

Psychophysiologische Profile in nutzerzentrierten Mensch-Maschine-Systemen: Extraktion kardialer und elektrodermalen Profile zur Bewertung der mentalen Beanspruchung

Zur Erlangung des akademischen Grades einer

**DOKTORIN DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)**

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

angenommene

Dissertation

von

Dipl.-Psych. Elena Wolf

Tag der mündlichen Prüfung:

24.09.2021

Hauptreferentin:

Prof. Dr.-Ing. Barbara Deml

Korreferent:

Prof. Dr.-Ing. Rainer Stiefelhagen

Zusammenfassung

Nutzerzentrierte Mensch-Maschine-Systemen sind darauf ausgerichtet, sich an die jeweiligen Anforderungen, Bedürfnisse und Gefühlszustände der nutzenden Person anzupassen. Dabei sollen adaptive Assistenzsysteme dem Menschen dann Unterstützung anbieten, sobald die kognitiven Ressourcen der Person an ihre Grenzen stoßen. Die kognitive Ressourcenauslastung wird durch das Konstrukt der mentalen Beanspruchung repräsentiert. Mit der physiologischen Messmethode kann die mentale Beanspruchung kontinuierlich und störungsfrei an der Schnittstelle zwischen Mensch und Maschine erfasst werden. Welche der zahlreichen physiologischen Parameter die mentale Beanspruchung zuverlässig abbilden, ist bis heute nicht eindeutig geklärt. Um valide von einer gemessenen physiologischen Aktivität auf die mentale Beanspruchung einer Person schließen zu können, fordern Cacioppo und Kollegen (Cacioppo & Tassinary, 1990; Cacioppo, Tassinary & Berntson, 2000, 2007, 2017) in ihrem theoretischen Rahmenmodell die Neustrukturierung physiologischer Einzelparameter zu physiologischen Profilen. Eine solche Neustrukturierung physiologischer Parameter wird in der vorliegenden Arbeit mit Hilfe von drei empirischen Studien umgesetzt. Hierfür werden mittels mehrdimensionaler Analyseverfahren kardiale und elektrodermale Profile aus den kardialen und elektrodermalen Einzelparametern abgeleitet. Diese sagen die mentale Beanspruchung nicht nur bedeutsam und spezifisch vorher, sondern sind auch über verschiedene mentale Belastungsfaktoren generalisierbar. Die physiologischen Profile werden in die Taxonomie von Cacioppo et al. (2000, 2007, 2017) eingeordnet und die Potentiale bei der Verwendung der Profile in einer nutzerzentrierten Mensch-Maschine-Schnittstelle dargelegt.

Abstract

User-centered human-machine systems are oriented on the user's requirements, needs and emotional states. In this context, adaptive assistance systems are designed to offer personal support as soon as the imposed task demands exceed the person's cognitive resources. The imposed task demands on the person's information processing resources are represented by the construct of mental workload. To identify a person's mental workload, the physiological measuring method can be used. With this method the mental workload can be measured continuously and without interference at the interface between human and machine. There are numerous physiological parameters which are discussed as potential indicators for mental workload. However, it has not yet been conclusively defined which of these physiological parameters can be used as valid representors of mental workload. According to the framework of Cacioppo and colleagues (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017) single physiological parameters should be restructured into physiological profiles to validly infer a person's mental workload from the physiological measure. In the present thesis, such a restructuring of physiological parameters is performed on the basis of three empirical studies. This is done by deriving cardiac and electrodermal profiles from single cardiac and electrodermal parameters using multidimensional analysis methods. The resulting profiles predict a person's mental workload significantly and specifically, but can also be generalized across different mental workload domains. Finally, the physiological profiles are mapped to the taxonomy of Cacioppo et al. (2000, 2007, 2017) and the potentials of using the profiles in a user-centered human-machine interface are outlined.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, um mich bei all denjenigen Menschen zu bedanken, die wesentlich zum Gelingen der vorliegenden Arbeit beigetragen haben.

Zunächst möchte ich Frau Prof. Dr.-Ing. Barbara Deml meinen herzlichen Dank aussprechen für die fachliche Betreuung und motivationale Unterstützung bei der Entstehung dieser Arbeit. Insbesondere möchte ich mich für das Vertrauen an die hervorgebrachten Forschungsideen und für die Förderung bei der Umsetzung der Forschungsbestrebungen bedanken. Mein herzlicher Dank gebührt auch Herrn Prof. Dr.-Ing. Rainer Stiefelhagen für die Übernahme des Zweitgutachtens und für die wertvollen Impulse bei der gemeinsamen Projektarbeit im Rahmen des Projekts KonsensOP.

Ein großes Dankeschön gebührt allen Kolleginnen und Kollegen des ifab. Der fachliche Austausch und die kollegiale Unterstützung waren bereichernd und einzigartig. Ihr habt nicht nur meine Forschungsarbeit, sondern auch mich positiv geprägt. Persönlich bedanken möchte ich mich bei Fabian Ries, Jonas Imbsweiler und Katrin Linstedt. Danke für eure kreativen Ideen, für eure großzügige Unterstützung und vor allem für den Spaß, den ihr in den Alltag gebracht habt. Danke für eure Freundschaft. Ein großes Dankeschön möchte ich auch allen wissenschaftlichen Hilfskräften aussprechen. Insbesondere Marco Wetter und Markus Döllken haben durch ihre Zuarbeit wesentlich zum Erfolg dieser Arbeit beigetragen.

Bedanken möchte ich mich zudem bei allen Kolleginnen und Kollegen des Projekts KonsensOP. Danke, dass ich so viel für meine Forschungsarbeit

von euch lernen durfte und dafür, dass trotz vorkommender Hindernisse der Zusammenhalt und das kollegiale Engagement großartig waren.

Ein weiteres persönliches Dankeschön gilt meinem Partner Max Spindler für die ununterbrochene Unterstützung in den letzten Jahren. Dein kritisches Korrekturlesen und deine effektiven Lösungsvorschläge haben diese Arbeit wesentlich mit beeinflusst. Du bist der Beste! Auch bei Georg Wolf und Maria Parfentev bedanke ich mich herzlichst für das Korrekturlesen.

Karlsruhe, im April 2021

Elena Wolf

Inhaltsverzeichnis

Zusammenfassung	i
Abstract	iii
Danksagung	v
Inhaltsverzeichnis	vii
Abbildungsverzeichnis	xvii
Tabellenverzeichnis	xxvii
Abkürzungen und Symbole	xxxv
Anmerkungen	xliii
1 Einleitung	1
2 Theoretische Grundlagen: Mentale Beanspruchung	7
2.1 Abgrenzung theoretischer Begriffe	7
2.1.1 Belastung und Beanspruchung	8
2.1.2 Kognitive und mentale Beanspruchung	10
2.1.3 Mentale und emotionale Beanspruchung	11
2.1.4 Beanspruchung und Stress	13
2.2 Modelle mentaler Beanspruchung	14
2.2.1 Mentale Beanspruchung in ressourcentheoretischen Modellen	16

2.2.2	Mentale Beanspruchung in aktivierungstheoretischen Modellen	19
2.3	Zusammenfassung und Eingrenzung	23
3	Theoretische Grundlagen: Psychophysiologische Indikatoren	27
3.1	Regulation durch das autonome Nervensystem	28
3.2	Kardiale Aktivität	32
3.2.1	Physiologische Grundlagen kardialer Aktivität	32
3.2.2	Messung kardialer Aktivität	35
3.2.3	Quantifizierung des kardialen Signals	37
3.2.4	Befundlage zur kardialen Aktivität	42
3.3	Elektrodermale Aktivität	45
3.3.1	Physiologische Grundlagen elektrodermalen Aktivität	45
3.3.2	Messung elektrodermalen Aktivität	47
3.3.3	Quantifizierung des elektrodermalen Signals	48
3.3.4	Befundlage zur elektrodermalen Aktivität	50
3.4	Multivariate Betrachtung psychophysiologischer Indikatoren	52
3.4.1	Psychophysiological Computing	54
3.4.2	Schwächen des Psychophysiological Computing	58
3.5	Zusammenfassung	60
4	Herleitung der Forschungsfragen	63
4.1	Rückschlüsse von der physiologischen Aktivität auf die mentale Beanspruchung	63
4.1.1	Rahmenmodell psychophysiologischer Zusammenhänge	65
4.1.2	Herleitung eines Eins-zu-eins-Zusammenhangs	68
4.1.3	Taxonomie psychophysiologischer Zusammenhänge	69
4.1.4	Überlappung und Abgrenzung zur psychometrischen Validität	73
4.2	Extraktion physiologischer Profile mentaler Beanspruchung	75

4.2.1	Dimensionsreduzierende Verfahren in der psychophy- siologischen Forschung	76
4.2.2	Methodische Herausforderungen dimensionsreduzie- render Verfahren	78
4.3	Forschungsziele und Forschungsprogramm	81
5	Methode	89
5.1	Messung der kardialen Aktivität	89
5.1.1	Apparatur zur Messung der kardialen Aktivität . . .	89
5.1.2	Aufbereitung und Parametrisierung der kardialen Mess- daten	90
5.2	Messung der elektrodermalen Aktivität	96
5.2.1	Apparatur zur Messung der elektrodermalen Aktivität	96
5.2.2	Aufbereitung und Parametrisierung der elektroderma- len Messdaten	97
5.3	Störeinflüsse und Kontrollvariablen physiologischer Messung	102
5.4	Messung der subjektiven mentalen Beanspruchung	103
5.4.1	Der National Aeronautics and Space Administration Task Load Index (NASA-TLX)	104
5.5	Messung der leistungsbezogenen mentalen Beanspruchung .	106
5.6	Statistische Auswertemethoden	109
5.6.1	Mehrebenenanalyse	110
5.6.2	Statistische Methoden zur Ableitung psychophysiologi- scher Profile	111
6	Studie I	123
6.1	Forschungsziel der Studie I	123
6.2	Methode	124
6.2.1	Operationalisierung der mentalen und emotionalen Beanspruchung	125
6.2.2	Versuchsdesign	131
6.2.3	Versuchsablauf	134

6.2.4	Abschlussbefragung und Kontrollvariablen	135
6.2.5	Manipulationskontrolle	135
6.2.6	Stichprobe	139
6.3	Ergebnisse	140
6.3.1	Kontrollfragestellung KF1: Ist die experimentelle Induktion mentaler und emotionaler Beanspruchung in Studie I gelungen?	140
6.3.2	Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	155
6.3.3	Forschungsfrage F1.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?	161
6.3.4	Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	165
6.3.5	Forschungsfrage F2.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?	168
6.3.6	Forschungsfrage F3.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung zuverlässig repräsentieren?	171
6.3.7	Forschungsfrage F3.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Valenz zuverlässig repräsentieren?	174

6.3.8	Forschungsfrage F4.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung zuverlässig repräsentieren? . . .	177
6.3.9	Forschungsfrage F4.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Valenz zuverlässig repräsentieren?	180
6.3.10	Forschungsfrage F5.1: Können die kardialen Profile der mentalen und emotionalen Beanspruchung bedeutsam voneinander abgegrenzt werden?	183
6.3.11	Forschungsfrage F5.2: Können die elektrodermalen Profile der mentalen und emotionalen Beanspruchung bedeutsam voneinander abgegrenzt werden?	188
6.4	Diskussion	191
6.4.1	Zusammenfassung und Interpretation der Manipulationskontrolle	191
6.4.2	Zusammenfassung und Interpretation der kardialen Profile	195
6.4.3	Zusammenfassung und Interpretation der elektrodermalen Profile	200
6.4.4	Limitationen der Studie	204
7	Studie II	209
7.1	Forschungsziel der Studie II	209
7.2	Methode	210
7.2.1	Operationalisierung der mentalen Beanspruchung . .	210
7.2.2	Versuchsdesign	213
7.2.3	Versuchsablauf	213
7.2.4	Abschlussbefragung und Kontrollvariablen	214
7.2.5	Manipulationskontrolle	214
7.2.6	Stichprobe	216

7.3	Ergebnisse	217
7.3.1	Kontrollfragestellung KF2: Ist die experimentelle Induktion mentaler Beanspruchung in Studie II gelungen?	218
7.3.2	Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	223
7.3.3	Forschungsfrage F1.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?	226
7.3.4	Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	230
7.3.5	Forschungsfrage F2.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?	233
7.3.6	Forschungsfrage F6.1: Sind die in Studie II extrahierten kardialen Profile mentaler Beanspruchung kongruent zu den in Studie I extrahierten Profilen mentaler Beanspruchung?	236
7.3.7	Forschungsfrage F6.2: Sind die in Studie II extrahierten elektrodermalen Profile mentaler Beanspruchung kongruent zu den in Studie I extrahierten Profilen mentaler Beanspruchung?	239
7.4	Diskussion	242
7.4.1	Zusammenfassung und Interpretation der Manipulationskontrolle	242
7.4.2	Zusammenfassung und Interpretation der kardialen Profile	243

7.4.3	Zusammenfassung und Interpretation der elektrodermalen Profile	246
7.4.4	Limitationen	250
8	Studie III	253
8.1	Forschungsziel der Studie III	253
8.2	Methode	254
8.2.1	Operationalisierung der mentalen Beanspruchung	255
8.2.2	Versuchsdesign	266
8.2.3	Versuchsablauf	267
8.2.4	Abschlussbefragung und Kontrollvariablen	268
8.2.5	Manipulationskontrolle	268
8.2.6	Stichprobe	270
8.3	Ergebnisse	271
8.3.1	Kontrollfragestellung KF3: Ist die experimentelle Induktion mentaler Beanspruchung in Studie III gelungen?	272
8.3.2	Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	276
8.3.3	Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?	279
8.3.4	Forschungsfrage F7.1: Sind die aufgabenübergreifenden kardialen Profile mentaler Beanspruchung der Studie III kongruent zu den kardialen Profilen mentaler Beanspruchung der Studien I und II?	283
8.3.5	Forschungsfrage F7.2: Sind die aufgabenübergreifenden elektrodermalen Profile mentaler Beanspruchung der Studie III kongruent zu den elektrodermalen Profilen mentaler Beanspruchung der Studien I und II?	286

8.4	Diskussion	289
8.4.1	Zusammenfassung und Interpretation der Manipulationskontrolle	289
8.4.2	Zusammenfassung und Interpretation der kardialen Profile	290
8.4.3	Zusammenfassung und Interpretation der elektrodermalen Profile	293
8.4.4	Limitationen	295
9	Psychophysiologische Profile in der beanspruchungssensitiven Mensch-Maschine-Schnittstelle	299
9.1	Psychometrische Bewertung der extrahierten Profile	299
9.1.1	Bewertung kardialer Profile	301
9.1.2	Bewertung kardialer Einzelparameter	308
9.1.3	Bewertung elektrodermalen Profile	312
9.1.4	Bewertung elektrodermalen Einzelparameter	319
9.2	Theoretische Verortung der extrahierten Profile	322
9.3	Anwendungsmöglichkeiten der physiologischen Profile in der Mensch-Maschine-Schnittstelle	326
9.4	Ausblick	331
9.5	Fazit	333
	Literaturverzeichnis	335
A	Anhang	363
A.1	Theoretische Verortung des Konstrukts Emotionen	363
A.2	Güte psychophysiologischer Parameter	365
A.2.1	Sensitivität	366
A.2.2	Diagnostizität bzw. Spezifität	367
A.2.3	Reliabilität	368
A.2.4	Validität	368
A.3	Psychometrische Reliabilität und Validität des NASA-TLX	372
A.4	Voraussetzungen der Mehrebenenanalyse	372

A.5	Unterschiede zwischen einer PARAFAC und einer generalisier-	
	ten PCA	373
A.6	Anhänge zu Studie I	374
	A.6.1 Deskriptive Statistik der kardialen Parameter	374
	A.6.2 Deskriptive Statistik der elektrodermalen Parameter .	375
	A.6.3 Deskriptive Statistik der Kriteriumsvariablen	376
	A.6.4 Bedingungs-Faktorladungen der NPLS-Regression .	379
A.7	Anhänge zu Studie II	383
	A.7.1 Deskriptive Statistik der kardialen Parameter	383
	A.7.2 Deskriptive Statistik der elektrodermalen Parameter .	384
	A.7.3 Deskriptive Statistik der Kriteriumsvariablen	385
	A.7.4 Bedingungs-Faktorladungen der NPLS-Regression .	387
A.8	Anhänge zu Studie III	391
	A.8.1 Deskriptive Statistik der kardialen Parameter	391
	A.8.2 Deskriptive Statistik der elektrodermalen Parameter .	392
	A.8.3 Deskriptive Statistik der Kriteriumsvariable	393
	A.8.4 Analyse der leistungsbezogenen mentalen Beanspru-	
	chung	394
	A.8.5 Bedingungs-Faktorladungen der NPLS-Regression .	401
A.9	Vorhersage der mentalen Beanspruchung	403

Abbildungsverzeichnis

1.1	Dargestellt sind die Inhalte der vorliegenden Arbeit als schematischer Aufbau.	4
2.1	Beanspruchung und Belastung dargestellt als Modell der mechanischen Feder (modifiziert nach Schlick, Bruder & Luczak, 2010, S. 39).	9
2.2	Darstellung des Multiple-Ressourcen-Modells als Würfel (modifiziert nach Wickens, Hollands, Banbury & Parasuraman, 2013, S. 330).	18
3.1	Die zweidimensionale Ebene des Autonomic Space Modells. Je nach Aktivierung des sympathischen und parasympathischen Systems kann sich eine Entkopplung, Koaktivierung oder Reziprozität ergeben. Grafik aus Berntson, Cacioppo und Quigley (1991, S. 469).	30
3.2	Frontalansicht des menschlichen Herzens. Der Verlauf des Blutstroms ist durch die Pfeile gekennzeichnet. Rot repräsentiert sauerstoffreiches und blau sauerstoffarmes Blut. Grafik aus Feigenspan (2017, S. 251).	32
3.3	Erregungsausbreitung eines Herzzyklus im Elektrokardiogramm (EKG). Grafik modifiziert nach Feigenspan (2017, S. 280). . . .	34
3.4	Ableitung der kardialen Aktivität an den Extremitäten nach Einthoven (a) und an der Brustwand nach Wilson (b). V_{LA} = Potential des linken Arms; V_{RA} : Potential des rechten Arms; V_{LL} : Potential des linken Beins. Grafik modifiziert nach Gupta, Mitra und Bera (2014, S. 8–9).	35

3.5	Der HRV-Dreiecksindex (Triang) ergibt sich aus der Division der Dichteverteilung der NN-Intervalle (D) mit dem Maximum der Dichteverteilung (Y). Grafik modifiziert nach Malik et al. (1996, S. 357).	39
3.6	Spektrale Leistungsdichte (SLD) der Herzperiode. Grafik modifiziert nach Berntson, Quigley, Norman und Lozano (2017, S. 196).	40
3.7	Schematischer Ausschnitt der Haut und der ekkrinen Schweißdrüse. Grafik modifiziert nach Boucsein (2001, S. 13).	46
3.8	Exemplarischer Verlauf einer EDR und der daraus abgeleiteten Parameter. Grafik aus Boucsein (2012, S. 154).	48
3.9	Typische Analyseschritte des Psychophysiological Computings. Grafik modifiziert nach Novak, Mihelj und Munih (2012, S. 155).	55
4.1	Gemäß dem Rahmenmodell von Cacioppo et al. (siehe Cacioppo & Tassinary, 1990; Cacioppo et al., 2007) ergeben sich fünf allgemeine Zusammenhangsmuster zwischen den psychologischen Elementen in der Teilmenge Ψ und den physiologischen Elementen in der Teilmenge Φ . Grafik modifiziert nach Berntson und Cacioppo (2002, S. 12).	66
4.2	Ein Viele-zu-viele-Zusammenhang kann durch die Neustrukturierung der Elemente aus Φ zu einem Profil Φ' in einen Viele-zu-eins-Zusammenhang und durch die Neustrukturierung der Elemente aus Φ' zu einem Profil Φ'' in einen Eins-zu-eins-Zusammenhang umgewandelt werden.	68
4.3	Die Gütekriterien Spezifität, Generalisierbarkeit und Sensitivität bilden die drei Achsen eines Oktaeders. Auf der horizontalen Ebene werden die vier Zusammenhangsklassen Auswirkung, Marker, Begleiterscheinung und Invariante in Abhängigkeit von den Kriterien Spezifität und Generalisierbarkeit abgegrenzt. Grafik modifiziert nach Cacioppo et al. (2007, S. 11).	70
4.4	Die PCA fasst beobachtbare Parameter (Rechtecke) zu nicht beobachtbaren Komponenten (Ovale) zusammen.	77

4.5	Projektion der Komponentenlösung auf das psychologische Konstrukt der mentalen Beanspruchung. Der empirische Zusammenhang zwischen der gemessenen mentalen Beanspruchung und der Komponente 1 ist maximal.	81
4.6	Einordnung der Forschungsfragen F in die Studien I, II und III. .	87
5.1	Beispiel für ein EKG-Signal, entnommen aus Studie III über ein Zeitintervall von 60 s. Oben: Gefiltertes EKG-Signal mit den detektierten R-Zacken (magenta Linien). Mitte: Herzschlagfrequenz über die Zeit als Anzahl der Schläge pro Minute. Unten: Übereinander gelegte Abschnitte des EKG-Signal bestehend aus P- und T-Wellen sowie QRS-Komplex.	91
5.2	Beispiel für ein EDA-Signal, entnommen aus Studie III über ein Zeitintervall von 60 s. Oben: Gefiltertes EDA-Signal als Hautleitwert. Mitte: Tonische Komponente des EDA-Signals. Unten: Phasische Komponente des EDA-Signals mit detektierten EDRs. Amplitudenmaximum eines EDR ist als vertikale rote Linie dargestellt.	98
5.3	Berechnung der Fläche unter der EDR nach Traxel (1957 zitiert nach Boucsein, 2012, S. 165). Grafik aus Boucsein (2012, S. 164).	101
5.4	Bei der Partial Least Squares Regression (PLS) werden die Prädiktormatrix \mathbf{X} und die Kriteriumsmatrix \mathbf{Y} in Faktorwerte und Faktorladungen zerlegt. Die Prämisse ist dabei die Maximierung der Kovarianz zwischen \mathbf{X} und \mathbf{Y}	112
5.5	Die trilineare Zerlegung der PARAFAC.	114
5.6	Die NPLS-Regression als Erweiterung der PLS auf mehrdimensionale Datenstrukturen.	115
6.1	Balltrackingaufgabe der Studie I mit grauem Spielfeld. Das Ziel der Aufgabe ist es, den Zielball (blauer Kreis) mit der dem Mauszeiger (weißer Kreis) zu verfolgen und zu treffen. Die verbleibende Zeit wurde oben links im Spielfeld angezeigt.	125

6.2	Balltrackingaufgabe mit Leistungsfeedback. Der Punktestand (Treffer) wird oben in der Bildschirmmitte angezeigt. Links: Negatives Leistungsfeedback mit rotem Bildschirmhintergrund. Rechts: Positives Leistungsfeedback mit gelbem Bildschirmhintergrund. . . .	127
6.3	Dargestellt ist der SAM bestehend aus den Skalen Valenz (oben), Erregung (Mitte) und Dominanz (unten). Grafik aus Bradley und Lang (1994).	131
6.4	Versuchsdesign mit den Blöcken 1 bis 3. Innerhalb der Blöcke sind die Bedingungen randomisiert.	132
6.5	Schematische Darstellung des Versuchsablaufs aus Studie I. . .	133
6.6	Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.	142
6.7	Mittelwert und Standardabweichung der Trefferquote in Prozent in den Versuchsbedingungen.	147
6.8	Mittelwert und Standardabweichung der Erregung in den Versuchsbedingungen.	149
6.9	Mittelwert und Standardabweichung der Valenz in den Versuchsbedingungen.	152
6.10	Faktorladungen a für das NASA-RTLX und die kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	157
6.11	Faktorladungen a für die Trefferquote und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.	163
6.12	Faktorladungen a für das NASA-RTLX und die elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	167

6.13	Faktorladungen a für die Trefferquote und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.	169
6.14	Faktorladungen a für die Erregung und die kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Erregung in den Bedingungen (Mitte) angegeben.	172
6.15	Faktorladungen a für die Valenz und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Valenz in den Bedingungen (Mitte) angegeben.	175
6.16	Faktorladungen a für die Erregung und die elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Erregung in den Bedingungen (Mitte) angegeben.	178
6.17	Faktorladungen a für die Valenz und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Valenz in den Bedingungen (Mitte) angegeben.	182
6.18	Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der kardialen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der kardialen Komponenten aus Studie I.	197
6.19	Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der elektrodermalen Komponenten aus Studie I.	202
7.1	Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.	219
7.2	Mittelwert und Standardabweichung der Trefferquote in Prozent in den Versuchsbedingungen.	221

7.3	Faktorladungen a für die NASA-RTLX und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	225
7.4	Faktorladungen a für die Trefferquote und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.	228
7.5	Faktorladungen a für die NASA-RTLX und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	232
7.6	Faktorladungen a für die Trefferquote und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.	235
7.7	Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der kardialen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der kardialen Komponenten aus Studie II.	244
7.8	Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der elektrodermalen Komponenten aus Studie II.	248
8.1	Dargestellt sind Ausschnitte aus einer Balltrackingaufgabe mit einem Distraktorball. Der Mauszeiger ist als weißer Kreis dargestellt. Links: Ohne Zielkontakt haben der Ziel- und Distraktorball dieselbe blaue Farbe. Rechts: Bei Zielkontakt leuchtet der Zielball hellblau auf.	256
8.2	Dargestellt sind Ausschnitte aus einer Gedächtnisaufgabe der Schwierigkeitsstufe G2: Präsentation einer Ziffer (links), Eingabemaske für die Antwort (Mitte) und Leistungsrückmeldung (rechts).	258

8.3	Dargestellt sind Ausschnitte aus der Konzentrationsaufgabe: Präsentation eines Zielobjekts (links) und Leistungsrückmeldung bei inkorrektter Reaktion (rechts).	260
8.4	Leichte Matrize (oben): ein Objekt pro Matrizelement und eine einzige Regel (jedes Objekt ist ein Mal in einer Zeile oder Spalte vertreten). Schwere Matrize (unten): mehrere Objekten pro Matrizelement (Stern mit Punkten und Gerade) und separaten Objektregeln (Stern mit Punkten drehen sich im Uhrzeigersinn und die Gerade dreht sich gegen den Uhrzeigersinn). Die Lösung ist im rechten Auswahlbereich in Grün markiert.	262
8.5	Dargestellt sind Ausschnitte aus einer schweren Matrizenaufgabe: Präsentation der Matrize (links), Auswahl eines Antwortelements (Mitte) und Leistungsrückmeldung bei inkorrektter Lösungsantwort (rechts).	263
8.6	Beispiel für Matrizen aus der Ruhebedingung (M1).	263
8.7	Links: Übereinstimmung zwischen Farbwort und Schriftfarbe (Kongruenz). Rechts: Keine Übereinstimmung zwischen Farbwort und Schriftfarbe (Inkongruenz).	265
8.8	Dargestellt sind Ausschnitte aus der Stroopaufgabe mit der Schwierigkeitsstufe S3: Präsentation des inkongruenten Farbwords mit Antwortfeld (links) und Leistungsrückmeldung bei korrekter Reaktion (rechts).	265
8.9	Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.	273
8.10	Faktorladungen a für das NASA-RTLX und die kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	278
8.11	Faktorladungen a für das NASA-RTLX und die elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.	281

8.12	Zusammenhang zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente und der kardialen Komponente der Studie III.	291
8.13	Zusammenhang zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente und der elektrodermalen Komponente der Studie III.	294
9.1	Strukturmodell für den Zusammenhang der kardialen Profile mit den Konstrukten der mentalen Beanspruchung (hellblau; repräsentiert durch das NASA-RTLX der Studien I bis III und durch die Trefferquote der Studie II) und emotionalen Beanspruchung (lila; repräsentiert durch die Emotionsdimension Valenz der Studie I). Die Höhe des Zusammenhangs ist am Regressionsgewicht B bzw. dessen Durchschnitt \bar{B} und am Anteil erklärter Varianz R_{PLS}^2 bzw. dessen Durchschnitt \bar{R}_{ns}^2 ablesbar (val bezieht sich auf das kreuzvalidierte R_{PLS}^2). Die Profile sind mit den für sie bedeutsamen oder marginalen Parametern verbunden. Profilspezifische Parameter sind farblich dargestellt. C beschreibt die Kongruenz der Profile. . . .	307
9.2	Strukturmodell für den Zusammenhang der elektrodermalen Profile mit den Konstrukten der mentalen Beanspruchung (hellblau; repräsentiert durch das NASA-RTLX der Studie III) und der emotionalen Beanspruchung (lila; repräsentiert durch die Emotionsdimensionen Erregung und (invertierte) Valenz der Studie I). Die Höhe des Zusammenhangs ist am Regressionsgewicht B bzw. dessen Durchschnitt \bar{B} und am Anteil erklärter Varianz R_{PLS}^2 bzw. dessen Durchschnitt \bar{R}_{ns}^2 ablesbar (val bezieht sich auf das kreuzvalidierte R_{PLS}^2). Die gekreuzten Regressionsgewichte (grau gestrichelte Pfeile) sind nicht bekannt. Die Profile sind mit den für sie bedeutsamen oder marginalen Parametern verbunden. Profilspezifische Parameter sind farblich dargestellt. C beschreibt die Kongruenz der Profile.	317

A-1	Mittelwert und Standardabweichung der Trefferquote in den Versuchsbedingungen.	398
A-2	Vorhergesage personenspezifischer NASA-RTLX-Werte durch die elektrodermalen Parameter für 10 Personen der Studie III. A: Vorhersage durch den Mittelwert der Parameter. B: Vorhersage durch das elektrodermale Profil (Parameter-Faktorladungen). C: Vorhersage durch das elektrodermale Profil und die Bedingungs-Faktorladungen.	404

Tabellenverzeichnis

2.1	Zuordnung physiologischer Kenngrößen zu den vier Aktivierungssystemen (nach Backs & Boucsein, 2009; Boucsein, 2012, S. 462)	22
5.1	Übersicht der verwendeten kardialen Parameter. Alle Parameter beziehen sich auf ein Zeitintervall von 60 s.	92
5.2	Übersicht der verwendeten elektrodermalen Parameter. Alle Parameter beziehen sich auf ein Zeitintervall von 60 s. Alle Parameter werden aus dem Hautleitwert berechnet	99
6.1	Die Experimentalbedingungen repräsentieren die vier Kombinationen aus Leistungsfeedback und Aufgabenschwierigkeit	128
6.2	Mittelwert und Standardabweichung für jedes Bilderset aus Studie I. Die statistischen Kennzahlen wurden anhand des normierten Datensatzes aus Lang, Bradley und Cuthbert (2008) berechnet .	130
6.3	Deskriptive Statistiken der Kontrollvariablen	141
6.4	Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX in Block 1 und 2	143
6.5	Parameter für das Modell $M_{S \times F}$ des NASA-RTLX	144
6.6	Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX in Block 3	145
6.7	Parameter für das Modell M_E des NASA-RTLX	145
6.8	Parameter für das Modell M_{Block} des NASA-RTLX	146
6.9	Modellfit-Indizes für die geschachtelten Modelle der Trefferquote	147
6.10	Parameter für das Modell M_F der Trefferquote	148

6.11	Modellfit-Indizes für die geschachtelten Modelle der Erregung in Block 1 und 2	149
6.12	Parameter für das Modell $M_{S \times F}$ der Erregung	150
6.13	Modellfit-Indizes für die geschachtelten Modelle der Erregung in Block 3	151
6.14	Parameter für das Modell M_E der Erregung	151
6.15	Modellfit-Indizes für die geschachtelten Modelle der Valenz in Block 1 und 2	153
6.16	Parameter für das Modell M_F der Valenz	154
6.17	Modellfit-Indizes für die geschachtelten Modelle der Valenz in Block 3	154
6.18	Parameter für das Modell $M_{V \times E}$ der Valenz	155
6.19	NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	156
6.20	NPLS-Regression der Trefferquote auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	162
6.21	NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	166
6.22	NPLS-Regression der Trefferquote auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	168
6.23	NPLS-Regression der Erregung auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	171

6.24	NPLS-Regression der Valenz auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	174
6.25	NPLS-Regression der Erregung auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	177
6.26	NPLS-Regression der Valenz auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	181
6.27	Kongruenz zwischen der kardialen Komponente 1 des NASA-RTLX (Na), der kardialen Komponente 1 der Trefferquote (Tq), der kardialen Komponente 2 der Erregung (Er) und der kardialen Komponente 2 der Valenz (Va). Die Komponenten werden auf Ebene der Personen, Bedingungen und Parameter verglichen . .	183
6.28	Differenz zwischen den Parameter-Faktorladungen der kardialen Komponente 1 des NASA-RTLX (Na), der kardialen Komponente 1 der Trefferquote (Tq), der kardialen Komponente 2 der Erregung (Er) und der kardialen Komponente 2 der Valenz (Va)	185
6.29	Parameter-Faktorladungen a für die kardialen Profile des NASA-RTLX, der Trefferquote, der Erregung und der Valenz	186
6.30	Kongruenz zwischen der elektrodermalen Komponente 1 des NASA-RTLX (Na), der elektrodermalen Komponente 1 der Trefferquote (Tq), der elektrodermalen Komponente 1 der Erregung (Er) und der elektrodermalen Komponente 1 der Valenz (Va). Die Komponenten werden auf Ebene der Personen, Bedingungen und Parameter verglichen	188

6.31	Differenz zwischen den elektrodermalen Parameter-Faktorladungen der elektrodermalen Komponente 1 des NASA-RTLX (Na), der elektrodermalen Komponente 1 der Trefferquote (Tq), der elektrodermalen Komponente 1 der Erregung (Er) und der elektrodermalen Komponente 1 der Valenz (Va)	189
6.32	Parameter-Faktorladungen a für die elektrodermalen Profile des NASA-RTLX, der Trefferquote, der Erregung und der Valenz . .	190
6.33	Übersicht der kardialen Profile und deren Vorhersagegüte aus Studie I	196
6.34	Übersicht der elektrodermalen Profile und deren Vorhersagegüte aus Studie I	201
7.1	Deskriptive Statistiken der Kontrollvariablen	218
7.2	Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX	220
7.3	Modellfit-Indizes für die geschachtelten Modelle der Trefferquote	221
7.4	Parameter für das Modell $M_{S \times K \times B}$ der Trefferquote	222
7.5	NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	224
7.6	NPLS-Regression der Trefferquote auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	227
7.7	NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	230
7.8	NPLS-Regression der Trefferquote auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	233

7.9	Differenz zwischen den Parameter-Faktorladungen der kardialen Komponente des NASA-RTLX (Na) und der kardialen Komponente der Trefferquote (Tq) aus Studie I und II	237
7.10	Parameter-Faktorladungen a für die kardialen Komponenten des NASA-RTLX und der Trefferquote jeweils aus Studie I und II	238
7.11	Differenz zwischen den Parameter-Faktorladungen der elektrodermalen Komponente des NASA-RTLX (Na) aus Studie I und II, der Trefferquote (Tq) aus Studie II und der Valenz (Va) aus Studie I	240
7.12	Parameter-Faktorladungen a für die elektrodermalen Komponenten des NASA-RTLX (aus Studie I und II), der Trefferquote (aus Studie II) und der Valenz (aus Studie I)	241
7.13	Übersicht der kardialen Profile und deren Vorhersagegüte aus Studie II	243
7.14	Übersicht der elektrodermalen Profile und deren Vorhersagegüte aus Studie II	247
8.1	Deskriptive Statistiken der Kontrollvariablen	272
8.2	Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX	274
8.3	Parameter für das Modell ($M_{S \times A}$) des NASA-RTLX	275
8.4	NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	277
8.5	NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)	280
8.6	Parameter-Faktorladungen a für die kardialen Komponenten des NASA-RTLX (aus Studie I, II und III) und der Trefferquote (aus Studie I)	284

8.7	Differenz zwischen den Parameter-Faktorladungen der kardialen Komponenten des NASA-RTLX (Na) aus Studie I, II und III sowie der kardialen Komponente der Trefferquote (Tq) aus Studie I . . .	285
8.8	Differenz zwischen den Parameter-Faktorladungen der elektrodermalen Profile des NASA-RTLX (Na) aus Studie III, II und I und der Valenz (Va) aus Studie I	286
8.9	Parameter-Faktorladungen a für die elektrodermalen Komponenten des NASA-RTLX (aus Studie I, II und III) und der Valenz (aus Studie I)	287
9.1	Kardiale Profile des NASA-RTLX (Na), der Trefferquote (Tq), der Erregung (Er) und der Valenz (Va) aus den Studien I bis III zugeordnet zur Art der Beanspruchung. Zu sehen sind die Parameter-Faktorladungen a , deren Kongruenzkoeffizient C , die Bestimmtheitsmaße $\text{val}R_{\text{PLS}}^2$ und $\text{val}R_{\text{PLS}}^2$ sowie das Regressionsgewicht B	302
9.2	Elektrodermale Profile des NASA-RTLX (Na), der Trefferquote (Tq), der Erregung (Er) und der Valenz (Va) aus den Studien I bis III zugeordnet zur Art der Beanspruchung. Zu sehen sind die Parameter-Faktorladungen a , deren Kongruenzkoeffizient C , die Bestimmtheitsmaße $\text{val}R_{\text{PLS}}^2$ und $\text{val}R_{\text{PLS}}^2$ sowie das Regressionsgewicht B	313
A-1	Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie I	374
A-2	Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie I	375
A-3	Deskriptive Statistiken des NASA-RTLX jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I	376
A-4	Deskriptive Statistiken der Trefferquote jeweils für die 6 Bedingungen der Balltrackingaufgabe sowie über alle Bedingungen	377

A-5	Deskriptive Statistiken der Erregung jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I	377
A-6	Deskriptive Statistiken der Valenz jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I	378
A-7	Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter	379
A-8	Bedingungs-Faktorladungen der Trefferquoten- und der kardialen Komponente aus der NPLS-Regression der Trefferquote auf kardiale Parameter	379
A-9	Bedingungs-Faktorladungen der Erregungs- und der kardialen Komponente aus der NPLS-Regression der Erregung auf kardiale Parameter	380
A-10	Bedingungs-Faktorladungen der Valenz- und der kardialen Komponente aus der NPLS-Regression der Valenz auf kardiale Parameter	380
A-11	Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter	381
A-12	Bedingungs-Faktorladungen der Trefferquoten- und der elektrodermalen Komponente aus der NPLS-Regression der Trefferquote auf elektrodermale Parameter	381
A-13	Bedingungs-Faktorladungen der Erregungs- und der elektrodermalen Komponente aus der NPLS-Regression der Erregung auf elektrodermale Parameter	382
A-14	Bedingungs-Faktorladungen der Valenz- und der elektrodermalen Komponente aus der NPLS-Regression der Valenz auf elektrodermale Parameter	382
A-15	Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie II	383
A-16	Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie II	384

A-17 Deskriptive Statistiken des NASA-RTLX jeweils für die 12 Bedingungen sowie über alle Bedingungen der Studie II	385
A-18 Deskriptive Statistiken der Trefferquote jeweils für die 12 Bedingungen sowie über alle Bedingungen der Studie II	386
A-19 Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter	387
A-20 Bedingungs-Faktorladungen der Trefferquoten- und der kardialen Komponente aus der NPLS-Regression der Trefferquote auf kardiale Parameter	388
A-21 Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter	389
A-22 Bedingungs-Faktorladungen der Trefferquoten- und der elektrodermalen Komponente aus der NPLS-Regression der Trefferquote auf elektrodermale Parameter	390
A-23 Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie III	391
A-24 Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie III	392
A-25 Deskriptive Statistiken des NASA-RTLX jeweils für die 20 Bedingungen sowie über alle Bedingungen der Studie III	393
A-26 Deskriptive Statistiken der Trefferquote jeweils für die 15 Bedingungen sowie über alle Bedingungen der Studie III	397
A-27 Modellfit-Indizes für die geschachtelten Modelle der Trefferquote	398
A-28 Parameter für das Modell ($M_{S \times A}$) der Trefferquote	399
A-29 Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter	401
A-30 Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter	402

Abkürzungen und Symbole

Abkürzungen

AIC	Akaike Informationskriterium
ANOVA	Varianzanalyse (engl.: Analysis of Variance)
ANS	Autonome Nervensystem
EDA	Elektrodermale Aktivität
EDR	Elektrodermale Reaktion
EDL	Elektrodermales Level, Niveauwert der Hautleitwerts
EKG	Elektrokardiogramm
GT	Generalisierbarkeitstheorie
HF	Spektrale Leistungsdichte im hohen Frequenzbereich
HRV	Herzratenvariabilität
IAPS	International Affective Picture Systems
ifab	Institut für Arbeitswissenschaft und Betriebsorganisation
KIT	Karlsruher Institut für Technologie
KonsensOP	Kontextsensitive Assistenz im aufmerksamen OP
KNN	Künstliche neuronale Netze
KTT	Klassische Testtheorie
LF	Spektrale Leistungsdichte im niedrigen Frequenzbereich
$M_{(\cdot)}$	Mehrebenenmodell
Max	Maximum

MEA	Mehrebenenanalyse, hierarchische Regression
Min	Minimum
MTMM	Multitrait-Multimethod (Analyse)
NASA-TLX	National Aeronautics and Space Administration Task Load Index
NASA-RTLX	Rohversion des National Aeronautics and Space Administration Task Load Index (ohne Gewichtung der Beanspruchungsstruktur)
NN	Um Extrazystolen bereinigter Abstand zwischen zwei R-Zacken des Elektrokardiograms
NS-EDR	Unspezifische elektrodermale Reaktion
NPLS	Multilineare Regression der partiellen kleinsten Quadrate
PARAFAC	Parallele Faktorenanalyse
PCA	Hauptkomponentenanalyse (engl.: Principal Component Analysis)
PLS	Regression der partiellen kleinsten Quadrate (engl.: Partial Least Squares)
ran(·)	Aus Zufallsdaten (engl.: random data) gewonnener Kennwert
RMSE	Wurzel der mittleren Fehlerquadratsumme (engl.: Root Mean Square Error)
RR	Abstand zwischen zwei R-Zacken des Elektrokardiograms
RSA	Respiratorische Sinusarrhythmie
SAM	Self-Assessment Manikin (Fragebogen)
SLD	Spektrale Leistungsdichte
SR	Hautwiderstand (engl.: Skin Resistance)
SQR	Summe quadrierter Residuen
SVM	Support Vector Machine

val(\cdot)	Kreuzvalidierte Kenngröße der multilineareren Regression der partiellen kleinsten Quadrate
ZNS	Zentrale Nervensystem

Symbole

$-2LL$	-2 log-likelihood, Devianz
a	Faktorladung, Komponentenladung
a_{Δ}	Differenz zwischen den Faktorladungen
B	Regressionsgewicht
C	Kongruenzkoeffizient
CI	Konfidenzintervall
d	Effektstärke Cohen's d
df	Freiheitsgrade
G	Leitwert
I	Strom
M	Arithmetischer Mittelwert
N	Größe der Stichprobe
p	Signifikanzwert
r	Bravais-Pearson'scher Korrelationskoeffizient
R	Widerstand
R^2_{GLMM}	Marginales Pseudo-Bestimmtheitsmaß der Mehrebenenanalyse
R^2_{PLS}	Annäherendes Bestimmtheitsmaß der multilineareren Regression der partiellen kleinsten Quadrate
SD	Standardabweichung (engl.: Standard Deviation)
SE	Standardfehler (engl.: Standard Error)
t	t-Wert (des t-Tests)

U	Spannung
α	Signifikanzniveau
$\Delta(\cdot)$	Differenz (Zuwachs) der Kenngrößen
Φ	Teilmenge der physiologischen Parameter
χ^2	Chi-Quadrat-Wert (des Chi-Quadrat-Anpassungstest)
Ψ	Teilmenge der psychologischen Konstrukte

Abkürzungen der physiologischen Parameter

Kardiale Parameter

CVNN	Verhältnis der Standardabweichung der NN-Intervalle zur mittleren Herzperiode
CVSD	Verhältnis des quadratischen Mittels der Unterschiede zwischen zwei benachbarten NN-Intervallen zur mittleren Herzperiode
DFA1	Trendbereinigte Fluktuation der Kurzzeit-Zyklusschwankungen (4–16 Herzschläge)
EKG.HF	Spektrale Leistungsdichte im hohen Frequenzbereich (0.15 Hz bis 0.4 Hz)
EKG.LF	Spektrale Leistungsdichte im niedrigen Frequenzbereich (0.04 Hz bis 0.15 Hz)
EKG.LF/HF	Verhältnis der spektralen Leistungsdichte im niedrigen Frequenzbereich zur spektrale Leistungsdichte im hohen Frequenzbereich
EKG.LFn	Normierte spektrale Leistungsdichte im niedrigen Frequenzbereich
EKG.TP	Spektrale Leistungsdichte im gesamten Frequenzbereich (0.0033 Hz bis 0.4 Hz)

HP	Mittlere Herzperiode, mittleres Schlagintervall, mittleres NN-Intervall
HR	Mittlere Herzschlagfrequenz
MQSD	Quadratisches Mittel der Unterschiede zwischen zwei benachbarten NN-Intervallen
pNN20	Prozentualer Anteil sukzessiver NN-Intervall-Differenzen größer als 20 ms
pNN50	Prozentualer Anteil sukzessiver NN-Intervall-Differenzen größer als 50 ms
SampEn	Sample Entropie als Komplexitätsmaß der mittleren Herzperiode über die Zeit
SDHR	Standardabweichung der Herzschlagfrequenz
SDNN	Standardabweichung der NN-Intervalle
Shannon.h	Shannon-Entropie auf Basis der Klassenwahrscheinlichkeiten der NN-Intervall-Dichteverteilung
SVDEn	Entropie der Singulärwertzerlegung (engl.: Singular Value Decomposition) als Indikator der Dimensionalität der NN-Intervall-Reihe
T.ampl	Mittlere Amplitude der T-Welle
T.area	Mittlere Fläche unter der T-Welle
Triang	Integral der NN-Intervall-Dichteverteilung geteilt durch das Maximum der Dichteverteilung

Elektrodermale Parameter

EDA.HF	Spektrale Leistungsdichte des Hautleitwertsignals im hohen Frequenzbereich (0.15 Hz bis 0.4 Hz)
EDA.LF	Spektrale Leistungsdichte des Hautleitwertsignals im niedrigen Frequenzbereich (0.04 Hz bis 0.15 Hz)

EDA.VHF	Spektrale Leistungsdichte des Hautleitwertsignals im sehr hohen Frequenzbereich (0.4 Hz bis 0.5 Hz)
EDL.D	Absolute Änderung der tonischen Komponente zwischen Aufgabenbeginn und -ende
EDL.M	Mittlerer tonischer Niveauwert des Hautleitwerts
EDR.amp	Mittlere Amplitude der unspezifischen elektrodermalen Reaktionen
EDR.area	Mittlere Fläche unter den unspezifischen elektrodermalen Reaktionen
EDR.N	Anzahl der unspezifischen elektrodermalen Reaktionen
EDR.peaks	Mittlere Zeit zwischen dem Reaktionsmaximum aufeinanderfolgender unspezifischen elektrodermalen Reaktionen
EDR.rec.t/2	Mittlere Zeit zwischen Reaktionsmaximum und dem Punkt eines 50 %-igen Abfalls der Amplitude der unspezifischen elektrodermalen Reaktionen
EDR.rec.t/c	Mittlere Zeit zwischen Reaktionsmaximum und dem Punkt eines 63 %-igen Abfalls der Amplitude der unspezifischen elektrodermalen Reaktionen
EDR.rise.t	Mittlere Zeit zwischen Reaktionsbeginn und Reaktionsmaximum der unspezifischen elektrodermalen Reaktionen
EDR.rise.v	Mittlere Geschwindigkeit des Anstiegs der unspezifischen elektrodermalen Reaktionen
Phasic.SD	Standardabweichung der phasischen Komponente des Hautleitwerts

Notation der Experimentalbedingungen

Studie I

S1	Geringe Aufgabenschwierigkeit
S2	Hohe Aufgabenschwierigkeit
-	Negatives Leistungsfeedback
+	Positives Leistungsfeedback
Eh	Hohe Erregung
En	Niedrige Erregung
V-	Negative Valenz
V+	Positive Valenz

Studie II

S1	Geringe Aufgabenschwierigkeit
S2	Hohe Aufgabenschwierigkeit
b	Farbige Zeichnungen
p	Farbige Fotos
z	Schwarz-weiße Zeichnungen
-	Geringe Komplexität
+	Hohe Komplexität

Studie III

B	Balltrackingaufgabe
G	Gedächtnisaufgabe
K	Konzentrationsaufgabe

M	Matrizenaufgabe
S	Stroopaufgabe
1	Ruhebedingung
2	Geringe Schwierigkeitsstufe
3	Mittlere Schwierigkeitsstufe
4	Hohe Schwierigkeitsstufe

Anmerkungen

Die Schreibweise von Zahlen orientiert sich in dieser Arbeit an den Richtlinien der Deutschen Gesellschaft für Psychologie (Pfetsch, 2019, S. 66–67). Dabei werden die Dezimalstellen von den ganzen Zahlen durch einen Punkt getrennt (z. B. 2.1). Bei Zahlen kleiner als Eins wird die führende Null weggelassen, wenn die Zahl nicht größer als Eins bzw. nicht kleiner als -1 werden kann.

Aus Gründen der besseren Lesbarkeit wird an Stellen, an denen eine geschlechtsneutrale Formulierung aus grammatikalischer oder stilistischer Sicht unpassend erscheint, das generische Maskulinum verwendet. Dabei werden weibliche und anderweitige Geschlechteridentitäten ausdrücklich mitgemeint.

1 Einleitung

Unsere Arbeitswelt und unser Alltag ist von der Digitalisierung und Technisierung durchdrungen. Die steigende Vernetzung technischer Systeme (z. B. Internet of Things, Cloud) bei gleichzeitiger Dezentralisierung der Daten, machen die Informationen für den Einzelnen zwar zugänglicher, bergen jedoch die Gefahr, den Menschen mit einer immer wachsenden Menge an Informationen zu überfordern. So führen steigende Anforderungen im Umgang mit modernen Kommunikations- und Informationstechnologien, aber auch die Bedienung immer komplexerer Maschinen im beruflichen Umfeld, zu einer zunehmenden Auslastung der individuellen Ressourcen. So belegen Meyer, Tisch und Hünefeld (2019) sowie S. Kirchner (2015), dass die steigende Digitalisierung und Technologisierung des Arbeitsplatzes mit einer zunehmenden Arbeitsintensivierung und -belastung einhergeht. Nicht selten wird eine situative Überforderung von Fehlern begleitet (z. B. Collins, 2020; Sheridan, 1981). In bestimmten Berufsgruppen (z. B. im Bereich der Luftraumüberwachung oder Medizin) kann dies sogar Menschenleben kosten. Zur optimalen und effizienten Bewältigung von Aufgaben ist es zweckmäßig, dem Menschen dort Hilfe anzubieten, wo die individuellen Ressourcen an ihre Grenzen stoßen. Die Einbindung solcher Unterstützungen in eine informationsverarbeitende Technologie ist heute bereits Realität.

Technische Unterstützungssysteme begleitet den Menschen schon heute im Alltag. So finden sich vom Antilockiersystem im Kraftfahrzeug bis hin zum persönlichen Assistenten im Smartphone (z. B. Siri von Apple) oder im eigenen Zuhause (z. B. Amazon Alexa oder Smart Home) zahlreiche Mensch-Maschine-

Systeme¹, die dem Menschen die Aufgaben erleichtern oder gar ganz abnehmen sollen. Auch in der Arbeitswelt kommen Assistenzsysteme vermehrt zum Einsatz. Längst gehen neue Systeme aus der Medizintechnik oder Industrie, z. B. das motion EAP², über reine Informationssysteme hinaus. Das Ziel solcher Assistenzsysteme ist neben der Verkürzung des Arbeitsprozesses die Verbesserung und Sicherung der Arbeitsqualität bei gleichzeitiger Entlastung der Angestellten. Doch nicht jedes gut gemeinte Assistenzsystem, das auf den Markt kommt, empfängt auch eine positive Resonanz. Bspw. musste Microsoft den ab Version 97 im Microsoft Office auftauchenden Assistenten mit dem englischen Namen „Clippy“ ab Version 2007 aufgrund starker Unbeliebtheit wieder aus dem Programm entfernen. Die Schwäche dieses Assistenzsystems lag in der mangelnden Anpassung an die Bedürfnisse der Nutzerinnen und Nutzer. Dieses Negativbeispiel verdeutlicht die Wichtigkeit, solche Mensch-Maschine-Systeme an den Menschen auszurichten. Ein Beispiel dafür sind intelligente Assistenzsysteme, die sich adaptiv an die Anforderungen, Bedürfnisse und Zustände des Nutzers anpassen (Dix, Finlay, Abowd & Beale, 2004). Das Ziel solcher nutzerzentrierten Mensch-Maschine-Systeme ist es, den Bedarf an Unterstützung zu erkennen und darauf zu reagieren. Dabei gerät die Erfassung von Nutzerzuständen, darunter Emotionen, Beanspruchung oder Stress, für die Schnittstelle zwischen Mensch und Maschine immer mehr in den Fokus der Entwicklung (z. B. Liao, Zhang, Zhu, Ji & Gray, 2006; Ochs, Pelachaud & Sadek, 2007; Sarrafzadeh, Alexander, Dadgostar, Fan & Bigdeli, 2008). Die Nutzung vor allem negativer Nutzerzustände kann nicht nur helfen, Fehler zu vermeiden oder Aufgaben erfolgreich zu bewältigen, sondern auch die präsentierte Informationsmenge an die Voraussetzungen der Situation anzupassen,

¹ Das Mensch-Maschine-System beschreibt die zielgerichtete Wechselwirkung (Interaktion) mindestens eines technischen Systems, der Maschine, mit einer oder mehreren Personen. Das übergeordnete Ziel liegt dabei in der erfolgreichen Bewältigung einer bestimmten Arbeitsaufgabe (Johannsen, 1993; Timpe & Jürgensohn, 2002).

² System zur Effizienzsteigerung und Assistenz bei Produktionsprozessen in Unternehmen auf Basis von Bewegungserkennung und Projektion.

um ein positives Erleben der Technologie im Sinne der User Experience zu gewährleisten (Liao et al., 2006).

Ein vielversprechendes und in der Arbeitswissenschaft bekanntes Konstrukt, um einen potentiellen Unterstützungsbedarf zu ermitteln, ist die mentale Beanspruchung. Zur Erfassung der mentalen Beanspruchung wurden in den letzten Jahrzehnten zahlreiche Verfahren entwickelt, die dieses Konstrukt als subjektives, leistungsbezogenes oder physiologisches Maß abbilden. Aufgrund der geringen Intrusion und der Möglichkeit zur kontinuierlichen Messung sind besonders physiologische – z. B. kardiale und elektrodermale – Maße für den Einsatz in einer Mensch-Maschine-Schnittstelle geeignet. Hierfür müssen diejenigen physiologischen Parameter gewählt werden, welche die mentale Beanspruchung robust und valide abbilden. Doch genau diese Fragestellung ist trotz jahrzehntelanger Forschung bis heute nicht zufriedenstellend geklärt. Dies liegt unter anderem daran, dass in der psychophysiologischen Forschung die Korrelation einer physiologischen Messgröße mit der mentalen Beanspruchung als Rechtfertigung für dessen Indikatorfunktion herangezogen wird. Der Korrelationsnachweis reicht allerdings nicht aus, um sicherzustellen, dass die erfasste physiologische Messgröße tatsächlich die mentale Beanspruchung – und kein anderes Konstrukt – wiedergibt. Immerhin werden in der Literatur dieselben physiologischen Messgrößen als Indikatoren unterschiedlicher Konstrukte (z. B. physische Beanspruchung, Emotionen, Müdigkeit) herangezogen. Eine weitere Herausforderung bildet die Fülle physiologischer Parameter, die als potentielle Indikatoren in Frage kommen. An dieser Stelle fehlt eine Übersicht sowie Bewertung der Parameter hinsichtlich ihrer Bedeutung bei der Vorhersage mentaler Beanspruchung.

Um die oben aufgeführten Herausforderungen an die beanspruchungssensitive Mensch-Maschine-Schnittstelle zu bewältigen, kann das Rahmenmodell von Cacioppo und Kollegen (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000) herangezogen werden. Cacioppo und Kollegen fordern die Neustrukturierung von physiologischen Einzelparametern zu Profilen. Auf diese Weise kann die Voraussetzung geschaffen werden, um von der gemessenen physiologischen

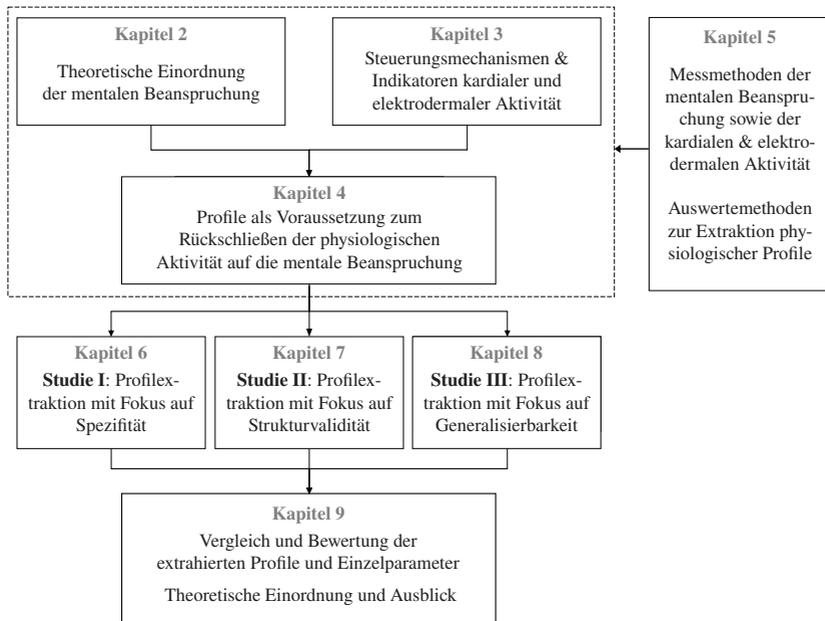


Abbildung 1.1: Dargestellt sind die Inhalte der vorliegenden Arbeit als schematischer Aufbau.

Aktivität auf die mentale Beanspruchung zu schließen. Das Ziel der vorliegenden Arbeit ist die Anwendung des Rahmenmodells von Cacioppo et al. auf die kardialen und elektrodermalen Parameter zur Vorhersage der mentalen Beanspruchung. Der Fokus richtet sich dabei auf die Herleitung physiologischer Profile, die sensitiv und spezifisch für die mentale Beanspruchung sind und über verschiedene mentale Belastungsfaktoren generalisiert werden können. Zudem sollte sich dieselbe Parameterkombination über Forschungsstudien hinweg ergeben (Strukturvalidität). Damit leistet die Arbeit einen Beitrag zur Identifikation valider Indikatoren für eine beanspruchungssensitive Mensch-Maschine-Schnittstelle.

Die vorliegende Arbeit gliedert sich in mehrere Kapitel. Der schematische Aufbau ist in Abbildung 1.1 dargestellt. Zu Beginn wird in Kapitel 2 die mentale Beanspruchung definiert und zu verwandten Forschungsstrukturen

abgegrenzt. Die gängigen theoretischen Erklärungsansätze der menschlichen Informationsverarbeitung werden vorgestellt und die mentale Beanspruchung in ressourcentheoretische und aktivierungstheoretische Modelle eingeordnet. Kapitel 3 beleuchtet die Messung der mentalen Beanspruchung als physiologisches Maß. Hier werden die autonomen Steuerungsmechanismen sowie gängige kardiale und elektrodermale Parameter vorgestellt. Anhand einschlägiger Forschungsbefunden wird die Eignung der Parameter als Indikatoren mentaler Beanspruchung bewertet. Darüber hinaus wird ein Einblick in die Teildisziplin des Psychophysiological Computing gegeben. Diese Teildisziplin wendet – im Gegensatz zur psychophysiologicalen Forschung – regelmäßig multivariate Analyseverfahren zur Vorhersage mentaler Beanspruchung an. In beiden Forschungsdisziplinen mangelt es jedoch an fundierten Validitätsnachweisen, um die physiologische Messgröße als einen Indikator mentaler Beanspruchung zu interpretieren. Diese Diskrepanz greift Kapitel 4 auf und definiert die Voraussetzungen für einen validen physiologischen Indikator mentaler Beanspruchung anhand des Rahmenmodells von Cacioppo und Kollegen (Cacioppo & Tassinary, 1990; Cacioppo et al., 2017). Am Ende von Kapitel 4 werden die zentralen Forschungsfragen der Arbeit vorgestellt. Das Ziel ist hierbei die Ableitung physiologischer Profile mentaler Beanspruchung.

In Kapitel 5 werden die studienübergreifenden Methoden zur Messung der physiologischen Signale sowie zur Messung der subjektiven und leistungsbezogenen mentalen Beanspruchung beschrieben. Darüber hinaus werden die Auswertemethoden mit dem Fokus auf die Ableitung psychophysiologicaler Profile vorgestellt. Die darauffolgenden Kapitel 6, 7 und 8 stellen die Studien I, II und III ausführlich vor. Alle Studien verfolgen das studienübergreifende Ziel, aus den physiologischen Parametern belastbare Profile mentaler Beanspruchung zu extrahieren. Zudem werden in den jeweiligen Studien studienspezifische Ziele adressiert, welche die Spezifität, Strukturvalidität und Generalisierbarkeit der Profile prüfen. Im letzten Kapitel 9 werden alle extrahierten Profile gegenübergestellt und anhand von Gütekriterien bewertet. Dabei wird eine Einordnung in das Rahmenmodell von Cacioppo et al. (Cacioppo & Tassi-

nary, 1990; Cacioppo et al., 2017) vorgenommen. Für die physiologischen Einzelparameter wird eine Übersicht in Hinblick auf ihre Relevanz zur Vorhersage mentaler Beanspruchung erstellt. Im Anschluss werden die latenten Wirkmechanismen physiologischer Profile diskutiert. Des Weiteren werden die Anwendungsmöglichkeiten in einer Mensch-Maschine-Schnittstelle beleuchtet. Am Ende erfolgt eine Diskussion zukünftiger Forschungspotentiale, die sich aus den physiologischen Profilen ergeben.

2 Theoretische Grundlagen: Mentale Beanspruchung

Zur Klärung der theoretischen Konstrukte der Arbeit werden in Kapitel 2.1 die zentralen Begriffe auf dem Gebiet der Beanspruchungsforschung definiert und von verwandten Begriffen abgegrenzt. Dieses Kapitel gibt somit die Definitionen und Notationen an, in welcher die vorliegende Arbeit verfasst ist. Um die mentale Beanspruchung bewerten zu können, ist es zunächst wichtig, die Wirkfaktoren sowie Entstehungsprozesse der menschlichen Informationsverarbeitung zu verstehen. Zu diesem Zweck werden in Kapitel 2.2 die unterschiedlichen Definitionen und theoretischen Konzepte der mentalen Beanspruchung vorgestellt.

2.1 Abgrenzung theoretischer Begriffe

In den letzten Jahrzehnten haben verschiedene Forschungsbereiche eine Vielzahl an unterschiedlichen Begriffen auf dem Gebiet der menschlichen Beanspruchung eingebracht. Zu diesen gehören u. a. die Begriffe: Belastung, Stress, mentale Beanspruchung, kognitive Beanspruchung und emotionale Beanspruchung. Die Definitionen dieser Begriffe sind einem ständigen Wandel unterworfen. Dabei ist stets zu beachten, dass keine Begriffsdefinition den Anspruch erhebt, widerspruchsfrei, allumfassend oder allgemein gültig zu sein. Um die in dieser Arbeit verwendeten Begriffe zu präzisieren, werden im Folgenden vermeintlich verwandte Konstrukte im Lichte der Forschungsliteratur differenziert bzw. theoretische Überlappungen beleuchtet.

2.1.1 Belastung und Beanspruchung

Beschäftigt man sich mit der mentalen Beanspruchung, so stößt man unweigerlich auf den Begriff *Belastung*. Um die Begriffe voneinander abzugrenzen, eignet sich der aus der Arbeitspsychologie stammende Belastungs-Beanspruchungs-Ansatz. Dieser ist heutzutage integraler Bestandteil der deutschen und europäischen Norm für psychische Arbeitsbelastung (EN ISO 10075-1, DIN Deutsches Institut für Normung e. V., 2000) und Ergonomie (EN ISO 26800, CEN Europäisches Komitee für Normung, 2011). Angelehnt an der technischen Mechanik unterscheidet dieser Ansatz zwischen Ursache und Wirkung: Hierbei wirken die aus der äußeren Umwelt stammenden Belastungen auf die Person ein und lösen dadurch im menschlichen Organismus ein Beanspruchungsreaktionen aus (EN ISO 10075-1, DIN Deutsches Institut für Normung e. V., 2000; EN ISO 26800, CEN Europäisches Komitee für Normung, 2011; Rohmert, 1984; Schlick et al., 2010, S. 38–43). Als Analogie verwendet Rohmert das Beispiel einer mechanischen Feder, dessen innere Federspannung (Beanspruchung) durch eine äußere Kraft (Belastung) auf- bzw. abgebaut werden kann (siehe Abbildung 2.1).

Im Belastungs-Beanspruchungs-Konzept ist die Belastung definiert als die Gesamtheit aller objektiv erfassbaren Einflüsse, welche auf die Person einwirken. Äußere Einflüsse (externe Faktoren) ergeben sich bspw. aus den Bedingungen und Eigenschaften der Aufgabe (z. B. Dauer der Tätigkeit), der physikalischen Umwelt (z. B. Raumbeleuchtung) oder der sozialen und gesellschaftlichen Umgebung (z. B. Konflikte oder kulturelle Normen). Die Beanspruchung ist definiert als die Auswirkung der Gesamtbelastung auf eine Person. Diese hängt allerdings nicht nur von den externen Faktoren, sondern auch von den individuellen Eigenschaften der Person (interne Faktoren) ab (J.-H. Kirchner, 1986; Rohmert, 1984). Darunter fallen sowohl zeitvariable Zustände, wie Motivation oder aktuelle Verfassung, aber auch über die Zeit stabile Personenmerkmale, wie Charaktereigenschaften oder Fähigkeiten. Durch die möglichen Kombinationen externer und interner Einflussfaktoren ergeben sich folgende Implikationen für

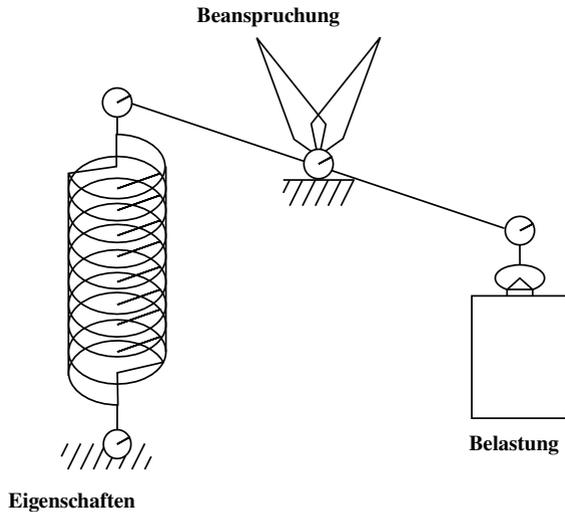


Abbildung 2.1: Beanspruchung und Belastung dargestellt als Modell der mechanischen Feder (modifiziert nach Schlick et al., 2010, S. 39).

das Zusammenspiel von Belastung und Beanspruchung: (1) unterschiedliche Belastungen können bei derselben Person zu einer unterschiedlichen, aber auch zur gleichen Beanspruchung führen; (2) eine identische Belastung kann bei derselben Person zu unterschiedlichen Zeiten eine unterschiedliche Beanspruchung hervorrufen; und (3) eine identische Belastung kann bei unterschiedlichen Personen zur gleichen Zeit eine unterschiedliche Beanspruchung aufweisen.

Die mentale Beanspruchung fällt im Rahmen des Belastungs-Beanspruchungs-Konzepts unter die psychische Beanspruchung. Diese wird nach der EN ISO 10075-1 (DIN Deutsches Institut für Normung e. V., 2000) definiert als die „Auswirkung der psychischen Belastung im Individuum in Abhängigkeit von seinen jeweiligen überdauernden und augenblicklichen Voraussetzungen, einschließlich der individuellen Bewältigungsstrategien“ (S. 3). Hierbei ist zu beachten, dass jede Art der Tätigkeit, ob körperlich oder geistig, die psychische Beanspruchung beeinflussen können.

Das Belastungs-Beanspruchungs-Konzepts unterteilt die Beanspruchung in eine zeitbezogene und eine valenzbezogene Dimension. Erstere bezieht sich auf die Unterscheidung kurzfristiger und langfristiger Beanspruchungsfolgen (EN ISO 26800, CEN Europäisches Komitee für Normung, 2011; Rohmert, 1984). Da kurzfristige Belastungen nicht zwangsläufig zu einer Gesundheitsgefährdung führen, eignen sie sich gut für den Einsatz in Laborstudien. Die Valenz-Dimension unterscheidet positive und negative Auswirkungen der Belastung (EN ISO 26800, CEN Europäisches Komitee für Normung, 2011; Rohmert, 1984). Somit führt eine Belastung nicht zwangsläufig zu einer Beeinträchtigung, wie Ermüdung, sondern sie kann durchaus auch anregend und aktivierend wirken.

2.1.2 Kognitive und mentale Beanspruchung

Sowohl die kognitive als auch die mentale Beanspruchung beziehen sich auf die Informationsverarbeitung im menschlichen Gehirn. Sowohl die Annahmen als auch die Effekte kognitiver und mentaler Beanspruchung ähneln sich stark (vgl. Galy, Cariou & Mélan, 2012). Der Begriff kognitive Beanspruchung ist eher im pädagogischen und didaktischen Bereich angesiedelt. Dabei wird davon ausgegangen, dass das Lernen mit kognitiver Beanspruchung einhergeht. Der Fokus liegt dabei auf der Analyse kognitiver Belastungsfaktoren, um das Lernen zu erleichtern. Bspw. unterscheidet die *Cognitive Load Theorie* (Sweller, van Merriënboer & Paas, 1998) zwischen extrinsischer (Darstellung des Lernmaterials), intrinsischer (Schwierigkeit des Lernmaterials) und lernbezogener Belastung (Lernaufwand), die je nach Ausprägung den Wissenserwerb erschweren oder erleichtern kann. In der vorliegenden Arbeit werden die Begriffe kognitiv und mental synonym verwendet, wobei lernbezogene Aspekte in dieser Arbeit keine Beachtung finden.

2.1.3 Mentale und emotionale Beanspruchung

Die emotionale Beanspruchung bezieht sich auf das Konstrukt der Emotionen. Eine einheitliche Definition dieses Konstrukts gibt es bis dato nicht. Vielmehr ist es eine Sammlung verschiedener theoretischer Richtungen und semantisch verwandter Begriffe – wie Affekt, Stimmung, Befindlichkeit und emotionale Beanspruchung – die im Laufe der Zeit hervorgegangen sind. Auch wenn bereits Versuche unternommen worden sind, diese Begriffe semantisch voneinander zu trennen, findet sich bis heute keine einheitliche Nutzung in der wissenschaftlichen Literatur. Da die vorliegende Arbeit nicht den Anspruch erhebt, eine Begriffsklarheit im Bereich der Emotionsforschung zu schaffen, werden die zuvor erwähnten Emotionsbegriffe synonym verwendet. Ein Abriss zur theoretischen Verortung des Konstrukts Emotionen ist im Anhang A.1 zu finden.

Die emotionale Beanspruchung wird in der vorliegenden Arbeit im Rahmen dimensionaler Emotionsmodelle definiert. Dimensionale Ansätze nach Wundt (1905) versuchen das emotionale Erleben durch wenige zugrunde liegende Dimensionen zu erklären. Wundt führt das Erleben von Gefühlen auf drei Grundgefühle zurück: „Spannung“, „Lust“ und „Erregung“. Darauf aufbauend gehen Bradley und Lang (1994) von den drei Dimensionen *Erregung* (erregend bis beruhigend; engl.: Arousal), *Valenz* (angenehm bis unangenehm) und *Dominanz* (souverän bis unterlegen) aus, um emotionales Empfinden zu beschreiben. Die ersten beiden Dimensionen Erregung und Valenz entsprechen den drei von Wundt (1905) beschriebenen Emotionsdimensionen. Die dritte Dimension Dominanz beschreibt die Kontrolle bzw. Stärke über eine Situation. In der Emotionsforschung wurden jedoch nur die Dimensionen Erregung und Valenz als unabhängige und stabile Dimensionen bestätigt (Sokolowski, 2002). Aus diesem Grund betrachtet die vorliegende Arbeit nur die ersten beiden Emotionsdimensionen Erregung und Valenz.

