abb. 6.2: NASA-TLX Bewertungen der einzelnen Anforderungsniveaus: *HA1* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

Das Frustrationsitem des NASA-TLX korreliert für die mentale Arithmetik signifikant mit der Fehlerquote der Nebenaufgabe ($r = 0,53$, $p < 0,01$). Für den visuellen Reiz wird dieser Zusammenhang nicht signifikant ($r = 0,19$, *n. s.*), was den Effekt der sozialen Erwünschtheit bestätigen würde. Allerdings wurde das Frustrationsitem für die Arithmetik ($M = 5,79$, $SD = 5,53$) nur leicht und nicht signifikant höher bewertet als für den visuellen Reiz ($M = 3,64$, $SD = 3,12$). Beide Bewertungen sind dazu im Mittel sehr gering, weshalb der Effekt der sozialen Erwünschtheit hier nicht weiter fokussiert werden soll. Die Ergebnisse zeigen aber, dass der Effekt durchaus einen wichtigen Aspekt für die Verwendung von arithmetischen Nebenaufgaben darstellen kann. Des Weiteren zeigt sich ein deutlicher Unterschied zwischen den beiden Hauptaufgaben

(HA1: 1 Kugel, HA2: 2 Kugeln). In der Bearbeitung ohne Nebenaufgaben ist das subjektive Beanspruchungsempfinden für zwei Kugeln signifikant höher als für eine Kugel ($t(38) = -10,91, p < 0,01$). Dieser Effekt bleibt unabhängig von der Nebenaufgabe sowohl in der arithmetischen ($t(38) = -7,04, p < 0,01$), als auch in der visuellen ($t(38) = -8,81, p < 0,01$) Nebenaufgabenbedingung konstant. Demnach sind die Anforderungsniveaus der Hauptaufgaben sehr gut hinsichtlich des subjektiven Beanspruchungsempfindens differenzierbar. Für die Hauptaufgabe mit einer Kugel (HA1) zeigen sich signifikante Anstiege des Beanspruchungsempfindens unter Hinzunahme der visuellen ($t(38) = -4,89, p < 0,01$) und der arithmetischen ($t(38) = -3,21, p < 0,01$) Nebenaufgabe. Diese Unterschiede zeigen sich bei den Hauptaufgaben mit zwei Kugeln (HA2) nicht, hier bleibt das Beanspruchungsempfinden nahezu konstant. Dies kann auf das instruierte Subsidiary Task Paradigm zurückgeführt werden. Während der leichteren Hauptaufgaben können die Nebenaufgaben noch relativ gut mitbearbeitet werden, wohingegen bei den schwereren Hauptaufgaben die Nebenaufgaben mehr und mehr vernachlässigt werden. Die Hauptaufgabe ist im Fokus und die Nebenaufgabe wird nur dann bearbeitet, wenn der Proband noch über freie Kapazitäten verfügt. Die Auswertung der Leistung in den Hauptaufgaben bestätigt diesen Eindruck (Abbildung 6.3). Insgesamt liegen signifikante Unterschiede zwischen den einzelnen Aufgabenteilen in der mittleren Kugelposition vor ($F(5, 190) = 80,06, p < 0,01$), wobei die Leistung mit nur einer Kugel (HA1) sowohl ohne Nebenaufgabe ($t(38) = -12,87, p < 0,01$), als auch mit arithmetischer ($t(38) = -10,15, p < 0,01$) und mit visueller Nebenaufgabe ($t(38) = -10,97, p < 0,01$) signifikant besser ist als mit zwei Kugeln (HA2). Die gute Differenzierbarkeit der Anforderungsniveaus zeigt sich demnach auch in der Leistung der Hauptaufgabe. Bei den Versuchsteilen mit einer Kugel unterscheiden sich die Leistungen nicht zwischen den beiden Nebenaufgabenbedingungen, wohingegen bei den Versuchsteilen mit zwei Kugeln die mittlere Kugelposition in der visuellen Nebenaufgabenbedingung signifikant höher ist als in der arithmetischen ($t(38) = -3,92, p < 0,01$).

Dies kann unter anderem auch auf die unterschiedlichen Reaktionen auf einen visuellen Reiz im peripheren Sichtfeld zurückgeführt werden. Probanden ver-

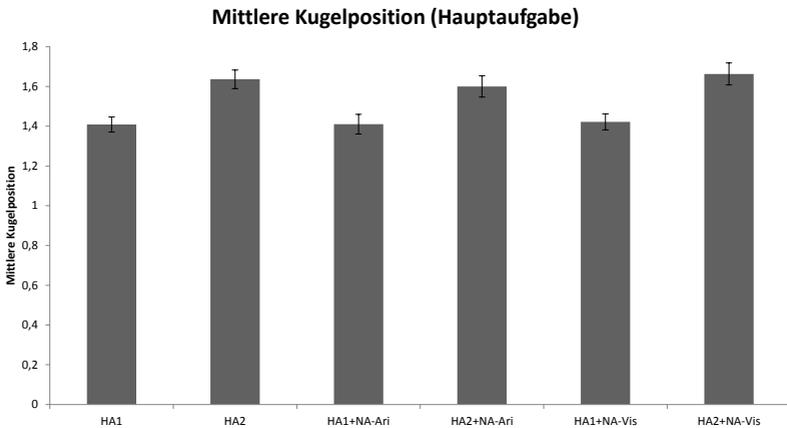


Abb. 6.3: Leistung in der Hauptaufgabe - mittlere Kugelposition von 0 (Mitte) bis 4 (Randbereich): *HA1* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

schieben möglicherweise kurzzeitig ihre Aufmerksamkeit auf den Reiz im peripheren Sichtfeld. Ein solcher Aufmerksamkeitswechsel wirkt sich bei einer Kugel noch nicht aus, da die zu verarbeitende Informationsmenge wahrscheinlich noch handhabbar ist. Bei zwei Kugeln wirkt sich dieser Wechsel in einer Latenzzeit aus, die der Proband braucht, um wieder den Zustand beider Kugeln zu erfassen. Diese Erkenntnisse entsprechen den Interaktionseffekten innerhalb des Würfelmodells (Abbildung 2.11) nach Wickens et al. (2013). In den Versuchsteilen, in denen nur die Nebenaufgabe bearbeitet wurde, zeigen sich diese Aufmerksamkeitswechsel auch in signifikanten Unterschieden im NNI ($t(38) = -2,28, p < 0,05$) und den Fixationsdauern ($t(38) = -2,59, p < 0,05$) zwischen der arithmetischen und der visuellen Nebenaufgabe. Obwohl die Probanden instruiert wurden, sich auf das weiße Kreuz in der Mitte des Bildschirms zu konzentrieren, hatte das Blickverhalten bei der visuellen Aufgabe eine signifikant höhere Entropie (NNI) und war somit deutlich unstrukturierter. In den Versuchsteilen mit einer Hauptaufgabe zeigt sich dieser Unterschied zwischen den Nebenaufgabenbedingungen nicht mehr. Daraus lässt sich ableiten, dass der

Effekt von Aufmerksamkeitswechsellern in der visuellen Nebenaufgabenbedingung durch die Anwesenheit einer Hauptaufgabe deutlich verringert wird. Die bisher erfasste Fokussierung der Probanden auf die Hauptaufgabe, entsprechend des Subsidiary Task Paradigm, wird durch dieses Ergebnis und den signifikanten Zusammenhang zwischen der Leistung in der Hauptaufgabe und dem subjektiven Beanspruchungsempfinden ($r = 0,49, p < 0,01$) nochmals bestätigt. Weiter bestätigt auch die Analyse der Leistungsdaten aus den Nebenaufgaben die bisherigen Ergebnisse bezüglich der Kalibrierungsaufgabe (Abbildung 6.4). Die Fehler in der arithmetischen ($F(2, 76) = 3,64, p < 0,01$) und der visuellen ($F(2, 76) = 27,15, p < 0,01$) Nebenaufgabe unterscheiden sich signifikant zwischen den Anforderungsniveaus. Die Fehlerhäufigkeit der arithmetischen Nebenaufgabe ist in der Hauptaufgabenbedingung mit zwei Kugeln signifikant höher als mit nur einer Kugel ($t(38) = -2,21, p < 0,01$) und ohne Hauptaufgabe ($t(38) = -2,35, p < 0,01$). Diese Unterschiede bestätigen die Erkenntnisse aus der Auswertung des subjektiven Beanspruchungsempfindens ($r = 0,37, p < 0,01$). Während bei der Hauptaufgabe mit einer Kugel noch ausreichend Kapazitäten für die Bearbeitung der Nebenaufgabe zur Verfügung stehen, wird bei zwei Kugeln jedoch die Nebenaufgabe zugunsten der Leistungserhaltung in der Hauptaufgabe vernachlässigt. Der Zusammenhang mit der Hauptaufgabe ist zwar signifikant ($r = 0,25, p < 0,05$), besitzt aber dennoch eine eher geringe Varianzaufklärung, was ein weiterer Indikator für die geringeren Interaktionseffekte der verschiedenen Darstellungsmodalitäten ist. Die Anzahl der Fehler in der arithmetischen Nebenaufgabe ist jedoch insgesamt sehr gering.

In Kombination mit der Hauptaufgabe mit zwei Kugeln wurden im Mittel $M = 1,05$ ($SD = 1,21$) Fehler gemacht. Als Beanspruchungsindikator ist die arithmetische Nebenaufgabe damit zu leicht, weshalb die leichten Additionsaufgaben in der nächsten Studie (Kapitel 7) durch mittelschwere Subtraktionsaufgaben ersetzt werden sollen. Die geringere Fehlerhäufigkeit erklärt darüber hinaus auch die bessere Leistung in der zweiten Hauptaufgabe, verglichen mit der visuellen Nebenaufgabe (Abbildung 6.3), da sie relativ einfach nebenbei ausgeführt werden konnte. Auch die Fehlerhäufigkeit der visuellen Nebenaufgabe ist in der Hauptaufgabenbedingung mit zwei Kugeln signifikant höher als mit nur einer

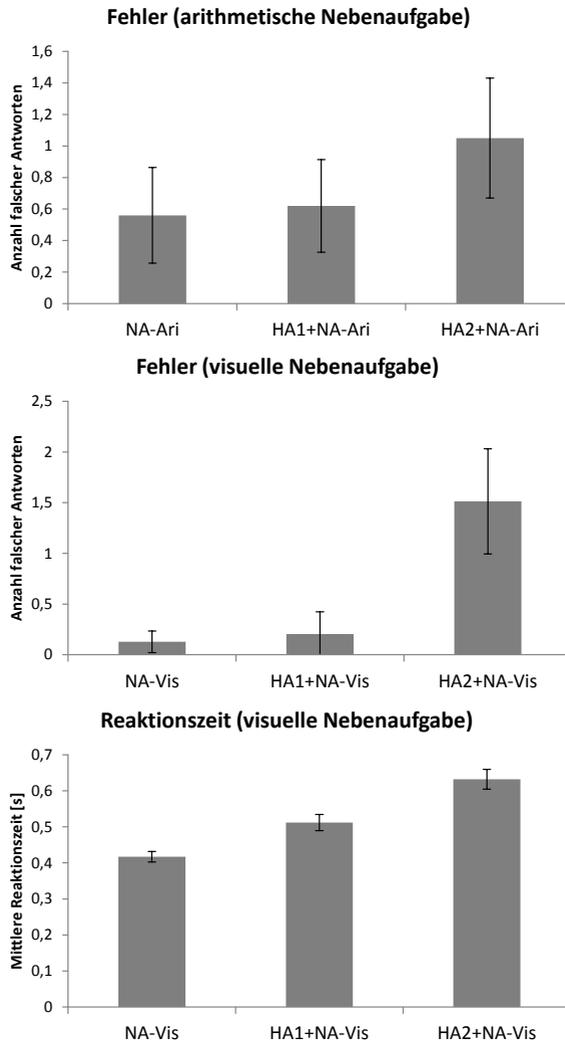


Abb. 6.4: Leistung in der Nebenaufgabe: *HA1* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

Kugel ($t(38) = -5,76, p < 0,01$) und ohne Hauptaufgabe ($t(38) = -5,2, p < 0,01$). Damit bestätigen auch diese Ergebnisse des visuellen Reizes die Erkenntnisse aus der NASA-TLX Auswertung ($r = 0,39, p < 0,01$). Weiter bekräftigt die starke Interaktion mit der Leistung in der Hauptaufgabe ($r = 0,54, p < 0,01$) die Vermutung über kurze Aufmerksamkeitswechsel zu den Reizen im peripheren Sichtfeld und damit auch die theoretischen Ausführungen des Würfelmodells (Abbildung 2.11) nach Wickens et al. (2013). Weiter unterscheiden sich auch die Reaktionszeiten der visuellen Nebenaufgabe signifikant zwischen den Anforderungsniveaus ($F(2, 76) = 209,24, p < 0,01$). Die Reaktionszeiten auf einen visuellen Reiz erhöhen sich sowohl durch die Hinzunahme der Hauptaufgabe mit einer Kugel ($t(38) = 9,95, p < 0,01$), als auch dann noch einmal durch die zweite Kugel ($t(38) = -11,87, p < 0,01$) signifikant. Die Reaktionszeit ist ein hoch sensibles Maß und zeigt ebenfalls die starken Interaktionseffekte mit der Hauptaufgabe ($r = 0,7, p < 0,01$), wie auch mit dem NASA-TLX ($r = 0,51, p < 0,01$).

Zusammenfassend erfüllt die Operationalisierung der Flugaufgabe über die beiden Kugeln in der Kalibrierungsaufgabe und die Differenzierung der Anforderungsniveaus die definierten Anforderungen (Kapitel 5.1). Sowohl die subjektiven als auch die leistungsbasierten Beanspruchungsindikatoren zeigen starke Unterschiede zwischen den Anforderungsniveaus und untereinander widerspruchsfreie, starke Zusammenhänge. Insgesamt bewegt sich die induzierte Beanspruchung innerhalb der gewünschten Bereiche A1 bis A3 gemäß der Definition nach De Waard (1996). Die Anforderungen der arithmetischen Nebenaufgabe müssen jedoch angepasst werden, da eine mittlere Fehlerhäufigkeit von $M = 1,05$ ($SD = 1,21$) bei zwölf Aufgaben nicht ausreichend ist, um auch in dieser Nebenaufgabenbedingung über die Grenze der Bereiche D und A1 hinauszukommen (Kapitel 2.2.1). Durch die in der Auswertung deutlich gewordenen unterschiedlich starken Interaktionseffekte zwischen Haupt- und Nebenaufgaben in den beiden Nebenaufgabenbedingungen, ist es möglich, zwei Aufgabenarten zu unterscheiden. Die beiden Bedingungen stellen unterschiedliche Anforderungen an den Anwender. Diese Differenzierbarkeit ist notwendig,

um eine Aufgabenabhängigkeit von physiologischen Beanspruchungsindikatoren innerhalb der individuellen Parameterauswahl (Kapitel 5.3.2) feststellen zu können. Da die Kalibrierungsaufgabe in der Anwendung vor den zu analysierenden Tätigkeiten durchgeführt werden soll, ist es eine wichtige Anforderung, den Probanden durch die Aufgaben nicht zu erschöpfen oder Stress auslösen. Das Frustrationsitem des NASA-TLX ist im Mittel mit $M = 6,34$ ($SD = 4,71$) über alle Versuchsteile hinweg relativ gering ausgeprägt. Es konnte auch kein Zusammenhang zwischen der Versuchsdauer und den weiteren NASA-TLX Items festgestellt werden, somit erfüllt die Kalibrierungsaufgabe auch in dieser Hinsicht die Anforderungen.

6.2.2 Auswertung der physiologischen Parameter

Wie in der ersten Studie werden die physiologischen Parameter zunächst auf der generellen Ebene hinsichtlich ihrer Diagnostizität und Sensitivität untersucht. Nachfolgend werden die beiden vorgestellten Auswertungskonzepte zur Erstellung der individuellen Beanspruchungsfunktionen (Kapitel 5.3) auf der individuellen Ebene angewendet und verglichen. Als Grundlage für die individuelle Beanspruchungsmodellierung dienen lediglich die vier letzten Versuchsteile, die der später anzuwendenden Kalibrierungsaufgabe entsprechen.

Generelle Ebene

Zunächst soll die Diagnostizität der einzelnen Parameter bezüglich der subjektiven und leistungsbasierten Parameter bewertet werden. Tabelle 6.1 fasst die Ergebnisse der Korrelationsanalysen beider Nebenaufgabenbedingungen zusammen. Eine Nebenaufgabenbedingung umfasst dabei den Versuchsteil ohne Hauptaufgabe und die beiden Versuchsteile mit Hauptaufgabe (eine Kugel und zwei Kugeln).

Bei der Korrelationsanalyse fällt zunächst ein großer Unterschied zwischen den bewussten und unbewussten physiologischen Parametern hinsichtlich ihrer Diagnostizität auf: Die bewussten Parameter (NNI, Fixationsdauer, Sakkaden-

Tab. 6.1: Korrelationsanalyse physiologischer Parameter Studie 2: NA = Nebenaufgabe, RT = Reaktionszeit (eigene Tabelle)

	arithmetische Nebenaufgabe			visuelle Nebenaufgabe			
	NASA-TLX	Fehler NA	Fehler HA	NASA-TLX	Fehler NA	RT NA	Fehler HA
Blinzelrate	$r = -0,04$ <i>n. s.</i>	$r = 0,08$ <i>n. s.</i>	$r = 0,31$ $p < 0,01$	$r = 0$ <i>n. s.</i>	$r = -0,21$ $p < 0,05$	$r = -0,05$ <i>n. s.</i>	$r = 0,18$ $p < 0,05$
Blinzeldauer	$r = -0,03$ <i>n. s.</i>	$r = -0,05$ <i>n. s.</i>	$r = 0,18$ $p < 0,05$	$r = 0$ <i>n. s.</i>	$r = 0,01$ <i>n. s.</i>	$r = 0,06$ <i>n. s.</i>	$r = 0,11$ <i>n. s.</i>
PERCLOS	$r = 0,06$ <i>n. s.</i>	$r = 0,05$ <i>n. s.</i>	$r = 0,33$ $p < 0,01$	$r = 0,08$ <i>n. s.</i>	$r = -0,02$ <i>n. s.</i>	$r = 0,02$ <i>n. s.</i>	$r = 0,25$ $p < 0,01$
Pupillendurchschn.	$r = 0,05$ <i>n. s.</i>	$r = 0,04$ <i>n. s.</i>	$r = 0,01$ <i>n. s.</i>	$r = -0,01$ <i>n. s.</i>	$r = 0,11$ <i>n. s.</i>	$r = 0,13$ <i>n. s.</i>	$r = -0,01$ <i>n. s.</i>
NNI	$r = 0,41$ $p < 0,01$	$r = 0,2$ $p < 0,05$	$r = 0,46$ $p < 0,01$	$r = 0,36$ $p < 0,01$	$r = 0,42$ $p < 0,01$	$r = 0,48$ $p < 0,01$	$r = 0,48$ $p < 0,01$
Fixationsdauer	$r = -0,13$ <i>n. s.</i>	$r = -0,14$ $p < 0,05$	$r = -0,22$ $p < 0,01$	$r = -0,18$ $p < 0,05$	$r = -0,22$ $p < 0,01$	$r = -0,41$ $p < 0,01$	$r = -0,28$ $p < 0,01$
Sakkadenzlängen	$r = 0,37$ $p < 0,01$	$r = 0,13$ <i>n. s.</i>	$r = 0,53$ $p < 0,01$	$r = 0,46$ $p < 0,01$	$r = 0,46$ $p < 0,01$	$r = 0,6$ $p < 0,01$	$r = 0,53$ $p < 0,01$

längen) haben eine deutlich höhere Varianzaufklärung, unabhängig von der Nebenaufgabenbedingung oder der Referenzgröße. Die Pupillendurchschnittsvariabilität zeigt keine signifikanten Zusammenhänge, während die Blinzelrate, Blinzeldauer und der PERCLOS in der arithmetischen Nebenaufgabenbedingung lediglich signifikante Zusammenhänge mit der Leistung in der Hauptaufgabe aufweisen. Dies kann auf der einen Seite die Ergebnisse bezüglich des hohen Einflusses der Hauptaufgabe auf die Gesamtbeanspruchungshöhe (Kapitel 6.2.1), wie es auch von der Operationalisierung der Kalibrierungsaufgabe beabsichtigt ist, bestätigen. Auf der anderen Seite ist die Varianzaufklärung der drei signifikanten Zusammenhänge auf Stichprobenebene sehr gering, weshalb dieser Rückschluss, basierend auf den unbewussten Parametern, nicht sicher getroffen werden kann. Für die visuelle Nebenaufgabenbedingung zeigt sich ein ähnliches Bild. Es existiert nur ein signifikanter Zusammenhang zwischen der Blinzelrate und der Fehlerhäufigkeit in der Nebenaufgabe. Beide Zusammenhänge zwischen der Blinzelrate und den Leistungsparametern in der visuellen Nebenaufgabe sind antiproportional, wohingegen alle anderen Zusammenhänge der Blinzelrate proportional sind. Aufgrund dessen, dass die drei Referenzgrößen

über beide Nebenaufgabenbedingungen hinweg signifikante, aber proportionale, Zusammenhänge aufweisen (Kapitel 6.2.1), kann der Vorzeichenwechsel in der Blinzelrate als weiterer Indikator für die starke Individualität der physiologischen Reaktionen des Menschen betrachtet werden. Die bewussten Parameter zeigen für beide Nebenaufgabenbedingungen eine deutlich höhere Diagnostizität. Insgesamt zeigt jeder bewusste Parameter ein homogenes Bild bezüglich der Richtung (Proportionalität) von Zusammenhängen über die gesamte Stichprobe hinweg. In der arithmetischen Nebenaufgabenbedingung zeigen sich deutlich stärkere Zusammenhänge mit der Hauptaufgabe als mit der Nebenaufgabe, dies bestätigt erneut den höheren Einfluss der Hauptaufgabe auf die Gesamtbeanspruchung und somit die Aufgabengestaltung der Kalibrierungsaufgabe. In der visuellen Nebenaufgabenbedingung zeigen sich signifikante Zusammenhänge aller drei bewussten Parameter mit allen Referenzgrößen. Weiter fällt eine deutliche Steigerung der Varianzaufklärung bezüglich der Leistungsparameter der Nebenaufgabe auf. Dieses Ergebnis indiziert die höheren Interaktionseffekte mit der Hauptaufgabe, verglichen mit der arithmetischen Nebenaufgabenbedingung. Bei der positiven Bewertung der Diagnostizität von bewussten Parametern muss allerdings ein wichtiger Einflussfaktor berücksichtigt werden. Die Anforderungssteigerung in der Hauptaufgabe ist durch die Hinzunahme einer zweiten, simultanen Aufgabe und nicht durch die Veränderung einer einzigen Aufgabe umgesetzt. Somit ergibt sich allein durch die Aufgabengestaltung eine Veränderung des Blickverhaltens. Die stärkeren Zusammenhänge bewusster Parameter mit der visuellen Nebenaufgabe bestätigen diesen Einflussfaktor, der sich auch in der Betrachtung der Sensitivität wiederfinden lässt. Abbildung 6.5 stellt dazu beispielhaft den NNI über die verschiedenen Versuchsteile dar.