Gemäß dem Belastungs-Beanspruchungs-Konzept wird die psychische Beanspruchung in eine mentale und eine emotionale Komponente unterteilt (EN

ISO 10075-1, DIN Deutsches Institut für Normung e. V., 2000; EN ISO 26800; CEN Europäisches Komitee für Normung, 2011; Rohmert, 1984). Allerdings ist die Differenzierung in eine mentale und emotionale Beanspruchung in der Forschungspraxis nicht trivial. In der Forschungspraxis finden sich gehäuft theoretische und definatorische Vermischungen zwischen mentaler und emotionaler Beanspruchung. Bspw. werden im Bereich kognitiver Emotionstheorien Emotionen auf ihre kognitiven Bewertungsprozesse reduziert, sodass Kognition und Emotion begrifflich nicht mehr eindeutig voneinander abgegrenzt werden können. Eine aus theoretischer Sicht zufriedenstellende Differenzierung ist bis heute nicht vorhanden. Aus diesem Grund wird häufig ein pragmatischer Ansatz verfolgt und zur Abgrenzung auf die spezifischen Belastungsfaktoren verwiesen, welche die Beanspruchung auslösen (Manzey, 1998). In diesem Sinne gilt derjenige Anteil der psychischen Gesamtbeanspruchung als mental, welcher von einer informatorischen Belastung ausgelöst wird und somit Anforderungen an die menschliche Informationsaufnahme, -verarbeitung und -umsetzung stellt (Manzey, 1998; Schlick et al., 2010, S. 393–394). Bestimmt wird die mentale Belastung von der Schwierigkeit und Komplexität einer primär kognitiven Aufgabe, wie bspw. Problemlösen, Überwachen oder Gedächtnisabruf. Als emotionale Beanspruchung gilt derjenige Anteil, welcher auf ausführungsspezifische Belastungen, wie soziale Konflikte, Lärm, Klima, Zeitdruck oder Gefahren, zurückzuführen ist und sich als motivationale oder affektive Reaktion, wie Angst, Hilflosigkeit oder Langeweile, äußert (Manzey, 1998; Schlick et al., 2010, S. 394). Zwar spezifiziert die beschriebene Differenzierung die Begriffe der mentalen und emotionalen Beanspruchung auf Ebene der Ursache (Belastung), bietet aber dennoch keine Lösung auf Ebene der Wirkung (Beanspruchung). Denn häufig treten emotionale und mentale Belastungsfaktoren bei der Ausführung einer Tätigkeit gemeinsam auf und führen somit zu einer Konfundierung auf der psychophysiologischen Wirkungsebene (Rüchel, 2013). Bspw. können komplexe Aufgabe nicht nur zu einer hohen Auslastung der menschlichen Informationsverarbeitung, sondern gleichzeitig zu einem Frustrationsgefühl führen. Auf der anderen Seite beeinflusst der aktuelle emotionale Zustand die Bearbeitung einer Aufgabe und somit die interne Wahrnehmung

der mentalen Beanspruchung (Hussain, Calvo & Chen, 2014). Folglich kann die Trennung in mentale und emotionale Anteile, analog einer Varianzaufklärung, in der Forschungspraxis nur in kontrollierten Laborbedingungen durch gezielte Operationalisierung mentaler und emotionaler Belastungsfaktoren realisiert werden. Phänomenologisch ist eine Trennung affektiver und kognitiver Auswirkungen jedoch aufgrund ihrer funktionalen Wechselbeziehung kaum möglich (Huang & Tettegah, 2010; Hussain et al., 2014; Immordino-Yang & Damasio, 2007). So schreibt auch Edelman (2000, S. 242): „Beim Prozess der Informationsverarbeitung handelt es sich eigentlich immer um ein Zusammenwirken kognitiver und emotionaler Prozesse. Emotionale (und motivationale) Faktoren sind selbst bei den abstraktesten Formen intellektueller Leistung beteiligt.“

2.1.4 Beanspruchung und Stress

Der Begriff Stress umfasst eine Bandbreite unterschiedlicher psychologischer Phänomene. Zu diesen zählen bspw. Frustration, Angst und Überforderung, aber auch lebensbedrohliche Erfahrungen. Zumeist wird Stress als Überbegriff für eine negativ behaftete Erfahrung mit hoher Intensität verwendet (vgl. Lazarus & Folkman, 1984). Die Forschungsfelder zur mentalen Beanspruchung, zu Emotionen sowie zu Stress sind stark miteinander verwandt, wodurch Überlappungen zwischen Konstruktdefinitionen und deren Operationalisierungen bestehen. Gleichermaßen verhält es sich mit den Konstrukten Distress und Eustress, definiert als negativ und positiv erlebter Stress. Analog zur Beanspruchung, welche als Auswirkung äußerer Belastungsfaktoren definiert ist, gilt Stress als eine Konsequenz äußerer Stressfaktoren. Eine klare Abgrenzung, ob es sich bei einem Reiz um einen Auslösefaktor für mentale Beanspruchung, Stress oder Emotionen handelt, gibt es nicht. Folglich werden ähnliche, wenn nicht identische, Auslösebedingungen für unterschiedliche Forschungsfelder verwendet (z. B. emotions- bzw. stressauslösende Bilder). Einen Unterschied gibt es zumeist nur im Forschungsschwerpunkt. Während die Emotionsforschung überwiegend an der Differenzierung verschiedener Emotionen interessiert ist, ist Stress als ein

eher übergreifendes Konstrukt einer negativ behafteten, jedoch unspezifischen Emotion zu verstehen (Boucsein, 2012, S. 369, 381). Im Gegensatz dazu sind negative und positive Emotionen bei der Erforschung mentaler Beanspruchung nur am Rande interessant.

2.2 Modelle mentaler Beanspruchung

Das theoretische Konzept der mentalen Beanspruchung (engl.: mental load, mental workload) fand bereits in der ersten Hälfte des 20-sten Jahrhunderts den Weg in die arbeitspsychologische Forschung und hat sich seit dieser Zeit zu einem der bekanntesten Konzepte in der Arbeitswissenschaft entwickelt (z. B. Manzey, 1998; Hancock & Meshkati, 1988; Wickens et al., 2013). Nicht zuletzt erklärt sich die Bekanntheit der mentalen Beanspruchung durch die angestrebten Bemühungen von Arbeitswissenschaftlern, die Leistungsfähigkeit des Menschen in bestehenden oder zukünftigen Mensch-Maschine-Systemen zu bewerten und zu optimieren. Vor allem im Bereich informatorischer Arbeit bildet die mentale Beanspruchung einen zentralen Faktor der Leistungsfähigkeit. Bekannte Beispiele sind die hohe Arbeitslast bestimmter Berufsgruppen, z. B. Callcenter Mitarbeiter oder Chirurgen, aber auch die simultane Ausführung mehrerer Aufgaben (Multitasking), wie das Führen eines Fahrzeugs und gleichzeitiges Telefonieren, zählen dazu. Überschreitet der Anspruch einer oder mehrerer Arbeitsaufgaben die individuellen Ressourcen menschlicher Informationsverarbeitung, erhöht sich auch die Gefahr von Fehlern, geringerer Produktivität und Erschöpfungserscheinungen. Trotz der Bedeutsamkeit dieses Konstrukts konnte sich bis heute keine einheitliche Definition für die mentale Beanspruchung durchsetzen. Vielmehr findet man eine Bandbreite an Definitionen, Erklärungsmodellen und Messmethoden, die je nach Forschungsrichtung eine mehr oder minder differenzierte Trennung bestimmter Prozesse und Begriffe vornehmen. Allen gemeinsam ist der Versuch, die menschliche Informationsverarbeitung in einem Modell abzubilden und gleichzeitig einen messtheoretischen

Zugang zur Bewertung mentaler Tätigkeiten und der resultierenden mentalen Beanspruchung zu bieten.

In der Literatur zur mentalen Beanspruchung findet man neben der Beschreibung der strukturellen Aufteilung menschlicher Informationsverarbeitung in Teilsysteme auch Ansätze, welche die Intensität der Informationsverarbeitung (vgl. Kahneman, 1973) in den Vordergrund stellen, die ihrerseits von der Komplexität und Schwierigkeit der Aufgabe, aber auch von den individuellen Personenmerkmalen, determiniert werden (Fahrenberg, 2001; Gaillard, 1993; Manzey, 1998; O'Donnell & Eggemeier, 1986). Über die Jahre hinweg entwickelten sich je nach Forschungsrichtung unterschiedliche Modellvorstellungen menschlicher Informationsverarbeitung, die Schlick et al. (2010, S. 286–312) in zwei Theoriegruppen gliedern: die (1) phänomenologisch-empirischen bzw. biologischen Modelle – welche den Prozess und die Funktionsstruktur von Wahrnehmung und Kognition beschreiben – und die (2) mathematisch-funktionalen bzw. technischen Modelle – welche versuchen informatorisch-mentale Prozesse mittels Gleichungen abzubilden. Für die vorliegende Arbeit wird die erste Modellgruppe betrachtet. Schlick et al. unterteilen die phänomenologisch-empirischen Modelle weiter in (a) sequentielle Modelle, welche den Fluss der Informationen durch den Organismus abbilden und (b) Kapazitätsmodelle, welche den Energieeinsatz bzw. die kognitive Kapazität bei der Informationsverarbeitung betrachten und die messtheoretische Bewertung mentaler Beanspruchung anstreben. Da die Kapazitätsmodelle den empirischen sowie messtheoretischen Fokus der vorliegenden Arbeit, nämlich die Analyse psychophysiologischer Indikatoren mentaler Beanspruchung widerspiegeln, werden diese nachfolgend vertieft dargestellt. Kapazitätsmodelle werden in der Literatur – trotz bestehender Gemeinsamkeit – in weitere Theoriegruppen unterteilt (hier nach Manzey, 1998): (1) die ressourcentheoretischen Modelle (siehe Abschnitt 2.2.1) und (2) die aktivierungstheoretischen Modelle (siehe Abschnitt 2.2.2).

Nicht jedes Modell kann eindeutig nur einer Theoriegruppe zugeordnet werden. Aufgrund bestehender Überschneidungen zu benachbarten Theoriegruppen erhebt die folgende Unterteilung kein Anspruch auf Allgemeingültigkeit und ist

als Betonung unterschiedlicher Schwerpunkte zu verstehen. Die anschließenden Kapitel stellen einen Ausschnitt aus der Gesamtheit existierender Modelle beider Theoriegruppen dar und schaffen eine theoretische Grundlage für die in der vorliegenden Arbeit angestrebte psychophysiologische Untersuchung mentaler Beanspruchung.

2.2.1 Mentale Beanspruchung in ressourcentheoretischen Modellen

Diese kognitionspsychologisch orientierte Theoriegruppe stellt die strukturellen Aspekte menschlicher Informationsverarbeitung in den Vordergrund (Manzey, 1998). Ressourcentheoretische Modelle stellen alle in der einen oder anderen Form Kapazitätsmodelle dar. Diesen liegt die gemeinsame Annahme zugrunde, dass die menschliche Informationsverarbeitung eine begrenzte Kapazität besitzt und die mentale Beanspruchungsreaktion eine Aktivierung bzw. Ausschöpfung dieser Kapazität darstellt (Schlick et al., 2010, S. 291–299). Der Ursprung ressourcentheoretischer Modelle liegt in der Interferenzforschung. Diese versucht, den Effekt des Leistungsabfalls bei der simultanen Bewältigung mehrerer Aufgaben (Multitasking) zu erklären. Solche Leistungseinbußen bei Doppeltätigkeiten können gemäß dem ressourcentheoretischen Ansatz dann auftreten, wenn bspw. die Ausschöpfung der begrenzten Verarbeitungskapazität bei der Bewältigung einer Aufgabe so stark ist, dass nur noch eine geringe Restkapazität (Kahneman, 1973) zur Bearbeitung der zweiten Aufgabe übrigbleibt.

Das Ressourcenmodell nach Kahneman (1973)

Kahneman (1973) legte mit seinem Ansatz einen wichtigen Grundstein für die Weiterentwicklung der Kapazitätsmodelle. Kahneman geht in seinem Ansatz von einer begrenzten Verarbeitungskapazität der menschlichen Informationsverarbeitung aus. Die Kapazität selbst ist unspezifisch und wird folglich mit Konzepten wie Aufmerksamkeit und Anstrengung gleichgesetzt (Manzey, 1998).

Zudem ist die Kapazität allgemein, sodass jegliche Art der Informationsverarbeitungsaufgabe, ob visuelles Suchen oder Kopfrechenaufgaben, auf dieselbe Kapazität zugreifen. Die Menge der Kapazität, welche zur Bewältigung einer Aufgabe akquiriert wird, kann bis zu einem individuellen Maximum, z. B. durch verstärkte Anstrengung, angehoben werden. Wie die Kapazität zugewiesen wird, hängt dabei von unterschiedlichen Merkmalen ab: (1) der Schwierigkeit der Aufgabe, welche die Kapazitätsauslastung beeinflusst, (2) der individuellen Disposition der Person, welche die Bereitstellung von Kapazitäten bei unwillkürlichen Reaktionsmustern (z. B. Orientierungsreaktion) sicherstellt und (3) den Intentionen der Person, welche eine motivationale Verteilung der Kapazitäten ermöglichen (Manzey, 1998).

Befunde aus der Interferenzforschung widersprachen allerdings der Annahme einer unspezifischen Verarbeitungskapazität, wonach schwere Aufgaben (mit einer hoher Ressourcenauslastung) stärker miteinander interferieren sollten als leichte Aufgaben (Schlick et al., 2010, S. 298). Die Stärke der Interferenz wurde jedoch wesentlich durch die Ähnlichen in der Aufgabenstruktur determiniert (Wickens, 1984). Unter Berücksichtigung dieser Kritik entwickelte Wickens (2002, 2008) seine Theorie multipler Ressourcen, die im Folgenden beschrieben wird.

Das Multiple-Ressourcen-Modell von Wickens (2002, 2008)

Die Theorie multipler Ressourcen von Wickens (2002, 2008) geht von mehreren Informationsverarbeitungssystemen aus. Die Ressourcen der jeweiligen Verarbeitungssysteme werden ebenfalls als begrenzt angenommen. Die Ressourcen können jedoch unabhängig voneinander abgerufen werden. Interferenzeffekte sind gemäß der Theorie multipler Ressourcen nur dann anzutreffen, wenn die zu bearbeitenden Aufgaben auf dieselbe Verarbeitungsstruktur bzw. -ressource zurückgreifen. Im Vordergrund der Theorie steht die strukturelle Abbildung menschlicher Informationsverarbeitung, die auf einer umfangreichen Analyse der Befunde aus der Interferenzforschung aufbaut.

Dargestellt wird das Modell als ein dreidimensionaler Würfel (siehe Abbildung 2.2). Im Modell unterscheidet Wickens zwischen vier Dimensionen der Informationsverarbeitung: (1) den Verarbeitungsstufen als Abbild der Verarbeitungsprozesse der Wahrnehmung, Kognition und Antwort; (2) der verbalen oder räumlichen Verarbeitungskodierung; (3) den Modalitäten als auditive, visuelle oder taktile Wahrnehmung; und (4) dem visuellen Kanal als die fokale oder periphere visuelle Wahrnehmung. Die Interferenz zweier Aufgaben ist umso größer, je stärker die besetzten Würfelzellen überlappen und somit um dieselbe Ressource konkurrieren.

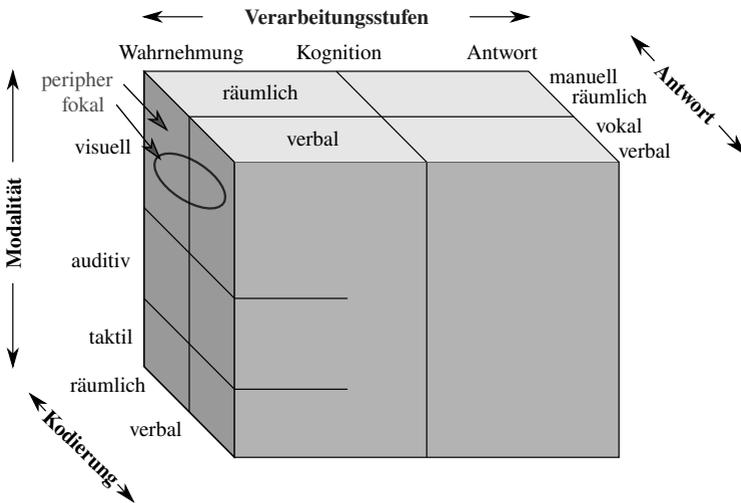


Abbildung 2.2: Darstellung des Multiple-Ressourcen-Modells als Würfel (modifiziert nach Wickens et al., 2013, S. 330).

Wenngleich Wickens von spezifischen Ressourcen ausgeht, nimmt er auch die Existenz übergeordneter unspezifischer Ressourcen an, die für unterschiedliche Anforderungen akquiriert werden können. Gemäß der Multiple-Ressourcen-Theorie wird mentale Beanspruchung als ein mehrdimensionales Konstrukt verstanden, welches sich in unterschiedliche mentale Beanspruchungsaspekte gliedern lässt. Auf dieser Grundlage bietet das Multiple-Ressourcen-Modell eine erste deskriptive Taxonomie, die u. a. optimal zur Operationalisierung

unterschiedlicher mentaler Beanspruchungszustände in experimentellen Untersuchungen genutzt werden kann.

2.2.2 Mentale Beanspruchung in aktivierungstheoretischen Modellen

Diese psychophysiologisch orientierte Theoriegruppe hebt den aktivierungstheoretischen bzw. energetischen Aspekt menschlicher Informationsverarbeitung hervor (Manzey, 1998). Die mentale Beanspruchung wird dabei mit dem psychophysischen Aufwand, der zur erfolgreichen Aufgabenbewältigung benötigt wird, gleichgesetzt (Manzey, 1998).

Das Konzept der zentralen Aktivierung nach Bartenwerfer (1969)

Ursprünglich basiert der aktivierungstheoretische Ansatz auf der Annahme eines eindimensionalen Kontinuums zentraler Aktivierung (Bartenwerfer, 1969), die innerhalb einer Person unterschiedliche Erregungs- bzw. Anspannungsstufen annehmen kann. Die zentrale Aktivierung spiegelt nach diesem Ansatz die energetische Basis der Informationsverarbeitung wieder (Manzey, 1998). Die Erweiterung dieses Konzepts durch das Yerkes-Dodson-Gesetz (Yerkes & Dodson, 1908) steuerte dazu bei, dass von einem U-förmigen Zusammenhang der zentralen Aktivierung und der Leistungsfähigkeit einer Person sowie einem negativen Zusammenhang zwischen dem optimalen Aktivierungsniveau und der Aufgabenschwierigkeit ausgegangen wurde (Manzey, 1998). Gemäß diesem eindimensionalen Ansatz konnte der Zustand der mentalen Beanspruchung nur hinsichtlich ihrer Stärke differenziert werden. Diese Beanspruchungsstärke galt es, mit Hilfe von physiologischen (z. B. Herzrate oder Hautleitfähigkeit) und subjektiven Maßen, den sog. Aktivierungsindikatoren, zu erfassen. Doch die empirischen Befunde zahlreicher Beanspruchungsstudien zeigen auf, dass die Eindimensionalität des Konzepts keine ausreichende Erklärungsgrundlage für die spezifischen Reaktionsmuster (Motivationsspezifität,

Stimuluspezifität, Individualspezifität) bestimmter Aktivierungsindikatoren bietet (vgl. Fahrenberg, 1983; Manzey, 1998). Um dieser Kritik zu begegnen, wurden das eindimensionale Aktivierungskonzept zu einem mehrdimensionalen Konstrukt in nachfolgenden Modellen ausdifferenziert.

Das dreidimensionale Aktivierungsmodell von Pribram und McGuinness (1975)

Ein mehrdimensionaler Ansatz des Aktivierungskonzepts ist das an der Neuropsychologie orientierte Modell von Pribram und McGuinness (1975). In ihrem Modell unterscheiden Pribram und McGuinness drei Funktions- bzw. Aktivierungssysteme, welche durch interaktive Prozesse die Aktivierung des zentralen Nervensystems (ZNS) und somit die Aufmerksamkeit kontrollieren. Das (1) *Erregungs-System* (engl.: Arousal-System) ist verantwortlich für die Steuerung der selektiven Aufmerksamkeit. Bspw. löst dieses System eine kurzlebige und inputbezogene Aktivierung, wie die Orientierungsreaktion auf neuartige Reize, aus. Das zweite System ist das (2) *Aktivierungs-System* (engl.: Activation-System), welches die motorischen Prozesse kontrolliert. Dieses sorgt für eine langlebige und reaktionsbezogene Aktivierung (z. B. erhöhte Reaktionsbereitschaft). Das (3) *Anstrengungs-System* (engl.: Effort-System) stellt die Koordinationseinheit der Funktionssysteme dar. Diese Koordinationseinheit kann die Reiz-Reaktions-Verknüpfungen zwischen dem Erregungs-System und dem Aktivierungs-System koppeln und entkoppeln. Im Gegensatz zum Erregungs- und Aktivierungs-System als überwiegend reizgesteuerte und unwillkürliche Funktionssysteme unterliegt das übergeordnete Anstrengungs-System einer willkürlichen Steuerung, die über Feedback-Prozesse die Anstrengung regulieren kann.

Das 3-Arousal- und das 4-Arousal-Modell von Backs und Boucsein (2009)

Aufbauend auf dem dreidimensionalen Aktivierungsmodell von Pribram und McGuinness (1975) erarbeiten Boucsein und Backs das 3-Arousal-Modell und später das 4-Arousal-Modell (Backs & Boucsein, 2009; Boucsein, 2012, S. 353–362). Hierbei werden die zuvor bei Pribram und McGuinness (1975) vorgestellten drei Funktionssysteme als Prozesse der emotionalen, tätigkeitsbezogenen und informatorischen Aktivierungssteuerung postuliert und um ein weiteres, unspezifisches Aktivierungssystem erweitert. Das (1) *Affekt-Erregungs-System* ist für die Aufmerksamkeitsreaktion im Sinne einer Orientierungsreaktion, aber auch für die Fluchtreaktion, zuständig. Hier werden überwiegend Reize für die emotionsbezogene Aktivierungssteuerung verarbeitet. Das (2) *Anstrengungs-System* ist für die Kopplung- und Entkopplung des Informationseingangs und -ausgangs des ZNS zuständig. Dieses System selektiert und vergleicht die eingehende Information mit vergangenen (gespeicherten) und zukünftigen (geplanten) Informationen und Handlungsabsichten. Die physiologische Aktivität, welche durch dieses System hervorgerufen wird, ist nach Boucsein (2012, S. 353) eine Begleiterscheinung der zentralen Informationsverarbeitung und kommt somit der Definition der mentalen Beanspruchung am nächsten. Das (3) *Vorbereitende-Aktivierungs-System* (engl.: Preparatory-Activation-System) ist für die Verhaltensaktivierung zuständig und sorgt für eine Reaktionsbereitschaft. Das (4) *allgemeine Erregungs-System* stellt ein unspezifisches Aktivierungssystem dar.

Basierend auf einer Literaturrecherche identifizierten Backs und Boucsein (2009) typische physiologische Indikatoren der vier Funktionssysteme. Wie in Tabelle 2.1 dargestellt, werden diese entsprechend der vier Aktivierungssysteme in die Gruppen 1 bis 4 eingeteilt. Das Affekt-Erregungs-System zeigt sich durch eine Reaktion der Gruppe 1 Indikatoren. Diese tritt immer dann auf, wenn bestimmte Auslösereize die Aufmerksamkeit in Richtung des neuen Reizes lenken. Als Folge stellt das Vorbereitende-Aktivierungs-System ein erhöhtes Bereitschaftspotential zur Verhaltensaktivierung bereit, welches sich

Tabelle 2.1: Zuordnung physiologischer Kenngrößen zu den vier Aktivierungssystemen (nach Backs & Boucsein, 2009; Boucsein, 2012, S. 462)

Aktivierungssystem	Physiologische Indikatoren
1 Affektive Erregung	- Anstieg der Frequenz unspezifischer elektrodermalen Reaktionen - Phasische Veränderungen der Herzrate
2 Anstrengung	- Abfall der Herzratenvariabilität - Anstieg der Abfallzeit der elektrodermalen Reaktionen - Anstieg der Sakkaden- und Fixationsdauer
3 Vorbereitende Aktivierung	- Moderater Anstieg der tonischen Herzrate - Anstieg der Amplitude unspezifischer elektrodermalen Reaktionen
4 Allgemeine Erregung	- Deutlicher Anstieg der Herzrate - Deutlicher Anstieg der tonischen elektrodermalen Aktivität

in der Reaktion der Gruppe 3 Indikatoren niederschlägt. Dieser beschriebene Reiz-Reaktions-Zyklus kann durch das Anstrengungs-System zum Zwecke einer Informationsintegration entkoppelt werden. Eine auf dieser Ebene stattfindende kognitive Informationsverarbeitung schlägt sich im Reaktionsmuster der Gruppe 2 nieder. Die allgemeine Erregung bildet die Reaktionsgruppe 4. Starke interne oder externe Reize können dazu führen, dass die physiologischen Reaktionen der Gruppe 4 alle anderen physiologischen Reaktionsgruppen überlagern (Backs & Boucsein, 2009; Boucsein, 2012, S. 463). Die Zuweisung physiologischer Aktivierungsindikatoren zu den Funktionssystemen bezeichnet Manzey als spekulativ (Manzey, 1998, S. 806). Aber auch Boucsein (2012, S. 360) verweist auf die ausstehende multivariate psychophysiologische Forschung, um die zugeordneten Aktivierungsindikatoren als solche zu verifizieren.

2.3 Zusammenfassung und Eingrenzung

Zur Eingrenzung des Begriffs Beanspruchung wird in der vorliegenden Arbeit das Belastungs-Beanspruchungs-Modell herangezogen. Die Beanspruchung beschreibt dabei die komplexe Wechselwirkung zwischen den äußeren Anforderungen der Aufgabe und den Merkmalen der Person. Die physische, mentale und emotionale Beanspruchung wird durch die Ableitung der Belastungsfaktoren differenziert. Dabei bildet die physische Beanspruchung die muskulären (z. B. Bewegungsaufwand), die mentale Beanspruchung die aufgabenspezifischen (z. B. Aufgabenschwierigkeit) und die emotionale Beanspruchung die ausführungsspezifischen (z. B. Erfolg bei der Aufgabenbewältigung) Belastungsfaktoren der Aufgabe ab. Die Abgrenzung der mentalen Beanspruchung zum Stress ist nur eingeschränkt möglich. Im Vergleich zur Stressforschung, werden in der mentalen Beanspruchungsforschung und in der vorliegenden Arbeit nur diejenigen Ursachen fokussiert, die sich auf die menschliche Informationsverarbeitung auswirken. Im Vordergrund der vorliegenden Arbeit stehen zudem kurzfristige und tendenziell negative Auswirkungen der Belastung, welche als Momentaufnahme mentaler Beanspruchung im Laborkontext operationalisiert und erfasst werden.

Aus dem Kapitel 2.2 ist ersichtlich, dass die mentale Beanspruchung durch unterschiedliche theoretische Modelle abbildbar ist. Zunächst einmal ist anzumerken, dass die beschriebenen Modelle der Einordnung der mentalen Beanspruchung in das theoretische Gefüge dienen. Eine Bestätigung oder Widerlegung der oben aufgeführten Modelle (deduktives Vorgehen) ist kein Ziel der vorliegenden (induktiv ausgerichteten) Arbeit. Vor dem Hintergrund der psychophysiologischen Erfassung mentaler Beanspruchung sind empirisch und messtheoretisch ausgerichtete Modelle, wie die Kapazitätsmodelle, von Bedeutung. Diese definieren die mentale Beanspruchung als eine Auslastung bzw. Aktivierung der zentralnervösen Kapazität (repräsentiert als Ressource oder Energie) bei der menschlichen Informationsverarbeitung.

Die ressourcentheoretischen Modelle bieten aufgrund der Betonung struktureller Aspekte eine solide theoretische Grundlage für die Operationalisierung und Strukturierung mentaler Beanspruchung im Experimentaldesign der vorliegenden Arbeit. Unter der Annahme einer begrenzten Verarbeitungskapazität nach Kahneman (1973) kann die zur Aufgabenbewältigung benötigte Kapazitätmenge durch eine Zunahme der Aufgabenschwierigkeit angehoben werden. Diese höhere Kapazitätsauslastung spiegelt die mentale Beanspruchung der Person wider. Das Multiple-Ressourcen-Modell von Wickens (2002, 2008) stellt eine zusätzliche Taxonomie zur Verfügung, welche als Orientierungshilfe bei der Operationalisierung mentaler Beanspruchung in unterschiedlichen Aufgabentypen herangezogen werden kann.

Ein Erklärungsansatz für die Manifestation mentaler Beanspruchung in der physiologischen Aktivität bieten die aktivierungstheoretischen Modelle. Die mentale Beanspruchung ist dabei definiert als die energetische Aktivierung zentraler Informationsverarbeitungssysteme, die erforderlich ist, um eine Aufgabe zu bewältigen. Aktivierungstheoretische Modelle bilden die Grundlage, um die mentale Beanspruchungsstärke als eine Ausprägung physiologischer Aktivität zu interpretieren (z. B. Bartenwerfer, 1969). Zudem verdeutlichen die aktivierungstheoretischen Modelle von Pribram und McGuinness (1975) sowie von Backs und Boucsein (2009) bzw. Boucsein (2012, S. 353–362), dass das Konstrukt der mentalen Beanspruchung auf zentraler und physiologischer Ebene nicht als ein eindimensionales Konstrukt zu verstehen ist. So können (zumindest auf theoretischer Basis) die physiologischen Prozesse mentaler Beanspruchung von den physiologischen Prozessen motorischer oder emotionaler Prozesse abgegrenzt und deren Wechselwirkung dargestellt werden (siehe Backs & Boucsein, 2009; Boucsein, 2012, S. 353–362). Die von Backs und Boucsein (2009) vorgenommene Zuweisung physiologischer Aktivierungsindikatoren zu den Aktivierungssystemen ist jedoch höchst zweifelhaft (Manzey, 1998) und wird in der vorliegenden Arbeit nicht als Grundwahrheit angenommen. Nichtsdestoweniger soll der theoretische Ansatz von Backs und Boucsein (2009) dahingehend verfolgt werden, dass physiologische Aktivierungsindikatoren der

mentalener Beanspruchung sowie benachbarten Konstrukten (z. B. Emotionen) auf empirischer Basis zugeordnet werden.

3 Theoretische Grundlagen: Psychophysiologische Indikatoren

Die Verwendung der mentalen Beanspruchung für eine Schnittstelle zwischen Mensch und Maschine ist an die Herausforderung geknüpft, den nicht direkt einsehbaren Prozess der menschlichen Informationsverarbeitung zuverlässig zu messen. Eine zu diesem Zweck geeignete Methode ist die physiologische Messung der mentalen Beanspruchung. Mit diesem Thema setzt sich dieses Kapitel auseinander. Neben der physiologischen Messmethode, kann die mentale Beanspruchung auch mit der subjektiven Methode (z. B. Fragebogen) oder der leistungsbezogenen Methode (z. B. Anzahl der Ausführungsfehler) gemessen werden. Gemäß O'Donnell und Eggemeier (1986) birgt jede Methodenklasse ihre eigenen Vor- und Nachteile. Im Rahmen einer beanspruchungssensitiven Mensch-Maschine-Schnittstelle ist die physiologische Messmethode der subjektiven und leistungsbezogenen Messmethode aufgrund folgender Vorteile vorzuziehen: Physiologische Messungen können ohne die Unterbrechung der Tätigkeit oder des Arbeitsprozesses durchgeführt werden (vgl. Kramer, 1991). Auch eine Leistungsmessung erfüllt dieses Kriterium. Diese erfordert allerdings das Ausüben einer Tätigkeit, deren Leistung objektiv beurteilbar ist. Letzteres ist für die physiologische Messung nicht erforderlich und macht eine Messung auch in Abwesenheit eines beobachtbaren Verhaltens möglich. Des Weiteren ermöglicht die physiologische Messung eine kontinuierliche Erfassung mit kurzzeitigen Zustandsänderungen, die im Bereich weniger Sekunden bis Minuten liegen. Dadurch lässt sich ein zeitlicher Verlauf der mentalen Beanspruchung abbilden.

Wenngleich die Messung mentaler Beanspruchung mit Hilfe physiologischer Messgrößen bereits seit Jahrzehnten gängige Forschungspraxis ist, ist die Suche nach geeigneten physiologischen Indikatoren menschlicher Informationsverarbeitung noch lange nicht abgeschlossen. Die Schwierigkeit der physiologischen Messung mentaler Beanspruchung liegt nicht zuletzt an der komplexen Steuerung physiologischer Aktivität durch das zentrale und periphere Nervensystem, welches einleitend in Abschnitt 3.1 beschrieben wird. Über die letzten Jahrzehnte entstand eine Vielzahl an physiologischen Messverfahren und Parametern, die als potentielle Indikatoren mentaler Beanspruchung gelten. In Abschnitt 3.2 und 3.3 werden die gängigen kardialen und elektrodermalen Parameter in Hinblick auf die physiologischen Grundlagen, die Messung und Quantifizierung des physiologischen Signals sowie die Forschungsbefunde beschrieben. Die Forschungsbefunde der physiologischen Parameter werden dabei hinsichtlich ihrer Eignung als mentale Beanspruchungsindikatoren anhand der Gütekriterien Sensitivität, Diagnostizität bzw. Validität (siehe Anhang A.2) bewertet. Kapitel 3.4 gibt einen Überblick der aktuellen Forschungslage auf dem Gebiet der multivariaten Analyse physiologischer Aktivität sowie zu der relativ jungen Forschungsdisziplin des Psychophysiological Computing.

3.1 Regulation durch das autonome Nervensystem

Warum der menschliche Körper bei übermäßiger mentaler (aber auch emotionaler und körperlicher) Belastung mit einer physiologischen Reaktion (z. B. dem Anstieg der Herzschlagfrequenz) reagiert, kann mit Hilfe autonomer Regulationsmechanismen des peripheren Nervensystems erklärt werden. Damit sich der Körper auf verändernde Belastungen und Anforderungen der Situation anpassen kann, besitzt er das vegetative Nervensystem, auch als autonomes Nervensystem (ANS) bezeichnet (Schmidt, Thews & Lang, 2000). Das ANS hat die Aufgabe, das Gleichgewicht lebenswichtiger Körperfunktionen, bspw. Atmung und Stoffwechsel, aufrechtzuerhalten (Homöostase) und somit optimale Bedingungen für die Zellfunktionen zu schaffen (Schmidt et al., 2000). Dies

geschieht durch eine automatische und nicht bewusst wahrnehmbare Steuerung von Organen und Körperfunktionen. Das ANS besteht aus drei Teilsystemen: dem Darmnervensystem, dem Sympathikus und dem Parasympathikus. Vor dem Hintergrund physiologischer Beanspruchungsreaktionen wird nur auf die letzten zwei Teilsysteme eingegangen. Während der Sympathikus den Körper in Aktionsbereitschaft versetzt, um diesen bspw. bei Bedrohungen auf Angriff, Flucht (Fight-or-Flight Reaktion) oder andere geforderten Anstrengungen vorzubereiten, ist der Parasympathikus für die Steuerung der Erholungs- und Regenerationsphasen zuständig (Kramer, 1991). Zu diesem Zweck werden die inneren Organe vom sympathischen und parasympathischen Nervensystem in ihrer Aktivität angeregt oder gehemmt. Bspw. steigert der Sympathikus aktivierende Prozesse wie die Herzschlagfrequenz und hemmt gleichzeitig andere, nicht erforderliche Körperfunktionen, wie die Verdauung. Während manche Organe und Organsysteme nur vom Sympathikus oder Parasympathikus innerviert werden (z. B. rein sympathische Innervation der Schweißdrüsen), werden andere Organsysteme, wie der Herzmuskel, durch beide Teilsysteme reguliert (Überblick siehe Schmidt et al., 2000). Nach Schmidt et al. (2000) sind die Effekte bei der Innervation durch beide Teilsysteme überwiegend antagonistischer Natur. Bei der antagonistischen Steuerung bedeutet die Erregung eines Teilsystems gleichzeitig die Hemmung des anderen Teilsystems. Der Antagonismus schafft eine autonome Balance im Sinne einer homöostatischen Anpassung. Mittlerweile ist die Annahme, dass der autonome Zustand der Organsysteme sich auf einem eindimensionalen, bipolaren Kontinuum von sympathischer bis parasympathischer Dominanz bewegt, veraltet. Aus der Perspektive der homöostatischen Regulation ist solch eine antagonistische Aktivierung naheliegend, da dieses Aktivierungsmuster (im Vergleich zur simultanen Aktivierung beider Teilsysteme) die Ressourcen zur Erhaltung eines homöostatischen Sollwerts maximiert (Berntson, Cacioppo & Grossman, 2007). So erscheint es nicht verwunderlich, dass das homöostatische Konzept über Jahrzehnte nicht nur das Bild zur Funktionsweise autonomer Steuerung, sondern auch psychophysiologischer Indikatoren geprägt hat (Berntson et al., 2007).

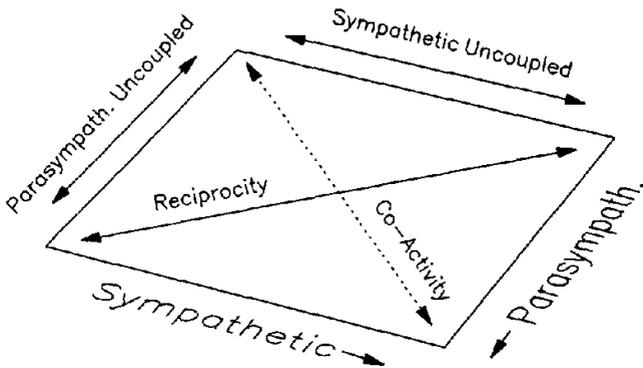


Abbildung 3.1: Die zweidimensionale Ebene des Autonomic Space Modells. Je nach Aktivierung des sympathischen und parasympathischen Systems kann sich eine Entkopplung, Koaktivierung oder Reziprozität ergeben. Grafik aus Berntson et al. (1991, S. 469).

Neuere Ansätze der autonomen Steuerung verweisen darauf, dass neben einer antagonistischen Aktivierung Belege für die Koaktivierung (gleichzeitige Aktivierung des Sympathikus und Parasympathikus) und Entkopplung (alleinige Aktivierung des Sympathikus oder Parasympathikus) sympathischer und parasympathischer Systeme existieren. Hier muss mindestens eine zweidimensionale Ebene zur Repräsentation der Komplexität autonomer Kontrollmechanismen aufgespannt werden (Berntson et al., 1991; Berntson, Cacioppo, Quigley & Fabro, 1994; Berntson et al., 2007; Berntson, Norman, Hawkley & Cacioppo, 2008). In Abbildung 3.1 ist die im *Autonomic Space Modell* von Berntson et al. (1991) beschriebene zweidimensionale Ebene, bestehend aus den beiden Achsen sympathischer (engl.: sympathetic) und parasympathischer (engl.: parasympathetic) Aktivierung veranschaulicht. Gemäß diesem Modell sind mehr als zwei antagonistische Formen der autonomen Steuerung möglich. Diese sympathische und parasympathische Flexibilität der autonomen Steuerung wird nach Sterling und Eyer (1988) als allostatistische bzw. nach Berntson et al. (2007) als allodynamische Steuerung bezeichnet (Vertiefung siehe Berntson et al., 2007). Die autonome Steuerung der Organe wird dabei nicht alleine durch

das Prinzip der autonomen Balance (Homöostase) im Sinne einer antagonistischen Verschaltung zwischen Sympathikus und Parasympathikus bestimmt. Die autonome Steuerung bildet vielmehr eine flexible Regulation (Allostase bzw. Allodynamic) aus Antagonismus, Koaktivierung und Entkopplung. Diese Flexibilität ermöglicht es dem Körper, adaptiv auf verschiedene innere und äußere Belastungen und Anforderungen zu reagieren (Berntson et al., 2007, 2008).

Die Tatsache, dass Sympathikus und Parasympathikus unabhängig voneinander auf ein Organsystem einwirken können, spielt vor allem bei der Interpretation und Messung derjenigen Organe eine Rolle, die von beiden autonomen Teilsystemen innerviert werden (z. B. das Herz). Ein gängiges Beispiel hierfür ist die Messung der Herzschlagfrequenz. Eine beobachtbare Erhöhung dieser kann gleichzeitig zurückgeführt werden auf:

- (1) Eine antagonistische Steuerung durch sympathische Aktivierung bei gleichzeitiger parasympathischer Hemmung.
- (2) Eine entkoppelte Steuerung durch eine reine sympathische Aktivierung bzw. eine reine parasympathische Hemmung.
- (3) Eine gleichgerichtete Steuerung (Koaktivierung) mit verstärkter sympathischer Aktivierung bzw. verstärkter parasympathischer Hemmung.

Aber auch wenn keine Veränderung in der Herzschlagfrequenz zu beobachten ist, kann sich eine nicht sichtbare Veränderung der autonomen Aktivierung vollzogen haben. Nach Berntson et al. (2007) ist eine verschiedenartige Wirkung der unterschiedlichen Kombinationen aus sympathischer und parasympathischer Aktivierung auf psychische Prozesse nicht auszuschließen. Gleichzeitig kann sie sich hinter inkonsistenten Befunden in der psychophysiologischen Forschung verbergen.

3.2 Kardiale Aktivität

In den folgenden Kapiteln wird ein Überblick zu dem physiologischen Hintergrund (Kapitel 3.2.1), der Erfassung (Kapitel 3.2.2), der Quantifizierung (Kapitel 3.2.3) und den empirischen Befunden (Kapitel 3.2.4) der kardialen Aktivität gegeben. Die folgende Darstellung erfolgt in Anlehnung an Berntson et al. (2017), Fahrenberg (2001), Feigenspan (2017, S. 250–284) und Malik et al. (1996).

3.2.1 Physiologische Grundlagen kardialer Aktivität

Das kardiale System ist eng mit dem vaskulären System verbunden. Gemeinsam bilden sie das kardiovaskuläre System, welches dafür verantwortlich ist, dass der Körper mit Blut versorgt wird. Das Herz hat hierbei die Funktion einer Um-

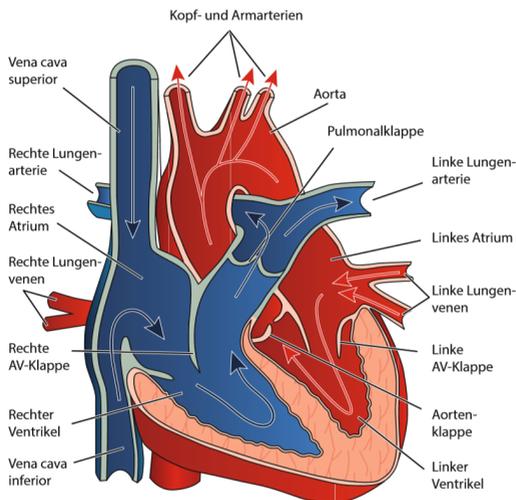


Abbildung 3.2: Frontalansicht des menschlichen Herzens. Der Verlauf des Blutstroms ist durch die Pfeile gekennzeichnet. Rot repräsentiert sauerstoffreiches und blau sauerstoffarmes Blut. Grafik aus Feigenspan (2017, S. 251).

wälzpumpe, die für die Bewegung des Blutes in den Gefäßen sorgt. Anatomisch besteht das Herz aus einer linken und einer rechten Herzhälfte, zwei muskulösen Hohlorganen, die wiederum durch die Herzklappen (Atrioventrikularklappen) in eine Herzkammer (Ventrikel) und einen vorgeschalteten Vorhof (Atrium) unterteilt sind (siehe Abbildung 3.2). Die Pumpaktivität des Herzens entsteht durch eine wiederkehrende Erschlaffung (Diastole) und Kontraktion (Systole) der Herzkammern, dem sog. kardialen Zyklus. In der Diastole werden die Herzkammern mit Blut gefüllt und während der Kontraktion wird das Blut in die Arterien abgeleitet. Weitere Herzklappen fungieren als Ventile und verhindern einen Rückfluss des Blutes in die Herzkammern.

Der rhythmische Herzzyklus aus Diastole und Systole, inklusive der zeitlichen Abfolge von Vorhof- und Herzkammeraktivität, unterliegt einem bioelektrischen Erregungsablauf des Herzens. Dieser hat für gewöhnlich seinen Ursprung im sog. Sinusknoten des rechten Vorhofs (siehe Abbildung 3.3). In Ruhe gibt der Sinusknoten eine Frequenz von ca. 60 bis 90 Schlägen pro Minute vor und fungiert dadurch als der primäre Schrittmacher des Herzzyklus. Ausgehend vom Sinusknoten durchläuft die Erregung in ca. 100 ms die Arbeitsmuskulatur beider Vorhöfe. Da die Grenze zwischen den Vorhöfen und den Herzkammern aus nicht erregbarem Bindegewebe besteht, erfolgt die Erregung der Herzkammern zeitverzögert über das ventrikuläre Erregungsleitsystem, bestehend aus AV-Knoten (Atrioventrikularknoten) und Tawara-Schenkel (siehe Abbildung 3.3).

In Bezug auf die autonome Steuerung sind sowohl sympathische als auch parasympathische Einflüsse feststellbar (Schmidt et al., 2000, S. 481). Eine sympathische Aktivierung erhöht und eine parasympathische Aktivierung verringert die Herzschlagfrequenz. Es ist anzunehmen, dass im Ruhezustand ein ständiger Erregungszustrom des Parasympathikus (Vagustonus) erfolgt, da die autonome Frequenz des Herzens (keine sympathische oder parasympathische Innervation) über der Ruhefrequenz des Herzens liegt (Schmidt et al., 2000, S. 482). Die Trennung sympathischer und parasympathischer Anteile in der Herzschlagfrequenz ist aufgrund der dualen Innervation sowie des komplexen

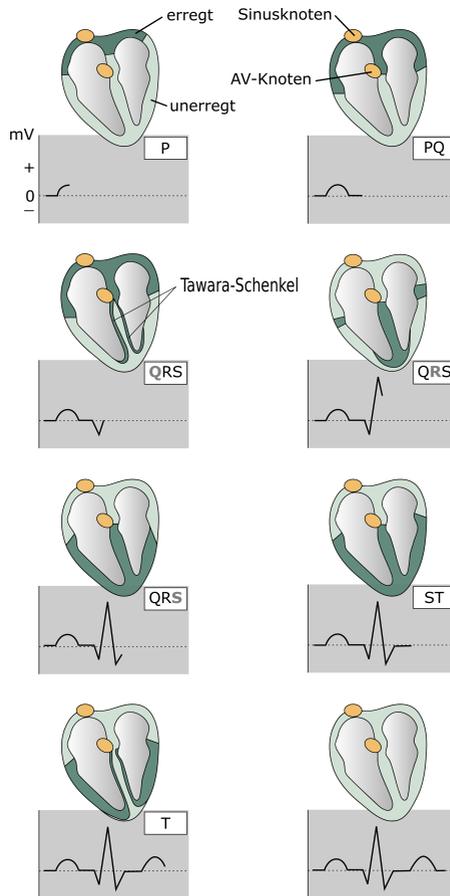


Abbildung 3.3: Erregungsausbreitung eines Herzzyklus im Elektrokardiogramm (EKG). Grafik modifiziert nach Feigenspan (2017, S. 280).

Zusammenspiels zwischen Sympathikus und Parasympathikus (siehe Kapitel 3.1) nicht ohne weiteres möglich.

3.2.2 Messung kardialer Aktivität

Ein nicht invasives Verfahren zur Messung kardialer Aktivität ist das Elektrokardiogramm (EKG). Dieses gibt die bioelektrischen Vorgänge bei der Erregung des Herzens wieder. Die Spannungsunterschiede zwischen erregten und unerregten Zellstrukturen, welche sich aus Generierung, Ausbreitung und Abbau elektrischer Impulse am Herzen ergeben, werden bis an die Körperoberfläche geleitet. Hier können die Spannungsänderungen mit Hilfe von Kontaktelektroden gemessen werden. Im Grunde repräsentiert das EKG die elektrischen Potentialunterschiede zwischen zwei Elektroden über die Zeit. Die Elektroden werden dabei an definierte Körperstellen platziert. Über die Jahre wurden unterschiedliche EKG-Ableitsysteme entwickelt. Diese unterscheiden sich nicht nur in der Platzierung der Elektroden an der Körperoberfläche, sondern auch in der Anzahl der Elektroden und somit der Anzahl der EKG-Ableitungen. Als internationaler Standard sorgen standardisierte Ableitsysteme für eine Vergleichbarkeit der Ergebnisse. Unterschieden werden die Ableitsysteme in Extremitäten-Ableitungen (Platzierung der Elektroden an Armen und Beinen) und Brustwandableitungen (Platzierung der Elektroden auf der Brust), aber auch bipolare Ableitungen (Potentialdifferenz zwischen zwei gleichberechtigten Elektroden) und unipolare Ableitungen (Potentialdifferenz zum Potentialmit-

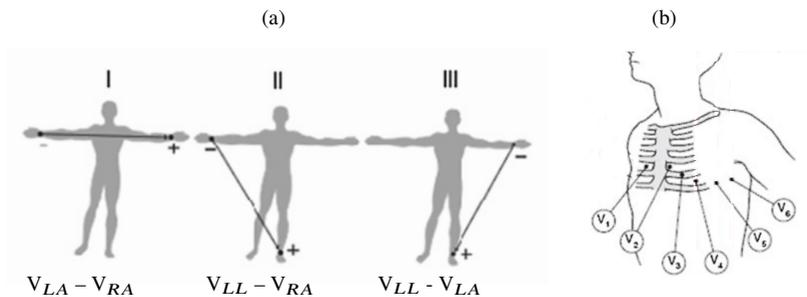


Abbildung 3.4: Ableitung der kardialen Aktivität an den Extremitäten nach Einthoven (a) und an der Brustwand nach Wilson (b). V_{LA} : Potential des linken Arms; V_{RA} : Potential des rechten Arms; V_{LL} : Potential des linken Beins. Grafik modifiziert nach Gupta et al. (2014, S. 8–9).

telwert der anderen Elektroden). In Abbildung 3.4 ist die Standardableitung nach Einthoven als Beispiel einer bipolaren Extremitäten-Ableitung und die Wilson-Ableitung als Beispiel einer unipolaren Brustwandableitung dargestellt.

In Abhängigkeit von der aktuellen Richtung des elektrischen Potentials des Herzens können positive und negative Ausschläge gemessen werden. Da von der negativen zur positiven Elektrode abgeleitet wird, deutet ein positiver Ausschlag im EKG darauf hin, dass sich der elektrische Impuls in Richtung der Elektrode bewegt, und umgekehrt. Zur Veranschaulichung der Erregungsausbreitung im Verlauf eines Herzzyklus im EKG soll auf die charakteristischen Wellen und Zacken im Signal in Anlehnung an Feigenspan (2017, S. 279–280) eingegangen werden. Diese sind in Abbildung 3.3 dargestellt. Die im Folgenden beschriebene Ausführung bezieht sich auf eine Ableitung nach Einthoven. Die Ausbreitung der Erregung über die beiden Vorhöfe, beginnend am Sinusknoten, ist als P-Welle im EKG zu erkennen. Die Dauer für die Überleitung der Erregung auf die Herzkammern über den AV-Knoten wird als PQ-Intervall (Beginn P-Welle bis Beginn Q-Zacke) bezeichnet und dauert in etwa 200 ms. Die Spannungsunterschiede innerhalb der Vorhöfe sinken aufgrund der vollständigen Erregung auf Null ab. Im Anschluss erfolgt der QRS-Komplex, welcher die Erregungsausbreitung in den Herzkammern kennzeichnet. Hierbei repräsentiert die Q-Zacke die Erregung der Herzscheidewand, die R-Zacke die Erregung der Muskulatur der Herzkammer in Richtung der Herzspitze und die S-Zacke überwiegend die Erregung der Muskulatur der linken Herzkammer. Die große Spannungsamplitude von 1 bis 2 mV im R-Zacken ergibt sich vor allem aufgrund der großen Muskelmasse des linken Ventrikels. Der mit QRS-Komplex zeitgleiche Erregungsabbau der Vorhöfe ist im EKG nicht sichtbar. Sind beide Herzkammern erregt, ist die Spannungsdifferenz wieder Null. Im EKG wird diese Zeit als ST-Intervall gekennzeichnet. Die T-Welle repräsentiert die Erregungsrückbildung der Herzkammern. Damit ist der Herzzyklus abgeschlossen und beginnt wieder von neuem.

3.2.3 Quantifizierung des kardialen Signals

Zur Bewertung der kardialen Aktivität ist es zunächst einmal notwendig, aussagekräftige Parameter aus dem Rohsignal des EKG zu extrahieren. Eines der charakteristischsten Merkmale des EKG-Signals sind die R-Zacken, welche für die Berechnung einer Reihe von Parametern genutzt werden. Der am häufigsten anzutreffende Parameter ist die Herzschlagfrequenz, auch Herzrate (HR) bezeichnet. Berechnet wird die HR als Anzahl der Herzschläge pro Minute. Die reziproke Kenngröße der Herzschlagfrequenz ist die Herzperiode (HP), auch als Schlagintervall oder RR-Intervall bezeichnet. Sie gibt den Abstand zwischen zwei R-Zacken in ms an. Beide Parameter lassen sich durch Quotientenbildung ($HR = 60000/HP$) ineinander überführen. Zusätzlich können Parameter aus den natürlich vorkommenden physiologischen Schwankungen der HP bzw. HR, die sog. Herzratenvariabilität (HRV)¹, berechnet werden. Die Variabilität wird durch das parasympathische und sympathische Nervensystem beeinflusst. Dabei schlägt sich der Einfluss des parasympathischen Nervensystems bedeutend schneller in der HR bzw. HP nieder (Reaktionszeit von 0.2 bis 0.6 s) als eine rein sympathische Innervation (Reaktionszeit von wenigen Sekunden). Die Schwankungen der HR bzw. HP werden auf schnelle und langsame Anteile hin parametrisiert und untersucht. Insgesamt gibt es weit über 100 unterschiedliche Parameter der HRV (z. B. 115 Parameter bei Smith, Owen & Reynolds, 2013). Einen Überblick bieten z. B. Berntson et al. (1997), Fahrenberg (2001), Malik et al. (1996) und Shaffer und Ginsberg (2017). Aufgrund dieser Vielfalt wird im Folgenden nur auf diejenigen Parameter der HRV eingegangen, die in der Forschungsliteratur gebräuchlich sind. Eine tabellarische Übersicht der in dieser Arbeit verwendeten Parameter findet sich in Unterkapitel 5.1.2.

Die HRV kann als eine Metrik aus dem Zeitbereich, dem Frequenzbereich oder als nichtlineare Kenngröße quantifiziert werden. Zudem werden die HRV-

¹ In der vorliegenden Arbeit wird der Begriff Herzratenvariabilität (Berechnung basierend auf den RR-Intervallen) synonym zum Begriff der Herzperiodenvariabilität (Berechnung basierend auf der HP) verwendet.

Parameter dahingehend unterschieden, ob sie aus einer Langzeit- (24 Stunden), Kurzzeit- (ca. 5 min) oder Ultra-Kurzzeitmessung (< 5 min) hervorgehen. Langzeit- und Kurzzeit- bzw. Ultra-Kurzzeitparameter sind nicht deckungsgleich. Zum einen ist die kardiale Aktivität in einem 24-Stunden-Intervall einem größeren Reiz- und Reaktionsspektrum ausgesetzt als in einem kurzen Zeitintervall. Zum anderen werden bei einer Langzeitmessung neben den schnellen auch sehr langsame (niederfrequente) Zyklusschwankungen (z. B. circadianer Rhythmus) miterfasst (vgl. Malik et al., 1996; Shaffer & Ginsberg, 2017). Für die vorliegende Arbeit sind vor allem die Ultra-Kurzzeitmaße (ca. 1 min) interessant, da eine adaptive Mensch-Maschine-Schnittstelle nur mit einer möglichst geringen Zeitverzögerung auf eine Zustandsänderung der Person reagieren sollte.

HRV-Parameter aus dem Zeitbereich

Zur Quantifizierung der HRV im Zeitbereich werden neben statistischen Verfahren auch geometrische Methoden verwendet. Beide Methoden geben das Ausmaß der Variabilität der RR-Intervalle in ms an. Für die Berechnung der HRV wird empfohlen, die RR-Intervalle vorab um Extrazystolen² zu bereinigen (vgl. Malik et al., 1996; Shaffer & Ginsberg, 2017). Die bereinigten RR-Intervalle werden als NN-Intervalle (normal to normal) bezeichnet. Statistische Maße der HRV sind u. a. die Standardabweichung der NN-Intervalle (SDNN), das mittlere Quadrat der Abweichung aufeinanderfolgender NN-Intervalle (MQSD) sowie der prozentuale Anteil aufeinanderfolgender NN-Intervall-Differenzen, die kleiner als 50 ms (pNN50) bzw. 20 ms (pNN20) sind. Zu den geometrischen Maßen zählt u. a. der HRV-Dreiecksindex (Triang). Dieser setzt sich aus dem Integral der NN-Intervall-Dichteverteilung (also dem Histogramm aller NN-Intervalle) dividiert durch das Maximum der Dichteverteilung zusammen (vgl. Malik et al., 1996). Abbildung 3.5 veranschaulicht die Herleitung des

² Herzschläge außerhalb des physiologischen Herzrhythmus.

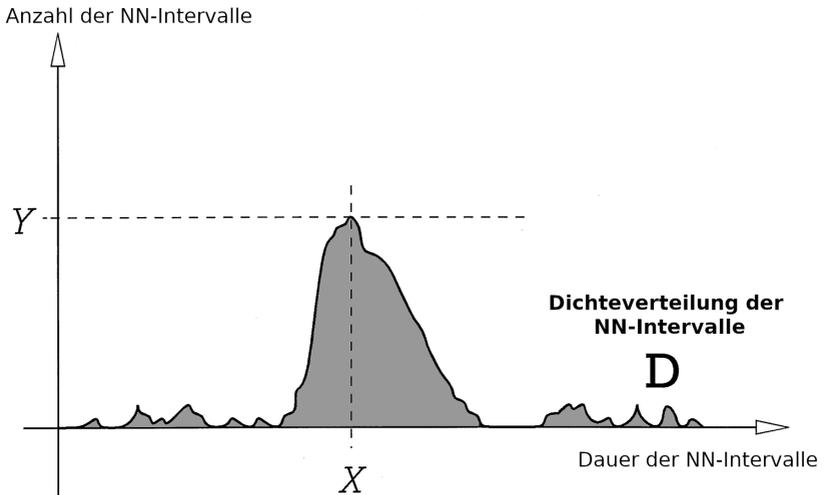


Abbildung 3.5: Der HRV-Dreiecksindex (Triang) ergibt sich aus der Division der Dichteverteilung der NN-Intervalle (D) mit dem Maximum der Dichteverteilung (Y). Grafik modifiziert nach Malik et al. (1996, S. 357).

HRV-Dreiecksindex. Die Berechnung der Parameter HR, HP, MQSD und pNN50 als Ultra-Kurzzeitmaße ist auch für Zeitabschnitte < 1 min valide möglich. Für den Parameter Triang ist die Validität für einen so kurzen Zeitbereich nicht nachgewiesen (Nussinovitch et al., 2011; Salahuddin, Cho, Jeong & Kim, 2007; Shaffer, Shearman & Meehan, 2016; zsf. Shaffer & Ginsberg, 2017).

HRV-Parameter aus dem Frequenzbereich

Im Frequenzbereich wird die HRV in drei Frequenzkomponenten zerlegt und für die jeweilige Ausprägung die spektrale Leistungsdichte berechnet (siehe Abbildung 3.6). Diese werden als Absolutwert in ms^2 oder als Relativwert in ν (normalisierte Einheit mit dem Wertebereich 0 bis 1; engl.: normalised units)

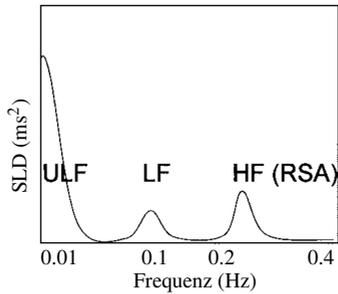


Abbildung 3.6: Spektrale Leistungsdichte (SLD) der Herzperiode. Grafik modifiziert nach Berntson et al. (2017, S. 196).

angegeben.³ Die Very-Low-Frequency Komponente (VLF) liegt im Bereich 0,0033 bis 0,04 Hz. Zur validen Erfassung ist es ratsam, einen Messbereich von mind. 270 s zu verwenden (Shaffer et al., 2016), weshalb dieser Parameter in der vorliegenden Arbeit nicht weiter berücksichtigt wird. Die Low-Frequency Komponente (LF bzw. der Relativwert LFn) liegt im Bereich 0,04 bis 0,15 Hz. Einige Autoren (Baek, Cho, Cho & Woo, 2015; Shaffer et al., 2016; Shaffer & Ginsberg, 2017) empfehlen einen Messbereich von 90 s, wobei auch kürzere Messbereiche von 50 s berichtet werden (z. B. Salahuddin et al., 2007). In der Vergangenheit wurde die LF als ein Indikator sympathischer Aktivität diskutiert. Neuere Erkenntnisse zeigen allerdings auf, dass die LF ebenfalls der parasympathischen Regulation unterworfen ist.

Die High-Frequency Komponente (HF bzw. der Relativwert HFn) ist im Bereich von 0,15 bis 0,40 Hz angesiedelt und kann in einem Zeitintervall von 60 s gemessen werden (Shaffer & Ginsberg, 2017). Diese Frequenzkomponente ist von der Atemaktivität abhängig und gilt gemeinhin als ein Repräsentant parasympathischer Aktivität (Fahrenberg, 2001; Shaffer & Ginsberg, 2017). Die HF-Komponente korrespondiert mit der respiratorischen Sinusarrhythmie (RSA), welche die atemsynchronen Schwankungen der HP bzw. HR abbildet. Die RSA entsteht durch die Beschleunigung der HR beim Einatmen

³ Der Relativwert ergibt sich aus dem Quotienten unterschiedlicher Frequenzkomponenten.

und der Verlangsamung beim Ausatmen. Es ist zu beachten, dass die HF-Komponente nicht den Vagustonus, sondern lediglich die parasympathische Kontrolle der HR bzw. HP widerspiegelt.

Ausgehend von der Annahme, dass die LF-Komponente überwiegend sympathischen und die HF-Komponente überwiegend parasympathischen Einflüssen unterworfen ist, wurde das Verhältnis zwischen der LF- und HF-Komponente (LF/HF) lange Zeit als ein Indikator für das Gleichgewicht zwischen sympathischen und parasympathischen Einflüssen angesehen (z. B. Malik et al., 1996; Pagani et al., 1986; Shaffer & Ginsberg, 2017). Diese Annahme ist allerdings aufgrund folgender Kritikpunkte umstritten: Erstens ist die LF-Komponente nicht nur sympathischen Einflüssen unterworfen. Zweitens kann das Zusammenspiel zwischen Sympathikus und Parasympathikus unterschiedliche Interaktionsformen annehmen. Drittens können bestimmte Einflussfaktoren, z. B. die Atmung, nicht gänzlich kontrolliert oder ausgeschlossen werden (Shaffer & Ginsberg, 2017). Die ursprüngliche Berechnung der LF/HF basierte auf einer Langzeitmessung. Nichtsdestoweniger geben einige Autoren auch kürzere Zeitintervalle von 20 bis 180 s an (z. B. Salahuddin et al., 2007; Shaffer et al., 2016).

Nichtlineare HRV-Parameter

Goldberger (1991) stellte fest, dass Herzzyklusschwankungen Eigenschaften komplex dynamischer Systeme aufweisen, die herkömmliche lineare Methoden nicht aufdecken können. Um die Komplexität und Unvorhersagbarkeit der HRV zu extrahieren, wurden nichtlineare Methoden entwickelt. Hierunter fällt die Berechnung der Entropie, wie der Sample-Entropie (SampEn). Sie repräsentiert die Vorhersagbarkeit der HRV über die Zeit (Shaffer & Ginsberg, 2017). Hohe Werte deuten auf eine geringe und niedrige Werte auf eine hohe Vorhersagbarkeit und Regelmäßigkeit der HRV hin (Beckers, Ramaekers & Aubert, 2001). Eine weitere Methode ist die trendbereinigende Fluktuationsanalyse von Peng, Havlin, Stanley und Goldberger (1995). Diese ermittelt die Korrelation zwischen

aufeinanderfolgenden NN-Intervallen über kürzere (DFA1) und längere (DFA2) Zeitabschnitte.

Morphologische EKG-Parameter

Neben den oben beschriebenen, rhythmischen Parametern, können auch morphologische Charakteristika des EKG-Signals, wie die T-Welle, genutzt werden. Vorgeschlagen wird die T-Wellen-Amplitude (T.amp) als ein Maß der sympathischen Aktivität. Die Verringerung der Amplitude wird als eine verstärkte sympathische Aktivierung gedeutet und ist mit anderen Maßen sympathischer Aktivität assoziiert (Kline, Ginsburg & Johnston, 1998; van Lien, Neijts, Willemssen & de Geus, 2015). In der Literatur besteht allerdings eine kontroverse Diskussion bezüglich der sympathischen Indikatorfunktion, da die T.amp indirekt auch durch den Parasympathikus moduliert werden kann (Annala, Yli-Hankala & Lindgren, 1994; Fahrenberg, 2001).

3.2.4 Befundlage zur kardialen Aktivität

Der Zusammenhang mentaler Beanspruchung mit der kardialen Aktivität wird in vielen Forschungsarbeiten mit einer verstärkten sympathischen Aktivierung des Herzens in mentalen Belastungssituationen erklärt. Dieser Zusammenhang erweckte bereits Anfang des 20-sten Jahrhunderts das Interesse der Forscher, weshalb über die Zeit eine Fülle an empirischen Studien entstand. Aus diesem Grund wird im Folgenden zusammenfassend auf besonders relevante Forschungsergebnisse im Rahmen der mentalen Beanspruchung eingegangen.

Der am häufigsten untersuchte kardiale Parameter ist die HR bzw. HP. Die HR bzw. HP gilt gemeinhin als ein Indikator für die körperliche und mentale Beanspruchung (Mulder, Mulder, Meijman, Veldman & van Roon, 2000). Nachgewiesen wurde ein signifikanter Anstieg der HR bzw. Abfall der HP bei steigender mentaler Beanspruchung und Anstrengung für ein breites Spektrum

an Aufgaben unter Labor- und Feldbedingungen (z. B. Backs, Lenneman & Sicard, 1999; Heine et al., 2017; Lenneman & Backs, 2007; Mukherjee, Yadav, Yung, Zajdel & Oken, 2011; Moya-Albiol et al., 2001; Salahuddin et al., 2007). Ein Überblick findet sich bei Ribback (2003). Zahlreiche Untersuchungen zeigen somit auf, dass die HR und HP sensitiv auf eine Steigerung der mentalen Beanspruchung reagieren. Nichtsdestoweniger gibt es ein paar Studien, die keinen Effekt mentaler Beanspruchung auf die HR bzw. HP finden (z. B. Fairclough & Houston, 2004; Wierwille, Rahimi & Casali, 1985). Im Gegensatz zur Sensitivität ist nach Manzey (1998) keine Diagnostizität der HR bzw. HP sichtbar. Die HR und HP zeigen eine Reaktion nicht nur auf mentale, sondern auch auf emotionale oder körperliche Belastungsfaktoren (zsf. Manzey, 1998).

Im Gegensatz zur HR ist ein Anstieg der HRV mit einem Abfall der mentalen Beanspruchung assoziiert (z. B. Porges, 1992; Stuiver & Mulder, 2014; van Roon, Mulder, Althaus & Mulder, 2004). Zahlreiche Studien belegen diesen negativen Zusammenhang bei Aufgaben unter Labor- und Feldbedingungen (z. B. Backs & Seljos, 1994; Backs et al., 1999; Mukherjee et al., 2011; Salahuddin et al., 2007; Sauvet et al., 2009; F. Schneider et al., 2017; zsf. siehe Ribback, 2003). Solche Befunde finden sich überwiegend für die klassischen Parameter aus dem Zeit- und Frequenzbereich (z. B. Bernardi et al., 2000; Boonnithi & Phongsuphap, 2011; Duschek, Muckenthaler, Werner & del Paso, 2009; Salahuddin et al., 2007; Stuiver, Brookhuis, de Waard & Mulder, 2014). Vor allem die LF-Komponente galt lange Zeit als das sensibelste Korrelat mentaler Beanspruchung. Diese Interpretation ist aus heutiger Sicht nicht mehr zweifelsfrei vertretbar. Zum einen, weil nicht nur sympathische, sondern auch parasympathische Einflüsse auf die LF-Komponente wirken können. Zum anderen, weil auch entgegengesetzte bis gar keine Zusammenhänge mit der mentaler Beanspruchung gefunden wurden (vgl. Fahrenberg, 2001). Zudem ist für die LF-Komponente eine zuverlässige Differenzierung mentaler Beanspruchungszustände lediglich zwischen Ruhe- und Aufgabenbedingung, jedoch nicht zwischen den feineren Abstufungen der mentalen Beanspruchung möglich (siehe Chanel, Rebetez, Bétrancourt & Pun, 2011; Manzey, 1998).

Die Diagnostizität der HRV-Parameter ist eingeschränkt. Hier finden sich die gleichen Einflussfaktoren, wie bereits bei der HR bzw. HP. Aus diesem Grund gelten auch die HRV-Parameter als unspezifische Beanspruchungsindikatoren (Manzey, 1998). Die Ursachen für die mangelnde Diagnostizität der HR bzw. HP und der HRV-Parameter (z. B. Fairclough & Houston, 2004) sind divers und reichen von Störvariablen, wie Emotionen oder externen Umweltfaktoren, bis hin zu inhärenten psychophysiologischen Prozessen, wie der allodynamischen ANS-Steuerung (Berntson et al., 2007; siehe Kapitel 3.1).

Im Vergleich zu den zeit- und frequenzbezogenen Parametern ist die Anzahl der empirischen Befunden zum Zusammenhang nichtlinearer HRV-Maße mit der mentalen Beanspruchung begrenzt. Ein signifikanter Anstieg des Parameters $SamEn$ sowie ein Abfall der Parameter $DFA1$ und $DFA2$ wird bei Tätigkeiten mit höherer subjektiver mentalen Beanspruchung belegt (z. B. Castaldo et al., 2018; Martin et al., 2016; Rieger, Stoll, Kreuzfeld, Behrens & Weippert, 2014). Im Vergleich dazu konnte die Studie von Mukherjee et al. (2011) keinen Zusammenhang zwischen der mentalen Beanspruchung und den nichtlinearen HRV-Parametern nachweisen. Aufgrund des Mangels an konsistenten Befunden besteht bisher kein Konsens darüber, welche der nichtlinearen Methoden die mentale Beanspruchung sensitiv und spezifisch abbilden kann (vgl. F. Schneider et al., 2017; Voss, Schulz, Schroeder, Baumert & Caminal, 2009).

Ebenfalls ist die Bewertung der T-Wellen-Amplitude bezüglich ihrer Sensitivität nur begrenzt möglich. Hinweise auf die Sensitivität im Sinne der Verringerung der T-Wellen-Amplitude aufgrund steigender mentaler Beanspruchung liefern die Studien von Heine et al. (2017), Heslegrave und Furedy (1979), Myrtek et al. (1994), Scher, Furedy und Heslegrave (1984) und van Lien et al. (2015). Allerdings besteht aufgrund von divergierenden Befunden (vgl. Furedy, Szabo & Peronnet, 1996) bislang kein Konsens darüber, ob dieser Parameter die mentale Beanspruchung valide abbildet.

Zusammenfassend existieren ausreichend Befunde, welche die Sensitivität der meisten kardialen Parameter für mentale Beanspruchungszustände belegen.

Noch nicht umfänglich erforscht, jedoch vielversprechend, sind die nichtlinearen HRV-Parameter. Eine Diagnostizität sowie Validität ist für die kardialen Parameter nicht gegeben. Folglich ist bei den kardialen Parametern von unspezifischen Beanspruchungsindikatoren auszugehen.

3.3 Elektrodermale Aktivität

Die elektrodermale Aktivität (EDA) beschreibt die Änderung des elektrischen Leitwerts bzw. elektrischen Widerstands der Hautoberfläche. Neben der kardialen Aktivität ist die EDA eines der am häufigsten verwendeten physiologischen Signale in der psychophysiologischen Forschung. In Anlehnung an Boucsein (2001), Boucsein (2012) und Dawson, Schell und Filion (2017) wird die EDA in Hinblick auf die physiologischen Grundlagen (Kapitel 3.3.1), die Erfassungsmethode (Kapitel 3.3.2), die Quantifizierung (Kapitel 3.3.3) und den Zusammenhang mit der mentalen Beanspruchung (Kapitel 3.3.4) beschrieben.

3.3.1 Physiologische Grundlagen elektrodermalen Aktivität

Die für die messbare EDA zentralen anatomischen Strukturen der Haut sind die Schweißdrüsen (sog. exokrine Drüsen), die ihr Sekret direkt an die Hautoberfläche abgeben. Diese werden in zwei Arten unterteilt: die apokrinen und die ekkrinen Schweißdrüsen, wobei nur die letzteren für die psychophysiologische Forschung von Interesse sind. Die Hauptfunktion der ekkrinen Schweißdrüsen besteht in der Regulation der Körpertemperatur, aber auch des Wasser- und Elektrolytenhaushaltes. Obwohl sich an den Handflächen (palmar) und Fußsohlen (plantar) die meisten Schweißdrüsen befinden, ist es umstritten, ob die sich hier befindenden ekkrinen Schweißdrüsen überhaupt am thermoregulierenden Schwitzen beteiligt sind. Forschungsbefunde deuten auf eine Beteiligung psychologischer Prozesse an der palmaren und plantaren Schweißdrüsenaktivität hin (Boucsein, 2001, 2012, S. 347–390; Dawson et al., 2017).

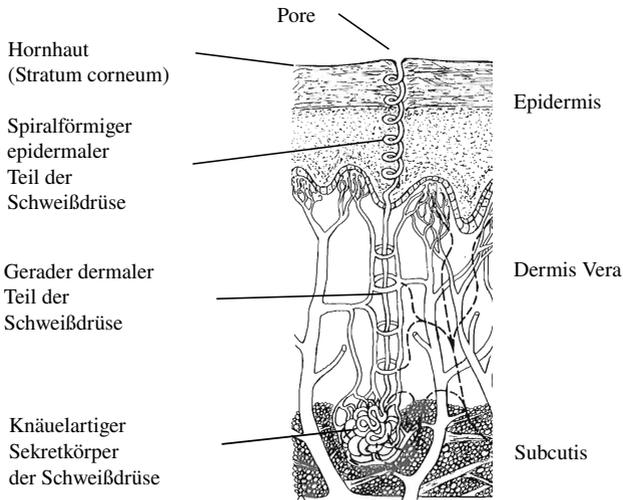


Abbildung 3.7: Schematischer Ausschnitt der Haut und der ekkrinen Schweißdrüse. Grafik modifiziert nach Boucsein (2001, S. 13).

Die anatomische Einbettung der ekkrinen Schweißdrüse in die vertikalen Schichten der Haut (Subcutis, Dermis Vera und Epidermis) ist in Abbildung 3.7 dargestellt. Die ekkrine Schweißdrüse besteht aus einem knäuelartigen Körper, welcher als der sekretorische Teil im Subcutis liegt. Ausgehend vom Sekretkörper verläuft ein gerader und anschließend ein gewundener Ausführungsgang (Ductus) an die Hautoberfläche, wo er in einer Pore endet. Die Steuerung der ekkrinen Schweißdrüse erfolgt durch das sympathische Nervensystem. Aus diesem Grund wird die EDA als ein Indikator sympathischer Aktivität betrachtet. Aber auch übergeordnete zentralnervöse Mechanismen, die an der Steuerung der EDA beteiligt sind, werden in der Literatur beschrieben (siehe Boucsein, 2001; Boucsein, 2012; Dawson et al., 2017, S. 32–42).

Eine ganze Reihe an Modellen wurde vorgeschlagen, um die elektrische Aktivität und Reaktivität der Haut zu erklären (Überblick siehe Boucsein, 2012, S. 71–84). Zur Veranschaulichung der wesentlichen Mechanismen sollen an dieser Stelle lediglich die passiven elektrischen Eigenschaften der Haut betrach-

tet werden. Aus elektrophysiologischer Sicht bilden die Subcutis, die Dermis Vera und die nicht verhornten Bereiche der Epidermis eine gut leitende Schicht (geringer elektrischer Widerstand), während die äußere Hornhaut (Stratum corneum) als schlecht leitende Schicht (hoher Widerstand) eine Art elektrische Barriere darstellt. Die Ducti der ekkrinen Schweißdrüsen bilden veränderliche Widerstände. Werden die ekkrinen Schweißdrüsen sympathisch erregt, so füllen sich die Ducti mit Schweiß. Dadurch verringert sich der elektrische Widerstand in den einzelnen Ductus-Widerständen. Eine Veränderung der Schweißdrüsenaktivität bedingt also eine Änderung des veränderlichen elektrischen Widerstands und somit auch eine messbare Änderung der EDA.