Der NNI zeigt eine hohe Sensitivität bezüglich der Anforderungsniveaus in den Hauptaufgaben. Die Unordnung des Blickverhaltens ist bei zwei Kugeln signifikant höher als bei einer Kugel. Dieser Unterschied ist sowohl in der Bedingung ohne Nebenaufgabe ($t(38) = -10,23, p < 0,01$), als auch mit der arithmetischen ($t(38) = -8,61, p < 0,01$) und der visuellen Nebenaufgabe ($t(38) = -6,79, p < 0,01$) signifikant. Eine Sensitivität des NNIs bezüglich der Nebenaufgabe kann jedoch nicht festgestellt werden. Die hohe Sensitivität gegenüber

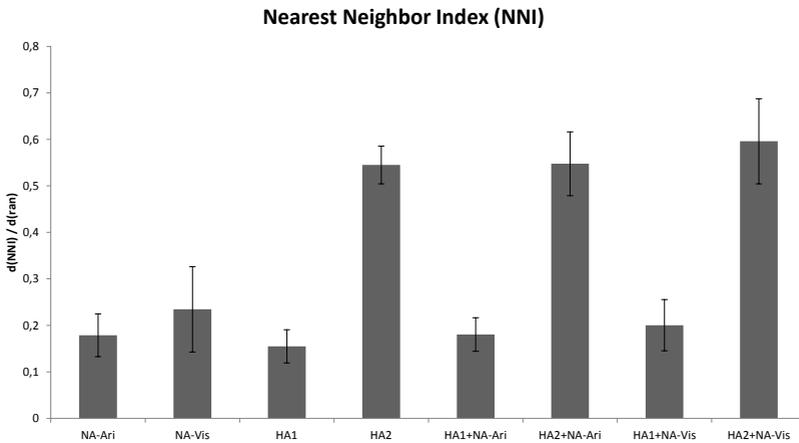


Abb. 6.5: NNI über alle Anforderungsniveaus: *HA1* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

der Hauptaufgabe bestätigt die Auswirkungen des veränderten Blickverhaltens durch die Hinzunahme einer zweiten, simultanen Aufgabe. Die Fixationsdauern und die Sakkadenlängen weisen ähnlich starke Unterschiede auf. Die entsprechenden Diagramme sind im Anhang (S. 214) dieser Arbeit zu finden. Die hohe Diagnostizität und Sensitivität zeigen das große Potential der bewussten Parameter für die Beschreibung der mentalen Beanspruchung auf. Diese Bewertung muss allerdings eingeschränkt werden, da sich viele Zusammenhänge und Unterschiede im Blickverhalten allein durch das Aufgabendesign ergeben. Diese Einschränkung zeigt damit erneut die hohe Bedeutung der Aufgabenunabhängigkeit für die Eignung eines Parameters für die Beanspruchungsmessung, die in dem Verfahren zur individuellen Parameterklassifikation geprüft wird (Kapitel 5.3.2).

Die Blinzelrate zeigt, im Gegensatz zu den bewussten Parametern, eine Sensitivität gegenüber der Nebenaufgabenbedingung (Abbildung 6.6). In den Bedingungen ohne Hauptaufgabe ist die Blinzelrate während der mentalen Arithmetik signifikant höher als während der Aufgabe mit dem visuellen Reiz ($t(38) = 5,27$,

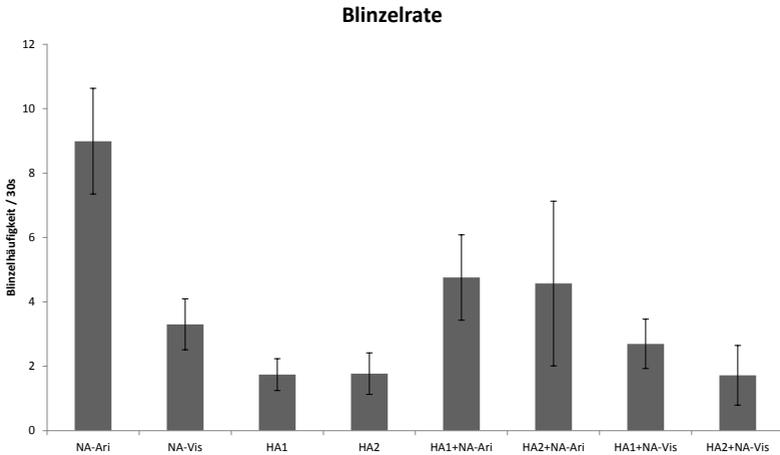


Abb. 6.6: Blinzelrate über alle Anforderungsniveaus: *HA1* = Hauptaufgabe mit einer beweglichen Kugel, *HA2* = Hauptaufgabe mit zwei beweglichen Kugeln, *NA-Ari* = arithmetische Nebenaufgabe, *NA-Vis* = visuelle Nebenaufgabe (eigene Abbildung)

$p < 0,01$). Dieser Unterschied ist auch in Kombination mit den Hauptaufgaben erkennbar, wenn auch nicht signifikant. Die Blinzelrate zeigt also eher eine Sensitivität gegenüber der Nebenaufgabe, obwohl die Diagnostizität hinsichtlich der Leistung in der Nebenaufgabe sehr gering ist (Tabelle 6.1). Dieser starke Unterschied zwischen der Diagnostizität und der Sensitivität zeigt erneut, dass die Auswertung der physiologischen Parameter auf Stichprobenebene keine zuverlässige Bewertung der mentalen Beanspruchung zulässt. Die inter- und intraindividuellen Unterschiede in den physiologischen Reaktionen auf mentale Anforderungen erschweren eine einheitliche Auswertung auf der generellen Ebene, wodurch eine individuelle Auswertung einmal mehr notwendig für eine zuverlässige Beanspruchungsbewertung erscheint.

Individuelle Ebene

Die individuelle Auswertung basiert auf den Daten der letzten vier Versuchsteile (Kombinationen aus Haupt- und Nebenaufgaben), die zusammen die eigentliche

Kalibrierungsaufgabe bilden. Dabei werden beide, in Kapitel 5.3 entwickelten, Auswertungsverfahren angewendet und hinsichtlich der Diagnostizität miteinander verglichen. Abbildung 6.7 stellt die transformierten (Fischers' z -Transformation), über alle Probanden gemittelten, und rücktransformierten Korrelationskoeffizienten r zwischen den individuellen physiologischen Parametern beider Auswertungsverfahren und der Leistung in der Nebenaufgabe dar. Zusätzlich sind die Korrelationskoeffizienten r von der Blinzelrate und dem NNI als Vergleich über die letzten vier Aufgabenteile aufgetragen. Über alle Probanden hinweg ergeben sich hoch signifikante mittlere Korrelationen für die individuellen Regressionsfunktionen ($r_Z = 0,99$, $\sigma_Z = 0,02$, $p < 0,01$) sowie für die individuelle Parameterklassifikation ($r_Z = 0,63$, $\sigma_Z = 0,18$, $p < 0,01$) mit der Leistung in der Nebenaufgabe. Beide Verfahren zur individuellen Beanspruchungsbewertung zeigen eine deutlich höhere Diagnostizität als die beiden beispielhaften Parameter (Blinzelrate, NNI) auf der Stichprobenebene (Tabelle 6.1). Weiter fällt auf, dass das Verfahren der individuellen Regressionen eine signifikant höhere Diagnostizität hinsichtlich der Leistung in der Nebenaufgabe aufweist, als die individuelle Parameterklassifikation ($t(35) = 11,54$, $p < 0,01$). Der signifikante Unterschied in der Diagnostizität zwischen den beiden individuellen Auswertungsverfahren kann im Wesentlichen auf die unterschiedlichen Kriterien zur Auswahl der physiologischen Parameter zurückgeführt werden. Bei der individuellen Regressionsbildung (Kapitel 5.3.1) werden die jeweils am höchsten mit der Referenzinformation korrelierenden Parameter ausgewählt. Darüber hinaus werden keine weiteren Bedingungen an die Parameter gestellt. Die hohe mittlere Korrelation, bei gleichzeitig sehr geringer Streuung, bestärkt die individuelle Herangehensweise, da es für jeden Probanden mindestens einen Parameter gibt, der die Referenzgröße hoch zuverlässig vorhersagen kann, was zu einer erheblichen Steigerung der Diagnostizität führt. Allerdings kann an dieser Stelle keine Aussage darüber getroffen werden, ob die extrahierten Parameter auch aufgabenübergreifend stabil bleiben. Dies soll in Studie 3 (Kapitel 7) geprüft werden. Aus dieser Unsicherheit bezüglich der Stabilität dieses Auswahlverfahrens heraus entstand das Verfahren zur individuellen Parameterklassifikation (Kapitel 5.3.2). Auch dieser Ansatz bietet eine

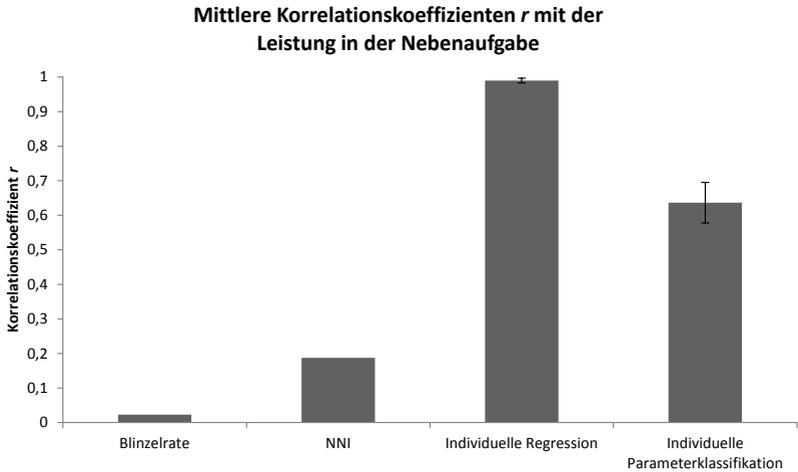


Abb. 6.7: Mittlere Korrelationskoeffizienten mit der Leistung in der Nebenaufgabe (eigene Abbildung)

sehr hohe Diagnostizität, die aber signifikant geringer ist. In dem schrittweisen Auswahlverfahren wird eine Vielzahl von Anforderungen an einen möglichen physiologischen Parameter gestellt, um die Aufgaben-, Anforderungs- und Zeitabhängigkeit zu prüfen. Weiter werden die individuellen Adaptionstrategien der Probanden berücksichtigt. Durch diese starke Reduktion möglicher Parameter unterscheiden sich die individuell ausgewählten Parameter beider Verfahren deutlich. Für $n = 26$ der $N = 39$ Probanden stimmt keiner der beiden, für die Regression ausgewählten, Parameter mit dem aus der Parameterklassifikation überein. Für $n = 3$ Probanden erfüllt kein Parameter die definierten Anforderungen der Parameterklassifikation. Nach dieser Betrachtung erfüllen die höchst korrelierenden Parameter nur bei $n = 10$ Probanden auch die Anforderungen der Parameterklassifikation, was die Vermutung aus den theoretischen Vorüberlegungen (Kapitel 2) bestätigt. Diese sagen aus, dass die Varianzaufklärung als Kriterium allein nicht ausreichend für die individuelle Parameterauswahl ist. Die Kalibrierungsaufgabe bietet aber in der vorliegenden Form die notwendigen Randbedingungen (Differenzierbarkeit, Anforderungsvariation), um eine

individuelle Beanspruchungsfunktion abzuleiten. Die Anforderungshöhe hat einen großen Einfluss, insbesondere in der individuellen Parameterklassifikation, weshalb sich durch die geplante Anpassung der Schwierigkeit in der arithmetischen Nebenaufgabe positive Veränderungen hinsichtlich der Zuverlässigkeit der Prüfung auf Anforderungsabhängigkeit ergeben können.

6.3 Diskussion der Studie 2

Das Ziel der zweiten Studie war es zum einen, die Operationalisierung der entwickelten Kalibrierungsaufgabe (Kapitel 5) empirisch zu untersuchen und zum anderen, die beiden Ansätze zur individuellen Auswertung und Modellierung der physiologischen Reaktionen (Kapitel 5.3) zu vergleichen. Als wichtige Voraussetzungen für die Ableitung einer individuellen Beanspruchungsfunktion, basierend auf der Ausführung der Kalibrierungsaufgabe, sollten insbesondere die verschiedenen Schwierigkeitsstufen hinsichtlich der Interaktionseffekte zwischen Haupt- und Nebenaufgaben, der Anforderungshöhe und deren Differenzierbarkeit analysiert werden. Sowohl der NASA-TLX als auch die Leistungsparameter in Haupt- und Nebenaufgaben zeigen eine sehr gute Differenzierbarkeit der Anforderungsniveaus. Dabei ist der Effekt bezüglich der Leistung in den Hauptaufgaben deutlich stärker als bezüglich der Nebenaufgaben. Dies bestätigt die Umsetzung der Kalibrierungsaufgabe nach dem Subsidiary Task Paradigm. Die Leistung in Haupt- und Nebenaufgabe zeigt darüber hinaus auch, dass die Anforderungshöhe innerhalb des optimalen Bereichs zur Erfassung der mentalen Beanspruchung nach De Waard (1996) liegt (Abbildung 6.3 und 6.4). Lediglich die arithmetische Nebenaufgabe ist mit einer mittleren Fehlerhäufigkeit von $M = 1,05$ ($SD = 1,21$) in Kombination mit zwei Kugeln deutlich zu leicht, sodass das Anforderungsniveau für die nächste Studie angehoben werden soll. In den Bedingungen ohne Hauptaufgaben zeigt sich eine signifikant höher empfundene Beanspruchung bei der arithmetischen Aufgabe. Zusammen mit der tendenziell höher empfundenen Frustration kann dies auf den Effekt der sozialen Erwünschtheit hinweisen (Bortz & Schuster, 2010). Aufgrund der insgesamt eher niedrig bewerteten Frustration ($M = 6,34$, $SD = 4,71$) wurde

der Effekt in dieser Studie nicht weiter berücksichtigt. Der Effekt sollte aber, insbesondere vor dem Hintergrund der zu leichten arithmetischen Nebenaufgabe und der geplanten Anpassung auf mittelschwere Subtraktionsaufgaben, in der nächsten Studie geprüft werden. Die Umsetzung der Anforderungsniveaus der Hauptaufgabe durch Hinzunahme einer zweiten, simultanen Aufgabe hat einen erheblichen Einfluss auf die Sensitivität der bewussten physiologischen Parameter. Die hohe Sensitivität hinsichtlich der Hauptaufgabe ist zu großen Anteilen auf das Aufgabendesign zurückzuführen, welches das Blickverhalten deutlich beeinflusst (Abbildung 6.5). Hier liegt ein wichtiger Ansatz für weitere Forschungen hinsichtlich weiterer möglicher Kalibrierungsaufgaben: die zusätzliche Variation der Anforderung ohne Nebenaufgaben, beispielsweise durch Variation der Kugelgeschwindigkeit. Der hohe Einfluss der Aufgabengestaltung wirkt sich auf der individuellen Ebene jedoch nicht mehr so stark aus. Im Falle der individuellen Regressionsfunktionen wird beispielsweise der NNI bei $n = 28$ Probanden als Prädiktor ausgewählt. Bei der individuellen Parameterklassifikation, die eine mögliche Aufgabenabhängigkeit prüft, wird der NNI lediglich für $n = 12$ Probanden als Prädiktor ausgewählt, wobei sich der NNI für diese 12 Probanden auch nicht signifikant zwischen den Anforderungsniveaus der Hauptaufgabe unterscheidet. Für die restliche Stichprobe wird der NNI als aufgabenabhängig klassifiziert und damit ausgeschlossen. Dieses Ergebnis deckt sich mit den Befunden auf der Stichprobenebene und zeigt, dass die verwendete Methode der Procrustes-Analyse durchaus in Lage ist, die Abhängigkeiten eines Parameters hinsichtlich der Aufgabe, der Anforderungshöhe oder der Aufgabendauer zu bewerten. Insgesamt liefern aber beide Verfahren eine sehr hohe Diagnostizität, wobei sich die ausgewählten Parameter doch sehr stark unterscheiden. Nur für $n = 10$ Probanden erfüllt einer der beiden höchst korrelierenden Parameter auch die Anforderungen aus der individuellen Parameterklassifikation. Die Vorhersagekraft der individuellen Regressionsfunktionen für eine weitere Aufgabe muss daher weiterhin in Frage gestellt werden.

6.4 Zusammenfassung der Studie 2

Die zweite Studie hat gezeigt, dass die entwickelte Kalibrierungsaufgabe mit gut differenzierbaren Anforderungsniveaus sehr gut für die Herleitung einer individuellen Beanspruchungsfunktion geeignet ist. Die Aufgabe erfüllt alle überprüften Anforderungen, auch hinsichtlich der Gesamtbeanspruchung der Probanden und der gesetzten zeitlichen Restriktionen durch die geplante Einsatzfähigkeit vor Felduntersuchungen. Lediglich die Anforderungen in der arithmetischen Nebenaufgabe müssen aufgrund der sehr niedrigen Fehlerhäufigkeiten auf mittelschwere Subtraktionsaufgaben angehoben werden. Mit Blick auf die nächste Studie sollte hierbei allerdings die Entwicklung der empfundenen Frustration, vor dem Hintergrund des Effekts der sozialen Erwünschtheit, beobachtet werden. Die Auswertung der physiologischen Parameter auf genereller Ebene hat gezeigt, dass die bewussten Parameter stark durch das Aufgabendesign der Hauptaufgabe beeinflusst werden. Durch das Verfahren zur individuellen Parameterklassifikation und die darin enthaltene Prüfung auf Aufgabenabhängigkeit eines Parameters wurde der NNI beispielsweise bei $n = 27$ Probanden als aufgabenabhängig klassifiziert. Für die $n = 12$ Probanden für die der NNI als nicht aufgabenabhängig klassifiziert wurde, ist die Sensitivität gegenüber der Hauptaufgabe auch deutlich geringer und die Unterschiede sind nicht signifikant. Bei dem Verfahren der individuellen Regressionsfunktion wird der NNI allerdings für $n = 28$ Probanden als Prädiktor ausgewählt. Dieses Ergebnis ist sehr wichtig für die Auswertung der nächsten Studie, in der die Kalibrierungsaufgabe zur Vorhersage anwendungsnaher Aufgaben verwendet werden soll. Die Überlegungen aus Kapitel 5, dass die Varianzaufklärung als alleiniges Auswahlkriterium für zuverlässige physiologische Parameter nicht ausreicht, werden bestätigt und in der nächsten Studie explizit getestet. Insgesamt zeigen beide Verfahren zur individuellen Auswertung eine erhebliche Steigerung der Diagnostizität im Vergleich zu der Auswertung auf Stichprobenebene, wodurch die individuelle Vorgehensweise nochmals motiviert wird.