3.3.2 Messung elektrodermalen Aktivität

Zur Erfassung der EDA werden üblicherweise zwei Elektroden auf der Hautoberfläche angebracht und die elektrische Potentialdifferenz zwischen den beiden Elektroden gemessen. Dabei werden zwei Messmethoden unterschieden: (1) die exosomatische Messung, bei der eine geringe elektrische Spannung (ca. 0.5 V) an der Haut appliziert wird; und (2) die endosomatische Messung, welche die intern erzeugten Potentialunterschiede an der Hautoberfläche misst. Zur exosomatischen Messung der EDA kann entweder ein Gleichstrom (bzw. Gleichspannung) oder ein Wechselstrom (bzw. Wechselspannung) verwendet werden. Die exosomatische Messung mit Gleichstrom ist die gängigste Methode und wird auch in der vorliegenden Arbeit verwendet. Diese wird im Folgenden in Anlehnung an Boucsein (2001) und Boucsein (2012, S. 121–126) vertieft.

Das Prinzip hinter der exosomatischen Gleichstrommessung ist das Ohmsche Gesetz $R = U/I$, wobei der Hautwiderstand R gleichgesetzt ist mit dem Quotienten zwischen der an der Haut applizierten Spannung U und dem Strom I . Wird der Strom konstant gehalten, dann erfolgt eine Messung des elektrischen Hautwiderstands (Skin Resistance; SR), wohingegen beim Konstanthalten der Spannung der elektrische Leitwert gemessen wird. Oder anders beschrieben: Da U einen festen Wert besitzt (z. B. 0.5 V), verändert sich der Strom proportional

zu $1/R$. Dies ist der Kehrwert des Leitwerts G mit der Einheit Siemens (S) und $1\text{ S} = 1/\Omega$. Da der elektrische Hautwiderstand hohe Werte annimmt, wird dieser in $\text{k}\Omega$ oder $\text{M}\Omega$ und der elektrische Hautleitwert in μS ($1\ \mu\text{S} = 1/1000000\ \Omega$) angegeben.⁴

3.3.3 Quantifizierung des elektrodermalen Signals

Um ein gemessenes Hautleitwertsignal bewerten zu können, werden aus diesem Parameter extrahiert. In der Literatur findet sich eine Vielzahl an Parametern. EDA-Parameter werden dahingehend unterschieden, ob sie eine phasische oder

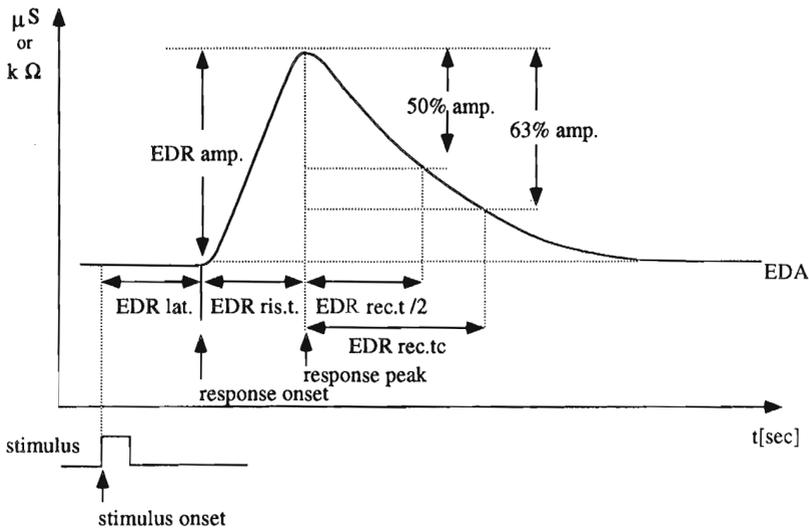


Abbildung 3.8: Exemplarischer Verlauf einer EDR und der daraus abgeleiteten Parameter. Grafik aus Boucsein (2012, S. 154).

⁴ Der elektrische Leitwert ist nicht gleich der elektrischen Leitfähigkeit. Allerdings ist in der deutschen Literatur der Begriff Hautleitfähigkeit gebräuchlich, obwohl der Hautleitwert gemeint ist. Auch in der vorliegenden Arbeit wird der Begriff Hautleitfähigkeit synonym zum Hautleitwert verwendet.

tonische Veränderung der EDA abbilden. Während tonische Parameter sich auf langsamere Veränderungen über die Zeit beziehen, repräsentieren phasische EDA-Parameter kurz andauernde Einzelereignisse, die auf dem tonischen Verlauf aufsetzen (Boucsein, 2012, S. 150).

Phasische Veränderungen werden als elektrodermale Reaktionen (Electrodermal Reaction; EDR) bezeichnet. Sowohl bekannte externe Auslösereize als auch nicht explizit erkennbare interne Reize können eine EDR auslösen. Folglich wird eine Trennung in eine spezifische (explizite Auslösereize) und nicht-spezifische (implizite Auslösereize) EDR vorgenommen. Letztere werden mit NS-EDR abgekürzt und als tonische Veränderungen parametrisiert. Unabhängig vom expliziten oder impliziten Auslösereiz zeigen alle EDRs einen typischen Signalverlauf (siehe Abbildung 3.8). Eine EDR dauert im Durchschnitt zwischen 4.5 und 6.5 s. Der Anstieg bis zum Reaktionsmaximum dauert zwischen 0.5 und 5 s (EDR Rise Time; EDR.rise.t)⁵ und der Abstieg zwischen 4 und 10 s (EDR Recovery Time; EDR.rec.t). Die Differenz des Reaktionsmaximums zum elektrodermalen Level (EDL) – dem Fußpunkt der EDA – ergibt die Amplitude der EDR (EDR.amp). Diese wird als ein Ausschlusskriterium für das Auftreten einer EDR genutzt. Reaktionsmaxima von 0.01 bis 0.05 μS werden dabei als bedeutsam erachtet (Boucsein, 2012, S. 157). Häufig kann beim Auslaufen einer EDR das Ausgangsniveau nicht mehr erreicht werden, weshalb die EDR.rec.t als Halbwertszeit – EDR.rec.t/2 (50 %) – oder als Zeitkonstante – EDR.rec.t/c (63 %) – ermittelt wird. Im Falle eines expliziten Auslösereizes kann die Latenzzeit (EDR.lat) vom Reiz- bis zum Reaktionsbeginn der EDR (EDR Onset; EDR.ons) bestimmt werden. Als weiterer Parameter wird die Fläche unter der EDR (EDR.area) betrachtet. Ein wesentliches Problem bei der Ermittlung von EDR-Parametern ist die Überlagerung elektrodermalen Reaktionen (Vertiefung siehe Boucsein, 2012, S. 152–155). Die Trennung tonischer

⁵ Bei der Notation der EDA-Parameter wird in Anlehnung an Boucsein (2012) ein Punkt zur Abgrenzung der Signalkomponenten verwendet. Beispiel: EDR Rise Time wird mit EDR.rise.t abgekürzt.

und phasischer Überlagerungen kann heutzutage mit Hilfe von mathematischen Modelle bewerkstelligt werden (siehe Unterkapitel 5.2.2).

Ein Parameter tonischer Veränderungen ist das elektrodermale Level (EDL). Das EDL wird üblicherweise als Mittelwert des EDA-Niveauwerts über ein bestimmtes Zeitintervall bestimmt. Zusätzlich zum mittleren EDL kann auch dessen Differenz zwischen verschiedenen Zeitintervallen (EDA.D) bestimmt werden. Weitere Parameter tonischer Veränderungen werden aus den phasischen Parametern berechnet. Ein typisches Beispiel hierfür ist die Anzahl unspezifischer Fluktuationen pro Minute (EDR.N) (Boucsein, 2012, S. 174–175). Zudem werden tonische Parameter als Mittelwert der oben beschriebenen, phasischen Parametern ermittelt (siehe Boucsein, 2001). Somit ist jeder phasische Parameter auch als tonischer Parameter quantifizierbar: z. B. Mittelwert aller nicht-spezifischen EDR-Amplituden in einem Zeitintervall (EDR.amp). Eine tabellarische Übersicht der in dieser Arbeit verwendeten Parameter findet sich in Unterkapitel 5.2.2.

3.3.4 Befundlage zur elektrodermalen Aktivität

Die Durchführung, aber auch bereits die Antizipation nahezu jeder Aufgabe führt zumindest initial zu einer Steigerung der tonischen EDA im Vergleich zum Ruhezustand. In zahlreichen Studien gelang der Nachweis eines signifikanten Anstiegs tonischer EDA-Parameter (EDL, EDR.N, EDR.amp und EDA.area) mit zunehmender mentaler Beanspruchung und Anstrengung für unterschiedliche Labor- und Feldaufgaben (z. B. Chanel et al., 2011; Collet, Salvia & Petit-Boulanger, 2014; Engström, Johansson & Östlund, 2005; Kohlisch & Schaefer, 1996; Mehler, Reimer & Coughlin, 2012; Moya-Albiol et al., 2001; Nourbakhsh, Chen, Wang & Calvo, 2017; Setz et al., 2010; Visnovcova, Mestanik, Gala, Mestanikova & Tonhajzerova, 2016; Widyanti, Hanna, Muslim & Satalaksana, 2017; Überblick siehe Ribback, 2003). Die EDA ist zwar sensitiv für den Unterschied zwischen Ruhe- und Aufgabenbedingung, allerdings kann nur schwer zwischen den feineren Stufen mentaler Beanspruchung unterschieden werden

(z. B. Widyanti et al., 2017). Eine feinere Differenzierung unterschiedlicher Schwierigkeitsstufen einer Arbeitsgedächtnisaufgabe gelang bspw. Mehler et al. (2012). Es gibt aber auch Untersuchungen, die keine signifikanten Unterschiede zwischen einer experimentellen Ruhe- und Aufgabenbedingung in denselben EDA-Parametern finden (z. B. Renaud & Blondin, 1997). Wenngleich für die Parameter EDR.rec.t wenige Belege vorliegen, so zeigen diese einen bedeutsamen Zusammenhang mit der mentalen Beanspruchung. Berichtet wird ein Abfall der EDR.rec.t im Vergleich zur Ruhebedingung bei unterschiedlichen mentalen Aufgaben (Edelberg, 1972).

Die oben aufgeführten Studien belegen zwar, dass ein Zusammenhang elektrodermalen Parameter mit der mentalen Beanspruchung zu erwarten ist, jedoch gibt es weitere psychologische Konstrukte, welche der EDA den Status als Indikator mentaler Beanspruchung streitig machen (Überblick siehe Boucsein, 2012, Kapitel 3). Besonders hervorzuheben ist an dieser Stelle die Emotionsforschung, welche eine Fülle an Belegen für den Zusammenhang der EDA mit unterschiedlichen Emotionen vorzeigt (zsf. Boucsein, 2012, S. 370–381). In diesem Bereich wird sowohl ein Zusammenhang der EDA-Parameter, wie z. B. EDR.amp oder EDR.N, mit steigender Emotionsintensität (z. B. Traxel, 1960; Nourbakhsh et al., 2017) als auch ein systematisches Zusammenhangsmuster für unterschiedliche Emotionen und Emotionsregulationsmechanismen belegt (z. B. Stemmler, 1989; Silvestrini & Gendolla, 2007). Folglich ist die Diagnostizität vieler EDA-Parameter als eher gering einzustufen. Eine höhere Diagnostizität wird für die Amplitude sowie die Rückbildungszeit der EDRs vermutet (Kramer, 1991; Manzey, 1998).

Inwieweit die unterschiedlichen EDA-Parameter redundante oder unabhängige Informationsquellen darstellen, ist bis heute nicht vollständig geklärt. Dawson et al. (2017) vermuten anhand des Ausmaßes und der Größe der Interkorrelation elektrodermalen Parameter, dass unterschiedliche EDA-Parameter unabhängige bzw. teilweise unabhängige Informationsquellen darstellen. Venables und Christie (1980, zitiert nach Boucsein, 2012, S. 220) gehen von einer differentiellen Validität des EDL und der EDR.N aus. Gleichzeitig schreiben sie den Parame-

tern EDR.rise.t und EDR.rec.t eine Redundanz zu (Venables & Christie, 1980, zitiert nach Boucsein, 2012, S. 228). Im Gegensatz dazu ordnet Edelberg (1972) den Parametern EDR.rise.t und EDR.rec.t zwei voneinander unabhängige Kontrollsysteme zu: ein Kontrollsystem für die Schweißreaktion (repräsentiert durch EDR.rise.t) und ein Kontrollsystem für die Schweißabsorption (repräsentiert durch EDR.rec.t).

Zusammenfassend ist eine valide Zuordnung der tonischen EDA-Parameter zu einem spezifischen psychologischen Konstrukt bisher nicht ohne weiteres möglich. Folglich ist insgesamt von einer geringen Diagnostizität und Validität der EDA-Parameter in Bezug auf die mentale Beanspruchung auszugehen. Nichtsdestoweniger zeigen die EDA-Parameter eine zufriedenstellende Sensitivität für unterschiedliche mentale Beanspruchungszustände. Zudem bringt die EDA den Vorteil mit sich, dass sie alleinig durch die sympathische Aktivität moduliert wird. Im Gegensatz zu anderen autonomen Signalen (z. B. Herzrate) besitzt die EDA keine Ambiguität in der Zuordnung zum sympathischen oder parasymphathischen System.

3.4 Multivariate Betrachtung psychophysiologischer Indikatoren

Frühe Arbeiten in der psychophysiologischen Forschung konzentrierten sich auf einzelne Parameter. Diese wurden auf ihre Güte zur Vorhersage psychischer Zustände, wie der mentalen Beanspruchung, untersucht. Aber auch heute noch finden sich zahlreiche Forschungsarbeiten, die sich auf einzelne Parameter fokussieren. Die Güte solcher Einzelparameter im Kontext mentaler Beanspruchungsforschung wurde in den Kapiteln 3.2.4 und 3.3.4 zusammengefasst. Ersichtlich ist, dass nicht alle Studien ein erwartungskonformes Ergebnis finden. Aus diesem Grund sind physiologische Parameter bis heute zwar vielversprechende, jedoch nicht gänzlich validierte Kenngrößen für die mentale Beanspruchung. Fahrenberg (2001) stellt die Validität und Generalisierung

physiologischer Parameter als Indikatoren mentaler Beanspruchung gänzlich in Frage. Die theoretische Begründung der Indikation ist hier ein besonderes Problem, da sich bspw. allein für die Herzschlagfrequenz eine Vielzahl an Bedeutungen und Interpretationen in der Forschungsliteratur finden. Diese reichen von psychischer oder physischer Beanspruchung, Emotionen, allgemeiner Aktiviertheit, metabolischer und thermoregulatorischer Muskelaktivität, Stimulusintensität bis hin zur funktionellen Störung und körperlicher Erkrankung (Fahrenberg, 2001). Die mangelnde Validität ist u. a. darin zu vermuten, dass die physiologischen Kenngrößen in der mentalen Beanspruchungsforschung nur unzureichend psychometrisch geprüft werden (siehe Matthews, Reinerman-Jones, Barber & Abich, 2015). Eine Ursache hierfür ist, dass die Validitätsbestimmung für psychophysiologische Parameter alles andere als trivial ist (Tsang & Wilson, 1997). Dies ist vermutlich auch ein Grund, warum bis heute lediglich eine schwache Validitätsform, die Kriteriumsvalidität (siehe A.2.4), als Rechtfertigung für die Verwendung physiologischer Maße herangezogen wird. Die in Kapitel 3.2.4 und 3.3.4 aufgeführten Studienergebnisse zur kardialen und elektrodermalen Aktivität stellen den Versuch zur Prüfung der Kriteriumsvalidität in der Forschungsliteratur dar. Das Vorliegen einer Kriteriumsvalidität ist allerdings zum Beweis der Validität nicht ausreichend. Eine höherwertigere Validität, wie die Konstruktvalidität (siehe Anhang A.2.4), wurde bis heute für keinen der Parameter hinreichend belegt. Vorliegende Befunde einer konsistenten und hohen Korrelation bedeuten nicht, dass ein physiologischer Parameter und die mentale Beanspruchung nur ausschließlich miteinander zusammenhängen. Eine solche fälschliche Annahme der Isomorphie (Eins-zu-eins-Zusammenhang) ist nach Cacioppo und Tassinari (1990) nicht nur gängige Praxis, sondern führt auch zu fehlerhaften Schlussfolgerungen. Darüber hinaus ist nicht geklärt, welche Parameter komplementäre und welche redundante Kenngrößen derselben physiologischen Aktivität darstellen. Dieser Sachverhalt spiegelt sich in den hohen Korrelationen physiologischer Parameter untereinander wider und führt unweigerlich zu der Frage nach der inkrementellen Validität einzelner Parameter (Fahrenberg, 2001).

Die Erkenntnis, dass einzelne physiologische Indikatoren oder einfache Schwellenwerte nicht ausreichen, um Rückschlüsse auf psychologische Zustände, wie die mentale Beanspruchung, zu ziehen, ist heute in der psychophysiologischen Forschung angekommen. Gefordert sind physiologische Aktivitätsprofile, die nur aus multivariaten Analysen gewonnen werden können. Trotz der immer lauter werdenden Forderung nach multivariaten Untersuchungen in der klassischen Psychophysiologie blieb der Anteil der Analysemethoden, die eine simultane Betrachtung mehrerer physiologischer Parameter erlauben, im Zeitraum von 2000 bis 2014 unter 10 % (Jennings & Allen, 2017). Im Gegensatz zur klassischen Psychophysiologie zeigte die Informationstechnik ab der Jahrtausendwende ein reges Interesse an der multivariaten Untersuchung psychophysiologischer Zusammenhänge. Über die Jahre entwickelten sich hier an die Psychophysiologie angelehnte Teildisziplinen, die unter den Namen Psychophysiological Computing, Affective Computing oder auch Cognitive Computing bekannt sind. Auf diese wird im Folgenden eingegangen.

3.4.1 Psychophysiological Computing

Spätestens seit den Arbeiten von Rosalind Picard (z. B. Picard, Vyzas & Healey, 2001) hat die Psychophysiologie auch Einzug in die informationstechnische Forschungsdisziplin erhalten. Das Ziel, psychologische Zustände, wie die mentale Beanspruchung, mit Hilfe physiologischer Reaktionen abzubilden, steht auch hier im Vordergrund. Und auch wenn sich die verfügbaren, statistischen Analysemethoden der klassischen psychophysiologischen Forschung und des Psychophysiological Computing überschneiden, ist das Vorgehen und die Wahl der Methoden meist eine andere. Die überwiegende Anzahl an Forschungsarbeiten im Bereich des Psychophysiological Computing nutzt komplexe Klassifikationsverfahren, u. a. Support Vector Machine (SVM), Diskriminanzanalyse oder Entscheidungsbäume, um eine kategoriale Zuordnung zu psychischen Zuständen vorzunehmen (Überblick siehe Novak et al., 2012). Im Gegensatz dazu kommen Schätzverfahren, wie lineare Regression, Fuzzy

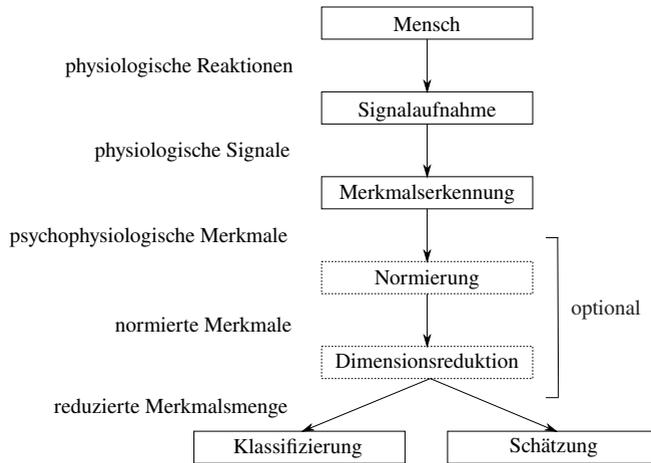


Abbildung 3.9: Typische Analyseschritte des Psychophysiological Computings. Grafik modifiziert nach Novak et al. (2012, S. 155).

Logik oder Varianzanalyse, seltener vor (Novak et al., 2012). Novak et al. (2012) sieht dies nicht zuletzt darin begründet, dass die Güte einer Klassifikation anhand der Treffsicherheit eindeutig bestimmt werden kann, während es bei Schätzverfahren mehrere Gütemaße, z. B. Standardfehler oder Varianz, ohne festgelegten Standard gibt (Novak et al., 2012). Die typischen Analyseschritte des Psychophysiological Computings sind in Abbildung 3.9 aus Novak et al. (2012) entnommen und werden im Folgenden zusammenfassend vorgestellt. Die ersten Schritte bestehen in der Erfassung und Parametrisierung des physiologischen Signals. Im Anschluss kann optional eine Normalisierung und/oder eine Dimensionsreduktion durchgeführt werden. Ersteres dient der Reduktion intra- oder interindividueller Varianz. Für die Dimensionsreduktion stehen unterschiedliche Verfahren, z. B. die Feature Subset Selection oder die Hauptkomponentenanalyse (Principal Component Analysis; PCA), zur Verfügung. Im Anschluss werden die physiologischen Parametern bzw. die reduzierten Merkmalsdaten mit Hilfe von Klassifikations- oder Schätzverfahren analysiert. Zum Schluss werden die Ergebnisvariablen bestimmten Konstruktausprägungen zugeordnet (z. B. hohe und niedrige mentale Beanspruchung). Das Ziel

besteht darin, eine maximale Treffsicherheit zu erreichen. Eine Validierung der Klassifikationsergebnisse erfolgt meist nur in Hinblick auf die Robustheit des Analyseverfahrens. Zu diesem Zweck werden die Gesamtdaten in mindestens zwei Datensätze unterteilt, die dann als Trainings- und Testdatensätze fungieren. Diese sog. Kreuzvalidierung gibt nur Aufschluss über die Anpassungsgüte der geschätzten Werte an die gemessenen Werte, jedoch nicht über die Richtigkeit der Interpretation. Oder anders gesagt, es überprüft nicht, ob die einer Ergebnisvariable zugeordnete Konstruktausprägung (z. B. hohe mentale Beanspruchung) zutreffend ist und somit auch nicht die psychometrische Validität. Das Psychophysiological Computing hat aufgrund seines reichen Methodenspektrums und teilweise vielversprechender Resultate in den letzten Jahren an Popularität gewonnen und zu einem regelrechten Boom an Forschungsarbeiten auf diesem Gebiet beigetragen (Überblick siehe Novak et al., 2012). Die Vorteile dieses Vorgehens liegen klar auf der Hand: Mehrere physiologische Parameter können gleichzeitig in die Analyse aufgenommen und zur Vorhersage der mentalen Beanspruchung genutzt werden. Eine solche multivariate Betrachtung und Vorhersage mentaler Beanspruchung ist der Analyse einzelner physiologischer Parameter nicht nur in Hinblick auf die Vorhersagegüte überlegen, sondern bietet zudem die Möglichkeit, komplexe Parameterinteraktionen in der Vorhersage zu berücksichtigen.

Auch das Konstrukt der mentalen Beanspruchung wurde bereits mehrfach mit unterschiedlichen Klassifikationsverfahren untersucht. Im Gegensatz dazu finden sich im Psychophysiological Computing nur vereinzelt Untersuchungen, die ein Schätzverfahren, wie die multivariate Varianzanalyse (z. B. Fairclough & Venables, 2006), verwenden. Der Vorteil der Schätzverfahren ist, dass Zusammenhänge mit Fragebögen oder Leistungsmaßen der mentalen Beanspruchung quantifiziert werden können und somit keine künstliche Einteilung in hoch oder niedrige Beanspruchungskategorien notwendig ist.

Ein Beispiel für Klassifikationsverfahren sind die Arbeiten von Wilson und Russell (2003a, 2003b, 2007). Mit Hilfe von Diskriminanzanalysen und KNN nehmen sie eine Klassifikation in unterschiedliche Beanspruchungskategorien

(gering, mittel, hoch bzw. Überbeanspruchung, Nicht-Überbeanspruchung) vor. Klassifiziert werden dabei 37 bis 88 physiologische Parameter. Die Klassifikation wird von Wilson und Russel für jede Person individuell berechnet. Zudem nehmen sie in zwei Studien (Wilson & Russell, 2003a, 2003b) für jede Versuchsperson die individuelle Reduktion der physiologischen Prädiktoren solange vor, bis eine maximale individuelle Treffsicherheit erreicht ist. Hierbei wird die Menge physiologischer Parameter um bis zu 90 % reduziert. Berichtet werden (über Personen gemittelte) Trefferquoten zwischen 55.9 % und 99 %. Für Einzelpersonen fallen die Trefferquoten vereinzelt auch gering aus (z. B. 28.7 % bei Wilson & Russell, 2003b). Nur in einer einzigen Studie geben Wilson und Russell (2003b) eine Übersicht zur Art sowie Relevanz der einzelnen physiologischen Parameter an. Aus dieser lassen sich zwei Sachverhalte erkennen: Erstens, für unterschiedliche Versuchspersonen sind unterschiedliche Vorhersageparameter relevant. Zweitens, einige Parameter (z. B. Herzperiode) besitzen für fast alle Versuchspersonen eine hohe Vorhersagekraft und andere Parameter (z. B. Lidschlaglatenz) nur für einzelne Versuchspersonen.

Individuelle Vorhersagemodelle werden auch von S. Chen und Epps (2013) mit dem Gaussian Mixture Model Klassifikator (für okulomotorische Parameter) oder von Haapalainen, Kim, Forlizzi und Dey (2010) mit dem Naive Bayes-Klassifikator (für Parameter kardialer, elektrodermalen, okulomotorischer und hirnelektrischer Signale) berechnet. Auch in diesen Studien wird für jede Person eine eigene Parameterselektion von zwei bis sieben Parametern vorgenommen. Dabei wird durchschnittlich eine Treffsicherheit von 71.1 % bzw. 83.2 % zur Unterscheidung von zwei Beanspruchungskategorien erreicht. Für einzelne Personen liegt diese allerdings nicht über der Ratewahrscheinlichkeit.

Die Trefferquoten fallen schlechter aus, wenn man personenunabhängige Klassifikationen vornimmt. Bspw. kommt Koenig et al. (2011) in seiner Diskriminanzanalyse mit elektrodermalen, kardialen, respiratorischen sowie temperaturbezogenen Parametern lediglich auf eine Treffsicherheit von 34 % bis 70 % bei der Unterscheidung von vier Schwierigkeitsstufen einer mentalen Aufgabe. Gleichermaßen berichten Chanel et al. (2011) eine Treffsicherheit von nur

46 % bis 54 % mit 22 Parametern der elektrodermalen, kardialen, respiratorischen Aktivität und der Temperatur. Eine Verbesserung der Vorhersage auf 59 % erreichten Chanel et al. (2011) durch eine gezielte Selektion relevanter Vorhersagevariablen. Die verschiedenen Selektionsmethoden ergeben unterschiedlich große Variablengruppen (5 bis 12 Parameter) mit nur drei sich überschneidenden Parametern. Die Ursachen dieser unterschiedlichen Ergebnismuster werden nicht hinterfragt und lediglich das Ergebnis mit der höchsten Treffsicherheit als Repräsentant gewählt.

Ob die berichteten Ergebnisvariablen der verschiedenen Klassifikationsverfahren tatsächlich nur die mentale Beanspruchung und kein anderes Konstrukt (z. B. Emotionen oder aufgabenspezifische Anforderungen) abbilden, muss in zukünftigen Analysen verifiziert werden (Nachweis der Konstruktvalidität). Anzumerken ist, dass mit ähnlichen Analyseverfahren und physiologischen Daten Trefferquoten in ähnlicher Höhe für inhaltlich verwandte Konstrukte erreicht werden. Beispiele hierfür sind die Konstrukte Erregung und Valenz (Chanel, Krongeg, Grandjean & Pun, 2006 mit 30 % bis 60 %; Baltaci & Gokcay, 2016 mit 65.5% bis 88.3 %), Stress (z. B. Healey & Picard, 2005 mit 97.4 %) und spezifische Emotionskategorien, wie Langeweile, Wut, Trauer, Frustration (z. B. Kapoor, Burleson & Picard, 2007; K. H. Kim, Bang & Kim, 2004; Shen, Wang & Shen, 2009 mit 58.3 % bis 86.3 %). Eine Fülle weiterer Beispiele findet sich zusammengefasst bei Afzal und Robinson (2011), Xu, Nwe und Guan (2015) sowie Novak et al. (2012).

3.4.2 Schwächen des Psychophysiological Computing

Die Schwächen des Psychophysiological Computing liegen in der mangelnden theoretischen Fundierung und Validierung der Forschungsergebnisse. Alle oben aufgeführten Arbeiten haben gemein, dass die Validität (bzw. Diagnostizität) im Sinne der Psychometrie nicht hinreichend gesichert ist und höchstens von einer Kriteriumsvalidität auszugehen ist. Oft besteht das Forschungsziel lediglich darin, eine maximale Treffsicherheit zu erzielen, ohne die Hintergründe und

Klassifikationsmuster physiologischer Daten zu hinterfragen. Zweitrangig ist die Frage, warum die Treffsicherheit erzielt wurde, also welche latente Struktur für das Zustandekommen der Ergebnisse verantwortlich ist und ob diese latente Struktur im Sinne der Konstruktvalidität das Konstrukt auch inhaltlich widerspiegelt. Nicht zuletzt ist die mangelhafte theoretische Validierung darin begründet, dass psychometrische Kriterien, wie Reliabilität und Validität oder gar die anwendungsorientierten Gütekriterien, wie Sensitivität und Diagnostizität, noch keinen Einzug in die Teildisziplin des Psychophysiological Computing erhalten haben. Im Vordergrund stehen informationstechnische Kriterien, wie Robustheit, rechnerischer Zeitaufwand, Transparenz oder Implementierungsaufwand des Klassifikationsverfahrens (siehe Novak et al., 2012). Dass diese Kriterien nicht ausreichen, um mit Hilfe physiologischer Daten psychologische Zustände vorherzusagen, zeigt die Tatsache, dass trotz einer hohen Literaturdichte es bis heute kein allgemein akzeptiertes Verfahren zur Klassifikation bzw. Schätzung psychischer Zustände gibt (Novak et al., 2012).

Einen ersten Schritt in Richtung Abgrenzung mentaler Beanspruchung von verwandten Konstrukten (psychosozialer Stress) unternehmen Setz et al. (2010). Hierfür untersuchen sie 16 elektrodermale Parameter mit sechs unterschiedlichen Klassifikationsverfahren (u. a. Diskriminanzanalyse und SVM). Wie bereits bei den oben beschriebenen Studien wird auch hier eine Parameterselektion vorgenommen. Die Parameterkombination mit der höchsten Treffsicherheit wird als Repräsentant für das Klassifikationsverfahren ausgewählt. Interessant ist, dass trotz gänzlich verschiedener Parameterkombinationen alle Klassifikationsverfahren mit einer Treffsicherheit von ca. 80 % zwischen mentaler Beanspruchung und psychosozialen Stress differenzieren können. Dieses Ergebnis wirft die Frage auf, warum bei den verschiedenen Klassifikationsverfahren unterschiedliche Parameterkombinationen entstehen. Zu erwarten wäre, dass ähnlich gute Klassifikationsverfahren dieselben physiologischen Parameter zur Vorhersage der mentalen Beanspruchung verwenden. Da solche Fragestellungen allerdings im Wettlauf um die höchste Treffsicherheit untergehen, bleibt eine Diskussion und Analyse solcher Fragen aus. Ist es nun ein einzelner Parameter

oder die Interaktion mehrerer unabhängiger Parameter, welche das latente Konstrukt der mentalen Beanspruchung abbildet? Oder sind es individuell spezifische Parameterkombinationen, die je nach Personengruppe oder Individuum spezifisch zusammengesetzt sind? Solche stabile, interindividuelle Unterschiede postuliert bspw. M. Schneider (2017). Interindividuelle Unterschiede sind hier nicht eine störende oder gar zu vernachlässigende Varianzquelle, sondern zu modulierende Faktoren in der Vorhersage mentaler Beanspruchung. Mit der Berücksichtigung interindividueller Unterschiede bei der Klassifikation psychophysiologischer Indikatoren wird eine Brücke zur Persönlichkeitsforschung im Sinne der Erforschung stabiler, psychophysiologischer Dispositionen geschaffen. Und obwohl einige Studien die Klassifikation für jede Person individuell berechnen (z. B. Wilson & Russell, 2003a, 2003b), wird das Vorgehen weder theoretisch begründet noch das Ergebnis hinterfragt. Insgesamt lassen die Ergebnisse des Psychophysiological Computing trotz vielversprechender Resultate noch eine Fülle an Fragen unbeantwortet.

3.5 Zusammenfassung

Die Messung mentaler Beanspruchung mit Hilfe physiologischer Parameter ist seit Jahrzehnten gängige Forschungspraxis. Und so finden sich in der Literatur eine Fülle an Studien, die den Zusammenhang mentaler Beanspruchung mit kardialen und elektrodermalen Parametern belegen. Trotz der Menge an Ergebnissen konnte bisher kein physiologischer Einzelparameter als robuster und valider Indikator mentaler Beanspruchung identifiziert werden. Zum einen, weil die autonome Steuerung kardialer und elektrodermalen Aktivität komplexer ist, als man es ursprünglich angenommen hatte. Zum anderen, weil bisher lediglich eine schwache Validitätsform, die Kriteriumsvalidität, als Rechtfertigung für die Verwendung physiologischer Parameter herangezogen wird. So werden dieselben physiologischen Parameter in der Literatur als Indikatoren unterschiedlicher Konstrukte (z. B. physische Beanspruchung, Emotionen, Aktiviertheit) herangezogen. Zudem gibt es keine Übersicht und Klarheit darüber, welche

physiologischen Parameter komplementär und welche redundant bei der Vorhersage mentaler Beanspruchung sind. Um Klarheit zu verschaffen, wurde die Forderung nach multivariaten Analyseverfahren immer lauter. Bis heute werden multivariate Verfahren in der klassischen psychophysiologischen Forschung nur vereinzelt eingesetzt. Dafür aber umso mehr in der informationstechnologischen Teildisziplin des Psychophysiological Computing. In diesem Forschungsfeld werden vor allem Klassifikationsverfahren angewandt. Das Ziel der meisten Forschungsarbeiten auf diesem Gebiet ist es, eine maximale Treffsicherheit zu erreichen. Diese reicht allerdings nicht aus, um von einer gemessenen physiologischen Aktivität auf die zugrundeliegende mentale Beanspruchung zu schließen. Denn sind es in der klassischen psychophysiologischen Forschung einzelne physiologische Parameter, die nachweislich unterschiedlichen Einflussfaktoren unterworfen sind, so sind es beim Psychophysiological Computing die Ergebnisvariablen, die auf Konstruktvalidität untersucht werden müssen. Und auch wenn der Nachweis der Kriteriumsvalidität durchaus gängig ist, so wird nur selten eine statistische Abgrenzung zu konkurrierenden Konstrukten berichtet. Auch scheint es bislang nicht von Interesse zu sein, aus welchem Grund unterschiedliche Untersuchungen mit gleichen, aber auch gänzlich verschiedenen physiologischen Parametern eine ähnlich hohe Treffsicherheit erreichen. Ist es nun ein einzelner Parameter, die Interaktion mehrerer unabhängiger Parameter oder gar eine individuell spezifische Parameterkombination, welche das latente Konstrukt der mentalen Beanspruchung abbilden?

Aufgrund der schwachen Validität bzw. Diagnostizität sind die Rückschlüsse, die aus den Forschungsbefunden gezogen werden können, nicht hinreichend, um valide Rückschlüsse von einer gemessenen physiologischen Aktivität auf die zugrundeliegende mentale Beanspruchung zu ziehen. Dies ist allerdings eine erforderliche Voraussetzung für die Verwendung psychophysiologischer Parameter in beanspruchungssensitiven Mensch-Maschine-Systemen. Wie dieses Ziel im Rahmen der psychophysiologischen Forschung erreicht werden kann, wird in den folgenden Kapiteln diskutiert und untersucht.

4 Herleitung der Forschungsfragen

Vor dem Hintergrund einer beanspruchungssensitiven Mensch-Maschine-Schnittstelle sind Forscher an der Identifikation valider physiologischer Indikatoren interessiert. Diese Indikatoren sollen Auskunft über den aktuellen mentalen Beanspruchungszustand der Person geben. Viel zu selten rückt die Frage in den Fokus, welche psychometrischen Voraussetzungen überhaupt erfüllt sein müssen, damit ein potentieller Indikator als valider Repräsentant der mentalen Beanspruchung gilt. Diese Frage wird in Kapitel 4.1 mit Hilfe des theoretischen Rahmenmodells von Cacioppo et al. (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017) beantwortet. Im Rahmenmodell werden physiologische Profile der psychologischen Konstrukte, wie der mentalen Beanspruchung, gefordert. Wie die Erstellung von physiologischen Profilen methodisch gelingen kann, wird in Kapitel 4.2 dargestellt. Anschließend wird in Kapitel 4.3 das Forschungsziel und -programm der vorliegenden Arbeit vorgestellt.

4.1 Rückschlüsse von der physiologischen Aktivität auf die mentale Beanspruchung

Um von einer gemessenen physiologischen Reaktion auf die dahinter liegende mentale Beanspruchung zu schließen, ist aus psychometrischer Perspektive nicht trivial. Es erfordert fundierte Kenntnisse über physiologische Korrelate und Störeinflüsse der mentalen Beanspruchung. Unterschiedliche theoretische Modelle und eine Vielzahl an potentiellen physiologischen Indikatoren mentaler Beanspruchung stellen Forscher vor die „Qual der Wahl“. Gleichzeitig existieren

lediglich Annahmen zur Indikatorfunktion der meisten physiologischen Einzelparameter. Obwohl zahlreiche Studien den Zusammenhang physiologischer Reaktionen und psychologischer Konstrukte (sei es mentale Beanspruchung, Emotionen oder Stress) untersuchen, gibt es kaum Forschungsarbeiten, die einen direkten Beitrag zur Validierung physiologischer Parameter als Indikatoren mentaler Beanspruchung leisten. Nach wie vor werden in der klassischen psychophysiologischen Forschung einzelne physiologische Parameter in den Vordergrund gerückt, ohne deren Güte oder Interpretation als Indikator mentaler Beanspruchung hinreichend zu hinterfragen. Methodisch wird zumeist der psychologische Kontext manipuliert und das physiologische Korrelat gemessen. Ein gängiges Beispiel aus der mentalen Beanspruchungsforschung ist die Einteilung in Kontroll- und Experimentalbedingungen durch die Manipulation der Aufgabenschwierigkeit. Unterschiede in der physiologischen Variable (z. B. Herzschlagfrequenz) werden dann auf eine Veränderung der mentalen Beanspruchung zurückgeführt. Die Ergebnisse solcher Untersuchungen reichen nach Cacioppo et al. (2007) allerdings nicht aus, um die physiologische Reaktion als validen Indikator eines psychischen Zustands, wie der mentalen Beanspruchung, zu interpretieren:

However, one cannot logically conclude that a processing stage or state has definitely been detected simply because a physiological response found previously to vary as a function of a psychological processing stage or state has been observed. (The logical flaw in this form of inference is termed affirmation of the consequent.)
(S. 8)

Um den Nachweis zu erbringen, dass eine beobachtbare Veränderung in einem physiologischen Parameter einen Indikator der mentalen Beanspruchung darstellt, müssen nach Cacioppo et al. (2007) nicht nur die Alternativerklärungen für den Zusammenhang ausgeschlossen (Nachweis der Spezifität), sondern auch die kausale Wirkung des physiologischen Parameters auf die mentale Beanspruchung bewiesen werden. Letzteres bedeutet, dass im Gegensatz zur gängigen

Forschungspraxis die physiologische Aktivität experimentell manipuliert und die psychologischen Reaktionen gemessen werden sollte (Cacioppo et al., 2007). Daraus leiten sich unmittelbar die Fragen ab, welche Schlussfolgerungen und Interpretationen die bisherigen psychophysiologischen Forschungsergebnisse überhaupt erlauben und wie eine valide Indikation der mentalen Beanspruchung zukünftig gelingen kann. Oder wie Berntson und Cacioppo (2002, S. 11) formulieren: „More generally, the question arises as to how we are to understand psychophysiological relations.“ Um diese Frage zu beantworten, entwickelt das Team um John T. Cacioppo ein Rahmenmodell (siehe Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017), welches die Voraussetzungen und Grenzen möglicher Rückschlüsse und Interpretationen aus psychophysiologischen Zusammenhängen beschreibt. Dieses Rahmenmodell und seine Implikationen werden in den folgenden Kapiteln 4.1.1, 4.1.2 und 4.1.3 näher erläutert.

4.1.1 Rahmenmodell psychophysiologischer Zusammenhänge

Zur Veranschaulichung der unterschiedlichen Zusammenhänge zwischen psychologischen Konstrukten und physiologischen Reaktionen beschreiben Cacioppo et al. (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017) psychologische Konstrukte als Elemente einer Teilmenge Ψ und physiologische Reaktionen als Elemente einer Teilmenge Φ . Ausgehend von der Annahme, dass sich psychologische Prozesse in körperlichen Reaktionen manifestieren, wird angenommen, dass alle psychologischen Element in Ψ mehr oder weniger physiologische Bezugselemente in Φ haben. In Abbildung 4.1 werden diese Beziehungen als Verbindungen zwischen den Elementen der Teilmengen dargestellt. Aufbauend auf diesen Annahmen leiten Cacioppo et al. fünf allgemeine Zusammenhangsmuster zwischen den Elementen von Ψ und Φ ab (siehe Abbildung 4.1):

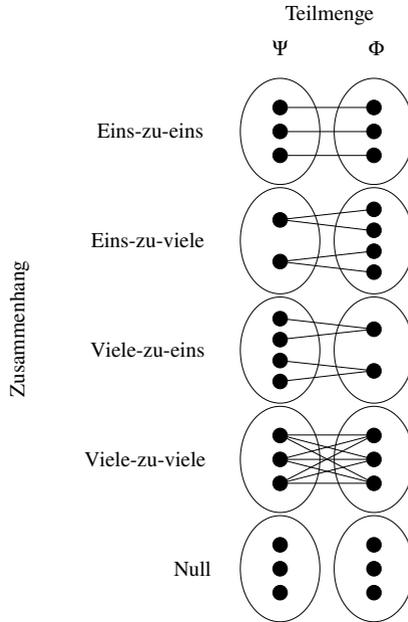


Abbildung 4.1: Gemäß dem Rahmenmodell von Cacioppo et al. (siehe Cacioppo & Tassinary, 1990; Cacioppo et al., 2007) ergeben sich fünf allgemeine Zusammenhangsmuster zwischen den psychologischen Elementen in der Teilmenge Ψ und den physiologischen Elementen in der Teilmenge Φ . Grafik modifiziert nach Berntson und Cacioppo (2002, S. 12).

1. Eins-zu-eins-Zusammenhang (Isomorphie): Die Assoziation eines Elements aus Ψ mit nur einem Element aus Φ (und umgekehrt),
2. Eins-zu-viele-Zusammenhang: Die Assoziation eines psychologischen Elements mit vielen physiologischen Elementen,
3. Viele-zu-eins-Zusammenhang: Die Assoziation vieler psychologischer Elemente mit demselben physiologischen Element,
4. Viele-zu-viele-Zusammenhang: Die Assoziation mehrerer psychologischer Elemente mit den gleichen oder überlappenden physiologischen Elementen,

5. Null-Zusammenhang: Das nicht Vorhandensein einer Assoziation zwischen psychischen und physischen Elementen.

Von den fünf gelisteten Zusammenhangsmustern erlaubt es nur der Eins-zu-eins- sowie der Viele-zu-eins-Zusammenhang von den physiologischen Reaktionen auf die psychischen Zustände der Person zu schließen; oder anders ausgedrückt, psychologische Elemente als Funktion physiologischer Elemente zu spezifizieren¹: $\Psi = f(\Phi)$ (siehe Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007). Wichtig ist dieser Sachverhalt in Anbetracht gängiger Forschungspraxis, psychologische Konstrukte (Ψ) experimentell zu manipulieren und physiologische Reaktionen (Φ) zu messen. Nur selten werden Elemente aus Φ manipuliert und die Elemente aus Ψ gemessen. Ein gängiges Beispiel ist die Untersuchung der Herzschlagfrequenz (Element aus Φ) in Abhängigkeit von der Aufgabenschwierigkeit (Element aus Ψ). Oft werden solche Befunde nicht nur voreilig als Beleg für die Indikatorfunktion der Herzschlagfrequenz (Φ) für die mentale Beanspruchung (Ψ) interpretiert, sondern auch so behandelt, als ob man die Herzschlagfrequenz (Φ) direkt manipuliert und die mentale Beanspruchung (Ψ) gemessen hätte. Eine solche Interpretation spiegelt die implizite Annahme wider, dass Φ und Ψ isomorph zueinander sind (Berntson & Cacioppo, 2002; Cacioppo et al., 2007; Sarter, Berntson & Cacioppo, 1996).

Um die gängigen psychophysiologischen Befunde richtig interpretieren zu können, ist somit entweder der Nachweis der Isomorphie zwischen Ψ und Φ zu erbringen oder zumindest die Abweichung von der Isomorphie zu berücksichtigen. Möchte man trotz eines fehlenden Eins-zu-eins-Zusammenhangs auf die zugrunde liegenden psychologischen Konstrukte schließen, dann muss ein Weg gefunden werden, den Zusammenhang zwischen Ψ und Φ als einen Eins-zu-eins-Zusammenhang oder mindestens als einen Viele-zu-eins-Zusammenhang

¹ Hypothesen, die sich auf die Beteiligung physiologischer Aktivität (Φ), wie der Herzschlagfrequenz, bei der Empfindung mentaler Beanspruchung (Ψ) beziehen – bzw. Forschungsbestreben, die physiologische Parameter als Indikatoren mentaler Beanspruchung untersuchen – zielen auf die Schlussfolgerung $\Psi = f(\Phi)$ ab. Dieses bedeutet, dass das Eintreten von Φ eine hinreichende Bedingung für Ψ ist (Berntson & Cacioppo, 2002).

zu spezifizieren (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017).

4.1.2 Herleitung eines Eins-zu-eins-Zusammenhangs

Angewandt auf die Erfassung mentaler Beanspruchung mit physiologischen Parametern bedeuten die oben aufgeführten Überlegungen, dass valide Rückschlüsse auf die mentale Beanspruchung gezogen werden können, wenn bestenfalls ein Eins-zu-eins-Zusammenhang nachgewiesen wird. Leider sind solche Eins-zu-eins-Zusammenhänge in der psychophysiologischen Forschung wohl eher die Ausnahme. Ein Blick in die Fachliteratur verdeutlicht dabei die Komplexität der Ausgangslage. Es hängen nicht nur mehrere psychologische Konstrukte, wie mentale Beanspruchung oder Emotionen, mit denselben physiologischen Parametern, sondern auch mehrere physiologische Parameter mit denselben psychologischen Konstrukten zusammen (siehe Kapitel 3). Insgesamt ergibt dieses Muster einen Viele-zu-viele-Zusammenhang, der auf den ersten Blick keine hinreichend validen Rückschlüsse erlaubt. Ein Lösungsansatz für dieses Problem besteht nach Cacioppo et al. (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017) in der Neustrukturierung der Elemente aus Φ und Ψ . Dieses kann erreicht werden, indem die physiologischen Elemente aus Φ in ein Profil höherer

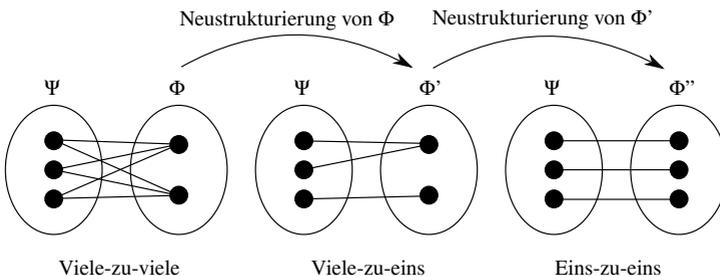


Abbildung 4.2: Ein Viele-zu-viele-Zusammenhang kann durch die Neustrukturierung der Elemente aus Φ zu einem Profil Φ' in einen Viele-zu-eins-Zusammenhang und durch die Neustrukturierung der Elemente aus Φ' zu einem Profil Φ'' in einen Eins-zu-eins-Zusammenhang umgewandelt werden.

Ordnung Φ' umgewandelt werden. Φ' bildet dabei eine Kombination bzw. ein Profil der Elemente aus Φ (siehe Abbildung 4.2). Bspw. zwei verschiedene Kombinationen kardialer Parameter, die zwei unterschiedliche Konstrukte, wie mentale Beanspruchung und Emotionen, abbilden. Mit diesem Vorgehen kann ein Viele-zu-viele-Zusammenhang zu einem Viele-zu-eins-Zusammenhang oder im besten Fall zu einem Eins-zu-eins-Zusammenhang vereinfacht werden. Analog ist diese Methode auch für die psychologischen Elemente in Ψ anwendbar: Wird bspw. der Zusammenhang zwischen Ψ und Φ' (bzw. Ψ und Φ'') durch individuelle Unterschiede beeinflusst, dann können diese individuellen Unterschiede als interagierende Elemente höherer Ordnung in Ψ' definiert werden (Cacioppo & Tassinari, 1990). Mit diesem Vorgehen wird nicht nur schrittweise eine Annäherung von einem Viele-zu-eins-Zusammenhang (oder gar Viele-zu-viele-Zusammenhang) hin zu einem Eins-zu-eins-Zusammenhang erreicht, sondern auch die Voraussetzung für die Interpretation physiologischer Signale als Funktion psychologischer Konstrukte geschaffen.

4.1.3 Taxonomie psychophysiologischer Zusammenhänge

Inwieweit eine physiologische Reaktion als Indikator eines psychischen Zustands interpretiert werden kann, wird nach Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007) mit Hilfe der Gütekriterien Spezifität, Generalisierbarkeit und Sensitivität bestimmt. Diese werden im Nachfolgenden in Bezug auf das Rahmenmodell näher erläutert.

Die *Spezifität* bezieht sich auf die Art des Zusammenhangs zwischen Φ und Ψ . Der Zusammenhang wird auf einem Kontinuum von einem angestrebten Eins-zu-eins- bis Viele-zu-eins-Zusammenhang eingeordnet. Im Sinne der analytischen Spezifität (siehe Kapitel A.2.2) wird hier der Grad überprüft, inwieweit eine physiologische Reaktion (bzw. ein physiologisches Reaktionsprofil) assoziiert ist mit einem oder mehreren psychologischen Zielzuständen (z. B. mentale Beanspruchung), jedoch gleichzeitig keine Assoziation mit anderen psychologischen Zuständen (z. B. Emotionen) aufweist.

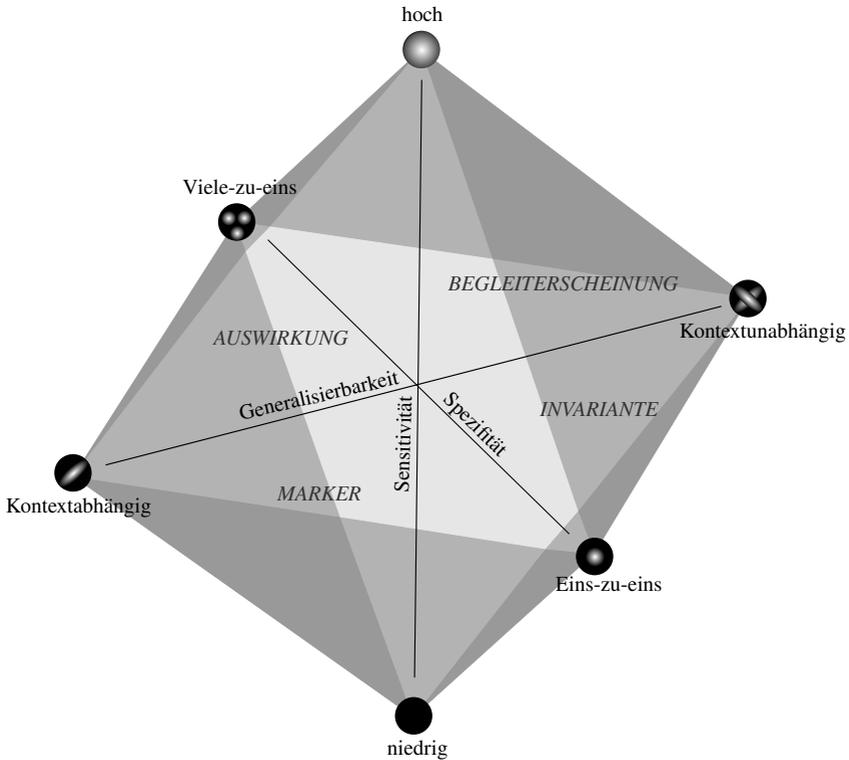


Abbildung 4.3: Die Gütekriterien Spezifität, Generalisierbarkeit und Sensitivität bilden die drei Achsen eines Oktaeders. Auf der horizontalen Ebene werden die vier Zusammenhangsklassen Auswirkung, Marker, Begleiterscheinung und Invariante in Abhängigkeit von den Kriterien Spezifität und Generalisierbarkeit abgegrenzt. Grafik modifiziert nach Cacioppo et al. (2007, S. 11).

Neben der Spezifität berücksichtigen Cacioppo et al. die *Generalisierbarkeit* (externe Validität) psychophysiologischer Zusammenhänge über Situationen oder Personen hinweg. Vor allem die Untersuchung psychophysiologischer Zusammenhänge im Forschungslabor stellt die externe Validität psychophysiologischer Forschungsergebnisse in Frage. Folglich ist die Übertragbarkeit der Ergebnisse bei der Interpretation psychophysiologischer Zusammenhänge stets zu berücksichtigen.

Die *Sensitivität* gibt Aufschluss darüber, inwieweit Veränderungen in der physiologischen Reaktion Veränderungen im zugrunde liegenden psychischen Zustand widerspiegeln (siehe Kapitel A.2.1). Sie bezieht sich somit auf die Wahrscheinlichkeit, dass eine physiologische Reaktion aus Φ mit irgendeinem psychologischen Element aus Ψ einhergeht.

Aufbauend auf der orthogonalen Anordnung der drei Gütekriterien leiten Cacioppo et al. vier Zusammenhangsklassen ab: Auswirkung, Marker, Begleiterscheinung und Invariante. Diese sind in Abbildung 4.3 dargestellt. Die drei Gütekriterien bilden die Form eines Oktaeders, wobei die Sensitivität die vertikale Achse repräsentiert. Auf der horizontalen Ebene werden die Zusammenhangsklassen je nach Ausprägung der Generalisierbarkeit und Spezifität abgegrenzt. Die horizontale Ebene in Abbildung 4.3 gibt das optimale Sensitivitätsniveau an, mit dem eine maximale Differenzierung zwischen den vier Zusammenhangsklassen erreicht werden kann (siehe antagonistische Beziehung zwischen Spezifität und Sensitivität in Kapitel A.2.2). Im Folgenden werden die einzelnen Klassen bezüglich ihrer Implikationen für die Interpretation eines psychophysiologischen Zusammenhangs beschrieben.

Eine psychophysiologische **Auswirkung** (engl.: Outcome) ist im Idealfall ein situationsspezifischer Viele-zu-eins-Zusammenhang. Der Nachweis, dass eine Variation in der psychologischen Variable eine Variation in der physiologischen Variable nach sich zieht (Φ als eine Funktion von Ψ), ist zumeist einer der ersten Forschungsziele in Laborexperimenten. Da der Fokus solcher ersten Studien nicht auf der Spezifität, der Generalisierbarkeit oder einer Neustrukturierung der Elemente in Φ liegt, stellen Auswirkungen die erste und häufigste Form psychophysiologischer Befunde dar (Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007). Es ist zu beachten, dass Auswirkungen keine hinreichend validen Rückschlüsse physiologischer Reaktionen auf psychologische Zustände erlauben. Bspw. kann ein Anstieg in der Herzschlagfrequenz als ein Auswirkungs-Zusammenhang mit der mentalen Beanspruchung beschrieben werden. Dieser erlaubt allerdings nicht, von einem beobachteten Anstieg der Herzschlagfrequenz auf die mentale Beanspruchung zu schließen.

Ein psychophysiologischer **Marker** ist im Idealfall ein situationspezifischer Eins-zu-eins-Zusammenhang. Dieser Zusammenhang sagt aus, dass das Auftreten einer physiologischen Reaktion das Auftreten einer psychologischen Reaktion valide vorhersagen kann. Ob dieser Zusammenhang extern valide ist, d. h. ob dieser aufgrund einer bestimmten Situation (z. B. kontrollierter Laborkontext oder systematische Störvariable) zustande kommt oder inhärent ist, bleibt bei einem Marker ungeklärt. Wichtig ist es hier die Randbedingungen (Situationsbedingungen) festzulegen, unter denen ein valider Rückschluss auf das psychologische Zielkonstrukt gezogen werden kann. Bei kleinen Abweichungen vom Isomorphismus besteht dennoch ein nützlicher, wenn auch kein idealer Marker-Zusammenhang (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007).

Eine psychophysiologische **Begleiterscheinung** (engl.: Concomitant) stellt im Idealfall einen situationsunabhängigen Viele-zu-eins-Zusammenhang dar. Hierbei kann ein physiologisches Element bzw. Profil das psychologische Zielkonstrukt zwar valide und situationsübergreifend vorhersagen, allerdings ist das Auftreten einer physiologischen Reaktion nicht zwangsläufig mit dem Vorliegen des psychologischen Zielkonstrukts verbunden (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007).

Eine psychophysiologische **Invariante** (engl.: Invariant) ist im Idealfall ein situationsübergreifender Eins-zu-eins-Zusammenhang. Im Falle einer Invariante darf nach Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007) das physiologische Element bzw. Profil nur dann eine Veränderung anzeigen, wenn auch eine Veränderung in der psychologischen Zielvariable zu beobachten ist (und umgekehrt). Zudem muss diese Voraussetzung über alle Situationen hinweg erfüllt sein. Von allen Zusammenhangsklassen ist es nur für den Invariante-Zusammenhang möglich, Studienergebnisse, bei denen Ψ variiert und Φ gemessen wurde, mit Studienergebnissen, bei denen Φ variiert und Ψ gemessen wurde, gleichzusetzen. Allerdings wird in der gängigen Forschungspraxis ein Isomorphismus a priori angenommen, ohne diesen vorher zu überprüfen (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007). Die

Folgen solcher Vorannahmen fassen Cacioppo und Tassinari (1990, S. 26) wie folgt zusammen:

Hence, assuming rather than establishing that a relation is invariant can lead to erroneous inferences and misleading theoretical “advances” [...] This practice can also create the illusion that psychophysiological relations are inherently unreliable and that there is nothing psychologically significant to be inferred from physiological signals.

Auch wenn sich die oben aufgeführten Zusammenhangsklassen in ihrer Interpretationsstärke als psychophysiologischer Indikator unterscheiden, bedeutet es nicht, dass bestimmte Zusammenhangsklassen, wie Auswirkungen, wertlos sind. Wichtig ist in diesem Zusammenhang, die Grenzen der Spezifität, Sensitivität und Generalisierbarkeit bei der Interpretation psychophysiologischer Ergebnisse zu beachten. Die hier vorgestellte Taxonomie bietet eine fundierte Grundlage, um psychophysiologische Zusammenhänge zu analysieren und einzuordnen.

4.1.4 Überlappung und Abgrenzung zur psychometrischen Validität

Neben dem Rahmenmodell von Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007, 2017) bietet auch ein fundierter Nachweis der psychometrischen Validität (siehe Kapitel A.2.4) eine Grundlage für die valide Interpretation einer physiologischen Reaktion als Indikator psychischer Zustände (Vertiefung siehe Richter & Slade, 2017). Vor diesem Hintergrund ist allerdings die Form des Validitätsnachweises mit entscheidend für die Interpretationsstärke des psychophysiologischen Zusammenhangs. So wiegt bspw. ein Nachweis der Kriteriumsvalidität – wie dieser in der mentalen Beanspruchungsforschung gängig ist – nicht so stark wie der relativ aufwendige Nachweis der Konstruktvalidität. Nichtsdestoweniger ist der Validitätsnachweis als psychometrisches Kriterium und das Rahmenmodell von Cacioppo et al.

komplementär zu verstehen. So kann der Nachweis eines Zusammenhangs zwischen der physiologischen Reaktion und dem psychischen Zielzustand sowohl als Nachweis der Kriteriumsvalidität bzw. der konvergenten Validität und auch als erster Schritt im Rahmenmodell von Cacioppo et al. gewertet werden (Richter & Slade, 2017). Der Fokus im Rahmenmodell von Cacioppo et al. richtet sich auf die Spezifität physiologischer Variablen. Das vorrangige Ziel besteht hierbei in der Herleitung eines Eins-zu-eins-Zusammenhangs (Marker oder Invariante), sodass von einer gemessenen physiologischen Reaktion auf den dahinter liegenden psychischen Zustand geschlossen werden kann. Analog zur konvergenten und diskriminanten Validität ist der entscheidende Schritt bei der Validierung eines physiologischen Markers bzw. einer Invariante der Nachweis, dass die physiologische Variable nur mit dem psychischen Zielzustand und nicht mit anderen psychologischen Zuständen assoziiert ist.

Das Rahmenmodell von Cacioppo et al. bietet auch Vorzüge im Vergleich zur Konstruktvalidität. Denn der Nachweis der Konstruktvalidität für physiologische Maße ist mit gängigen Methoden wie der Multitrait-Multimethod-Analyse (MTMM-Analyse; Campbell & Fiske, 1959) schwierig. Der Grund liegt darin, dass physiologische Maße a priori keine Konstruktrennung erlauben. Im Vergleich dazu liegen für subjektive Maße konstruktsspezifische Messinstrumente vor (z. B. Emotionsfragebogen oder Fragebogen zur mentalen Beanspruchung). Um also die MTMM-Methode auf physiologische Maße anzuwenden, müssten vorab Parameterprofile ermittelt werden, die spezifischen Konstrukten (z. B. mentale Beanspruchung) zugeordnet sind. Zukünftige Untersuchungen zur Validität physiologischer Maße erfordern demnach die Ermittlung von physiologischen oder sogar psychologischen Profilen, um eine über die Kriteriumsvalidität hinausgehende Erkenntnis gewinnen zu können. Somit leistet das Rahmenmodell von Cacioppo et al. auch einen Mehrwert zur Konstruktvalidierung psychophysiologischer Reaktionen.

4.2 Extraktion physiologischer Profile mentaler Beanspruchung

Einen einzelnen physiologischen Parameter zu finden, der die mentale Beanspruchung valide abbildet, ist in Anbetracht der Komplexität psychophysiologischer Zusammenhänge nicht realisierbar. Die Befundlage zeigt auf, dass die mentale Beanspruchung gleichzeitig in mehreren physiologischen Parametern ein Reaktionsmuster aufweist, sodass mindestens von einem Eins-zu-viele-Zusammenhang ausgegangen werden muss. Empirisch betrachtet eröffnet sich sogar ein Viele-zu-viele-Zusammenhang. Dieser sollte zumindest auf einen viele-zu-eins-Zusammenhang reduziert werden, um die Grundlage für valide Rückschlüsse zu schaffen. Folglich ist es wichtig, sich mit der Frage zu beschäftigen, wie mehrere physiologische Parameter simultan analysiert und sinnvoll zu physiologischen Profilen zusammengeführt werden können. Physiologische Profile sind nach Blascovich, Vanman, Berry Mendes und Dickerson (2011) den Einzelparametern in ihrer Spezifität und Sensitivität überlegen.

Die Herleitung physiologischer Profile der mentalen Beanspruchung birgt noch weitere Vorzüge. Die resultierenden Profile können als Aggregat inhaltlich ähnlicher Variablen – vergleichbar mit der Aggregation mehrerer Fragebogen-Items zu einer Skala – verstanden werden. Aus Sicht der klassischen Testtheorie (KTT) werden durch die Kombination mehrerer Messwerte desselben Konstrukts zufällige Fehlerquellen eliminiert und somit die Reliabilität gesteigert (vgl. Strube, Newman, Lord & Nguyen, 2017). Physiologische Profile würden somit nicht nur einen Beitrag zur Validität, sondern auch zur Reliabilität physiologischer Messverfahren leisten.

Zukünftiges Forschungsbestreben sollte sich somit auf die Zusammenfassung einzelner Parameter zu physiologischen Profilen richten. Ein solches Vorgehen ist im Bereich des Psychophysiological Computing kein Novum. Wenngleich diese Teildisziplin nicht am theoretischen Rahmenmodell von Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007, 2017) ausgerichtet

ist, so berücksichtigt das klassische Vorgehen (zsf. Novak et al., 2012; siehe Kapitel 3.4.1) bereits die Erstellung von Profilen aus mehreren physiologischen Parametern. Allerdings betrachtet das Psychophysiological Computing die Profilbildung lediglich als einen optionalen Schritt der Datenvorverarbeitung und zeigt kein Interesse an den latenten Merkmalsstrukturen hinter den Profilmustern. Diese sind simpel gesprochen nur ein Mittel zum Zwecke der Variablenreduktion und Genauigkeitssteigerung (vgl. Novak et al., 2012). Aus diesem Grund wurden in diesem Bereich weder geeignete physiologische Parameter definiert (jede Studie bedient sich einer eigenen Variablenauswahl), noch die Interkorrelationen bzw. Redundanzen bestimmter physiologischer Variablengruppen vertieft thematisiert. Welche physiologischen Variablengruppen sich zur Vorhersage der mentalen Beanspruchung eignen, war hier bisher nicht von Belang. Nichtsdestoweniger bieten die modernen Auswertemethoden dieser Teildisziplin eine vielversprechende Grundlage, um multivariate Analyseverfahren auf physiologische Daten anzuwenden. Bspw. können dimensionsreduzierende Analyseverfahren, wie die Hauptkomponentenanalyse (engl.: Principal Component Analysis; PCA), im Sinne der Profilbildung nach Cacioppo et al. eingesetzt werden.

4.2.1 Dimensionsreduzierende Verfahren in der psychophysiologischen Forschung

Die PCA bietet nicht nur den Vorteil einer Dimensionreduktion multivariater Daten, sondern sie ermöglichen es auch, die Struktur psychophysiologischer Daten zu erschließen. Einfach ausgedrückt, fasst die PCA vorhandene Informationen aus den physiologischen Daten (möglichst ohne Informationsverlust) in wenige Komponenten zusammen (siehe Abbildung 4.4). Ganz im Sinne des Rahmenmodells von Cacioppo et al. können zusammengehörige sowie unabhängige physiologische Parameter identifiziert und zu Profilen zusammengefasst werden. Ob eine resultierende Komponente nun als mentale Beanspruchung, Angst, Bewegung oder sympathische bzw. parasympathische Aktivität zu interpretieren

ist, hängt vom theoretischen Rahmenmodell und der Operationalisierung der zu untersuchenden Konstrukte ab.

Die PCA als ein dimensionsreduzierendes und strukturaufdeckendes Verfahren wurde bereits in den 90-ern von Backs (1995, 1998) in psychophysiologischen Studien genutzt. Mit Hilfe der PCA extrahierte Backs aus unterschiedlichen kardialen Parametern (des EKG und der Impedanzkardiographie) zwei Komponenten, die er als sympathische und parasymphatische Aktivität interpretierte. Sein Ziel bestand jedoch nicht in der Erstellung physiologischer Profile mentaler Beanspruchung, sondern in der Aufklärung der beobachtbaren Dissoziation zwischen einer gemessenen Herzschlagfrequenz und den subjektiven sowie leistungsbezogenen Maßen mentaler Beanspruchung. Hierbei stützt sich Backs auf das Modell des Autonomic Space (Berntson et al., 1991; siehe Kapitel 3.1). Die Dissoziation begründet Backs (1995) mit den verschiedenen Formen der autonomen Steuerung (Antagonismus, Koaktivierung und Entkopplung). Dabei versuchen Backs et al. (Backs, 2000; Backs, Rohdy & Barnard, 2005; Lenneman & Backs, 2003) nicht nur die Form der autonomen Steuerung aus den zwei Komponenten zu bestimmen, sondern diese auch unterschiedlichen kognitiven Verarbeitungsprozessen analog Wickens (2002, 2008) (siehe Kapitel 2.2.1)

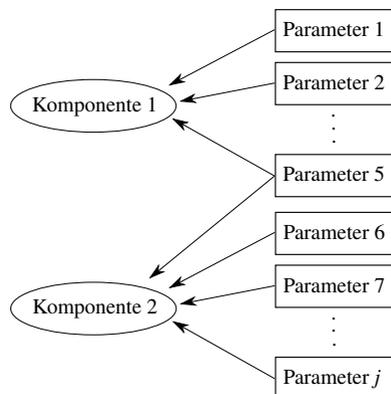


Abbildung 4.4: Die PCA fasst beobachtbare Parameter (Rechtecke) zu nicht beobachtbaren Komponenten (Ovale) zusammen.

zuzuordnen. Die Form der autonomen Steuerung leiten Backs et al. (Backs, 1995, 1998; Backs et al., 2005; Lenneman & Backs, 2007) aus den Faktorwerten der sympathischen und parasympathischen Komponente ab.

Ob die von Backs und Kollegen extrahierten Komponenten tatsächlich als sympathische und parasympathische Aktivierung interpretiert werden dürfen, ist hierbei nicht erwiesen. Auch wenn Backs (1995, 1998) eine gewisse inhaltliche Validität und Stabilität der Faktorenlösung über Experimente hinweg annimmt, so fehlt dennoch der Beleg der Konvergenz mit validen Indikatoren sympathischer und parasympathischer Aktivität. Ebenso fehlt ein Nachweis der Divergenz mit konkurrierenden oder konfundierenden Konstrukten. Eine solche mangelnde Validierung der Interpretation von Komponentenlösungen ist keine Seltenheit. Sie bildet eine der methodischen Herausforderungen dimensionsreduzierender Verfahren in der psychophysiologischen Forschung. Diese Herausforderungen und deren Lösungen werden im nachfolgenden beschrieben.

4.2.2 Methodische Herausforderungen dimensionsreduzierender Verfahren

Erste Herausforderungen dimensionsreduzierender Verfahren sind die Variationsquellen, die durch das Verfahren abgebildet werden können. Die drei klassischen Variationsquellen psychophysiologischer Forschung (vgl. Stemmler, 2001) sind (1) die Unterschiede zwischen den Personen (interindividuell), (2) die Unterschiede zwischen den Situationen (intraindividuell) und (3) die Unterschiede zwischen den Variablen (z. B. Items in einem Fragebogen oder physiologische Einzelparameter). Die PCA ist in der Lage, nur zwei Variationsquellen abzubilden. Ist man bspw. an interindividuellen (über Situationen stabilen) Unterschieden, sog. Dispositionen, interessiert, erfasst man in der Regel von mehreren Personen eine bestimmte Anzahl an Variablen (z. B. Items im Fragebogen). Forschungsdisziplinen, wie die differentielle oder diagnostische Psychologie, können sich dabei auf eine natürlich gegebene interindi-

viduelle Variation des zu untersuchenden psychologischen Konstrukts (z. B. Persönlichkeitsunterschiede) stützen. Für den Forschungsbereich der mentalen Beanspruchung ist dies allerdings nur bedingt möglich, da die mentale Beanspruchung ein situationsabhängiges Merkmal ist. Die mentale Beanspruchung manifestiert sich als Zustand und steht somit im Gegensatz zur Disposition. Oft muss die interessierende intraindividuelle Variation durch eine Manipulation der Situation künstlich hergestellt oder die Situation vorab erfasst werden. In solchen Fällen ist die Validität für die Variationsquelle der Situation zu prüfen. Die Analyse intraindividuelle Unterschiede mit der PCA bedeutet gleichzeitig, dass die Variationsquelle der Personen durch die Variationsquelle der Situation ausgetauscht werden muss.

Dies bringt eine weitere Herausforderung mit sich, nämlich die interindividuelle Varianz physiologischer Daten. Während diese Varianzquelle für interindividuelle Ansätze erwünscht und erkenntnisbringend ist, wird diese Varianz bei der Untersuchung intraindividuelle Unterschiede als störend empfunden. Folglich wurde bis in die Gegenwart versucht, diese auszupartialisieren, zu mitteln oder als irrelevante Variation zu kontrollieren. Nicht zuletzt werden nicht konforme Forschungsergebnisse auf diese Varianzquelle zurückgeführt. Über die Jahrzehnte ist jedoch eine Trendwende zu beobachten, welche die Relevanz interindividuelle Variation bei der Erforschung psychophysiologische Zusammenhänge in den Fokus rückt (vgl. Stemmler, 2001). Die Trendwende führte sogar dazu, dass psychophysiologische Zusammenhänge für jede Person einzeln analysiert werden (z. B. S. Chen & Epps, 2013; Haapalainen et al., 2010; Wilson & Russell, 2003a, 2003b).

Insgesamt wird deutlich, dass die mentale Beanspruchung als situationsabhängiger Personenzustand unter Berücksichtigung nur zweier Variationsquellen nicht vollständig abgebildet werden kann. Benötigt wird eine simultane Betrachtung von drei Variationsquellen: die Variation über Situationen, über Personen und über physiologische Parameter hinweg. Letzteres, um auf Basis der Parametervariation und -kovariation unterschiedliche physiologische Profile zu erstellen und inhaltlich voneinander abzugrenzen. Dimensionsreduzierende

Analyseverfahren, die dieser Herausforderung gewachsen sind, existieren bereits. Zu diesen zählen u. a. die Parallele Faktorenanalyse (PARAFAC) oder die Tucker-Analyse (Übersicht siehe Kroonenberg, 2008). Diese Methoden bilden eine Erweiterungen zur PCA, um eine Dimensionsreduktion bzw. Komponentenextraktion für Daten mit drei oder mehr Varianzquellen durchzuführen. Eine Vertiefung zu den Prinzipien der mehrdimensionalen Analyse wird in Unterkapitel 5.6.2 gegeben. In der psychophysiologischen Forschung werden diese mehrdimensionalen Verfahren bisher kaum angewandt. Bekannt sind lediglich Studien aus dem Fachgebiet der Neurowissenschaften (z. B. Beckmann & Smith, 2005; Miwakeichi et al., 2004; Morup, Hansen, Herrmann, Parnas & Arnfred, 2006; Schmitz et al., 2014). Bspw. untersuchen Schmitz et al. (2014) die Synchronität neuronaler Aktivitätsmuster mit Hilfe der PARAFAC-Analyse. Die PARAFAC-Analyse wird ebenfalls von Miwakeichi et al. (2004) verwendet, um spektrale und topographische Profile im EEG zu erstellen.

Eine weitere Herausforderung besteht in der Interpretation der Komponenten. Denn die aus der PCA, der PARAFAC- oder der Tucker-Analyse gewonnenen Komponenten geben keine Auskunft über die dahinter liegenden latenten Merkmale (theoretischen Konstrukte). In diesem Fall ist durch logische Überlegungen zu schließen, ob die Komponenten das gewünschte Konstrukt oder gar eine Störvariable repräsentieren. Den Bezug zum psychologischen Konstrukt herstellen zu können, ist allerdings wesentlich, um inhaltlich sinnvolle und interpretierbare physiologische Profile mentaler Beanspruchung zu definieren. Eine Lösung für diese Herausforderung ist die Projektion der Komponentenlösung auf eine latente Merkmalsstruktur (vgl. Abdi, 2010). Einfach ausgedrückt, werden die Komponenten in der Art extrahiert, dass ein maximaler Zusammenhang mit dem psychologischen Konstrukt (gemessene mentale Beanspruchung) besteht (Abbildung 4.5). Erreicht werden kann das für zweidimensionale Datenstrukturen mit dem dimensionsreduzierenden Verfahren der *Partial Least Square Regression* (PLS-Regression) und für mehrdimensionale Datenstrukturen mit dem dimensionsreduzierenden Verfahren der *Multilinear Partial Least Squares Regression* (NPLS-Regression; Vertiefung siehe Kapitel 5.6.2). Beide Verfah-

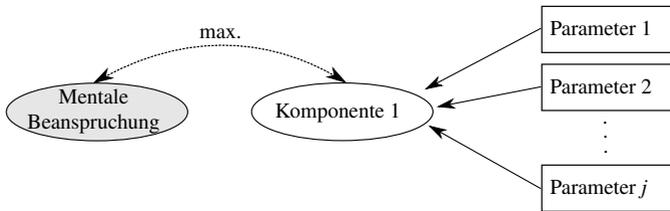


Abbildung 4.5: Projektion der Komponentenlösung auf das psychologische Konstrukt der mentalen Beanspruchung. Der empirische Zusammenhang zwischen der gemessenen mentalen Beanspruchung und der Komponente 1 ist maximal.

ren ermöglichen es nicht nur, physiologische Profile zu extrahieren, die am repräsentativsten für die mentale Beanspruchung sind, sondern auch physiologische Profile konkurrierender Konstrukte, wie Emotionsprofile, zu diesen abzugrenzen.