7 Studie 3: Vorhersage der Beanspruchung in anwendungsnahen Aufgaben

Nachdem die Operationalisierung der in Kapitel 5 entwickelten Kalibrierungsaufgabe in der zweiten Studie (Kapitel 6) erfolgreich untersucht wurde, wird die individuelle Kalibrierung im Rahmen dieser Studie zur Vorhersage der mentalen Beanspruchung in zwei anwendungsnahen Aufgaben verwendet. Dabei wird lediglich die Anforderungshöhe der mentalen Arithmetik auf mittelschwere Subtraktionsaufgaben angepasst. Die Auswertung der physiologischen Parameter auf individueller Ebene in der vorherigen Studie hat gezeigt, dass beide Auswertungsverfahren (Kapitel 5.3) eine erhebliche Steigerung der Diagnostizität im Vergleich zu der Auswertung auf genereller Ebene bewirken. Weiterhin wurde herausgestellt, dass es für jeden Probanden einen spezifischen physiologischen Indikator gibt, der die mentale Beanspruchung zuverlässig beschreiben kann. Die bisherigen Ergebnisse bestätigen damit die theoretischen Vorüberlegungen über eine stark individuelle, bewusst und unbewusst ablaufende, Informationsverarbeitung. Diese wirkt sich ebenso individuell auf die mentale Beanspruchung und damit auch auf die physiologischen Reaktionen des menschlichen Körpers aus. Diese Studie soll prüfen, ob diese Erkenntnisse und Ergebnisse auch praktisch anwendbar sind. Dafür wird die individuelle Beanspruchungsfunktion aus der Kalibrierungsaufgabe, die vor der eigentlichen Aufgabe ausgeführt wird, dazu verwendet, die entsprechende Referenzgröße (subjektives Beanspruchungsempfinden, Leistung in der Nebenaufgabe) in den

anwendungsnahen Aufgaben vorherzusagen. Es werden insgesamt zwei Aufgaben verwendet: eine Aufgabe besteht aus einer manuellen Tätigkeit der Qualitätssicherung, in der Leiterplatten auf ihre korrekte Montage geprüft werden sollen. Die andere Aufgabe umfasst eine Fahrsimulation aus der Transportlogistik, in der ein Lastkraftwagen (Lkw) durch verschiedene Szenarien gesteuert werden soll. Beide Aufgaben enthalten eine Nebenaufgabe, deren Leistungsdaten zum Abgleich mit der Beanspruchungsfunktion erfasst werden. Weiter füllt der Proband auch nach jedem Durchgang den NASA-TLX-Fragebogen aus. Zusätzlich werden die Effekte der Messwiederholung der Kalibrierungsaufgabe untersucht, weshalb die beiden genannten Aufgaben auf zwei Versuchstage verteilt werden. An beiden Tagen wird jeweils zu Beginn die Kalibrierung durchgeführt. Darüber hinaus sollen die deutlichen Unterschiede zwischen den zwei individuellen Auswertungsverfahren aus Studie 2 (Kapitel 6.2.2) vor dem Hintergrund der zuverlässigen Prädiktion der mentalen Beanspruchung in einer anderen Aufgabe untersucht werden. Die zentrale Fragestellung dieser Studie kann wie folgt umrissen werden: Ist eine zuverlässige Vorhersage der mentalen Beanspruchung in anwendungsnahen Aufgaben basierend auf den individuellen Beanspruchungsfunktionen aus der Kalibrierungsaufgabe möglich? Welche Auswirkungen hat die wiederholte Ausführung der Kalibrierungsaufgabe?

7.1 Experimentaldesign der Studie 3

In dieser Studie werden $N = 36$ Probanden, überwiegend Studenten, unter kontrollierten, aber anwendungsnahen Laborbedingungen untersucht. Die Probanden bearbeiten zwei Aufgaben (Qualitätssicherung, Transportlogistik) mit jeweils drei fest definierten Anforderungsniveaus an zwei Versuchstagen. Sowohl die Reihenfolge der Aufgaben, als auch die der Anforderungsniveaus werden randomisiert, um Reihenfolgeeffekte zu vermeiden. An beiden Versuchstagen führen die Probanden die Kalibrierungsaufgabe vor der jeweiligen Aufgabe durch, wobei auch hier die Reihenfolge der Nebenaufgabenbedingungen randomisiert wird. Die mentale Beanspruchung wird entsprechend dem trimodalen Ansatz (Kapitel 3.1), wie auch in den Studien 1 und 2, über subjektivi-

ve, leistungsbasierte und physiologische Parameter analysiert. Das subjektive Beanspruchungsempfinden wird über den NASA-TLX erfasst (Kapitel 3.1.1), welcher den Probanden nach jedem Anforderungsniveau gereicht wird. Für die Aufzeichnung der physiologischen, okularen Beanspruchungsparameter (Kapitel 3.1.3) tragen die Probanden während des gesamten Versuches ebenfalls das kopfbasierte Blickerfassungssystem Dikablis Professional (binokular, 60Hz) der Firma Ergoneers GmbH. Die leistungsbasierten Messungen unterscheiden sich zwischen den Aufgaben und werden folgend im Rahmen der detaillierten Aufgabenbeschreibungen näher erläutert. Zusätzlich werden die physiologischen Parameter auf der individuellen Ebene entsprechend der beiden entwickelten Auswertungsverfahren (Kapitel 5.3) ausgewertet und für die Vorhersage der Referenzinformationen in den beiden Aufgaben genutzt. Die Studie dauert je Versuchstag inklusive der Instruktionen und der Bearbeitung der Aufgaben zirka 60 min. Bevor die Probanden mit der Bearbeitung der Aufgaben beginnen, erhalten sie eine umfassende Einführung in die Messinstrumente und die Erklärung der Kalibrierungsaufgabe. Anschließend erfolgt die Erläuterung des jeweils direkt folgenden Versuchsteils durch den Versuchsleiter. Der Versuchsleiter führt die Instruktionen wieder nach einem ausgearbeiteten Skript durch, um Versuchsleitereffekte als Störvariable zu minimieren. Das Blickerfassungssystem wird vor jedem der insgesamt 14 Versuchsteile (beide Tage zusammen) neu kalibriert, um Messartefakte, die durch das Verrutschen des Systems entstehen können, zu verhindern. Nachfolgend sollen die anwendungsnahen Aufgaben und die verwendeten Nebenaufgaben detailliert beschrieben werden.

Aufgabe A1: Qualitätssicherung

Die erste Aufgabe besteht aus einer manuellen Tätigkeit in der Qualitätsprüfung von Leiterplatten. Die Aufgabe wird bewusst nicht an einem Computer bearbeitet, um die Vorhersagemöglichkeiten der Beanspruchungsfunktionen aus den Kalibrierungsaufgaben auch bei sehr dynamischen Tätigkeiten mit einer Vielzahl an Bewegungsartefakten zu untersuchen. Der Arbeitsplatz des Probanden ist dabei wie folgt aufgebaut (Abbildung 7.1): Vor dem Probanden

befindet sich der Arbeitsbereich einschließlich einer Musterleiterplatte, mit der die Leiterplatten aus einem Puffer auf der rechten Seite abgeglichen werden sollen. Sind alle Widerstände auf einer Leiterplatte korrekt montiert, werden diese auf dem Feld für Gutteile auf der linken Seite abgelegt. Ist die Leiterplatte fehlerhaft, so wird diese auf das Feld für Ausschuss gelegt. Jeder Widerstand auf der Leiterplatte besteht aus vier kleineren, farblich kodierten Widerständen, die in Reihe geschaltet sind. Ein Widerstand funktioniert nur dann korrekt, wenn die richtigen farblichen Widerstände in der richtigen Reihenfolge verbaut wurden.



Abb. 7.1: Versuchsaufbau Aufgabe A1: Qualitätssicherung (eigene Abbildung)

Sobald ein kleiner Widerstand nicht mit der Musterleiterplatte übereinstimmt, ist die Leiterplatte Ausschuss. Um in Hinblick auf die Vergleichbarkeit der Probanden sicherzustellen, dass jeder Proband die gesamte Leiterplatte kontrolliert und diese nicht bei dem ersten Fehler aussortiert, werden die Probanden gebeten ein Fehlerprotokoll auszufüllen. In diesem Fehlerprotokoll soll die Anzahl falscher Widerstände für jede inkorrekt montierte Leiterplatte dokumentiert werden. Den Probanden wird dazu erklärt, dass in naher Zukunft ein Qualitätsaudit stattfindet, weshalb ein genaues Protokoll der Fehlerzahlen benötigt wird. Weiter sind alle zwölf zu kontrollierenden Leiterplatten im Vergleich zu dem Muster gespiegelt. Erklärt wird diese Steigerung der Schwierigkeit mit den individuellen Kunden-

wünschen in dieser Charge. Die Taktzeit für die Kontrolle einer Leiterplatte ist auf 15 Sekunden festgelegt und wird den Probanden über einen akustischen Piepton mitgeteilt. Dabei wird diese Taktzeit aber lediglich als Referenzzeit und nicht als feste Grenze instruiert. Damit soll eine unvollständige Kontrolle der Leiterplatten aus zeitlichen Gründen verhindert werden, aber dennoch ein Zeitdruck aufgebaut werden. Die Taktzeit ist für alle drei Anforderungsniveaus gleich. Die Anforderungsniveaus definieren sich über die Anzahl der Widerstände auf der Leiterplatte: drei, vier oder fünf (Abbildung 7.2). Die Aufgabe endet, wenn die 12 x 15 Sekunden (3 Minuten) abgelaufen sind. Als Leistungsmaß in der Hauptaufgabe werden die falsch zugeordneten Leiterplatten und die mittlere Bearbeitungsdauer pro Leiterplatte ausgewertet. Zusätzlich zu dieser Hauptaufgabe werden die Probanden gebeten, akustisch präsentierte, mittelschwere Subtraktionsaufgaben verbal und so schnell als möglich mit dem richtigen Ergebnis zu beantworten. Diese arithmetische Nebenaufgabe wird mit einer Taktzeit von fünf Sekunden präsentiert. Die Aufgabe soll ebenfalls nach dem Subsidiary Task Paradigm bearbeitet werden, was durch die Taktzeit von 15 Sekunden in der Hauptaufgabe unterstützt wird. Die Probanden haben keine Zeit, den Fokus auf die Nebenaufgabe zu legen, da sie sonst aus dem Takt fallen. Die Fehler in der Nebenaufgabe werden als leistungsbasierter Beanspruchungsindikator erfasst. Die Nebenaufgabe bleibt über alle drei Versuchsteile gleich, die Anforderungsniveaus definieren sich allein über die Anzahl der Widerstände (Abbildung 7.2). Die einzelnen Anforderungsniveaus bearbeitet der Proband in randomisierter Reihenfolge.

Aufgabe A2: Transportlogistik

Die zweite Aufgabe umfasst Fahraufgaben aus dem Bereich der Transportlogistik mit einer computergestützten Lkw-Fahrsimulation (Euro Truck Simulator 2, SCS Software). Innerhalb der drei Anforderungsniveaus sollen die Probanden verschiedene Strecken mit dem Lkw zurücklegen. Der Lkw wird dem Personenkraftwagen (Pkw) in dieser Aufgabe vorgezogen, da kein Proband Erfahrungen mit dem Führen eines Lastkraftwagens hat und somit geringere

Assoziationen mit der realen Situation geweckt werden. Trotz der Verwendung eines Lenkrades und der entsprechenden Pedalerie, unterscheidet sich das Fahrverhalten der Simulation doch noch stark von dem realen Führen eines Fahrzeuges. Zu Beginn der Aufgabe bekommt der Proband die Möglichkeit, sich in einer Eingewöhnungsfahrt an die Fahrweise eines Lkw mit Auflieger zu gewöhnen. Alle drei Aufgabenteile werden mit dem gleichen Lkw inklusive Auflieger durchgeführt. Die Probanden werden instruiert mit oberster Priorität darauf zu achten, dass der Lkw unbeschädigt das Ziel erreicht. Innerhalb der vorgeschriebenen Geschwindigkeitsbegrenzungen soll allerdings auch so schnell als möglich gefahren werden, da die Ware dringend am Zielort benötigt wird. Durch diese beiden Instruktionen werden simultane Anforderungen an die Geschwindigkeit und die Präzision der Bearbeitung gestellt. Die drei Anforderungsniveaus definieren sich über vorgegebene Fahrtstrecken (Abbildung 7.2). In der leichten Bedingung wird der Proband gebeten, dem Navigationssystem folgend, aus einer Stadt heraus und auf einer gut ausgebauten Landstraße zu einem definierten Zielort zu fahren. Das Verkehrsaufkommen ist dabei sehr gering, es müssen zwei Lichtsignalanlagen überfahren werden und an einer Kreuzung einem anderen Verkehrsteilnehmer die Vorfahrt gewährt werden. Die Aufgabe wird nach fünf Minuten Fahrzeit beendet. In der mittleren Bedingung steht das Fahrzeug des Probanden zu Beginn auf einem Berg, der über eine Serpentinstraße hinab gefahren werden soll. Die Straße enthält eine Vielzahl enger Kurven, in denen insbesondere über die Rückspiegel auf das unbeschädigte Herumführen des Aufliegers geachtet werden soll. Die Aufgabe endet nach dem Überfahren einer Ziellinie im Tal. Das höchste Anforderungsniveau umfasst eine Fahrt durch ein Dorf mit sehr schmalen Straßen. Die Fahrt beginnt unmittelbar vor der Einfahrt in das Dorf und endet beim Überfahren einer Ziellinie nach dem Ortsausgang. Die Straße innerhalb der Ortsschilder ist durch mehrere sehr enge Kurven und Engstellen charakterisiert. Während der Durchführung ist der Proband gezwungen, mehrere Informationsfenster auf dem Bildschirm (Sicht aus dem Fenster und Sicht über den Rückspiegel) zu erfassen, um die Aufgabe korrekt ausführen zu können. In allen drei Aufgabenteilen werden die Anzahl der Fehler (Verstöße gegen die Straßenverkehrsordnung,

Kollisionen, Fahren auf der falschen Fahrbahn) und die Dauer der Fehler erfasst. Zusätzlich soll während aller drei Fahrten eine visuelle mentale Rotation als Nebenaufgabe auf einem Zusatzbildschirm bearbeitet werden (Kapitel 3.1.2). Die verwendeten Geometrien entsprechen mit acht Konturpunkten der leichten Kategorie (Abbildung 3.2). Die beiden zu vergleichenden Geometrien werden in einem zeitlichen Abstand von fünf Sekunden für je drei Sekunden angezeigt. Die Probanden müssen innerhalb der drei Sekunden entscheiden, ob die Geometrien gleich oder gespiegelt sind. Da der Lkw über ein automatisches Getriebe verfügt, werden die Eingaben für die mentale Rotationsaufgabe über zwei Schalter (Grün = gleich, Rot = gespiegelt) an der Position des Schalthebels getätigt. Eine zusätzliche Belastung durch eine motorische Aktion soll gering gehalten werden. Die Leistung in der Nebenaufgabe wird über die Anzahl der Fehler quantifiziert. In allen drei Versuchsteilen, deren Reihenfolge randomisiert wird, liegt die Priorität auf der Hauptaufgabe (Subsidiary Task Paradigm).

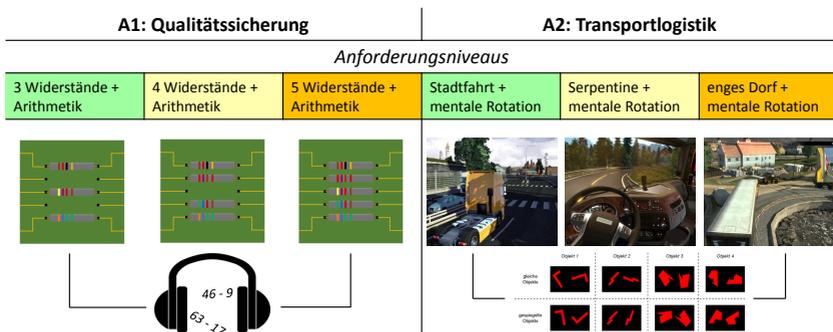


Abb. 7.2: Übersicht der Anforderungsniveaus der Aufgabe A1 (Qualitätssicherung) und A2 (Transportlogistik) (eigene Abbildung)

Die vorstehende Abbildung 7.2 fasst die Anforderungsniveaus beider Aufgaben zusammen.

7.2 Ergebnisse der Studie 3

Die Stichprobe umfasst $N = 36$ (26 männliche, 10 weibliche) Probanden, im Wesentlichen Studenten, die im Mittel 25,08 Jahre alt sind ($SD = 3,63$). Alle Probanden haben ein uneingeschränktes oder korrigiert-uneingeschränktes Sehvermögen (10 Brillenträger, 1 Kontaktlinsenträger). Im Mittel liegen 8,39 Tage ($SD = 6,63$) zwischen den beiden Versuchstagen der Probanden. Die Ergebnisse der Studie werden in drei Abschnitten beschrieben: Zunächst werden die Ergebnisse der Qualitätssicherung (Kapitel 7.2.1) und der Transportlogistik (Kapitel 7.2.2) hinsichtlich des subjektiven Beanspruchungsempfindens, der leistungs-basierten Daten und der physiologischen Parameter auf genereller Ebene betrachtet. Im Anschluss werden beide Kalibrierungsaufgaben analysiert und bezüglich der Lerneffekte und Wiederholungsgenauigkeit untersucht (Kapitel 7.2.3). Abschließend werden dann die individuellen Beanspruchungsfunktionen nach beiden Auswertungsverfahren (Kapitel 5.3) anhand der Kalibrierungsaufgaben entwickelt und zur Vorhersage der Beanspruchung in den beiden Anwendungsaufgaben genutzt.

7.2.1 A1: Qualitätssicherung

Das subjektive Beanspruchungsempfinden unterscheidet sich signifikant zwischen den drei Anforderungsniveaus ($F(2,70) = 20,32, p < 0,01$) der Qualitätssicherungsaufgabe (Abbildung 7.3). Die NASA-TLX Bewertungen für die Leiterplatten mit fünf Widerständen sind signifikant höher als für Leiterplatten mit drei ($t(35) = -5,63, p < 0,01$) und vier Widerständen ($t(35) = -4,39, p < 0,01$). Darüber hinaus zeigt auch das mittlere Anforderungsniveau eine signifikant höher empfundene Beanspruchung als die leichte Bedingung ($t(35) = -2,71, p < 0,01$).

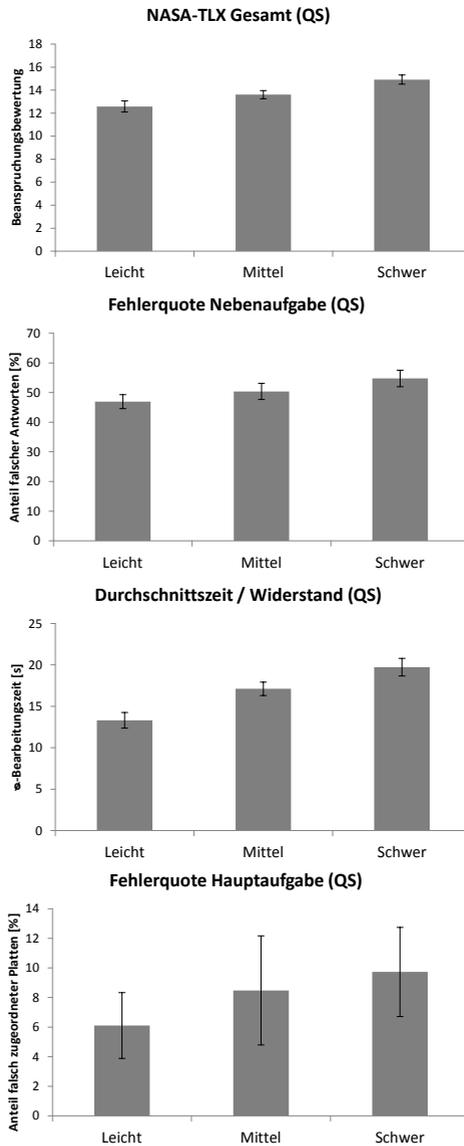


Abb. 7.3: Subjektive und leistungs-basierte Beanspruchungsindikatoren in der Qualitätssicherungsaufgabe (eigene Abbildung)