4.3 Forschungsziele und Forschungsprogramm

Beanspruchungssensitive Mensch-Maschine-Schnittstellen stehen vor der Herausforderung, die gemessenen physiologischen Parameter valide auf den mentalen Beanspruchungszustand der Person zurückzuführen. Nach dem Rahmenmodell von Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007, 2017) ist ein valider Rückschluss nur dann möglich, wenn im Idealfall ein Eins-zu-eins-Zusammenhang oder zumindest ein Viele-zu-eins-Zusammenhang zwischen den physiologischen Elementen in Φ und den psychologischen Elementen in Ψ vorliegt. Nach dem heutigen Stand der Forschung ist allerdings von einem Viele-zu-viele-Zusammenhang auszugehen. Durch eine Neustrukturierung der Elemente aus Φ in ein Profile höherer Ordnung Φ' kann der Viele-zu-viele-Zusammenhang in einen Viele-zu-eins-Zusammenhang und bestenfalls in einen Eins-zu-eins-Zusammenhang überführt werden. Ausgehend von der Ausprägung der Spezifität, Generalisierbarkeit und Sensitivität ist es möglich, die physiologischen Profile als Auswirkung, Marker, Begleit-

erscheinung oder Invariante zu spezifizieren (vgl. Cacioppo & Tassinary, 1990; Cacioppo et al., 2000, 2007, 2017). Im Sinne des Rahmenmodells von Cacioppo et al. ist ein Ziel der vorliegenden Arbeit die Herleitung physiologischer Profile mentaler Beanspruchung. Damit wird der erste Schritt zur Neustrukturierung der Elemente aus Φ in ein Profile höherer Ordnung Φ' geschaffen. Zudem soll überprüft werden, ob die ermittelten Profile von physiologischen Profilen konkurrierender Konstrukte abgegrenzt und auf unterschiedliche Aufgabentypen generalisiert werden können.

In der vorliegenden Arbeit werden physiologische Profile separat für die kardiale und elektrodermale Aktivität abgeleitet. Auf eine Profilbildung über alle physiologische Maße hinweg wird verzichtet. Dies ist damit begründet, dass eine simultane Analyse kardialer und elektrodermalen Parameter nicht nur zu einer unüberschaubaren Menge an Parametern in der Analyse führen, sondern auch die Gefahr irrelevanter Einflüsse auf das Extraktionsverfahren, z. B. Methodeneffekte, erhöhen würde. Ebenfalls wird auf eine Vorselektion physiologischer Parameter verzichtet, da zum aktuellen Forschungsstand keine empirischen Selektionskriterien vorliegen. Notwendig hierfür sind empirische Ergebnisse darüber, welche physiologischen Parameter im Verbund eines Profils für welches Konstrukt bedeutsam sind.

Für die Zusammenfassung physiologischer Einzelparameter zu physiologischen Profilen werden in der vorliegenden Arbeit moderne, dimensionsreduzierende Analyseverfahren eingesetzt. Diese ermöglichen es, alle notwendigen Varianzquellen (Personen, Situationen und Parameter) simultan in die Analyse zu integrieren. Dadurch wird der mentalen Beanspruchung als situationsabhängiger Zustand Rechnung getragen und die starke interindividuelle Variation physiologischer Aktivität berücksichtigt. Zudem werden die physiologischen Profile so extrahiert, dass sie maximal mit der gemessenen mentalen Beanspruchung zusammenhängen und somit bestenfalls als Repräsentanten der mentalen Beanspruchung interpretiert werden können.

Das Vorgehen der vorliegenden Arbeit ist als induktiv bzw. datengetrieben zu verstehen. Diese Vorgehensweise ist damit begründet, dass bisher keine fundierte theoretische Basis zur Gruppierung bestimmter physiologischer Parameter zu Profilen vorliegt. Aufgrund der induktiven Vorgehensweise werden keine Hypothesen, sondern Fragestellungen (F) formuliert. Die ersten beiden Forschungsfragen F1 und F2 gelten studienübergreifend und werden jeweils in zwei Teilfragestellungen F1.1 und F1.2 bzw. F2.1 und F2.2 unterteilt. Diese lauten:

F1 Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive (F1.1) und leistungsbezogene (F1.2) mentale Beanspruchung zuverlässig repräsentieren?

F2 Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive (F2.1) und leistungsbezogene (F2.2) mentale Beanspruchung zuverlässig repräsentieren?

Zur Beantwortung dieser Fragestellungen werden drei Studien durchgeführt. Das übergeordnete Ziel aller drei Studien ist die Ableitung physiologischer Profile mentaler Beanspruchung. Dieses Ziel wird durch die Forschungsfrage F1 und F2 abgebildet. Durch die Beantwortung ein und derselben Fragestellung kann überprüft werden, ob es sich bei den extrahierten Profilen mentaler Beanspruchung um zufällige oder konsistente Befunde handelt. Nichtsdestoweniger sind die drei Studien nicht identisch. In Studie I liegt der Fokus auf der Spezifität, also der Abgrenzung physiologischer Profile der mentalen Beanspruchung von denen der emotionalen Beanspruchung. Der Fokus der Studie II ist die Strukturvalidität, also die Sicherstellung einer Kongruenz zwischen den physiologischen Profilen mentaler Beanspruchung bei unabhängigen Experimenten. Die Studie III fokussiert gleichzeitig die Strukturvalidität und die Generalisierbarkeit der physiologischen Profile mentaler Beanspruchung über verschiedene Aufgaben hinweg. Im folgenden Überblick werden das Forschungsprogramm sowie die studienspezifischen Fragestellungen vorgestellt.

Studie I: In der Studie I (Kapitel 6) werden erste kardiale und elektrodermale Profile der mentalen Beanspruchung extrahiert. Die mentale Beanspruchung wird als subjektive mentale Beanspruchung im NASA-RTLX und als leistungsbezogene mentale Beanspruchung in einer Sekundäraufgabe operationalisiert. Das spezifische Ziel der Studie I ist die Analyse der Spezifität der extrahierten kardialen und elektrodermalen Profile. Überprüft wird dabei die Abgrenzung der Profile mentaler Beanspruchung von den Profilen der emotionalen Beanspruchung. Zu diesem Zweck werden erste kardiale und elektrodermale Profile der subjektiven Erregung und Valenz extrahiert und mit den physiologischen Profilen der subjektiven und leistungsbezogenen mentalen Beanspruchung verglichen. Im Experiment wird die mentale Beanspruchung durch unterschiedliche Schwierigkeitsstufen einer computerbasierten Primär- und Sekundäraufgabe realisiert. Die emotionale Empfindung wird durch ein fingiertes Feedback sowie durch emotionale Bilder hervorgerufen.

Da neben der mentalen Beanspruchung die Extraktion physiologischer Profile emotionaler Beanspruchung im Fokus steht, werden für die Studie I folgende studienspezifische Forschungsfragen und Teilfragestellungen formuliert:

- F3** Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung (F3.1) und subjektive Valenz (F3.2) zuverlässig repräsentieren?
- F4** Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung (F4.1) und subjektive Valenz (F4.2) zuverlässig repräsentieren?
- F5** Können die kardialen (F5.1) und elektrodermalen (F5.2) Profile der mentalen und emotionalen Beanspruchung bedeutsam voneinander abgegrenzt werden?

Studie II: Wie bereits in Studie I werden auch in Studie II (Kapitel 7) kardiale und elektrodermale Profile der subjektiven sowie leistungsbezogenen mentalen

Beanspruchung ermittelt. Diese Studie verfolgt das Ziel, die Strukturvalidität der in Studie I extrahierten kardialen und elektrodermalen Profile mentaler Beanspruchung zu prüfen. Somit kann in Studie II die Frage beantwortet werden, ob sich die kardialen und elektrodermalen Profile der subjektiven und leistungsbezogenen mentalen Beanspruchung in einem weiteren, unabhängigen Experiment wiederfinden. Die studienspezifische Fragestellung der Studie II lautet:

F6 Sind die in Studie II extrahierten kardialen (F6.1) und elektrodermalen (F6.2) Profile mentaler Beanspruchung kongruent zu den in Studie I extrahierten Profilen mentaler Beanspruchung?

In Studie II werden zudem noch offene Aspekte aus den Ergebnissen der Studie I aufgegriffen. Erstens wird in Studie II die Körperbewegung der Versuchspersonen auf ein Minimum reduziert. Zweitens wird die Leistung als Primäraufgabenmaß erfasst. Operationalisiert wird die mentale Beanspruchung in Form einer auditiven Rechenaufgabe mit zwei Schwierigkeitsstufen.

Studie III: In Studie III (Kapitel 8) werden kardiale und elektrodermale Profile der subjektiven, jedoch nicht der leistungsbezogenen, mentalen Beanspruchung abgeleitet. Somit wird in III nur die Teilfragestellung 1.1 und 2.1 beantwortet. Der Verzicht auf die Erfassung der leistungsbezogenen mentalen Beanspruchung ist durch den studienspezifischen Fokus der Studie III bedingt. Dieser besteht in der Extraktion kardialer und elektrodermalen Profile mentaler Beanspruchung über verschiedene Aufgabentypen hinweg. Da die Primäraufgabenleistung sich zwischen den verschiedenen Aufgabentypen sowohl qualitativ als auch quantitativ unterscheidet, wird auf die Extraktion physiologischer Profile der leistungsbezogenen mentalen Beanspruchung verzichtet.²

² Auf die Verwendung derselben Sekundäraufgabe über verschiedene Aufgabentypen hinweg wurde verzichtet. Ein solches Vorgehen würde im Widerspruch zum studienspezifischen Ziel, nämlich der Extraktion physiologischer Profile über verschiedene Aufgaben hinweg, stehen.

Mit Hilfe der Studie III wird geprüft, ob die physiologischen Profile mentaler Beanspruchung über verschiedene Aufgabentypen hinweg generalisierbar sind. Zu diesem Zweck werden im Experiment fünf computerbasierte Aufgabentypen in ihrer Schwierigkeit variiert. Die Aufgabentypen bilden jeweils unterschiedliche kognitive Belastungsfaktoren ab. Zudem wird in Studie III ein weiteres Mal die Strukturvalidität geprüft, indem die physiologischen Profile aus Studie I und II mit den aufgabenübergreifenden physiologischen Profilen aus Studie III verglichen werden. Die studienspezifische Fragestellung der Studie III lautet:

- F7** Sind die aufgabenübergreifenden kardialen (F7.1) und elektrodermalen (F7.2) Profile mentaler Beanspruchung der Studie III kongruent zu den kardialen und elektrodermalen Profilen mentaler Beanspruchung der Studien I und II?

Aus den Studienbeschreibungen ist zu entnehmen, dass es sich bei allen drei Studien um Experimente handelt, bei denen die mentale Beanspruchung (Studie I bis III), die emotionale Beanspruchung (Studie I) und der Aufgabentyp (Studie III) experimentell manipuliert werden. Die Analyse der Fragestellungen F1 bis F7 ist somit an die Voraussetzung geknüpft, dass die experimentelle Manipulation der Konstrukte gelungen ist. Dieses Gelingen soll vorab statistisch geprüft werden. Die zur Manipulationsprüfung gehörenden Annahmen werden spezifisch für Studie I in Kapitel 6.2.5, für Studie II in Kapitel 7.2.5 und für Studie III in Kapitel 8.2.5 in Form von Hypothesen und Unterfragestellungen beschrieben. Diese beziehen sich auf die folgenden Kontrollfragestellungen (KF):

- KF1** Ist die experimentelle Induktion mentaler und emotionaler Beanspruchung in Studie I gelungen?
- KF2** Ist die experimentelle Induktion mentaler Beanspruchung in Studie II gelungen?
- KF3** Ist die experimentelle Induktion mentaler Beanspruchung in Studie III gelungen?

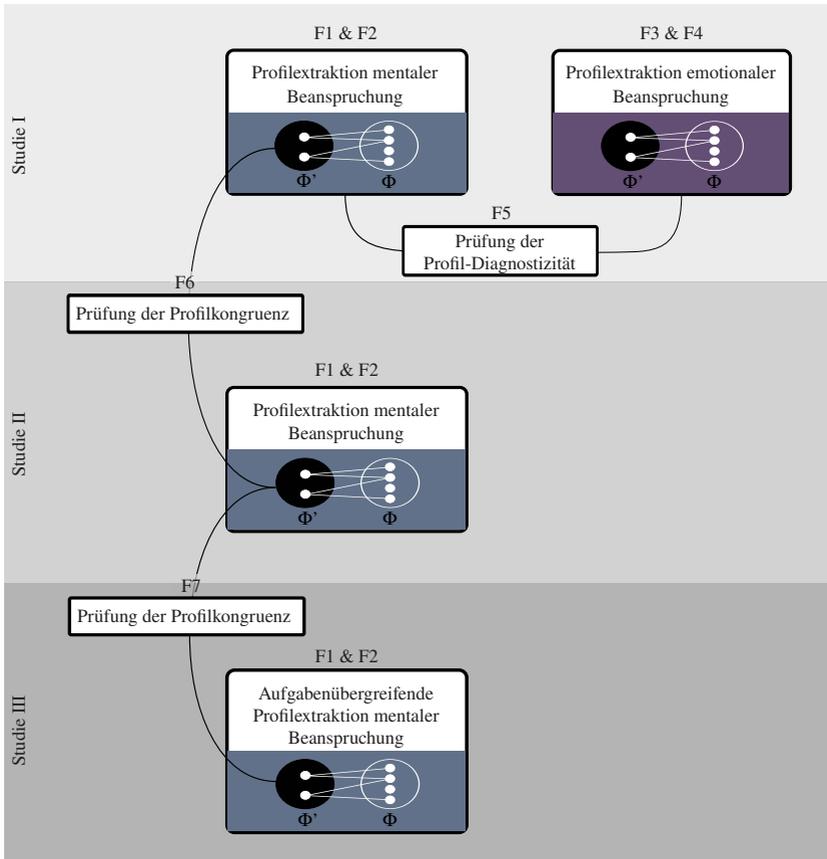


Abbildung 4.6: Einordnung der Forschungsfragen F in die Studien I, II und III.

Eine Übersicht aller Forschungsfragen sowie deren Einordnung in die Studien I, II und III ist in Abbildung 4.6 zu sehen. Abschließend werden die kardialen und elektrodermalen Profile mentaler Beanspruchung aus allen drei Studien gegenübergestellt und miteinander verglichen. Auf diese Weise können Hinweise zur Konsistenz und Generalisierbarkeit der extrahierten physiologischen Profile über die Experimente hinweg gewonnen werden. Unter Hinzunahme weiterer Gütekriterien werden in Kapitel 9.1 die gewonnenen kardialen und elektroder-

malen Profile mentaler Beanspruchung in die Taxonomie von Cacioppo und Tassinary (1990) sowie Cacioppo et al. (2000, 2007, 2017) eingeordnet.

5 Methode

Die Methoden werden an mehreren Stellen der Arbeit beschrieben. Einmal in diesem Kapitel zur Darstellung der studienübergreifenden Methoden und darüber hinaus in den Kapiteln 6, 7 und 8 als studienspezifische Methodenkapitel. In den studienübergreifenden Methoden werden die verwendeten Mess- und Auswertemethoden der physiologischen Signale (Kapitel 5.1 und 5.2) sowie deren Störeinflüsse (Kapitel 5.3) aufgeführt. Zudem werden die studienübergreifenden Methoden zur Messung der subjektiven (Kapitel 5.4) und der leistungsbezogenen mentalen Beanspruchung (Kapitel 5.5) dargestellt. Zum Schluss erfolgt eine Beschreibung der studienübergreifenden statistischen Auswertung der Messdaten. Studienspezifische Methoden wie studienspezifische Materialien, Stichproben, Versuchsdesign und -durchführung sowie studienspezifische Auswertemethoden werden in den jeweiligen Methodenkapiteln der Studien I, II und III erläutert.

5.1 Messung der kardialen Aktivität

Im Folgenden wird die Messapparatur und die Aufbereitung sowie Parametrisierung der kardialen Aktivität für die Studien I, II und III vorgestellt.

5.1.1 Apparatur zur Messung der kardialen Aktivität

In allen drei Studien der vorliegenden Arbeit wurde zur Messung der kardialen Aktivität ein EKG verwendet. Das Messprinzip dieses Verfahrens wurde bereits

in Kapitel 3.2.2 beschrieben. Für die Messung des kardialen Signals wurden zwei verschiedene Messapparaturen eingesetzt.¹ Diese werden nachfolgend spezifiziert.

Für die Studie I erfolgte die Registrierung der kardialen Aktivität mit Hilfe der Shimmer3 ECG Unit der Firma Shimmer. Die Abtastrate des verwendeten EKG betrug 204.8 Hz. Verwendet wurde die bipolare Ableitung $V_{LL}-V_{RA}$ in Anlehnung an die Einthoven-Ableitung. Im Gegensatz zur klassischen Einthoven-Ableitung erfolgte die Platzierung der Elektroden gemäß den Empfehlungen des Handbuchs (Shimmer, 2016a, S. 8) an der Brustwand. Eingesetzt wurden selbstklebende Einweg-Gel-Elektroden.

In Studie II und III wurde die kardiale Aktivität mit Hilfe des varioport-B Biosignal Rekorders der Firma Becker Meditec aufgezeichnet. Die Abtastrate betrug 500 Hz. Abweichend zu den vorgestellten Ableitungen in Kapitel 3.2.2 empfiehlt das Varioport Handbuch eine Ableitung des oberen Ende des Brustbeins gegen den linken untersten Rippenbogen (Becker Meditec, 2016, S. 15). Gemäß dieser Empfehlungen wurden selbstklebende Einweg-Gel-Elektroden an der Brustwand platziert.

5.1.2 Aufbereitung und Parametrisierung der kardialen Messdaten

Für die Vorverarbeitung und Parametrisierung der EKG-Daten wurden die Python-Bibliotheken BioSPPy Version 0.5.1 (Carreiras et al., 2015) und NeuroKit Version 0.2.5 (Makowski, 2016) genutzt. Zur Beseitigung der Störfrequenzen aus dem Rohsignal des EKG wurde ein Bandpassfilter mit einem Grenzbereich

¹ Die Verwendung unterschiedlicher Sensoren ist durch den Projektkontext der Studie I begründet. Die Studie I wurde im Rahmen des Projekts KonsensOP (Kontextsensitive Assistenz im aufmerksamen OP) durchgeführt. Um das Operationspersonal bei der Ausführung ihrer Tätigkeit nicht zu behindern, wurden drahtlose Sensoren (Shimmer3 ECG Unit) verwendet. In Studie II und III wurde hingegen auf kabelgebundene Sensoren zurückgegriffen, um die Robustheit bei der Datenerfassung zu erhöhen.

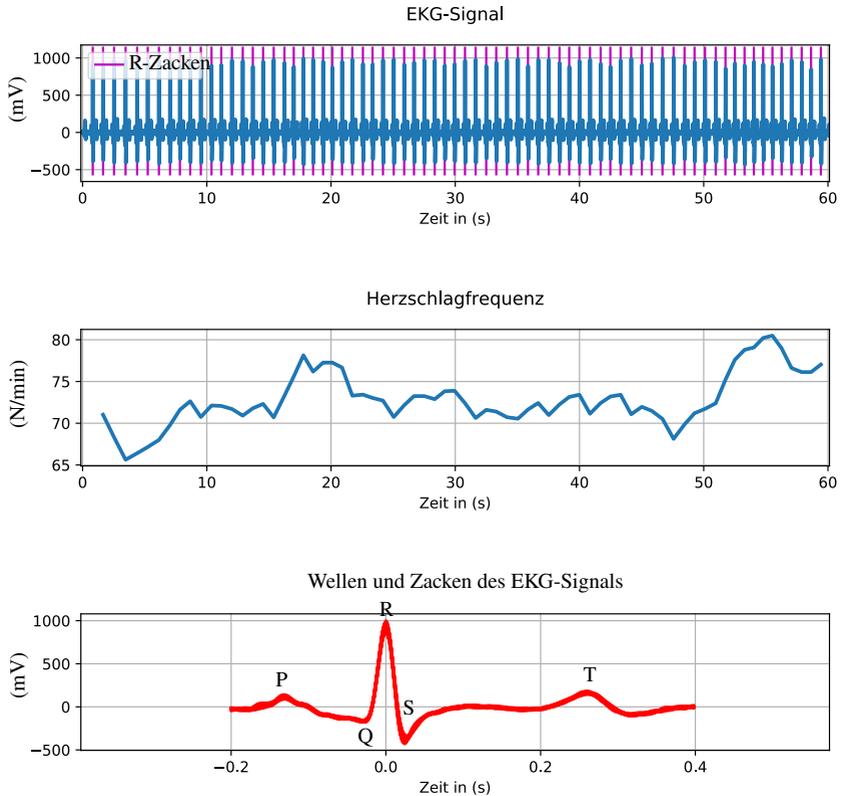


Abbildung 5.1: Beispiel für ein EKG-Signal, entnommen aus Studie III über ein Zeitintervall von 60 s. Oben: Gefiltertes EKG-Signal mit den detektierten R-Zacken (magenta Linien). Mitte: Herzschlagfrequenz über die Zeit als Anzahl der Schläge pro Minute. Unten: Übereinander gelegte Abschnitte des EKG-Signals bestehend aus P- und T-Wellen sowie QRS-Komplex.

zwischen 3 Hz und 45 Hz verwendet. Zur Bestimmung der RR-Intervalle wurden die QRS-Komplexe in BioSPPy extrahiert (siehe Abbildung 5.1). Die extrahierten RR-Intervalle dienen als Grundlage für die Quantifizierung unterschiedlicher EKG-Metriken. Zur Ableitung von NN-Intervallen (vgl. Lippman, Stein & Lerman, 1994; Salo, Huikuri & Seppänen, 2001) wurden in der Bibliothek NeuroKit all diejenigen RR-Intervalle ausgeschlossen, die 25 % größer oder kleiner waren als das vorangegangene RR-Intervall. Aus den resultierenden

NN-Intervallen wurden die in Kapitel 3.2.3 beschriebenen Parameter berechnet. Eine Beschreibung aller verwendeter Parameter findet sich in Tabelle 5.1. Hierbei werden die Parameter gemäß Malik et al. (1996) bzw. Shaffer und Ginsberg (2017) in zeitbezogene, frequenzbezogene und nichtlineare Maße eingeteilt. Als zusätzliche Kategorie werden morphologische Maße verwendet. Aufgrund des geringen Messintervalls von 60 s fallen alle kardialen Parameter in den Bereich der Ultra-Kurzzeitmessung.

Tabelle 5.1: Übersicht der verwendeten kardialen Parameter. Alle Parameter beziehen sich auf ein Zeitintervall von 60 s.

Parameter	Gemessen in	Beschreibung
Maße aus dem Zeitbereich (statistisch)		
HR	1/min	Mittlere Herzschlagfrequenz
SDHR	1/min	Standardabweichung der Herzschlagfrequenz
HP	ms	Mittlere Herzperiode (mittleres NN-Intervall)
SDNN	ms	Standardabweichung der NN-Intervalle
MQSD	ms	Quadratisches Mittel der Unterschiede zwischen zwei benachbarten NN-Intervallen
pNN50	%	Prozentualer Anteil sukzessiver NN-Intervall-Differenzen größer als 50 ms
pNN20	%	Prozentualer Anteil sukzessiver NN-Intervall-Differenzen größer als 20 ms
CVSD		Verhältnis MQSD zur HP
CVNN		Verhältnis SDNN zur HP
Maße aus dem Zeitbereich (geometrisch)		
Triang		Integral der NN-Intervall-Dichteverteilung geteilt durch das Maximum der Dichteverteilung

Shannon.h		Shannon-Entropie auf Basis der Klassenwahrscheinlichkeiten der NN-Intervall-Dichteverteilung
Maße aus dem Frequenzbereich		
EKG.TP	ms ²	Spektrale Leistungsdichte im gesamten Frequenzbereich (0.0033–0.4 Hz)
EKG.LF	ms ²	Spektrale Leistungsdichte im niedrigen Frequenzbereich (0.04–0.15 Hz)
EKG.HF	ms ²	Spektrale Leistungsdichte im hohen Frequenzbereich (0.15–0.4 Hz)
EKG.LFn		Normierte spektrale Leistungsdichte im niedrigen Frequenzbereich
EKG.LF/HF		Verhältnis der LF zur HF
Nichtlineare Maße		
DFA1		Trendbereinigende Fluktuationsanalyse der Kurzzeit-Zyklusschwankungen (4–16 Herzschläge)
SampEn		Sample-Entropie als Komplexitätsmaß der HP über die Zeit
SVDEn		Entropie der Singulärwertzerlegung als Indikator der Dimensionalität der NN-Intervall-Reihe
Morphologische Maße		
T.amp	mV	Mittlere Amplitude der T-Welle
T.area	mV·ms	Mittlere Fläche unter der T-Welle

Zusätzlich zu den in Kapitel 3.2.3 beschriebenen zeitbezogenen Parametern wurde der Variationskoeffizient aufeinanderfolgender Zyklusunterschiede als $CVSD = MQSD / HP$ und $CVNN = SDNN / HP$ berechnet (vgl. Voss et al.,

1996; Voss, Schroeder, Heitmann, Peters & Perz, 2015). Bei der Herleitung geometrischer Maße ist es notwendig, die NN-Intervalle in eine diskrete Skala mit einer optimalen Klassenlänge zu konvertieren, um eine geglättete Dichteverteilung für die weitere Analyse zu erhalten. Die Bibliothek NeuroKit verwendet hierbei die empfohlene Klassenbreite von 8 ms (siehe Malik et al., 1996). Aus dieser Dichteverteilung wird für den Parameter Triang das Integral berechnet. Dieselbe Dichteverteilung wird in NeuroKit zur Berechnung des Parameter Shannon.h (als die Entropie der NN-Dichteverteilung) genutzt. Dabei wird die Shannon-Entropie (Shannon, 1948) auf Basis der Klassenwahrscheinlichkeiten der NN-Intervall-Dichteverteilung berechnet (vgl. Voss et al., 1996, 2015). Damit wird die Unsicherheit bzw. der Informationsgehalt der Dichteverteilung quantifiziert. Da die Berechnung der Entropie in diesem Fall auf Basis der NN-Dichteverteilung (also dem geometrischen Muster der NN-Intervalle) erfolgt, wird der Parameter Shannon.h als geometrisches Maß eingeordnet.²

Vor der Berechnung frequenzbasierter Maße wurden die NN-Intervalle durch Interpolation der Messpunkte in eine äquidistante Zeitreihe überführt. Als Methode der Frequenzanalyse wurde die Welch-Methode gewählt. Die spektrale Leistungsdichte wurde entsprechend der in Kapitel 3.2.3 beschriebenen Frequenzkomponenten (LF und HF) berechnet (siehe Tabelle 5.1). Die VLF ist für Zeitintervalle < 60 s nicht geeignet und wurde folglich aus der Analyse ausgeschlossen. Als zusätzlicher frequenzbezogener Parameter wurde die Total-Power (TP) mit in die Analyse aufgenommen. Diese stellt gemäß Malik et al. (1996) die gesamte spektrale Leistungsdichte von 0.0033 Hz bis 0.4 Hz dar. Der Relativwert der HF-Komponente (HF_{fn}) wurde nicht in die weitere Analyse einbezogen. Dieser ist aufgrund der verwendeten Normierung (siehe Gleichung 5.1) redundant zum LF_{fn} (siehe Gleichung 5.2). Normiert wurde gemäß den Empfehlungen von Malik et al. (1996).

² Eine Einordnung des Parameters Shannon.h als nichtlineares Maß ist ebenfalls möglich.

$$\text{LFn} = \frac{\text{LF}}{(\text{TP} - \text{VLF})} \cdot 100 \quad (5.1)$$

$$\text{HFn} = 1 - \text{LFn} \quad (5.2)$$

Als nichtlineare Maße wurden die in Kapitel 3.2.3 beschriebenen Parameter SampEn und DFA1 mit Hilfe der Bibliothek NeuroKit berechnet. DFA2 wurde ausgeschlossen, da eine Betrachtung von Langzeit- Zyklusschwankungen (16 bis 64 Herzschläge) in einem Zeitintervall von 60 s nicht sinnvoll ist. Zusätzlich wurde der Parameter SVDEn (engl.: Singular Value Decomposition Entropy) aus der Bibliothek NeuroKit verwendet. Dieser gibt Auskunft über die Anzahl benötigter orthogonaler Vektoren, die zur Erklärung der Daten benötigt werden. Je höher der Wert, desto mehr orthogonale Vektoren (mehr Dimensionen) werden zur Erklärung der NN-Intervall-Reihe benötigt. Forschungsbefunde zu diesem Parameter liegen nur für den Bereich der klinischen Diagnostik von Herzerkrankungen vor (Li, Yang, Li & Cai, 2008; Shao et al., 2020). Li et al. (2008) berichtet eine geringe SVDEn bei Personen mit einer Herzinsuffizienz sowie mit steigendem Alter. Der Zusammenhang dieses Parameters mit der mentalen oder emotionalen Beanspruchung wurde bisher nicht untersucht.

Zur Quantifizierung morphologischer Maße (T.amp und T.area) wurde zunächst die T-Welle in der Bibliothek NeuroKit identifiziert. Die Differenz des T-Wellen-Maximums zum Fußpunkt des EKG-Signals wurde als Amplitude der T-Welle (T.amp) definiert. Die Fläche der T-Welle wurde als die durch den Fußpunkt eingeschlossene Fläche berechnet (vgl. Heine et al., 2017). Sowohl aus den Amplituden als auch aus den Flächen wurde ein Mittelwert für ein 60 s Intervall berechnet.

5.2 Messung der elektrodermalen Aktivität

Dieses Kapitel beschreibt die Messapparatur, die Aufbereitung sowie die Parameterisierung der elektrodermalen Aktivität für die Studien I, II und III.

5.2.1 Apparatur zur Messung der elektrodermalen Aktivität

In allen drei Studien der vorliegenden Arbeit wurde die elektrodermale Aktivität mit der exosomatischen Methode erfasst (siehe Unterkapitel 3.3.2). In den einzelnen Studien kamen unterschiedliche Messapparaturen zum Einsatz.³

Für die Studie I erfolgte die Registrierung der elektrodermalen Aktivität mit der Shimmer3 GSR+ Unit der Firma Shimmer (Messbereich: 10 k Ω bis 4,7 M Ω bzw. 0,2 μ S bis 100 μ S). Gemessen wurde der Hautwiderstand, welcher in den Hautleitwert umgerechnet wurde. Für die Messung wurden zwei mit Silber-Silberchlorid (Ag/AgCl) überzogene Trockenelektroden (GSR Electrodes von Shimmer Sensing) gemäß den Empfehlungen von Venables und Christie (1980, zitiert nach Boucsein, 2012, S. 220) an den mittleren Fingergliedern (Phalangen) des Zeige- und Ringfingers angebracht und eine Spannung von 0,5 V am Körper angelegt. Die Abtastrate bei der Datenerfassung betrug 51,2 Hz. Die Rohwerte der Shimmer3 GSR+ Unit wurden in k Ω ausgegeben. Diese wurden für die weitere Datenanalyse in μ S umgewandelt.

In Studie II und III wurde die elektrodermale Aktivität mit Hilfe des varioport-B Biosignal Rekorders von Becker Meditec (Messbereich: 0 μ S bis 50 μ S; Auflösung: 0,002 μ S) aufgezeichnet. Die Abtastrate betrug 500 Hz. Für die Messung wurden zwei mit Silber-Silberchlorid (Ag/AgCl) überzogene, selbstklebende

³ Die Verwendung unterschiedlicher Sensoren ist durch den Projektkontext der Studie I begründet. Die Studie I wurde im Rahmen des Projekts KonsensOP durchgeführt. Um das Operationspersonal bei der Ausführung ihrer Tätigkeit nicht zu behindern, wurden drahtlose Sensoren (Shimmer3 GSR+ Unit) verwendet. In Studie II und III wurde hingegen auf kabelgebundene Sensoren zurückgegriffen, um die Robustheit bei der Datenerfassung zu erhöhen.

Elektroden gemäß den Empfehlungen von Venables und Christie (1980, zitiert nach Boucsein, 2012, S. 220) ebenfalls an den mittleren Fingergliedern (Phalangen) des Zeige- und Ringfingers platziert und auch eine Spannung von 0.5 V am Körper angelegt. Um den Kontakt zwischen den Elektroden und der Hautoberfläche zu verbessern sowie eine Mindestfeuchtigkeit der Haut zu gewährleisten, wurde eine isotonische Paste (0.5 % NaCl) auf die Messfläche der Elektroden aufgetragen.

5.2.2 Aufbereitung und Parametrisierung der elektrodermalen Messdaten

Die Verarbeitung der elektrodermalen Signale erfolgte mit den Python-Bibliotheken BioSPPy Version 0.5.1 (Carreiras et al., 2015), NeuroKit Version 0.2.5 (Makowski, 2016) und cvxEDA (Greco, Valenza, Lanata, Scilingo & Citi, 2016). Zur Entfernung von Störfrequenzen (z. B. durch Bewegung der Versuchspersonen) aus dem Hautleitwertsignal wurde gemäß den Empfehlungen des Nutzerhandbuchs (Shimmer, 2016b, S. 11) ein Tiefpassfilter mit einer Grenzfrequenz von 5 Hz verwendet. Im Anschluss erfolgte die Zerlegung des EDA-Signals in eine tonische und eine phasische Komponente. Zu diesem Zweck wurden in den letzten Jahrzehnten eine Reihe von mathematischen Modellen entwickelt. Ein erst kürzlich entwickeltes Modell von Greco und Kollegen nutzt eine Kombination verschiedener mathematischer Methoden (vgl. Greco et al., 2016). Neben der Konvolution zur Modulierung der phasischen Komponente wird im Modell die bayessche Statistik sowie der Ansatz der konvexen Optimierung sowie der Sparsity zur Zerlegung des EDA-Signals in eine tonische, phasische sowie eine Störkomponente genutzt. Erste Untersuchungen bestätigen eine zufriedenstellende Vorhersagevalidität des Verfahrens (vgl. Greco et al., 2016; Greco et al., 2017). Für die vorliegende Signalanalyse wurde das implementierte Verfahren in der Bibliothek cvxEDA genutzt. Die Abbildung 5.2 zeigt ein Beispiel für die Aufteilung des elektrodermalen Signals (oben) in eine tonische (Mitte) und phasische Komponente (unten).

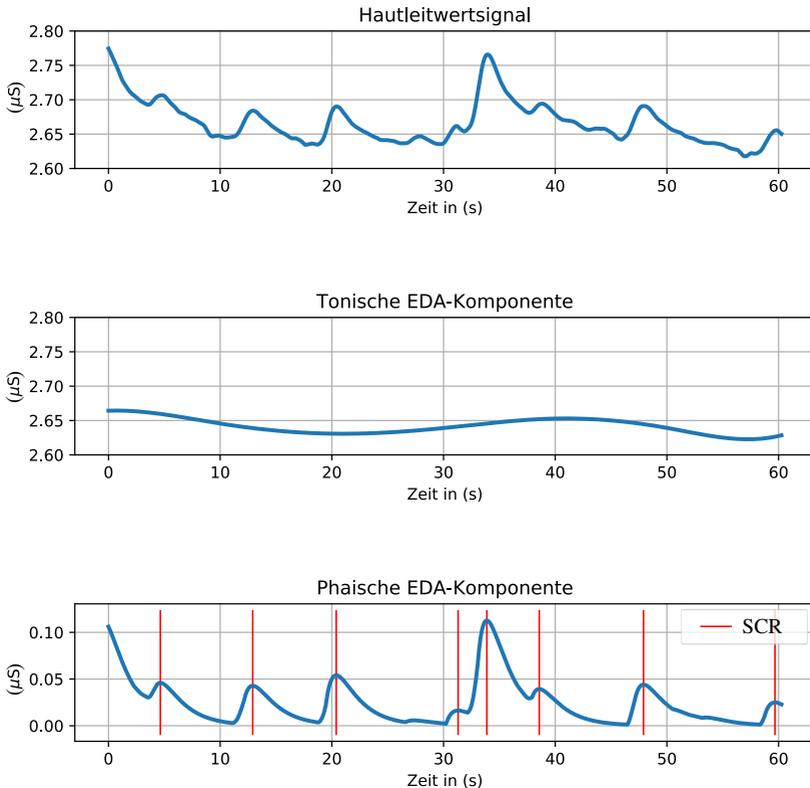


Abbildung 5.2: Beispiel für ein EDA-Signal, entnommen aus Studie III über ein Zeitintervall von 60 s. Oben: Gefiltertes EDA-Signal als Hautleitwert. Mitte: Tonische Komponente des EDA-Signals. Unten: Phaische Komponente des EDA-Signals mit detektierten EDRs. Amplitudenmaximum eines EDR ist als vertikale rote Linie dargestellt.

Die tonische EDA-Komponente dient zur Berechnung des mittleren elektrodermalen Aktivitätslevels (EDL.M) über die gesamte Dauer der Aufgabe. Zudem wurde aus der tonischen EDA-Komponente die absolute Änderung des EDL zwischen Aufgabenbeginn und -ende (EDL.D) berechnet. Aus der phaischen Komponente des Signals wurden die nicht-spezifischen elektrodermalen Reaktionen (NS-EDRs) mit Hilfe eines gradientenbasierten Verfahrens der Bibliothek NeuroKit extrahiert. Der Amplitudengrenzwert wurde gemäß Boucsein (2012,

S. 157) auf $0.01 \mu\text{S}$ festgelegt. Auf Grundlage der NS-EDRs wurden weitere EDA-Parameter berechnet. Diese sind in Tabelle 5.2 beschrieben. Alle Parameter beziehen sich auf ein Zeitintervall von 60 s.

Tabelle 5.2: Übersicht der verwendeten elektrodermalen Parameter. Alle Parameter beziehen sich auf ein Zeitintervall von 60 s. Alle Parameter werden aus dem Hautleitwert berechnet

Parameter	Gemessen in	Beschreibung
Maße aus dem Zeitbereich (basierend auf der tonischen Komponente)		
EDL.M	μS	Mittlerer tonischer Niveauwert des Hautleitwerts
EDL.D	μS	Absolute Änderung der tonischen Komponente zwischen Aufgabenbeginn und -ende
Maße aus dem Zeitbereich (basierend auf den NS-EDRs bzw. der phasischen Komponente)		
Phasic.SD	μS	Standardabweichung der phasischen Komponente des Hautleitwerts
EDR.N		Anzahl der NS-EDRs
EDR.amp	μS	Mittlere Amplitude der NS-EDRs
EDR.rise.t	s	Mittlere Zeit zwischen Reaktionsbeginn und Reaktionsmaximum der NS-EDRs
EDR.rec.t/2	s	Mittlere Zeit zwischen Reaktionsmaximum und dem Punkt eines 50 %-igen Abfalls der Amplitude der NS-EDRs
EDR.rec.t/c	s	Mittlere Zeit zwischen Reaktionsmaximum und dem Punkt eines 63 %-igen Abfalls der Amplitude der NS-EDRs
EDR.area	$\mu\text{S}\cdot\text{ms}$	Mittlere Fläche unter den NS-EDRs
EDR.rise.v	$\mu\text{S}/\text{s}$	Mittlere Geschwindigkeit des Anstiegs der NS-EDRs

EDR.peaks	s	Mittlere Zeit zwischen dem Reaktionsmaximum aufeinanderfolgender NS-EDRs
-----------	---	--

Maße aus dem Frequenzbereich

EDA.LF	μS^2	Spektrale Leistungsdichte des Hautleitwert-Signals im niedrigen Frequenzbereich (0.04–0.15 Hz)
EDA.HF	μS^2	Spektrale Leistungsdichte des Hautleitwert-Signals im hohen Frequenzbereich (0.15–0.4 Hz)
EDA.VHF	μS^2	Spektrale Leistungsdichte des Hautleitwert-Signals im sehr hohen Frequenzbereich (0.4–0.5 Hz)

Tonische Parameter, die aus den NS-EDR abgeleitet wurden (z. B. EDR.amp) sind von dem Vorhandensein einer NS-EDR abhängig. Ist dies innerhalb des Zeitintervalls von 60 s nicht der Fall, weißt der Parameter für diese Variable einen fehlenden Wert auf. Die Fläche unter der EDR (EDR.area) wird analog Traxel (1957 zitiert nach Boucsein, 2012, S. 165) durch Gleichung 5.3:

$$\text{EDR.area} = \text{amp}_{\text{max}} \cdot T \quad (5.3)$$

angenähert (siehe Abbildung 5.3). Neben den beschriebenen Parametern aus Kapitel 3.3.3 wurden zwei weitere Parameter aus den NS-EDRs abgeleitet. Dazu gehört die durchschnittliche Geschwindigkeit eines NS-EDR Anstiegs (EDR.rise.v). Dieser Parameter berücksichtigt nicht nur die Zeit zwischen Reaktionsbeginn und Reaktionsmaximum, sondern gleichzeitig auch die Amplitude des Reaktionsmaximums. Der zweite Zusatzparameter, EDR.peaks, bildet die mittlere Zeit zwischen den einzelnen NS-EDRs ab. Darüber hinaus wurde aus der phasischen Komponente des EDA-Signals explorativ der Parameter Phasic.SD berechnet. Dieser beschreibt die phasische Variabilität der EDA als

Standardabweichung der phasischen Komponente. Forschungsbefunde zum Zusammenhang dieses Parameters mit der mentalen Beanspruchung liegen bisher nicht vor.

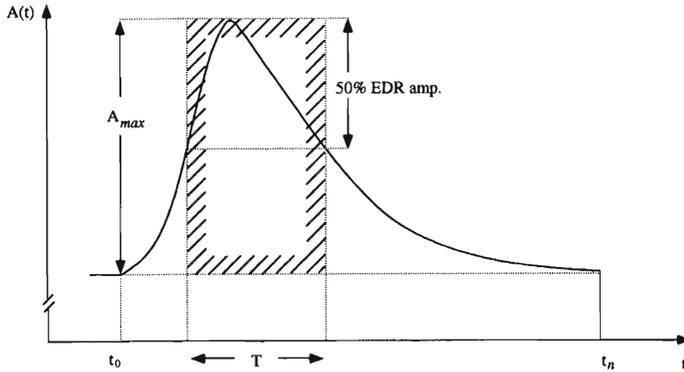


Abbildung 5.3: Berechnung der Fläche unter der EDR nach Traxel (1957 zitiert nach Boucsein, 2012, S. 165). Grafik aus Boucsein (2012, S. 164).

Alle bisher aufgeführten Parameter bilden tonische Maße aus dem Zeitbereich ab. In der vorliegenden Arbeit wurden zusätzlich die tonischen Kenngrößen aus dem Frequenzbereich ermittelt. Hierfür wurde die spektrale Leistungsdichte des elektrodermalen Signals berechnet (vgl. F. Chen et al., 2016, S. 90–91; Ghaderyan & Abbasi, 2016; Posada-Quintero et al., 2016a; Posada-Quintero, Florian, Orjuela-Cañón & Chon, 2016b). Einen Konsens bezüglich eines bewährten Frequenzbereichs (wie für die EKG-Parameter) gibt es für die EDA-Frequenzparameter bisher allerdings noch nicht. In der Literatur findet sich sowohl eine grobe Einteilung des Spektrums in gleich große Bereiche für 0–1 Hz oder 0.05–5 Hz, als auch eine Orientierung an den Frequenzkomponenten des EKG (z. B. F. Chen et al., 2016, S. 90–91; Ghaderyan & Abbasi, 2016). Ein wichtiger Unterschied zu den frequenzbezogenen Maßen der kardialen Aktivität ist, dass die Frequenzkomponenten der EDA nur die Aktivität des Sympathikus – jedoch nicht des Parasympathikus – repräsentieren. Erste Ergebnisse der elektrodermalen Frequenzparameter zeigen einen positiven Zusammenhang mit der kognitiven Aufgabenanforderung (Posada-Quintero et

al., 2016a, 2016b). Dies wird vor allem den unteren Frequenzbereichen (0.08 Hz bis 0.24 Hz) zugesprochen (Posada-Quintero et al., 2016a, 2016b). In der vorliegenden Arbeit wurde das elektrodermale Signal über einen Tiefpassfilter mit einer Grenzfrequenz von 5 Hz gefiltert und anschließend die spektrale Leistungsdichte mit Hilfe der Welch-Methode berechnet.⁴ Berechnet wurden dieselben Frequenzbereiche, die bereits für die kardialen Frequenzparameter verwendet wurden. Zusätzlich wurde analog Posada-Quintero et al. (2016a) die Very-High-Frequency Komponente (VHF) mit einem Frequenzbereich von 0.4–0.5 Hz berechnet.

5.3 Störeinflüsse und Kontrollvariablen physiologischer Messung

Es gibt eine Vielzahl an Störeinflüssen, welche zu einer Verzerrung der kardialen und elektrodermalen Signale führen können. Zu den Störeinflüssen der kardialen und elektrodermalen Messung zählen bspw. Bewegungsartefakte, Artefakte durch starkes oder unregelmäßiges Einatmen und Husten. Ebenso kann Sprechen, Druck oder Zug an den Elektroden sowie eine instabile Platzierung der Elektroden auf der Hautoberfläche die Messung beeinflussen. Um solche Störeinflüsse zu vermeiden, wurde bereits bei der Messung im Experiment auf die Kontrolle und Erfassung bestimmter Faktoren geachtet. Zur Vermeidung von Bewegungsartefakten im elektrodermalen Signal, wurde ausschließlich auf der nicht-dominanten Hand (überwiegend links) gemessen. Die Hand wurde im Versuchsablauf auf einer weichen Unterlage mit der Handfläche nach oben platziert und sollte nicht bewegt werden. Zudem wurden alle Versuchspersonen dazu instruiert, während der Aufgabendurchführung nicht zu sprechen und sich möglichst nicht zu bewegen. Eine ordnungsgemäße Platzierung der Elektroden

⁴ Eine Aufteilung in eine tonische und eine phasische Komponente fand hier nicht statt.

wurde vor der Datenerfassung überprüft. Die Kontrolle der Atemaktivität konnte im Rahmen des Experiments nicht durchgeführt werden.⁵

Weiterhin kann sich die Luftfeuchtigkeit und -temperatur auf die Hautleitwertmessung auswirken, weshalb diese bei Experimentalstart und -ende als Kontrollvariable erfasst wurden. Zudem ist zu berücksichtigen, dass die Elektrode die Schweißverdunstung an der Hautoberfläche mindern kann, wodurch es zu einem kontinuierlichen Anstieg der Hydratation und somit auch des Hautleitwerts über die Zeit kommen kann. Vor allem bei lang andauernden Experimenten sollte daher die Zeit beim Vergleich von unterschiedlichen Abschnitten berücksichtigt werden. Als Kontrollvariable wurde die Reihenfolge der Experimentalbedingungen im zeitlichen Verlauf mit erfasst.

5.4 Messung der subjektiven mentalen Beanspruchung

Die subjektive Bewertung der mentalen Beanspruchung spiegelt das Erleben der Person wider und besitzt deshalb nach Hart und Staveland (1988) einen hohen Stellenwert in der Beanspruchungsmessung. Das ist u. a. einer der Gründe, weshalb subjektive Messverfahren ein beliebtes Kriterium sind, mit dem andere Messverfahren validiert werden (Kriteriumsvalidität). Aber auch subjektive Messverfahren sind bestimmten Limitationen unterworfen. Hierzu gehören u. a. Erinnerungseffekte (vor allem bei nachgelagerter Erfassung). Zudem können nur Dinge berichtet werden, die einem zum Zeitpunkt der Befragung bewusst sind. Eine Übersicht der Limitationen subjektiver Beanspruchungsmaße findet sich bspw. bei O'Donnell und Eggemeier (1986) oder Tsang und Wilson

⁵ Es ist anzumerken, dass die Atemaktivität, wie bereits die elektrodermale und kardiale Aktivität, durch die mentale Beanspruchung beeinflusst wird (vgl. Grassmann, Vlemincx, Leupoldt, Mittelstädt & Van den Bergh, 2016). Versucht man nun die kardialen und elektrodermalen Signale um sog. Atemartefakte im Signal zu bereinigen, steht man vor dem Problem der Differenzierung der Atem-Artefakte von wahren Effekten der mentalen Beanspruchung.

(1997). Folglich ist es wichtig, bewährte und standardisierte Messverfahren zu verwenden. Ein solches Verfahren ist der National Aeronautics and Space Administration Task Load Index (NASA-TLX) von Hart und Staveland (1988).

5.4.1 Der National Aeronautics and Space Administration Task Load Index (NASA-TLX)

Der NASA-TLX ist ein von Hart und Staveland (1988) entwickelter Fragebogen zur subjektiven Bewertung der mentalen Beanspruchung. Es ist eines der am meist eingesetzten Messverfahren und wurde bereits in vielen Anwendungsszenarien der Labor- und Feldforschung erprobt (siehe Hart, 2006). Eine deutsche Übersetzung des NASA-TLX wurde bspw. von Sepehr (1988) vorgelegt. Der NASA-TLX ist ein multidimensionales Messinstrument, welches in seiner ursprünglichen Form die mentale Beanspruchung aus dem gewichteten Mittelwert der sechs Unterskalen berechnet. Diese sind die (1) mentale Anforderung, (2) körperliche Anforderung, (3) zeitliche Anforderung, (4) Anstrengung, (5) Ausführung der Aufgabe und (6) Frustration.

Es ist zu beachten, dass die sechs Skalen des NASA-TLX nicht anhand eines theoretischen Rahmenmodells (vgl. Kapitel 2.2) entwickelt wurden (deduktives Vorgehen), sondern induktiv aus den empirischen Ergebnissen mehrerer Studien von Hart und Staveland (1988) abgeleitet sind. Das Ziel bestand dabei in der Identifikation relevanter Einflussfaktoren, die zur subjektiven Bewertung der allgemeinen mentalen Beanspruchung von Personen herangezogen werden. Für Hart und Staveland (1988) ist die mentale Beanspruchung keine inhärente Eigenschaft der Situation bzw. Aufgabe, sondern ein Resultat der Interaktion unterschiedlicher Einflussfaktoren: bspw. Aufgabenanforderungen, Erwartungen oder Wahrnehmungen der Person, welche sich personen- oder aufgabenspezifisch unterscheiden können. Diese Einflussfaktoren werden von den Autoren in drei Faktorkategorien aufgeteilt: Die (a) aufgabenbezogenen Faktoren beziehen sich auf die Bewertung objektiver Kriterien der Aufgabenanforderung und wer-

den im NASA-TLX durch die Skalen der mentalen, körperlichen und zeitlichen Anforderung repräsentiert. Die (b) verhaltensbezogenen Faktoren werden zur Bewertung der Aufgabenbewältigung herangezogen. Im NASA-TLX sind diese Faktoren mit der Anstrengung und der Ausführung der Aufgabe vertreten. Die (c) personenbezogenen Faktoren repräsentieren die Auswirkungen der Aufgabenanforderungen, des Verhaltens und der Leistung auf die Person. Zu diesen gehört die Skala Frustration. Eine Fülle an empirischen Studien bestätigen eine zufriedenstellende psychometrische Reliabilität und Validität des NASA-TLX (siehe Anhang A.3).

Die sechs Skalen des NASA-TLX werden für gewöhnlich nach Bearbeitung einer Aufgabe auf einer 20-stufigen Ratingskala von 1 (gering) bis 20 (hoch) bewertet (Erfassung der Beanspruchungshöhe). Bei der Auswertung wird die 20-stufige Ratingskala mit einem Inkrement von 5 berechnet, sodass sich ein Wertebereich zwischen 5 und 100 ergibt. Als nächstes werden die Personen dazu aufgefordert, die sechs Skalen mit Hilfe von 15 paarweisen Vergleichen in Hinblick auf ihren Einfluss auf die empfundene mentale Beanspruchung zu gewichten (Erfassung der Beanspruchungsstruktur). Die Gewichtung der Skalen soll Aufschluss über die personenspezifischen Definitionen, aber auch über die aufgabenspezifischen Ursachen der mentalen Beanspruchung geben (Hart & Staveland, 1988). Auf diese Gewichtung bzw. Erfassung der Beanspruchungsstruktur wird in der Rohversion des NASA-TLX, dem NASA-RTLX (Byers, Bittner & Hill, 1989), verzichtet. Byers et al. (1989) fanden in fünf verschiedenen Studien Korrelationswerte von .96 bis .98 zwischen dem NASA-TLX und dem NASA-RTLX und schlossen auf eine gleich hohe Sensitivität beider Versionen. Hendy, Hamilton und Landry (1993) berichten sogar eine höhere Sensitivität des NASA-RTLX. Nygren (1991) sieht in der Gewichtung der Skalen sogar psychometrische Schwächen und empfiehlt diese wegzulassen. Ein praktischer Grund, der gegen die Verwendung der Gewichtung spricht, ist die Zeitersparnis in der Versuchsdurchführung.

In der vorliegenden Arbeit wurde der NASA-RTLX favorisiert und somit nur die Beanspruchungshöhe und nicht die Beanspruchungsstruktur erfasst. Gerechnet

wurde mit dem ursprünglichen Wertebereich von 1–20, da eine Umwandlung der Skalenwerte in den Wertebereich 5–100 keinerlei psychomotorische Vorteile mit sich bringt. Bedingt durch den Fokus der vorliegenden Arbeit auf die mentalen Aufgabenanforderungen wurde die Skala körperliche Anforderung ausgeschlossen. Mit diesem Vorgehen wurde der Einfluss dieser Skala auf den Gesamtwert der mentalen Beanspruchung herausgenommen. Die Befragung fand nach jeder Experimentalaufgabe statt. Bewertet wurde das subjektive Erleben der vorangegangenen Aufgabe.

5.5 Messung der leistungsbezogenen mentalen Beanspruchung

Leistungsmaße gehören zu den Verhaltensmaßen und bilden die Qualität der Aufgabenausführung ab. Aus diesem Grund sind sie eine beliebte Methode zur Messung mentaler Beanspruchung. Leistungsmaße erlauben es nicht nur, die mentale Beanspruchung kontinuierlich zu erfassen, sondern sind auch mit geringem Aufwand verbunden. Verwendet werden sowohl Maße der Effektivität (z. B. Fehlerrate) als auch der Effizienz (z. B. Reaktionszeit).

In der Literatur (vgl. Schlick et al., 2010, S. 409–414) werden Leistungsmaße in zwei Klassen unterteilt. Die erste Klasse ist die Leistungsmessung in einer Primäraufgabe. Dies ist die Aufgabe, für welche die mentale Beanspruchung zu quantifizieren ist. Hier liegt die Annahme einer limitierten Informationsverarbeitungskapazität (siehe ressourcentheoretische Modelle, Kapitel 2.2.1) zugrunde. Die limitierte Informationsverarbeitungskapazität wird durch die Aufgabenbewältigung aufgebraucht (Tsang & Wilson, 1997). Eine steigende mentale Beanspruchung spiegelt sich in der Güte der Aufgabenausführung wider (Stanton, 2013, S. 290). Auch wenn diese Annahme eine gewisse Augenscheinvalidität besitzt, ist die Differenzierung der mentalen Beanspruchung von alternativen Erklärungen (z. B. Motivation) nicht immer möglich (geringe Diagnostizität nach Schlick et al., 2010, S. 409–414). Zudem kann sich die

Aufgabenleistung auch bei gleich bleibender mentaler Beanspruchung deutlich unterscheiden, und umgekehrt. Wickens et al. (2013, S. 323) erklären letzteren Effekt durch die sog. Reservekapazität bzw. Restkapazität, welche bei Bedarf, z. B. durch erhöhte Anstrengung, abgerufen werden kann. Eine Leistungsminderung wäre demnach erst im Falle einer Überlastung bzw. Überschreitung vorhandener Ressourcen zu beobachten. Folglich ist anzunehmen, dass die Sensitivität einer Primäraufgabenleistung nur für mittlere bis hohe mentale Beanspruchungszustände gegeben ist (Tsang & Wilson, 1997). Eine endgültige Bewertung der Reliabilität und Validität von Leistungsmaßen einer Primäraufgabe ist nach Stanton (2013, S. 293) schwierig, da diese jeweils an die spezifische Aufgabe und den Kontext angepasst und somit nicht standardisiert sind.

Die Sekundäraufgabenleistung bildet die zweite Klasse der Leistungsmaße und bezieht sich auf die Leistungsmessung in einer zusätzlichen Aufgabe oder Nebenaufgabe. Eine gängige Technik ist das *Subsidiary Task Paradigma*. Hier bearbeitet die Person eine konkurrierende Aufgabe unter der Priorisierung der Primäraufgabe, bspw. durch Instruktion oder monäteren Anreiz (Wickens et al., 2013, S. 331). Sekundäraufgabenmaße entstammen den Untersuchungen zum Multi-Tasking, wobei die Ressourcenmodelle die theoretische Basis bilden (siehe Wickens et al., 2013, S. 321–345). Hier liegt die Annahme zugrunde, dass die Leistung in der Sekundäraufgabe die Restkapazität, die bei der Bearbeitung der Primäraufgabe übrig bleibt, misst. Somit bildet die Leistung die Ressourcenauslastung der Primäraufgabe ab. Trotz zufriedenstellender Sensitivität und Diagnostizität (zsf. Stanton, 2013, S. 293) ist die Anwendung von Sekundäraufgabenmaßen nicht trivial und an gewisse Voraussetzungen und Vorannahmen geknüpft (O'Donnell & Eggemeier, 1986; Schlick et al., 2010, S. 413; zsf. Stanton, 2013, S. 292). Bspw. darf die Gesamtmenge der zur Verfügung stehenden Ressourcen einer Person nicht über verschiedene Schwierigkeitsstufen der Primäraufgabe variieren. Des weiteren darf die Sekundäraufgabe keinen Einfluss auf die Bearbeitung der Primäraufgabe (Interferenz) haben. Allerdings ist die Intrusion der Sekundäraufgabe und die damit einhergehenden zusätzli-

chen Kapazitätsanforderungen, z. B. durch Koordination der beiden Aufgaben (siehe Tulga & Sheridan, 1980), nur schwer kontrollierbar. Aus diesem Grund ist die richtige Wahl der Sekundäraufgabe eine Prämisse für die Validität der Messung. Bspw. muss die Sekundäraufgabe dieselbe kognitive Ressource (siehe Multiple-Ressourcen-Modell, Kapitel 2.2.1) in Anspruch nehmen wie auch die Primäraufgabe. Eine falsche Wahl (unterschiedliche kognitive Ressourcen) führt in so einem Fall zur Fehleinschätzung der zugrunde liegenden mentalen Beanspruchung (Tsang & Wilson, 1997).

In den Studien der vorliegenden Arbeit wurden je nach Studie unterschiedliche Leistungsmaße verwendet. Die Unterschiede ergeben sich durch die unterschiedlichen Aufgabenstellungen und -typen. Die Berechnung der Leistung erfolgte für jede Bedingung separat und bezog sich auf eine Zeitspanne von 60 s. In Studie I wurde die Effektivität der Aufgabenausführung als Anzahl richtiger Antworten in einer Sekundäraufgabe verwendet. Die Leistungsmessung in der Primäraufgabe war aufgrund der spezifischen Aufgabenstellung nicht möglich. Demgegenüber wurde in Studie II und III die Leistung in der Primäraufgabe gemessen. Auch hier wurde die Effektivität als Anteil der korrekten Aufgabenausführung operationalisiert. Auf die Messung der Sekundäraufgabe in Studie II und III wurde aus den folgenden Gründen verzichtet: Erstens wurden sowohl zwischen den Studien als auch innerhalb der Studie III unterschiedliche Aufgabentypen konzipiert. Dieselbe Sekundäraufgabe für unterschiedliche Aufgabentypen zu verwenden, ist allerdings nicht ratsam und auch nicht immer möglich. Die Entwicklung verschiedener Sekundäraufgaben sowie der empirische Nachweis der oben aufgeführten Voraussetzungen (vor allem Intrusionsfreiheit) birgt einen hohen Aufwand, der im Rahmen des vorliegenden Fokus (auf die physiologische Messung mentaler Beanspruchung) nicht gerechtfertigt wäre. Zweitens sollte eine Dissoziation zwischen subjektiven und leistungsbezogenen Maßen minimiert werden. Diese kommt durch die Sensitivität subjektiver Maße auf die Anzahl der parallel zu bearbeitenden Aufgaben zustanden. So beschreiben Yeh und Wickens (1988) eine höhere subjektive mentale Beanspruchung bei zwei leichten Aufgaben und guter Leistung im Vergleich zu einer

einzigsten schweren Aufgabe mit schlechtem Leistungsergebnis. Begründet wird dieses durch die Spezifität der jeweiligen Maße: Während leistungsbezogene Maße sensitiv auf das Ausmaß der Konkurrenz zwischen kognitiven Ressourcen reagieren, sind subjektive Maße sensitiv für die Gesamtmenge investierter Ressourcen (Yeh & Wickens, 1988). Im Fokus der vorliegenden Arbeit steht die Gesamtmenge investierter Ressourcen. Folglich wurde auf die Operationalisierung einer Sekundäraufgabe in den Studien II und III verzichtet. Eine Ausnahme bildet die Studie I, in der die Messung einer Primäraufgabenleistung bedingt durch das Studiendesign nicht möglich war und auf das Sekundäraufgabenmaß zurückgegriffen werden musste. Alle Leistungsmaße wurden als Mittelwert über die Dauer einer Experimentalaufgabe berechnet.

5.6 Statistische Auswertemethoden

Die Auswertung der Studien I bis III folgt stets demselben Vorgehen. In einem ersten Schritt werden Ausreißer aus den Messdaten eliminiert. Dabei werden alle Datenpunkte (Parameterausprägung in einer Bedingung), die mehr als $\pm 3 SD$ vom Gesamtmittelwert und vom personenspezifischen Mittelwert des jeweiligen Parameter abweichen, als Ausreißer definiert und aus der weiteren Analyse ausgeschlossen. In einem nächsten Schritt wird mit Hilfe der *Mehrebenenanalyse* (MEA) überprüft, ob die Manipulation der mentalen Beanspruchung durch die Experimentalbedingungen gelungen ist. Der Hauptfokus der Analyse richtet sich auf die Identifikation psychophysiologischer Profile. Zu diesem Zweck wird die NPLS-Regression (Multilinear Partial Least Squares Regression) berechnet. Die einzelnen Analysemethoden werden in den folgenden Kapiteln näher spezifiziert.

5.6.1 Mehrebenenanalyse

Mit der Mehrebenenanalyse (MEA) wird geprüft, ob die Manipulation der subjektiven und leistungsbezogenen mentalen Beanspruchung (sowie weiterer, studienspezifischer Konstrukte) durch die Versuchsbedingungen gelungen ist. In der MEA werden die Kriteriumsvariablen (z. B. NASA-RTLX-Werte) durch die Prädiktoren (Experimentalbedingungen) vorhergesagt. Auf eine Normierung oder Standardisierung der Daten wird aufgrund der ausschließlich nominalskalierten Prädiktoren verzichtet. Weit wichtiger ist in solchen Fällen die Wahl der richtigen Dummy- oder Effektkodierung, welche in den studienspezifischen Methodenkapiteln spezifiziert ist.

Bei den Studien der vorliegenden Arbeit handelt es sich um ein messwiederholtes Design.⁶ Die abhängige Datenstruktur der Messwiederholung wird durch die explizite Modellierung von fixen und zufälligen Effekten in der MEA berücksichtigt (siehe Nezlek, Schröder-Abé & Schütz, 2006). Im vorliegenden Fall werden die Experimentalbedingungen als feste Effekte und die Personen als zufällige Effekte spezifiziert. Feste Effekte bedeuten somit eine Effekt-Konstanz über Personen hinweg, wohingegen zufällige Effekte für interindividuelle Varianz stehen. Variieren können dabei sowohl die Steigungen als auch die Ordinatenabschnitte (engl.: Intercept), wobei in der vorliegenden Arbeit nur variierende Ordinatenabschnitte angenommen werden.

Bei der Durchführung einer MEA werden zuerst mittels der Maximum-Likelihood-Schätzung ein Modell und dessen Regressionsparameter ermittelt. Im zweiten Schritt wird der Gesamteffekt eines Prädiktors mit Hilfe von Modellvergleichen (χ^2 -Anpassungstest) geschachtelter Modelle bestimmt (Modellvergleich für jeden hinzukommenden Prädiktor). Von einem bedeutsamen

⁶ In solchen Fällen wäre auch eine Analyse mit der messwiederholten Varianzanalyse (ANOVA) möglich. Die MEA bietet allerdings zahlreiche Vorteile (z. B. keine Voraussetzung der Sphärizität und flexibler Umgang mit fehlenden Werten), weshalb Autoren wie Maurissen und Vidmar (2017) oder Quené und van den Bergh (2004) eine klare Empfehlung für die Verwendung der MEA geben. Wie die ANOVA gehört auch die MEA zum Allgemeinen Linearen Modell.

Effekt des Prädiktors ist dann auszugehen, wenn sich der Modellfit bei Hinzunahme des Prädiktors signifikant verbessert. Als Modellfitindizes werden die Devianz als $-2 \log\text{-likelihood}$ ($-2LL$) und das Akaike Informationskriterium (AIC) herangezogen. Zusätzliche Informationen bietet auch das marginale Pseudo- R^2 (R^2_{GLMM}), definiert als der Anteil der erklärten Varianz der festen Effekte an der Gesamtvarianz. Das am wenigsten restriktive, aber noch signifikante Modell ist das zu interpretierende Modell. Geprüft wird, wenn nicht anders vermerkt, auf einem Signifikanzniveau von $\alpha = .05$.

Möchte man valide Schlüsse aus den Ergebnissen der MEA ziehen, müssen gewisse Voraussetzungen erfüllt sein. Die Überprüfung der Voraussetzungen in einer MEA erfolgt nicht im Vorhinein, sondern im Nachhinein für das zu interpretierende Modell (Singer & Willett, 2003, S. 128). Die Voraussetzungen sind in Anhang A.4 beschrieben. Diese wurden im Rahmen der vorliegenden Arbeit geprüft. Auf eine explizite Darstellung der Ergebnisse wird aus Platzgründen verzichtet.

Die MEA wurde mit der Bibliothek lme4 (Version 1.1-21; Bates, Mächler, Bolker & Walker, 2015), MuMIn (Version 1.41.0; Bartoń, 2019) und lmerTest (Version 3.1-0; Kuznetsova, Brockhoff & Christensen, 2017) in der Statistiksoftware R (Version 3.5.2; R Core Team, 2018) berechnet sowie mit Hilfe der Bibliotheken ggplot2 (Version 3.1.1; Wickham, 2016) und papaja (Version 0.1.0.9842; Aust & Barth, 2018) grafisch dargestellt.

5.6.2 Statistische Methoden zur Ableitung psychophysiologischer Profile

Im Fokus der vorliegenden Arbeit steht die NPLS-Regression. Das Ziel dieser Analyse ist die Erstellung erster kardialer und elektrodermalen Profile mentaler Beanspruchung. Zur Veranschaulichung dieser Methode wird zunächst auf die Partial Least Squares Regression (PLS; dt.: Regression der partiellen kleinsten

$$\mathbf{X} = \mathbf{T} \mathbf{W}^T + \mathbf{E} \quad \mathbf{X} = \mathbf{T} \mathbf{W}^T + \mathbf{E}$$

$$\mathbf{Y} = \mathbf{U} \mathbf{Q}^T + \mathbf{F} \quad \mathbf{Y} = \mathbf{U} \mathbf{Q}^T + \mathbf{F}$$

$$\mathbf{U} = \mathbf{B} \cdot \mathbf{T}$$

Abbildung 5.4: Bei der Partial Least Squares Regression (PLS) werden die Prädiktormatrix \mathbf{X} und die Kriteriumsmatrix \mathbf{Y} in Faktorwerte und Faktorladungen zerlegt. Die Prämisse ist dabei die Maximierung der Kovarianz zwischen \mathbf{X} und \mathbf{Y} .

Quadrate) und im Anschluss auf die Multilinearität der Daten sowie auf die konkrete Umsetzung in der vorliegenden Arbeit eingegangen.

Partial Least Squares Regression

Die Partial Least Squares Regression (PLS) ist ein multivariates Verfahren, das nicht nur die Struktur bzw. Komponenten mehrerer Prädiktorvariablen aufdeckt (vergleichbar mit der PCA), sondern auch eine Regression der Kriteriumsvariable zulässt. Das Grundprinzip dieser Methode ist in Abbildung 5.4 dargestellt. Analog zur PCA wird eine Prädiktormatrix \mathbf{X} in die Matrizen \mathbf{T} und \mathbf{W}^T und eine Kriteriumsmatrix \mathbf{Y} in die Matrizen \mathbf{U} und \mathbf{Q}^T zerlegt. Diese enthalten jeweils die Faktorwerte (\mathbf{T} und \mathbf{U}) und -ladungen (\mathbf{W}^T und \mathbf{Q}^T). Zudem ergeben sich die Residualmatrizen \mathbf{E} und \mathbf{F} . Bei der PLS wird nun

versucht, die Residualmatrizen zu minimieren und gleichzeitig die Kovarianz zwischen \mathbf{X} und \mathbf{Y} durch Gleichung 5.4:

$$\mathbf{U} = \mathbf{B} \cdot \mathbf{T} \quad (5.4)$$

zu maximieren. Anders ausgedrückt, die PLS extrahiert die Komponenten derart, dass die Kovarianz zwischen den Faktorwerten von \mathbf{X} und \mathbf{Y} maximal ist. Letzteres unterscheidet die PLS wesentlich von der PCA. Das Regressionsgewicht \mathbf{B} gibt an, um wie viel die Faktorwerte \mathbf{U} ab- oder zunehmen, wenn die Faktorwerte \mathbf{T} um 1 steigen.

Im Vergleich zur multiplen Regression ergibt sich der Mehrwert einer PLS vor allem dann, sobald die Prädiktoren hoch miteinander korrelieren (Multikollinearität). Somit bietet diese Methode optimale Voraussetzungen, um diejenigen latenten Strukturen (Komponenten) aus der Vielzahl physiologischer Parameter zu extrahieren, welche mit dem Konstrukt der mentalen Beanspruchung zusammenhängen. Finden sich solche Komponenten, dann bilden die Faktorladungen der physiologischen Parameter die gesuchten physiologischen Profile der mentalen Beanspruchung ab. Aber auch die physiologischen Profile weiterer Konstrukte (z. B. Emotionsdimensionen: Erregung und Valenz) können mit dieser Methode abgeleitet und mit den physiologischen Profilen der mentalen Beanspruchung verglichen werden.

Multilineare Daten

Multilinearität bezieht sich auf Daten mit mehr als zwei Varianzquellen (Dimensionen). Dies bedeutet, dass die Strukturen (Komponenten) einer Datenmatrix \mathbf{X} , nicht wie in Abbildung 5.4 dargestellt, nur über zwei Dimensionen (z. B. verschiedene Variablen bei unterschiedlichen Personen), sondern über mindestens eine weitere Dimension (z. B. unterschiedliche Versuchsbedingungen) extrahiert werden können (vgl. Abbildung 5.5). Eine mehrdimensionale Datenstruktur findet sich auch in der vorliegenden Arbeit: Hier werden unterschiedliche

physiologische Parameter mehrerer Versuchspersonen über unterschiedliche Versuchsbedingungen erfasst. In der Literatur finden sich unterschiedliche Extraktionsmethoden und Algorithmen merdimensionaler Daten, die sich in ihren Ergebnissen und Interpretationen unterscheiden (Übersicht siehe Kroonenberg, 2008). Das in der NPLS-Regression angewandte Verfahren zur multilinearen Zerlegung wird analog zur sog. Parallelen Faktorenanalyse (PARAFAC) berechnet und interpretiert. Zum besseren Verständnis wird diese Methode kurz umrissen (Vertiefung siehe Kroonenberg, 2008; Bro, 1997).

Die PARAFAC ist im Grunde eine generalisierte PCA auf mehrdimensionalen Daten. Nichtsdestoweniger gibt es markante Unterschiede (siehe Bro, 1997). Diese Unterschiede sind in Anhang A.5 aufgeführt. Das Ergebnis der PARAFAC besteht, wie bei einer PCA, aus Faktorwerten und -ladungen. Insgesamt kommt bei einer PARAFAC pro zusätzliche Datendimension eine Faktorladungsmatrix⁷ hinzu (siehe Abbildung 5.5). Eine solche trilineare Zerlegung der Prädiktormatrix \mathbf{X} wird auch in der NPLS-Regression durchgeführt. Hier mit dem Unterschied, dass die Komponenten (analog der PLS) derart extrahiert werden, dass die Kovarianz zwischen den Faktorwerten \mathbf{T} und \mathbf{U} maximal ist (siehe Abbildung 5.6). Somit ist die NPLS-Regression eine Erweiterung der PLS auf multilineare Daten (hier: multilineare Prädiktormatrix).

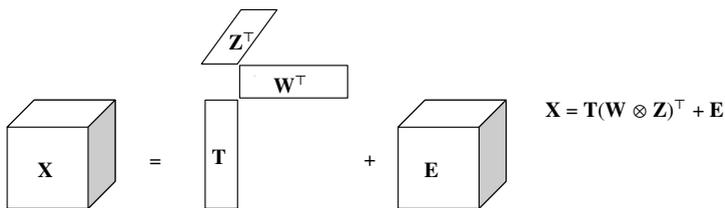


Abbildung 5.5: Die trilineare Zerlegung der PARAFAC.

⁷ Genau genommen wird bei multilinearen Analysen nicht zwischen Faktorwerten und -ladungen unterschieden.

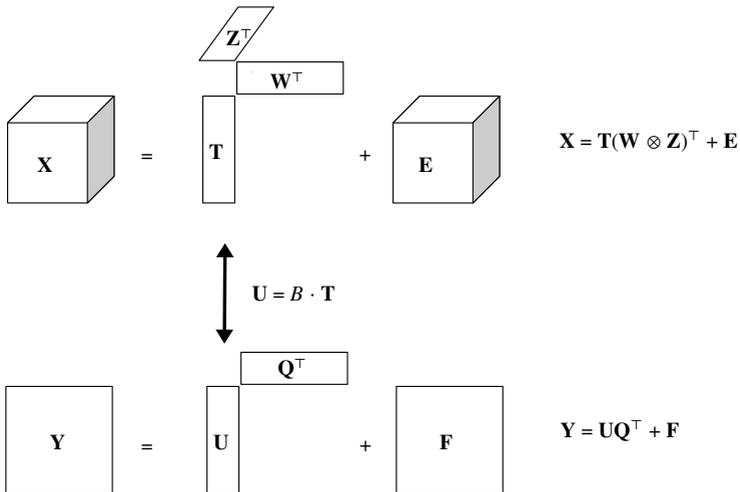


Abbildung 5.6: Die NPLS-Regression als Erweiterung der PLS auf mehrdimensionale Datenstrukturen.