In der schweren Bedingung (fünf Widerstände) liegt die NASA-TLX Bewertung im Mittel bei $M = 14,92$ ($SD = 2,26$), dieses liegt allerdings im oberen Drittel, induziert aber noch keine Überforderung. Das Item „Anstrengung“ des NASA-TLX ist in der schweren Bedingung mit $M = 16,53$ ($SD = 2,43$) sehr hoch, zeigt aber keine Deckeneffekte. Zusammen mit der Frustration, die bei fünf zu kontrollierenden Widerständen bei $M = 12,36$ ($SD = 4,19$) im mittleren Bereich liegt, kann die Qualitätssicherungsaufgabe als sehr beanspruchend, aber nicht überfordernd bewertet werden. Insgesamt indizieren die Ergebnisse des NASA-TLX sehr gut differenzierbare Anforderungsniveaus. Die Fehlerquote in der Nebenaufgabe bestätigt die Ergebnisse des subjektiven Empfindens nur teilweise ($F(2, 70) = 5,76, p < 0,01$). Die Fehlerquote ist in der schweren Bedingung signifikant höher als in der leichten Bedingung ($t(35) = -3,50, p < 0,01$). Sowohl zwischen den Bedingungen mit drei und vier zu kontrollierenden Widerständen, als auch zwischen vier und fünf Widerständen sind ebenfalls Anstiege der Fehlerquote in der Nebenaufgabe zu erkennen, diese sind aber nicht signifikant (Abbildung 7.3). Dieses Ergebnis führt zu der Vermutung, dass das Subsidiary Task Paradigm in dieser Konstellation aus Haupt- und Nebenaufgabe nicht eindeutig von den Probanden umgesetzt wurde. Der geringe Anstieg der Fehlerquote in der Nebenaufgabe könnte zum einen auf das generell sehr hohe Anforderungsniveau, aber auch auf eine Fokussierung der Probanden auf die Nebenaufgabe zurückgeführt werden. Weiter fällt bei der Betrachtung der Daten auf, dass die Fehlerquote in der Hauptaufgabe in allen drei Bedingungen sehr niedrig ist und sich darüber hinaus nicht signifikant unterscheidet (Abbildung 7.3). Die richtige Zuordnung der Leiterplatten zu Gutteilen und Ausschuss scheint demnach ebenfalls eine sehr hohe Priorität für die Probanden gehabt zu haben. Lediglich die durchschnittliche Bearbeitungszeit pro Leiterplatte unterscheidet sich signifikant zwischen allen drei Anforderungsniveaus ($F(2, 70) = 30,12, p < 0,01$). Die Probanden akzeptieren möglicherweise eine längere Bearbeitungszeit, um die Zuordnung und die arithmetische Nebenaufgabe gut zu machen. Natürlich muss an dieser Stelle die soziale Erwünschtheit als begleitender Effekt berücksichtigt werden, aber eine detaillierte Analyse der Tätigkeit über Videoaufzeichnungen zeigt, dass die Probanden die Zeit, in der

die verglichene Leiterplatte abgelegt und eine neue aufgenommen wird nutzen, um die arithmetische Nebenaufgabe zu bearbeiten. Die Mehrheit der richtigen Antworten auf die Nebenaufgabe wurde genau in diesen Wechselzeiten gegeben. In der leichten Bedingung liegen $M = 64,11\%$ ($SD = 9,06$), in der mittleren $M = 65,54\%$ ($SD = 9,04$) und in der schweren Bedingung $M = 70,08\%$ ($SD = 10,65$) der richtigen Antworten in den Wechselphasen. Dabei ist der Anteil in der schweren Bedingung signifikant höher als in der leichten ($t(35) = -2,53$, $p < 0,05$). Der Effekt verstärkt sich demnach mit steigenden Anforderungen in der Hauptaufgabe, was den signifikanten Anstieg der durchschnittlichen Bearbeitungszeit erklärt. Die Nebenaufgabe ist durch diese Aufteilung sehr gut mit der Hauptaufgabe kombinierbar und lediglich die Bearbeitungszeit indiziert den Anforderungsanstieg. Der Prioritätenwechsel zwischen Haupt- und Nebenaufgabe im zeitlichen Verlauf der Aufgabe bestätigt die theoretischen Vorüberlegungen zu dem Vorhandensein von Strategiewechseln auch innerhalb eines Anforderungsniveaus (Kapitel 2.2.2). Der Prioritätenwechsel kann jedoch nicht eindeutig konkreten Zeitpunkten zugeordnet werden. Es kann durchaus zu Überlappungen der beiden Aufgaben kommen, weshalb zum einen die prozentualen Anteile der richtigen Antworten in den Wechselzeiten nicht noch höher sind und zum anderen keine signifikanten Interaktionseffekte zwischen der Haupt- und der Nebenaufgabe vorliegen. Der NASA-TLX korreliert sowohl signifikant mit der Leistung in der Nebenaufgabe ($r = 0,19$, $p < 0,05$), als auch mit der Leistung in der Hauptaufgabe ($r = 0,26$, $p < 0,01$). Am deutlichsten ist der Zusammenhang mit der durchschnittlichen Bearbeitungszeit ($r = 0,52$, $p < 0,01$). Diese Erkenntnisse motivieren weitere Forschungsarbeiten in dem Bereich der Strategie- oder Prioritätenwechsel. Durch eine echtzeitfähige Auswertung der okularen Parameter könnten die Zeitpunkte der Prioritätenwechsel mit Veränderungen in den physiologischen Reaktionen verglichen werden. Dadurch ließe sich der Bereich der Überlappungen in Zusammenhang mit der Aufmerksamkeit noch weiter beleuchten. Bezüglich der physiologischen Indikatoren zeigt kein Parameter signifikante Unterschiede zwischen den Anforderungsniveaus auf Stichprobenebene. Die geringe Sensitivität kann zu einem gewissen Anteil auf die individuell unterschiedlich ausgeprägten Prioritätenwechsel zwischen Haupt-

und Nebenaufgabe zurückgeführt werden. Die mittlere Anzahl der richtigen Antworten in Folge auf die arithmetische Nebenaufgabe unterscheidet sich nicht zwischen den Anforderungsniveaus, weist aber eine große individuelle Streuung auf, was ein weiterer Indikator für die hohe Individualität der Strategiewechsel ist. Bei der Betrachtung der Diagnostizität fällt auf, dass die Blinzelrate signifikante Zusammenhänge mit dem NASA-TLX ($r = -0,28, p < 0,01$) und der Leistung in der Nebenaufgabe ($r = -0,41, p < 0,01$) aufweist. Die Pupillendurchschnittsvariabilität hingegen zeigt signifikante Zusammenhänge mit dem NASA-TLX ($r = 0,19, p < 0,05$) und der Leistung in der Hauptaufgabe ($r = 0,18, p < 0,05$). Alle weiteren Parameter weisen keine signifikante Diagnostizität auf. Da beide genannten Parameter zu den unbewussten Indikatoren gehören, zeigen die Unterschiede in der Diagnostizität den starken individuellen Einfluss der Aufgabenart auf spezifische Parameter.

7.2.2 A2: Transportlogistik

Das subjektive Beanspruchungsempfinden unterscheidet sich auch signifikant zwischen den Anforderungsniveaus der Aufgabe aus der Transportlogistik ($F(2, 70) = 14,63, p < 0,01$). Die NASA-TLX Bewertungen bezüglich der Fahrt durch das Dorf (schwer) liegen signifikant höher als in der leichten Fahrt durch die Stadt ($t(35) = -6,12, p < 0,01$) und in der mittleren Fahrt die Serpentine hinunter ($t(35) = -3,27, p < 0,01$). Weiter zeigt auch die Serpentinestrecke im Vergleich zu der Stadtfahrt ein höheres subjektives Beanspruchungsempfinden ($t(35) = -4,82, p < 0,01$).

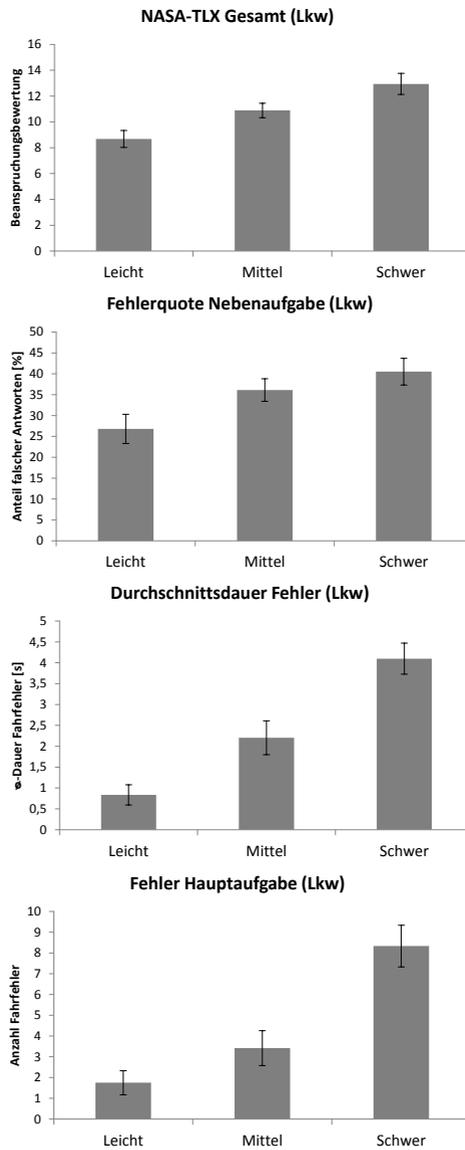


Abb. 7.4: Subjektive und leistungs-basierte Beanspruchungsindikatoren in der Lkw-Aufgabe (eigene Abbildung)

Die NASA-TLX Bewertungen liegen in der schweren Bedingung im Mittel bei $M = 12,94$ ($SD = 2,81$) und damit im mittleren Beanspruchungsbereich. Die Fehlerhäufigkeit ($F(2, 70) = 44,21, p < 0,01$) und die durchschnittliche Dauer der Fehler ($F(2, 70) = 57,26, p < 0,01$) in der Hauptaufgabe bestätigen die Ergebnisse des subjektiven Beanspruchungsempfindens (Abbildung 7.4). Beide Leistungsparameter zeigen ebenso signifikante Unterschiede zwischen jeweils allen drei Anforderungsniveaus. Die entsprechenden statistischen Kennwerte sind im Anhang (S. 215) dieser Arbeit zu finden. Die hoch signifikanten Korrelationen zwischen dem NASA-TLX und der Fehlerhäufigkeit ($r = 0,53, p < 0,01$) sowie der Durchschnittsdauer der Fehler ($r = 0,48, p < 0,01$) zeigen die hohe Bedeutung der Hauptaufgabe für das subjektive Empfinden verglichen mit der Nebenaufgabe ($r = 0,30, p < 0,01$). Das Subsidiary Task Paradigm wurde von den Probanden gut umgesetzt. Im Vergleich zu der Qualitätssicherungsaufgabe stellt die Fahraufgabe jedoch kontinuierlichere Anforderungen. Die Anforderungen schwanken zwar auch hier durch die Streckenführung, aber auch auf geraden Streckenabschnitten muss das Fahrzeug kontinuierlich durch, wenn auch kleine, Lenkimpulse geführt werden. Darüber hinaus sind die Interaktionseffekte der visuellen Nebenaufgabe mit den Fehlern ($r = 0,37, p < 0,01$) und der durchschnittlichen Dauer der Fehler ($r = 0,28, p < 0,01$) in der Hauptaufgabe signifikant. Verglichen mit der Qualitätssicherung konkurrieren hier die Ziele der Haupt- und Nebenaufgabe noch stärker, sodass ein Prioritätswechsel weg von der Hauptaufgabe nicht ohne Leistungseinbuße möglich war. Die Fehlerquote in der Nebenaufgabe steigt zwischen der leichten und mittleren Bedingung signifikant an ($t(35) = -3,40, p < 0,01$), wobei sich die mittlere und die schwere Bedingung nicht signifikant voneinander unterscheiden. Der starke Anstieg der Fehlerquote von der leichten zu der mittleren Bedingung kann mit den steigenden Anforderungen an die Kontrolle des Aufliegers über die Spiegel (Informationsaufnahme mehrerer Anzeigeelemente) erklärt werden. Durch die vermehrte Kontrolle der Rückspiegel steigt die Priorität dieser Informationsquellen stark an, sodass diese in das Blickverhalten integriert werden müssen. Die Veränderung des Blickverhaltens zusammen mit dem Anstieg an den zu verarbeitenden Informationen erfordert mehr Ressourcen zu Las-

ten der Leistung in der Nebenaufgabe. Durch den starken Interaktionseffekt zwischen Haupt- und Nebenaufgabe verstärkt sich diese Verschiebung der Ressourcen zwischen dem leichten und dem mittleren Anforderungsniveau. Der NNI bestätigt das veränderte Blickverhalten. Der NNI ist während der Stadtfahrt signifikant höher als während der Fahrt die Serpentine hinunter ($t(35) = 2,26, p < 0,05$) oder durch das enge Dorf ($t(35) = 2,18, p < 0,05$). Es fällt auf, dass sich der NNI antiproportional zu den anderen Beanspruchungsindikatoren (NASA-TLX, Leistung) verhält. Das Blickverhalten ist in den beiden höheren Anforderungsniveaus eher gruppiert und weniger zufällig. Die Spiegel werden offensichtlich in ein regelmäßiges Kontrollmuster des Blickverhaltens integriert. Dieser Effekt untermauert die hohe Aufgabenabhängigkeit des NNI auf der generellen Auswertungsebene, wie auch in Studie 2 (Kapitel 6.2.2) festgestellt wurde. Dies motiviert erneut die individuelle Auswertung, zumal alle anderen physiologischen Beanspruchungsindikatoren keine signifikanten Unterschiede zwischen den Anforderungsniveaus zeigen.

7.2.3 Vergleich der beiden Kalibrierungsaufgaben

In diesem Abschnitt werden beide Kalibrierungsaufgaben ausgewertet und miteinander verglichen. In der ersten Kalibrierungsaufgabe zeigen sich signifikante Unterschiede im subjektiven Beanspruchungsempfinden ($F(3, 105) = 75,95, p < 0,01$) und in der Fehlerquote der Nebenaufgaben ($F(3, 105) = 61,84, p < 0,01$) zwischen den vier Anforderungsniveaus (Abbildung 7.5). Die paarweisen, Bonferroni-korrigierten, t-Tests zeigen, dass sich jeweils alle Anforderungsniveaus signifikant hinsichtlich der beiden Beanspruchungsindikatoren unterscheiden. Lediglich die Hauptaufgabe mit einer Kugel und arithmetischer Nebenaufgabe unterscheidet sich nicht von der Hauptaufgabe mit zwei Kugeln und visueller Nebenaufgabe. Die entsprechenden statistischen Kennwerte sind im Anhang (S. 217) dieser Arbeit zu finden. Die ähnliche Sensitivität (Abbildung 7.5) wird durch den signifikanten Zusammenhang beider Indikatoren bestätigt ($r = 0,67, p < 0,01$). Die arithmetische Nebenaufgabenbedingung zeigt in beiden Beanspruchungsindikatoren höhere Anforderungen als die visuelle

Nebenaufgabenbedingung. In Studie 2 lagen beide noch auf einem ähnlichen Anforderungsniveau. Die Anforderungssteigerung in der arithmetischen Nebenaufgabe, durch die Verwendung von Subtraktionsaufgaben, zeigt demnach den gewünschten Effekt. Die Fehlerquote in Kombination mit zwei Kugeln liegt jetzt mit $M = 49,07\%$ ($SD = 19,29$) im mittleren, angestrebten Bereich.

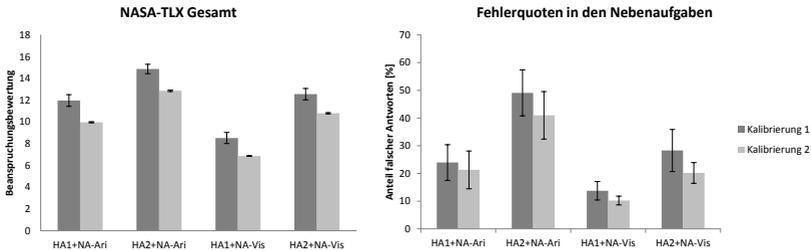


Abb. 7.5: Subjektive und leistungs-basierte Beanspruchungsindikatoren in den Kalibrierungsaufgaben 1 und 2: HA1 = Hauptaufgabe mit einer beweglichen Kugel, HA2 = Hauptaufgabe mit zwei beweglichen Kugeln, NA-Ari = arithmetische Nebenaufgabe, NA-Vis = visuelle Nebenaufgabe (eigene Abbildung)

Darüber hinaus konnte durch die Anpassung die Differenzierbarkeit der vier Anforderungsniveaus gesteigert werden, was die Prüfung der Aufgabenabhängigkeit von physiologischen Parametern durch den breiteren Anforderungsbereich verbessert. Wichtig ist an dieser Stelle die Betrachtung des Frustrationsitems im NASA-TLX. Hier muss sichergestellt werden, die Probanden durch die Ausführung der Kalibrierungsaufgabe nicht so starkem Stress auszusetzen, dass die eigentlich zu analysierende Aufgabe hierdurch beeinflusst wird. Über alle Versuchsteile der Kalibrierung hinweg ist das Frustrationsitem mit $M = 9,94$ ($SD = 4,68$) im Vergleich zu Studie 2 mit Additionsaufgaben (Kapitel 6.2.1) zwar leicht, aber nicht signifikant höher. Der Stressfaktor der Kalibrierungsaufgabe kann demnach, auch unter Verwendung der Subtraktionsaufgaben, für den geplanten Einsatzzweck vor Anwendungsaufgaben als angemessen betrachtet werden. Auch in der zweiten Kalibrierungsaufgabe liegen signifikante Unterschiede in der NASA-TLX Bewertung ($F(3, 105) = 61,00, p < 0,01$) und in der Fehlerquote der Nebenaufgaben ($F(3, 105) = 44,70, p < 0,01$) zwischen

den Anforderungsniveaus vor. Die paarweisen, Bonferroni-korrigierten t-Tests liefern exakt die gleichen Differenzierungen wie in der ersten Kalibrierung für beide Beanspruchungsindikatoren (Abbildung 7.5). Die entsprechenden statistischen Kennwerte sind im Anhang dieser Arbeit zu finden. Auch der hoch signifikante Zusammenhang zwischen dem NASA-TLX und der Fehlerquote in der Nebenaufgabe zeigt sich in der zweiten Kalibrierung ($r = 0,58, p < 0,01$). Die gute Differenzierbarkeit der Anforderungsniveaus in der Kalibrierungsaufgabe, entsprechend dieser Ergebnisse, bleibt auch bei einer wiederholten Ausführung der Kalibrierung erhalten. Allerdings zeigt sich ein Adaptionseffekt, der sich insbesondere auf die subjektiv empfundene Beanspruchung, aber auch die Fehlerquote in der Nebenaufgabe, auswirkt (Abbildung 7.5). Die NASA-TLX Bewertungen der zweiten Kalibrierung liegen sowohl in der ersten ($t(35) = 5,18, p < 0,01$) als auch in der zweiten Hauptaufgabe ($t(35) = 5,47, p < 0,01$) mit arithmetischer Nebenaufgabe, als auch in der ersten ($t(35) = 3,85, p < 0,01$) und zweiten Hauptaufgabe ($t(35) = 3,74, p < 0,01$) mit visueller Nebenaufgabe signifikant unter den Bewertungen der ersten Kalibrierung. Für die Fehlerquote in der Nebenaufgabe zeigt sich die gleiche Tendenz, wenn auch nicht signifikant. Bei der Betrachtung von Abbildung 7.5 fällt auf, dass die Streuung der NASA-TLX Bewertungen in den Anforderungsniveaus der zweiten Kalibrierung sehr klein ist. Im Vergleich mit der ersten Kalibrierung zeigt sich die Auswirkung von individuellen Erfahrungen und Konzepten. Während die Aufgaben bei der ersten Kalibrierung noch völlig unbekannt sind und jeder Proband individuelle Erfahrungen mitbringt, wissen die Probanden in der zweiten Kalibrierung wie die Aufgaben ablaufen. Die Erfahrungen und Konzepte bezüglich der Kalibrierungsaufgaben haben sich demnach aufgrund des erlebten, ersten Versuchstages angeglichen. Dieser Effekt wirkt sich jedoch nicht signifikant auf die Leistung aus, was die Umsetzung der Kalibrierungsaufgabe nochmals bestätigt. Die Leistungs- oder Adaptionkurve wird durch die hohe Dynamik der Hauptaufgaben in Kombination mit dem Subsidiary Task Paradigm anscheinend sehr flach gehalten. Die Probanden haben eine Vorkenntnis und wissen was sie erwartet, sodass die NASA-TLX-Bewertungen signifikant niedriger sind und deutlich weniger streuen. Die gezeigten Unterschiede zwischen der

ersten und zweiten Kalibrierung sind für den NASA-TLX nicht abhängig von der Zeit zwischen den Versuchen ($r = -0,16, n. s.$). Die Veränderungen in der Leistung zeigen hingegen eine signifikante Abhängigkeit von der Zeit zwischen den Versuchstagen ($r = -0,43, p < 0,01$). Je mehr Zeit zwischen den Versuchen liegt, desto geringer ist der Anstieg in der Leistung. Die Aussage bezüglich des Adaptionsverlaufs muss demnach korrigiert, beziehungsweise erweitert werden. Die Probanden entwickeln ihre Strategie der Aufgabenbearbeitung während der ersten Kalibrierung weiter. Die neue Strategie kann bei einer kürzeren Zeit zwischen den Versuchen sofort abgerufen werden, was zu einem Abfall der Fehlerquote in der Nebenaufgabe führt. Durch eine längere Zeit zwischen den Versuchen, ist die Strategie nicht mehr sofort abrufbar, sondern erfordert ein erneutes Erproben. Es ist demnach ein Adaptionseffekt vorhanden, der die Leistungskurve beeinflusst, aber durch die hohe Dynamik der Aufgabe verkürzt wird. Die Betrachtung der physiologischen Parameter auf Stichprobenebene zeigt eine sehr hohe Diagnostizität des NNI und der Sakkadenlängen in beiden Kalibrierungen (Tabelle 7.1). Alle anderen Parameter weisen keine signifikanten Zusammenhänge auf, wodurch die Ergebnisse der zweiten Studie bestätigt werden (Kapitel 6.2.2). Zwischen der ersten und zweiten Kalibrierung liegen keine signifikanten Unterschiede zwischen den beiden okularen Parametern (NNI, SK) vor. In der ersten Kalibrierung zeigt der NNI signifikante Unterschiede zwischen den Anforderungsniveaus ($F(3, 105) = 28,53, p < 0,01$). Allerdings liegt dabei lediglich eine Sensitivität gegenüber der Hauptaufgabe vor.

Tab. 7.1: Vergleich der Diagnostizität von NNI und Sakkadenlängen zwischen Kalibrierung 1 und 2: NA = Nebenaufgabe (eigene Tabelle)

	Kalibrierung 1		Kalibrierung 2	
	NASA-TLX	Fehlerquote NA	NASA-TLX	Fehlerquote NA
NNI	$r = 0,29,$ $p < 0,01$	$r = 0,26,$ $p < 0,01$	$r = 0,40,$ $p < 0,01$	$r = 0,27,$ $p < 0,01$
Sakkaden- länge	$r = 0,18,$ $p < 0,05$	$r = 0,23,$ $p < 0,01$	$r = 0,39,$ $p < 0,01$	$r = 0,29,$ $p < 0,01$

Der NNI ist in beiden Bedingungen mit zwei Kugeln sowohl mit arithmetischer ($t(35) = -6,25, p < 0,01$), als auch mit visueller Nebenaufgabe ($t(35) = -7,22, p < 0,01$) signifikant höher als mit einer Kugel. Zwischen den Nebenaufgabenbedingungen liegen keine Unterschiede vor. In der zweiten Kalibrierung liegen signifikante Unterschiede zwischen den Anforderungsniveaus sowohl für den NNI ($F(3, 105) = 45,78, p < 0,01$) als auch für die Sakkadenlängen ($F(3, 105) = 82,87, p < 0,01$) vor. Beide Parameter zeigen exakt die gleichen Sensitivitäten wie der NNI in Kalibrierung 1. Die statistischen Kennwerte sind im Anhang dieser Arbeit zu finden. Die Ergebnisse bestätigen die hohe Aufgabenabhängigkeit des NNI, aber auch der Sakkadenlängen, auf Stichprobenebene erneut.

7.2.4 Individuelle Beanspruchungsvorhersage

Das folgende Kapitel beinhaltet die Analyse der individuellen Beanspruchungsvorhersage durch die in Kapitel 5.3 entwickelten Auswertungsverfahren. Dabei stehen vier Fragen im Fokus: Unterscheiden sich die ausgewählten Parameter für die individuellen Beanspruchungsfunktionen ähnlich stark wie in Studie 2? Ist eine Prognose der Beanspruchung aufgabenunabhängig möglich? Ist vor jeder zu analysierenden Aufgabe eine Kalibrierung für einen Probanden notwendig oder reicht eine einmalige Kalibrierung aus? Wie hoch ist die Vorhersagegenauigkeit beider Verfahren? Um diese Fragen zu beantworten, wird zunächst das Verfahren der individuellen Regressionen (Kapitel 5.3.1) betrachtet (Abbildung 7.6).

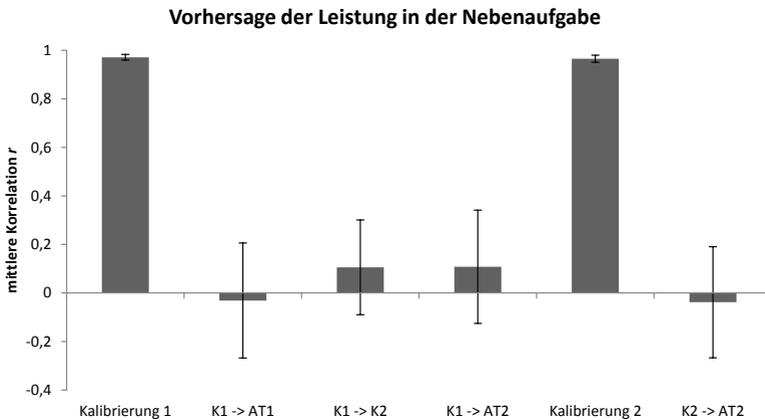


Abb. 7.6: Ergebnisse des Verfahrens individueller Regressionen: $K1$ = Kalibrierung 1, $K2$ = Kalibrierung 2, $AT1$ = Anwendungsaufgabe an Tag 1, $AT2$ = Anwendungsaufgabe an Tag 2 (eigene Abbildung)

In Abbildung 7.6 sind die transformierten (Fischers' z-Transformation), über alle Probanden gemittelten und schließlich rücktransformierten Korrelationskoeffizienten zwischen den individuellen Regressionsfunktionen und der Referenzgröße (Fehlerquote in der Nebenaufgabe) aufgetragen. Bei „Kalibrierung 1“ ($r_Z = 0,97$, $\sigma_Z = 0,04$, $p < 0,01$) und „Kalibrierung 2“ ($r_Z = 0,97$, $\sigma_Z = 0,05$, $p < 0,01$) sind dabei die mittleren Korrelationen innerhalb der beiden Kalibrierungsaufgaben dargestellt, aus denen die Regressionsfunktionen gebildet wurden. Die Ergebnisse der zweiten Studie werden demnach bestätigt, denn in beiden Kalibrierungsaufgaben existieren für jeden Probanden individuelle, physiologische Parameter, die in der Lage sind, die mentale Beanspruchung zuverlässig zu beschreiben. Werden die individuellen Beanspruchungsfunktionen aus der ersten Kalibrierungsaufgabe aber genutzt, um die direkt im Anschluss bearbeiteten Anwendungsaufgaben (QS, Lkw) des ersten Versuchstages vorherzusagen (K1 -> AT1), dann zeigt sich lediglich eine Korrelation nahe Null (Abbildung 7.6). Die Vorhersage einer direkt folgenden Anwendungsaufgabe ist unabhängig von dem Inhalt der Aufgabe nicht möglich. Auch eine Aufteilung nach Qualitätssicherungsaufgabe und Lkw-Aufgabe steigert die Diagnostizität

nicht. Es gibt insgesamt 18 Probanden, für die es mit der Regressionsfunktion möglich ist, die Leistung vorherzusagen. Bei den weiteren 18 Probanden ist die Korrelation nahe Null oder hoch signifikant, aber genau antiproportional. Für einen Probanden, für den beispielsweise in der Kalibrierungsaufgabe gilt: je höher die Beanspruchung, desto höher die Blinzelrate, für diesen gilt in der Anwendungsaufgabe: je höher die Beanspruchung, desto niedriger die Blinzelrate. Dieses Verhalten der individuellen physiologischen Parameter wird bei 16 Probanden beobachtet. Das Ergebnis bestärkt die Erkenntnisse aus der zweiten Studie, dass die Varianzaufklärung als alleiniges Kriterium zur Parameterauswahl nicht ausreichend ist. Allerdings reicht es nicht aus, allein die Aufgabenabhängigkeit als Grund für die schlechte Vorhersagegenauigkeit zu nennen. Auch die Vorhersage der Leistung in der zweiten Kalibrierungsaufgabe durch die Beanspruchungsfunktion der ersten Kalibrierung ($K1 \rightarrow K2$) ist ebenfalls nicht möglich, obwohl die Aufgaben exakt die gleichen sind. Die Vielzahl der individuellen Einflussfaktoren auf die menschliche Informationsverarbeitung und damit auch auf die mentale Beanspruchung aus den theoretischen Vorüberlegungen wird durch dieses Ergebnis nochmals bestätigt. Weiter ist die Vorhersage der Beanspruchung in den Anwendungsaufgaben am zweiten Versuchstag weder durch die erste, noch durch die zweite Kalibrierung möglich. Vor der detaillierten Analyse der Parameter und dem Vergleich mit dem Verfahren der Parameterklassifikation kann zusammenfassend festgehalten werden, dass durch die beiden Kalibrierungsaufgaben individuell hoch zuverlässige, physiologische Beanspruchungsindikatoren extrahiert werden können, die allerdings nicht für eine Vorhersage der mentalen Beanspruchung geeignet sind. Die Ergebnisse für das Verfahren der individuellen Parameterklassifikation (Kapitel 5.3.2) sind in Abbildung 7.7 zusammengefasst.

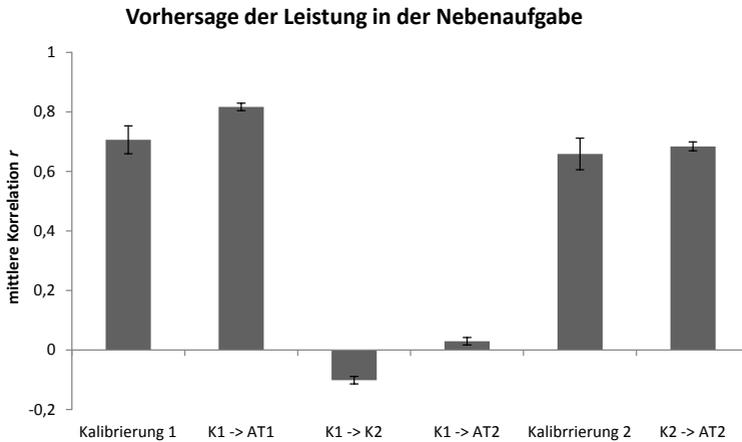


Abb. 7.7: Ergebnisse des Verfahrens zur individuellen Parameterklassifikation: *K1* = Kalibrierung 1, *K2* = Kalibrierung 2, *AT1* = Anwendungsaufgabe an Tag 1, *AT2* = Anwendungsaufgabe an Tag 2 (eigene Abbildung)

Auch mit dem Verfahren der individuellen Parameterklassifikation können sowohl mit der ersten ($r_Z = 0,71$, $\sigma_Z = 0,14$, $p < 0,01$) als auch mit der zweiten Kalibrierung ($r_Z = 0,66$, $\sigma_Z = 0,16$, $p < 0,01$) individuelle Parameter ausgewählt werden, welche die mentale Beanspruchung zuverlässig beschreiben. Es fällt auf, dass mit der ersten Kalibrierung die folgende Anwendungsaufgabe des ersten Tages (K1 -> AT1) hoch zuverlässig ($r_Z = 0,90$, $\sigma_Z = 0,19$, $p < 0,01$) vorhergesagt werden kann. Dabei werden 32 Probanden hoch zuverlässig hinsichtlich der Leistung in der Nebenaufgabe vorhergesagt. Für vier Probanden existiert kein Parameter, der die definierten Anforderungen erfüllt. Dies wird dem Versuchsleiter direkt nach der Kalibrierung zurückgemeldet, sodass eine fehlerhafte Interpretation der zu analysierenden Aufgabe vermieden wird. Diese vier Probanden werden durch das Verfahren der individuellen Regressionen fehlerhaft vorhergesagt. Darüber hinaus können die Anwendungsaufgaben am zweiten Versuchstag hoch zuverlässig durch die individuellen Parameter aus der zweiten Kalibrierung (K2 -> AT2) vorhergesagt werden ($r_Z = 0,68$, $\sigma_Z = 0,04$, $p < 0,01$). Durch Verwendung des Verfahrens der individuellen Parameterklas-

sifikation, basierend auf der Procrustes-Analyse, kann die Beanspruchung in einer direkt folgenden Anwendungsaufgabe hoch zuverlässig vorhergesagt und quantifiziert werden. Eine zuverlässige Vorhersage der Beanspruchung in der zweiten Kalibrierungsaufgabe als auch in den zweiten Anwendungsaufgaben (zweiter Versuchstag) durch die erste Kalibrierung ist nicht zuverlässig möglich. Es fällt auf, dass die Diagnostizität in der Vorhersage der mentalen Beanspruchung durch die zweite Kalibrierung (K2 -> AT2) eine geringere Diagnostizität aufweist, als die Vorhersagen durch Kalibrierung 1 (K1 -> AT1). Basierend auf der zweiten Kalibrierungsaufgabe konnten nur für 23 Probanden physiologische Parameter gefunden werden, die die definierten Anforderungen erfüllen. Für weitere acht Probanden konnte kein Modell generiert werden und bei insgesamt fünf Probanden funktioniert die Vorhersage mit dem ausgewählten Parameter nicht zuverlässig. Hier verändert sich die Proportionalität des Parameters zu der Referenzgröße zwischen der Kalibrierung und der Anwendungsaufgabe. Folgend werden diese fünf Probanden näher betrachtet, um zu verstehen, warum die Parameterklassifikation nicht zuverlässig anwendbar ist. Im ersten Schritt werden die Tage zwischen den Versuchen analysiert (Abbildung 7.8 links), da der Vergleich der Kalibrierungen bereits einen Einfluss auf die Leistung gezeigt hat (Kapitel 7.2.3). Die Analyse der Daten zeigt, dass die Dauer zwischen den beiden Versuchstagen bei den fünf Probanden signifikant kürzer ist ($t(14,15) = 3,30, p < 0,01$) als bei der restlichen Stichprobe. Dieses Ergebnis wird durch den signifikanten Zusammenhang zwischen der Vorhersagegenauigkeit in der zweiten Kalibrierung und der Zeit zwischen beiden Versuchsteilen für diese fünf Probanden bestätigt ($r = 0,57, p < 0,01$). Auf Ebene der Stichprobe liegt dieser Effekt nicht vor ($r = 0,09, n. s.$).

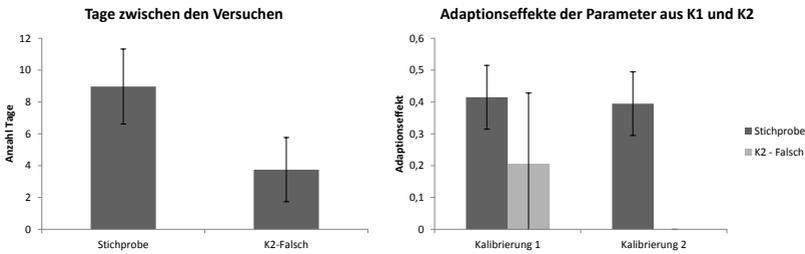


Abb. 7.8: Analyse der Probanden ohne zuverlässige Vorhersage (eigene Abbildung)

Die in Kapitel 7.2.3 aufgestellte Vermutung über eine zeitlich beschränkte Adaption der Leistungskurve kann durch die genannten Ergebnisse verifiziert werden. Die Adaptionseffekte (Kapitel 5.3.2), die sich auf die physiologischen Beanspruchungsindikatoren auswirken, unterscheiden sich auf Ebene der Stichprobe nicht signifikant zwischen den beiden Kalibrierungen (Abbildung 7.8 rechts). Für die fünf Probanden, die durch die zweite Kalibrierung nicht zuverlässig vorhergesagt werden können (K2 -> AT2), liegen in der der zweiten Kalibrierung auch signifikant kleinere Adaptionseffekte vor ($t(4) = 1,66, p < 0,05$). Das proaktive und reaktive Verhalten der fünf Probanden ist in der zweiten Kalibrierung signifikant gemindert, sie adaptieren somit weniger ihre Strategie als noch in Kalibrierung K1. Überdies können die fünf Probanden hoch zuverlässig mit den ausgewählten Parametern der ersten Kalibrierung in den zweiten Anwendungsaufgaben vorhergesagt werden ($r_Z = 0,70, \sigma_Z = 0,13, p < 0,01$). Aus der vorstehenden Auswertung kann folglich die Hypothese abgeleitet werden, dass ein Proband die Kalibrierungsaufgabe nur wiederholen soll, wenn eine gewisse Zeit seit der letzten Kalibrierung vergangen ist. Wie genau dieses Zeitfenster definiert ist, kann mit den erhobenen Daten nicht quantifiziert werden. An dieser Stelle bietet sich ein guter Anknüpfungspunkt für weitere Forschungsarbeiten, um die Leistungskurve durch die Kalibrierungsaufgabe noch weiter zu detaillieren und eine klare Empfehlung für die Praxis abzuleiten. Da anwendungsorientierte Studien nur selten in vergleichbar kurzen Zeitintervallen mit den gleichen Probanden durchgeführt werden, wird dieser Punkt hier nicht weiter beleuchtet. Nachfolgend werden an dieser Stelle die Unterschiede beider

Verfahren noch näher betrachtet. Für beide Verfahren wird die Pupillendurchschnittsvariabilität sehr häufig für die individuelle Beanspruchungsfunktion ausgewählt. Für das Verfahren der individuellen Regressionen wird allerdings der NNI am häufigsten in beiden Kalibrierungen ausgewählt, was einen großen Unterschied zwischen den beiden Verfahren darstellt. Auch in der zweiten Studie ist dieser Unterschied durch die große Aufgabenabhängigkeit des NNI bereits aufgefallen. Innerhalb des Verfahrens der individuellen Parameterklassifikation werden die unbewussten Parameter in der zweiten Kalibrierung für $n = 19$ Probanden (67,86%) ausgewählt. Dies kann zu einem gewissen Teil auf die deutlich steigende Diagnostizität der bewussten Parameter (insbesondere NNI und Sakkadenlängen) von Kalibrierung K1 zu K2 auf der Stichprobenebene zurückgeführt werden. Die bewussten Parameter unterliegen für einige Probanden einer geringen Aufgabenabhängigkeit, die durch die Kriterien der Parameterklassifikation ausgeschlossen wird. In der ersten Kalibrierung erfüllt nur für $n = 11$ Probanden mindestens ein ausgewählter, höchst korrelierender Parameter auch die Anforderungen der Parameterklassifikation. In der zweiten Kalibrierung sinkt die Anzahl dieser Probanden auf $n = 5$. Für die individuellen Regressionen bleibt bei $n = 15$ Probanden mindestens ein physiologischer Parameter in der Regression aus K2 gleich im Vergleich zu K1. Bei der Parameterklassifikation bleibt nur bei $n = 2$ Probanden der physiologische Prädiktor konstant. Dieses Ergebnis verdeutlicht, dass neben der Aufgabenabhängigkeit auch weitere Einflussfaktoren eine wichtige Rolle für die zuverlässige Beanspruchungsmessung spielen. Insgesamt konnte durch die Kriterien der Parameterklassifikation eine hoch zuverlässige, individuelle Beanspruchungsmessung entwickelt werden, die für den Einsatz in anwendungsorientierten Aufgaben geeignet ist.

7.3 Diskussion der Studie 3

Im Rahmen der dritten Studie wurde die mentale Beanspruchung auf individueller Ebene in zwei Anwendungsaufgaben an zwei Versuchstagen, ausgehend von den jeweils direkt davor ausgeführten Kalibrierungsaufgaben, vorhergesagt. Eine Aufgabe der Transportlogistik (Lkw-Simulator) und eine Qualitätssicherungsaufgabe bildeten die Anwendungsaufgaben. Die individuellen Beanspruchungsfunktionen wurden durch die beiden Auswertungsverfahren nach Kapitel 5.3 berechnet und hinsichtlich der Zuverlässigkeit der Vorhersage verglichen. Die Ergebnisse zeigen, dass eine Vorhersage der mentalen Beanspruchung in einer Anwendungsaufgabe durch den Ansatz der individuellen Regressionen nicht möglich ist (Abbildung 7.6). Die ausgewählten individuellen Beanspruchungsindikatoren verändern ihr Verhalten sowohl bei der Bearbeitung anderer Aufgaben, als auch bei dem wiederholten Ausführen der Kalibrierungsaufgabe sehr stark, bis hin zu einer völligen Umkehr des Zusammenhangs mit der Referenzgröße. Die Vermutung, dass die Varianzaufklärung als alleiniges Auswahlkriterium für die physiologischen Parameter unzureichend ist, wurde durch diese Ergebnisse bestätigt, da die physiologischen Reaktionen des Menschen neben der Aufgabe auch von den individuellen Zielen, Erfahrungen und der Motivation abhängig sind. Das Verfahren der individuellen Parameterklassifikation nimmt eine Auswahl der Parameter anhand einer Vielzahl an Kriterien vor. Die Auswertung dieser Studie zeigt, dass die Vorhersage einer anwendungsnahen Aufgabe durch einen physiologischen Parameter, der die definierten Kriterien erfüllt, sehr zuverlässig möglich ist (Abbildung 7.7). Allerdings muss hier eine Einschränkung vorgenommen werden: Die Vorhersage der Aufgaben des zweiten Versuchstages durch die erste Kalibrierung ist nicht möglich. Darüber hinaus konnten auch fünf Probanden nicht durch die zweite Kalibrierung in der zweiten Anwendungsaufgabe vorhergesagt werden. Bei diesen fünf Probanden lagen signifikant weniger Tage zwischen den Versuchsteilen, als bei der restlichen Stichprobe, was sich auch deutlich negativ auf die Stärke der physiologischen Adaptionseffekte (Strategiewechsel) auswirkte. Es gibt folglich eine Adaption der Leistungskurve in der Kalibrierungsaufgabe, die scheinbar mit zunehmender

Zeit nach der Durchführung wieder sehr schnell abnimmt. Gründe dafür können in der hohen Dynamik der Hauptaufgabe, aber auch in der kurzen Ausführungszeit der Kalibrierungsaufgabe liegen (Kahneman, 1973). Die Ergebnisse zeigen demnach einen Einfluss der Zeit zwischen den Versuchstagen, der in weiteren Studien genauer untersucht werden muss. Insbesondere in Hinblick auf die Verwendung der Kalibrierungsaufgabe zur Analyse tatsächlicher Arbeitssysteme, in denen die zu betrachtende Aufgabe den Probanden sehr gut bekannt ist, ergibt sich ein weiterer wichtiger Diskussionspunkt. In dieser Studie unterscheiden sich die Versuchstage deutlich, da an dem ersten Tag sowohl die Kalibrierung, als auch die Anwendungsaufgabe unbekannt sind. Am zweiten Versuchstag ist die Kalibrierungsaufgabe jedoch bekannt, die Anwendungsaufgabe jedoch wieder unbekannt. In der Analyse tatsächlicher Arbeitssysteme kehrt sich das Verhältnis um: die Kalibrierung ist in der Regel unbekannt, aber die Anwendungsaufgabe ist sehr gut bekannt. In Kombination mit den Ergebnissen der Vorhersage ergibt sich hier ein kritischer Punkt, der mit Blick auf eine praktische Anwendung in weiteren Studien analysiert werden muss. Die Verbesserung der Leistung und die Verminderung des subjektiven Beanspruchungsempfindens in der Wiederholung der Kalibrierungsaufgabe bestätigen die vorstehenden Überlegungen (Abbildung 7.5). Die guten Vorhersageergebnisse der Parameterklassifikation lassen allerdings vermuten, dass es zwar einen Einfluss der aktuellen Verfassung des Probanden gibt, eine Vorhersage einer bekannten Aufgabe durch eine unbekannte Kalibrierung aber möglich ist. Durch die Prüfung der Aufgabenabhängigkeit hat die Art der Aufgabe keinen Einfluss auf den ausgewählten physiologischen Parameter. In Kombination mit der Betrachtung des Drifts (Kapitel 5.3.2) wird auch eine mögliche Abnahme der Beanspruchung, beispielsweise durch die Bekanntheit einer Aufgabe, in der Parametersauswahl berücksichtigt. Ein weiterer wichtiger Diskussionspunkt dieser Studie betrifft die Aufgabe der Qualitätssicherung. In der Auswertung der subjektiven und leistungs-basierten Beanspruchungsindikatoren zeigten sich temporäre Strategiewechsel über die Bearbeitungszeit hinweg (Kapitel 7.2.1). Beide Indikatoren indizieren Strategiewechsel ausgehend von dem aktuellen Inhalt der Hauptaufgabe. Die Probanden nutzten die Wechselzeiten zwischen den Vergleichen

der Leiterplatte, um verstärkt die arithmetische Nebenaufgabe zu bearbeiten. Allerdings konnten die genauen Zeitpunkte der Strategie- oder Prioritätswechsel, aufgrund der geringen zeitlichen Auflösung der Leistungsparameter, nicht identifiziert werden. Die Strategiewechsel sind ein wichtiger Bestandteil und Einflussfaktor für die individuelle Bildung einer Beanspruchungsfunktion. An dieser Stelle sind weitere Untersuchungen, auch insbesondere der physiologischen Parameter, notwendig, um die Übergangsphasen besser zu verstehen und die Ergebnisse in der Auswertung der Kalibrierungsaufgaben berücksichtigen zu können.