Zentrierung und Normierung multilinearere Daten

Bei multilinearen Daten ist die Wahl der Zentrierungs- und Normierungsmethode entscheidend für die spätere Interpretation der Ergebnisse. Im Gegensatz zu den meisten psychometrischen Analysemethoden, bei denen die Zentrierung und Normierung der Daten innerhalb derselben Variable üblich ist, wird für mehrdimensionale Daten die Zentrierung zwischen den Einheiten einer Dimension (Faser-Zentrierung; engl.: fibre centering) und die Normierung innerhalb der Einheiten einer Dimension (Platten-Zentrierung; engl.: slab normalization) empfohlen (siehe Bro, 1997; Harshman & Lundy, 1994; Kroonenberg, 1983, S. 127–167; Kroonenberg, 2008, S. 109–142).⁸

⁸ Die Auswirkungen der Zentrierung können am besten veranschaulicht werden, indem man sich diese als Haupt- und Interaktionseffekte einer Varianzanalyse vor Augen führt: Bei einem dreidimensionalen Datensatz bestehend aus Person \times Situation \times Variable sind insgesamt drei Haupteffekte (Person, Situation und Variable), drei Zweifachinteraktionen (Person \times Situation,

In der vorliegenden Arbeit werden die Daten zwischen den Experimentalbedingungen zentriert und innerhalb der physiologischen Parameter sowie der Personen normiert. Durch das Zentrieren werden die Daten um den Haupteffekt der Person, der Parameter sowie der Interaktion $\text{Person} \times \text{Parameter}$ bereinigt. Dies hat den Vorteil, dass die Ergebnisse leichter zu interpretieren sind, da die Faktorklösungen auf eine mittlere Ausprägung der abhängigen Variable (z. B. der mentalen Beanspruchung gemessen im NASA-RTLX) referenziert werden können. Es ist allerdings zu beachten, dass durch die Zentrierung der Anteil der erklärten Varianz deutlich reduziert wird (Kroonenberg, 2008, S. 121; Kroonenberg, 1983, S. 146). Die Normierung innerhalb der physiologischen Parameter ist erforderlich, weil die einzelnen physiologischen Parameter ein unterschiedliches Variations- und Skalenspektrum aufweisen. Die Normierung innerhalb der Personen erlaubt es, personenspezifisches Reaktionsverhalten oder Antwortverhalten im Fragebogen (z. B. Tendenz zur Mitte) auszugleichen. Aufgrund der Zentrierung und Normierung der Daten sind die in der vorliegenden Arbeit berichteten Regressionsgewichte B der NPLS-Regression als standardisierte Regressionsgewichte (Wertebereich -1 bis $+1$) zu verstehen. Diese geben somit einen Hinweis auf die Größe des vorliegenden Zusammenhangs der extrahierten Komponenten.

Anwendung der NPLS-Regression auf die Messdaten der Studien

In der vorliegenden Arbeit wird die NPLS-Regression mit der N-way Toolbox (Version 3.20; Andersson & Bro, 2000) in Matlab (Version 9.5.0 R2018b, MATLAB, 2018) gerechnet. Dabei wird die zweidimensionale Kriteriumsma-

Variable \times Situation, Person \times Variable) und eine Dreifachinteraktion (Person \times Situation \times Variable) möglich. Die Zentrierung einer Dimension, bspw. zwischen den Situationen, bewirkt die Eliminierung derjenigen Haupt- und Interaktionseffekte, die keinen Bezug zur Situation haben. In diesem Beispiel also die Eliminierung des Haupteffekts der Person und der Variable sowie der Zweifachinteraktionen Person \times Variable. Dies hat den Vorteil, dass die nicht im Fokus stehenden Personen- und Variableneffekte die restlichen (im Fokus stehenden) Effekte nicht mehr überdecken (Vertiefung siehe Kroonenberg, 2008, S. 132–134).

trix (Person \times Versuchsbedingung) auf die dreidimensionale Prädiktormatrix (Person \times Versuchsbedingung \times physiologischer Parameter) regressiert. Die Komponenten werden dabei so extrahiert, dass die Kovarianz zwischen den Faktorwerten der physiologischen Prädiktormatrix und der Kriteriumsmatrix maximal ist. Die Maximierung erfolgt auf der Ebene der Bedingungen. Übertragen auf die Abbildung 5.6 bedeutet dies, dass \mathbf{U} das Kriterium (z. B. NASA-RTLX-Werte) und \mathbf{T} den Prädiktor (z. B. kardiale Aktivität) über die Bedingungen darstellt. \mathbf{Q}^T und \mathbf{W}^T repräsentieren in diesem Fall die Ausprägungen der Personen und \mathbf{Z} die unterschiedlichen Parameter (z. B. verschiedene kardiale Parameter). Für jede abhängige Variable der jeweiligen Studie wird ein eigenes NPLS-Modell separat für die kardialen und elektrodermalen Parameter berechnet. Fehlende Werte werden durch den Expectation-Maximization-Algorithmus geschätzt und ersetzt.

Als Anhaltspunkt der Modellgüte wird ein annäherndes Bestimmtheitsmaß $R_{n,s}^2$ berechnet. Dieses ergibt sich aus dem Anteil der durch die extrahierten Komponenten erklärten Quadratsumme und der zu erklärenden Quadratsumme. Im Fall der NPLS-Regression liegt das Bestimmtheitsmaß sowohl für die Prädiktor- als auch die Kriteriumsmatrix vor. Um die Gültigkeit des Modells zu belegen, wird eine Kreuzvalidierung durchgeführt. Bei der Leave-One-Out-Methode werden die Datensätze in Test- und Trainingsdaten unterteilt, indem die Daten einer Person p als Testdatensatz und die Daten der restlichen Personen als Trainingsdatensatz genutzt werden. Der Trainingsdatensatz dient zur Berechnung der Modellparameter und der anschließenden Vorhersage der Testdaten. Dadurch wird überprüft, ob die extrahierten Komponenten (aus dem Trainingsdatensatz) zu einer Reduktion der Quadratsumme der Residuen in den Prädiktoren des Testdatensatzes führen. Gleichzeitig werden die Komponenten der Prädiktoren im Testdatensatz dazu genutzt, um die Kriteriumsvariable (z. B. NASA-RTLX-Werte) der Testdaten vorherzusagen. Je geringer die Abweichung der vorhergesagten Werte zu den tatsächlichen Werten, desto besser gelingt es, die Person p vorherzusagen. Insgesamt werden im Rahmen der Kreuzvalidierung so viele Trainings- und Testdatensätzen generiert

und NPLS-Regressionen berechnet, wie es Versuchspersonen im Datensatz gibt. Diese dienen wiederum als Grundlage zur Berechnung kreuzvalidierter (val) Bestimmtheitsmaße ($\text{val}\Delta R_{\text{PLS}}^2$) für jede zusätzlich aufgenommene Komponente⁹.

Wie viele und welche Komponenten am Ende aufgenommen werden, wird in der vorliegenden Arbeit anhand von drei Voraussetzungen entschieden. Die ersten beiden Voraussetzungen beziehen sich auf den Zuwachs im kreuzvalidierten Bestimmtheitsmaß $\text{val}\Delta R_{\text{PLS}}^2$. Ob ein Zuwachs bedeutsam ist, wird empirisch über das Konfidenzintervall zufälliger ΔR_{PLS}^2 , den sog. $\text{rand}\Delta R_{\text{PLS}}^2$, ermittelt. Zu diesem Zweck werden 500 Zufalldatensätze – mit der gleichen Größe und Dimensionalität wie der zu untersuchende Datensatz – generiert und über die NPLS-Regression kreuzvalidiert. Liegt das validierte ΔR_{PLS}^2 ($\text{val}\Delta R_{\text{PLS}}^2$) bei Hinzunahme einer Komponente in das Modell außerhalb des 99 %-Konfidenzintervalls (99 % *CI*) der $\text{rand}\Delta R_{\text{PLS}}^2$ -Verteilung¹⁰, dann gilt dieser Zuwachs als bedeutsam. Ein bedeutsamer Zuwachs im $\text{val}\Delta R_{\text{PLS}}^2$ muss sowohl für die Prädiktormatrix (1. Voraussetzung) als auch für die Kriteriumsmatrix (2. Voraussetzung) vorliegen. Die dritte Voraussetzung prüft, ob die Vorhersage des Kriteriums durch die Hinzunahme einer Komponente signifikant verbessert wird. Zur Beurteilung der Vorhersagegüte wird die Quadratsumme der Residuen aus der Kreuzvalidierung (valSQR) herangezogen. Das valSQR wird mit Hilfe eines nicht parametrischen Verfahrens auf eine signifikante Reduktion bei Hinzunahme einer Komponente ins Modell geprüft. Randall (1995) empfiehlt, für die PLS-Regression den Permutationstests gemäß van der Voet (1994) zu verwenden. Geprüft wird einseitig mit einem Signifikanzniveau von 5 %.

Es ist zu beachten, dass im Originalmodell alle drei Dimensionen (Person \times Versuchsbedingung \times physiologischer Parameter) und im kreuzvalidierten Modell

⁹ Obwohl das Bestimmtheitsmaß als Anteil erklärter Varianz interpretiert wird, kann der Koeffizient auch negative Werte annehmen. Im Fall der Kreuzvalidierung ist das dann der Fall, wenn die vorhergesagten Kriteriumswerte des Testdatensatzes so stark von den tatsächlichen Werten abweichen, dass die vorhergesagte Fehlerquadratsumme größer ist als die zu erklärende Fehlerquadratsumme. In der vorliegenden Arbeit wird darauf verzichtet, das R_{PLS}^2 in solchen Fällen auf Null zu setzen.

¹⁰ Die $\text{rand}\Delta R_{\text{PLS}}^2$ -Verteilung folgt einer Normalverteilung.

nur zwei Dimensionen (Versuchsbedingung \times physiologischer Parameter) zur Berechnung der vorhergesagten Werte genutzt werden können. Somit bietet die hier verwendete Kreuzvalidierung eine Angabe darüber, wie valide das Modell die mentale Beanspruchung von unbekanntem Personen vorhersagen kann.¹¹ Im Gegensatz dazu gibt die Vorhersagegüte des Originalmodells Auskunft darüber, wie gut eine bereits bekannte Person vorhergesagt werden kann.¹² Da in der Kreuzvalidierung weniger Vorhersageparameter genutzt werden können, ist mit einer deutlichen Reduzierung des erklärten Varianzanteils für die kreuzvalidierte Vorhersage zu rechnen. Das Ausmaß der Reduktion zeigt gleichzeitig auf, wie stark die Vorhersage durch die personenspezifische Varianz beeinflusst wird.

Interpretation der Faktorladungen

Die physiologischen Profile werden durch die Faktorladungen a (Wertebereich: -1 bis 1) der Parameterdimension abgebildet. Gleichzeitig bieten die Faktorladungen der physiologischen Parameter auf den Komponenten wertvolle Informationen darüber, wie hoch die Parameter mit der Komponente zusammenhängen und diese somit repräsentieren. Je näher die Faktorladung an -1 oder 1 liegt, desto höher ist deren Zusammenhang mit dem physiologischen Profil. Wie hoch eine Faktorladung sein sollte, um als bedeutsam interpretiert zu werden, kann statistisch nicht endgültig bestimmt werden. So schreibt Peterson (2000, S. 263): „[T]here is no consensus as to what constitutes a “high” or “low” factor loading, both in general and in specific research conditions“. Gemäß der Metaanalyse von Peterson (2000) werden Grenzwerte im Bereich von $|\cdot 20|$ bis $|\cdot 70|$ in der Forschung angewandt. Merenda (1997) nennt einen Richtwert

¹¹ Möglich wäre auch eine Kreuzvalidierung über die Versuchsbedingungen hinweg. Diese Art der Kreuzvalidierung würde allerdings die Personenunabhängigkeit der extrahierten Profile vernachlässigen (siehe Kapitel 9.3).

¹² Im Gegensatz zum kreuzvalidierten Modell ist es im Originalmodell möglich, die vorhergesagten Werte mit der personenspezifischen Ausprägung der physiologischen Aktivität und des Antwortverhaltens (z. B. im Fragebogen) zu gewichten.

ab $|a| \geq .30$, um eine Variable zu einer Komponente gehörig zu interpretieren. Hair, Tatham, Anderson und Black (1998, S. 111) gehen ebenfalls von einem Minimum von $|.30|$ und einer praktischen Signifikanz ab $|a| \geq .50$ aus. Für mehrdimensionale Daten sind allerdings bislang keine Richtwerte aus der Literatur bekannt. Als Orientierung können lediglich mehrdimensionale Analysebeispiele aus Kroonenberg (2008) herangezogen werden. Hierbei wird $|a| > .35$ als hoch und $.25 \leq |a| \leq .35$ als mittel hoch interpretiert (siehe Kroonenberg, 2008, S. 252). Für die vorliegende Arbeit wird $|a| \geq .30$ als eine bedeutsame und $.20 \leq |a| < .30$ als eine marginale Ladung des Parameters auf der Komponente interpretiert.

Zur Bewertung der Spezifität (Diagnostizität) physiologischer Profile wird der Kongruenzkoeffizient C nach Tucker (1951, zitiert nach Lorenzo-Seva & Ten Berge, 2006) verwendet. Der Koeffizient gibt eine Auskunft darüber, wie ähnlich zwei extrahierte Komponenten unterschiedlicher Stichproben sind. Der Wert des Kongruenzkoeffizienten liegt zwischen -1 und $+1$. Eine statistische Absicherung des Kongruenzkoeffizienten existiert allerdings nicht. Nach Lorenzo-Seva und Ten Berge (2006) gelten Komponenten mit einem C zwischen $.85$ bis $.94$ als ähnlich und mit einem C über $.95$ als identisch. Welche Variablen genau zu der Ähnlichkeit bzw. Unterschiedlichkeit der Komponenten beitragen, kann anhand der Abweichungen (Δ) der Faktorladungen $|a_{\Delta}|$ beurteilt werden. Für die Bewertung liegen ebenfalls nur Richtwerte aus der Literatur vor. Demnach muss zwischen den Parameter-Faktorladungen (verschiedener Komponenten) eine Differenz von mind. $.20$ bestehen, um als unterschiedlich zu gelten (Gaskin, 2012, zitiert nach Ibrahim, Shiratuddin & Wong, 2015). Allerdings gelten auch diese Richtwerte bisher nur für zweidimensionale Ansätze. Für die vorliegende Arbeit wird $|a_{\Delta}| \geq .20$ als bedeutsam interpretiert. Zudem werden Unterschiede mit $.15 \leq |a_{\Delta}| < .20$ als marginal bedeutsam und mit $|a_{\Delta}| < .15$ als nicht bedeutsam interpretiert.

Ermittlung der Stichprobengröße

Es gehört zur guten Praxis eine Analyse der Teststärke¹³ vor der Datenerhebung durchzuführen, um eine geeignete Stichprobengröße für den Test zu ermitteln. Allerdings gibt es bisher keine methodischen Ansätze, um eine Teststärke für multilineare Daten bzw. die NPLS-Regression zu berechnen. Aus diesem Grund wird in den folgenden Studien I, II und III eine Stichprobengröße von ca. 50 Personen angestrebt.

Für die MEA wird die Teststärke im Nachgang an die Analyse bestimmt. Eine Vorabbestimmung der Stichprobengröße durch eine Teststärken-Analyse der MEA wäre irreführend, da diese lediglich zur Prüfung der Manipulationskontrolle und nicht der Hauptfragestellungen dient. Zur Berechnung der Teststärke wurde die Bibliothek `simr` (Version 1.0.5; Green & MacLeod, 2016) in der Statistiksoftware R (Version 3.5.2; R Core Team, 2018) genutzt. Dabei werden 1000 Monte-Carlo-Simulationen berechnet und mit dem z-Test ($\alpha = .05$) geprüft.

¹³ Wahrscheinlichkeit einer richtigen Entscheidung zugunsten der Alternativhypothese.

6 Studie I

In diesem Kapitel werden der studienspezifische Hintergrund der Studie I in Abschnitt 6.1 und die studienspezifischen Methoden in Abschnitt 6.2 beschrieben. Im Anschluss werden in Abschnitt 6.3 die Ergebnisse für die Kontrollfragestellung KF1 und die Forschungsfragen F1 bis F5 berichtet. Zum Schluss erfolgt in Abschnitt 6.4 eine Zusammenfassung und Interpretation aller Ergebnisse sowie die Beleuchtung der Limitationen der Studie I.

6.1 Forschungsziel der Studie I

Die Studie I hat zum Ziel, kardiale und elektrodermale Profile mentaler Beanspruchung zu bestimmen. Dieses Ziel wurde bereits ausführlich in Kapitel 4.3 beschrieben und bezieht sich auf die Forschungsfragen F1 und F2. Des Weiteren werden in Studie I die Forschungsfragen F3 und F4 adressiert. Diese prüfen, ob kardiale und elektrodermale Profile emotionsbezogener Beanspruchung extrahiert werden können. Die emotionale Beanspruchung wird dabei durch die Emotionsdimensionen Erregung und Valenz (siehe Abschnitt 2.1.3) repräsentiert. Zum Schluss gilt es die kardialen und elektrodermalen Profile mentaler und emotionaler Beanspruchung im Sinne der Spezifität voneinander abzugrenzen (Forschungsfrage F5). Es ist anzumerken, dass der Versuch der Studie I im Rahmen des Projekts Kontextsensitive Assistenz im aufmerksamen OP (KonsensOP) durchgeführt wurde (siehe Wolf, Martinez, Roitberg, Stiefelhagen & Deml, 2018).

Dass eine Abgrenzung mentaler und emotionaler Beanspruchung auf der Wirkungsebene nicht trivial ist, wurde bereits im Unterkapitel 2.1.3 beschrieben. Folglich wird gemäß Manzey (1998) und Schlick et al. (2010) die Differenzierung zunächst auf Ebene der Ursache (Belastung) vorgenommen. Dieser Ansatz liefert die Basis für eine experimentelle Operationalisierung der Beanspruchung. Dabei kann die mentale Beanspruchung durch eine informatorische Belastung (z. B. Aufgabenschwierigkeit) ausgelöst werden, während die emotionale Beanspruchung durch ausführungsspezifische (z. B. Ausführungserfolg bzw. Misserfolg) oder durch rein affektive Belastungen (z. B. emotionsauslösende Stimuli) hervorgerufen wird. Im Idealfall würde sich diese Trennung mentaler und emotionaler Ursachen auf der psychologischen Wirkungsebene widerspiegeln. Erfasst wird die psychologische Wirkungsebene mit Hilfe von leistungsbezogenen und subjektiven Maßen. Ob auf der physiologischen Wirkungsebene eine Trennung mentaler und emotionaler Beanspruchung möglich ist, ist zu prüfen. Zu diesem Zweck werden kardiale und elektrodermale Profile der subjektiven Erregung und Valenz extrahiert und mit den physiologischen Profilen der subjektiven und leistungsbezogenen mentalen Beanspruchung verglichen.

6.2 Methode

In diesem Abschnitt werden die studienspezifischen Methoden der Studie I dargestellt. Diese beziehen sich auf die Operationalisierung der studienspezifischen Konstrukte (Abschnitt 6.2.1), das Versuchsdesign (Abschnitt 6.2.2), die Versuchsdurchführung (Abschnitt 6.2.3), die Kontrollvariablen (Abschnitt 6.2.4), die Manipulationskontrolle (Abschnitt 6.2.5) und die Beschreibung der Stichprobe (Abschnitt 6.2.6).

6.2.1 Operationalisierung der mentalen und emotionalen Beanspruchung

Die Operationalisierung der mentalen und emotionalen Beanspruchung bezieht sich zum einen auf die experimentelle Induktion in den Versuchsbedingungen und zum anderen auf die Messung der mentalen und emotionalen Beanspruchung. Diese wird in den folgenden Abschnitten vertieft.

Experimentelle Induktion mentaler Beanspruchung

Die mentale Beanspruchung wurde durch informatorische Belastungsfaktoren in Form einer Leistungsaufgabe operationalisiert. Die Leistungsaufgabe setzte sich aus einer Primäraufgabe und einer Sekundäraufgabe zusammen. In der Primäraufgabe wurde die Versuchsperson dazu aufgefordert, einen sich zufällig über den grauen Computerbildschirm bewegenden blauen Ball mit dem Maus-

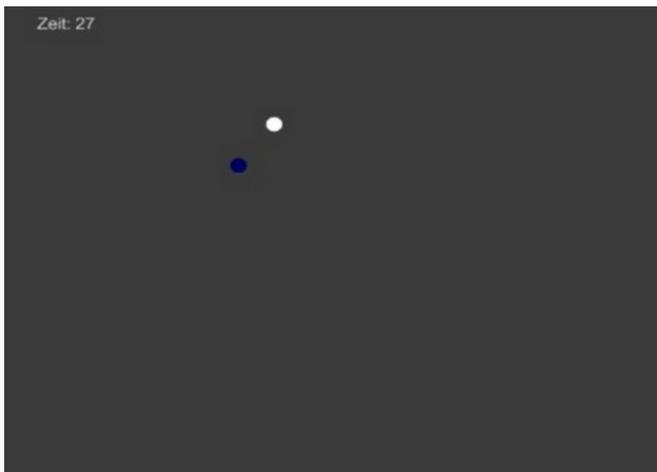


Abbildung 6.1: Balltrackingaufgabe der Studie I mit grauem Spielfeld. Das Ziel der Aufgabe ist es, den Zielball (blauer Kreis) mit der dem Mauszeiger (weißer Kreis) zu verfolgen und zu treffen. Die verbleibende Zeit wurde oben links im Spielfeld angezeigt.

zeiger zu verfolgen (siehe Abbildung 6.1). Um den Versuchspersonen einen Treffer zu signalisieren, leuchtete der Zielball hellblau auf, wenn der Mauszeiger den Ball berührte. Das Aufgabenziel bestand darin, den Ball so häufig wie möglich zu treffen. Insgesamt wurden für die Balltrackingaufgabe zwei Schwierigkeitsstufen a 60 s umgesetzt. In der leichten Schwierigkeitsstufe S1 war die Geschwindigkeit des Balls gering. In der schweren Schwierigkeitsstufe S2 war die Geschwindigkeit des Balls um den Faktor 3 größer als in S1. Zudem wurde in S2 die Mausbewegung umgekehrt, sodass die Bewegungsrichtung der Computermaus und des Mauszeigers auf dem Bildschirm sowohl in der vertikalen als auch in der horizontalen Richtung invertiert war. Alle Versuchspersonen wurden instruiert, die Primäraufgabe (Balltrackingaufgabe) bei der Bearbeitung zu priorisieren. Die bereits verstrichene Zeit für die Aufgabe wurde oben links im Bildschirm angezeigt.

Die Sekundäraufgabe wurde in Form einer auditiven Subtraktionsaufgabe dargeboten. Diese wurde parallel zur Balltrackingaufgabe für 60 s bearbeitet. In dieser Zeit wurden den Versuchspersonen acht Subtraktionsaufgaben auditiv vorgegeben. Alle Subtraktionsaufgaben bestanden aus drei Operanden (z. B.: $56 - 13 - 2$). Die Lösungseingabe erfolgte mit den Maustasten. Mit diesem Vorgehen wurde eine gleichzeitige Bearbeitung der Balltracking- und Subtraktionsaufgabe gewährleistet. Ergab die Lösung der Subtraktionsaufgabe die Zahl 7, dann musste die linke Maustaste gedrückt werden. Für die Lösung 9, musste die rechte Maustaste und für jede andere Lösung (nicht 7 und nicht 9) musste die mittlere Maustaste (Mausrad) gedrückt werden. Pro Experimentalbedingung hatten drei von den insgesamt acht Subtraktionsaufgaben die Zahl 7 oder 9 als Lösung. Für die Eingabe der Lösung hatten die Versuchspersonen 3 s Zeit, bevor die nächste Subtraktionsaufgabe präsentiert wurde. Die Versuchsperson erhielt keine Rückmeldung zur Korrektheit der Lösung.

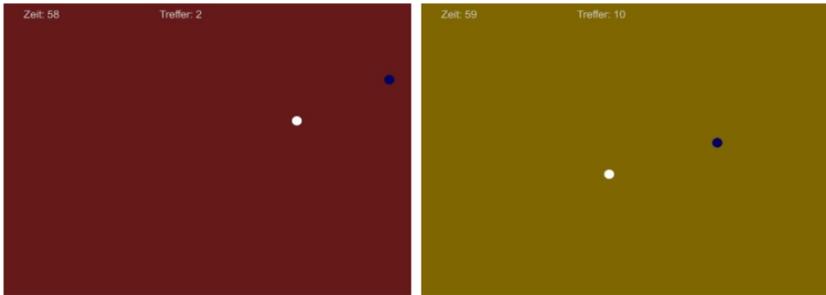


Abbildung 6.2: Balltrackingaufgabe mit Leistungsfeedback. Der Punktestand (Treffer) wird oben in der Bildschirmmitte angezeigt. Links: Negatives Leistungsfeedback mit rotem Bildschirmhintergrund. Rechts: Positives Leistungsfeedback mit gelbem Bildschirmhintergrund.

Experimentelle Induktion von Erregung und Valenz durch ausführungsspezifische Belastung

Die Operationalisierung der emotionalen Beanspruchung erfolgte zunächst über ausführungsspezifische Belastungsfaktoren in Form eines fingierten Leistungsfeedbacks. Dabei konnte die ausführungsspezifische Belastungsinduktion parallel zur mentalen Beanspruchungsinduktion (Balltrackingaufgabe) durchgeführt werden. Hierfür diente die Studie von J. Kim, André, Rehm, Vogt und Wagner (2005) als Orientierung. Die Valenz wurde dabei durch positives oder negatives Feedback variiert, die Erregung durch die Interaktion des Feedbacks mit der Aufgabenschwierigkeit. Demnach wäre bei negativem Feedback gepaart mit hoher Aufgabenschwierigkeit eine geringe Erregung (erwartete schlechte Leistung) und gepaart mit niedriger Aufgabenschwierigkeit eine hohe Erregung zu erwarten (unerwartete schlechte Leistung). Umgekehrt wäre bei positivem Feedback eine geringe Erregung zu erwarten, wenn die Aufgabe leicht ist (erwartete gute Leistung), jedoch eine hohe Erregung zu erwarten, wenn die Aufgabe schwer ist (unerwartete gute Leistung).

Im vorliegenden Experiment wurde der Punktestand – also die relative Anzahl der Treffer (Quotient aus Anzahl tatsächlicher Treffer und Anzahl möglicher

Gesamttreffer) – in der Balltrackingaufgabe systematisch variiert und den Versuchspersonen oben in der Bildschirmmitte angezeigt (siehe Abbildung 6.2). Jedes Mal, wenn der Mauszeiger den Zielball berührte, erhöhte sich der Punktestand. Dieser bewegte sich im Wertebereich von 0 (kein Ballkontakt) bis 100 (ununterbrochener Ballkontakt). Zur Steigerung der emotionalen Involviertheit wurde den Versuchspersonen ein monetärer Gewinn von 2 Euro in Aussicht gestellt. Der Gewinn konnte erzielt werden, sobald ein vorab definierter Punktestand in der Aufgabe erreicht wurde. Der zu erreichende Punktestand richtete sich nach der Schwierigkeitsstufe der Balltrackingaufgabe. Bei der einfachen Balltrackingaufgabe S1 waren 10 Punkte und bei der schweren Balltrackingaufgabe S2 30 Punkte zu erreichen. Die Farbe des Bildschirmhintergrunds passte sich an den Punktestand an: Bei gelber Hintergrundfarbe wurde ein Überschreiten des Zielpunktestands und bei roter Hintergrundfarbe ein Unterschreiten des Zielpunktestands signalisiert. Mit diesem Vorgehen wurde sichergestellt, dass die Versuchspersonen sich des Gewinns bzw. Verlusts bewusst waren.

Tabelle 6.1: Die Experimentalbedingungen repräsentieren die vier Kombinationen aus Leistungsfeedback und Aufgabenschwierigkeit

	Aufgabenschwierigkeit	Leistungsfeedback
S1+	niedrig	positiv
S2+	hoch	positiv
S1-	niedrig	negativ
S2-	hoch	negativ

Insgesamt wurden vier Experimentalbedingungen konstruiert. Diese sind in Tabelle 6.1 dargestellt. Die vier Bedingungen bilden eine Mischung aus mentaler und emotionaler Beanspruchungsinduktion. Um das Erreichen bzw. Unterschreiten des Punktestands für das Feedback systematisch variieren zu können, wurde der Punktestand sowie die Mauszeigerbewegung durch die Experimentalssoftware in eine vorgegebene Richtung manipuliert. So wurde ein höherer

oder niedrigerer Punktestand als der tatsächliche angezeigt. Zudem wurde die Mauszeigebewegung bei der Verfolgung des Zielballs erschwert oder durch eine Automatisierung erleichtert. Die Versuchspersonen wurden nicht über diese Leistungsmanipulation informiert. Am Ende gewann jede Versuchsperson zwei Mal (insgesamt 4 Euro).

Experimentelle Induktion von Erregung und Valenz durch rein affektive Belastung

Die rein affektiven Belastungsfaktoren zur experimentellen Induktion von Erregung und Valenz wurden mit Hilfe des International Affective Picture Systems (IAPS; Lang et al., 2008) umgesetzt. Das IAPS ist eine Sammlung von Photographien, welche auf den Emotionsdimensionen Erregung, Valenz und Dominanz normiert ist. Das IAPS ist ein standardisiertes Instrument zur Emotionsinduktion und findet weltweit Anwendung. Insgesamt wurden 40 Bilder aus dem IAPS ausgesucht und in vier Bildersets anhand der Ausprägung auf den normierten Skalen Erregung und Valenz gruppiert. Das erste Bilderset EnV+ bestand aus Bildern mit einer niedrigen Ausprägung in der Erregung und einer positiven Valenz, das zweite Bilderset EhV+ aus einer hohen Erregung und positiver Valenz, das dritte Bilderset EnV- aus geringer Erregung und negativer Valenz sowie das vierte Bilderset EhV- aus hoher Erregung und negativer Valenz. Jedes Bilderset beinhaltete 10 Bilder. Tabelle 6.2 zeigt für jedes Bilderset die mittlere (normierte) Ausprägung der Emotionsdimension sowie deren Standardabweichung. Die Präsentation jedes Bildersets dauerte 60 s, wobei jedes Bild für 6 s gezeigt wurde. Die Reihenfolge der Bilder innerhalb eines Bildersets wurde über Versuchspersonen hinweg konstant gehalten.

Messung der mentalen Beanspruchung

In Studie I wurde die subjektive mentale Beanspruchung mit dem NASA-RTLX (ohne die Skala körperliche Anforderung) gemessen (siehe Unterkapitel 5.4.1).

Tabelle 6.2: Mittelwert und Standardabweichung für jedes Bilderset aus Studie I. Die statistischen Kennzahlen wurden anhand des normierten Datensatzes aus Lang et al. (2008) berechnet

	Erregung		Valenz	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
EnV+	3.15	1.94	6.47	1.94
EhV+	6.68	1.82	6.96	1.58
EnV-	4.08	2.01	2.96	1.56
EhV-	6.89	1.92	2.08	1.35

Zudem wurde die leistungsbezogene mentale Beanspruchung als Effektivitätsmaß der Sekundäraufgabe abgebildet (siehe Kapitel 5.5).¹ Berechnet wurde hierfür die Anzahl richtiger Antworten in der auditiven Subtraktionsaufgabe in Prozent pro Experimentalbedingung (Trefferquote). Die NASA-RTLX-Werte und die Trefferquote der Sekundäraufgabe bilden jeweils die Kriterien in der NPLS-Regression ab.

Messung der emotionalen Beanspruchung

Die emotionale Beanspruchung wurde als subjektives Maß mit Hilfe eines Fragebogens, des Self-Assessment Manikins (SAM), erfasst. Lang (1980 zitiert nach Bradley & Lang, 1994) entwickelte ein sprachfreies Instrument zur Erfassung der Emotionsdimensionen Erregung, Valenz und Dominanz. Das SAM bildet mit Hilfe stilisierter Figuren, den sog. Manikins, die entsprechenden emotionalen Erregungen ab. Pro Dimension werden fünf Manikins auf einer fünfstufigen Skala von positiv bis negativ (Valenz), von erregt bis ruhig

¹ Die Primäraufgabenleistung (Balltrackingaufgabe) war aufgrund der experimentellen Manipulation der Aufgabenausführung verfälscht und wurde folglich nicht für die weitere Analyse berücksichtigt.

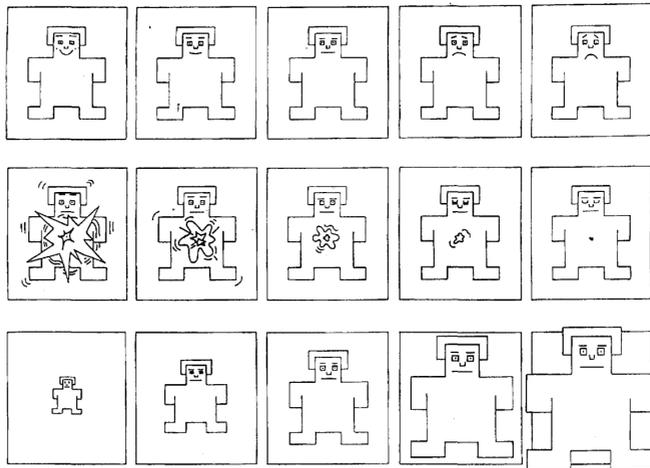


Abbildung 6.3: Dargestellt ist der SAM bestehend aus den Skalen Valenz (oben), Erregung (Mitte) und Dominanz (unten). Grafik aus Bradley und Lang (1994).

(Erregung) und von unterlegen bis überlegen (Dominanz) dargeboten (siehe Abbildung 6.3). Berichtet wird eine Überlegenheit des SAM im Vergleich zu sprachgebundenen Maßen (vgl. Bradley & Lang, 1994). Das Instrument wurde bereits in zahlreichen empirischen Studien erfolgreich eingesetzt (z. B. Betella & Verschure, 2016; Bradley, Codispoti, Cuthbert & Lang, 2001). In der vorliegenden Arbeit wurde der SAM-Fragebogen nach jeder Experimentalbedingung den Versuchspersonen vorgelegt. Für die weitere Analyse wurden nur die ersten beiden Dimensionen Erregung und Valenz berücksichtigt. Die Erregungs- und Valenz-Werte des SAM bilden jeweils die Kriterien in der NPLS-Regression ab.

6.2.2 Versuchsdesign

Das Experiment der Studie I setzt sich aus 10 Experimentalbedingungen zusammen. Diese Experimentalbedingungen werden in drei Blöcke unterteilt. Zu Block 1 gehören die Experimentalbedingungen S1 und S2. Dieser Block repräsentiert die Bedingungen mit reiner mentaler Beanspruchungsinduktion

(keine Emotionsinduktion). In Block 2 wird sowohl die mentale als auch die emotionale Beanspruchung induziert. Zu diesem zählen die Bedingungen S1+, S2+, S1- und S2-. Block 3 repräsentiert den Block der rein emotionalen Beanspruchung mit den Bedingungen EnV+, EhV+, EnV- und EhV-. Das Versuchsdesign der Blöcke 1 bis 3 ist in Abbildung 6.4 als Ablaufdiagramm dargestellt.

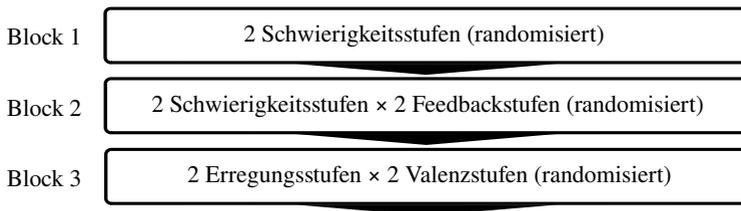


Abbildung 6.4: Versuchsdesign mit den Blöcken 1 bis 3. Innerhalb der Blöcke sind die Bedingungen randomisiert.

Alle Teilnehmerinnen und Teilnehmer durchliefen die Blöcke in einer festen, aufsteigenden Reihenfolge. Da es sich um ein messwiederholtes Design handelte, durchlief jede Versuchsperson alle Experimentalbedingungen. Der Vorteil eines messwiederholten Designs ist die Berücksichtigung von interindividuellen Unterschieden. Nachteile entstehen durch Sequenzeffekte, wie Ermüdungs- oder Lernprozesse, welche die Daten systematisch verzerren können. Zur Vermeidung von Sequenzeffekten wurden die Experimentalbedingungen innerhalb der Blöcke randomisiert vorgegeben. Da auf ein vollständiges Ausbalancieren aller Experimentalbedingungen (über Blöcke hinweg) verzichtet wurde, handelt es sich um ein quasi-experimentelles Design.²

Insgesamt ergeben sich für die Studie I vier Kriterien: die NASA-RTLX-Werte, die Erregungs-Werte, die Valenz-Werte und die Trefferquote. Jedes Kriterium bildet jeweils eine Kriteriumsmatrix im Rahmen der NPLS-Regression ab.

² Auf eine vollständige Ausbalancierung wurde verzichtet, da die Versuchsblöcke in sich geschlossene Operationalisierungskonzepte darstellten.

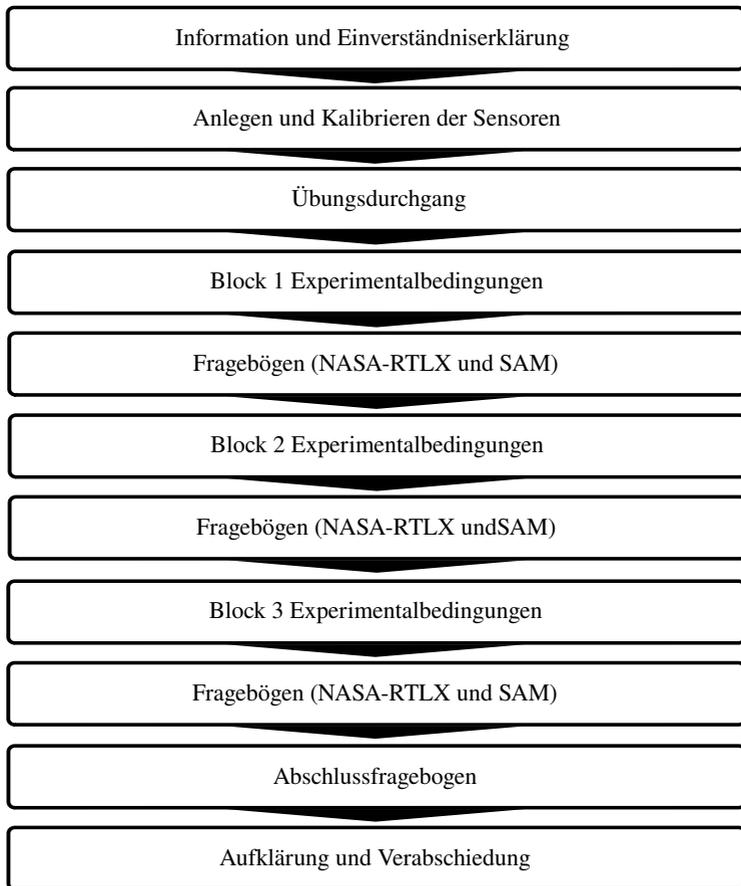


Abbildung 6.5: Schematische Darstellung des Versuchsablaufs aus Studie I.

Die Prädiktormatrix besteht jeweils aus den kardialen und elektrodermalen Parametern.

6.2.3 Versuchsablauf

Das Experiment der Studie I fand im Labor statt. Vor Beginn des Experiments wurden alle Versuchspersonen mündlich und schriftlich über den Zweck der Studie sowie über die Erfassung der Daten informiert. Danach wurde eine Einverständniserklärung zur Teilnahme am Experiment unterzeichnet. Im Anschluss wurden die Elektroden zur Messung der kardialen und elektrodermalen Aktivität von der Versuchsleitung angebracht und mit der Messsensorik verbunden. Vor dem Start des Experiments wurden alle Sensoren auf ihre Funktionstüchtigkeit sowie Signalplausibilität von der Versuchsleitung geprüft. Daran anschließend instruierte die Versuchsleitung die Versuchsperson über die Aufgaben sowie die Fragebögen (NASA-RTLX und SAM). Das Experiment fand automatisiert am Versuchsrechner statt, um Versuchsleitereffekte zu vermeiden. Als Experimentalsoftware wurde die Blender Game Engine Version 2.76 (Blender Online Community, 2015) verwendet. Eine schematische Darstellung des Versuchsablaufs ist in Abbildung 6.5 zu sehen

Am Anfang durchliefen die Versuchspersonen einen Übungsdurchgang, um sich an die Aufgaben (Balltrackingaufgabe und Subtraktionsaufgabe) zu gewöhnen. Als nächstes wurden die Experimentalbedingungen aus Block 1 bearbeitet. Danach folgte Block 2. Hierbei wurden die Versuchspersonen im Glauben gelassen, dass das Feedback ihre tatsächliche Leistung wiedergibt. Den Versuchspersonen wurde mitgeteilt, dass sie in Block 2 um einen Geldtopf von 8 Euro spielen. Beim Übertreffen des Punktestands in einer Aufgabe können sie 2 Euro aus dem Geldtopf behalten, beim Untertreffen verlieren sie 2 Euro aus dem Geldtopf. Durch die Feedbackmanipulation gewann und verlor jede Versuchsperson jeweils in zwei Aufgaben und ging somit mit einem Zusatzgewinn von 4 Euro raus. Zuletzt wurden die Versuchspersonen in Block 3 dazu aufgefordert, verschiedene Bildersets anzuschauen.

Die Bearbeitung jeder Experimentalbedingung dauerte 60 s. Nach jeder Experimentalbedingung beantworteten die Versuchspersonen den NASA-RTLX und anschließend den SAM. Nach Beendigung der drei Versuchsblöcke beant-

worteten die Versuchspersonen einen Abschlussfragebogen (Kontrollvariablen). Zudem folgte eine Aufklärung und Befragung der Versuchspersonen über die Manipulation der Leistung in Block 2. Nach dem Ablegen aller Sensoren wurden die Versuchspersonen für ihre Teilnahme vergütet und verabschiedet. Der gesamte Versuch dauerte ca. 90 min.

6.2.4 Abschlussbefragung und Kontrollvariablen

Der Abschlussfragebogen beinhaltet die relevanten soziodemografischen Daten der Versuchspersonen. Hierzu gehören die Variablen Alter, Geschlecht und beruflicher Hintergrund.

Als Kontrollvariablen wurden die Temperatur und die relative Luftfeuchtigkeit zu Anfang und am Ende des Experiments gemessen. Des Weiteren wurden am Ende des Experiments drei Kontrollfragen zur Feedbackmanipulation in Block 2 erfasst. Die folgenden Frageitems konnten auf einer 5-stufigen Antwortskala von „nein gar nicht“ bis „ja genau“ beantwortet werden:

1. Die Trefferquote entsprach meiner tatsächlichen Leistung.
2. Die Trefferquote empfand ich als glaubhaft.
2. Die Trefferquote empfand ich als gerecht.

Die drei Items wurden gemittelt und als Glaubwürdigkeit des Feedbacks interpretiert. Alle drei Items zeigen eine bedeutsame Interkorrelation ($r = .34$ bis $r = .74$, $p < .02$) und eine zufriedenstellende interne Konsistenz (Cronbach's Alpha) von $.79$.

6.2.5 Manipulationskontrolle

Zur Beantwortung der Forschungsfragen F1 bis F3 ist es erforderlich, dass die vier Kriteriumsvariablen (NASA-RTLX, Erregung, Valenz und Trefferquote)

über die Experimentalbedingungen hinweg erwartungskonform variieren. Aus diesem Grund wird die experimentelle Manipulation statistisch geprüft. Zu diesem Zweck werden Hypothesen (H) innerhalb der Kontrollfragestellung KF1 formuliert.

Hypothesen der Manipulationskontrolle

Zunächst werden die Effekte auf die subjektive und leistungsbezogene mentale Beanspruchung geprüft. Es ist zu erwarten, dass die subjektive (NASA-RTLX-Werte) und die leistungsbezogene (Trefferquote) mentale Beanspruchung mit zunehmender Aufgabenschwierigkeit sowohl in Block 1 als auch in Block 2 steigt.

H1.1 Je höher die Aufgabenschwierigkeit, desto höher sind die NASA-RTLX-Werte.

H1.2 Je höher die Aufgabenschwierigkeit, desto geringer ist die Trefferquote.

In Block 2 sind zusätzlich zu den mentalen Belastungsfaktoren die emotionalen Belastungsfaktoren durch das Leistungsfeedback operationalisiert. Da emotionale und mentale Belastungsfaktoren auf der Wirkungsebene nicht eindeutig voneinander trennbar sind (siehe Abschnitt 2.1.3), ist von einem Einfluss des Feedbacks auf die NASA-RTLX-Werte auszugehen. Zusätzlich wird die Interaktion zwischen der Trefferquote und dem Feedback geprüft. Die Hypothesen für die erwarteten Haupteffekte lauten:

H1.3 Positives Feedback führt zu geringeren und negatives Feedback zu höheren NASA-RTLX-Werten im Vergleich zu keinem Feedback.

H1.4 Positives Feedback führt zu einer besseren und negatives Feedback zu einer geringeren Trefferquote im Vergleich zu keinem Feedback.

Aufgrund der anzunehmenden Wechselbeziehung mentaler und emotionaler Beanspruchung ist für Block 3 der Einfluss emotionaler Belastung auf die

mentale Beanspruchung zu prüfen. Da für diesen Block keine Leistungsmessung möglich war, wird die Prüfung nur für die subjektive mentale Beanspruchung durchgeführt. Hierzu werden folgende Hypothesen postuliert:

H1.5 Je höher die Erregung des IAPS-Bildersets, desto höher sind die NASA-RTLX-Werte.

H1.6 Je geringer (negativer) die Valenz des IAPS-Bildersets, desto höher sind die NASA-RTLX-Werte.

Als Nächstes werden die Effekte auf die Emotionsdimensionen Erregung und Valenz untersucht. In Block 1 gab es kein Feedback und somit keine explizite Emotionsinduktion. Aus diesem Grund werden für die Experimentalbedingungen dieses Blocks eine tendenziell neutrale Erregung und Valenz angenommen. Neutral bezieht sich hierbei auf einen Wertebereich, der zwischen den Endpunkten des positiven und negativen Feedbacks liegt. In Block 2 wird für die Erregung gemäß den Annahmen von J. Kim et al. (2005) eine Wechselwirkung zwischen der Aufgabenschwierigkeit und der Art des Feedbacks zugrunde gelegt. Dabei ist von einer disordinalen Interaktion (gegenläufiger Trend der Effekte) auszugehen. In einem solchen Fall ist der Effekt der Aufgabenschwierigkeit auf die subjektive Erregung davon abhängig, ob das Leistungsfeedback positiv oder negativ ist. Wie sich das neutrale bzw. kein Feedback (Block 1) in Relation dazu verhält ist nicht bekannt.

H1.7 Es gibt einen signifikanten Interaktionseffekt zwischen dem Feedback und der Schwierigkeit der Aufgabe auf die Erregungs-Werte im SAM: Für positives Feedback gibt es einen positiven und für negatives Feedback einen negativen Effekt der Schwierigkeit auf die Erregungs-Werte im SAM.

In Block 2 ist davon auszugehen, dass negatives Leistungsfeedback zu einer geringen (d. h. negativen) Valenz und positives Leistungsfeedback zu einer hohen (d. h. positiven) Valenz im SAM führt. Da in Block 1 kein Feedback gegeben wurde, ist hier von einer mittleren Valenz auszugehen. Die Ausprägung

der subjektiven Valenz für diesen Block sollte zwischen der des positiven und negativen Leistungsfeedbacks liegen. Zudem wird der Haupteffekt der Aufgabenschwierigkeit und dessen Interaktion mit dem Feedback untersucht. Diesbezüglich werden keine Vorannahmen getroffen.

H1.8 Die Valenz des Feedbacks hat einen signifikanten positiven Haupteffekt auf die Valenz-Werte im SAM.

In Block 3 wird geprüft, ob sich die Erregungs- und Valenzausprägung der IAPS-Bildersets in der subjektiven Erregung und Valenz widerspiegeln. Geprüft wird zudem der Interaktionseffekt zwischen der Erregungs- und Valenzausprägung der Bildersets.

H1.9 Je höher die Erregung des IAPS-Bildersets, desto höher sind die Erregungs-Werte im SAM.

H1.10 Je höher die Valenz des IAPS-Bildersets, desto höher sind die Valenz-Werte im SAM.

Auswertungsmethoden der Manipulationskontrolle

Die Hypothesen der Manipulationskontrolle werden mit Hilfe von Mehrebenenanalysen (MEA) geprüft. Zur Beantwortung der Hypothesen H1.1 bis H1.4 sowie H1.7 bis H1.8 wird eine gemeinsame MEA aus Block 1 und 2 berechnet. Für jede Kriteriumsvariable wird eine separate MEA mit den Haupteffekten Schwierigkeit (S), Feedback (F) und dem Interaktionseffekt Schwierigkeit \times Feedback (S \times F) berechnet. Die Aufgabenschwierigkeit mit den Stufen leicht (S1) und schwer (S2) sowie das Aufgabenfeedback mit den drei Stufen positiv (+), negativ (-) und kein Feedback () werden aufgrund des Nominalskalenniveaus dummykodiert. Darüber hinaus erfolgt eine Kontrolle der Glaubwürdigkeit des Feedbacks, um einen Einfluss auf die Kriteriumsvariablen auszuschließen. Zu diesem Zweck wird ein erweitertes Mehrebenenmodell mit dem zusätzlichen festen Effekt Glaubwürdigkeit (G) berechnet.

Die Hypothesen H1.5 bis H1.6 sowie H1.9 bis H1.10 werden ebenfalls mit Hilfe eines gemeinsamen Mehrebenenmodells geprüft. Die Unterschiede zwischen den Emotionsstufen der IAPS-Bildersets – bestehend aus niedriger (En) oder hoher (Eh) Erregung und positiver (V+) oder negativer (V-) Valenz – werden dummykodiert. Getestet werden sowohl die Haupteffekte (Erregung und Valenz des Bildersets) als auch der Interaktionseffekt Erregung \times Valenz (E \times V). Eine niedrige Erregung und eine positive Valenz im Bilderset bilden jeweils die Referenzgruppe der Dummykodierung.

Des Weiteren wird mit der MEA überprüft, ob ein bedeutsamer Unterschied in den Kriteriumsvariablen zwischen den Blöcken 1, 2 und 3 besteht. Hierfür werden die drei Blöcke dummykodiert, wobei der Block 3 (IAPS-Bedingungen) als Referenz dient.

Für die Hypothesentests der Regressionsgewichte wird eine Bonferroni-Korrektur durchgeführt, um eine Alphafehlerkumulierung zu vermeiden. Das korrigierte Alphaniveau wird in den entsprechenden Abschnitten der Ergebnisbeschreibung berichtet.

6.2.6 Stichprobe

Am Experiment der Studie I nahmen insgesamt 50 Versuchspersonen teil, die überwiegend im universitärem Umfeld rekrutiert wurden. Folglich waren 84.0 % Studenten. Davon gaben 70.0 % das männliche und 30.0 % das weibliche Geschlecht an. Der Altersdurchschnitt lag bei 25.0 ($SD = 7.8$).

Da bei der Aufzeichnung von physiologischen Daten Sensorausfälle nicht vollständig vermieden werden konnten, gelten die folgenden Stichproben: Die Analyse der kardialen Aktivität wurde mit 44 Personen durchgeführt, da sechs Versuchspersonen aufgrund fehlerhafter Aufnahmen oder Sensorausfälle ausgeschlossen werden mussten. Die Analyse der elektrodermalen Aktivität erfolgte für 48 Personen. Zwei Versuchspersonen wurden aufgrund von Sensorausfällen ausgeschlossen.

6.3 Ergebnisse

In den folgenden Abschnitten werden nacheinander die Ergebnisse zu den Forschungsfragen berichtet. Zu Anfang wird in Abschnitt 6.3.1 geprüft, ob die Induktion mentaler und emotionaler Beanspruchung im Experiment gelungen ist. Im Anschluss folgt die Extraktion kardialer Profile mentaler Beanspruchung in Abschnitt 6.3.2 und 6.3.3, die Extraktion elektrodermalen Profile mentaler Beanspruchung in Abschnitt 6.3.4 und 6.3.5, die Extraktion kardialer Profile emotionaler Beanspruchung in Abschnitt 6.3.6 und 6.3.7 und die Extraktion elektrodermalen Profile emotionaler Beanspruchung in Abschnitt 6.3.8 und 6.3.9. Zu diesem Zweck werden NPLS-Regressionen der zweidimensionalen Kriteriumsmatrix (Bedingung \times Person) auf die dreidimensionale Prädiktormatrix (Bedingung \times Person \times Parameter) durchgeführt. Zum Schluss wird im Sinne der Spezifität geprüft, ob die kardialen (Abschnitt 6.3.10) und elektrodermalen (Abschnitt 6.3.11) Profile der mentalen Beanspruchung von den entsprechenden Profilen der emotionalen Beanspruchung bedeutsam abgegrenzt werden können. Die deskriptive Statistik kardialer und elektrodermalen Parameter ist aus Platzgründen im Anhang A.6.1 und A.6.2 dargestellt.

6.3.1 Kontrollfragestellung KF1: Ist die experimentelle Induktion mentaler und emotionaler Beanspruchung in Studie I gelungen?

In den nachfolgenden Abschnitten werden zunächst die Kontrollvariablen berichtet und analysiert. Im Anschluss erfolgt nacheinander die Analyse der NASA-RTLX-Werte, der Trefferquote sowie der Erregungs- und Valenz-Werte aus dem SAM.

Analyse der Kontrollvariablen

Die deskriptiven Statistiken der Kontrollvariablen sind in Tabelle 6.3 dargestellt. Die Glaubwürdigkeit des Feedbacks liegt im mittleren Wertebereich ($M = 3.09$). Ausreißer ($\pm 3 SD$) liegen nicht vor.

Tabelle 6.3: Deskriptive Statistiken der Kontrollvariablen

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
Glaubwürdigkeit	50	3.09	1.01	1.33	4.67	-0.12	-1.32
Anfangstemperatur	49	22.62	0.87	20.4	23.8	-0.88	-0.11
Endtemperatur	48	23.34	0.66	21.7	24.6	-0.53	-0.27
Luftfeuchtigkeit (Beginn)	49	42.00	7.36	27.0	55.0	0.11	-1.07
Luftfeuchtigkeit (Ende)	48	41.94	6.76	28.0	53.0	0.07	-1.04

Für die Raumtemperatur und die relative Luftfeuchtigkeit fehlen jeweils ein bis zwei Messpunkte. Ein t-Test für abhängige Stichproben³ deckt auf, dass es einen signifikanten Anstieg in der Raumtemperatur ($t(47) = -11.39$, $p < .001$, $d = 1.64$), jedoch nicht in der relativen Luftfeuchtigkeit ($t(47) = 1.32$, $p = .192$) zum Ende des Experiments gibt. Der mittlere Anstieg in der Temperatur beträgt $0.72\text{ }^{\circ}\text{C}$.

Analyse des NASA-RTLX

Die NASA-RTLX-Werte für die zehn Experimentalbedingungen sind in Abbildung 6.6 dargestellt. Eine Tabelle mit den deskriptiven Statistiken des NASA-RTLX findet sich im Anhang A-3. Ausreißer ($\pm 3 SD$) oder fehlende Daten liegen nicht vor. Deskriptiv ist bereits ein erwartungskonformer Trend im Sinne einer Zunahme der subjektiven mentalen Beanspruchung im

³ Signifikanzniveau von 5 % (zweiseitig).

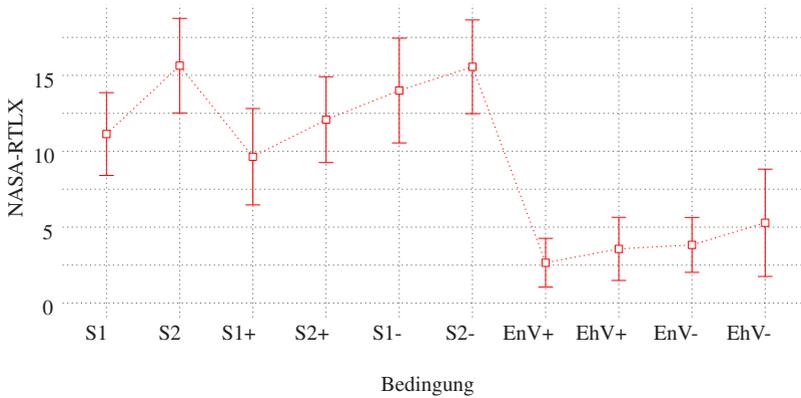


Abbildung 6.6: Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.

NASA-RTLX bei hoher Aufgabenschwierigkeit sowie bei negativem Feedback zu erkennen. Eine geringe Erhöhung ist auch bei den Bildersets mit hoher Erregung und negativer Valenz feststellbar. Insgesamt bewegen sich die mittleren NASA-RTLX-Werte bei Bearbeitung der Balltrackingaufgaben zwischen 9.64 und 15.56 (Standardabweichung zwischen 2.73 und 3.45) und sind damit insgesamt höher als bei den IAPS-Bildersets: Mittelwert zwischen 2.66 und 5.29 und Standardabweichung zwischen 1.61 und 3.54.

Ob die deskriptiv beobachteten Trends der Aufgabenschwierigkeit und des Aufgabenfeedbacks inferenzstatistisch bestätigt werden können, wird mit der ersten MEA überprüft. Diese deckt die Hypothesen H1.1 und H1.3 der Manipulationskontrolle ab (siehe Abschnitt 6.2.5).

Die Ergebnisse des Modellvergleichs in Tabelle 6.4 zeigen eine signifikante Reduktion der Devianz ($-2LL$) und des AIC sowohl für die Haupteffektmodelle der Schwierigkeit (M_S) und des Feedback (M_F) als auch für deren Interaktion ($M_{S \times F}$). Der Anteil zusätzlich erklärter Varianz ist für die Haupteffekte mit 14.1 % und 18.4 % größer als für den Interaktionseffekt mit 2.6 %. Die Glaubwürdigkeit des Feedbacks hat keinen bedeutsamen Effekt auf die subjek-

Tabelle 6.4: Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX in Block 1 und 2

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	1608.65	1602.65			0.00
+ Schwierigkeit (M_S)	4	1540.19	1532.19	70.45***	1	14.14
+ Feedback (M_F)	6	1405.51	1393.51	138.68***	2	32.62
+ Schwierigkeit \times Feed- back ($M_{S \times F}$)	8	1381.25	1365.25	28.26***	2	35.28
+ Glaubwürdigkeit (M_G)	9	1383.0	1365.0	0.26	1	35.50

*** $p < .001$, ** $p < .01$, * $p < .05$.

tive mentale Beanspruchung im NASA-RTLX. Die Regressionsgewichte der einzelnen Faktorstufen sind in Tabelle 6.5 für das Modell $M_{S \times F}$ dargestellt.⁴ Alle Regressionseffekte sind signifikant und bestätigen die zuvor beobachteten Trends: signifikante Zunahme der NASA-RTLX-Werte bei hoher (im Vergleich zur geringen) Aufgabenschwierigkeit und bei negativem Aufgabenfeedback (im Vergleich zu keinem Aufgabenfeedback). Zudem zeigt sich ein signifikanter Abfall der NASA-RTLX-Werte bei positivem Aufgabenfeedback.

Der signifikante ordinale Interaktionseffekt (siehe Abbildung 6.6) repräsentiert den höheren Unterschied (steilerer Anstieg) zwischen den Bedingungen ohne Feedback (S1 und S2) im Vergleich zu denselben Schwierigkeitsstufen mit Aufgabenfeedback: flacherer Anstieg zwischen S1+ und S2+ bzw. S1- und S2-. Der Anstieg zwischen den Schwierigkeitsstufen mit Aufgabenfeedback unterscheidet sich nicht voneinander (siehe nicht signifikanter Interaktionseffekt in Tabelle 6.5, letzte Zeile). Die Teststärken für die bedeutsamen Effekte des Modells $M_{S \times F}$ liegen zwischen 96.0 % und 100.0 %.

In einer zweiten MEA werden die Effekte der IAPS-Bildersets auf die NASA-RTLX-Werte in Block 3 untersucht. Diese beziehen sich auf die

⁴ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .007$.

Tabelle 6.5: Parameter für das Modell $M_{S \times F}$ des NASA-RTLX

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	11.14***	0.43	25.89	109.80	2.33
Schwierigkeit: 2 – 1	4.50***	0.39	11.55	250.00	
Feedback: pos. – neut.	-1.50**	0.39	-3.83	250.00	
Feedback: neg. – neut.	2.86***	0.39	7.34	250.00	
Feedback: neg. – pos.	4.36***	0.39	11.17	250.00	
Schwierigkeit (2 – 1) × Feedback (pos. – neut.)	-2.06**	0.55	-3.73	250.00	
Schwierigkeit (2 – 1) × Feedback (neg. – neut.)	-2.94***	0.55	-5.33	250.00	
Schwierigkeit (2 – 1) × Feedback (neg. – pos.)	-0.88	0.55	-1.59	250.00	

*** $p < .0001$, ** $p < .001$, * $p < .007$.

Hypothesen H1.5 und H1.6 (siehe Abschnitt 6.2.5). Die Ergebnisse der Modellvergleiche sind in Tabelle 6.6 abgebildet. Der AIC sowie die Devianz zeigen einen signifikanten Haupteffekt der Valenz der Bildersets (M_V) sowie einen signifikanten Haupteffekt der Bilderset-Erregung (M_E). Beide Faktoren klären zusammen 13.6 % der Varianz auf. Die Regressionskoeffizienten der beiden Haupteffekte finden sich in Tabelle 6.7.⁵ Diese belegen einen signifikanten Anstieg der NASA-RTLX-Werte für diejenigen IAPS-Bildersets mit negativer (im Vergleich zu positiver) Valenz und hoher (im Vergleich zu niedriger) Erregung. Die Teststärken für die beiden Haupteffekte des Modells M_E liegen zwischen 99.5 % und 100.0 %.

⁵ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .025$

Tabelle 6.6: Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX in Block 3

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	925.97	919.97			0.00
+ Valenz (M_V)	4	903.06	895.06	24.91***	1	8.17
+ Erregung (M_E)	5	885.75	875.75	19.31***	1	13.63
+ Valenz \times Erregung ($M_{V \times E}$)	6	886.65	874.65	1.09	1	13.92

*** $p < .001$, ** $p < .01$, * $p < .05$.

In einer weiteren MEA werden die Unterschiede zwischen den Blöcken 1 bis 3 untersucht. Das Ergebnis des Modellvergleichs zeigt eine bedeutsame Reduktion der Devianz für das Hauptfaktormodell M_{Block} , $\chi^2_{diff}(2) = 573.2$, $p < .001$, $R^2_{GLMM} = 64.64$. Die Regressionsgewichte⁶ (siehe Tabelle 6.8) belegen signifikant geringere NASA-RTLX-Werte in Block 3 (IAPS-Bildersets) im Vergleich zu den Blöcken 1 und 2 (Balltrackingaufgabe). Für beide Effekte liegen die Teststärken bei 100.0 %. Mit $B = 8.99$ und 9.55 sind die hier vorgefundenen Effekte größer als alle anderen Effekte, die für das NASA-RTLX

Tabelle 6.7: Parameter für das Modell M_E des NASA-RTLX

	Feste Effekte				Zufallseffekte
	B	SE	t	df	SD
(Ordinatenabschnitt)	2.52***	0.31	8.20	110.51	1.47
Valenz: neg. – pos.	1.45***	0.26	5.55	150.00	
Erregung: hoch – niedrig	1.18***	0.26	4.54	150.00	

*** $p < .001$, ** $p < .01$, * $p < .025$.

⁶ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .016$.

vorgefunden wurden (z. B. Effekt der Aufgabenschwierigkeit mit $B = 4.50$). Bedeutsame Unterschiede in den NASA-RTLX-Werten zwischen den Blöcken 1 und 2 finden sich nicht.

Tabelle 6.8: Parameter für das Modell M_{Block} des NASA-RTLX

	Feste Effekte				Zufallseffekte
	B	SE	t	df	SD
(Ordinatenabschnitt)	3.84***	0.30	12.72	96.96	1.54
Block: 2 – 3	8.99***	0.30	30.40	450.00	
Block: 1 – 3	9.55***	0.36	26.38	450.00	
Block: 2 – 1	-0.57	0.36	-1.56	450.00	

*** $p < .001$, ** $p < .01$, * $p < .016$.

Analyse der Trefferquote

Eine Leistungserfassung mit Hilfe der Trefferquote wurde ausschließlich in den Balltrackingaufgabe durchgeführt und kann nur für diese sechs Bedingungen ausgewertet werden. Die Trefferquote für die sechs Experimentalbedingungen ist in Abbildung 6.7 dargestellt. Eine Tabelle mit den deskriptiven Statistiken der Trefferquote findet sich im Anhang A-4. Ausreißer ($\pm 3 SD$) oder fehlende Daten liegen nicht vor. Im Gesamtdurchschnitt wurden 28.9 % ($SD = 25.5$) der arithmetischen Aufgaben richtig gelöst.

Mit Hilfe der MEA werden die Hypothesen H1.2 und H1.4 (siehe Abschnitt 6.2.5) geprüft. Die Ergebnisse der geschachtelten Modellvergleiche in Tabelle 6.9 zeigen eine signifikante Reduktion der Devianz und des AIC für die Modelle mit den Haupteffekten Schwierigkeit (M_S) und Feedback (M_F). Das Modell M_F erklärt dabei einen größeren Anteil (15.5 %) an der Gesamtvarianz als das Modell M_S (5.6 %). Die Glaubwürdigkeit des Feedbacks hat keinen bedeutsamen Einfluss auf die Trefferquote. Die Regressionsgewichte

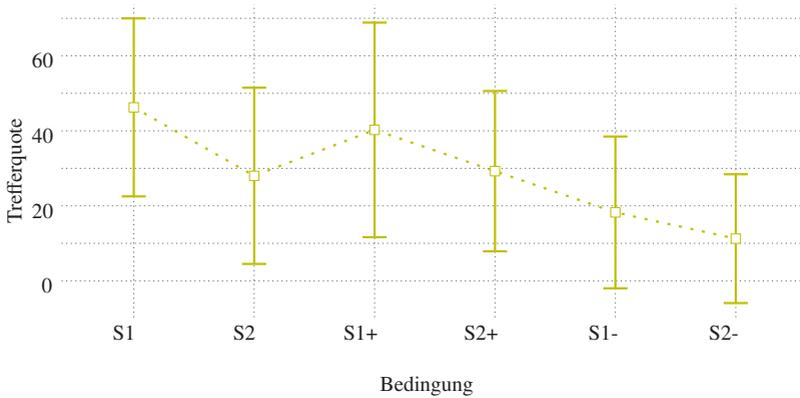


Abbildung 6.7: Mittelwert und Standardabweichung der Trefferquote in Prozent in den Versuchsbedingungen.

der einzelnen Faktorstufen sind in Tabelle 6.10 für das Modell M_F dargestellt.⁷ Ein signifikanter Abfall der Trefferquote findet sich bei zunehmender Aufgabenschwierigkeit sowie bei negativem Aufgabenfeedback im Vergleich zum positiven und keinem Feedback (Teststärken bei 100.0 %). Im Gegensatz

Tabelle 6.9: Modellfit-Indizes für die geschachtelten Modelle der Trefferquote

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	2758.00	2752.00			0.00
+ Schwierigkeit (M_S)	4	2734.88	2726.88	25.12***	1	5.65
+ Feedback (M_F)	6	2652.70	2640.70	86.18***	2	21.23
+ Schwierigkeit \times Feedback ($M_{S \times F}$)	8	2651.10	2635.10	5.60	2	22.07
+ Glaubwürdigkeit (M_G)	9	2650.6	2632.6	2.51	1	24.01

*** $p < .001$, ** $p < .01$, * $p < .05$.

⁷ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .013$.

zur subjektiven mentalen Beanspruchung im NASA-RTLX findet sich für die Trefferquote kein Interaktionseffekt und kein bedeutsamer Unterschied in der Leistung zwischen positivem und keinem Feedback.

Tabelle 6.10: Parameter für das Modell M_F der Trefferquote

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	43.17***	2.87	15.02	113.46	14.74
Schwierigkeit: 2 – 1	-12.08***	1.98	-6.11	250.00	
Feedback: pos. – neut.	-2.38	2.42	-0.98	250.00	
Feedback: neg. – neut.	-22.38***	2.42	-9.23	250.00	
Feedback: neg. – pos.	-20.00***	2.42	-8.25	250.00	

*** $p < .001$, ** $p < .005$, * $p < .013$

Analyse der Erregung

Die Erregungs-Werte des SAM sind für die zehn Experimentalbedingungen in Abbildung 6.8 dargestellt. Eine Tabelle mit den deskriptiven Statistiken der Erregung findet sich im Anhang A-5. Ausreißer ($\pm 3 SD$) oder fehlende Daten liegen nicht vor. Der deskriptive Trend der Erregungs-Werte zeigt für Block 1 und 2 einen tendenziellen Anstieg für die Schwierigkeitsstufe 2 sowie für negatives Leistungsfeedback. Ein ähnlicher Verlauf wurde bereits für die NASA-RTLX-Werte beschrieben. Im Gegensatz zu den NASA-RTLX-Werten lässt sich in den Erregungs-Werten jedoch kein Unterschied zwischen den Bedingungen Block 3 und den beiden anderen Blöcken (Block 1 und 2) erkennen.

Ob sich die beobachteten Trends der Aufgabenschwierigkeit und des Aufgabenfeedbacks inferenzstatistisch bestätigen, wird mit der ersten MEA für Block 1 und 2 geprüft. Diese bezieht sich auf die Hypothese H1.7 (siehe Abschnitt 6.2.5).

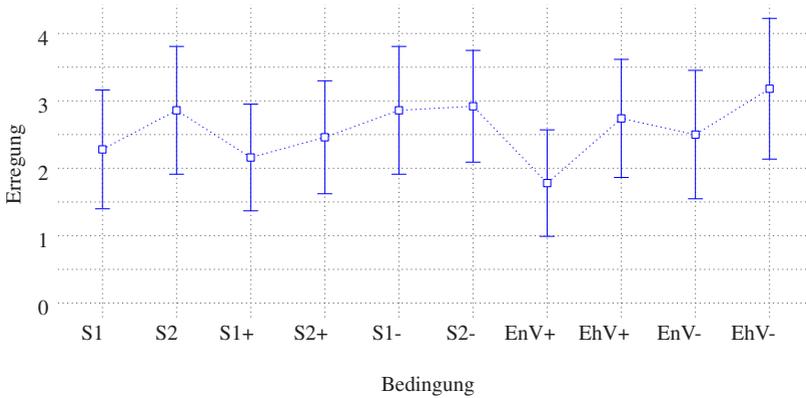


Abbildung 6.8: Mittelwert und Standardabweichung der Erregung in den Versuchsbedingungen.

Die Ergebnisse des Modellvergleichs in Tabelle 6.11 zeigen eine signifikante Reduktion der Devianz und des AIC sowohl für die Haupteffekte Schwierigkeit (M_S) und Feedback (M_F) als auch deren Interaktion ($M_{S \times F}$) auf. Der Anteil zusätzlich erklärter Varianz ist für den Interaktionseffekt mit 1.3 % am geringsten, gefolgt vom Haupteffekt der Schwierigkeit mit 2.9 % und mit 6.7 % am

Tabelle 6.11: Modellfit-Indizes für die geschachtelten Modelle der Erregung in Block 1 und 2

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	690.12	684.12			0.00
+ Schwierigkeit (M_S)	4	673.54	665.54	18.57***	1	2.92
+ Feedback (M_F)	6	628.90	616.90	48.64***	2	9.63
+ Schwierigkeit \times Feedback ($M_{S \times F}$)	8	621.89	605.89	11.01**	2	10.97
+ Glaubwürdigkeit (M_G)	9	623.51	605.51	0.38	1	11.42

*** $p < .001$, ** $p < .01$, * $p < .05$.

Tabelle 6.12: Parameter für das Modell $M_{S \times F}$ der Erregung

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	2.28***	0.12	18.62	107.42	0.67
Schwierigkeit: 2 – 1	0.58***	0.11	5.29	250.00	
Feedback: pos. – neut.	-0.12	0.11	-1.09	250.00	
Feedback: neg. – neut.	0.58***	0.11	5.29	250.00	
Feedback: neg. – pos.	0.70***	0.11	6.38	250.00	
Schwierigkeit (2 – 1) × Feedback (pos. – neut.)	-0.28	0.16	-1.80	250.00	
Schwierigkeit (2 – 1) × Feedback (neg. – neut.)	-0.52**	0.16	-3.35	250.00	
Schwierigkeit (2 – 1) × Feedback (neg. – pos.)	-0.24	0.16	-1.55	250.00	

*** $p < .0001$, ** $p < .001$, * $p < .007$.

höchsten für das Feedback. Die Glaubwürdigkeit des Feedbacks zeigt keinen bedeutsamen Effekt auf die Erregung im SAM.

Die Regressionsgewichte der einzelnen Faktorstufen sind in Tabelle 6.12 für das Modell $M_{S \times F}$ dargestellt.⁸ Hier zeigt sich ein bedeutsamer Anstieg in der Erregung für die Schwierigkeitsstufe 2 sowie beim negativem (im Vergleich zu neutralem oder positivem) Feedback. Beide Effekte zeigen eine Teststärke von 100.0 %. Zwischen dem neutralen und positiven Feedback findet sich kein signifikanter Unterschied. Der Interaktionseffekt kann anhand der Abbildung 6.8 verdeutlicht werden. Hier zeigt sich eine Steigung der Erregung zwischen den Bedingungen S1 zu S2 (Steigerung der Schwierigkeitsstufe ohne Feedback). Diese Steigung ist bei negativem Feedback (Bedingung S1- zu

⁸ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .007$.

Tabelle 6.13: Modellfit-Indizes für die geschachtelten Modelle der Erregung in Block 3

Modell	<i>df</i>	AIC	-2LL	χ^2_{diff}	<i>df</i> _{diff}	R^2_{GLMM}
Nullmodell	3	575.98	569.98			0.00
+ Valenz (M_V)	4	556.02	548.02	21.96***	1	7.77
+ Erregung (M_E)	5	501.24	491.24	56.78***	1	23.28
+ Valenz \times Erregung ($M_{V \times E}$)	6	501.21	489.21	2.03	1	23.73

*** $p < .001$, ** $p < .01$, * $p < .05$.

S2-) nicht gegeben. Folglich unterscheiden sich beide Steigungen signifikant voneinander (siehe Tabelle 6.12, Zeile 7; Teststärke bei 92.1 %). Die Steigung zwischen den Bedingungen S1+ zu S2+ liegt dazwischen und zeigt deshalb keine bedeutsamen Unterschiede zu den Steigungen innerhalb der neutralen und negativen Feedbackbedingungen (siehe Tabelle 6.12, Zeile 6 und 8). Insgesamt widerspricht das vorgefundene Interaktionsmuster der Annahme einer disordinalen Interaktion in H1.7.