7.4 Zusammenfassung der Studie 3

Der Schwerpunkt dieser Studie liegt auf der Untersuchung der Möglichkeit, die mentale Beanspruchung, operationalisiert über eine wählbare Referenzgröße (subjektives Empfinden, Leistung, mentale Effizienz), zuverlässig in zwei anwendungsnahen Aufgaben (Qualitätssicherung, Transportlogistik) vorherzusagen. Basierend auf dem Ausführen einer Kalibrierungsaufgabe wurden individuelle Beanspruchungsfunktionen nach zwei Verfahren berechnet: dem Ansatz individueller Regressionen (Kapitel 5.3.1) und dem Ansatz der Parameterklassifikation (Kapitel 5.3.2). Die Kalibrierung wurde jeweils vor der zu bearbeitenden Anwendungsaufgabe an beiden Versuchstagen durchgeführt, sodass weiter die Effekte der Wiederholung der Kalibrierungsaufgabe analysiert werden konnten. Die Ergebnisse zeigen, dass eine zuverlässige Vorhersage der mentalen Beanspruchung lediglich mit dem Ansatz der individuellen Parameterklassifikation möglich ist. Die Varianzaufklärung als alleiniges Auswahlkriterium in der Auswertung über individuelle Regressionen ist nicht ausreichend, um die vielfältigen Einflussfaktoren auf die individuelle Informationsverarbeitung abdecken zu können. Es zeigen sich allerdings sowohl auf der generellen, als auch der individuellen Ebene der physiologischen Beanspruchungserfassung Unterschiede in der Diagnostizität zwischen den beiden Kalibrierungsaufgaben. Ein möglicher Einflussfaktor ist dabei die Bekanntheit der Kalibrierungsaufgabe am zweiten Versuchstag. Dieser Effekt bietet einen guten Anschlusspunkt für weite-

re Forschungsarbeiten, da in der späteren Anwendung zumeist die Kombination aus der unbekanntem Kalibrierungsaufgabe mit einer bekannten Anwendungsaufgabe vorliegen wird. Auch die beobachteten Strategie- und Prioritätswechsel in der Qualitätssicherungsaufgabe zeigen einen Forschungsbedarf auf Ebene der physiologischen Indikatoren, um das Verhalten eines solchen Übergangs besser zu verstehen. Insgesamt zeigt die Studie, dass ein individueller Ansatz der Beanspruchungsmessung über eine Kalibrierungsaufgabe zu einer erheblichen Steigerung der Zuverlässigkeit in der Verwendung physiologischer Indikatoren führen kann. Darüber hinaus ermöglicht diese Art der Beanspruchungsmessung eine zuverlässige Bewertung von sozio-technischen Systemen, in denen der Einsatz von leistungsbasierten oder subjektiven Beanspruchungsindikatoren nicht möglich ist.

8 Diskussion

Im Rahmen dieser Arbeit wurden drei empirische Studien durchgeführt. Die zentralen Forschungsfragen umfassten die Entwicklung einer zuverlässigen mentalen Beanspruchungsmessung in sozio-technischen Systemen über physiologischen Parameter nach einem individuellen Ansatz. Das Vorhaben zur Entwicklung einer Kalibrierungsaufgabe für die Herleitung individueller Beanspruchungsfunktionen basiert dabei auf den theoretischen und empirischen Erkenntnissen aus dem Bereich der menschlichen Informationsverarbeitung (Kapitel 2.1). Die Reaktionen eines Menschen auf Reize werden dabei sehr stark durch individuelle Erfahrungen, Ziele oder den Trainingsgrad beeinflusst. Die hohe Individualität der menschlichen Kognition ist ein wesentlicher Grund für die Verhaltensvariabilität, die in der Interaktion verschiedener Menschen mit dem gleichen System beobachtet werden kann. Die Forschungsarbeiten von Dehaene und Changeux (2005) sowie Kiefer et al. (2012) zeigen, dass auch Ausprägungen unbewusster, hoch individueller Informationsverarbeitungsprozesse kognitive Ressourcen verbrauchen, was zu einer steigenden Bedeutung von differenziellen und individuellen Modellen für die zuverlässige Beanspruchungsmessung führt. Zusätzlich verändert der Mensch während der Bearbeitung einer Aufgabe sowohl auf bewusster als auch auf unbewusster Ebene seine Strategien, was wiederum die Betrachtung der mentalen Beanspruchung auf Stichprobenebene zusätzlich erschwert (Matthews & Campbell, 1998). Diese Adaptionsprozesse finden nicht nur aufgrund von Veränderungen der Aufgaben oder des Anforderungsniveaus, sondern auch durch Veränderungen der individuellen Zielstellungen (Sträter, 2005) oder zur Optimierung der Anstrengung

sowie der Leistung statt (Sperling & Doshier, 1986). Zusammenfassend zeigt der theoretische Hintergrund dieser Arbeit (Kapitel 2) die starken inter- und intraindividuellen Unterschiede in der menschlichen Informationsverarbeitung. Insbesondere die Studie von Fels et al. (2015) zeigt, dass die Zuverlässigkeit der physiologischen Beanspruchungsmessung bei der Betrachtung auf einer differenziellen Ebene (Gruppenebene) im Vergleich zu der generellen Ebene (Stichprobenebene) erheblich gesteigert werden kann (Kapitel 2.2.2). Gagnon et al. (2011) postulieren, basierend auf ihren empirischen Befunden, dass die Möglichkeit einer zuverlässigen Beanspruchungsmessung über physiologische Parameter ausschließlich auf der individuellen Ebene funktionieren kann. Diese Hypothese aufgreifend, umfasste die erste Studie dieser Arbeit eine Analyse der Diagnostizität und Sensitivität verschiedener physiologischer Beanspruchungsindikatoren auf genereller, differenzieller und individueller Ebene. Dazu wurden zwei Aufgaben mit jeweils drei Anforderungsniveaus verwendet: eine Flugaufgabe und eine mentale Rotation (Kapitel 4.1). Die Ergebnisse zeigten einige wichtige Diskussionspunkte bezüglich der Operationalisierung der Aufgaben auf. Die in der Flugaufgabe verwendete n-back Nebenaufgabe (Kapitel 3.1.2) war durch mehrmals auftretende Folgefehler geprägt, was die Diagnostizität der Leistungsparameter negativ beeinflusste. Darüber hinaus zeigte sich, dass die Anzahl der Konturpunkte von Geometrien als einziges Kriterium für die Definition der Anforderungsniveaus in der mentalen Rotation nicht ausreichend ist. Die Formgebung hat entscheidenden Einfluss, da verschiedene individuelle Assoziationen mit den Formen geweckt werden können, die das Erkennen von Fehlern erleichtern. Die Auswertung auf genereller Ebene zeigte das Potential der physiologischen (okularen) Parameter durch eine gute Sensitivität. Anders als beispielsweise in den Studien von Schwalm (2009) oder Di Nocera et al. (2007) auf genereller Ebene zeigen die okularen Parameter lediglich eine sehr geringe Diagnostizität. Die fehlende Diagnostizität erlaubt keine zuverlässige Interpretation des Beanspruchungszustandes. Die Analyse auf differenzierter Gruppenebene führt zu einer deutlichen Steigerung der Diagnostizität, wobei die Varianzaufklärung hinsichtlich des subjektiven Beanspruchungsempfindens ($r^2 = 0,23$) und der Leistung ($r^2 = 0,17$) noch sehr gering ist. Zusammen mit der

Instabilität der statistischen Gruppierungsverfahren bei wiederholter Ausführung ist eine sichere und zuverlässige Interpretation auf differenzierter Ebene noch nicht möglich. Bei mehrfachem Wiederholen der Clusterverfahren (Kapitel 3.2.2) verändern sich die Anzahl und die Zusammensetzung der Gruppen teilweise erheblich. Diese Veränderungen beeinflussen wiederum die physiologischen Parameter, die am besten zur Beschreibung der mentalen Beanspruchung in den einzelnen Gruppen geeignet sind. Darüber hinaus können viele Einflussfaktoren (z. B. Aufgabenabhängigkeit) nicht geprüft werden. Es lassen sich somit keine eindeutigen Aussagen hinsichtlich der Vorhersagezuverlässigkeit der gewählten Parameter treffen. Die Auswertung der physiologischen Reaktionen auf der individuellen Ebene weist sehr großes Potential durch eine sehr hohe Diagnostizität auf (Kapitel 4.2.3). Es konnte für jeden Probanden ein physiologischer Indikator identifiziert werden, der zuverlässige Rückschlüsse auf den mentalen Beanspruchungszustand erlaubt. Allerdings war diese Auswertung nur mit erheblichem, nicht praktikablem Aufwand möglich, da die physiologischen Reaktionen der einzelnen Probanden nur schwer miteinander verglichen und in Relation zueinander gesetzt werden konnten. Um die großen Vorteile der individuellen Auswertung durch eine Senkung des Aufwandes und die Schaffung einer Vergleichbarkeit zwischen Probanden auch für die praktische Verwendung zur Verfügung zu stellen, wurde eine Kalibrierungsaufgabe mit zwei Auswertungsverfahren entwickelt. Die Kalibrierungsaufgabe umfasst zwei Aufgabenarten mit jeweils zwei Anforderungsniveaus. Die Aufgabenarten werden über zwei verschiedene Nebenaufgaben definiert: ein visueller Reiz im peripheren Sichtfeld und eine auditiv präsentierte mentale Arithmetik (Kapitel 5.1). Aufgrund der unterschiedlichen Informationskanäle, durch die die Reize der Nebenaufgaben aufgenommen werden, ergeben sich gemäß dem Ressourcenmodell nach Wickens (1984) unterschiedliche Interaktionseffekte mit der visuellen Hauptaufgabe. Die Ergebnisse der dritten Studie zeigen, dass die Fehlerquote in der visuellen Nebenaufgabe signifikant geringer ist als in der arithmetischen Nebenaufgabe (Kapitel 7.2.3). Dieser Unterschied wird nicht nur durch die verschiedenen Interaktionseffekte, sondern auch durch die unterschiedliche Verarbeitung der Reize beeinflusst. Die visuelle Nebenaufgabe besteht aus

einer einfachen Stimulus-Response Aufgabe mit nur einer Antwortmöglichkeit. Die mentale Arithmetik erfordert eine komplexere Verarbeitung des Reizes aufgrund von mehreren Verarbeitungsschritten und Auswahlmöglichkeiten. Die Parameterauswahl der beiden Auswertungsverfahren zur Bildung einer individuellen Beanspruchungsfunktion basiert demnach auf zwei Aufgabenarten, deren Anforderungen sich stark in der Wahrnehmung und Verarbeitung unterscheiden. Durch die Hinzunahme einer dritten Bedingung, die sich in einem mittleren Anforderungsbereich befindet, könnte die Robustheit der Prognose noch weiter gesteigert werden. Die aktuellen Bedingungen erzeugen aufgrund der unterschiedlichen Komplexität der Reize verschieden hohe Beanspruchungen, was die Bildung eines funktionalen Zusammenhangs erschwert. Beide Auswertungsverfahren (Kapitel 5.3.2) dienen der Auswahl eines individuellen physiologischen Prädiktors für die mentale Beanspruchung, operationalisiert über die Leistung, das subjektive Empfinden oder die mentale Effizienz. Das erste Auswertungsverfahren ist der Ansatz individueller Regressionen (Kapitel 5.3.1). Basierend auf den Ansätzen von Jo et al. (2012) und Ryu und Myung (2005) werden jeweils der bewusste und unbewusste Parameter mit der höchsten Varianzaufklärung in einer Regressionsfunktion kombiniert. Jedoch finden individuelle Adaptionsprozesse nicht nur aufgrund von Veränderungen der Aufgaben oder des Anforderungsniveaus statt. Auch Veränderungen der individuellen Zielstellungen (Sträter, 2005) oder eine Optimierung der Anstrengung sowie der Leistung führen zu Strategieanpassungen (Sperling & Doshier, 1986). Aus diesem Grund kann es bei der Verwendung der Varianzaufklärung (aufgabenabhängig) als einziges Kriterium zu Problemen bei der Vorhersage von anderen Aufgaben führen. Die entsprechende Analyse ist Gegenstand der dritten Studie. Der Ansatz der individuellen Parameterklassifikation beschreibt die gezielte Kombination mehrerer methodischer Verfahrensweisen, um die Aufgaben-, Anforderungs- und Zeitabhängigkeit zu prüfen sowie das Adaptionsverhalten und einen möglichen Drift zu beschreiben (Kapitel 5.3.2). Die Procrustes-Analyse (Kapitel 3.2.1) und eine kombinierte Korrelationsanalyse auf drei zeitlichen Ebenen bilden den Kern der Parameterauswahl. Für die Auswahl der geeigneten Verfahren anhand eines Testdatensatzes wurden noch

weitere Verfahren aus Mustererkennung und Statistik berücksichtigt (Logistische Regression, Diskriminanzanalyse, Support-Vector-Machines, Neuronale Netze). Lediglich die beschriebene Kombination der Verfahren innerhalb der individuellen Parameterklassifikation (Kapitel 5.3.2) zeigte zuverlässige Ergebnisse in der Beanspruchungsprognose und Quantifizierung. Innerhalb der zweiten Studie wurden die Operationalisierung der Kalibrierungsaufgabe sowie die Möglichkeit, individuell zuverlässige Beanspruchungsfunktionen mit beiden Auswertungsverfahren zu extrahieren, überprüft. Sowohl das subjektive Beanspruchungsempfinden, als auch die Leistung in Haupt- und Nebenaufgabe zeigen eine sehr gute Differenzierbarkeit der Anforderungsniveaus. Darüber hinaus liegt auch die Anforderungshöhe in dem nach De Waard (1996) optimalen Bereich zur Beanspruchungsmessung. Lediglich die arithmetische Nebenaufgabe war deutlich zu leicht und wurde durch die Umstellung von Additions- auf Subtraktionsaufgaben für die dritte Studie angepasst. Die Operationalisierung der Anforderungsniveaus der Hauptaufgabe über Hinzunahme einer zweiten, simultanen Aufgabe (zwei Kugeln) hat allerdings einen erheblichen Einfluss auf die Sensitivität der bewussten Parameter (insbesondere NNI). Die hohe Sensitivität hinsichtlich der Hauptaufgabe auf genereller Ebene kann zum größten Teil auf das Aufgabendesign zurückgeführt werden, welches das Blickverhalten deutlich beeinflusst. Ein möglicher Forschungsansatz wäre an dieser Stelle beispielsweise die Umsetzung der Anforderungsniveaus durch die Variation der Kugelgeschwindigkeit. Durch die Verwendung des Verfahrens der individuellen Regressionen wird der NNI bei $n = 28$ Probanden (71,79%) als Prädiktor ausgewählt. Bei der individuellen Parameterklassifikation wird die Aufgabenabhängigkeit geprüft und der NNI lediglich für $n = 12$ Probanden (30,77%) ausgewählt. Dieser Unterschied zeigt, vor dem Hintergrund der Ergebnisse bezüglich der Hauptaufgabe, dass eine Auswertung basierend auf der Procrustes-Analyse durchaus in der Lage ist, die Aufgabenabhängigkeit eines Parameters zu prüfen. Jedoch wird auch die vermutete Problematik mit der Varianzaufklärung als einziges Kriterium bestärkt. Nur für $n = 10$ Probanden (25,64%) erfüllt einer der höchst korrelierenden Parameter auch die Anforderungen der individuellen Parameterklassifikation. In der dritten Studie sollte

schließlich geprüft werden, inwieweit eine Vorhersage der mentalen Beanspruchung in den Anwendungsaufgaben über die beiden Auswertungsverfahren möglich ist. Die Probanden bearbeiteten dazu an zwei Versuchstagen jeweils die Kalibrierungsaufgabe und danach eine Anwendungsaufgabe (Qualitätssicherung, Transportlogistik). Eine Vorhersage der mentalen Beanspruchung in den Anwendungsaufgaben ist zuverlässig lediglich durch das Verfahren der individuellen Parameterklassifikation möglich. Die Parameter mit der höchsten Varianzaufklärung verändern ihr Verhalten sowohl bei der Bearbeitung einer anderen Aufgabe, als auch bei dem wiederholten Ausführen der Kalibrierungsaufgabe bei vielen Probanden sehr stark. Für die erste Kalibrierung erfüllt auch nur für $n = 11$ Probanden (30,56%) mindestens einer der höchst korrelierenden Parameter die Anforderungen der Parameterklassifikation. In der zweiten Kalibrierung sinkt die Anzahl sogar auf $n = 5$ Probanden (13,89%). Allerdings ist eine Vorhersage der Aufgaben am zweiten Versuchstag durch die erste Kalibrierung auch mit dem Verfahren der individuellen Parameterklassifikation nicht möglich. Ein möglicher Grund kann in dem großen Unterschied zwischen Versuchstag 1 und 2, aufgrund der erfahrungsbasierten Adaption der individuellen Leistungskurve, liegen. Die Kalibrierungsaufgabe ist am ersten Tag noch unbekannt, am zweiten Tag aber bekannt. Die Anwendungsaufgaben sind dagegen an beiden Tagen unbekannt. Dieser Unterschied verdeutlicht sich durch fünf Probanden, die in der zweiten Anwendungsaufgabe auch durch das Modell aus der zweiten Kalibrierung falsch vorhergesagt werden. Die Zeit zwischen den Versuchstagen ist für diese fünf Probanden signifikant kürzer als für die restliche Stichprobe. Auch sind die Adaptionen der physiologischen Reaktionen dieser fünf Probanden in der zweiten Kalibrierung signifikant weniger als in der ersten Kalibrierung, während bei der restlichen Stichprobe keine Unterschiede vorliegen. Die fünf Probanden können zuverlässig auch in der zweiten Anwendungsaufgabe durch die erste Kalibrierung vorhergesagt werden. Auch auf Stichprobenebene zeigen sich signifikante Zusammenhänge zwischen der Zeit zwischen den Versuchstagen und dem Leistungsanstieg in Kalibrierung 2 im Vergleich zu Kalibrierung 1. Der Einfluss der Zeit zwischen den Kalibrierungen sollte in nachfolgenden Studien noch weiter untersucht werden, um die

Verläufe der individuellen Leistungskurven genau nachzuvollziehen. Aus den Erkenntnissen können dann Empfehlungen für die Praxis abgeleitet werden, die den Zeitraum definieren, ab dem die Kalibrierung bei messwiederholten Versuchsdesigns wiederholt werden sollte. Insgesamt zeigen die Ergebnisse der individuellen Parameterklassifikation, dass die mentale Beanspruchung über die Kalibrierung unmittelbar vor einer Anwendungsaufgabe hoch zuverlässig vorhergesagt und quantifiziert werden kann. Auch die untersuchten Aufgabekombinationen, bestehend aus unbekanntem Anwendungsaufgaben, jeweils einmal mit einer bekannten und einmal mit einer unbekanntem Kalibrierungsaufgabe, sollten mit Blick auf die praktische Anwendung erweitert werden. In der Praxis wird sehr wahrscheinlich mit einer unbekanntem Kalibrierung eine sehr gut bekannte Tätigkeit analysiert. Dieser Fall wurde jedoch von den Studien nicht abgedeckt, weshalb die Effekte in weiteren Studien analysiert werden sollten. Weiter sollten auch die, bei der Qualitätssicherungsaufgabe identifizierten, Strategie- und Prioritätswechsel im zeitlichen Verlauf der Aufgabe in weiteren Studien detailliert analysiert werden. Die Erkenntnisse können dann in die Klassifikation des Adaptionsverhaltens einfließen und die Zuverlässigkeit der Kalibrierung weiter erhöhen. Insgesamt konnten durch die drei durchgeführten Studien die Forschungsfragen dieser Arbeit (Kapitel 2.3) beantwortet werden. Es wurde eine Kalibrierungsaufgabe entwickelt, die alle definierten Anforderungen einer praktischen Anwendung erfüllt. Die Kalibrierung ermöglicht die zuverlässige Vorhersage der mentalen Beanspruchung über die physiologischen Reaktionen des Menschen auf individueller Ebene. Das Auswertungsverfahren der individuellen Parameterklassifikation berücksichtigt mögliche Einflussfaktoren auf die Parameter, sowie die individuellen Adaptionseffekte, die über die kombinierte Korrelationsanalyse erfolgreich identifiziert werden können. Die Ergebnisse der Arbeit ermöglichen eine hoch zuverlässige Auswertung der mentalen Beanspruchung auf der individuellen Ebene, wobei der Aufwand der Auswertung minimiert und Vergleichbarkeit deutlich gesteigert werden konnte. Die Vergleichbarkeit resultiert dabei aus der eindeutigen Quantifizierung der mentalen Beanspruchung im Wertebereich von Null (keine Beanspruchung) bis Eins (maximale Beanspruchung). Aufgrund der Berücksichtigung der In-

dividualität physiologischer Reaktionen ist diese Quantifizierung auch in der Prognose der Beanspruchung aufgabenunabhängig sehr zuverlässig. Schließlich erlaubt die Teilautomatisierung der Auswertung in Kombination mit der kurzen Ausführungsdauer der Kalibrierungsaufgabe die individuelle Auswertung mit vergleichsweise geringem Aufwand.