Die zweite MEA untersucht die Effekte innerhalb der IAPS-Bedingungen (Block 3). Sie bezieht sich auf die Hypothese H1.9. Die Ergebnisse der Modell-

Tabelle 6.14: Parameter für das Modell M_E der Erregung

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	1.85***	0.12	15.49	105.77	0.59
Valenz: neg. – pos.	0.58***	0.10	5.88	150.00	
Erregung: hoch – niedrig	0.82***	0.10	8.31	150.00	

*** $p < .001$, ** $p < .01$, * $p < .025$.

vergleiche sind in Tabelle 6.13 dargestellt. Der AIC sowie die Devianz zeigen einen signifikanten Haupteffekt der Valenz (M_V) sowie Erregung (M_E). Der Faktor Erregung leistet dabei einen größeren Beitrag (15.5 %) zur Aufklärung der Gesamtvarianz als der Faktor Valenz (7.7 %). Die Regressionskoeffizienten der beiden Haupteffekte sind in Tabelle 6.14 dargestellt.⁹ Diese belegen eine signifikante Steigung der Erregungs-Werte im SAM für IAPS-Bilder mit negativer (im Vergleich zu positiver) Valenz und für IAPS-Bilder mit hoher (im Vergleich zu niedriger) Erregung. Die Teststärke beider Haupteffekte liegt bei 100.0 %.

In einer letzten MEA wird der Unterschied zwischen den Versuchsblöcken 1, 2 und 3 untersucht. Hierbei finden sich keine bedeutsamen Effekte, $\chi^2_{\text{diff}}(2) = 0.4$, $p = .817$.

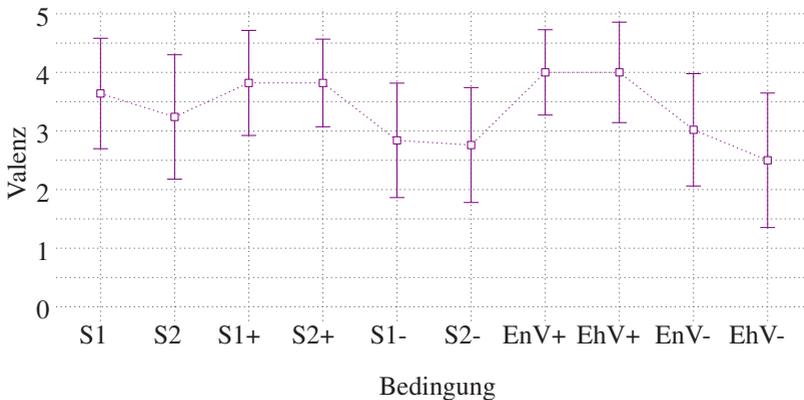


Abbildung 6.9: Mittelwert und Standardabweichung der Valenz in den Versuchsbedingungen.

Analyse der Valenz

Die Valenz-Werte des SAM sind für die zehn Experimentalbedingungen in Abbildung 6.9 dargestellt. Eine Tabelle mit den deskriptiven Statistiken der Valenz findet sich im Anhang A-6. Ausreißer ($\pm 3 SD$) oder fehlende Daten liegen nicht vor.

Zunächst werden die Bedingungsunterschiede innerhalb der Balltrackingaufgabe (Block 1 und 2) in Bezug auf die Hypothese H1.8 (siehe Abschnitt 6.2.5) untersucht. Die Ergebnisse des Modellvergleichs in Tabelle 6.15 zeigen eine signifikante Reduktion der Devianz und des AIC allein für den Haupteffekt des Feedbacks (M_F). Aus Tabelle 6.16 ist erkennbar, dass die Valenz-Werte bei positivem Feedback im Vergleich zu keinem Feedback signifikant steigen (positivere Valenz) und bei negativem Feedback im Vergleich zu keinem oder positivem Feedback signifikant sinken (negativere Valenz).¹⁰ Die Teststärken für die bedeutsamen Effekte des Feedbacks in Modell M_F liegen zwischen 96.7 % und 100.0 %.

Tabelle 6.15: Modellfit-Indizes für die geschachtelten Modelle der Valenz in Block 1 und 2

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	834.26	828.26			0.00
+ Schwierigkeit (M_S)	4	833.70	825.70	2.55	1	0.61
+ Feedback (M_F)	6	754.16	742.16	83.55***	2	17.44
+ Schwierigkeit \times Feedback ($M_{S \times F}$)	8	753.94	737.94	4.22	2	18.15
+ Glaubwürdigkeit (M_G)	9	752.97	734.97	2.97	1	20.48

*** $p < .001$, ** $p < .01$, * $p < .05$.

⁹ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .025$.

¹⁰ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .013$.

Tabelle 6.16: Parameter für das Modell M_F der Valenz

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	3.52***	0.12	29.90	123.08	0.58
Schwierigkeit: 2 – 1	-0.16	0.08	-1.89	250.00	
Feedback: pos. – neut.	0.38***	0.10	3.67	250.00	
Feedback: neg. – neut.	-0.64***	0.10	-6.18	250.00	
Feedback: neg. – pos.	-1.02***	0.10	-9.85	250.00	

*** $p < .001$, ** $p < .005$, * $p < .013$.

Innerhalb der zweiten MEA werden die Unterschiede zwischen den IAPS-Bedingungen (Block 3) für die Hypothese H1.10 (siehe Abschnitt 6.2.5) untersucht. Gemäß dem AIC sowie der Devianz in Tabelle 6.17 sind die Modelle M_V , M_E und $M_{V \times E}$ signifikant. Der Faktor Valenz (des Bildersets) leistet mit 30.2 % den größten Beitrag zur Varianzaufklärung (siehe Tabelle 6.17). Der Faktor Erregung (des Bildersets) und die Interaktion Valenz \times Erregung klären gemeinsam nur 2.7 % der Varianz in den Valenz-Werten des SAM auf. Die Regressionskoeffizienten sind in Tabelle 6.18 für das Modell $M_{V \times E}$

Tabelle 6.17: Modellfit-Indizes für die geschachtelten Modelle der Valenz in Block 3

Modell	<i>df</i>	AIC	-2LL	χ^2_{diff}	<i>df</i> _{diff}	R^2_{GLMM}
Nullmodell	3	618.80	612.80			0.00
+ Valenz (M_V)	4	530.15	522.15	90.65***	1	30.24
+ Erregung (M_E)	5	526.57	516.57	5.58	1	31.57
+ Valenz \times Erregung ($M_{V \times E}$)	6	522.78	510.78	5.79*	1	32.90

*** $p < .001$, ** $p < .01$, * $p < .05$.

Tabelle 6.18: Parameter für das Modell $M_{V \times E}$ der Valenz

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	4.00***	0.13	30.55	150.21	0.53
Valenz: neg. – pos.	-0.98***	0.15	-6.48	150.00	
Erregung: hoch – niedrig	0.00	0.15	0.00	150.00	
Valenz \times Erregung	-0.52*	0.21	-2.43	150.00	

*** $p < .001$, ** $p < .01$, * $p < .016$.

dargestellt.¹¹ Nur das Regressionsgewicht des Faktors Valenz erreicht dabei das korrigierte Signifikanzniveau von $\alpha = .016$. Hierbei ist ein signifikanter Abfall der Valenz-Werte für Bildersets mit negativer Valenz im Vergleich zu Bildersets mit positiver Valenz festzustellen (Teststärke bei 100.0 %).

In einer letzten MEA werden die Unterschiede zwischen den Versuchsblöcken 1, 2 und 3 untersucht. Dabei finden sich keine signifikanten Effekte der Blöcke auf die Valenz-Werte im SAM, $\chi^2_{\text{diff}}(2) = 1.3$, $p = .511$.

6.3.2 Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

In diesem Abschnitt werden ausführlich die Ergebnisse der NPLS-Regression beschrieben und erläutert. Exemplarisch wird dabei das Vorgehen bei der Analyse und Interpretation mehrdimensionaler Daten vermittelt. Die Ergebnisse

¹¹ Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .016$.

der restlichen Fragestellungen werden aufbauend auf dem hier vermittelten Wissen verkürzt dargestellt.

In einem ersten Schritt werden die Anzahl der Komponenten gemäß Abschnitt 5.6.2 bestimmt sowie die Komponentenmodelle kreuzvalidiert. Die Ergebnisse der Kreuzvalidierung sind für die NPLS-Regression vom NASA-RTLX auf die kardialen Parameter in Tabelle 6.19 dargestellt.¹² Nur das Komponentenmodell 1 erfüllt die in Abschnitt 5.6.2 definierten Voraussetzungen¹³ zur Bestimmung eines bedeutsamen Zuwachses an Vorhersagegüte. Folglich soll Komponente 1 weiter analysiert werden. Im Originalmodell erklärt die Komponente 1 23.4 % der Varianz in den kardialen Parametern. In den NASA-RTLX-Werten sagt die Komponente 1 73.2 % der Varianz vorher. Im kreuzvalidierten Modell verringert sich der Anteil erklärter Varianz im NASA-RTLX auf 56.8 %.

Tabelle 6.19: NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					673.21
1	23.39	21.80 [†]	73.17	56.78 [†]	171.15***
2	3.66	1.44 [†]	14.47	6.01 [†]	147.34
3	4.83	2.12 [†]	2.83	-0.66	149.95
4	3.87	1.02 [†]	2.28	0.83	146.66
5	2.96	0.51	2.65	1.52 [†]	140.62

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.94$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 1.10$.

¹² Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.48[0.16, 0.94]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -4.84[-13.89, 1.10]$.

¹³ Zu den Voraussetzungen gehört der über den Zufall hinausgehender Zuwachs im $\text{val}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix und der Kriteriumsmatrix sowie ein bedeutsamer Abfall der valSQR.

Die Ladungen der Personen, Bedingungen und kardialen Parameter auf der Komponente 1 sind in Abbildung 6.10 dargestellt.¹⁴ Zur Analyse der Ladungen wird zunächst die Dimension der Bedingungen herangezogen. Eine tabellarische Darstellung der Bedingungs-Faktorladungen findet sich im Anhang Tabelle A-7. Aufgrund der angewandten Zentrierung (innerhalb der Bedingungen) repräsentiert die Abszissenachse den durchschnittlichen Wert der kardialen Aktivität (Prädiktormatrix) bzw. der subjektiven mentalen Beanspruchung (Kriteriumsmatrix) über alle Bedingungen hinweg. Demnach repräsentieren positive Faktorladungen eine überdurchschnittliche und negative Faktorladungen eine unterdurchschnittliche kardiale Aktivität (Prädiktormatrix) bzw. subjektive mentale Beanspruchung (Kriteriumsmatrix) in der Bedingungsdimension. Betrachtet man die Bedingungs-Faktorladungen der NASA-RTLX-Komponente 1 in Abbildung 6.10 (Mitte, gestrichelte rote Linie), so erkennt man deutlich eine Übereinstimmung mit dem bereits bekannten Muster der NASA-RTLX-Mittelwerte (vgl. Abbildung 6.6). Die NASA-RTLX-Mittelwerte sind zum Vergleich ebenfalls in Abbildung 6.10 (Mitte) als gestrichelte dunkelgraue Linie dargestellt. So sind für die NASA-RTLX-Komponente 1 höhere Faktorladungen (überdurchschnittliche mentale Beanspruchung) für die höheren Schwierigkeitsstufen und für das negative Feedback erkennbar. Die negativen Faktorladungen der IAPS-Bedingungen repräsentieren eine unterdurchschnittliche subjektive mentale Beanspruchung. Die Übereinstimmung bestätigt, dass die aus dem NASA-RTLX extrahierte Komponente 1 tatsächlich die subjektive mentale Beanspruchung abbildet.

Die kardiale Komponente 1 wurde durch die NPLS-Regression so extrahiert, dass ihre Bedingungs-Faktorladungen maximal mit den zuvor beschriebenen Bedingungs-Faktorladungen der NASA-RTLX-Komponente 1 kovariieren.¹⁵

¹⁴ Für die kardiale Prädiktormatrix (dreidimensionale Matrix) ergeben sich Ladungen für jede Bedingung, jede Person und jeden Parameter. Die Kriteriumsmatrix des NASA-RTLX (zweidimensionale Matrix) besitzt Faktorladungen für jede Person und jede Bedingung.

¹⁵ Anders ausgedrückt, wurde aus der kardialen Prädiktormatrix der beste Repräsentant der subjektiven mentalen Beanspruchung über die Bedingungen hinweg bestimmt.

Das Regressionsgewicht B repräsentiert diesen Zusammenhang. Für die Komponente 1 ergibt sich: $B = 0.37$, $t(438) = 6.94$, $p < .001$. Die Bedingungs-Faktorladungen der kardialen Komponente 1 sind in Abbildung 6.10 (Mitte) als durchgezogene rote Linie zu sehen. Hier ist zu erkennen, dass diese Bedingungs-Faktorladungen überwiegend mit denen der NASA-RTLX-Komponente 1 übereinstimmen: Bspw. die negativen Faktorladungen (unterdurchschnittliche kardiale Aktivität) für die IAPS-Bedingungen und die positiven Faktorladungen (überdurchschnittliche kardiale Aktivität) für die Balltracking-Bedingungen. Allerdings zeigen sich auch Unterschiede in den Bedingungs-Faktorladungen der kardialen Komponente 1 und der NASA-RTLX-Komponente 1. Bedeutsame ($|a_{\Delta}| \geq .20$) und marginale ($|a_{\Delta}| \geq .15$) Faktorladungsunterschiede sind in der Abbildung 6.10 (Mitte) eingezeichnet. Hier ist zu sehen, dass die Faktorladungen der Bedingungen S2, S2+ und EhV- marginal und der Bedingung S1+ sogar bedeutsam voneinander abweichen. Dies bedeutet, dass für die kardiale Komponente 1 (im Gegensatz zur NASA-RTLX-Komponente 1) der Effekt der Aufgabenschwierigkeit bei den Bedingungen ohne Aufgabenfeedback (S1 und S2) kaum oder gar nicht vorhanden ist. Innerhalb der Feedback-Bedingungen ist der Anstieg zwischen den Schwierigkeitsstufen wiederum gegeben (vgl. S1+ und S1- mit S2+ und S2-). Zudem gibt es deskriptiv keinen Unterschied zwischen den Bedingungen mit negativem (S1- und S2-) und positivem (S1+ und S2+) Feedback für die kardiale Komponente 1 (im Gegensatz zur NASA-RTLX-Komponente 1). Auch zeigt die kardiale Komponente 1 keinen Anstieg mit steigender Erregung oder steigender Valenz der Bildersets. Insgesamt spiegelt die kardiale Komponente 1 auf deskriptiver Basis weder die im NASA-RTLX gefundenen Effekte des Aufgabenfeedbacks noch die Effekte der Erregung und Valenz der Bildersets wider.

Um zu verstehen, was genau mit dem Begriff der (über- bzw. unterdurchschnittlichen) kardialen Aktivität gemeint ist, wird die Dimension der Parameter (Abbildung 6.10, unten) näher betrachtet. Hier sind Ladungen der einzelnen Parameter auf der Komponente 1 zu sehen. Diese Ladungen bilden den Zu-

sammenhang der kardialen Parameter mit der Komponente 1 ab.¹⁶ Marginale Faktorladungen ($|a| \geq .20$) zeigen die Parameter HR, SDNN, pNN50, CVSD, Shannon.h, EKG.TP, EKG.LF, und SVDEn. Bedeutsame Faktorladungen haben hingegen die Parameter HP, MQSD, pNN20 und EKG.HF mit $|a| \geq .30$. Die hier dargestellte Ladungsstruktur der Parameter-Faktorladungen bildet ein erstes kardiales Profil ab. Eben dieses Profil – genauer gesagt die über Faktorladungen gewichtete Summe einzelner Parameter – definiert die kardiale Aktivität der Komponente 1. Wie hoch diese gewichtete Summe für die einzelnen Bedingungen des vorliegenden Datensatzes ist, wird an den Faktorladungen der Bedingungsdimension abgelesen. Sie beträgt bspw. .35 (also überdurchschnittlich) für die Bedingung S2+ und -.35 (also unterdurchschnittlich) für die Bedingung EnV+.¹⁷

Zum Schluss gilt es die Faktorladungen der Personen (siehe Abbildung 6.10, oben) und deren Bedeutung darzustellen. Zwar bilden die interindividuellen Unterschiede keinen Fokus der vorliegenden Arbeit, jedoch sind diese für die Vorhersagegüte des Modells und für das Verständnis der Ergebnisse relevant. Darum werden die Person-Faktorladungen in diesem Abschnitt ausführlich erläutert und in den weiteren Abschnitten zusammenfassend beschrieben. Bedingt durch die spezifische Zentrierung der Daten sind die Person-Faktorladungen für die Interaktionen Person \times Bedingung und Person \times Bedingung \times Parameter relevant. Einfacher ausgedrückt, geben die Person-Faktorladungen an, wie stark sich die Personen hinsichtlich ihrer durchschnittlichen kardialen Aktivität (durchgezogene rote Linie) bzw. ihrer subjektiven mentalen Beanspruchung

¹⁶ Es ist wichtig, die negativen und positiven Faktorladungen nicht mit einer unter- oder überdurchschnittlichen kardialen Aktivität (analog der Interpretation der Bedingungs-Faktorladungen) gleichzusetzen. Das Vorzeichen stellt hier lediglich die Richtung des Zusammenhangs kardialer Parameter mit der Komponente 1 bzw. untereinander dar.

¹⁷ Auch einzelne Parameter können interpretiert werden, wenn die Bedingung \times Parameter Interaktion mit berücksichtigt wird. Hier ein Beispiel für den Parameter HR: Aufgrund der positiven Faktorladung der HR wissen wir, dass die HR für die hoch beanspruchende Bedingungen S2- durch $a_{HR} \times a_{S2-}$ steigt und für die niedrig beanspruchende Bedingungen EnV+ durch $a_{HR} \times a_{EnV+}$ fällt. Ein umgekehrtes Muster, also ein Abfall für S2- und Anstieg für EnV+ zeigt sich für den Parameter MQSD mit seiner negativen Faktorladung a_{MQSD} .

(gestrichelte rote Linie) – aber auch deren Schwankungen über Bedingungen hinweg – unterscheiden (z. B. zeigt die Person 2 eine höhere durchschnittliche kardiale Aktivität bzw. Schwankung als die Person 5). Die in Abbildung 6.10 (oben) dargestellten Person-Faktorladungen lassen auf einen Blick erkennen, warum bisherige Versuche, die mentale Beanspruchung ohne die Berücksichtigung individueller Unterschiede vorherzusagen, zu eher mäßigem Erfolg führten: Denn während die NASA-RTLX-Komponente 1 sich nur gering zwischen den Personen unterscheidet (d. h. Personen schätzen ihre empfundene mentale Beanspruchung ähnlich ein), weist deren kardiale Aktivität sehr hohe Unterschiede zwischen den Personen auf. Die Relevanz dieser personenspezifischer Variabilität für die Vorhersagegüte der subjektiven mentalen Beanspruchung wird ebenfalls durch die Reduktion des $\text{val}\Delta R_{\text{PLS}}^2$ im Vergleich zum ΔR_{PLS}^2 in der Kriteriumsmatrix (siehe Tabelle 6.19) deutlich. Werden nämlich Personen vorhergesagt, die dem Komponentenmodell 1 unbekannt sind, dann verringert sich der Anteil erklärter Varianz um 16.4 %. Ist die Person hingegen bekannt, dann kann deren kardiale Aktivität mit dem personenspezifischen Faktor gewichtet werden und so zu einer besseren Vorhersagegüte der subjektiven mentalen Beanspruchung beitragen.

6.3.3 Forschungsfrage F1.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind in Tabelle 6.20 dargestellt.¹⁸ Hier zeigt sich ein bedeutsamer Zuwachs an Vorhersagegüte nur für die Komponente 1. Die Komponente 1 erklärt im Originalmodell 9.0 % der Varianz in den kardialen

¹⁸ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.86 [0.15, 1.79]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -7.37 [-26.0, 1.99]$.

Parametern und sagt 40.6 % der Varianz in der Trefferquote vorher. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 19.5 % der Varianz. Die Ladungen der Komponente sind in Abbildung 6.11 dargestellt.

Zur Interpretation der Werte wird die Dimension der Bedingungen (Abbildung 6.11, Mitte) betrachtet (tabellarische Darstellung siehe Anhang A-8). Da eine Leistungsmessung nur innerhalb der Balltracking-Bedingungen möglich war, erfolgt die Zentrierung der Leistung und der kardialen Aktivität nur innerhalb der Balltracking-Bedingungen. Folglich ist hier die Abszissenachse in Relation zu der Abszissenachse der NASA-RTLX-Faktorladungen (siehe Abbildung 6.10, Mitte) versetzt. Sie repräsentiert die durchschnittliche Trefferquote bzw. kardiale Aktivität über alle Balltracking-Aufgaben hinweg. Die Bedingungs-Faktorladungen der Trefferquoten-Komponente 1 (gestrichelte olivfarbene Linie) zeigen eine hohe Ähnlichkeit zu den Bedingungs-mittelwerten der Trefferquote (gestrichelte dunkelgraue Linie in Abbildung 6.11, Mitte). Damit ist die extrahierte Trefferquoten-Komponente 1 als ein Repräsen-

Tabelle 6.20: NPLS-Regression der Trefferquote auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR^2_{PLS}) und das kreuzvalidierte Modell ($\text{val}\Delta R^2_{\text{PLS}}$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	valSQR
0					267.46
1	9.01	6.48 [†]	4.58	19.51 [†]	177.07***
2	7.28	4.01 [†]	19.10	-5.41	188.97
3	9.15	7.07 [†]	14.43	-12.88	217.31
4	6.65	4.41 [†]	13.79	-9.17	237.48
5	2.41	1.32	12.11	-6.57	251.93

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 1.79$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 1.99$.

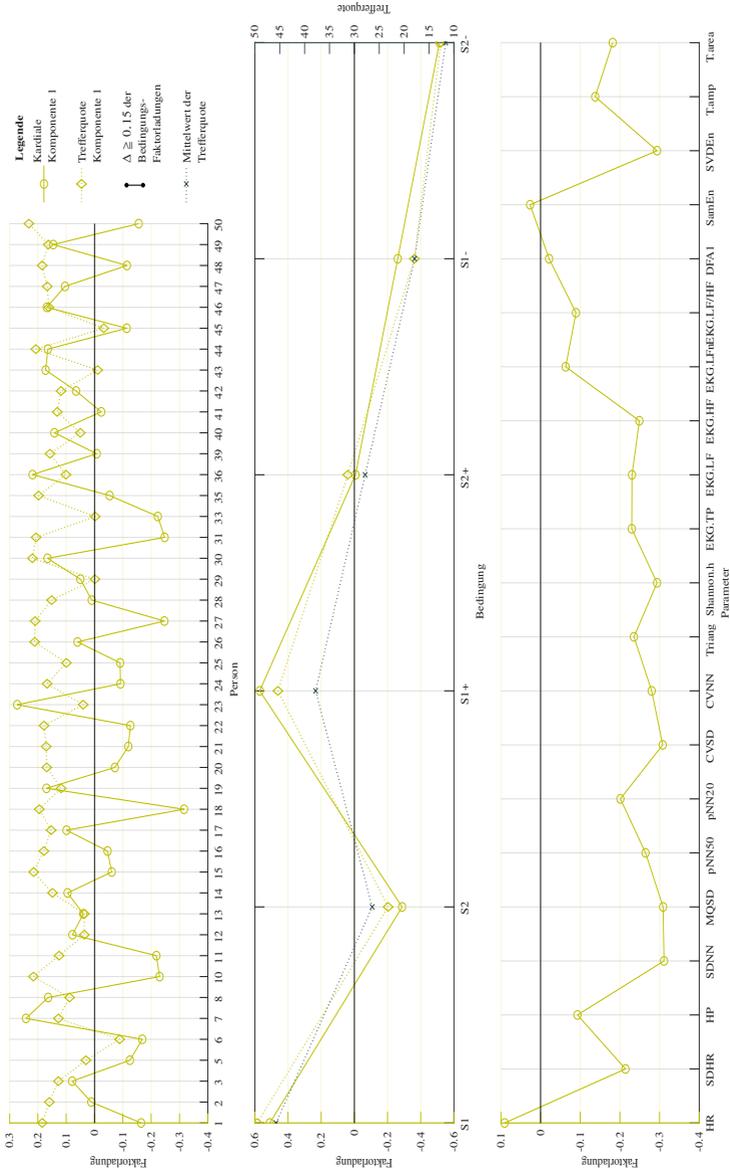


Abbildung 6.11: Faktorladungen α für die Treffferquote und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Treffferquote in den Bedingungen (Mitte) angegeben.

tant der leistungsbezogenen mentalen Beanspruchung zu interpretieren. Diese Trefferquoten-Komponente 1 kovariiert mit der kardialen Komponente 1 in der Bedingungsdimension und kann durch diese mit $B = -0.42$, $t(262) = -4.78$, $p < .001$ vorhergesagt werden. Die Bedingungs-Faktorladungen der kardialen Komponente 1 sind in Abbildung 6.11 (Mitte) als durchgezogene olivfarbene Linie abgebildet. Um den Zusammenhang zwischen der kardialen Komponente und der Trefferquoten-Komponente in Abbildung 6.11 sichtbar zu machen, wurden die Bedingungs-Faktorladungen der kardialen Komponente (zusammen mit den Person-Faktorladungen) invertiert.¹⁹ Negative Faktorladungswerte repräsentieren somit eine überdurchschnittliche und positive Faktorladungswerte eine unterdurchschnittliche kardiale Aktivität. In der Abbildung 6.11 (Mitte) ist zu erkennen, dass die invertierten Bedingungs-Faktorladungen der kardialen Komponente mit den Faktorladungen der Trefferquoten-Komponente übereinstimmen: Eine unterdurchschnittliche Trefferquote (bei hoher Aufgabenschwierigkeit oder negativem Leistungsfeedback) ist mit einer überdurchschnittlichen kardialen Aktivität assoziiert (und umgekehrt).

Die Parameter SDNN, MQSD und CVSD laden mit $|a| \geq .30$ bedeutsam auf der kardialen Komponente 1 (siehe Abbildung 6.11, unten). Marginal bedeutsame Faktorladungen ($|a| \geq .20$) besitzen die Parameter SDHR, pNN50, pNN20, CVNN, Triang, Shannon.h, EKG.TP, EKG.LF, EKG.HF und SVDEn. Für die Faktorladungen der Personen (siehe Abbildung 6.11, oben) ist zu erkennen, dass die Trefferquoten-Komponente deskriptiv eine geringere interindividuelle Varianz aufweist als die kardiale Komponente.

¹⁹ Eine Invertierung der Faktorladungen bedeutet eine Spiegelung an der Abszissenachse in Abbildung 6.11. Aufgrund der Dreidimensionalität (Person \times Bedingung \times Parameter) sind nur zwei Dimensionen gleichzeitig zu invertieren.

6.3.4 Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind für die Regression vom NASA-RTLX auf die elektrodermalen Parameter in der Tabelle 6.21 dargestellt.²⁰ Aus den Modellvergleichen ist zu folgern, dass nur die erste Komponente eine valide Vorhersage der NASA-RTLX-Werte bietet. Im Originalmodell erklärt diese Komponente 11.7 % der Varianz in den elektrodermalen Parametern und kann 81.0 % der Varianz in den NASA-RTLX-Werten vorhersagen. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 52.5 %. Die Ladungen der Komponente 1 sind in Abbildung 6.12 dargestellt. Eine tabellarische Darstellung der Bedingungs-Faktorladungen ist im Anhang A-11 zu finden.

Zunächst wird die Bedingungsdimension in Abbildung 6.12 (Mitte) betrachtet. Hier ist erkennbar, dass die aus der NASA-RTLX-Matrix extrahierte Komponente 1 (gestrichelte rote Linie) in ihren Bedingungs-Faktorladungen dem bereits bekannten Verlauf der NASA-RTLX-Mittelwerte (gestrichelte dunkelgraue Linie) folgt. Folglich kann diese Komponente als Repräsentant der subjektiven mentalen Beanspruchung interpretiert werden. In der Bedingungsdimension zeigt sich eine Zunahme der Faktorladungen der NASA-RTLX-Komponente um $B = 0.67$, $t(478) = 16.16$, $p < .001$ bei Zunahme der elektrodermalen Faktorladungen. Die Faktorladungen der NASA-RTLX-Komponente 1 sind in Abbildung 6.12 (Mitte) als durchgezogene rote Linie abgebildet. Bedeutsame Abweichungen zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der NASA-RTLX-Komponente gibt es keine. Nur für die

²⁰ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.67 [0.19, 1.37]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{KLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -5.62 [-16.65, 0.92]$.

Tabelle 6.21: NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	valSQR
0					513.28
1	11.67	10.15 [†]	80.97	52.50 [†]	205.22***
2	6.47	4.19 [†]	7.48	-2.71	216.91
3	5.11	2.91 [†]	3.47	-0.99	221.19
4	4.05	1.10	1.83	-0.19	222.00
5	4.20	0.27	1.75	-1.63	229.04

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $val\Delta R^2 > 1.37$, der Kriteriumsmatrix: [†] $val\Delta R^2 > 0.92$.

Bedingung S2- zeigt sich ein marginaler Unterschied: In Abbildung 6.12 (Mitte) ist zu erkennen, dass es bei der elektrodermalen Komponente im Vergleich zur NASA-RTLX-Komponente keinen Anstieg von der Bedingung S2- zur Bedingung S1- gibt.

Die Faktorladungen der Parameter (Abbildung 6.12, unten) bilden hier ein erstes elektrodermales Profil der subjektiven mentalen Beanspruchung ab. Die elektrodermale Komponente 1 ist am stärksten mit dem Parameter EDR.N ($a = .43$) assoziiert, gefolgt von den Parametern EDL.M, EDR.rect/2, EDR.rect/c und EDR.peaks mit $|a| > .30$. Marginal bedeutsam sind die Parameter EDL.D, EDR.rise.t, EDR.rise.v, EDA.HF und Phasic.SD mit $|a| \geq .20$. Die Faktorladungen der Personen sind in Abbildung 6.12 (oben) zu sehen. Die Person-Faktorladungen des NASA-RTLX (gestrichelte rote Linie) weisen kaum interindividuelle Unterschiede auf, wohingegen die elektrodermalen Person-Faktorladungen (durchgezogene rote Linie) hohe interindividuelle Unterschiede zeigen.

6.3.5 Forschungsfrage F2.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?

In Tabelle 6.22 zeigt sich, dass von den fünf Komponenten nur die erste Komponente die Trefferquote bedeutsam vorhersagt.²¹ Folglich soll Komponente 1 weiter analysiert werden. Sie erklärt im Originalmodell 8.6 % der Varianz in der elektrodermalen Parametermatrix und sagt 41.9 % der Varianz in der Trefferquote vorher. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 10.3 %.

Tabelle 6.22: NPLS-Regression der Trefferquote auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					326.11
1	8.61	6.38 [†]	41.91	10.32 [†]	215.22***
2	10.19	8.48 [†]	16.88	-7.94	234.28
3	5.01	2.76 [†]	15.29	-9.63	257.39
4	6.91	4.72 [†]	13.21	-4.75	268.79
5	5.85	3.53 [†]	12.69	-10.66	294.37

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 2.37$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 1.57$.

Zunächst wird die Dimension der Bedingungen (Abbildung 6.13, Mitte) analysiert. Eine Tabelle der Bedingungen-Faktorladungen findet sich im Anhang A-12.

²¹ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 1.19[0.19, 2.37]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -8.91[-29.51, 1.57]$.

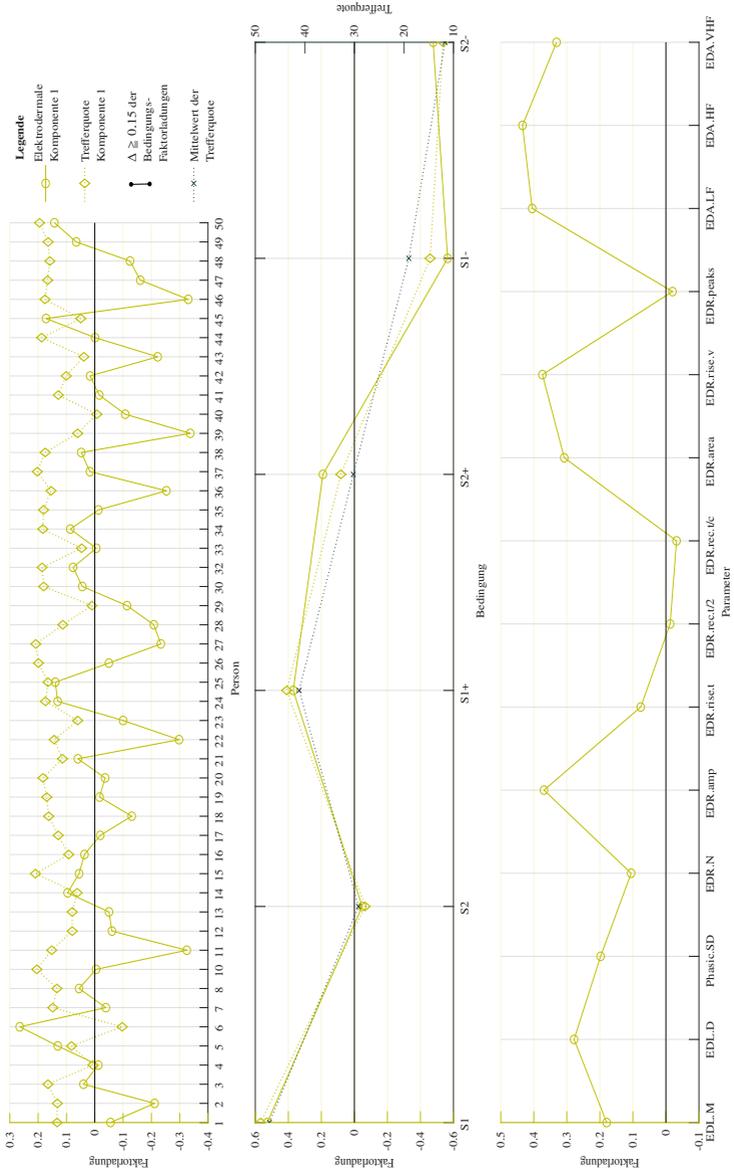


Abbildung 6.13: Faktorladungen α für die Trefferquote und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.

Es ist zu beachten, dass die Abszissenachse die durchschnittliche Trefferquote bzw. durchschnittliche elektrodermale Aktivität über alle Balltracking-Aufgaben (und nicht über alle Bedingungen) hinweg darstellt. Die Faktorladungen der Trefferquote (gestrichelte olivfarbene Linie in Abbildung 6.13, Mitte) sind dabei fast identisch zu den Mittelwerten der Trefferquote in den Bedingungen (gestrichelte dunkelgraue Linie in Abbildung 6.13, Mitte). Somit ist die extrahierte Trefferquoten-Komponente als ein Repräsentant der leistungsbezogenen mentalen Beanspruchung zu interpretieren. In der Bedingungsdimension kovariert die Trefferquoten-Komponente 1 mit der elektrodermalen Komponente 1. Dabei sagt die elektrodermale Komponente 1 die Trefferquoten-Komponente 1 mit $B = -0.54$, $t(286) = -6.04$, $p < .001$ vorher. Die Bedingungs-Faktorladungen der elektrodermalen Komponente 1 sind in Abbildung 6.13 (Mitte) als durchgezogene olivfarbene Linie abgebildet.²² Die (invertierten) elektrodermalen Bedingungs-Faktorladungen stimmen mit den Bedingungs-Faktorladungen der Trefferquoten-Komponente überein, sodass es keine bedeutsamen Unterschiede zwischen den Faktorladungen gibt.

Die Faktorladungen der Parameter (Abbildung 6.13, unten) lassen darauf schließen, dass die Parameter EDA.LF und EDA.HF mit $|a| \geq .40$, gefolgt von EDR.rise.v, EDR.amp, EDR.area und EDA.VHF mit $|a| \geq .30$ bedeutsam auf der elektrodermalen Komponente 1 laden. Marginal bedeutsame Faktorladungen ($|a| \geq .20$) besitzen zudem die Parameter EDL.D und Phasic.SD. Auf Ebene der Personen (siehe Abbildung 6.13, oben) ist deskriptiv eine geringere interindividuelle Varianz der Trefferquoten-Komponente (gestrichelte olivfarbene Linie) im Vergleich zur elektrodermalen Komponente (durchgezogene olivfarbene Linie) festzustellen.

²² Es ist zu beachten, dass die Person- und Bedingungs-Faktorladungen der elektrodermalen Komponente in Abbildung 6.13 invertiert sind: Negative Werte bedeuten eine über- und positive Werte eine unterdurchschnittliche elektrodermale Aktivität.

Tabelle 6.23: NPLS-Regression der Erregung auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					493.66
1	23.89	22.33 [†]	18.04	-18.04	456.81
2	4.03	1.52 [†]	31.00	15.40 [†]	397.22 ^{***}
3	3.80	1.51 [†]	13.51	-5.08	416.90
4	4.33	1.55 [†]	8.21	-2.66	427.20
5	2.61	0.81	9.25	-7.44	455.99

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.94$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 1.10$.

6.3.6 Forschungsfrage F3.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung zuverlässig repräsentieren?

Für die NPLS-Regression wurde eine Person aus der Analyse ausgeschlossen, da diese keine Variation in der subjektiven Erregung besaß. In Tabelle 6.23 ist erkennbar, dass nur die Komponente 2 einen bedeutsamen Beitrag zur Varianzaufklärung in der kardialen Parametermatrix leistet und gleichzeitig die Erregungs-Werte des SAM im kreuzvalidierten Modell bedeutsam vorhersagt.²³ Im Originalmodell kann die Komponente 2 nur einen geringen Varianzanteil von 4.0 % in den kardialen Parametern erklären und 31.0 % der Varianz in den Erregungs-Werten vorhersagen. In der kreuzvalidierten Vorhersage

²³ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.48 [0.16, 0.94]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -4.84 [-13.89, 1.10]$.

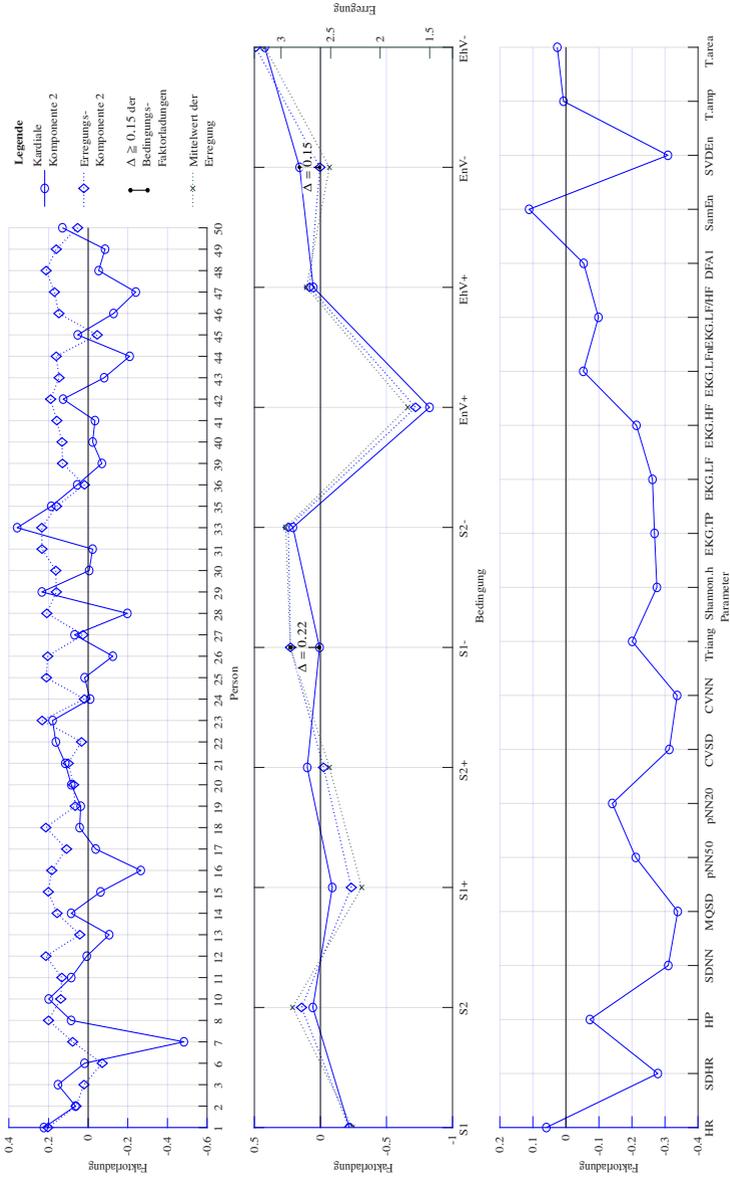


Abbildung 6.14: Faktorladungen α für die Erregung und die kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Erregung in den Bedingungen (Mitte) angegeben.

verringert sich der Anteil erklärter Varianz auf 15.4 %. Die Faktorladungen der Komponente 2 sind in Abbildung 6.14 dargestellt.

Die Erregungs-Komponente 2 folgt in ihren Bedingungs-Faktorladungen (siehe Abbildung 6.14, Mitte, gestrichelte blaue Linie) dem bereits bekannten Verlauf der Erregungs-Mittelwerte (siehe gestrichelte dunkelgraue Linie in Abbildung 6.14, Mitte) und ist somit als Repräsentant der subjektiven Erregung zu interpretieren. Die Erregungs-Komponente 2 kovariiert mit der kardialen Komponente 2 (Abbildung 6.14, Mitte, durchgezogene blaue Linie) in der Bedingungsdimension, $B = 0.49$, $t(428) = 5.04$, $p < .001$. Eine tabellarische Darstellung der Bedingungs-Faktorladungen befindet sich im Anhang A-9. Die Abbildung 6.14 (Mitte) zeigt für die Bedingung S1- eine bedeutsame Abweichung zwischen der Faktorladung der kardialen Komponente und der Faktorladung der Erregungs-Komponente. Dies kann darauf zurückgeführt werden, dass der Haupteffekt der Aufgabenschwierigkeit innerhalb der negativen Feedbackbedingungen für die Erregungs-Komponente nicht vorliegt. Im Gegensatz dazu ist für die kardiale Komponente ein Anstieg zwischen den Schwierigkeitsstufen 1 und 2 bei negativem Leistungsfeedback gegeben. In Abbildung 6.14 (Mitte) ist deskriptiv zu erkennen, dass die kardiale Komponente innerhalb der Feedbackbedingungen den Haupteffekt des Aufgabenfeedbacks sowie den ordinalen Interaktionseffekt nicht abbildet. Eine marginal bedeutsame Abweichung zwischen den Komponenten liegt zudem für die Faktorladung der Bedingung EnV- vor. In dieser Bedingung liegt die kardiale Aktivität unterhalb der subjektiven Erregung.

Auf der kardialen Komponente 2 laden die Parameter SDNN, MQSD, CVSD, CVNN und SVDEn bedeutsam ($|a| > .30$; siehe Abbildung 6.14, unten). Weitere, marginal bedeutsame Parameter sind SDHR, pNN50, CVSD, Triang, Shannon.h, EKG.TP, EKG.LF und EKG.HF mit $|a| \geq .20$. Für die Faktorladungen der Personen (Abbildung 6.14, oben) ist zu erkennen, dass die Erregungs-Komponente 2 eine geringere interindividuelle Varianz aufweist als die kardiale Komponente 2.

Tabelle 6.24: NPLS-Regression der Valenz auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR^2_{PLS}) und das kreuzvalidierte Modell ($\text{val}\Delta R^2_{\text{PLS}}$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	valSQR
0					481.96
1	22.97	21.15 [†]	12.32	-6.36	421.19*
2	4.19	1.88 [†]	33.20	25.31 [†]	320.96***
3	5.29	2.28 [†]	10.29	-3.53	334.92
4	2.40	0.95	10.84	-3.52	348.88
5	2.64	0.97	8.00	-8.00	380.56

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.94$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 1.10$.

6.3.7 Forschungsfrage F3.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Valenz zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind in Tabelle 6.24.²⁴ Eine bedeutsame Verbesserung der Vorhersagegüte der subjektiven Valenz gewährleistet nur die Komponente 2. Diese klärt im Originalmodell nur 4.2 % der Varianz in den kardialen Parametern auf und sagt bis zu 33.2 % der Varianz in der subjektiven Valenz vorher. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 25.3 %. Die Faktorladungen der Komponente 2 sind in Abbildung 6.15 dargestellt.

²⁴ Der mittlere zufällige Zuwachs $\text{rand}\Delta R^2_{\text{PLS}}$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.48[0.16, 0.94]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R^2_{\text{PLS}}$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -4.84[-13.89, 1.10]$.

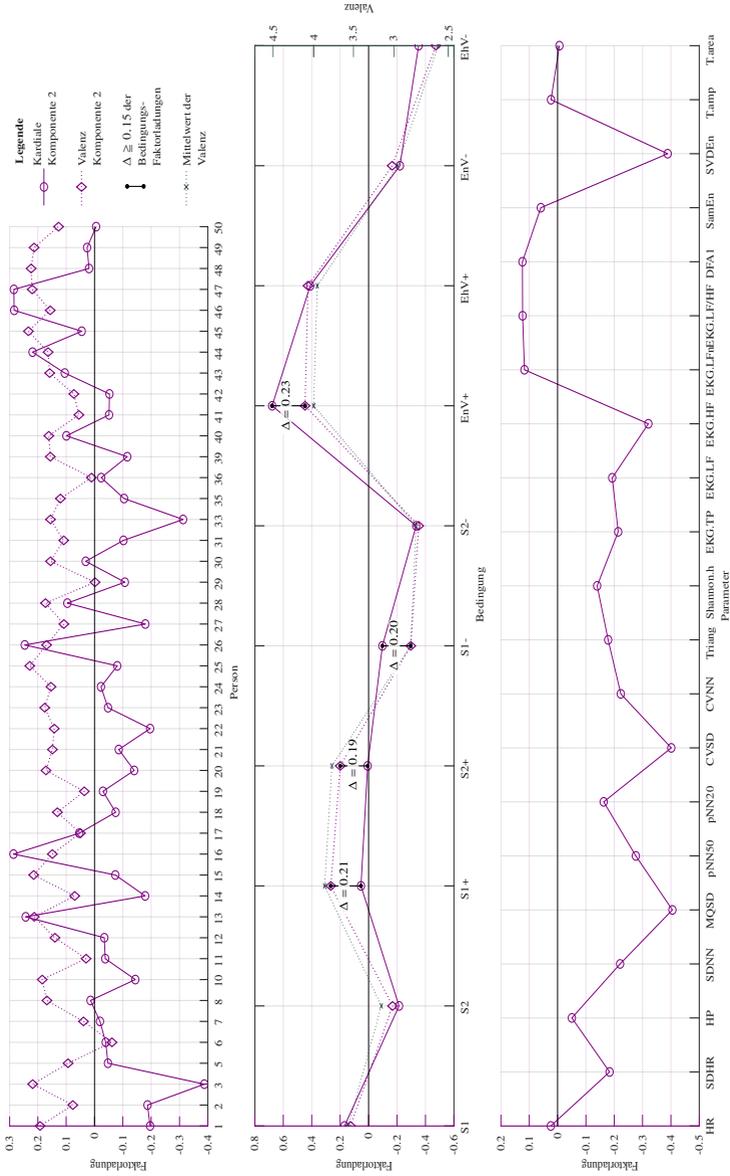


Abbildung 6.15: Faktorladungen α für die Valenz und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Valenz in den Bedingungen (Mitte) angegeben.

Zur Interpretation wird die Dimension der Bedingungen (siehe Abbildung 6.15, Mitte) herangezogen (Tabelle der Bedingungs-Faktorladungen siehe Anhang A-10). Um den Zusammenhang zwischen der kardialen Komponente und der Valenz-Komponente in Abbildung 6.15 sichtbar zu machen, wurden die Bedingungs-Faktorladungen der kardialen Komponente (zusammen mit den Person-Faktorladungen) invertiert. Negative Faktorladungswerte repräsentieren somit eine überdurchschnittliche und positive Faktorladungswerte eine unterdurchschnittliche kardiale Aktivität. In Abbildung 6.15 (Mitte) sind die Faktorladungen der Valenz-Komponente 2 als violett gestrichelte Linie abgebildet. In Abbildung 6.15 (Mitte) ist erkennbar, dass die Komponente 2 (gestrichelte violette Linie) ein nahezu identisches Verlaufsmuster zu den Valenz-Mittelwerten (gestrichelte dunkelgraue Linie) zeigt. Folglich wird die Valenz-Komponente 2 als Repräsentant der subjektiven Valenz interpretiert. Die Valenz-Komponente 2 kovariiert mit der kardialen Komponente 2 in der Bedingungsdimension und kann durch diese mit $B = -0.59$, $t(438) = -6.07$, $p < .001$ vorhergesagt werden. Die Bedingungs-Faktorladungen der kardialen Komponente 2 sind in Abbildung 6.15 (Mitte) als durchgezogene violette Linie zu sehen. Die bedeutsamen und marginalen Abweichungen der Faktorladungen für die Bedingungen S1+, S2+ und S1- zeigen auf, dass die Unterschiede zwischen positivem und negativem Feedback bei der kardialen Komponente 2 geringer sind als bei der Valenz-Komponente 2. Erst die Bedingung S2- zeigt einen deutlichen Abfall der kardialen Aktivität im Vergleich zu den anderen Balltrackingbedingungen. Folglich scheint der Haupteffekt des Aufgabenfeedbacks für die kardiale Komponente 2 geringer ausgeprägt zu sein als für die Valenz-Komponente 2. Im Gegensatz dazu weicht die Faktorladung der Bedingung EnV+ in der kardialen Komponente 2 stärker vom Durchschnitt ab als die Faktorladung der Valenz-Komponente 2.

Die kardiale Komponente 2 ist am stärksten mit den Parametern MQSD, CVSD, EKG.HF und SVDEn assoziiert ($|a| > .30$). Neben diesen drei bedeutsamen Parametern sind noch die Faktorladungen der Parameter SDNN, pNN50, CVNN und EKG.TP mit $|a| \geq .20$ marginal bedeutsam. Für die Faktorladungen

der Personen (siehe Abbildung 6.15, oben) ist insbesondere für die kardiale Komponente 2 eine hohe Variabilität erkennbar. Die im Vergleich dazu geringere Variabilität der Valenz-Komponente 2 zeigt auf, dass die subjektive Valenz sich zwischen den Personen geringer unterscheidet als die kardiale Aktivität.

6.3.8 Forschungsfrage F4.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Erregung zuverlässig repräsentieren?

Für die NPLS-Regression wurde eine Person aus der Analyse ausgeschlossen, da diese keine Variation in der subjektiven Erregung besaß. Die Kreuzvalidierung in Tabelle 6.25 zeigt auf, dass nur die Komponente 1 einen bedeutsamen Beitrag

Tabelle 6.25: NPLS-Regression der Erregung auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	valSQR
0					452.35
1	8.91	5.94 [†]	27.98	8.16 [†]	388.47**
2	9.68	7.97 [†]	24.15	-7.09	418.46
3	3.49	1.19	10.64	-2.58	429.36
4	4.25	1.66 [†]	7.11	-4.28	447.48
5	4.86	2.33 [†]	8.15	-1.28	452.88

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $val\Delta R^2 > 1.37$, der Kriteriumsmatrix: [†] $val\Delta R^2 > 0.92$.

zur Verbesserung der Vorhersagegüte leistet.²⁵ Im Originalmodell kann die Komponente 1 einen Varianzanteil von 8.9 % in den elektrodermalen Parametern erklären und 28.0 % der Varianz in der subjektiven Erregung vorhersagen. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 8.2 %. Die Ladungen der Komponente 1 sind in Abbildung 6.16 dargestellt.

Zur Interpretation der Ergebnisse wird die Bedingungsdimension in Abbildung 6.16 (Mitte) herangezogen (tabellarische Darstellung siehe Anhang A-13). Die Faktorladungen der Erregungs-Komponente 1 (gestrichelte blaue Linie) folgen dabei dem Verlauf der Erregungs-Mittelwerte (gestrichelte dunkelgraue Linie). Diese sind jedoch nicht identisch, da innerhalb der IAPS-Bedingungen für die Erregungs-Komponente 1 – im Vergleich zu den Mittelwerten der subjektiven Erregung – ein höherer Unterschied zwischen der niedrigen (EnV+ und EnV-) und hohen (EhV+ und EhV-) Bilderset-Erregung zu sehen ist. Dies bestätigt die Interpretation der Erregungs-Komponente 1 als einen Repräsentanten der subjektiven Erregung. In der Bedingungsdimension kovariiert die elektrodermale Komponente 1 (durchgezogene blaue Linie in Abbildung 6.16, Mitte) mit der Erregungs-Komponente 1. Dabei kann die elektrodermale Komponente 1 die Erregungs-Komponente 1 mit $B = 0.45$, $t(468) = 5.56$, $p < .001$ vorhersagen. In Abbildung 6.16 (Mitte) ist lediglich für die Bedingung EnV+ ein bedeutsamer Unterschied zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der Erregungs-Komponente zu verzeichnen. Bei dieser Bedingung befindet sich die elektrodermale Aktivität im durchschnittlichen und die subjektive Erregung im unterdurchschnittlichen Wertebereich.

Die Faktorladungen der Parameter (Abbildung 6.16, unten) bilden hier ein erstes Profil der subjektiven Erregung ab. Die Parameter EDR.amp, EDR.rise.v und EDA.LF ($a \geq .40$) laden am höchsten auf der elektrodermalen Komponente. Weitere bedeutsame Faktorladungen haben die Parameter EDR.area

²⁵ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.67 [0.19, 1.37]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -5.62 [-16.65, 0.92]$.

und EDA.HF mit $a \geq .30$. Marginale Bedeutsamkeit besitzen die Parameter EDA.VHF und Phasic.SD mit $|a| \geq .20$. Die Faktorladungen der Personen (siehe Abbildung 6.16, oben) zeigen sowohl für die Faktorladungen der Erregungs-Komponente 1 als auch der elektrodermalen Komponente 1 starke interindividuelle Unterschiede.

6.3.9 Forschungsfrage F4.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die emotionale Beanspruchung als subjektive Valenz zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung befinden sich in Tabelle 6.26.²⁶ Einen bedeutsamen Beitrag zur Vorhersagegüte leistet die Komponente 1. Diese erklärt im Originalmodell 7.9 % der Varianz in den elektrodermalen Parametern und kann 33.5 % der Varianz in der subjektiven Valenz vorhersagen. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz drastisch auf 4.2 %. Die Faktorladungen der Komponente sind in Abbildung 6.17 dargestellt.

Die extrahierte Valenz-Komponente 1 folgt in ihren Bedingungs-Faktorladungen (siehe Abbildung 6.17, Mitte, gestrichelte violette Linie) dem Verlauf der Valenz-Mittelwerte (gestrichelte dunkelgraue Linie in Abbildung 6.17, Mitte) und kann somit als Repräsentant der subjektiven Valenz interpretiert werden. In der Bedingungsdimension kann die elektrodermale Komponente 1 (Abbildung 6.17, Mitte, durchgezogene violette Linie) die Komponente 1 der subjektiven Valenz mit $B = -0.53$, $t(478) = -6.80$, $p < .001$ vorhersagen.²⁷

²⁶ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.67[0.19, 1.37]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -5.62[-16.65, 0.92]$.

²⁷ Es ist zu beachten, dass die Bedingungs- und Person-Faktorladungen der elektrodermalen Komponente in Abbildung 6.17 invertiert sind: Negative Werte bedeuten eine über- und positive Werte eine unterdurchschnittliche elektrodermale Aktivität.

Tabelle 6.26: NPLS-Regression der Valenz auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	valSQR
0					559.30
1	7.91	6.08 [†]	33.51	4.18 [†]	413.96**
2	10.53	8.56 [†]	18.33	0.49	411.86
3	4.58	1.98 [†]	10.42	-1.4	417.9
4	3.99	1.48 [†]	8.73	-2.14	427.12
5	5.56	2.36 [†]	7.3	-4.74	447.61

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $val\Delta R^2 > 1.37$, der Kriteriumsmatrix: [†] $val\Delta R^2 > 0.92$.

Gemäß Abbildung 6.17 (Mitte) weichen die Faktorladungen der Bedingungen EnV+ sowie EhV+ bedeutsam und die Faktorladungen der Bedingungen S2- sowie EhV- marginal voneinander ab (Tabelle der Bedingungs-Faktorladungen siehe Anhang A-14). Dabei liegen die Faktorladungen der elektrodermalen Komponente (im Vergleich zur Valenz-Komponente) bei den Bedingungen S2-, EnV+ und EhV+ näher am Durchschnitt (Abszissenachse). Hingegen ist die Faktorladung der elektrodermalen Komponente für die Bedingung EhV- weiter vom Durchschnitt entfernt als die Faktorladung der Valenz-Komponente.

Die Parameter-Faktorladungen sind in Abbildung 6.17 (unten) abgebildet. Bedeutsamen Ladungen auf der elektrodermalen Komponente besitzen die Parameter EDR.amp, EDR.rise.v EDA.LF und EDA.HF mit $a > .40$ gefolgt von Phasic.SD mit $a = .31$. Eine marginale Bedeutsamkeit haben die Parameter EDR.area und EDR.VHF. Die Faktorladungen der Personen (siehe Abbildung 6.17, oben) zeigen hohe interindividuelle Unterschiede sowohl für die Valenz-Komponente als auch für die elektrodermale Komponente.

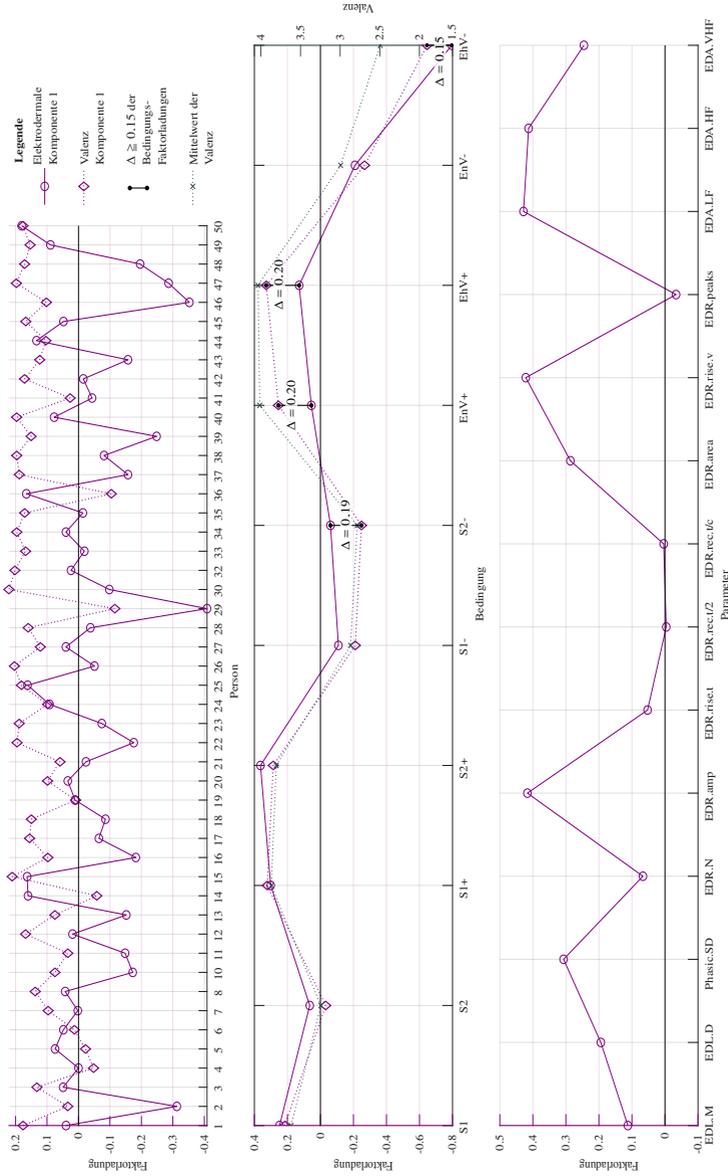


Abbildung 6.17: Faktorladungen α für die Valenz und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Valenz in den Bedingungen (Mitte) angegeben.

6.3.10 Forschungsfrage F5.1: Können die kardialen Profile der mentalen und emotionalen Beanspruchung bedeutsam voneinander abgegrenzt werden?

Zur Bewertung, welche der extrahierten kardialen Komponenten identische, ähnliche oder unterschiedliche Konstrukte abbilden, werden die Kongruenzkoeffizienten in Tabelle 6.27 betrachtet. Die Kongruenzkoeffizienten sind jeweils für die Faktorladungen der Personen, Bedingungen und Parameter aufgetragen. Relevant für die Beurteilung der Ähnlichkeit kardialer Profile ist der Kongruenzkoeffizient der Parameter-Faktorladungen. Die Kongruenzkoeffizienten der Personen und Bedingungen liefern zudem einen Hinweis darauf, ob eine Ähnlichkeit nur auf Ebene der Parameter-Faktorladungen besteht oder ob die Ähnlichkeit alle Ebenen der Komponente umfasst. In Tabelle 6.27 (letzte Spalte) ist eine Ähnlichkeit zwischen fast allen kardialen Profilen zu erkennen. Nur das kardiale Profil des NASA-RTLX und der Erregung sind in ihren Parameter-Faktorladungen unähnlich. Jedoch liegt deren Kongruenz ($C = .84$) knapp unter

Tabelle 6.27: Kongruenz zwischen der kardialen Komponente 1 des NASA-RTLX (Na), der kardialen Komponente 1 der Trefferquote (Tq), der kardialen Komponente 2 der Erregung (Er) und der kardialen Komponente 2 der Valenz (Va). Die Komponenten werden auf Ebene der Personen, Bedingungen und Parameter verglichen

Komponentenvergleich	Kongruenzkoeffizient C		
	Person	Bedingung	Parameter
Na – Tq	–.11	–.14	.87
Na – Er	.08	.11	.84
Na – Va	–.29	–.20	.89
Tq – Er	–.85	–.38	.95
Tq – Va	.83	.39	.88
Er – Va	–.84	–.70	.90

Identisch $|C| > .95$, **ähnlich** $.85 \leq |C| \leq .95$, **unterschiedlich** $|C| < .85$.

der Ähnlichkeitsschwelle von .85 und bietet somit eine schwache Differenzierung. Zudem zeigt das Profil der Erregung eine sehr hohe Ähnlichkeit zum kardialen Profil der Valenz und dem Profil der Trefferquote. So liegt es auf Ebene der Parameter mit $C = .95$ knapp unter der Identitätsgrenze zur kardialen Komponente der Trefferquote. Diese Ähnlichkeit wird auf Ebene der Personen mit $C = -.85$ bestätigt. Nur auf Ebene der Bedingungen besteht geringe Ähnlichkeit zwischen den Komponenten der Erregung und Trefferquote. Auch beim Vergleich der kardialen Komponenten des NASA-RTLX, der Trefferquote und der Valenz ist eine Ähnlichkeit auf Ebene der Parameter zu verzeichnen (siehe Tabelle 6.27, letzte Spalte).

Welche kardialen Parameter für die Ähnlichkeit bzw. Unähnlichkeit der Profile verantwortlich sind, zeigt Tabelle 6.28 anhand der Differenzen zwischen den Parameter-Faktorladungen. Gemäß Tabelle 6.28 unterscheiden sich die vier extrahierten Profile in nur vier von insgesamt 21 Parametern bedeutsam voneinander: HR, HP, EKG.LF/HF und T.area. Die Parameter HR und HP erlauben es, das kardiale Profil des NASA-RTLX vom kardialen Profilen der Trefferquote, dem Profil der Erregung und dem Profil der Valenz abzugrenzen. Zusätzlich kann das Profil des NASA-RTLX zum Profil der Erregung durch den Parameter EKG.LF/HF abgegrenzt werden. Eine bedeutsame Unterscheidbarkeit zwischen den kardialen Profilen der Trefferquote, der Erregung und der Valenz ist jeweils nur durch einen einzigen Parameter gegeben: EKG.LF/HF (Valenz zu Erregung und Valenz zu Trefferquote) oder T.area (Erregung zu Trefferquote). Es finden sich zudem zahlreiche marginale Ladungsunterschiede zwischen den kardialen Profilen für die Parameter (siehe Tabelle 6.28). Sie alle lassen jedoch nur eine marginale Differenzierung zu.

Einen Überblick über die Relevanz der kardialen Parameter für die jeweiligen kardialen Profile bietet Tabelle 6.29. Auch hier ist die Ähnlichkeit der extrahierten Profile sowie eine fehlende Einfachstruktur²⁸ erkennbar. Viele Parameter

²⁸ Bei einer Einfachstruktur liegen für gewisse Parameter hohe Faktorladungen auf einer Komponente und gleichzeitig niedrige Faktorladungen für alle anderen Komponenten vor.

Tabelle 6.28: Differenz zwischen den Parameter-Faktorladungen der kardialen Komponente 1 des NASA-RTLX (Na), der kardialen Komponente 1 der Trefferquote (Tq), der kardialen Komponente 2 der Erregung (Er) und der kardialen Komponente 2 der Valenz (Va)

Parameter	Na-Tq	Na-Er	Na-Va	Tq-Er	Tq-Va	Er-Va
HR	.19	.22	.26	.03	.06	.04
SDHR	.13	.19	.10	.06	-.03	-.09
HP	-.21	-.23	-.25	-.02	-.04	-.02
SDNN	.05	.05	-.04	.00	-.09	-.09
MQSD	-.01	.02	.09	.03	.10	.07
pNN50	.00	-.05	.01	-.06	.01	.06
pNN20	-.10	-.16	-.14	-.06	-.04	.02
CVSD	.02	.03	.12	.01	.09	.09
CVNN	.10	.15	.04	.06	-.06	-.11
Triang	.08	.05	.03	-.04	-.06	-.02
Shannon.h	.09	.07	-.06	-.02	-.15	-.14
EKG.TP	-.02	.02	-.03	.04	-.01	-.05
EKG.LF	.00	.03	-.04	.03	-.04	-.07
EKG.HF	-.05	-.09	.02	-.03	.08	.11
EKG.LFn	.16	.15	-.02	-.01	-.18	-.17
EKG.LF/HF	.19	.20	-.03	.01	-.22	-.22
DFA1	.12	.15	-.03	.03	-.15	-.18
SamEn	-.07	-.15	-.09	-.08	-.02	.05
SVDEn	.01	.03	.11	.02	.10	.08
T.amp	.11	-.04	-.05	-.15	-.17	-.02
T.area	.15	-.06	-.03	-.21	-.18	.03

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, **marginal** $|a_{\Delta}| \geq .15$, nicht bedeutsam $|a_{\Delta}| < .15$.

laden gleichzeitig auf den kardialen Profilen des NASA-RTLX, der Trefferquote, der Erregung und der Valenz. Der Parameter EKG.LF/HF wurden zwar als bedeutsam zur Differenzierung mehrerer kardialer Profile (NASA-RTLX zu Erregung, Valenz zu Trefferquote und Valenz zu Erregung) identifiziert, jedoch läßt dieser Parameter auf keinem der Profile bedeutsam oder marginal (siehe

Tabelle 6.29: Parameter-Faktorladungen a für die kardialen Profile des NASA-RTLX, der Trefferquote, der Erregung und der Valenz

Parameter	Kardiale Komponente			
	NASA-RTLX	Trefferquote	Erregung	Valenz
HR	.28	.09	.06	.02
SDHR	-.09	-.22	-.28	-.18
HP	-.30	-.09	-.07	-.05
SDNN	-.26	-.31	-.31	-.22
MQSD	-.32	-.31	-.34	-.41
pNN50	-.26	-.27	-.21	-.28
pNN20	-.30	-.20	-.14	-.16
CVSD	-.28	-.31	-.31	-.40
CVNN	-.19	-.28	-.34	-.22
Triang	-.15	-.24	-.20	-.18
Shannon.h	-.20	-.29	-.28	-.14
EKG.TP	-.25	-.23	-.27	-.21
EKG.LF	-.23	-.23	-.26	-.19
EKG.HF	-.30	-.25	-.21	-.32
EKG.LFn	.10	-.06	-.05	.12
EKG.LF/HF	.10	-.09	-.10	.12
DFA1	.10	-.02	-.05	.12
SamEn	-.04	.03	.11	.06
SVDen	-.28	-.29	-.31	-.39
T.amp	-.03	-.14	.01	.02
T.area	-.04	-.18	.03	-.01

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

Tabelle 6.29). Die Bedeutsamkeit für die Differenzierung kommt lediglich durch das invertierte Ladungsmuster zustande: positive Faktorladung auf den kardialen Profilen des NASA-RTLX ($a = .10$) und Valenz ($a = .12$), jedoch negative Faktorladung auf den kardialen Profilen der Erregung ($a = -.10$) und der Trefferquote ($a = -.09$). Gleiches gilt für den Parameter T.area. Dieser zeigt zwar eine bedeutsame Differenzierung zwischen dem Profil des NASA-RTLX und der Trefferquote, jedoch keine bedeutsame oder marginale Faktorladung auf einem der Profile (siehe Tabelle 6.29).

Alleinig die Parameter HR und HP zeigen eine Einfachstruktur der Faktorladungen. Diese laden hoch auf dem kardialen Profil des NASA-RTLX und gering auf den kardialen Profilen der Trefferquote, der Erregung und der Valenz. Um zu prüfen, ob die Einfachstruktur der HR und HP zu einer spezifischen Vorhersage der mentalen Beanspruchung bzw. der emotionalen Beanspruchung führt, werden gekreuzte Regressionen gerechnet. Regressiert wird die subjektive mentale Beanspruchung (NASA-RTLX) auf die kardiale Komponente der emotionalen Beanspruchung (Valenz) und die subjektive Valenz auf die kardiale Komponente der mentalen Beanspruchung (NASA-RTLX). Die kardiale Komponente der Valenz kann die subjektive mentale Beanspruchung bedeutsam in der Bedingungsdimension vorhersagen, $B = -0.47$, $t(438) = -5.24$, $p < .001$. Hierbei werden 22.4 % der Varianz im NASA-RTLX durch die kardiale Komponente erklärt. Im Gegensatz dazu ist die Vorhersage der subjektiven Valenz durch die kardiale Komponente des NASA-RTLX (in der Bedingungsdimension) nicht bedeutsam möglich, $B = -0.03$, $t(438) = -0.30$, $p = 1.0$. Der Anteil erklärter Varianz liegt hierbei unter einem Prozent.

6.3.11 Forschungsfrage F5.2: Können die elektrodermalen Profile der mentalen und emotionalen Beanspruchung bedeutsam voneinander abgegrenzt werden?

Die Kongruenzkoeffizienten in Tabelle 6.30 geben Auskunft darüber, welche elektrodermalen Komponenten identische, ähnliche oder unterschiedliche Konstrukte – jeweils auf Ebene der Personen, Bedingungen und Parameter – abbilden. Relevant ist hierbei vor allem der Kongruenzkoeffizient der Parameterfaktorladungen. Für die vier extrahierten elektrodermalen Komponenten zeigt sich nur für die Komponente des NASA-RTLX eine deutliche Abgrenzung zu den restlichen drei elektrodermalen Komponenten. Mit $|C| \leq .64$ unterscheidet sich die elektrodermale Komponente des NASA-RTLX auf Ebene der Personen, Bedingungen und Parameter zu den restlichen elektrodermalen Komponenten (siehe Tabelle 6.30). Im Gegensatz dazu sind die elektrodermale Komponente der Trefferquote, die elektrodermale Komponente der Erregung und die

Tabelle 6.30: Kongruenz zwischen der elektrodermalen Komponente 1 des NASA-RTLX (Na), der elektrodermalen Komponente 1 der Trefferquote (Tq), der elektrodermalen Komponente 1 der Erregung (Er) und der elektrodermalen Komponente 1 der Valenz (Va). Die Komponenten werden auf Ebene der Personen, Bedingungen und Parameter verglichen

Komponentenvergleich	Kongruenzkoeffizient C		
	Person	Bedingung	Parameter
Na – Tq	–.13	–.08	.59
Na – Er	–.63	–.30	.64
Na – Va	.46	.17	.56
Tq – Va	–.62	–.58	.97
Er – Va	.72	.61	.98
Er – Tq	–.94	–.88	.99

Identisch $|C| > .95$, **ähnlich** $.85 \leq |C| \leq .95$, **unterschiedlich** $|C| < .85$.

Tabelle 6.31: Differenz zwischen den elektrodermalen Parameter-Faktorladungen der elektrodermalen Komponente 1 des NASA-RTLX (Na), der elektrodermalen Komponente 1 der Trefferquote (Tq), der elektrodermalen Komponente 1 der Erregung (Er) und der elektrodermalen Komponente 1 der Valenz (Va)

Parameter	Na–Tq	Na–Er	Na–Va	Tq–Er	Tq–Va	Er–Va
EDL.M	.18	.24	.25	.06	.07	.00
EDL.D	–.06	.07	.03	.12	.08	–.04
Phasic.SD	.05	–.04	–.06	–.09	–.11	–.02
EDR.N	.33	.31	.37	–.02	.04	.06
EDR.amp	–.21	–.27	–.26	–.06	–.05	.01
EDR.rise.t	–.31	–.21	–.28	.10	.02	–.07
EDR.rec.t/2	–.32	–.26	–.33	.06	–.01	–.07
EDR.rec.t/c	–.29	–.26	–.33	.04	–.04	–.07
EDR.area	–.25	–.25	–.23	.00	.02	.02
EDR.rise.v	–.13	–.20	–.18	–.07	–.05	.02
EDR.peaks	–.34	–.29	–.32	.04	.01	–.03
EDA.LF	–.28	–.28	–.31	.00	–.02	–.03
EDA.HF	–.22	–.18	–.20	.04	.02	–.02
EDA.VHF	–.16	–.08	–.07	.08	.09	.01

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, **marginal** $|a_{\Delta}| \geq .15$, **nicht bedeutsam** $|a_{\Delta}| < .15$.

elektrodermale Komponente der Valenz in ihren Parameter-Faktorladungen identisch ($C \geq .97$). Zwischen der elektrodermalen Komponente der Erregung und Valenz ist die Ähnlichkeit auch auf Ebene der Bedingungen und Personen gegeben ($|C| \geq .88$). Die elektrodermale Komponente der Trefferquote unterscheidet sich auf diesen beiden Ebenen mit $|C| \leq .72$ von den elektrodermalen Komponenten der Erregung und Valenz.

Tabelle 6.31 zeigt die Differenzen zwischen den Parameter-Faktorladungen der elektrodermalen Komponenten auf. Hier wird die zuvor berichtete Übereinstimmung zwischen den elektrodermalen Profilen der Trefferquote, der Erregung und der Valenz bestätigt. Folglich handelt es sich bei diesen drei

Tabelle 6.32: Parameter-Faktorladungen α für die elektrodermalen Profile des NASA-RTLX, der Trefferquote, der Erregung und der Valenz

Parameter	Elektrodermale Komponente			
	NASA-RTLX	Trefferquote	Erregung	Valenz
EDL.M	.36	.18	.12	.11
EDL.D	.22	.28	.15	.19
Phasic.SD	.25	.20	.29	.31
EDR.N	.43	.11	.12	.07
EDR.amp	.16	.37	.42	.42
EDR.rise.t	-.23	.08	-.02	.05
EDR.rec.t/2	-.33	-.01	-.07	.00
EDR.rec.t/c	-.33	-.03	-.07	.00
EDR.area	.06	.31	.31	.29
EDR.rise.v	.25	.37	.44	.42
EDR.peaks	-.36	-.02	-.06	-.03
EDA.LF	.12	.40	.40	.43
EDA.HF	.21	.43	.39	.41
EDA.VHF	.17	.33	.25	.25

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

elektrodermalen Profilen um dasselbe Profil. Dieses kann zu dem elektrodermalen Profil des NASA-RTLX anhand von zehn Parametern bedeutsam abgegrenzt werden (siehe Tabelle 6.31). Nur die Parameter EDL.D, Phasic.SD und EDA.VHF bieten keinen bedeutsamen Beitrag zur Differenzierung. Die große Anzahl an bedeutsamen Differenzen deutet auf eine Einfachstruktur der Faktorladungsmatrix hin. Diese Annahme wird in Tabelle 6.32 bestätigt: Hier laden die Parameter EDL.M, EDR.N, EDR.amp, EDR.rec.t/2, EDR.rec.t/c, EDR.area, EDR.peaks, EDA.LF und EDA.VHF entweder auf der Komponente des NASA-RTLX oder auf den anderen drei (identischen) Komponenten. Die Parameter EDL.D, Phasic.SD, EDR.rise.v und EDA-HF können dagegen nicht eindeutig nur einer Komponente zugeordnet werden.

Mit Hilfe der gekreuzten Regression wird die Spezifität der Profile bei der Vorhersage mentaler und emotionaler Beanspruchung geprüft. Für die gekreuzte Regression der subjektiven Valenz auf die elektrodermale Komponente des NASA-RTLX ergeben sich keine bedeutsamen Effekte, $B = 0.07$, $t(478) = 0.69$, $p = 1.0$, $R^2 = 0.77$. Demgegenüber kann die subjektiven mentalen Beanspruchung (NASA-RTLX) durch die elektrodermale Komponente der Valenz mit $B = 0.24$, $t(478) = 2.66$, $p = 0.01$, $R^2 = 7.15$ vorhergesagt werden.

6.4 Diskussion

In den folgenden Unterkapiteln werden zunächst die Ergebnisse zur Manipulationskontrolle zusammengefasst und deren Implikationen beleuchtet (Abschnitt 6.4.1). Darauf folgend werden die extrahierten Profile kardialer Aktivität (Abschnitt 6.4.2) und elektrodermalen Aktivität (Abschnitt 6.4.3) zusammenfassend dargestellt und deren Relevanz interpretiert. In Abschnitt 6.4.4 werden die Einschränkungen der Studie I dargelegt.

6.4.1 Zusammenfassung und Interpretation der Manipulationskontrolle

Zusammenfassend ist für die Fragestellung KF1 eine gelungene Induktion mentaler und eine mehrheitlich gelungene Induktion emotionaler Beanspruchung im Experiment festzustellen. Es zeigt sich, dass die mentale und emotionale Beanspruchung sowohl durch mentale als auch durch emotionale Belastungsfaktoren hervorgerufen werden konnte. Dies wird in den folgenden Abschnitten zusammenfassend erörtert.

Induktion mentaler Beanspruchung

Die subjektive und leistungsbezogene mentale Beanspruchung wurde im Experiment bedeutsam und hypothesenkonform durch die Aufgabenschwierigkeit beeinflusst (Bestätigung der H1.1 und H1.2). Darüber hinaus gibt es einen signifikanten Haupteffekt des Feedbacks auf die subjektive und leistungsbezogene mentale Beanspruchung. Die subjektive mentale Beanspruchung wird bei negativem Feedback erhöht und bei positivem Feedback verringert (Bestätigung der H1.3). Demgegenüber wird die H1.4 nur teilweise bestätigt: Die leistungsbezogene mentale Beanspruchung wird nur bei negativem Feedback (im Vergleich zu keinem Feedback) signifikant erhöht, jedoch nicht wie erwartet bei positivem Feedback (im Vergleich zu keinem Feedback) verringert. Auch wenn die Aufgabenschwierigkeit und das Aufgabenfeedback sowohl auf die subjektive mentale Beanspruchung als auch die leistungsbezogene mentale Beanspruchung bedeutsame Einflüsse zeigen, erklären diese Prädiktoren unterschiedlich viel Varianz. In der subjektiven mentalen Beanspruchung erklärt die Aufgabenschwierigkeit 14.1 % der Varianz, in der leistungsbezogenen mentalen Beanspruchung jedoch nur 5.7 %. Das Aufgabenfeedback erklärt in der subjektiven mentalen Beanspruchung 18.4 % der Varianz und in der leistungsbezogenen mentalen Beanspruchung 15.5 %. Es fällt auf, dass das Aufgabenfeedback einen wesentlich größeren Beitrag zur Varianzaufklärung in der leistungsbezogenen mentalen Beanspruchung leistet als die Aufgabenschwierigkeit. Diese Ergebnisse decken auf, dass die emotionalen (ausführungsspezifischen) Belastungen die Varianz in der leistungsbezogenen mentalen Beanspruchung stärker beeinflussen als die mentalen (informativischen) Belastungen.

Für die subjektive mentale Beanspruchung ergibt sich zudem ein ordinaler Interaktionseffekt zwischen Aufgabenschwierigkeit und Aufgabenfeedback. Die Interaktion zeigt auf, dass eine Erhöhung der Aufgabenschwierigkeit dann den größten Zuwachs in der subjektiven mentalen Beanspruchung bewirkt, wenn ein Aufgabenfeedback fehlt. Derselbe Zuwachs fällt innerhalb der Bedingungen mit negativem oder positivem Feedback geringer aus. Eine mögliche Erklärung

für diesen Interaktionseffekt ist, dass positive oder negative Emotionen die Wahrnehmung der mentalen Beanspruchung in eine Richtung verzerren. Diese emotionale Verzerrung wird bei Steigerung der mentalen Belastung in die subjektive Bewertung der mentalen Beanspruchung mit einbezogen. Gemessen am Anteil erklärter Varianz spielt der bedeutsame Interaktionseffekt mit 2.6 % eine eher untergeordnete Rolle bei der Varianzaufklärung der subjektiven mentalen Beanspruchung. Des Weiteren zeigt sich, dass es auch bei einer reinen affektiven Belastung (IAPS-Bilder) einen bedeutsamen Effekt auf die subjektive mentale Beanspruchung gibt (Bestätigung der H1.5 und H1.6). Je höher die affektive Belastung (hohe Erregung und negative Valenz), desto höher wird auch die mentale Beanspruchung wahrgenommen.

Darüber hinaus zeigten die Ergebnisse eine bedeutsam geringere subjektive mentale Beanspruchung bei rein affektiven Bedingungen (Block 3) im Vergleich zu Bedingungen mit zusätzlichen oder reinen mentalen Anforderung (Block 1 und 2). Dies ist nicht überraschend, da für die Betrachtung der Bildersets keine hohe kognitive Kapazitätsauslastung zu erwarten war. Nichtsdestoweniger ist für die weitere Interpretation der physiologischen Profile zu berücksichtigen, dass dieser Effekt größer ist als alle anderen Effekte, die für die subjektive mentale Beanspruchung gefunden wurden.

Alles in einem kann von einer gelungenen Manipulation der subjektiven und leistungsbezogenen mentalen Beanspruchung gesprochen werden. Bei der Extraktion physiologischer Profile sollte jedoch der Einfluss der emotionalen Belastungsfaktoren sowie der große Beanspruchungsunterschied des Aufgabenblocks 3 zu den Blöcken 1 und 2 mit berücksichtigt werden.

Induktion emotionaler Beanspruchung

Die Induktion der Erregung in Block 1 und 2 durch eine systematische Variation der Aufgabenschwierigkeit und des Aufgabenfeedbacks ist nur bedingt gelungen. Die in H1.7 postulierte disordinale Interaktion ist nicht signifikant

und muss zurückgewiesen werden. Stattdessen wurde eine ordinale Interaktion festgestellt: Für positives sowie für neutrales Aufgabenfeedback gibt es einen positiven Effekt der Aufgabenschwierigkeit auf die wahrgenommene Erregung. Entgegen den Annahmen in H1.7²⁹, ist für negatives Feedback kein Effekt der Aufgabenschwierigkeit auf die wahrgenommene Erregung zu finden. Hier wäre für die Bedingung S2- eine geringere Ausprägung der subjektiven Erregung zu erwarten. Eine mögliche Ursache für diesen Befund könnte die Glaubwürdigkeit des Feedbacks sein. Diese liegt im Mittel nur bei 3.09 (Wertebereich 1.0 bis 5.0). Zwar kann ein Haupteffekt der Glaubwürdigkeit auf interindividueller Ebene ausgeschlossen werden, jedoch nicht auf intraindividuelle Ebene. Hierfür müsste die Glaubwürdigkeit des Feedbacks nach jeder Aufgabe erfasst werden. Eine weitere Ursache ist in der monetären Bindung des Feedbacks zu vermuten. Der monetäre Anreiz könnte den Effekt verursacht haben, indem die negative Erfahrung des Geldverlusts zu einer höheren Erregung bei der Bedingung S2- führte. Neben dem bedeutsamen Interaktionseffekt wurden bedeutsame Haupteffekte – nämlich der der Aufgabenschwierigkeit und des Aufgabenfeedbacks – auf die subjektive Erregung in Block 1 und 2 festgestellt. Bei Zunahme der Aufgabenschwierigkeit sowie bei negativem (im Vergleich zu neutralem und positivem) Feedback steigt die subjektive Erregung. Insgesamt können die zwei Haupteffekte und der Interaktionseffekt gemeinsam nur 11.0 % der Varianz in der subjektiven Erregung der Blöcke 1 und 2 erklären.

Die Induktion der subjektiven Erregung durch die IAPS-Bildersets führte zu den erwarteten Ergebnissen. So zeigte sich eine Steigung der subjektiven Erregung für die Bildersets mit hohem Erregungsniveau (Bestätigung der H1.9). Darüber hinaus wurde auch bei Bildersets mit negativer Valenz eine Steigung der subjektiven Erregung festgestellt. Dass die Erregung und Valenz nicht unabhängig voneinander sind, zeigten bereits vorangegangene Studien (z. B. Kron, Pilkiw, Banaei, Goldstein & Anderson, 2015). Der Anteil erklärter Varianz

²⁹ Hypothese H1.7 postuliert, dass es für negatives Feedback einen negativen Effekt der Schwierigkeit auf die Erregungs-Werte im SAM gibt.

liegt für den Block 3 bei 23.3 % wesentlich höher als für Block 1 und 2 (11.0 %). Dies ist nicht überraschend, da in Block 3 eine rein affektive Belastungsinduktion durchgeführt wurde. Im Gegensatz zur subjektiven mentalen Beanspruchung fand sich kein Unterschied der mittleren subjektiven Erregung zwischen Block 3 und Block 1 bis 2.

In Block 1 und 2 (Balltrackingaufgabe) wurde die subjektive Valenz nur durch das Aufgabenfeedback erwartungskonform beeinflusst (Bestätigung der H1.8). Hier führte positives Feedback zu einer positiven subjektiven Valenz und negatives Feedback zu einer negativen subjektiven Valenz. Das Aufgabenfeedback erklärte 16.8 % der Varianz. Auch in Block 3 (IAPS-Bildersets) konnte der erwartete positive Haupteffekt der Bilderset-Valenz auf die subjektive Valenz gefunden werden (Bestätigung der H1.10). Dieser Effekt erklärt 30.2 % der Varianz. Ein Effekt des Bilderset-Erregung auf die subjektive Valenz lag nicht vor. Wie bereits die subjektive Erregung zeigt auch die subjektive Valenz keinen Unterschied zwischen den Bedingungen mit reiner affektiven Belastung (Block 3) und den Bedingungen mit zusätzlicher oder reiner mentaler Belastung (Block 1 und 2).

Zusammenfassend konnten drei von vier Hypothesen zur Prüfung der emotionalen Beanspruchungsinduktion bestätigt werden. Lediglich die H1.7 musste teilweise verworfen werden. Unter Berücksichtigung der ordinalen Interaktion für die H1.7 kann von einer gelungenen Manipulation der subjektiven Erregung und der subjektiven Valenz im Experiment gesprochen werden.

6.4.2 Zusammenfassung und Interpretation der kardialen Profile

Bezug nehmend auf Forschungsfrage F1 konnte in Studie I ein bedeutsames kardiales Profil sowohl für das NASA-RTLX als auch für die Trefferquote gefunden werden. Ebenfalls finden sich kardiale Profile für die Emotionsdimensionen Erregung und Valenz (Forschungsfrage F3). Die Ergebnissen zur

Tabelle 6.33: Übersicht der kardialen Profile und deren Vorhersagegüte aus Studie I

Kriterium	B	R^2_{PLS}	$\text{val}R^2_{\text{PLS}}$	Interpretation
NASA-RTLX	0.37	73.17	56.78	Mentale Beanspruchung
Trefferquote	-0.42	40.58	19.51	} Mentale und emotionale Beanspruchung
Erregung	0.49	31.00	15.40	
Valenz	-0.58	33.20	25.31	Emotionale Beanspruchung

Anmerkungen. Das R^2_{PLS} und $\text{val}R^2_{\text{PLS}}$ beziehen sich auf den Anteil erklärter Varianz im Kriterium.

Forschungsfrage F5.1 zeigen auf, dass eine Abgrenzung der kardialen Profile untereinander nur eingeschränkt möglich ist. Es findet sich keine eindeutige Einfachstruktur der Profile. So unterscheiden sich die Profile in nur ein bis maximal zwei Parametern bedeutsam voneinander und sind als ähnlich zu interpretieren. Die Trennung der kardialen Wirkungsebene in emotionale und mentale Beanspruchung ist nur in eine Richtung möglich: Das kardiale Profil des NASA-RTLX sagt spezifisch die subjektive mentale Beanspruchung, jedoch nicht die subjektive Valenz, vorher. Demgegenüber kann das kardiale Profil der Valenz sowohl die subjektive Valenz als auch die subjektive mentale Beanspruchung vorhersagen. Dieser Befund deckt sich mit den Ergebnissen aus der Manipulationskontrolle: Eine negative Valenz führt zu einem bedeutsamen Anstieg der subjektiven mentalen Beanspruchung, aber eine hohe mentale Beanspruchung führt nicht zwangsläufig zu einem negativen Valenzempfinden.

Tabelle 6.33 fasst die statistischen Maße sowie die Interpretation der vier kardialen Profile zusammen. Hier ist am kreuzvalidierten Bestimmtheitsmaß ($\text{val}R^2_{\text{PLS}}$) erkennbar, dass die Vorhersagegüte für Personen, die dem Modell unbekannt sind, zum Teil stark abnimmt. Auch wenn der verbleibende Anteil erklärter Varianz immer noch bedeutsam und beachtlich ist, gehen 7.7 % bis 21.1 % erklärter Varianz verloren. Dies verdeutlicht die Relevanz der personenspezifischen Unterschiede für die Vorhersagegüte der mentalen und

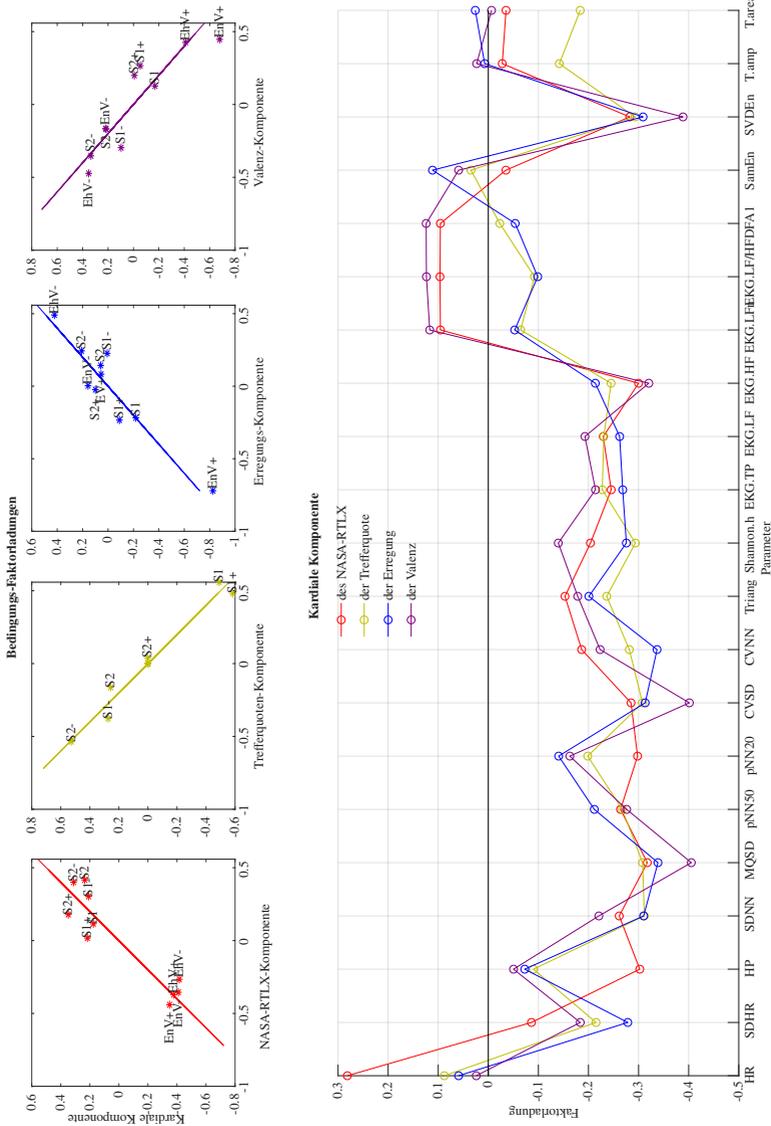


Abbildung 6.18: Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der kardialen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der kardialen Komponenten aus Studie I.

emotionalen Beanspruchung durch die kardialen Profile. Auf die einzelnen kardialen Profile wird im folgenden näher eingegangen.

Von den vier kardialen Profilen klärt das kardiale Profil des NASA-RTLX mit 73.2 % (Originalmodell) den größten Anteil der Varianz im jeweiligen Kriterium auf. Das im Vergleich dazu eher geringe Regressionsgewicht von $B = 0.37$ lässt vermuten, dass der hohe Varianzanteil durch den großen Beanspruchungsunterschied der beiden Bedingungs-Cluster (Balltracking- und IAPS-Bedingungen) zustande kommt. Diese Cluster sind in Abbildung 6.18 (oben, rote Linie) deutlich erkennbar. Somit wird der Zusammenhang zwischen dem kardialen Profil und der subjektiven mentalen Beanspruchung im NASA-RTLX wesentlich durch den großen Beanspruchungsunterschied zwischen den Balltracking- und den IAPS-Bedingungen moduliert. Des Weiteren ist zu beachten, dass das kardiale Profil des NASA-RTLX – im Gegensatz zur gemessenen subjektiven mentalen Beanspruchung – die emotionalen Belastungsfaktoren nicht abbildet. Ob diese Bereinigung emotionaler Einflüsse auf ein studienspezifisches Phänomen zurückzuführen ist, bleibt zu klären. Ungeklärt bleibt ebenfalls der fehlende Anstieg in den kardialen Faktorladungen bei steigender Aufgabenschwierigkeit in den Bedingungen ohne Aufgabenfeedback. Für die restlichen Bedingungen ist eine erwartungskonforme Steigung kardialer Faktorladungen bei steigender Aufgabenschwierigkeit gegeben. Alles in einem ist das kardiale Profil des NASA-RTLX als ein Repräsentant der subjektiven mentalen Beanspruchung der Studie I zu interpretieren. Von allen kardialen Profilen ist das kardiale Profil des NASA-RTLX das Einzige, für das eine Einfachstruktur der Ladungen – für die Parameter HR und HP – vorliegt. Die restlichen Parameter laden auf mehreren oder sogar allen kardialen Profilen gleichzeitig.

Das kardiale Profil der Trefferquote klärt mit $B = -0.42$ bis zu 40.6 % (Originalmodell) der Varianz in der Trefferquote auf. In Abbildung 6.18 (oben, olivfarbene Linie) ist zu erkennen, dass das kardiale Profil der Trefferquote auch die emotionalen Belastungsfaktoren innerhalb der Balltracking-Aufgabe abbildet. Dies ist wenig überraschend, wenn man bedenkt, dass der Effekt der emotionalen Belastungsfaktoren (15.5 % Varianzaufklärung) auf die gemessene

Trefferquote weit aus höher ist als der Effekt der mentalen Belastungsfaktoren (5.7 % Varianzaufklärung). Die fehlende Leistungsmessung in den Bedingungen des IAPS-Bildersets bietet eine weitere Erklärung dafür, dass das Profil des NASA-RTLX und das Profil der Trefferquote die mentale Beanspruchung unterschiedlich repräsentieren. Dadurch konnte bei der Extraktion des kardialen Profils der Trefferquote der große Beanspruchungsunterschied zwischen den Balltracking- und den IAPS-Bedingungen nicht berücksichtigt werden. Innerhalb der Balltracking-Bedingungen beeinflussen sich die mentale und die emotionale Beanspruchung jedoch gegenseitig. Diese fehlende Trennung mentaler und emotionaler Beanspruchung auf der psychologischen Wirkungsebene wirkt sich auf die Extraktion des kardialen Profils so aus, dass eine Mischform mentaler und emotionaler Beanspruchung resultiert. Auch das Parameter-Ladungsmuster des kardialen Profils der Trefferquote hat eine höhere Kongruenz mit dem kardialen Profil der Erregung als mit dem kardialen Profil des NASA-RTLX. Folglich wird das Profil der Trefferquote als ein Mischprofil aus mentaler und emotionaler Beanspruchung gewertet.

Das kardiale Profil der Erregung erklärt mit $B = 0.49$ bis zu 31.0 % (Originalmodell) der Varianz in der subjektiven Erregung. Dieses Profil liegt an der Identitätsschwelle zum kardialen Profil der Trefferquote. Beide Profile können nur durch einen einzigen Parameter, T.area, bedeutsam abgegrenzt werden. Allerdings besitzt T.area für keinen der beiden Profile eine bedeutsame Parameter-Faktorladung. Ein einziger und zudem nicht bedeutsamer Parameter genügt nicht zur Differenzierung beider Profile. Zudem wurde festgestellt, dass – im Vergleich zur Erregungs-Komponente – die kardiale Komponente der Erregung die mentalen Belastungsfaktoren stärker und die emotionalen Belastungsfaktoren schwächer abbildet (z. B. Anstieg zwischen den Schwierigkeitsstufen 1 und 2 bei negativem Leistungsfeedback; siehe Abbildung 6.18, oben, blaue Linie). Auf Grundlage dieser Ergebnisse ist das kardiale Profil der Erregung – wie bereits das kardiale Profil der Trefferquote – als Mischform mentaler und emotionaler Beanspruchung zu interpretieren.

Die kardialen Profile der Trefferquote und der Erregung unterscheiden sich zum kardialen Profil der Valenz in nur einem Parameter (EKG.LF/HF) bedeutsam. Auch hier besitzt der Parameter keine bedeutsame Parameter-Faktorladung auf den Profilen. Dass die kardialen Profile der Valenz und der Erregung eine gewisse Ähnlichkeit haben, ist in Anbetracht der Abhängigkeit der Emotionsdimensionen Erregung und Valenz (siehe Abschnitt 6.3.1) nicht überraschend. Nichtsdestoweniger bildet das kardiale Profil der Valenz (im Gegensatz zum kardialen Profil der Erregung) weniger die mentalen und mehr die emotionalen Belastungsfaktoren ab. Dies ist in Abbildung 6.18 (oben, violette Linie) erkennbar: Die Bedingungs-Faktorladungen bilden eine gleichmäßig verteilte Punktwolke entlang der Valenzausprägung der Bildersets sowie des Aufgabenfeedbacks. Folglich wird dieses kardiale Profil als Repräsentant der emotionalen Beanspruchung interpretiert. Auch wenn das kardiale Profil der Valenz nur 33.2 % der Varianz (im Originalmodell) aufklärt, so zeigt es über alle kardialen Profile hinweg ein hohes Regressionsgewicht von -0.58 .

In Abbildung 6.18 (unten) sind die Parameter-Faktorladungen der vier kardialen Profile abgebildet. Die Parameterwerte der HRV in den kardialen Profilen sinken mit Zunahme der mentalen oder emotionalen Beanspruchung. Im Gegensatz dazu nimmt der Parameter HP in dem kardialen Profil des NASA-RTLX mit Zunahme der mentalen Beanspruchung zu. Diese Zusammenhangsmuster sind konsistent mit den bisherigen Befunden (vgl. Abschnitt 3.2.4). Eine endgültige Bewertung der kardialen Einzelparameter sowie deren Zuordnung zu einem psychologischen Konstrukt erfolgt am Ende der Arbeit in Unterkapitel 9.1.2.

6.4.3 Zusammenfassung und Interpretation der elektrodermalen Profile

Insgesamt konnten in der Studie I vier elektrodermale Profile identifiziert werden. Diese vier Profile können auf zwei Profile reduziert werden (siehe Tabelle 6.34). Alle extrahierten elektrodermalen Profile sind in Abbildung 6.19

Tabelle 6.34: Übersicht der elektrodermalen Profile und deren Vorhersagegüte aus Studie I

Kriterium	B	R^2_{PLS}	$\text{val}R^2_{\text{PLS}}$	Interpretation
NASA-RTLX	0.67	80.97	52.50	Mentale Beanspruchung
Trefferquote	-0.54	41.91	10.32	
Erregung	0.45	27.98	8.16	Emotionale Beanspruchung
Valenz	-0.53	33.51	4.18	

Anmerkungen. Die Bestimmtheitsmaße R^2_{PLS} und $\text{val}R^2_{\text{PLS}}$ beziehen sich auf den Anteil erklärter Varianz im Kriterium.

(unten) zu sehen. Bezug nehmend auf Forschungsfrage F2 wurde ein bedeutsames elektrodermales Profil für die subjektive mentale Beanspruchung identifiziert. Ein elektrodermales Profil, welches zuverlässig die leistungsbezogene mentale Beanspruchung repräsentiert, konnte nicht gefunden werden, da dieses Profil verstärkt die emotionalen Belastungsfaktoren abbildet. Dies ist in Abbildung 6.19 (oben, olivfarbene Linie) erkennbar: Die Balltracking-Bedingungen verlaufen innerhalb der Punktwolke entlang des positiven und negativen Aufgabenfeedbacks (und kaum entlang der Aufgabenschwierigkeit). Vor allem innerhalb der Aufgaben mit positivem oder negativem Feedback bildet das elektrodermale Profil der Trefferquote den Unterschied zwischen negativem und positivem Aufgabenfeedback besser ab als die Unterschiede zwischen der leichten und der schweren Aufgabe. Die Modulation des Profils der Trefferquote durch emotionale Belastungsfaktoren wurde bereits beim kardialen Profil der Trefferquote festgestellt (siehe Abschnitt 6.4.2). Dies ist zum einen auf den höheren Einfluss der emotionalen – im Vergleich zu den mentalen – Belastungsfaktoren auf die gemessene Trefferquote zurückzuführen (siehe Abschnitt 6.3.1). Zum anderen liegt keine Unabhängigkeit mentaler und emotionaler Beanspruchung innerhalb der Balltracking-Bedingungen vor.

In Bezug auf die Forschungsfrage F4 konnte jeweils ein elektrodermales Profil der Erregung und ein Profil der Valenz extrahiert werden. Die Ergebnisse zur Forschungsfrage F5.2 zeigen auf, dass diese beiden Profile sowohl unterein-

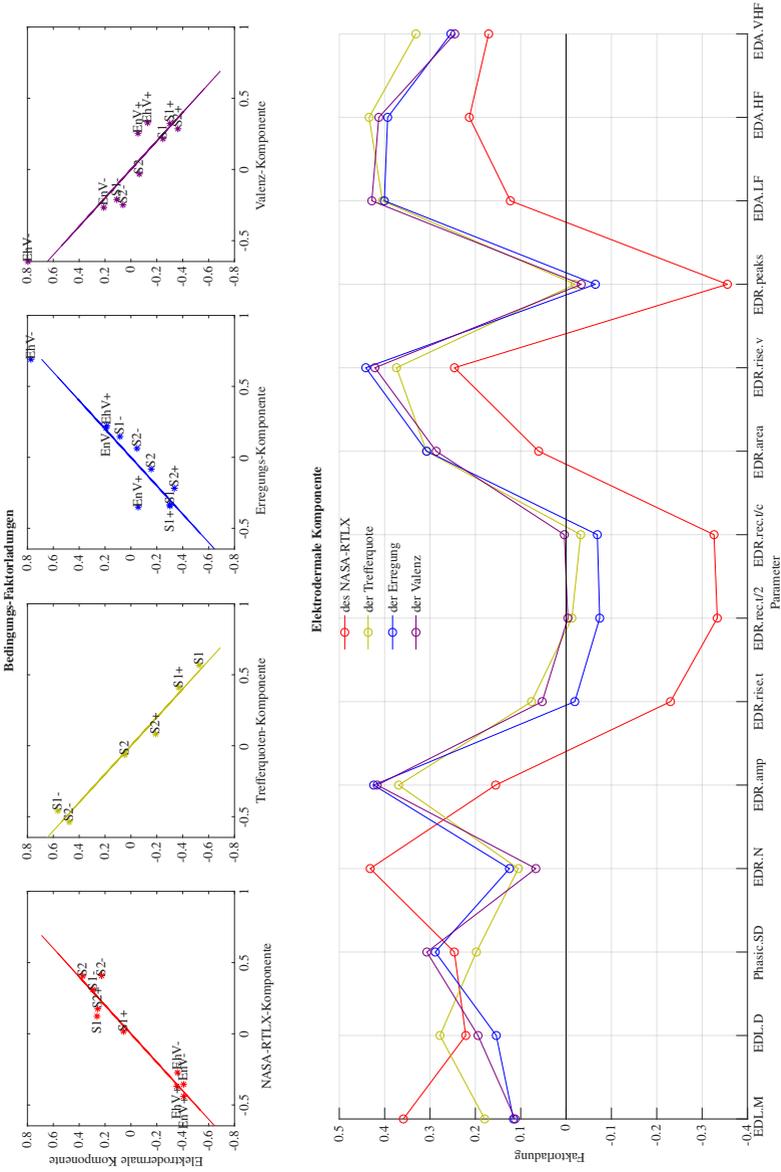


Abbildung 6.19: Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der elektrodermalen Komponenten aus Studie I.

ander als auch zum elektrodermalen Profil der Trefferquote identisch sind. Die identischen Profile werden als ein elektrodermales Profil der emotionalen Beanspruchung interpretiert. Das elektrodermale Profil der emotionalen Beanspruchung erklärt (im Originalmodell) 28.0 % der Varianz in der subjektiven Erregung und 33.5 % der Varianz in der subjektiven Valenz. Allerdings nimmt die Vorhersagegüte für Personen, die dem Modell nicht bekannt sind, sehr stark ab (siehe $\text{val}R_{\text{PLS}}^2$ in Tabelle 6.34). Dieser hohe Varianzverlust von 19.8 % bis 31.6 % ist unter anderem auf die hohen interindividuellen Unterschiede sowohl in der elektrodermalen Aktivität als auch in der subjektiven Erregung und in der subjektiven Valenz zurückzuführen. Die in den kreuzvalidierten Modellen verbleibenden, geringen Varianzanteile deuten darauf hin, dass die interindividuellen Gewichtungsfaktoren bei der Vorhersage der emotionalen Beanspruchung essentiell sind.

Gemäß den Ergebnissen aus Forschungsfrage F5.2 können die zur emotionalen Beanspruchung zugeordneten elektrodermalen Profile zum elektrodermalen Profil des NASA-RTLX abgegrenzt werden. Das elektrodermale Profil des NASA-RTLX weist mit $R_{\text{PLS}}^2 = 81.0$ (Originalmodell) und $B = 0.67$ die besten Vorhersagewerte auf und ist als Repräsentant der subjektiven mentalen Beanspruchung zu interpretieren. Es ist zu beachten, dass das elektrodermale Profil des NASA-RTLX wesentlich durch den Unterschied zwischen den Balltracking- und den IAPS-Bedingungen moduliert wird (siehe Bedingungs-Cluster in Abbildung 6.19, oben, rote Linie). Der große Beanspruchungsunterschied zwischen den beiden Bedingungs-Clustern (Balltracking- und IAPS-Bedingungen) bietet zudem eine Erklärung für die vergleichsweise hohen Vorhersagewerte dieses Profils. Der Anteil erklärter Varianz in der subjektiven mentalen Beanspruchung wird bei der Vorhersage unbekannter Personen um 28.8 % reduziert (siehe $\text{val}R_{\text{PLS}}^2$ in Tabelle 6.34). Auch wenn der verbleibende Anteil erklärter Varianz von 52.5 % beachtlich ist, können die interindividuellen Gewichtungsfaktoren einen wesentlichen Beitrag zu Vorhersagegüte leisten.

Im Gegensatz zu den kardialen Profilen liegt für die elektrodermalen Profile eine Einfachstruktur der Parameter-Ladungsmatrix vor. Diese erlaubt eine

Trennung zwischen dem Profil der mentalen Beanspruchung und dem Profil der emotionalen Beanspruchung. Allerdings ist es auch für die elektrodermalen Profile möglich, die subjektive mentale Beanspruchung durch das elektrodermale Profil der Valenz – jedoch nicht die subjektive Valenz durch das Profil der mentalen Beanspruchung – vorherzusagen. Im Vergleich zu den kardialen Profilen ist dieser Effekt für die elektrodermalen Profile geringer ausgeprägt: 22.4 % Varianzaufklärung durch das kardiale Profil der Valenz gegenüber 7.2 % Varianzaufklärung durch das elektrodermale Profil der Valenz.

Das elektrodermale Profil der mentalen Beanspruchung wird durch Veränderungen des mittleren tonischen Hautleitwerts (EDL.M), durch die die NS-EDR-Anzahl (EDR.N) und durch zeitabhängige Parameter der NS-EDRs (EDR.rec.t/2, EDR.rec.t/c und EDR.peaks) vertreten. Konform zu den bisherigen Forschungsbefunden nehmen die Parameter EDR.N und EDL.M mit steigender mentaler Beanspruchung zu, während EDR.rect/2 oder EDR.rect/c abnehmen (vgl. Abschnitt 3.3.4). Demgegenüber wird das elektrodermale Profil der emotionalen Beanspruchung durch Parameter aus dem Frequenzbereich (EDA.LF, EDA.HF und EDA.VHF) und durch amplitudenabhängigen Maße der NS-EDRs (EDR.amp, EDR.area und EDR.rise.v) repräsentiert. Diese Parameter nehmen mit steigender emotionaler Beanspruchung zu. Eine endgültige Bewertung der elektrodermalen Einzelparameter sowie deren Zuordnung zu einem psychologischen Kriterium erfolgt am Ende der Arbeit in Unterkapitel 9.1.4.

6.4.4 Limitationen der Studie

In der Studie I konnten mehrere physiologische Profile identifiziert und auf die mentale sowie emotionale Beanspruchung zurückgeführt werden. Allerdings war eine eindeutige Zuordnung nicht immer möglich. Einschränkungen in der Ergebnisinterpretation und unerwartete Ergebnisse sind in der Experimentalforschung gängig. Nichtsdestoweniger ist es wichtig diese ausführlich zu beleuchten und somit den Weg für weitere Experimente zu ebnet. In den

folgenden Abschnitten werden die Einschränkungen der vorliegenden Studie aufgeführt und diskutiert.

Einschränkungen der Messdaten

Nicht gänzlich geklärt ist die Reliabilität kardialer Kurzzeitmaße (siehe Abschnitt 3.2.3). Es ist zu beachten, dass für Metriken der Ultra-Kurzzeitmessung bisher nur wenige fundierte Validierungsstudien vorliegen, die kardiale Ultra-Kurzzeitmaße mit Parametern aus längeren Messintervallen vergleichen. Pecchia, Castaldo, Montesinos und Melillo (2018) verweisen in ihrem Review zur Validität kardialer Ultra-Kurzzeitmaße (aus 29 Studien) auf den Mangel an zuverlässigen und stringenten Validierungsmethoden. Sowohl Pecchia et al. (2018) als auch Shaffer und Ginsberg (2017) kommen zu dem Schluss, dass in Bezug auf die Ultra-Kurzzeitmessung eine endgültige Bewertung der Validität sowie eine Methodenstandardisierung noch aussteht.

Darüber hinaus wurde in der Studie I keine Kontrolle der Atemaktivität vorgenommen. Diese ist sowohl für die kardialen als auch für die elektrodermalen Parameter von Bedeutung. Bspw. können NS-EDRs auch durch tiefes Einatmen ausgelöst werden. Es ist anzumerken, dass die Atemaktivität – wie auch die elektrodermale und kardiale Aktivität – ebenfalls durch die mentale Beanspruchung beeinflusst wird (vgl. Grassmann et al., 2016). Versucht man nun die kardialen und elektrodermalen Signale um sogenannte Atemartefakte im Signal zu bereinigen, ergibt sich das Problem einer validen Differenzierung der Atemartefakte von wahren Effekten der mentalen Beanspruchung. Nichtsdestoweniger könnten zukünftige Arbeiten untersuchen, ob die Sensitivität und Diagnostizität physiologischer Profile durch die Bereinigung potentieller Atemartefakte verbessert werden kann.

Des Weiteren wurde ein signifikanter Anstieg in der Raumtemperatur um durchschnittlich 0.72 °C gefunden. Dieser Anstieg ist zwar gering, jedoch kann ein Einfluss der Raumtemperatur auf die Ergebnisse der elektrodermalen Profile

für die vorliegende Studie I nicht ausgeschlossen werden. Der quantitative und qualitative Einfluss der Raumtemperatur auf elektrodermale Profile ist eine eigenständige Forschungsfrage, welche im Rahmen der vorliegenden Arbeit nicht adressiert wird.

Einschränkungen in der Konstruktrennung

Eine Einschränkung der Studie I ist die Abhängigkeit der subjektiven mentalen und emotionalen Beanspruchung (Erregung und Valenz). Die Manipulationskontrolle in Abschnitt 6.3.1 zeigt, dass die mentale und emotionale Beanspruchung von den Belastungsfaktoren des jeweils anderen Konstrukts beeinflusst werden. Diese mangelnde Diagnostizität der psychologischen Konstrukte wirkt sich auch auf die Diagnostizität der extrahierten physiologischen Profile aus. So ist es nicht überraschend, dass die physiologischen Profile Einflüsse mentaler und emotionaler Beanspruchung oder sogar eine Mischung aus beiden abbilden. Optimal wäre es, physiologische Profile zu extrahieren, die auf eine reine – also von der emotionalen Beanspruchung unabhängige – mentale Beanspruchung zurückzuführen sind. Eine Möglichkeit wäre es, das konkurrierende Konstrukt (hier: emotionale Beanspruchung) aus der Kriteriumsvariable (hier: mentale Beanspruchung) auszupartialisieren. Hier ergäbe sich allerdings das Problem, wie ein auf dieser Basis extrahiertes Profil inhaltlich zu interpretieren ist. Dieser Lösungsansatz ist kein Fokus der vorliegenden Arbeit und wäre in zukünftigen Forschungsarbeiten zu klären.

Eine mangelnde Trennung mentaler und emotionaler Beanspruchung wurde in Studie I vor allem für das Leistungsmaß festgestellt. Sowohl das kardiale als auch das elektrodermale Profil der Trefferquote repräsentieren verstärkt die emotionale und weniger die mentale Beanspruchung. Im Gegensatz zu den physiologischen Profilen des NASA-RTLX wurden die Profile der Trefferquote nur aus den Bedingungen der Balltrackingaufgabe extrahiert. Innerhalb der Balltrackingaufgabe ist der Effekt der emotionalen Belastungsfaktoren auf die Trefferquote sogar stärker als der Effekt der mentalen Belastungsfaktoren (siehe

Abschnitt 6.3.1). Demnach ist es wenig überraschend, dass die extrahierten Profile diese Varianzquellen entsprechend nachbilden. Unklar bleibt, ob dieser Befund spezifisch für das vorliegende Experiment der Studie I ist oder ob es sich hierbei um eine generelle Eigenschaft des Leistungsmaßes handelt. So könnte die Ursache in der Verwendung des Sekundäraufgabenmaßes liegen.

Einschränkungen im Versuchsdesign

Zunächst ist der quantitative und qualitative Belastungsunterschied zwischen den Experimentalbedingungen der Balltrackingaufgabe und den IAPS-Bildersets zu erwähnen. Dieser Unterschied birgt potentielle Störfaktoren für die kardialen und elektrodermalen Profile. So kann nicht ausgeschlossen werden, dass der aufgabenspezifische Grad an Bewegung sich in den extrahierten physiologischen Profilen niederschlägt. Zwar wurden alle Versuchspersonen dazu aufgefordert, sich so wenig wie möglich im Versuchsablauf zu bewegen, jedoch war für die Balltrackingaufgabe im Vergleich zu den IAPS-Bildersets mindestens eine kontinuierliche Bewegung der Hand notwendig. Um diesen potentiellen Störfaktor auszuschließen, ist es wichtig, kardiale und elektrodermale Profile über verschiedene Aufgabentypen hinweg mit unterschiedlichen motorischen Anforderungen zu vergleichen. Dieser Aspekt wird im Rahmen der Studie II und Studie III beleuchtet.

Nicht zuletzt ist die fehlende Randomisierung zwischen den Versuchsblöcken 1 bis 3 zu erwähnen. Diese wurde vermieden, um emotionale Einflüsse durch die IAPS-Bildersets auf die Balltrackingbedingungen zu verhindern. Mit der intra- und interindividuell konstanten Abfolge der Blöcke können allerdings Reihenfolgeeffekte, wie Ermüdung und Motivation, auf die physiologischen Profile nicht ausgeschlossen werden.

Einschränkungen in der externen Validität

Bedingt durch den Laborkontext der Studie I ist eine Übertragbarkeit der Ergebnisse auf eine Aufgabe außerhalb der Laborumgebung oder sogar auf eine reale Situation nicht geklärt. Auch eine Übertragung auf andere Aufgabentypen ist mit dem aktuellen Forschungsstand nicht gesichert. Letzteres wird mit Hilfe der Studie III beleuchtet. Die externe Validität der vorliegenden Studie wird zusätzlich dadurch beschränkt, dass es sich nicht um eine probabilistische, sondern um eine anfallende Stichprobe mit einer selektiven Personengruppe handelt. Der hohe studentische Anteil stellt sowohl bezogen auf das Alter, den Bildungsgrad und viele weitere Merkmalen keine verteilungskonforme Abbildung der deutschen oder sogar der Weltbevölkerung dar. Es ist nicht auszuschließen, dass die in dieser Studie extrahierten physiologischen Profile für andere Bevölkerungsgruppen divergieren.

7 Studie II

In diesem Kapitel werden die Hintergründe der Studie II in Abschnitt 7.1 und die studienspezifischen Methoden in Abschnitt 7.2 dargelegt. Anschließend werden in Abschnitt 7.3 die Ergebnisse zur Kontrollfragestellung KF2 und zu den Forschungsfragen F1, F2 und F6 beschrieben. Zum Schluss werden in Abschnitt 7.4 die Ergebnisse interpretiert und deren Limitationen beleuchtet.

7.1 Forschungsziel der Studie II

Wie bereits in Studie I gezeigt, besteht auch in Studie II das wesentliche Ziel in der Ableitung kardialer und elektrodermalen Profile mentaler Beanspruchung (siehe Forschungsfrage F1 und F2 in Kapitel 4.3). Das spezifische Ziel der Studie II ist die Prüfung der Strukturvalidität der bereits in Studie I extrahierten kardialen und elektrodermalen Profile mentaler Beanspruchung (siehe Forschungsfrage F6 in Kapitel 4.3). Im Gegensatz zu Studie I stehen ausschließlich die subjektive und leistungsbezogene mentale Beanspruchung im Fokus der Studie II. Darüber hinaus werden zwei wesentliche Aspekte aus Studie I aufgegriffen. Erstens wird in Studie II die Leistung als Primäraufgabenmaß erfasst. Demgegenüber wurde in Studie I die Leistung in der Sekundäraufgabe gemessen. Zweitens wird in Studie II eine Aufgabe konzipiert, welche die Körperbewegung der Versuchspersonen minimiert. Eine Einschränkung der Studie I ergab sich nämlich aus den unterschiedlichen Bewegungsanforderungen zwischen den Balltracking- und den IAPS-Bedingungen. Denn während die IAPS-Bildersets nur betrachtet werden mussten, war für die Balltrackingaufgabe

eine kontinuierliche Bewegung der Hand notwendig. Dadurch konnte nicht ausgeschlossen werden, dass sich der aufgabenspezifische Grad an Bewegung in den extrahierten physiologischen Profilen der Studie I niedergeschlagen hat. Zusammenfassend dient die Studie II dazu, die physiologischen Profile mentaler Beanspruchung aus Studie I zu replizieren und gleichzeitig die oben beschriebenen Einschränkungen aus Studie I zu adressieren.

7.2 Methode

Dieser Abschnitt beschreibt die studienspezifischen Methoden. Dies umfasst die Operationalisierung der mentalen Beanspruchung (Abschnitt 7.2.1), das Versuchsdesign (Abschnitt 7.2.2), die Versuchsdurchführung (Abschnitt 7.2.3), die Kontrollvariablen (Abschnitt 7.2.4), die Manipulationskontrolle (Abschnitt 7.2.5) und die Beschreibung der Stichprobe (Abschnitt 7.2.6).

7.2.1 Operationalisierung der mentalen Beanspruchung

In diesem Abschnitt werden die Methoden erläutert, mit der die mentale Beanspruchung im Experiment ausgelöst und gemessen wurde.

Experimentelle Induktion mentaler Beanspruchung

Die Induktion der mentalen Beanspruchung erfolgte über informatorische Belastungsfaktoren in Form einer Multiplikationsaufgabe am Computer. Diese wurde als Primäraufgabenmaß umgesetzt. Insgesamt wurde die Aufgabe so konzipiert, dass sich die Versuchspersonen im Experiment möglichst wenig bewegen. Die Multiplikationsaufgaben wurde den Versuchspersonen akustisch mitsamt der Lösung (z. B. $8 \cdot 26 = 208$) über einen Lautsprecher dargeboten. Jeder Versuchsdurchlauf dauerte 60 s. Innerhalb dieser Zeitspanne mussten acht

Multiplikationsrechnungen auf ihre Richtigkeit hin überprüft werden. Pro Versuchsbedingung wurde den Versuchspersonen drei Multiplikationsrechnungen mit einem korrekten Ergebnis (z. B. $6 \cdot 2 = 12$) und fünf Multiplikationsrechnungen mit einem falschen Ergebnis (z. B. $6 \cdot 2 = 11$) vorgegeben. Jede Multiplikationsrechnung wurde für 7.5 s akustisch präsentiert. Die Reihenfolge korrekter und falscher Multiplikationsrechnungen wurde innerhalb einer Versuchsbedingung randomisiert. Die Versuchspersonen wurden dazu angeleitet, bei einem korrekten Multiplikationsergebnis die Eingabetaste auf der Computertastatur zu betätigen. Dafür hatten sie 4 s Zeit, bevor die nächste Multiplikationsrechnung vorgegeben wurde. Bei einem falschen Multiplikationsergebnis sollte keine Antwortreaktion (kein Tastendruck) erfolgen. Die Versuchsperson erhielt keine Rückmeldung darüber, ob ihre Antwortreaktion richtig oder falsch war.

Die mentale Beanspruchung wurde mit Hilfe von zwei Schwierigkeitsstufen manipuliert. Bei der leichten Schwierigkeitsstufe S1 enthielten die Multiplikationsrechnungen die Zahlen 0 bis 6 (z. B. $1 \cdot 4 = 4$). Für die schwere Schwierigkeitsstufe S2 wurden Rechnungen mit ein- und zweistelligen Zahlen verwendet (z. B. $9 \cdot 19 = 171$). Für die falschen Multiplikationsrechnungen wurden unterschiedliche Verfahren angewandt, um Lösungsstrategien der Versuchspersonen zu verhindern. Zum einen wurden Ergebnisse um die Differenz von 1 (z. B. $8 \cdot 17 = 137$) oder 10 (z. B. $8 \cdot 26 = 218$) verfälscht. Zum anderen wurde eine scheinbar logische Lösung, wie $0 \cdot 2 = 2$, präsentiert.

Alle Versuchspersonen wurden dazu instruiert, sich im Experiment auf die Multiplikationsaufgabe zu fokussieren. Parallel zur Multiplikationsaufgabe wurden in jeder Versuchsbedingung nacheinander 10 Bilder (jedes für 6 s) auf dem Computerbildschirm visuell präsentiert. Die Bilder wurden über die Versuchsbedingungen hinweg systematisch variiert. Mit diesem Vorgehen sollte nicht die Aufgabenschwierigkeit, sondern die visuelle Anforderung pro Versuchsbedingung manipuliert werden. Letzteres diente im Rahmen des Experiments dazu, den Einfluss der visuellen Komplexität auf okulomotorische Parameter zu untersuchen. Dieser Teilaspekt ist kein Bestandteil der vorliegenden Arbeit

und wird deshalb nicht weiter vertieft. Nichtsdestoweniger ist es wichtig, einen potentiellen Einfluss der unterschiedlichen Bilder auf die subjektive und leistungsbezogene mentale Beanspruchung zu prüfen bzw. auszuschließen. Die präsentierten Bilder wurden hinsichtlich des Bildmaterials und der visuellen Komplexität variiert. Das Bildmaterial bildete drei Faktorstufen ab: schwarz-weiße Zeichnung (z), farbige Zeichnung (b) und farbiges Foto (p) desselben Motivs. Hierfür wurden schwarz-weiße sowie bunte Zeichnungen aus dem Bilderset von Snodgrass und Vanderwart (1980) sowie Rossion und Pourtois (2004) verwendet. Die Fotografien wurden basierend auf einer Internetrecherche so ausgewählt, dass sie das Motiv der schwarz-weiß und bunten Zeichnungen möglichst ähnlich abbilden. Die visuelle Komplexität der Bilder unterteilte sich in eine geringe (-) und eine hohe (+) Komplexitätsstufe.¹

Messung der mentalen Beanspruchung

In Studie II wurde die subjektive mentale Beanspruchung mit dem NASA-RTLX (ohne die Skala körperliche Anforderung) erfasst (siehe Abschnitt 5.4.1). Zudem wurde die leistungsbezogene mentale Beanspruchung als Effektivitätsmaß der Multiplikationsaufgabe gemessen. Hierfür wurde die Anzahl korrekter Antwortreaktionen auf die präsentierten Multiplikationsrechnungen pro Experimentalbedingung in Prozent (Trefferquote) berechnet. Eine Reaktion galt als korrekt, wenn die Eingabetaste nach der Präsentation einer korrekten Multiplikationsrechnung gedrückt wurde oder keine Reaktion nach der Präsentation einer falschen Multiplikationsrechnung erfolgte.

¹ Die Unterteilung erfolgte anhand der berichteten, subjektiven visuellen Komplexität aus Snodgrass und Vanderwart (1980). Diese wurde auf einer 5-stufigen Likertskala von gering bis hoch bewertet. Für die verwendeten Bilder der Studie II lag der Mittelwert der geringen Komplexitätsstufe bei 1.5 und der hohen Komplexitätsstufe bei 4.0.

7.2.2 Versuchsdesign

Das Experiment der Studie II umfasst 12 Experimentalbedingungen: 2 Schwierigkeitsstufen \times 2 Komplexitätsstufen \times 3 Bildertypen. Bei dem messwiederholten Design durchlief jede Versuchsperson alle Experimentalbedingungen. Zur Vermeidung von Sequenzeffekten wurde die Reihenfolge der Experimentalbedingungen innerhalb der Personen randomisiert.

Insgesamt ergeben sich für die Studie II zwei Kriteriumsvariablen: die NASA-RTLX-Werte und die Trefferquote. Jede der Kriteriumsvariablen bildet eine Kriteriumsmatrix im Rahmen der NPLS-Regression. Die Prädiktormatrix der NPLS-Regression setzt sich jeweils aus den kardialen und elektrodermalen Parametern zusammen.

7.2.3 Versuchsablauf

Der Versuch der Studie II fand im Labor des Instituts für Arbeitswissenschaft und Betriebsorganisation (ifab) des KIT statt. Vor dem Versuchsstart informierte die Versuchsleitung alle Versuchspersonen mündlich und schriftlich über den Zweck der Datenerfassung. Danach unterzeichnete die Versuchsperson eine Einverständniserklärung zur Teilnahme am Experiment. Im Anschluss brachte die Versuchsleitung die Messelektroden der kardialen und elektrodermalen Messapparatur an der Versuchsperson an. Nachdem die Versuchsperson den Sitzplatz vor dem Versuchsrechner eingenommen hatte, wurde der Versuchsperson das Blickregistrierungsgerät² aufgesetzt. Vor dem Start des Experiments wurden alle Sensoren auf ihre Funktionstüchtigkeit sowie Signalplausibilität geprüft. Daran anschließend instruierte die Versuchsleitung die Versuchsperson über den Versuchsablauf, die Aufgaben sowie die Fragebögen. Alle Versuchspersonen wurden gebeten, sich im Experiment so wenig wie möglich zu bewegen. Nach

² Im Rahmen des Experiments wurden okulomotorische Daten erfasst. Diese bilden jedoch keinen Bestandteil der vorliegenden Arbeit.

der Klärung offener Fragen seitens der Versuchsperson, wurde das Experiment gestartet. Dieses fand automatisiert am Versuchsrechner statt, um Versuchsleitereffekte zu vermeiden. Als Experimentalsoftware wurde OpenSesame (Mathôt, Schreij & Theeuwes, 2011) verwendet. Zu Beginn des Experiments durchliefen die Versuchspersonen einen Übungsdurchlauf, um sich an die Experimentalumgebung zu gewöhnen und mögliche Fragen zu klären. Im Anschluss wurden die 12 Experimentalbedingungen in randomisierter Reihenfolge bearbeitet. Jede Experimentalbedingung dauerte 60 s. Nach jeder Experimentalbedingung beantworteten die Versuchspersonen den NASA-RTLX. Am Ende des Experiments beantwortete die Versuchsperson einen Abschlussfragebogen (Kontrollvariablen) und legte alle Sensoren ab. Anschließend wurde der Versuchsperson für die Teilnahme gedankt und sie wurde vom Versuchsleiter verabschiedet. Der gesamte Versuch dauerte ca. 60 min.

7.2.4 Abschlussbefragung und Kontrollvariablen

Der Abschlussfragebogen der Studie II umfasste die soziodemografischen Daten der Versuchspersonen und wurde im Anschluss an den Versuch erhoben. Hierzu gehörten die Variablen Alter, Geschlecht und beruflicher Hintergrund. Als zusätzliche Kontrollvariablen wurden die Temperatur und die relative Luftfeuchtigkeit zu Beginn und am Ende des Experiments erhoben.

7.2.5 Manipulationskontrolle

Zur Beantwortung der Forschungsfragen F1, F2 und F6 ist es notwendig, dass die Kriteriumsvariablen (NASA-RTLX-Werte und Trefferquote) über die Experimentalbedingungen hinweg variieren. Ob die Kriterien durch die experimentelle Manipulation erwartungskonform variieren, soll vorab statistisch geprüft werden. Zu diesem Zweck werden Hypothesen (H) für die Kontrollfragestellung KF2 im folgenden Abschnitt formuliert und deren Auswertemethode im Anschluss beschrieben.

Hypothesen der Manipulationskontrolle

Zunächst werden die Effekte auf die mentale Beanspruchung geprüft. Für die subjektive mentale Beanspruchung (NASA-RTLX-Werte) wird ein Anstieg mit zunehmender Aufgabenschwierigkeit angenommen. Ein Einfluss der visuellen Komplexität oder des Bildmaterials auf die NASA-RTLX-Werte wird nicht erwartet. Über die Haupteffekte hinaus wird explorativ der Interaktionseffekt Aufgabenschwierigkeit \times Komplexität \times Bildmaterial getestet. Da zur Art der Interaktion keine Annahmen getroffen werden können, wird diese im Rahmen einer Unterfragestellung (UF) der KF2 geprüft. Folgende Hypothesen und Unterfragestellung ergeben sich in Bezug auf die NASA-RTLX-Werte.

- H2.1** Je höher die Aufgabenschwierigkeit, desto höher sind die NASA-RTLX-Werte.
- H2.2** Es gibt keinen bedeutsamen Effekt der visuellen Komplexität auf die NASA-RTLX-Werte.
- H2.3** Es gibt keinen bedeutsamen Effekt des Bildmaterials auf die NASA-RTLX-Werte.
- UF2.1** Gibt es bedeutsame Interaktionseffekte zwischen der Aufgabenschwierigkeit, der visuellen Komplexität und dem Bildmaterial auf die NASA-RTLX-Werte?

Gleichgerichtete Annahmen werden für die leistungsbezogene mentale Beanspruchung (Trefferquote) getroffen:

- H2.4** Je höher die Aufgabenschwierigkeit, desto geringer ist die Trefferquote.
- H2.5** Es gibt keinen bedeutsamen Effekt der visuellen Komplexität auf die Trefferquote.
- H2.6** Es gibt keinen bedeutsamen Effekt des Bildmaterials auf die Trefferquote.

UF2.2 Gibt es bedeutsame Interaktionseffekte zwischen der Aufgabenschwierigkeit, der visuellen Komplexität und dem Bildmaterial auf die Trefferquote?

Auswertemethoden der Manipulationskontrolle

Die Hypothesen der Manipulationskontrolle werden mit Hilfe der Mehrebenenanalysen (MEA) geprüft. Hierbei wird für jedes Kriterium (NASA-RTLX-Werte und Trefferquote) eine separate MEA berechnet. Der Aufbau der MEA ist für jedes Kriterium identisch. Diese prüft auf Unterschiede zwischen den Schwierigkeitsstufen (2-stufig) in der arithmetischen Aufgabe. Gleichzeitig testet die MEA auf Unterschiede zwischen den Komplexitätsstufen (2-stufig) und den Bildmaterialstufen (3-stufig) des visuellen Materials. Zur Berechnung der MEA werden die Schwierigkeits-, Bildmaterial- und Komplexitätsstufen dummykodiert. Die Referenzgruppe bildet die leichte Schwierigkeitsstufe (S1) bei niedriger visueller Komplexität (-) und der bunten Zeichnung (b). Zusätzlich zu den Haupteffekten wird auf Interaktionseffekte zwischen der Schwierigkeit, der Komplexität und dem Bildmaterial geprüft. Zu diesem Zweck werden aus den Dummyvariablen Interaktionsvariablen gebildet. Insgesamt ergeben sich fünf Zweifach- und zwei Dreifachinteraktionen. Für die Hypothesentests der Regressionsgewichte wird aufgrund der Vielzahl an Dummyvariablen eine Bonferroni-Korrektur durchgeführt, um eine Alphafehlerkumulierung zu vermeiden. Das korrigierte Alphaniveau liegt hierfür bei $\alpha = .005$.

7.2.6 Stichprobe

Am Experiment nahmen 50 Versuchspersonen freiwillig teil. Diese Stichprobe bestand zu 90 % aus Studenten und zu 10 % aus Berufstätigen. 42 % waren männlichen und 58 % weiblichen Geschlechts. Der Altersdurchschnitt lag bei 25.0 ($SD = 3.6$) Jahren.

Bei den physiologischen Daten mussten Personen aufgrund von Sensorausfällen oder Artefakten aus der weiteren Analyse ausgeschlossen werden. Es gelten folgende Stichproben: Die Analyse der kardialen Aktivität wurde mit 49 Personen durchgeführt. Eine Person wurde aufgrund der Einnahme von Herz-Kreislauf-Medikamenten aus der weiteren Analyse ausgeschlossen. Die elektrodermale Aktivität wurde für 48 Personen analysiert. Hier mussten zwei Personen aufgrund von Sensorausfällen aus der weiteren Analyse ausgeschlossen werden.

7.3 Ergebnisse

In den folgenden Abschnitten werden die Ergebnisse zu den Forschungsfragen nacheinander berichtet. Zu Beginn wird in Abschnitt 7.3.1 geprüft, ob die Induktion mentaler Beanspruchung im Experiment gelungen ist. Es folgt die Extraktion kardialer Profile mentaler Beanspruchung (Abschnitt 7.3.2 und 7.3.3) und elektrodermalen Profile mentaler Beanspruchung (Abschnitt 7.3.4 und 7.3.5). Zu diesem Zweck werden mehrere NPLS-Regressionen der zweidimensionalen Kriteriumsmatrix (Bedingung \times Person) auf die dreidimensionale Prädiktormatrix (Bedingung \times Person \times Parameter) durchgeführt. Die Zentrierung und Normierung der Matrizen erfolgt gemäß den Vorgaben im Unterkapitel 5.6.2. Zum Schluss wird im Sinne der Strukturvalidität geprüft, ob die kardialen (Abschnitt 7.3.6) und elektrodermalen (Abschnitt 7.3.7) Profile mentaler Beanspruchung kongruent zu den physiologischen Profilen aus der Studie I sind. Die deskriptive Statistik kardialer und elektrodermalen Parameter ist aus Platzgründen im Anhang A.7.1 und A.7.2 beschrieben.

Tabelle 7.1: Deskriptive Statistiken der Kontrollvariablen

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
Anfangstemperatur	50	22.82	0.88	21.2	24.8	3.6	0.02
Endtemperatur	50	23.07	0.84	21.5	25.0	3.5	0.12
Luftfeuchtigkeit (Beginn)	50	37.90	3.98	31.0	44.0	13.0	-0.06
Luftfeuchtigkeit (Ende)	49	39.04	3.87	32.0	46.0	14.0	-0.11

7.3.1 Kontrollfragestellung KF2: Ist die experimentelle Induktion mentaler Beanspruchung in Studie II gelungen?

In den folgenden Abschnitten werden die Statistiken und Analysen für die Kontrollvariablen, die NASA-RTLX-Werte sowie die Trefferquote berichtet.

Analyse der Kontrollvariablen

Die deskriptiven Statistiken der Kontrollvariablen sind in Tabelle 7.1 dargestellt. Für die relative Luftfeuchtigkeit fehlt ein Messpunkt, da dieser nicht notiert wurde. Ein t-Test für abhängige Stichproben³ deckt auf, dass es sowohl eine signifikante Steigung in der Raumtemperatur ($t(47) = -5.88, p < .001, d = 0.83$) als auch in der relativen Luftfeuchtigkeit ($t(48) = -8.26, p < .001, d = 1.18$) zum Ende des Experiments gibt. Die Temperatur steigt im Mittel um 0.25 °C und die relative Luftfeuchtigkeit im Mittel um 1.1 %.

³ Signifikanzniveau von 5 % (zweiseitig).

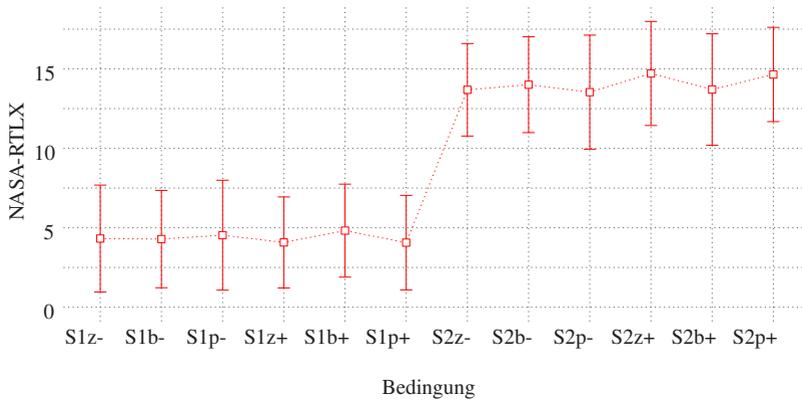


Abbildung 7.1: Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.

Analyse des NASA-RTLX

Die NASA-RTLX-Werte für die 12 Experimentalbedingungen sind in Abbildung 7.1 dargestellt. Eine Tabelle mit den deskriptiven Statistiken des NASA-RTLX findet sich im Anhang A-17. Ausreißer ($\pm 3 SD$) oder fehlende Datenpunkte liegen nicht vor. Deskriptiv ist bereits ein erwartungskonformer Trend im Sinne einer Zunahme der NASA-RTLX-Werte für die schweren (S2) im Vergleich zu den einfachen (S1) arithmetischen Aufgaben zu erkennen. Ob dieser Trends der Aufgabenschwierigkeit sich inferenzstatistisch bestätigt, wird mit der MEA überprüft. Diese deckt die Hypothesen H2.1 bis H2.3 und Unterfragestellungen UF2.1 innerhalb der Manipulationskontrolle ab (siehe Abschnitt 7.2.5).

Die Ergebnisse des Modellvergleichs in Tabelle 7.2 decken auf, dass das Haupteffektmodell der Aufgabenschwierigkeit sowie das Interaktionsmodell aus Schwierigkeit \times Komplexität \times Bildmaterial eine bedeutsame Reduktion der Devianz und des AIC aufweisen. Das Modell M_5 klärt dabei 70.2 % der Varianz auf. Der zuvor beschriebene Trend wird durch den signifikanten Anstieg

Tabelle 7.2: Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	3799.00	3793.00			0.00
+ Schwierigkeit (M_S)	4	2795.77	2787.77	1005.23***	1	70.21
+ Komplexität (M_K)	5	2795.43	2785.43	2.34	1	70.27
+ Bildmaterial (M_B)	7	2799.43	2785.43	0.00	2	70.27
+ Schwier. \times Kompl. \times Bildmat. ($M_{S \times K \times B}$)	14	2797.90	2769.90	15.56*	7	70.64

*** $p < .001$, ** $p < .01$, * $p < .05$.

der NASA-RTLX-Werte für die schweren im Vergleich zu den einfachen Experimentalaufgaben bestätigt, $B = 9.7$, $t(550) = 53.58$, $p < .001$. Die Teststärke für diesen Effekt liegt bei der vorliegenden Stichprobe bei 100.0 %. Im Modellvergleich (siehe Tabelle 7.2) zeigt das Interaktionsmodell $M_{S \times K \times B}$ zwar einen bedeutsamen Effekt, jedoch wird das korrigierte Signifikanzniveau von $\alpha = .005$ von keinem Regressionsgewicht der Interaktionsvariablen erreicht.

Analyse der Trefferquote

Die Trefferquote für die 12 Experimentalbedingungen ist in Abbildung 7.2 dargestellt. Eine Tabelle mit den deskriptiven Statistiken der Trefferquote findet sich im Anhang A-18. Ausreißer ($\pm 3 SD$) oder fehlende Datenpunkte liegen nicht vor. Im Gesamtdurchschnitt wurden ca. 85 % ($SD = 25.5$) der Multiplikationsaufgaben richtig gelöst. Deskriptiv folgt die Trefferquote in den Versuchsbedingungen einem gespiegelten Verlauf der zuvor beschriebenen NASA-RTLX-Werte. Zudem lassen sich Interaktionseffekte vermuten. Ob sich bedeutsame Effekte finden, wird mit der MEA überprüft. Diese deckt die Hypothesen H2.4 bis H2.6 sowie die Unterfragestellungen UF2.2 innerhalb der Manipulationskontrolle ab (siehe Abschnitt 7.2.5).

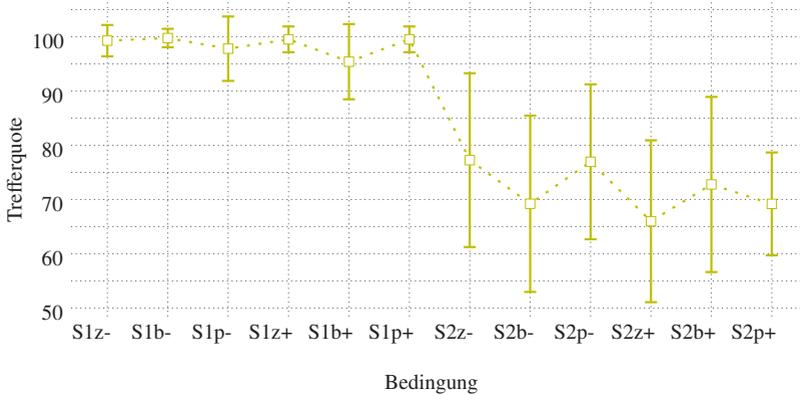


Abbildung 7.2: Mittelwert und Standardabweichung der Trefferquote in Prozent in den Versuchsbedingungen.

Die Ergebnisse der geschachtelten Modellvergleiche in Tabelle 7.3 zeigen eine signifikante Reduktion der Devianz und des AIC sowohl für die Haupteffekte Schwierigkeit (M_S) und Komplexität (M_K) als auch für die Dreifachinteraktion Schwierigkeit \times Komplexität \times Bildmaterial ($M_{S \times F \times B}$). Allerdings wird der Hauptanteil an erklärter Varianz (58.8 %) durch den Haupteffekt der Schwierigkeit vorhergesagt. Die anderen Effekte erklären nur 0.7 % bis 2.6 % zusätzlicher

Tabelle 7.3: Modellfit-Indizes für die geschachtelten Modelle der Trefferquote

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	5134.73	5128.73			0.00
+ Schwierigkeit (M_S)	4	4596.60	4588.60	540.13***	1	58.83
+ Komplexität (M_K)	5	4587.14	4577.14	11.46***	1	59.56
+ Bildmaterial (M_B)	7	4588.74	4574.74	2.40	2	59.71
+ Schwier. \times Kompl. \times Bildmat. ($M_{S \times K \times B}$)	14	4563.40	4535.40	39.31***	7	62.11

*** $p < .001$, ** $p < .01$, * $p < .05$.

Varianz in der Trefferquote. Die Regressionsgewichte der einzelnen Faktorstufen sind in Tabelle 7.4 für das Modell $M_{S \times F \times B}$ dargestellt. Erwartungskonform findet sich ein signifikanter Abfall der Trefferquote bei Zunahme der Aufgabenschwierigkeit (Teststärke bei 100.0 %). Zudem zeigen sich Effekte der visuellen Komplexität und des Bildmaterials in der Interaktion mit der Schwierigkeit (Schwierigkeit \times Bildmaterial_{b-z}, Schwierigkeit \times Bildmaterial_{b-p}, Schwierigkeit \times Komplexität \times Bildmaterial_{b-z} und Schwierigkeit \times Komplexität \times Bildmaterial_{b-p}; siehe Tabelle 7.4). Die Interaktionen können in Abbildung 7.2 veranschaulicht werden: Bei geringer Aufgabenschwierigkeit und geringer visueller Komplexität gibt es keine Unterschiede in der Trefferquote zwischen dem Bildmaterial (vgl. Bedingungen S1z-, S1b- und S1p-). Ist die visuelle Komplexität bei geringer Schwierigkeit allerdings hoch, dann

Tabelle 7.4: Parameter für das Modell $M_{S \times K \times B}$ der Trefferquote

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	99.76***	1.51	65.92	568.32	2.86
S: 2 - 1	-30.54***	2.06	-14.81	550.00	
K: (+) - (-)	-4.36	2.06	-2.11	550.00	
B: b - z	-0.48	2.06	0.23	550.00	
B: b - p	-1.96	2.06	-0.95	550.00	
S \times K	7.92	2.92	2.71	550.00	
S \times B _{b-z}	8.52*	2.92	2.92	550.00	
S \times B _{b-p}	9.70**	2.92	3.32	550.00	
K \times B _{b-z}	4.60	2.92	1.58	550.00	
K \times B _{b-p}	6.08	2.92	2.08	550.00	
S \times K \times B _{b-z}	-19.42***	4.13	-4.71	550.00	
S \times K \times B _{b-p}	-17.40***	4.13	-4.22	550.00	

*** $p < .0001$, ** $p < .001$, * $p < .005$.

Anmerkungen. S = Schwierigkeit, K = visuelle Komplexität, B = Bildmaterial.

zeigt sich ein geringer Abfall in der Trefferquote für die bunten Zeichnungen im Vergleich zum restlichen Bildmaterial (vgl. Bedingungen S1b+ mit S1z+ und S1p+). Bei hoher Aufgabenschwierigkeit ist die Trefferquote insgesamt geringer (Haupteffekt der Aufgabenschwierigkeit), jedoch ebenfalls abhängig von der Ausprägung der visuellen Komplexität und des Bildmaterials: Bei geringer visueller Komplexität ist die Trefferquote für die bunten Zeichnungen (im Vergleich zu den schwarz-weiß Zeichnungen oder zu den bunten Fotos) geringer (vgl. Bedingung S2b- mit S2z- und S2p-) und bei hoher visueller Komplexität höher (vgl. Bedingung S2b+ mit S2z+ und S2p+). Die Teststärken für die bedeutsamen Interaktionseffekte liegen zwischen 82.8 % und 99.6 %.

7.3.2 Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

In einem ersten Schritt wird die Anzahl der Komponenten bestimmt sowie die Komponentenmodelle kreuzvalidiert. Die Ergebnisse der Kreuzvalidierung sind für die Regression vom NASA-RTLX auf die kardialen Parameter in Tabelle 7.5 dargestellt.⁴ Aus den Modellvergleichen ist zu folgern, dass nur die erste Komponente einen signifikanten Zuwachs in der Vorhersagegüte der NASA-RTLX-Werte zeigt. Im Originalmodell erklärt diese Komponente 6.9 % der Varianz in den kardialen Parametern und 82.4 % der Varianz in den NASA-RTLX-Werten. In der kreuzvalidierten Vorhersage verringert sich dieser Anteil auf 55.8 %. Die Ladungen der Komponente sind in Abbildung 7.3 zu sehen.

⁴ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.40 [0.13, 0.77]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{KLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -4.14 [-12.45, 0.56]$.

Tabelle 7.5: NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR^2_{PLS}) und das kreuzvalidierte Modell ($\text{val}\Delta R^2_{\text{PLS}}$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	valSQR
0					810.82
1	6.86	5.34 [†]	82.37	55.81 [†]	238.16***
2	4.23	2.62 [†]	5.85	3.16 [†]	221.13
3	6.00	2.83 [†]	1.95	-0.30	222.75
4	5.03	3.76 [†]	1.95	-1.25	229.50
5	4.55	2.48 [†]	1.75	-0.82	233.92

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.77$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 0.56$.