9 Ausblick

Die vorliegende Arbeit konnte die zentralen Forschungsfragen (Kapitel 2.3) nach einer zuverlässigen, aufgabenunabhängigen Vorhersage der mentalen Beanspruchung beantwortet werden. Durch die Entwicklung der Kalibrierungsaufgabe und dem Auswertungsverfahren der individuellen Parameterklassifikation ist es möglich die individuelle Physiologie des Menschen zu berücksichtigen. Die individuelle Beanspruchungsfunktion, basierend auf den Ergebnissen der Kalibrierungsaufgabe, erlaubt die hoch zuverlässige Beanspruchungsbewertung über physiologische, blickbasierte Indikatoren, ohne subjektive oder leistungsorientierte Parameter als Referenzgröße in Anwendungsaufgaben simultan erfassen zu müssen. Die Entwicklungen dieser Arbeit leisten damit einen Beitrag zur menschenzentrierten und nachhaltigen Systemgestaltung sozio-technischer Systeme. Um die vielseitigen Ziele der Systemgestaltung (Kapitel 1) zu erreichen, ist es wichtig, die Systeminformationen sowie die individuellen Beanspruchungszustände und die Verhaltensvariabilität der Anwender zu erfassen und zu verstehen. Die Kalibrierungsaufgabe und das damit verbundene Auswertungsverfahren der individuellen Parameterklassifikation sind Bestandteile der Erfindungsmeldung EM16352, die am 30.09.2016 am Karlsruher Institut für Technologie (KIT) eingereicht wurde. Die Erfindungsmeldung umfasst darüber hinaus auch ein Verfahren zur echtzeitfähigen Erfassung der individuellen Strategien und von Strategiewechseln über die Auswertung der blickbasierten Daten. Diese Kombination erlaubt ein umfassendes Verständnis des Anwenders in sozio-technischen Systemen, sodass zuverlässig Gestaltungsempfehlungen abgeleitet werden können. Neben der Produktgestaltung und -entwicklung können die Verfahren auch

zur Auslegung und Gestaltung von Arbeitsplätzen in vielfältigen Domänen genutzt werden. In der Automobilbranche beispielsweise werden immer mehr Funktionen und Informationen für den Fahrer eines Fahrzeuges bereitgestellt. Diese werden über verschiedene Infotainment-Systeme mit unterschiedlichsten Eingabemöglichkeiten dargeboten. Die Ablenkungswirkungen solcher Systeme sind bereits weitreichend erforscht (Wickens et al., 2013). Doch wie hoch ist die Ablenkungswirkung im Detail und wie können durch die Gestaltung der Anzeigeelemente oder deren Positionierung (in Abhängigkeit von verschiedenen Fahrsituationen) Gestaltungsempfehlungen abgeleitet werden? Genau diese Fragestellung kann durch diese Arbeit in Kombination mit der Erfindungsmeldung beantwortet werden. Durch die Erfassung von Strategiewechseln kann eine Priorisierung des Infotainment-Systems in Abhängigkeit von der Situation erkannt werden und anhand des Beanspruchungszustandes bewertet werden. Durch den Vergleich mit anderen Betrachtungsintervallen können gestalterische Lösungen abgeleitet werden. So kann unter anderem der Informationsgehalt von Anzeigesystemen in bestimmten Situationen systematisch verringert werden, um potentiell häufige Blickzuwendungen und damit eine Priorisierung des Zusatzinformationssystems zu verhindern. Die allgemeine Luftfahrtindustrie entwickelt und konzipiert Luftfahrzeuge für private und kleinere geschäftliche Kunden im Bereich des Motor- und des Segelflugs. Durch die starke Digitalisierung der Instrumente und dem Einzug der rechnergestützten Navigation in die Cockpits ist mittlerweile fast jedes Cockpit in Flugzeugen der allgemeinen Luftfahrt ein Unikat, welches nach den Wünschen des Kunden gestaltet wird. Die Unfallzahlen in der allgemeinen Luftfahrt nehmen stetig zu und die Gründe sind in den meisten Fällen auf den Menschen und eine nicht-korrekte Bedienung oder Entscheidung zurückzuführen. Jede Verhaltensweise des Menschen wird jedoch durch das System, mit welchem er interagiert, stark beeinflusst (Sträter, 2005). Eine negative Verhaltensvariabilität des Menschen kann durch eine angepasste Systemgestaltung abgefangen werden, da die notwendigen Informationen in kritischen Situationen dort und in der Art dargestellt werden, wo und wie sie der individuelle Nutzer erwartet. Fliegt der Nutzer vor dem Kauf mit einem Blickerfassungssystem den Probeflug der Maschine, so können seine Strategien

und seine Beanspruchungen erfasst und ausgewertet werden. Negativen Veränderungen der Beanspruchung in verschiedenen Flugphasen kann durch die Anordnung und Darstellungsform der Instrumente entgegengewirkt werden. Die notwendigen Informationen für eine solche Systemgestaltung werden von den Entwicklungen dieser Arbeit und der Erfindungsmeldung bereitgestellt. Auch im Bereich der kommerziellen Luftfahrt kann die Erfindung in zwei sicherheitsrelevanten Bereichen zu einer erheblichen Steigerung der Systemzuverlässigkeit beitragen. Die Cockpits der Passagierflugzeuge zählen zu den komplexesten Arbeitssystemen mit dem höchsten Überwachungsaufwand (Parasuraman & Rizzo, 2007). Ein hoch technisiertes System, welches überwiegend automatisiert agiert, muss durch die Piloten ständig überwacht werden. In kritischen Situationen müssen Piloten schnellstmöglich die Fehlerquelle identifizieren und die geeignete Vorgehensweise zur Störungsbehebung einleiten. Hier reagiert jeder Pilot, trotz standardisierten Trainingsverfahren, aufgrund der individuellen Erfahrungen anders. Durch den Einsatz der Beanspruchungs- und Strategieerfassung innerhalb des Pilotentrainings (2x jährlich im Full-Flight-Simulator) können Piloten bedarfsgerecht auf den Umgang mit kritischen Situationen vorbereitet werden. Auf der technischen Seite können die Cockpits, ähnlich wie in der Automobilbranche, so gestaltet werden, dass sie die menschliche Verhaltensvariabilität berücksichtigen und so in kritischen Situationen optimal unterstützen. Auch der Bereich der Flugsicherung gilt als ein hoch sicherheitsrelevanter Bereich, in dem die Verfahren zur Gestaltung der Leitwarte, als auch zur Schulung und Fehleranalyse bei Fluglotsen eingesetzt werden könnten. Neben der Vielzahl an wirtschaftlichen Verwendungsmöglichkeiten der entwickelten Beanspruchungsmessung in Kombination mit der Strategieerfassung bieten die Ergebnisse dieser Arbeit auch sehr wichtige wissenschaftliche Anknüpfungspunkte für folgende Forschungsarbeiten. In den Diskussionskapiteln der einzelnen Studien werden einige Forschungsfragen aufgeworfen. Folgend sollen die wichtigsten noch einmal zusammengefasst dargestellt werden. Die Operationalisierung der Anforderungsniveaus der Hauptaufgabe innerhalb der Kalibrierungsaufgabe hat, durch das gewählte Aufgabendesign mit zwei simultanen Aufgaben, einen erheblichen Einfluss auf die Blickstrategie und damit auch auf die Sensitivität der

bewussten physiologischen Parameter. Ein Ansatzpunkt wäre es, die Schwierigkeit einer einzelnen Hauptaufgabe, durch beispielsweise die Variation der Geschwindigkeit der Kugel oder einer invertierten Steuerung, zu definieren. Allerdings würde so zu einem gewissen Maße die Repräsentativität verloren gehen, da in den meisten sozio-technischen Systemen, wie oben beschrieben, mehrere Aufgaben simultan bearbeitet werden müssen. Darüber hinaus sollten die Strategie- und Prioritätswechsel über die Dauer einer Aufgabe hinweg detailliert, auch unter Berücksichtigung der physiologischen Reaktionen, analysiert werden. Die Erkenntnisse darüber, wann genau und mit welchen Auswirkungen die Strategie gewechselt wird, kann wichtige Kriterien für die Klassifikation des Adaptionverhaltens innerhalb der individuellen Parameterauswahl bereitstellen, was wiederum die Zuverlässigkeit weiter erhöhen kann. In der Diskussion der dritten Studie zeigte sich der praxisorientierte Bedarf, die Effekte einer Kalibrierung mit einer, für den Probanden unbekannt, Kalibrierungsaufgabe für die Vorhersage einer bekannten Anwendungsaufgabe zu untersuchen. Diese Konstellation wird sich in den meisten praktischen Erhebungen wiederfinden, wurde aber in der Form in den aktuellen Studien noch nicht untersucht. Die identifizierten Unterschiede zwischen den beiden analysierten Versuchstagen in der dritten Studie zusammen mit dem Einfluss durch die Zeit zwischen den Versuchen motivieren diese Forschungsfrage. Weiter können infolge der Weiterentwicklung zu einer echtzeitfähigen Beanspruchungserfassung auch die Auswirkungen des Feedbacks über den Beanspruchungszustand an den Anwender selbst untersucht werden. Wie beeinflusst dieses Wissen das Verhalten oder die Adaptionen? Im Rahmen dieser Arbeit wurden lediglich sozio-technische Systeme, bestehend aus einem oder mehreren technischen Systemen und einem Menschen, analysiert. Die Möglichkeit der zuverlässigen Beanspruchungsmessung über die individuelle Auswertung, einhergehend mit der Erfassung der Strategien der Anwender, ermöglicht zukünftig auch die Analyse von sozio-technischen Systemen mit mehreren Menschen (z. B. Pilot und Co-Pilot). Wichtige Forschungsfragen bezüglich der Verteilung von Beanspruchungen innerhalb des Teams oder wann eine Aufgabe in welcher Form übergeben werden sollte, können hier im Fokus stehen. Zusammenfassend bieten die Ergebnisse dieser Arbeit sowohl vielseitige

wirtschaftliche Anwendungsmöglichkeiten, als auch wichtige und nachhaltige wissenschaftliche Anknüpfungspunkte.

Literaturverzeichnis

- Anand, S., Terken, J. & Hogema, J. (2013). The Effect of Cognitive Load on Adaption to Differences in Steering Wheel Force Feedback Level. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 158–164.
- Arenius, M. (2009). *Getting the Feeling - “Human Error” in an educational ship-handling simulator*. Magisterarbeit, University Linköping, Linköping.
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*(44), 75–106.
- Bainbridge, L. (1997). The change in concepts needed to account for human behavior in complex dynamic tasks. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 27(3), 351–359.
- Beatty, J. & Lucero-Wagoner, B. (2007). The Pupillary System. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 142–162). Cambridge [England] and New York: Cambridge University Press.
- Benedetto, S., Pedrotty, M., Minin, L., Baccino, T., Re, A. & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research*(14), 199–208.

- Bertrau-Pavy, D., Raber, J. & Piper, B. (2011). Contributions of age, but not sex, to mental rotation performance in a community sample. *PEBL Technical Report Series*(2).
- Bodala, I. P., Ke, Y., Mir, H., Thakor, N. V. & Al-Nashash, H. (2014). Cognitive workload estimation due to vague visual stimuli using saccadic eye movements. *Proceedings of IEEE 36th Annual International Conference of Engineering in Medicine and Biology Society (EMBC)*, 2993–2996.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollst. überarb. u. erw. Aufl. Aufl.). Berlin: Springer.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.
- Brookhuis, K. & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*(42), 898–903.
- Cardona, G. & Quevedo, N. (2014). Blinking and driving: the influence of saccades and cognitive Workload. *Current Eye Research*(39), 239–244.
- Chang, S. G., Yu, B. & Vetterli, M. (2000). Adaptive Wavelet Thresholding for Image Denoising and Compression. *IEEE Transactions on Image Processing*(9(9)), 1532–1546.
- Chen, S. & Epps, J. (2014). Using Task-Induced Pupil Diameter and Blink Rate to Infer Cognitive Load. *Human-Computer Interaction*(29(4)), 390–413.
- Clark, P. J. & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*(35(4)), 445–453.
- Coltheart, M. (2000). Deep Dyslexia Is Right-Hemisphere Reading. *Brain and Language*(71), 299–309.
- Cooper, G. E. & Harper, R. P. (1969). The Use of Pilot Rating in the Evaluation of Aircraft handling Qualities. *NASA Ames Technical Report NASA TN-D-5153*, 111–151.

- Daubechies, I. (Hrsg.). (2006). *Ten lectures on wavelets* (Bd. 61, 9. print Aufl.). Philadelphia, Pa.: Soc. for Industrial and Applied Mathematics.
- De Waard, D. (1996). *The measurement of drivers' mental workload*. Haren: The Traffic Research Centre VSC, University of Groningen.
- Dehaene, S. & Changeux, J.-P. (2005). Ongoing Spontaneous Activity Controls Access to Consciousness: A Neuronal Model for Inattentive Blindness. *PLOS Biology*(3), e141.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J. & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences*, 10(5), 204–211.
- Demberg, V. (2013). Pupillometry: The Index of Cognitive Activity in a dual-task study. In M. Knauff, M. Pauen, M. Sebanz & I. Wachsmuth (Hrsg.), *Proceedings of 35th Annual Conference of the Cognitive Science Society* (Bd. 35). Austin: Austin TX.
- Di Nocera, F., Camilli, M. & Terenzi, M. (2006). Using the Distribution of Eye Fixations to Assess Pilots' Mental Workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(1(3)), 271–285.
- Di Nocera, F., Camilli, M. & Terenzi, M. (2007). *A Random Glance on the Flight Deck: Pilots' Scanning Strategies and Real-Time Assessment of Mental Workload*. Rom.
- Di Nocera, F., Terenzi, M. & Camilli, M. (2006). Another look at scanpath: distance to nearest neighbour as a measure of mental workload. In D. De Waard, K. A. Brookhuis & A. Toffetti (Hrsg.), *Developments in human factors in transportation, design, and evaluation* (S. 295–303). Maastricht: Shaker Publishing.
- Di Stasi, L. L., Antolí, A. & Canas, J. J. (2013). Evaluating mental workload while interacting with computer-generated artificial environments. *Entertainment Computing*(4(1)), 63–69.

- Donoho, D. L. (1995). De-Noising by Soft-Thresholding. *IEEE Transactions on Information Theory*(41(3)), 613–627.
- Donoho, D. L. & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*(81(3)), 425–455.
- Donoho, D. L. & Johnstone, I. M. (1998). Minimax Estimation via Wavelet Shrinkage. *The Annals of Statistics*(26(3)), 879–921.
- Donoho, D. L., Johnstone, I. M., Kerkycharian, G. & Picard, D. (1996). Density Estimation by Wavelet Thresholding. *The Annals of Statistics*(24(2)), 508–539.
- Dovidio, J. F., Kawakami, K. & Johnson, C. (1997). On the Nature of Prejudice: Autonomic and Controlled Processes. *Journal of Experimental Social Psychology*(33), 510–540.
- Driver, M. J. & Mock, T. J. (1975). Human Information Processing, Decision Style Theory, and Accounting Information Systems. *The Accounting Review*(50), 490–508.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behaviour. *Psychological Review*(66), 183–201.
- Eggemeier, T. F. & Wilson, G. F. (1991). Performance-based and subjective assessment of workload in multi-task environments. In D. L. Damos (Hrsg.), *Multiple-task performance* (S. 217–275). London and Washington, DC: Taylor & Francis.
- Eilers, K., Nachreiner, F. & Hänecke, K. (1986). Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*(40(12)), 215–224.
- Engström, J., Åberg, N., Johansson, E. & Hammarbäck, J. (2005). Comparison between Visual and Tactile Signal Detection Tasks Applied to the Safety Assessment of in-Vehicle Information Systems. *Proceedings of*

-
- the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 232–240.
- Estes, S. (2015). The Workload Curve: Subjective Mental Workload. *Human Factors*(57(7)), 1174–1187.
- Eysenck, M. W. (1982). *Attention and Arousal: Cognition and Performance*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Faure, V., Lobjois, R. & Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 40, 78–90.
- Fels, M., Bauer, R. & Gharabaghi, A. (2015). Predicting workload profiles of brain–robot interface and electromyographic neurofeedback with cortical resting-state networks: Personal trait or task-specific challenge? *Journal of Neural Engineering*, 12(4), 046029.
- Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll* (4th edition Aufl.). Los Angeles: Sage.
- 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), 129–145.
- Gagnon, J., Durantin, G., Vachon, F., Causse, M., Tremblay, S. & Dehais, F. (2011). Anticipating Human Error before it happens: Towards a psychophysiological model for online prediction of mental workload. In *Proceedings of the Human Factors and Ergonomics Society 55th annual meeting* (S. 100–111). Santa Monica, CA: Human Factors and Ergonomics Society.
- Galy, E., Cariou, M. & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International journal*

of psychophysiology : official journal of the International Organization of Psychophysiology, 83(3), 269–275.

- Galy, E. & Mélan, C. (2015). Effects of Cognitive Appraisal and Mental Workload Factors on Performance in an Arithmetic Task. *Applied Psychophysiological Biofeedback*(40), 313–325.
- Goldstein, R., Walrath, L. C., Stern, J. A. & Stroock, B. D. (1985). Blink Activity in a Discrimination Task as a Function of Stimulus Modality and Schedule of Presentation. *Psychophysiological*(6), 629–635.
- Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*(40(1)), 33–51.
- Handl, A. (2010). *Multivariate Analysemethoden: Theorie und Praxis multivariater Verfahren unter besonderer Berücksichtigung von S-PLUS* (2. Aufl. Aufl.). Heidelberg: Springer.
- Harbluk, J. L., Burns, P. C., Hernandez, S., Tam, J. & Glazduri, V. (2013). Detection Response Tasks: Using Remote, Headmounted And Tactile Signals to Assess Cognitive Demand While Driving. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 78–84.
- Harris, R. L., Glover, B. J. & Spady, A. A. (1986). Analytical Techniques of Pilot Scanning Behaviour and Their Application. *NASA Technical Paper NASA-TP-2525*, 1–39.
- Harris, R. L., Tole, J. R., Stephens, A. T. & Ephrat, A. R. (1982). Visual scanning behaviour and pilot workload. *Aviation, Space, and Environmental Medicine*(53(11)), 1067–1072.
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Hrsg.), *Human Mental Workload*. Amsterdam: North Holland Press.

- Hick, W. E. (1952). One the rate of gain of information. *Quartely Journal of Experimental Psychology*(4(1)), 11–26.
- Hilburn, B. & Jorna, P. G. (2001). Workload and air traffic control. In P. A. Hancock & P. A. Desmond (Hrsg.), *Stress, Workload, and Fatigue // Stress, workload, and fatigue*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Hilburn, B., Jorna, P. G., Byrne, E. A. & Parasuraman, R. (1996). The effect of adaptive air traffic control (ATC) dicision aiding on controller mental workload. *Proceedings of the 2nd Automation Technology and Human Performance Conference*, 84–91.
- Hitch, G. J. (1978). The Role of Short-term Working Memory in Mental Arithmetic. *Cognitive Psychology*(10), 302–323.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*(45), 73–93.
- Hollnagel, E. (2004). *Barriers and accident prevention*. Aldershot, Hampshire, England and Burlington, VT: Ashgate.
- Hoover, A., Singh, A., Fishel-Brown, S. & Muth, E. (2012). Real-time detection of workload changes using heart rate variability. *Biomedical Signal Processing and Control*, 7(4), 333–341.
- Ichikawa, N. & Ohira, H. (2004). Eyeblink Activity as an Index of Cognitive Processing: Temporal Distribution of Eyeblinks as an Indicator of Expectancy in semantic Priming. *Perceptual and Motor Skills*(98), 131–140.
- Iqbal, S. T., Adamczyk, P. D., Zheng, X. S. & Bailey, B. P. (2005). Towards an Index of Opportunity: Understanding Changes in Mental Workload during Task Execution. *Interruptions & Attention*(1), 311–320.
- Jessee, M. S. (2010). Ocular activity as a measure of mental and visual workload. *Proceedings of the HFES 2010 Human Factors and Ergonomics Society*.

- Jo, S., Myung, R. & Yoon, D. (2012). Quantitative prediction of mental workload with the ACT-R cognitive architecture. *International Journal of Industrial Ergonomics*, 42(4), 359–370.
- Jong, T. de. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instr Sci*(38), 105–134.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kantowitz, B. H. & Knight, J. L. (1976). Testing tapping timesharing: I. Audio secondary Task. *Acta Psychologica*(40), 343–362.
- Katzenberger, L. F. (1967). *Auffassung und Gedächtnis: Psychologische und Pädagogische Aspekte*. München: Reinhardt Verlag.
- Ketelsen, K. & Welsh, M. (2010). Working memory and mental arithmetic: A case for dual central executive resources. *Brain and Cognition*(74), 203–209.
- Kiefer, M., Adams, S. & Zovko, M. (2012). Attentional sensitization of unconscious visual processing: Top-down influences on masked priming. *Advances in Cognitive Psychology*, 8(1), 50–61.
- Kihlstrom, J. F. (1987). The Cognitive Unconscious. *Science*(237), 1445–1452.
- Kohler, I. (1951). *Über Aufbau und Wandlungen der Wahrnehmungswelt: Insbesondere über bedingte Empfindungen*. Unver"offentlichte Dissertation, Österreichische Akademie der Wissenschaften, Wien.
- Kramer, A. F., Sirevaag, E. J. & Braune, R. (1987). A Psychophysiological Assessment of Operator Workload During Simulated Flight Missions. *Human Factors*(29), 145–160.
- Kramer, A. F., Wickens, C. D. & Donchin, E. (1983). An analysis of the processing requirements of a complex perceptual-motor task. *Human Factors*(25), 597–621.

- Laeng, B., Sirois, S. & Gredebäck, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*(7(1)), 18–27.
- Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in cognitive sciences*(7), 12–18.
- Lee, K.-M. & Kang, S.-Y. (2002). Arithmetic operation and working memory: differential suppression in dual tasks. *Cognition*(83), 63–68.
- Lin, Z. & He, S. (2009). Seeing the invisible: the scope and limits of unconscious processing in binocular rivalry. *Progress in neurobiology*, 87(4), 195–211.
- Loewenfeld, I. (1993). *The Pupil: Anatomy, Physiology, and clinical Applications*. Unveröffentlichte Dissertation, Iowa State University, Ames.
- Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rößler (Hrsg.), *Enzyklopädie der Psychologie* (Bd. 5, S. 799–864). Göttingen: Hogrefe.
- Marcel, A. J. (1983a). Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes. *Cognitive Psychology*(15), 238–300.
- Marcel, A. J. (1983b). Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition. *Cognitive Psychology*(15), 197–237.
- Marshall, S. P. (18.07.2000). *Method and Apparatus for Eye tracking and monitoring Pupil Dilation to evaluate cognitive Activity* (Nr. 6.090.051).
- Marshall, S. P. (2002). *The Index of Cognitive Activity: Measuring Cognitive Workload*. Scottsdale.
- Marshall, S. P. (2007). Identifying Cognitive State from Eye Metrics. *Aviation, Space, and Environmental Medicine*(78), 165–175.

- Marshall, S. P., Pleydell-Pearce, C. & Dickson, B. (2003). Integrating Psychophysiological Measures of Cognitive Workload and Eye-Movements to Detect Strategy Shifts. *Proceedings of the Thirth-Sixth Annual Hawaii International Conference on System Sciences*.
- Martins, R. & Carvalho, J. M. (2015). Eye blinking as an indicator of fatigue and mental load - a systematic review. In P. M. Arezes (Hrsg.), *Occupational safety and hygiene III* (S. 231–235). Leiden, The Netherlands: CRC Press/Balkema.
- Matthews, G. & Campbell, S. E. (1998). Task-Induced Stress and Individual Differences in Coping. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, 821–825.
- Matthews, G., Reinerman-Jones, L. E., Barber, D. J. & Abich, J. (2015). The Psychometrics of Mental Workload: Multiple Measures Are Sensitive but Divergent. *Human Factors*(57(1)), 125–143.
- May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P. & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta Psychologica*(75(1)), 75–89.
- Meister, D. (1976). *Behavioural foundations of system development*. New York: Wiley.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*(63), 81–97.
- Moray, N. (1979). *Mental Workload: Its Theory and Measurement* (Bd. 8). Boston, MA: Springer US and Imprint and Springer.
- Mulder, G. (1986). The concept and measurement of mental effort. In G. R. J. Hockey, A. W. K. Gaillard & M. G. H. Coles (Hrsg.), *Energetics and human information processing* (S. 175–198). Dordrecht: Martinus Nijhoff Publishers.

- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W.H. Freeman.
- Neudörfer, A. (Hrsg.). (2014). *Konstruieren sicherheitsgerechter Produkte: Methoden und systematische Lösungssammlungen zur EG-Maschinenrichtlinie* (6., aktual. Aufl. Aufl.). Berlin: Springer Vieweg.
- Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research*, 52(2-3), 207–215.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Niederl, T. (2007). Untersuchungen zu kumulativen psychischen und physiologischen Effekten des fliegerischen Personals auf der Kurzstrecke. *Forschungsbericht des Deutschen Zentrum für Luft- und Raumfahrt e. V.* 17.
- O'Donnell, R. D. & Eggemeier, T. F. (1986). Workload assessment methodology. In K. Boff, L. Kaufmann & J. Thomas (Hrsg.), *Handbook of Perception and Performance* (Bd. 2, S. (42)1–49). New York: Wiley.
- Ogden, R. T. (1997). *Essential wavelets for statistical applications and data analysis*. Boston: Birkhäuser.
- Ogden, R. T. & Parzen, E. (1996). Data dependent wavelet thresholding in nonparametric regression with change-point applications. *Computational Statistics*(22), 53–70.
- Osis, S. T., Hettinga, B. A., Macdonald, S. L. & Ferber, R. (2015). A novel method to evaluate error in anatomical marker placement using a modified generalized Procrustes analysis. *Computer methods in biomechanics and biomedical engineering*(10), 1108–1116.
- Paas, F. G. W. & VanMerriënboer, J. J. G. (1993). The Efficiency of Instructional Conditions: An Approach to Combine Mental Effort and Performance Measures. *Human Factors*(35(4)), 737–743.

- Packebusch, L. (2003). Psychische Belastung und Beanspruchung - Normung für die Praxis. In *Wirtschaftspsychologie aktuell* (Bd. 4, S. 32–36). Deutscher Psychologen Verlag GmbH.
- Parasuraman, R. & Rizzo, M. (Hrsg.). (2007). *Neuroergonomics: The brain at work*. Oxford, New York: Oxford University Press.
- Paus, T. (2000). Functional anatomy of arousal and attention systems in the human brain. *Progress in Brain Research*(126), 65–77.
- Porac, C. & Coren, S. (1976). The dominant eye. *Psychological Bulletin*(83(5)), 880–897.
- Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering* (Bd. 12). New York, N.Y.: North-Holland.
- Recarte, M. A., Perez, E., Conchillo, A. & Nunes, L. M. (2008). Mental workload and visual impairment: differences between pupil, blink and subjective rating. *The Spanish Journal of Psychology*(11), 374–385.
- Reid, G. B. & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Advances in Psychology*(52), 185–218.
- 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.
- Rößger, P. (1997). *Die Entwicklung der Pupillometrie zu einer Methode der Messung mentaler Beanspruchung in der Arbeitswissenschaft*. Dissertation, Sinzheim.
- Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*. Dissertation, RWTH Aachen, Aachen.

- Rubio, S., Díaz, E., Martín, J. & Puente, J. M. (2004). Evaluation of Subjective mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review*(53), 61–86.
- Ryu, K. & Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, 35(11), 991–1009.
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*(53), 61–97.
- Schlick, C., Bruder, R. & Luczak, H. (2010). *Arbeitswissenschaft* (3., vollst. überarb. und erw. Aufl. Aufl.). Berlin and Heidelberg: Springer.
- Schneider, M., Bringeland, S. & Deml, B. (in Review). Classification of psychophysiological measures for mental workload assessment in human-robot-interaction. *Theoretical Issues in Ergonomics Science*.
- Schneider, M. & Deml, B. (2016). Evaluation einer multimodalen Mensch-Roboter-Schnittstelle. In Jäger M. (Hrsg.), *Arbeit in komplexen Systemen - Digital, vernetzt, human?!* Dortmund: GFA Press.
- Schneider, W. & Shiffrin, R. M. (1977). Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention. *Psychological Review*(84), 2–67.
- Schnotz, W. & Kürschner, C. (2007). A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, 19(4), 469–508.
- Schubert, P. & Kirchner, M. (2014). Ellipse area calculations and their applicability in posturography. *Gait Posture*(39), 518–522.
- Schwalm, M. (2009). *Pupillometrie als Methode zur Erfassung mentaler Beanspruchungen im automotive Kontext*. Dissertation, University of Saarland, Saarbrücken.

- Shapiro, K. L., Arnell, K. M. & Raymond, J. E. (1997). The attentional blink. *Trends in cognitive sciences*(1(8)), 291–296.
- Shiffrin, R. M. & Schneider, W. (1977). Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory. *Psychological Review*(84), 69–133.
- Shingledecker, C. A. & Crabtree, M. S. (1982). Subsidiary radio communications tasks for workload assessment. *Research & Development Simulations: 11. Task sensitivity evaluation (Report No. AFAMRL-TR-82-57)*.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed. Aufl.). Cambridge, Mass.: MIT Press.
- Sirevaag, E. J., Kramer, A. F., Wickens, C. D., Reisweber, M., Strayer, D. & Grenell, J. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. *Ergonomics*(36), 1121–1140.
- Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford, New York: Pergamon Press.
- Sonnleitner, A., Simon, M., Kincses, W. E., Buchner, A. & Schrauf, M. (2012). Alpha spindles as neurophysiological correlates indicating attentional shift in a simulated driving task. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*(1), 110–118.
- Spelke, E., Hirst, W. & Neisser, U. (1976). Skills of divided attention. *Cognition*(4), 215–230.
- Sperandio, A. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*(21), 367–390.
- Sperling, G. & Doshier, B. A. (1986). Strategy and Optimization in Human Information Processing. In K. Boff, L. Kaufmann & J. Thomas (Hrsg.), *Handbook of Perception and Performance* (Bd. 1, S. 1–65, Kapitel 2). New York: Wiley.

- 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.
- Sträter, O. (2005). *Cognition and safety: An integrated approach to systems design and assessment*. Aldershot, England and Burlington, VT: Ashgate.
- Tamietto, M., Cauda, F., Corazzini, L. L., Savazzi, S., Marzi, C. A., Goebel, R., Weiskrantz, L. & Gelder, B. de. (2010). Collicular Vision Guides Nonconscious Behaviour. *Journal of Cognitive Neuroscience*(22(5)), 888–902.
- Thornton, D. C. (1985). An Investigation of the "Von Restorff" Phenomenon. *Proceedings of the Human Factors Society 29th Annual Meeting*, 760–764.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*(12), 242–248.
- Trepel, M. (2004). *Neuroanatomie: Struktur und Funktionen* (3., neu bearb. Aufl. Aufl.). München [u.a.]: Urban & Fischer.
- Trowbridge, F. A. & Collins McLaughlin, A. (2013). Perceptual Dual-Task Training via Simulation for Veterinary Students. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(57), 1780–1784.
- Tsang, P. S. & Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*(39(3)), 358–381.
- Tse, P. U., Martinez-Conde, S., Schlegel, A. A. & Macknik, S. L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proceedings of the National Academy of Sciences of the United States of America*, 102(47), 17178–17183.

- van Gaal, S. & Lamme, V. A. F. (2012). Unconscious High-Level Information Processing: Implication for Neurobiological Theories of Consciousness. *The Neuroscientist*(18), 287–301.
- Veltman, J. A. & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*(41), 656–669.
- Verner, M., Herrmann, M. J., Troche, S. J., Roebers, C. M. & Rammsayer, T. H. (2013). Cortical oxygen consumption in mental arithmetic as a function of task difficulty: a near-infrared spectroscopy approach. *Frontiers in Human Neuroscience*(7), 1–9.
- Vidulich, M. A. & Tsang, P. S. (1986). Techniques of subjective workload assessment: A comparison of SWAT and the NASA-Bipolar methods. *Ergonomics*(29 (11)), 1385–1398.
- Vidulich, M. A. & Wickens, C. D. (1986). Causes of dissociation between subjective workload measures and performance. Caveats for the use of subjective assessments. *Applied Ergonomics*(17), 291–296.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Hrsg.), *Varieties of attention* (S. 63–102). London: Academic Press.
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159–177.
- Wickens, C. D., Hollands, J. G., Banbury, S. & Parasuraman, R. (2013). *Engineering psychology and human performance* (Fourth edition Aufl.). Boston: Pearson.
- Wilkinson, W. A. & Cox, M. D. (1996). Discrete Wavelet Analysis of Power System Transients. *IEEE Transactions on Power Systems*(11(4)), 2038–2044.

- Wilson, G. F. (2009). An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. *The International Journal of Aviation Psychology*(12), 3–18.
- Xu, Y. & Franconeri, S. L. (2015). Capacity for Visual Features in Mental Rotation. *Psychological Science*(26(8)), 1241–1251.
- Yeh, Y. Y. & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*(30), 111–120.
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*(18), 459–482.
- Zhang, J., Yin, Z. & Wang, R. (2015). Recognition of Mental Workload Levels Under Complex Human–Machine Collaboration by Using Physiological Features and Adaptive Support Vector Machines. *IEEE Transactions on Human-Machine Systems*, 45(2), 200–214.
- Zhang, Y., Zheng, H., Duan, Y., Meng, L. & Zhang, L. (2015). An Integrated Approach to Subjective Measuring Commercial Aviation Pilot Workload. *Proceedings of IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, 1093–1098.

Anhang

Studie 1: ANOVA der k-Means Analyse

	ANOVA					
	Cluster		Fehler		F	Sig.
	Mittel der Quadrate	df	Mittel der Quadrate	df		
VP1	,036	7	,002	32	22,444	,000
VP2	,047	7	,004	32	11,299	,000
VP3	,113	7	,003	32	34,659	,000
VP4	,050	7	,002	32	21,867	,000
VP5	,087	7	,003	32	32,832	,000
VP6	,073	7	,005	32	13,722	,000
VP7	,023	7	,002	32	9,132	,000
VP8	,027	7	,002	32	13,338	,000
VP9	,017	7	,003	32	6,734	,000
VP10	,086	7	,005	32	16,153	,000
VP11	,081	7	,004	32	21,336	,000
VP12	,045	7	,002	32	20,366	,000
VP13	,027	7	,004	32	6,468	,000
VP14	,065	7	,005	32	12,438	,000
VP15	,068	7	,003	32	23,487	,000
VP16	,028	7	,001	32	19,013	,000
VP17	,015	7	,004	32	3,647	,005
VP19	,039	7	,006	32	6,691	,000
VP20	,078	7	,003	32	30,305	,000
VP21	,027	7	,002	32	13,160	,000
VP22	,054	7	,002	32	26,091	,000
VP23	,030	7	,002	32	15,860	,000
VP25	,038	7	,003	32	11,777	,000
VP26	,008	7	,003	32	2,311	,050
VP28	,025	7	,003	32	9,384	,000
VP29	,055	7	,002	32	25,587	,000
VP30	,122	7	,004	32	30,348	,000
VP31	,046	7	,005	32	10,126	,000
VP32	,061	7	,003	32	19,589	,000
VP34	,062	7	,004	32	14,370	,000
VP35	,016	7	,003	32	4,761	,001
VP36	,047	7	,002	32	23,385	,000
VP38	,020	7	,006	32	3,363	,008
VP39	,015	7	,005	32	2,929	,017
VP40	,061	7	,004	32	15,516	,000
VP41	,085	7	,003	32	30,730	,000
VP42	,032	7	,004	32	7,253	,000
VP43	,054	7	,006	32	9,135	,000
VP44	,034	7	,003	32	11,589	,000
VP46	,049	7	,003	32	15,638	,000
VP47	,005	7	,001	32	4,421	,002
VP48	,020	7	,004	32	4,798	,001

Die F-Tests sollten nur für beschreibende Zwecke verwendet werden, da die Cluster so gewählt wurden, daß die Differenzen zwischen Fällen in unterschiedlichen Clustern maximiert werden. Dabei werden die beobachteten Signifikanzniveaus nicht korrigiert und können daher nicht als Tests für die Hypothese der Gleichheit der Clustermittelwerte interpretiert werden.

Studie 3: t-Test Lkw-Aufgabe

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	99,86% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	TLXGesamt.8,0: TLX Gesamt - TLXGesamt.9,0: TLX Gesamt	-2,2111	2,7510	,4585	-3,8021	-,6201	-4,823	35	,000
Paaren 2	TLXGesamt.8,0: TLX Gesamt - TLXGesamt.10,0: TLX Gesamt	-4,2667	4,1854	,6976	-6,6873	-1,8460	-6,116	35	,000
Paaren 3	TLXGesamt.9,0: TLX Gesamt - TLXGesamt.10,0: TLX Gesamt	-2,0556	3,7665	,6278	-4,2339	,1228	-3,274	35	,002

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	99,86% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	LKW Fehler.8,0: LKW Fehler - LKW Fehler.9,0: LKW Fehler	-1,6667	3,1623	,5270	-3,4956	,1622	-3,162	35	,003
Paaren 2	LKW Fehler.8,0: LKW Fehler - LKW Fehler.10,0: LKW Fehler	-6,5833	4,2987	,7164	-9,0695	-4,0972	-9,189	35	,000
Paaren 3	LKW Fehler.9,0: LKW Fehler - LKW Fehler.10,0: LKW Fehler	-4,9167	5,3632	,8939	-8,0185	-1,8148	-5,500	35	,000

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	99,86% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	LKW Durchschnittsdauer Fehler.8,0: LKW Durchschnittsdauer Fehler - LKW Durchschnittsdauer Fehler.9,0: LKW Durchschnittsdauer Fehler	-1,36944	1,69577	,28263	-2,35019	-,38870	-4,845	35	,000
Paaren 2	LKW Durchschnittsdauer Fehler.8,0: LKW Durchschnittsdauer Fehler - LKW Durchschnittsdauer Fehler.10,0: LKW Durchschnittsdauer Fehler	-3,26417	1,46406	,24401	-4,11091	-2,41743	-13,377	35	,000
Paaren 3	LKW Durchschnittsdauer Fehler.9,0: LKW Durchschnittsdauer Fehler - LKW Durchschnittsdauer Fehler.10,0: LKW Durchschnittsdauer Fehler	-1,89472	2,26186	,37698	-3,20287	-,58658	-5,026	35	,000

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	99,86% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	LKW Fehlerquote NA 8,0: LKW Fehlerquote NA - LKW Fehlerquote NA 9,0: LKW Fehlerquote NA	-,093046801	,1641327661	,0273554610	-,187972906	,0018793039	-3,401	35	,002
Paaren 2	LKW Fehlerquote NA 8,0: LKW Fehlerquote NA - LKW Fehlerquote NA 10,0: LKW Fehlerquote NA	-,137136541	,1883173971	,0313862329	-,246049815	-,028223266	-4,369	35	,000
Paaren 3	LKW Fehlerquote NA 9,0: LKW Fehlerquote NA - LKW Fehlerquote NA 10,0: LKW Fehlerquote NA	-,044089740	,1464145961	,0244024327	-,128768550	,0405690705	-1,807	35	,079

Studie 3: t-Test Vergleich Kalibrierung 1 und 2

Test bei gepaarten Stichproben

		Gepaarte Differenzen				T	df	Sig. (2-seitig)	
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz				
					Untere				Obere
Paaren 1	K1TLXGesamt.1,0: K1 TLXGesamt - K2TLXGesamt.1,0: K2 TLXGesamt	2,0000	2,3186	,3864	1,2155	2,7845	5,175	35	,000
Paaren 2	K1TLXGesamt.2,0: K1 TLXGesamt - K2TLXGesamt.2,0: K2 TLXGesamt	2,0056	2,2000	,3667	1,2612	2,7499	5,470	35	,000
Paaren 3	K1TLXGesamt.3,0: K1 TLXGesamt - K2TLXGesamt.3,0: K2 TLXGesamt	1,6556	2,5806	,4301	,7824	2,5287	3,849	35	,000
Paaren 4	K1TLXGesamt.4,0: K1 TLXGesamt - K2TLXGesamt.4,0: K2 TLXGesamt	1,7556	2,8151	,4692	,8031	2,7081	3,742	35	,001
Paaren 5	K1Leistung/Nebenaufgabe e.1,0: K1 Leistung Nebenaufgabe - K2Leistung/Nebenaufgabe e.1,0: K2 Leistung Nebenaufgabe	,0264550265	,1137996719	,0189666120	-,012049243	,0649592958	1,395	35	,172
Paaren 6	K1Leistung/Nebenaufgabe e.2,0: K1 Leistung Nebenaufgabe - K2Leistung/Nebenaufgabe e.2,0: K2 Leistung Nebenaufgabe	,0810185185	,1470424179	,0245070697	,0312665221	,1307705149	3,306	35	,002
Paaren 7	K1Leistung/Nebenaufgabe e.3,0: K1 Leistung Nebenaufgabe - K2Leistung/Nebenaufgabe e.3,0: K2 Leistung Nebenaufgabe	,0350855776	,0566250555	,0094375092	,0159264153	,0542447399	3,718	35	,001
Paaren 8	K1Leistung/Nebenaufgabe e.4,0: K1 Leistung Nebenaufgabe - K2Leistung/Nebenaufgabe e.4,0: K2 Leistung Nebenaufgabe	,0811271164	,1453209116	,0242201519	,0319575940	,1302966389	3,350	35	,002

Die Berücksichtigung des kognitiven Zustandes und des Verhaltens von Anwendern ist eine der wichtigsten Voraussetzungen für die Gestaltung menschenzentrierter Schnittstellen und Arbeitssysteme. Dementsprechend liegt ein großer Schwerpunkt der aktuellen arbeitswissenschaftlichen Forschung auf der zuverlässigen Erfassung und Interpretation des kognitiven und physiologischen Zustandes von Anwendern in der Interaktion mit dem zu gestaltenden, technischen System. Im Rahmen dieser Arbeit wird ein Verfahren entwickelt, mit dessen Hilfe der kognitive Zustand, operationalisiert über die individuelle Leistungsfähigkeit und das subjektive Empfinden, allein unter Verwendung okularer Parameter, kontinuierlich erfasst und quantifiziert werden kann. Durch die Verwendung einer Kalibrierungsaufgabe vor der zu analysierenden Tätigkeit selbst, wird ein individuelles Beanspruchungsmodell erstellt, um die Schnittstelle anhand des kognitiven Zustandes und des Anwenderverhaltens zu optimieren.



ISSN 2363-7315

ISBN 978-3-7315-0809-0

Gedruckt auf FSC-zertifiziertem Papier