Die aus der NASA-RTLX-Matrix extrahierte Komponente 1 ist in ihren Bedingungs-Faktorladungen (siehe Abbildung 7.3, Mitte, gestrichelte rote Linie) identisch mit dem bereits bekannten Verlauf der NASA-RTLX-Mittelwerte (gestrichelte dunkelgraue Linie in Abbildung 7.3, Mitte). Dabei zeigen alle Bedingungen mit den schweren arithmetischen Aufgaben eine überdurchschnittliche (positive) und alle Bedingungen mit den leichten Multiplikationsrechnungen eine unterdurchschnittliche (negative) Faktorladung. Somit kann die NASA-RTLX-Komponente 1 als ein Repräsentant der subjektiven mentalen Beanspruchung interpretiert werden. In der Bedingungsdimension kovariiert die NASA-RTLX-Komponente 1 mit der kardialen Komponente 1 (durchgezogene rote Linie in Abbildung 7.3, Mitte). Hierbei kann die kardiale Komponente 1 die NASA-RTLX-Komponente 1 mit $B = 0.71$, $t(586) = 22.11$, $p < .001$ vorhersagen. Es gibt keine bedeutsamen ($|a_{\Delta}| \geq .20$) oder marginal bedeutsamen ($|a_{\Delta}| \geq .15$) Unterschiede zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente 1 und der kardialen Komponente 1 (tabellarische Darstellung der Bedingungs-Faktorladungen siehe Anhang A-19). Wie bereits bei der NASA-RTLX-Komponente 1 liegen auch bei der kardialen

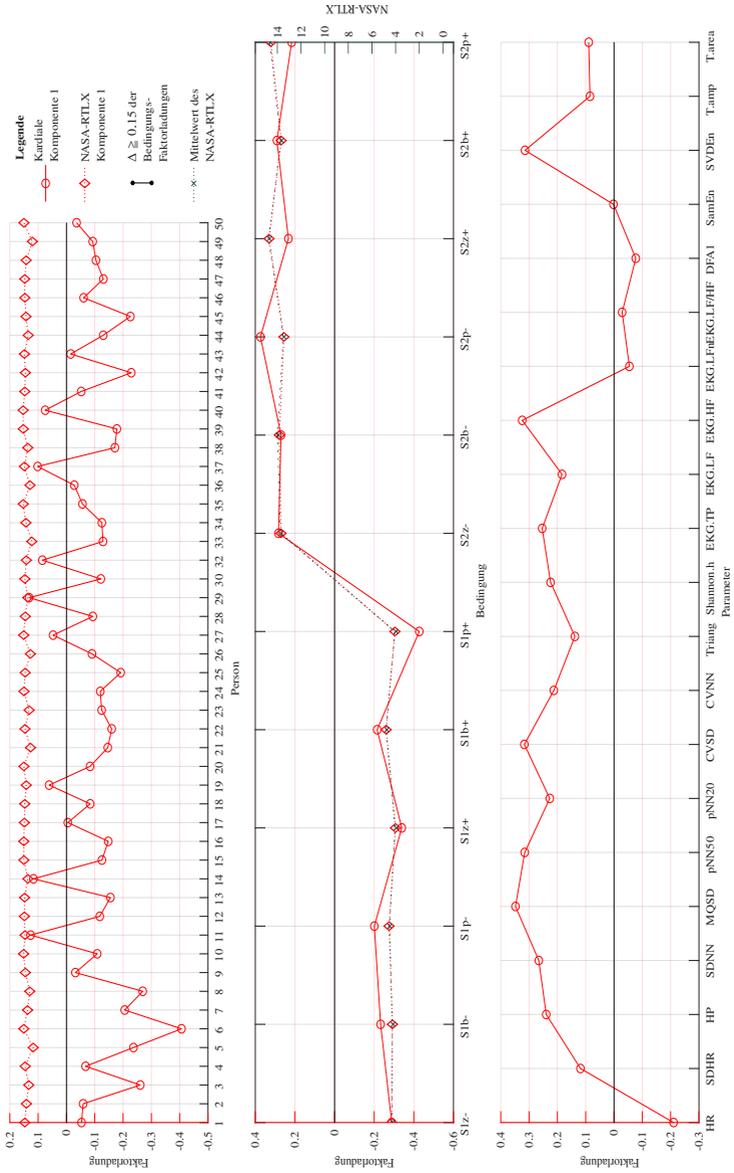


Abbildung 7.3: Faktorladungen α für die NASA-RTLX und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.

Komponente 1 die Faktorladungen für die Bedingungen mit schweren Multiplikationsrechnungen im überdurchschnittlichen und für die Bedingungen mit leichten Multiplikationsrechnungen im unterdurchschnittlichen Bereich (vgl. Abbildung 7.3, Mitte).

Die Faktorladungen der Parameter (Abbildung 7.3, unten) bilden das kardiale Profil der subjektiven mentalen Beanspruchung ab. Bedeutsame Faktorladungen auf der kardialen Komponente 1 ($|a| \geq .30$) besitzen die Parameter MQSD, pNN50, CVSD, EKG.HF und SVDEn. Parameter mit marginal bedeutsamen Faktorladungen ($|a| \geq .20$) sind HR, HP, SDNN, pNN20, CVNN, Shannon.h und EKG.TP. Die Faktorladungen der Personen sind in Abbildung 7.3 (oben) zu sehen. Die Person-Faktorladungen der NASA-RTLX-Komponente (gestrichelte rote Linie) weisen geringe interindividuelle Unterschiede auf, wohingegen die kardialen Person-Faktorladungen (durchgezogene rote Linie) hohe interindividuelle Unterschiede zeigen.

7.3.3 Forschungsfrage F1.2: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind in Tabelle 7.6 dargestellt.⁵ Hier zeigt sich ein bedeutsamer Zuwachs an erklärter Varianz in der kardialen Parametermatrix für alle fünf Komponenten. Von den fünf Komponenten kann jedoch nur die erste Komponente die Trefferquote signifikant vorhersagen. Sie erklärt im Originalmodell 6.7 % der Varianz in den kardialen Parametern und sagt 63.6 % der Varianz in der Trefferquote vorher. In der Kreuzvalidierten

⁵ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.40[0.13, 0.77]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -4.14[-12.45, 0.56]$.

Tabelle 7.6: NPLS-Regression der Trefferquote auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					623.49
1	6.66	5.11 [†]	63.61	40.36 [†]	321.44***
2	4.59	2.62 [†]	11.41	1.61 [†]	312.76
3	4.66	2.70 [†]	5.48	-2.05	323.82
4	4.69	2.40 [†]	4.74	-2.44	336.98
5	3.47	1.41 [†]	4.78	-3.78	357.35

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.77$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 0.56$.

Vorhersage verringert sich der Anteil erklärter Varianz für unbekannte Personen auf 40.4 %. Die Ladungen der Komponente 1 sind in Abbildung 7.4 dargestellt.

Zur Interpretation der Werte wird die Bedingungsdimension (Abbildung 7.4, Mitte) herangezogen. Eine Tabelle der Bedingungs-Faktorladungen findet sich im Anhang A-20. Die Faktorladungen der Trefferquote (gestrichelte olivfarbene Linie) sind in der Bedingungsdimension identisch zu den Bedingungs-mittelwerten der Trefferquote (gestrichelte dunkelgraue Linie). Somit ist die extra-hierte Trefferquoten-Komponente als ein Repräsentant der leistungsbezogenen mentalen Beanspruchung zu interpretieren. In der Bedingungsdimension kovariert die Trefferquoten-Komponente 1 mit der kardialen Komponente 1. Die Trefferquoten-Komponente 1 kann mit $B = -0.63$, $t(586) = -13.71$, $p < .001$ durch die kardiale Komponente 1 vorhergesagt werden. Die Bedingungs-Faktorladungen der kardialen Komponente 1 sind in Abbildung 7.4 (Mitte) als

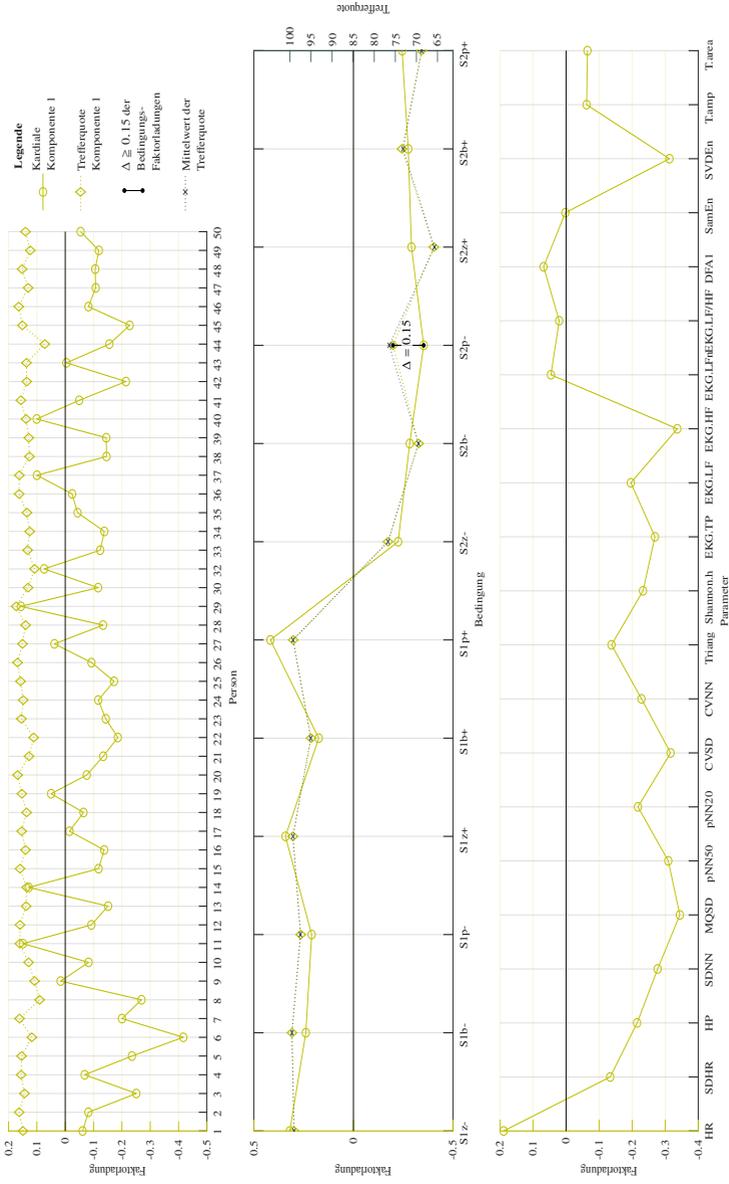


Abbildung 7.4: Faktorladungen α für die Trefferskote und kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferskote in den Bedingungen (Mitte) angegeben.

durchgezogene olivfarbene Linie zu sehen.⁶ Es besteht eine hohe Übereinstimmung zwischen den Bedingungs-Faktorladungen der (invertierten) kardialen Komponente und der Trefferquoten-Komponente. In beiden Komponenten ist der Haupteffekt der Aufgabenschwierigkeit erkennbar. So zeigen sich überdurchschnittliche Faktorladungen für die leichten Multiplikationsaufgaben und unterdurchschnittliche Faktorladungen für die schweren Multiplikationsaufgaben (siehe Abbildung 7.4, Mitte). Im Gegensatz zur Trefferquoten-Komponente 1 wird der Interaktionseffekt (Schwierigkeit \times Komplexität \times Bildmaterial) – welcher für die Trefferquote gefunden wurde (siehe Abschnitt 7.3.1) – von der kardialen Komponente 1 nur für die leichten Aufgaben (S1) und nicht für die schweren Aufgaben (S2) abgebildet. Letzteres ist an der marginalen Abweichung der Bedingung S2p- erkennbar.

Die Faktorladungen der Parameter (Abbildung 7.4, unten) bilden das kardiale Profil der Trefferquote. Bedeutsame Faktorladungen ($|a| \geq .30$) besitzen die Parameter MQSD, pNN50, CVSD, EKG.HF und SVDEn. Marginal bedeutsame Faktorladungen ($|a| \geq .20$) zeigen die Parameter HP, SDNN, pNN20, CVNN, Shannon.h, EKG.TP und EKG.LF. In Abbildung 7.4 (oben) ist anhand der Person-Faktorladungen zu erkennen, dass es hohe interindividuelle Unterschiede für die kardiale Komponente, jedoch nicht die Trefferquoten-Komponente, gibt.

Tabelle 7.7: NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das Kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					524.52
1	6.85	5.22 [†]	82.52	51.20 [†]	257.69***
2	1.44	0.77	5.75	-28.62	408.79
3	3.42	0.33	2.45	-0.08	409.21
4	1.14	-0.66	1.64	0.80 [†]	405.01
5	-0.28	-0.65	1.61	0.64 [†]	401.63

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.77$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 0.56$.

7.3.4 Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind für die Regression vom NASA-RTLX auf die elektrodermalen Parameter in Tabelle 7.7 dargestellt.⁷ Hier zeigt sich ein bedeutsamer Zuwachs an erklärter Varianz in der elektrodermalen Parametermatrix für die erste Komponente, welche im Originalmodell 6.7 % der Varianz erklärt. Trotz des geringen Varianzanteils in der Prädiktormatrix kann die elektrodermale Komponente 1 im Originalmodell 82.5% der Varianz

⁶ Es ist zu beachten, dass die Person- und Bedingungs-Faktorladungen der kardialen Komponente in Abbildung 7.4 invertiert sind: Negative Werte bedeuten eine über- und positive Werte eine unterdurchschnittliche kardiale Aktivität.

⁷ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.40[0.13, 0.77]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -4.14[-12.45, 0.56]$.

in den NASA-RTLX-Werten vorhersagen. In der kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz auf 51.2 %. Die Ladungen der Komponente 1 sind in Abbildung 7.5 dargestellt.

Die extrahierte Komponente 1 des NASA-RTLX zeigt in den Bedingungs-Faktorladungen (siehe Abbildung 7.5, Mitte, gestrichelte rote Linie) einen identischen Verlauf zu den Bedingungs-Mittelwerten des NASA-RTLX (gestrichelte dunkelgraue Linie). Somit ist die NASA-RTLX-Komponente 1 als ein Repräsentant der subjektiven mentalen Beanspruchung zu interpretieren. In der Bedingungsdimension kann die elektrodermale Komponente 1 (siehe Abbildung 7.5, Mitte, durchgezogene rote Linie) die NASA-RTLX-Komponente 1 mit $B = 0.92$, $t(574) = 25.24$, $p < .001$ vorhersagen. Eine Tabelle mit den Bedingungs-Faktorladungen der Komponenten findet sich im Anhang A-21. Wie in Abbildung 7.5 (Mitte) zu erkennen ist, liegt die elektrodermale Aktivität bei schweren Aufgaben über und bei leichten Aufgaben unter dem Durchschnitt. Marginale oder bedeutsamen Unterschiede zwischen den Bedingungs-Faktorladungen der NASA-RTLX- und elektrodermalen Komponente 1 gibt es nicht.

Die Faktorladungen der elektrodermalen Parameter (Abbildung 7.5, unten) definieren das elektrodermale Profil. Die Parameter Phasic.SD, EDR.N, EDR.amp, EDA.LF und EDA.HF laden mit $|a| \geq .30$ bedeutsam auf der elektrodermalen Komponente. Marginale Faktorladungen ($|a| \geq .20$) ergeben sich für die Parameter EDL.D, EDR.area, EDR.rise.v, EDR.peaks und EDA.VHF. An den Faktorladungen der Personen (Abbildung 7.5, oben) ist zu erkennen, dass für die NASA-RTLX-Komponente (gestrichelte rote Linie) geringe und für die elektrodermale Komponente (durchgezogene rote Linie) hohe interindividuelle Unterschiede vorliegen.

7.3.5 Forschungsfrage F2.2: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die leistungsbezogene mentale Beanspruchung zuverlässig repräsentieren?

Die Ergebnisse der Kreuzvalidierung sind in Tabelle 7.8 dargestellt.⁸ Von den fünf Komponenten kann nur die Komponente 1 die Trefferquote bedeutsam vorhersagen. Sie erklärt im Originalmodell 6.8 % der Varianz in den elektrodermalen Parametern und sagt 63.4 % der Varianz in der Trefferquote vorher. In der kreuzvalidierten Vorhersage sinkt der Anteil erklärter Varianz auf 37.2 % ab. Die Ladungen der Komponente 1 sind in Abbildung 7.6 dargestellt.

Tabelle 7.8: NPLS-Regression der Trefferquote auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($\text{val}\Delta R_{\text{PLS}}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	ΔR_{PLS}^2	$\text{val}\Delta R_{\text{PLS}}^2$	valSQR
0					753.21
1	6.79	5.12 [†]	63.36	37.15 [†]	331.86 ^{***}
2	2.55	1.27 [†]	6.83	1.60 [†]	323.42
3	4.13	0.64	7.39	-7.54	363.26
4	-0.50	0.21	7.13	-2.60	377.01
5	1.06	-0.83	5.40	-2.01	387.62

Bedeutsamer Abfall der valSQR: ^{***} $p < .001$, ^{**} $p < .01$, ^{*} $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.77$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 0.56$.

Zur Interpretation werden die Faktorladungen der Bedingungen (Abbildung 7.6, Mitte) herangezogen. Eine Tabelle der Bedingungs-Faktorladungen ist im

⁸ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktormatrix beträgt: $M [99 \% CI] = 0.40 [0.13, 0.77]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -4.14 [-12.45, 0.56]$.

Anhang A-22 zu finden. Die Faktorladungen der Trefferquote (gestrichelte olivfarbene Linie in Abbildung 7.6, Mitte) stimmen mit den Bedingungsmittelwerten der Trefferquote (gestrichelte dunkelgraue Linie) überein. Damit kann die extrahierte Trefferquoten-Komponente als ein Repräsentant der leistungsbezogenen mentalen Beanspruchung interpretiert werden. Die extrahierte elektrodermale Komponente 1 (Abbildung 7.6, Mitte, durchgezogene olivfarbene Linie) kann die Trefferquoten-Komponente 1 mit $B = -0.81$, $t(574) = 15.35$, $p < .001$ in der Bedingungsdimension vorhersagen.⁹ Bedeutsame oder marginale Unterschiede zwischen den Bedingungs-Faktorladungen der Trefferquoten-Komponente 1 und der (invertierten) elektrodermalen Komponente 1 gibt es nicht. Nichtsdestoweniger ist in Abbildung 7.6 (Mitte) erkennbar, dass die elektrodermale Komponente 1 zwar den Haupteffekt der Aufgabenschwierigkeit, jedoch nicht den Interaktionseffekt (Schwierigkeit \times Komplexität \times Bildmaterial) – der für die leistungsbezogene mentale Beanspruchung gefunden wurde – wiedergibt: fehlende Sprünge der elektrodermalen Komponente 1 für die Bedingungen mit hoher Aufgabenschwierigkeit.

Die Faktorladungen der elektrodermalen Parameter (Abbildung 7.6, unten) definieren das elektrodermale Profil der Trefferquote. Bedeutsame Faktorladungen ($|a| \geq .30$) haben die Parameter Phasic.SD, EDR.N, EDR.amp, EDA.LF und EDA.HF. Marginale Faktorladungen ($|a| \geq .20$) ergeben sich für die Parameter EDL.D, EDR.area, EDR.rise.v, EDR.peaks und EDA.VHF. Die Faktorladungen der Personen sind in Abbildung 7.6 (oben) abgebildet. Die Person-Faktorladungen der Trefferquoten-Komponente (gestrichelte olivfarbene Linie) besitzen kaum interindividuelle Unterschiede, wohingegen die elektrodermalen Person-Faktorladungen (durchgezogene olivfarbene Linie) hohe interindividuelle Unterschiede zeigen.

⁹ Es ist zu beachten, dass die Person- und Bedingungs-Faktorladungen der elektrodermalen Komponente in Abbildung 7.6 invertiert sind: Negative Werte bedeuten eine über- und positive Werte eine unterdurchschnittliche elektrodermale Aktivität.

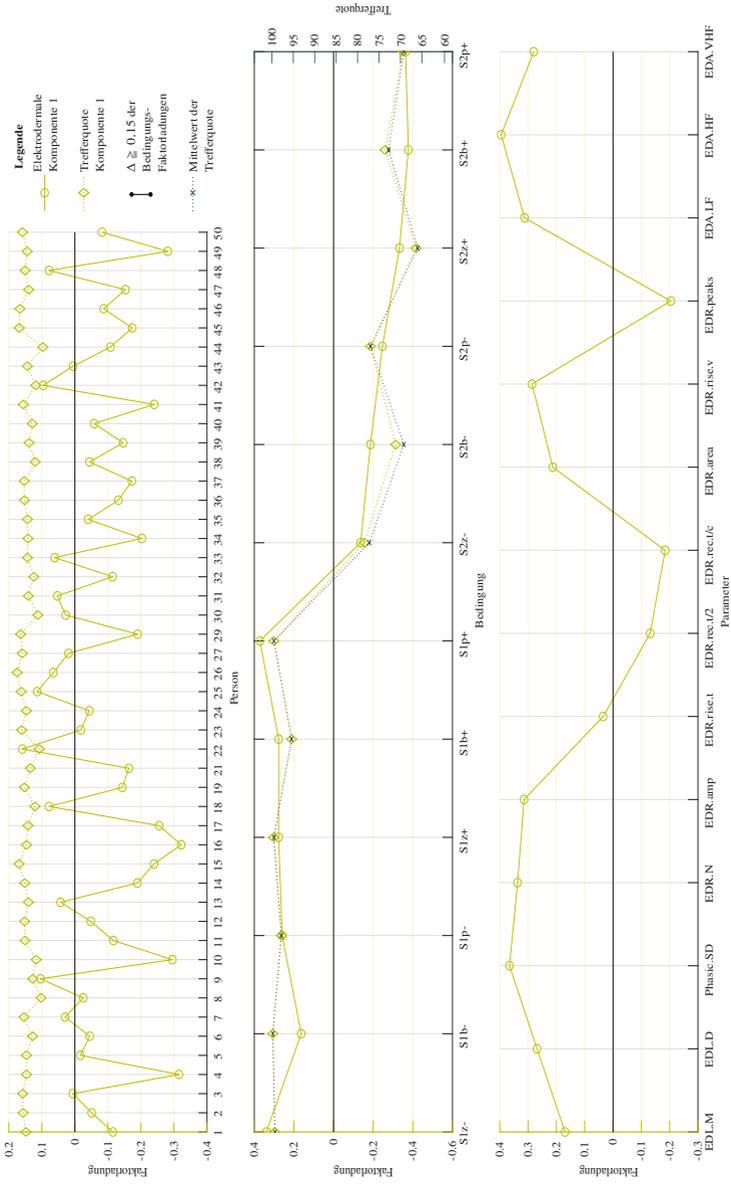


Abbildung 7.6: Faktorladungen α für die Trefferquote und elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte der Trefferquote in den Bedingungen (Mitte) angegeben.

7.3.6 Forschungsfrage F6.1: Sind die in Studie II extrahierten kardialen Profile mentaler Beanspruchung kongruent zu den in Studie I extrahierten Profilen mentaler Beanspruchung?

Zur Beantwortung der Fragestellung F6.1 wird zunächst die Kongruenz zwischen den kardialen Komponenten innerhalb der Studie II bestimmt. Die Kongruenzkoeffizienten zeigen auf, dass die kardiale Komponente des NASA-RTLX und die kardiale Komponente der Trefferquote für alle drei Dimensionen identisch zueinander sind (Person: $C = -.99$, Bedingung: $C = -.99$, Parameter: $C = 1.0$). Auch in Tabelle 7.9 (zweite Spalte) ist zu erkennen, dass sich die Parameterfaktorladungen der kardialen Komponenten aus Studie II um einen Betrag von maximale .03 unterscheiden. Auf Basis dieser Ergebnisse handelt es sich also um dasselbe kardiale Profil der mentalen Beanspruchung. Der Tabelle 7.10 ist zu entnehmen, dass die fünf Parameter MQSD, pNN50, CVSD, EKG.HF und SVDEn bedeutsam auf den zwei identischen Profilen laden.

Da die kardialen Profile des NASA-RTLX und der Trefferquote in Studie II identisch sind, wird nur das kardiale Profil des NASA-RTLX für die weitere Analyse der Kongruenz mit Studie I herangezogen. In Studie I wurde das kardiale Profil des NASA-RTLX als ein Repräsentant der mentalen Beanspruchung interpretiert (siehe Abschnitt 6.4.2). Das kardiale Profil der Trefferquote aus Studie I wurde hingegen als eine Mischung mentaler und emotionaler Beanspruchung gedeutet. Die folgende Kongruenzanalyse erfolgt mit diesen beiden kardialen Profilen aus Studie I. Der Kongruenzkoeffizient der Parameterfaktorladungen belegt mit $C = .98$ eine Identität zwischen den kardialen Profilen des NASA-RTLX aus Studie I und Studie II. Das kardiale Profil des NASA-RTLX aus Studie II und das kardiale Profil der Trefferquote aus Studie I liegen mit $C = .93$ unter der Identitätsschwelle von .95 und gelten demnach als ähnlich. Die hohe Ähnlichkeit bestätigt sich auch in Tabelle 7.9 (letzte Spalte): Hier zeigt nur der Parameter HP eine marginale Differenz von .15 auf. Die Faktorladungen der kardialen Profile aus Studie I und II sind in Tabelle 7.10

Tabelle 7.9: Differenz zwischen den Parameter-Faktorladungen der kardialen Komponente des NASA-RTLX (Na) und der kardialen Komponente der Trefferquote (Tq) aus Studie I und II

Parameter	Na (II) – Tq (II)	Na (I) – Na (II)	Tq (I) – Na (II)
HR	.02	.07	–.12
SDHR	.01	.03	–.10
HP	–.03	–.06	.15
SDNN	.01	.00	–.05
MQSD	–.01	.03	.04
pNN50	–.01	.05	.05
pNN20	–.01	–.07	.03
CVSD	.00	.03	.01
CVNN	.02	.03	–.07
Triang	.00	–.01	–.10
Shannon.h	.01	.02	–.07
EKG.TP	.02	.01	.03
EKG.LF	.02	–.05	–.05
EKG.HF	.02	.02	.08
EKG.LFn	.00	.04	–.12
EKG.LF/HF	.01	.07	–.12
DFA1	.01	.02	–.10
SamEn	.00	–.03	.04
SVDEn	.00	.03	.02
T.amp	–.02	.06	–.06
T.area	–.03	.05	–.09

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, **marginal** $|a_{\Delta}| \geq .15$, **nicht bedeutsam** $|a_{\Delta}| < .15$.

Tabelle 7.10: Parameter-Faktorladungen α für die kardialen Komponenten des NASA-RTLX und der Trefferquote jeweils aus Studie I und II

Parameter	Studie II		Studie I	
	NASA-RTLX	Trefferquote	NASA-RTLX	Trefferquote
HR	.21	.19	.28	.09
SDHR	-.12	-.13	-.09	-.22
HP	-.24	-.21	-.30	-.09
SDNN	-.27	-.28	-.26	-.31
MQSD	-.35	-.34	-.32	-.31
pNN50	-.32	-.31	-.26	-.27
pNN20	-.23	-.22	-.30	-.20
CVSD	-.32	-.32	-.28	-.31
CVNN	-.21	-.23	-.19	-.28
Triang	-.14	-.14	-.15	-.24
Shannon.h	-.22	-.23	-.20	-.29
EKG.TP	-.25	-.27	-.25	-.23
EKG.LF	-.18	-.20	-.23	-.23
EKG.HF	-.32	-.34	-.30	-.25
EKG.LFn	.05	.05	.10	-.06
EKG.LF/HF	.03	.02	.10	-.09
DFA1	.08	.07	.10	-.02
SamEn	.00	.00	-.04	.03
SVDen	-.31	-.31	-.28	-.29
T.amp	-.08	-.06	-.03	-.14
T.area	-.09	-.06	-.04	-.18

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

dargestellt. Hier ist eine hohe Überschneidung in der Ladungsstruktur für marginale Faktorladungen ($a \geq .20$), jedoch nicht für bedeutsame Faktorladungen ($a \geq .30$) sichtbar. Für die identischen kardialen Profile des NASA-RTLX aus Studie I und II (Spalte 2 und 4 in Tabelle 7.10) gibt es nur für die Parameter MQSD und EKG.HF eine überschneidende Bedeutsamkeit der Ladung.

7.3.7 Forschungsfrage F6.2: Sind die in Studie II extrahierten elektrodermalen Profile mentaler Beanspruchung kongruent zu den in Studie I extrahierten Profilen mentaler Beanspruchung?

Im ersten Schritt wird die Ähnlichkeit der in Studie II extrahierten elektrodermalen Komponenten bewertet. Hierbei sind die elektrodermale Komponente des NASA-RTLX und die elektrodermale Komponente der Trefferquote sowohl für die Dimension der Person ($C = -1.0$), der Bedingung ($C = -.99$) und der Parameter ($C = 1.0$) identisch (vgl. Tabelle 7.11, zweite Spalte). Beide elektrodermalen Komponenten bilden somit dasselbe elektrodermale Profil ab. Auf diesem laden gemäß Tabelle 7.12 die fünf Parameter Phasic.SD, EDR.N, EDR.amp, EDA.LF und EDA.HF bedeutsam. Aufgrund des Identitätsnachweises reicht es aus, wenn für die Kongruenzanalyse mit Studie I nur eins der beiden elektrodermalen Profile aus Studie II verwendet wird. Für die weitere Analyse wird somit nur das elektrodermale Profil des NASA-RTLX berücksichtigt.

In Studie I wurde das elektrodermale Profils des NASA-RTLX als ein Repräsentant der mentalen Beanspruchung identifiziert (siehe Abschnitt 6.4.3). Mit diesem wird das elektrodermale Profil des NASA-RTLX aus Studie II verglichen. Der Kongruenzkoeffizient deckt für die Parameter-Faktorladungen einen Unterschied zwischen den beiden Profilen auf, $C = .82$. In Tabelle 7.11 sind bedeutsame Faktorladungsunterschiede in den Parametern EDL.M und EDR.rise.t erkennbar. Zudem zeigen sechs weitere Parameter marginale Unterschiede in ihren Faktorladungen (siehe Tabelle 7.11, dritte Spalte). Dieser

Tabelle 7.11: Differenz zwischen den Parameter-Faktorladungen der elektrodermalen Komponente des NASA-RTLX (Na) aus Studie I und II, der Trefferquote (Tq) aus Studie II und der Valenz (Va) aus Studie I

Parameter	Na (II) – Tq (II)	Na (I) – Na (II)	Na (II) – Va (I)
EDL.M	-.01	.20	.05
EDL.D	-.01	-.04	.06
Phasic.SD	-.01	-.11	.05
EDR.N	-.01	.10	.27
EDR.amp	.01	-.17	-.10
EDR.rise.t	.03	-.29	.00
EDR.rec.t/2	-.02	-.19	-.14
EDR.rec.t/c	-.01	-.14	-.19
EDR.area	.01	-.16	-.06
EDR.rise.v	.00	-.04	-.14
EDR.peaks	-.01	-.15	-.17
EDA.LF	.00	-.19	-.12
EDA.HF	-.01	-.18	-.02
EDA.VHF	.00	-.11	.03

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, marginal $|a_{\Delta}| \geq .15$, nicht bedeutsam $|a_{\Delta}| < .15$.

Unterschied zwischen den Profilen ist ebenfalls in Tabelle 7.12 (vgl. Spalte 2 und 4) anhand der unterschiedlichen Ladungsstruktur erkennbar: Die Parameter EDL.M, EDR.N, EDR.rec.t/2, EDR.rec.t/c und EDR.peaks laden bedeutsam auf dem elektrodermalen Profil des NASA-RTLX aus Studie I, während die Parameter Phasic.SD, EDR.N, EDR.amp, EDA.LF und EDA.HF auf dem elektrodermalen Profil des NASA-RTLX aus Studie II bedeutsame Faktorladungen besitzen.

Da das elektrodermale Profile des NASA-RTLX aus Studie II keine Ähnlichkeit zum elektrodermalen Profil der mentalen Beanspruchung aus Studie I aufweist, wird zusätzlich für die Kongruenz mit dem elektrodermalen Profil der emotionalen Beanspruchung aus Studie I geprüft. In Studie I wurden die

Tabelle 7.12: Parameter-Faktorladungen a für die elektrodermalen Komponenten des NASA-RTLX (aus Studie I und II), der Trefferquote (aus Studie II) und der Valenz (aus Studie I)

Parameter	Studie II		Studie I	
	NASA-RTLX	Trefferquote	NASA-RTLX	Valenz
EDL.M	.16	.17	.36	.11
EDL.D	.26	.27	.22	.19
Phasic.SD	.36	.37	.25	.31
EDR.N	.33	.34	.43	.07
EDR.amp	.32	.31	.16	.42
EDR.rise.t	.06	.03	-.23	.05
EDR.rec.t/2	-.15	-.13	-.33	.00
EDR.rec.t/c	-.19	-.18	-.33	.00
EDR.area	.22	.21	.06	.29
EDR.rise.v	.29	.29	.25	.42
EDR.peaks	-.21	-.20	-.36	-.03
EDA.LF	.31	.31	.12	.43
EDA.HF	.39	.40	.21	.41
EDA.VHF	.28	.28	.17	.25

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

drei identischen elektrodermalen Profile der Trefferquote, der Erregung und der Valenz als Repräsentanten der emotionalen Beanspruchung interpretiert (siehe Abschnitt 6.4.3). Für die Kongruenzanalyse wird das elektrodermale Profil der Valenz herangezogen. Dieses Profil ist auf Ebene der Parameter ähnlich zum elektrodermalen Profil des NASA-RTLX aus Studie II, $C = .89$. Die letzte Spalte in Tabelle 7.11 zeigt auf, dass nur für den Parameter EDR.N ein bedeutsamer Unterschied in der Ladungsstruktur vorliegt. Dieser lädt hoch auf dem Profil des NASA-RTLX aus Studie II und gering auf dem Profil der Valenz (siehe Tabelle 7.12). Für die Parameter EDR.rec.t/2, EDR.rec.t/c und EDR.peaks liegen zudem marginale Faktorladungsunterschiede vor.

7.4 Diskussion

In diesem Kapitel werden die Ergebnisse zur Manipulationskontrolle (Abschnitt 7.4.1), zu den kardialen Profilen (Abschnitt 7.4.2) und zu den elektrodermalen Profilen (Abschnitt 7.4.3) zusammengefasst und interpretiert. Abschließend werden die Limitationen der Studie II sowie deren Implikationen in Abschnitt 7.4.4 beleuchtet.

7.4.1 Zusammenfassung und Interpretation der Manipulationskontrolle

In Bezug auf die Fragestellung KF2 ist eine gelungene Induktion mentaler Beanspruchung festzustellen. Die Ergebnisse belegen sowohl für die NASA-RTLX-Werte als auch für die Trefferquote eine erwartungskonforme und bedeutsame Veränderung durch die unterschiedlichen Schwierigkeitsstufen der Aufgabe. Die Hypothesen H2.1 und H2.4 können somit bestätigt werden. Bestätigt wurden ebenfalls die Annahmen der Hypothesen H2.2 und H2.3, sodass keine bedeutsamen Haupteffekte der visuellen Komplexität und des Bildmaterials auf die subjektive mentale Beanspruchung vorliegen. Zudem gibt es keinen bedeutsamen Interaktionseffekt auf die subjektive mentale Beanspruchung (Beantwortung der UF2.1). Insgesamt wirkt sich lediglich die Aufgabenschwierigkeit bedeutsam auf die subjektive mentale Beanspruchung aus.

Für die leistungsbezogene mentale Beanspruchung bestätigen sich zudem die Hypothesen H2.5 (kein Haupteffekt der visuellen Komplexität) und H2.6 (kein Haupteffekt des Bildmaterials). Im Gegensatz zur subjektiven mentalen Beanspruchung liegt allerdings eine bedeutsame Dreifachinteraktion aus Aufgabenschwierigkeit, visueller Komplexität und Bildmaterial vor (Beantwortung der UF2.2). Der Interaktionseffekt lässt darauf schließen, dass die schweren Aufgaben bei hoher visueller Komplexität die kognitive Leistungskapazität stärker auslasten, wenn farbige Zeichnungen (im Vergleich zu schwarz-weiß

Zeichnungen und farbigen Fotos) präsentiert werden. Wird die visuelle Komplexität bei den schweren Aufgaben verringert, dreht sich der Effekt um. In diesem Fall lasten die Farbinformationen die kognitive Leistungskapazität weniger aus als die schwarz-weißen Zeichnungen oder die farbigen Fotos. Allerdings erklärt der Interaktionseffekt im Vergleich zum Haupteffekt der Aufgabenschwierigkeit (58.8 % Varianzaufklärung) nur einen sehr geringen Varianzanteil von 2.6 % und spielt somit nur eine untergeordnete Rolle in der Varianz der leistungsbezogenen mentalen Beanspruchung.

7.4.2 Zusammenfassung und Interpretation der kardialen Profile

Die im Rahmen der Forschungsfrage F1.1 und F1.2 extrahierten kardialen Profile sind in Abbildung 7.7 aufgetragen: Insgesamt ergeben sich ein kardiales Profil für das NASA-RTLX (rot) und ein kardiales Profil für die Trefferquote (olivfarben). Die statistischen Kenngrößen sind in Tabelle 7.13 zusammengefasst. Beide kardialen Profile bilden die zwei Schwierigkeitsstufen der Aufgaben erwartungskonform ab (vgl. Abbildung 7.7, oben) und können als Repräsentanten der subjektiven bzw. leistungsbezogenen mentalen Beanspruchung interpretiert werden. Zudem sind beide kardialen Profile identisch und werden deshalb zu einem kardialen Profil der mentalen Beanspruchung zusammengefasst. Daraus ist abzuleiten, dass unter Verwendung des Primäraufgabenmaßes das kardiale Pro-

Tabelle 7.13: Übersicht der kardialen Profile und deren Vorhersagegüte aus Studie II

Kriterium	B	R_{PLS}^2	$valR_{PLS}^2$	Interpretation
NASA-RTLX	0.71	82.37	55.81	} Mentale Beanspruchung
Trefferquote	-0.63	63.60	40.36	

Anmerkungen. Die Bestimmtheitsmaße R_{PLS}^2 und $valR_{PLS}^2$ beziehen sich auf den Anteil erklärter Varianz im Kriterium.

fil der Trefferquote mit dem des NASA-RTLX übereinstimmt. Ob dieser Befund spezifisch für die vorliegende Studie ist oder generell für alle Primäraufgaben gilt, kann an dieser Stelle nicht beantwortet werden. Bedeutsame Ladungen auf dem kardialen Profil der mentalen Beanspruchung haben die Parameter MQSD, pNN50, CVSD, EKG.HF und SVDEn mit $|a| > .30$. Diese Parameter nehmen konform zu den bisherigen Forschungsbefunden (siehe Abschnitt 3.2.4) bei Zunahme der mentalen Beanspruchung ab. Eine endgültige Bewertung der kardialen Einzelparameter sowie deren finale Zuordnung zu einem psychologischen Kriterium erfolgt am Ende der Arbeit in Unterkapitel 9.1.2.

In der Übersichtstabelle 7.13 ist zu sehen, dass trotz der Identität weniger Varianz in der Trefferquote (63.6 %) als im NASA-RTLX (82.4 %) durch die kardialen Komponenten erklärt wird. Eine mögliche Ursache könnte der Interaktionseffekt Schwierigkeit \times Komplexität \times Bildmaterial sein. Dieser hat einen bedeutsamen Einfluss auf die leistungsbezogene mentale Beanspruchung, jedoch nicht auf die subjektive mentale Beanspruchung (siehe Abschnitt 6.4.1). Die kardiale Komponente der Trefferquote bildet diesen Interaktionseffekt allerdings nicht nach. Da die kardiale Komponente der Trefferquote diesen Varianzanteil in der leistungsbezogenen mentalen Beanspruchung nicht erklären kann, ist auch dessen Anteil erklärter Varianz geringer. Die Tabelle 7.13 zeigt auf, dass die Vorhersagegüte für Personen, die dem Modell nicht bekannt sind, stark abnimmt (siehe $\text{val}R_{\text{pl.s}}^2$). Auch wenn der verbleibende Anteil erklärter Varianz immer noch bedeutsam und beachtlich ist, gehen über 20.0 % verloren. Dies verdeutlicht die Relevanz der personenspezifischen Unterschiede für die Vorhersagegüte der mentalen Beanspruchung durch das kardiale Profil.

In Bezug auf Forschungsfrage F6.1 wurde die Kongruenz zwischen den extrahierten Profilen mentaler Beanspruchung der Studie I und II bestätigt. Dabei ist das in Studie II identifizierte kardiale Profil der mentalen Beanspruchung identisch zu demjenigen kardialen Profil, welches in Studie I als ein Repräsentant der mentalen Beanspruchung interpretiert wurde (Profil des NASA-RTLX). Zudem zeigt das kardiale Profil der mentalen Beanspruchung aus Studie II nur eine Ähnlichkeit zu demjenigen kardialen Profil, welches in Studie I als

ein Mischprofil mentaler und emotionaler Beanspruchung interpretiert wurde (Profil der Trefferquote). Hierbei unterscheiden sich die Faktorladungen nur für den Parameter HP marginal voneinander. Faktorladungsunterschiede des Parameters HP wurden bereits in Studie I als Differenzierungsmerkmal zwischen den mentalen und emotionalen Beanspruchungsrepräsentanten berichtet (siehe Abschnitt 6.3.10).

Insgesamt bestätigen die Übereinstimmungen zwischen den beiden Studien die Strukturvalidität des extrahierten kardialen Profils der mentalen Beanspruchung. Zudem wird die Vermutung, dass der aufgabenspezifische Grad an Bewegung sich im kardialen Profil der mentalen Beanspruchung der Studie I niedergeschlagen hat, durch die vorliegenden Befunde widerlegt. Im Gegensatz zur Studie I wurde in der Aufgabe der Studie II nur ein minimaler und über alle Aufgaben hinweg konstanter Grad an Bewegung gefordert. Trotzdem stimmen die extrahierten Profile mentaler Beanspruchung zwischen den Studien überein und bestätigen damit die bisherige Interpretation.

7.4.3 Zusammenfassung und Interpretation der elektrodermalen Profile

Aus den elektrodermalen Daten wurden zwei Profile, eines für das NASA-RTLX und eines für die Trefferquote (siehe Abbildung 7.8, unten), extrahiert. Beide Profile können die mentale Beanspruchung in den leichten und schweren Aufgaben zuverlässig vorhersagen (siehe Abbildung 7.8, oben). Des Weiteren sind beide elektrodermalen Profile identisch und können zu einem gemeinsamen Profil zusammengefasst werden. Die vorliegenden Ergebnisse bestätigen den bereits für die kardialen Profile beschriebenen Befund, dass unter Verwendung des Primäraufgabenmaßes das Profil der Trefferquote identisch mit dem Profil des NASA-RTLX ist. Auch hier kann keine Aussage darüber getroffen werden, ob dieser Befund spezifisch für die vorliegende Studie ist oder generell für alle Primäraufgabenmaße gilt. Insgesamt ergibt sich in Bezug

Tabelle 7.14: Übersicht der elektrodermalen Profile und deren Vorhersagegüte aus Studie II

Kriterium	B	R^2_{PLS}	$\text{val}R^2_{\text{PLS}}$	Interpretation
NASA-RTLX	0.92	82.52	51.20	} Mentale Beanspruchung
Trefferquote	-0.81	63.36	37.15	

Anmerkungen. Die Bestimmtheitsmaße R^2_{PLS} und $\text{val}R^2_{\text{PLS}}$ beziehen sich auf den Anteil erklärter Varianz im Kriterium.

auf die Forschungsfragen F2.1 und F2.2 ein gemeinsames elektrodermales Profil als Repräsentant der mentalen Beanspruchung. Die Tabelle 7.14 fasst die statistischen Kenngrößen der Profile zusammen. Hier ist eine deutliche Reduktion des kreuzvalidierten Bestimmtheitsmaßes um ca. 30 % im Vergleich zum Bestimmtheitsmaß des Originalmodells erkennbar. Dies verdeutlicht die Relevanz der personenspezifischen Unterschiede für die Vorhersagegüte der mentalen Beanspruchung.

Trotz der Identität beider elektrodermalen Profile aus Studie II ist die Vorhersagegüte für das elektrodermale Profil des NASA-RTLX besser als für das elektrodermale Profil der Trefferquote (vgl. R^2_{PLS} in Tabelle 7.14). Wie bereits für das kardiale Profil ist auch hier die Ursache im Interaktionseffekt Schwierigkeit \times Komplexität \times Bildmaterial zu vermuten, welcher durch das extrahierte Profil der Trefferquote nicht abgebildet wird. Bedeutsame Faktorladungen auf dem elektrodermalen Profil der mentalen Beanspruchung haben die Parameter EDR.N, EDR.amp, EDA.LF, EDA.HF und Phasic.SD mit $|a| > .30$. Sie alle nehmen mit Zunahme der mentalen Beanspruchung zu. Dieses Zusammenhangsmuster ist konform zu den bisherigen Forschungsbefunden (siehe Unterkapitel 3.3.4). Eine endgültige Bewertung der elektrodermalen Parameter sowie deren finale Zuordnung zu einem psychologischen Kriterium erfolgt am Ende der Arbeit in Unterkapitel 9.1.4.

In Bezug auf die Forschungsfrage F6.2 konnte keine Kongruenz zu demjenigen elektrodermalen Profil gefunden werden, welches in Studie I als Repräsentant

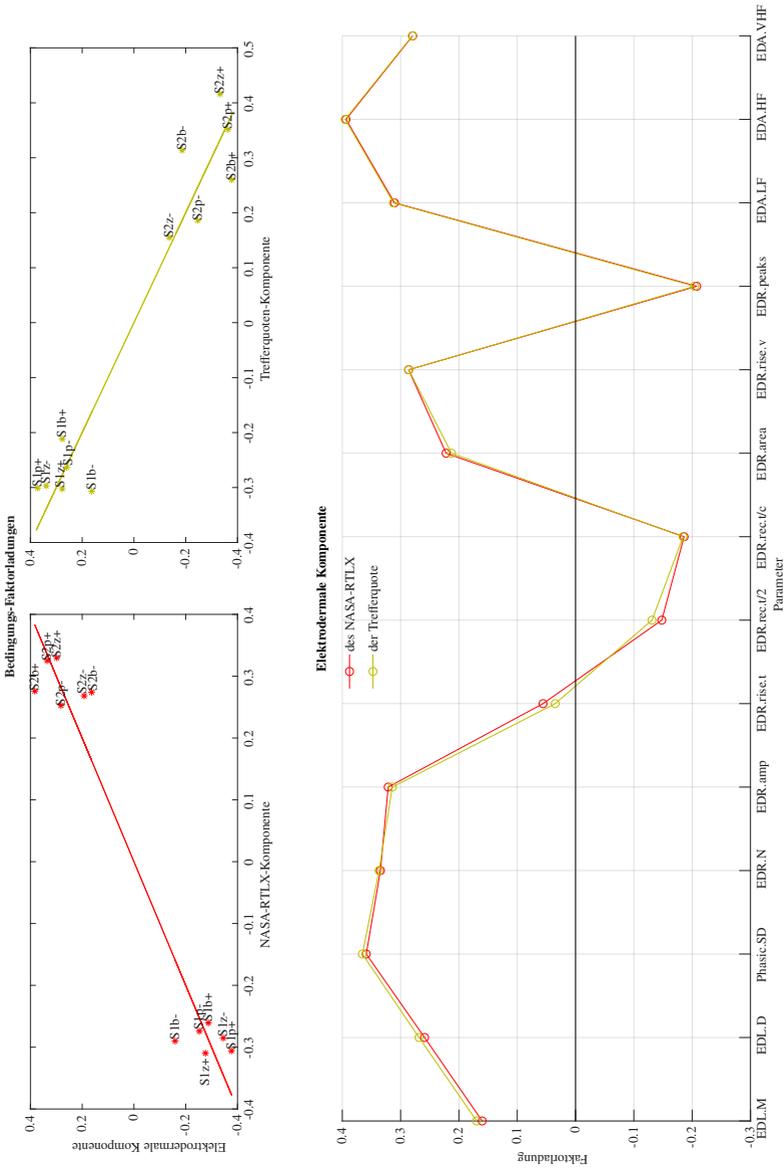


Abbildung 7.8: Oben: Zusammenhang zwischen den Bedingungs-Faktorladungen der elektrodermalen Komponente und der Kriterien-Komponente. Unten: Parameter-Faktorladungen der elektrodermalen Komponente aus Studie II.

der mentalen Beanspruchung interpretiert wurde (Profil des NASA-RTLX). Die Profile mentaler Beanspruchung aus Studie I und II zeigen unterschiedliche Faktorladungsmuster und weisen somit keine Strukturvalidität auf. Demgegenüber ergab sich eine Ähnlichkeit zu demjenigen elektrodermalen Profil, welches in Studie I als ein Repräsentant der emotionalen Beanspruchung interpretiert wurde (Profil der Valenz). Dieses Ergebnis wirft die Frage auf, ob das in Studie II extrahierte elektrodermale Profil zumindest teilweise die emotionale Beanspruchung mit abbildet. So könnte bspw. eine Konfundierung zwischen mentaler und emotionaler Beanspruchung in Studie II vorliegen. Immerhin erfolgte im Gegensatz zu Studie I keine explizite Trennung in mentale und emotionale Belastungsfaktoren. Dieser Ansatz erklärt allerdings nicht, warum das elektrodermale Profil aus Studie II keine Ähnlichkeit mit dem elektrodermalen Profil der mentalen Beanspruchung aus Studie I hat. Eine Erklärung dafür ist die Konfundierung der mentalen Beanspruchung in Studie I. So könnte der in Studie I vorliegende große Beanspruchungsunterschied zwischen der Balltracking- und der IAPS-Aufgabe das elektrodermale Profil des NASA-RTLX wesentlich mit geprägt haben. Durch die fehlende Strukturvalidität kann die vermutete Konfundierung des elektrodermalen Profils des NASA-RTLX durch den aufgabenspezifischen Grad an Bewegung in der Studie I nicht widerlegt werden. Im Gegensatz zu Studie I wurde in Studie II der Bewegungsgrad zwischen den Experimentalbedingungen minimiert und konstant gehalten, um den Einfluss dieser Störvariable auf die elektrodermalen Profile zu beseitigen. Eine weitere Erklärung für die fehlende Ähnlichkeit ist die Verwendung unterschiedlicher Messsensoren (siehe Abschnitt 7.4.4) und Aufgabentypen zwischen den Studien. Zudem ist zu beachten, dass in der Studie II nur eine einzige Aufgabe operationalisiert wurde. Folglich könnte das extrahierte elektrodermale Profil der Studie II diese Aufgabenspezifität widerspiegeln. Diese Fragestellung wird in der Studie III adressiert.

Insgesamt kann für die elektrodermalen Profile der Studien I und II keine Strukturvalidität festgestellt werden. Zudem kann mit den vorliegenden Befunden nicht hinreichend geklärt werden, ob das in Studie I oder das in Studie II

extrahierte elektrodermale Profil die mentale Beanspruchung zuverlässig repräsentiert. Diese offene Fragestellung ist in einer weiteren Studie zu klären.

7.4.4 Limitationen

Die Einschränkungen in der Ergebnisinterpretation der Studie II werden im folgenden diskutiert.

Einschränkungen der Messdaten

Für die Studie II ergeben sich Einschränkungen durch die ungeklärte Reliabilität kardialer Ultra-Kurzzeitmaße sowie durch die fehlende Kontrolle der Atemaktivität. Beide Limitationen wurden bereits im Rahmen der Studie I diskutiert und können im Unterkapitel 6.4.4 nachgelesen werden.

Der potentielle Einfluss der Raumtemperaturänderung auf elektrodermale Profile wurde ebenfalls bereits im Unterkapitel 6.4.4 der Studie I beschrieben. Eine geringe, jedoch bedeutsame Steigung liegt für die Studie II vor. Diese stieg um durchschnittlich 0.25 °C. Ein bedeutsamer Effekt findet sich ebenfalls für die relative Luftfeuchtigkeit. Diese stieg durchschnittlich um 1.1 %.

Eine Einschränkung bei der Bewertung der Kongruenz zwischen den physiologischen Profilen resultiert aus der Verwendung unterschiedlicher Messsensoren bei der Erhebung kardialer als auch elektrodermalen Aktivität. Die mangelnde Kongruenz zwischen den elektrodermalen Profile der Studien I und II könnte darauf zurückzuführen sein. In Studie I wurde die Shimmer3 GSR+ Unit (Abtastrate 51.2 Hz) und in Studie II der varioport-B Biosignal Rekorder (Abtastrate 500 Hz) verwendet (siehe Abschnitt 5.2.1). Daneben gibt es zwei bedeutende Unterschiede. Erstens wurde in Studie II der Hautleitwert und in Studie I der Hautwiderstand gemessen. Allerdings ist der Hautwiderstand unter bestimmten Bedingungen nicht linear in den Hautleitwert überführbar (Boucsein, 2012, S. 248). Zweitens wurden in Studie I Trockenelektroden verwendet, hingegen

in Studie II Elektroden mit einer isotonischen Paste. Beide Faktoren können zu quantitativen und qualitativen Unterschieden in den elektrodermalen Parametern und Profilen geführt haben.

Auch die kardiale Aktivität wurde mit unterschiedlichen Sensoren erfasst: in Studie I mit der Schimmer3 ECG Unit und in Studie II mit dem varioport-B Biosignal Rekorder (siehe Abschnitt 5.1.1). Diese unterscheiden sich nicht nur in der Art der Ableitungen am Brustkorb, sondern auch in den Abtastraten (204.8 Hz in Studie I und 500 Hz in Studie II). Trotz dieser Unterschiede sind die kardialen Profile des NASA-RTLX aus Studie I und II identisch. Allerdings wurden in Abschnitt 7.3.6 Unterschiede in den bedeutsamen Parameter-Faktorladungen festgestellt. So gab es für die identischen kardialen Profile des NASA-RTLX aus Studie I und II nur zwei Parameter (MQSD und EKG.HF) die in beiden Studien bedeutsam auf dem Profil laden. Eine Ursache hierfür könnte in der Verwendung unterschiedlicher Sensoren liegen.

Einschränkungen in der Konstruktrennung

Im Rahmen der Studie I wurde eine Konfundierung mentaler und emotionaler Beanspruchung bereits im Unterkapitel 6.4.4 diskutiert. In Studie I äußerte sich die Konfundierung durch die Extraktion von physiologischen Mischprofilen mentaler und emotionaler Beanspruchung. Eine Konfundierung könnte ebenfalls in Studie II für das elektrodermale Profil vorliegen. Immerhin zeigt dieses Profil keine Kongruenz mit dem elektrodermalen Profil der mentalen Beanspruchung aus Studie I, jedoch eine Ähnlichkeit zum elektrodermalen Profil der emotionalen Beanspruchung aus Studie I. Da eine psychometrische Trennung emotionaler und mentaler Beanspruchung kein Ziel der Studie II war, wurde im Versuchsdesign nur eine reine mentale Belastungsinduktion umgesetzt. Im Gegensatz dazu konnten die mentalen und emotionalen Belastungsfaktoren durch das Versuchsdesign der Studie I deutlich voneinander abgegrenzt werden. Dies wirft die Frage auf, ob ohne eine explizite Trennung im Versuchsdesign es zwangsläufig zur Extraktion von Mischprofilen kommt.

Eine weitere Limitation ist die Präsentation der Bildersets im Versuch. Die Variation der visuellen Komplexität und des Bildmaterials erzeugt potentielle Störfaktoren im Experiment. Ein Einfluss dieser Störfaktoren besteht nachweislich für die leistungsbezogene mentale Beanspruchung, jedoch nicht für die subjektive mentale Beanspruchung. Diese Störeinflüsse wurden als potentielle Ursache für die geringere Varianzaufklärung durch das Profil der Trefferquote im Vergleich zum Profil des NASA-RTLX bereits diskutiert. Es ist aber auch möglich, dass die visuelle Komplexität und das Bildmaterials die physiologische Aktivität und somit die extrahierten Profile verzerren. In Anbetracht der fehlenden Kongruenz der elektrodermalen Profile mentaler Beanspruchung aus Studie I und II ist ein solcher Einfluss nicht auszuschließen.

Einschränkungen in der externen Validität

Die Ergebnisse der Studie II sind auf den Laborkontext und den spezifischen Aufgabentyp der Studie beschränkt. Zudem handelt es sich um eine anfallende Stichprobe, die nicht repräsentativ für die Gesamtpopulation ist. Beide Limitationen wurden bereits im Rahmen der Studie I diskutiert (siehe Unterkapitel 6.4.4) und sind auf die vorliegende Studie II übertragbar.

8 Studie III

Dieses Kapitel beschreibt das Forschungsziel der Studie III in Abschnitt 8.1 sowie die studienspezifischen Methoden in Abschnitt 8.2. Anschließend werden die Ergebnisse der Kontrollfragestellung KF3 sowie der Forschungsfragen F1.1, F2.1, F7.1 und F7.2 in Abschnitt 8.3 vorgestellt. Am Ende werden in Abschnitt 8.4 die Ergebnisse interpretiert und deren Limitationen beleuchtet.

8.1 Forschungsziel der Studie III

Wie bereits in den ersten beiden Studien verfolgt auch die Studie III das Ziel, kardiale und elektrodermale Profile mentaler Beanspruchung abzuleiten. Der spezifische Fokus der Studie III besteht darin, die physiologischen Profile mentaler Beanspruchung über unterschiedliche Aufgaben hinweg zu extrahieren. Die Ergebnisse der Studie III sollen somit Aufschluss darüber geben, ob die physiologischen Profile mentaler Beanspruchung über verschiedene Aufgabentypen hinweg Gültigkeit besitzen. Insgesamt werden in Studie III fünf verschiedene Aufgabentypen verwendet, die in ihren kognitiven Belastungsfaktoren variieren. Alle Aufgabentypen adressieren dabei einen anderen Bereich der Kognition. Diese Bereiche sind: (i) die visuell-motorische Koordination (Augen-Hand-Koordination), (ii) der Abruf von Informationen aus dem Arbeitsgedächtnis, (iii) die selektive Aufmerksamkeit und Konzentration, (iv) das schlussfolgernde Denken und (v) die kognitive Inhibition irrelevanter Reize. Auf eine ausführliche Darstellung dieser kognitiven Bereiche wird verzichtet, da diese lediglich der Operationalisierung verschiedener kognitiver Belastungs-

faktoren dienen. Die daraus abgeleiteten physiologischen Profile dienen zum einen der Prüfung der Generalisierbarkeit über verschiedene Aufgaben hinweg. Zum anderen werden die aufgabenübergreifenden physiologischen Profile aus Studie III zur Prüfung der Strukturvalidität verwendet. Hierzu werden sie mit den aufgabenspezifischen extrahierten Profilen der mentalen Beanspruchung aus Studie I und II verglichen (siehe Forschungsfrage F7 Kapitel 4.3).

Die Extraktion physiologischer Profile wird in Studie III nur für die subjektive, jedoch nicht für die leistungsbezogenen, mentale Beanspruchung durchgeführt. Dies hängt damit zusammen, dass unterschiedliche Leistungsmaße für die Erfassung der fünf kognitiven Bereiche eingesetzt wurden.¹ Dadurch unterscheidet sich die gemessene Leistung qualitativ und quantitativ zwischen den Aufgabentypen. Ein Beleg hierfür wird im Anhang A.8.4 anhand der Analyse der Leistungsdaten aus Studie III erbracht. Die daraus extrahierten physiologischen Profile wären nicht mehr valide auf die mentale Beanspruchung zurückzuführen. Infolgedessen werden in Studie III nur die Teilfragestellung F.1.1 und F2.1 (siehe Kapitel 4.3) adressiert.

8.2 Methode

Dieser Abschnitt beschreibt die studienspezifischen Methoden der Studie III. Dies umfasst die Operationalisierung der mentalen Beanspruchung (Abschnitt 8.2.1), das Versuchsdesign (Abschnitt 8.2.2), die Versuchsdurchführung (Abschnitt 8.2.3), die Kontrollvariablen (Abschnitt 8.2.4), die Manipulationskontrolle (Abschnitt 8.2.5) und die Beschreibung der Stichprobe (Abschnitt 8.2.6).

¹ Eine einheitliche Leistungsmessung wäre in diesem Fall durch die Applikation einer aufgabenübergreifenden Sekundäraufgabe denkbar. Dies würde allerdings im Widerspruch zum studienspezifischen Ziel stehen, nämlich der Extraktion physiologischer Profile über verschiedene kognitive Belastungsbereiche hinweg.

8.2.1 Operationalisierung der mentalen Beanspruchung

In diesem Abschnitt werden die Methoden beschrieben, mit denen die mentale Beanspruchung im Experiment ausgelöst und gemessen wurde.

Experimentelle Induktion mentaler Beanspruchung

Die Induktion der mentalen Beanspruchung erfolgte über informatorische Belastungsfaktoren in Form von fünf unterschiedlichen Aufgabentypen am Computer. Jede der Aufgaben bildet eines der in Kapitel 8.1 aufgeführten Bereiche der Kognition ab. Die fünf Aufgabentypen sind:

- (i) die Balltrackingaufgabe mit Belastungsfaktoren im Bereich der visuell-motorischen Koordination,
- (ii) die Gedächtnisaufgabe mit Belastungsfaktoren im Bereich des Arbeitsgedächtnisses,
- (iii) die Konzentrationsaufgabe mit Belastungsfaktoren im Bereich der Aufmerksamkeit,
- (iv) die Matrizenaufgabe mit Belastungsfaktoren im Bereich des schlussfolgernden Denkens und
- (v) die Stroopaufgabe mit Belastungsfaktoren im Bereich der kognitiven Inhibition.

Innerhalb jedes Aufgabentyps wurde die Höhe der mentalen Belastung durch die Veränderung der Aufgabenschwierigkeit variiert. Die niedrigste Schwierigkeitsstufe (1) wurde stets als Ruhebedingung konzipiert. Dabei wurden den Versuchspersonen die jeweiligen Aufgabeninhalte (z. B. Zahlensequenz bei der Gedächtnisaufgabe) auf dem Bildschirm präsentiert, die sie nur betrachten sollten. Eine Lösung der Aufgaben war bei der Ruhebedingung weder notwendig noch möglich. Im Gegensatz dazu wurde bei den restlichen drei Schwierigkeits-

stufen (2, 3 und 4) eine Antwortreaktion von den Versuchspersonen gefordert. Nach jeder Antwortreaktion erfolgte eine Leistungsrückmeldung zur Richtigkeit der Lösung. Jede Experimentalbedingung dauerte 60 s. Um sicher zu gehen, dass die Versuchspersonen die Aufgabenstellung und -bedingung verstanden hatten, wurde am Anfang jedes neuen Aufgabentyps ein Übungsdurchgang durchgeführt. Bei Bedarf konnte dieser Wiederholt werden. Die Schwierigkeit der Aufgaben wurde durch aufgabenspezifische Belastungsfaktoren manipuliert. Diese werden in den folgenden Abschnitten vorgestellt.

Balltrackingaufgabe

Trackingaufgaben werden zur Messung der visuell-motorischen Koordination eingesetzt (z. B. Gauthier, Vercher, Ivaldi & Marchetti, 2004). Eine Variante der Trackingaufgabe wurde bereits in Studie I vorgestellt. Diese diente als Orientierung für die Balltrackingaufgabe (B) der Studie III. Das Ziel der Aufgabe war es, einen blauen Zielball mit dem Mauszeiger auf dem Computerbildschirm zu verfolgen und so häufig wie möglich zu treffen. Der Zielball bewegte sich dabei zufällig und in konstanter Geschwindigkeit über den Bildschirm. Neben dem

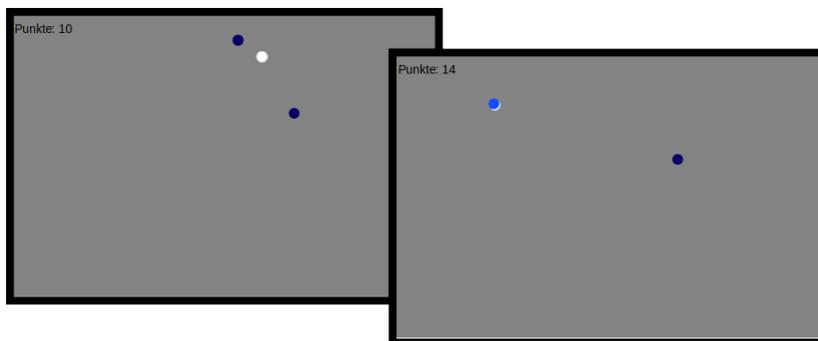


Abbildung 8.1: Dargestellt sind Ausschnitte aus einer Balltrackingaufgabe mit einem Distraktorball. Der Mauszeiger ist als weißer Kreis dargestellt. Links: Ohne Zielkontakt haben der Ziel- und Distraktorball dieselbe blaue Farbe. Rechts: Bei Zielkontakt leuchtet der Zielball hellblau auf.

Zielball bewegte sich ein gleichfarbiger Distraktorball mit gleicher Geschwindigkeit zufällig über den Bildschirm. Die Personen wurden dazu instruiert, den Distraktorball bei der Aufgabenausführung zu ignorieren. Um den Zielball vom Distraktorball differenzieren zu können, leuchtete der Zielball bei jedem Treffer mit dem Mauszeiger hellblau auf (siehe Abbildung 8.1). Dies ermöglichte den Versuchspersonen, den Zielball bei Berührung zu identifizieren. Bei jedem Aufgabenstart leuchtete der Zielball für 3 s hellblau auf, um diesen als solchen zu markieren. Während der gesamten Aufgabe wurde ein Leistungsfeedback in Form eines Punktestands in der oberen linken Bildschirmcke präsentiert. Der Punktestand wurde als die relative Anzahl der Treffer (Quotient aus Anzahl tatsächlicher Treffer und Anzahl möglicher Gesamttreffer) gerechnet. Jedes Mal, wenn der Mauszeiger den Zielball berührte, erhöhten sich der Punktestand. Der Punktestand bewegte sich im Bereich zwischen 0 (kein Ballkontakt) bis 100 (ununterbrochener Ballkontakt).

Die Schwierigkeitsstufen der Balltrackingaufgaben wurden durch eine aufsteigende Ballgeschwindigkeit sowie durch die Anzahl der Distraktorbälle operationalisiert. In der Ruhebedingung (B1) sollte der Bildschirm mit dem sich bewegenden Ziel- und Distraktorball betrachtet werden. Folglich wurden in der Bedingung B1 weder der Mauszeiger noch der Punktestand angezeigt. Beide Bälle bewegten sich mit einer niedrigen Geschwindigkeit über den Bildschirm. Bei der leichten Schwierigkeitsstufe (B2) musste der Zielball mit dem Mauszeiger verfolgt werden. Neben dem Zielball bewegte sich nur ein Distraktorball mit derselben niedrigen Geschwindigkeit über den Bildschirm. Für die mittlere Schwierigkeitsstufe (B3) wurde die Ballgeschwindigkeit um das fünffache erhöht, während die Gesamtanzahl an Distraktorbällen konstant bei Eins blieb. In der höchsten Schwierigkeitsstufe (B4) wurde die Ballgeschwindigkeit auf dem Niveau von B3 belassen und die Anzahl der Distraktorbälle erhöht. Hier bewegten sich 20 Distraktorbälle gleichzeitig über den Bildschirm. Vor dem Beginn der Experimentalbedingungen wurde ein Übungsdurchgang mit mittlerer Ballgeschwindigkeit und einem Distraktorball für 30 s bearbeitet.

Gedächtnisaufgabe

Die Gedächtnisaufgabe diente zur Operationalisierung des Informationsabrufs aus dem Arbeitsgedächtnis. Die in Studie III verwendete Gedächtnisaufgabe (G) beruht auf dem Test der Gedächtnisspanne, welcher häufig in Intelligenztest (z. B. Wechsler Adult Intelligence Scale; Petermann & Wechsler, 2012) Verwendung findet. Dieser Test misst die Speicherkapazität des menschlichen Arbeitsgedächtnisses, indem eine Sequenz von gesehenen oder gehörten numerischen Ziffern korrekt aus dem Gedächtnis abgerufen und wiedergegeben wird. Die durchschnittliche Gedächtnisspanne für Erwachsene beträgt 7 ± 2 Elemente bei einer vorwärts gerichteten Wiedergabe (Miller, 1956). Der Test existiert in verschiedenen Versionen, bspw. in einer vorwärts oder rückwärts gerichteten Wiedergabe der Sequenz.

In der Gedächtnisaufgabe der Studie III wurden den Versuchspersonen mehrere numerische Ziffernsequenzen (z. B. „3 4 2“) visuell präsentiert. Diese mussten im Anschluss wiedergeben werden. Für jede Ziffernsequenz erfolgte die Präsentation der Ziffern nacheinander auf dem Bildschirm. Dabei handelte es sich immer um ganze Zahlen im Bereich von 0 bis 9. Innerhalb einer Ziffernsequenz kam eine Zahl nur ein Mal vor. Jede Ziffer wurde für 1 s angezeigt. Am Ende der Ziffernsequenz erschien eine Eingabemaske, in welche die Versuchsperson ihre Antwort eingeben und mit einem Klick auf das Feld „ENTER“ bestätigen

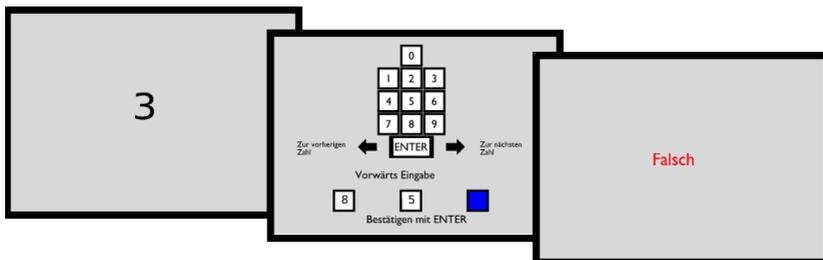


Abbildung 8.2: Dargestellt sind Ausschnitte aus einer Gedächtnisaufgabe der Schwierigkeitsstufe G2: Präsentation einer Ziffer (links), Eingabemaske für die Antwort (Mitte) und Leistungsrückmeldung (rechts).

konnte.² Abbildung 8.2 zeigt den zeitlichen Ablauf der Gedächtnisaufgabe im Experiment. Die Lösung der Aufgabe musste in einer vorgegebenen Zeit eingegeben werden, bevor der Test automatisch zur nächsten Ziffernsequenz überging. Pro Ziffer war eine maximale Bearbeitungszeit von 2.5 s vorgesehen. Nach jeder Antwort erschien für 0.5 s die Leistungsrückmeldung „Richtig“ bei korrekter und „Falsch“ bei inkorrekt reaktion in der Bildschirmmitte (siehe Abbildung 8.2).

Die Schwierigkeit der Gedächtnisaufgabe wurde zum einen durch die Länge der Ziffernsequenz und zum anderen durch die Eingaberichtung (Vorwärts- oder Rückwärtseingabe) manipuliert. Da die erste Schwierigkeitsstufe (G1) als Ruhebedingung konzipiert war, wurde hier eine zufällige Ziffernsequenz für 60 s präsentiert. Eine Eingabemaske für die Antwort war nicht vorhanden. Bei der zweiten Schwierigkeitsstufe (G2) bestand die Ziffernsequenz aus insgesamt drei Ziffern. Die Ziffern mussten dabei innerhalb von 7.5 s in der präsentierten Reihenfolge, also vorwärts, eingegeben werden. Für die dritte Schwierigkeitsstufe (G3) wurde eine fünfstellige Ziffernsequenz verwendet. Die Eingabe der Ziffernsequenz in das Antwortfeld musste hierbei jedoch in der umgekehrten Präsentationsreihenfolge (rückwärts) erfolgen. Für die Eingabe hatten die Versuchspersonen maximal 12.5 s. Die längste Ziffernsequenz, bestehend aus sieben Ziffern, wurde in der höchsten Schwierigkeitsstufe (G4) mit einer Rückwärtseingabe und einer maximalen Eingabezeit von 17.5 s umgesetzt. Im vorgelagerten Übungsdurchlauf wurden insgesamt vier Ziffernsequenzen präsentiert: zwei dreistellige und zwei fünfstellige Ziffernsequenzen mit jeweils einer vorwärts und einer rückwärts gerichteten Eingabe.

² Musste eine eingegebene Ziffer korrigiert werden, konnte das bereits befüllte Eingabefeld überschrieben werden.

Konzentrationsaufgabe

Die Konzentrationsaufgabe der Studie III ist an dem sogenannten d2-Test (Brickenkamp, 2002) zur Messung der selektiven Aufmerksamkeits- und Konzentrationsleistung (Verarbeitungsgeschwindigkeit) angelehnt. In der Originalfassung umfasst der d2-Test eine mehrzeilige Buchstabensequenz. Die Buchstaben sind jeweils mit ein bis zwei Strichen ober- oder unterhalb des Buchstabens markiert (z. B. d'' oder d'). Das Ziel bei diesem Test ist es, nur eine bestimmte Zielkombination zu markieren.

Bei der Konzentrationsaufgabe (K) der Studie III wurden den Versuchspersonen nacheinander einzelne Buchstaben einer Buchstabensequenz auf dem Bildschirm präsentiert (siehe Abbildung 8.3). Die Sequenz setzte sich aus den Buchstaben d , b , q und p zusammen. Zudem waren die Buchstaben oben und unten mit keinem bis maximal zwei Strichen markiert. Das Ziel der Aufgabe war es, auf den Buchstaben p mit zwei Strichen (p'' , p' und p') durch Betätigung der Eingabetaste zu reagieren. Auf ein p mit mehr oder weniger als zwei Strichen sowie bei allen anderen Buchstaben, unabhängig davon wie viele Striche diese besaßen, sollte die Versuchspersonen nicht reagieren. Der Anteil der Zielobjekte in einer Buchstabensequenz betrug 44.5 % (angelehnt an das

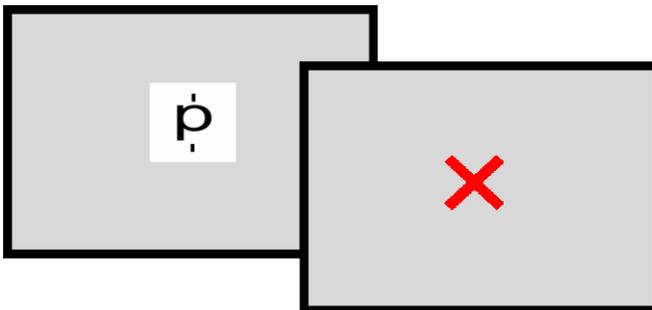


Abbildung 8.3: Dargestellt sind Ausschnitte aus der Konzentrationsaufgabe: Präsentation eines Zielobjekts (links) und Leistungsrückmeldung bei inkorrektter Reaktion (rechts).

Verhältnis im d2-Test). Die Reihenfolge der Stimulusobjekte innerhalb einer Buchstabensequenz war zufällig gewählt.

Die niedrigste Schwierigkeitsstufe (K1) enthielt als einzige Bedingung keine Zielobjekte, um eine unbewusste Reaktion der Versuchspersonen in der Ruhebedingung zu vermeiden. Die Reaktion auf das Zielobjekt musste in der Zeit erfolgen, in der das Stimulusobjekt auf dem Bildschirm präsentiert wurde. Die Anzeigzeit variierte dabei abhängig von der Schwierigkeitsstufe. In der ersten und zweiten Schwierigkeitsstufe (K1 und K2) wurden die Stimulusobjekte für 1.6 s, in der dritten Schwierigkeitsstufe für 0.9 s und in der höchsten Schwierigkeitsstufe (K4) für 0.5 s angezeigt. Da jeder Durchgang 60 s dauerte, wurden in der Bedingung K1 und K2 bis zu 35, in der Bedingung K3 bis zu 61 und in der Bedingung K4 bis zu 102 Stimulusobjekte präsentiert. Erfolgte eine inkorrekte Antwortreaktion auf einen Aufgabenreiz (z. B. Reaktion auf ein Nicht-Zielobjekt oder Nicht-Reaktion auf ein Zielobjekt), dann wurde für 0.15 s ein rotes Kreuz in der Bildschirmmitte eingeblendet (siehe Abbildung 8.3). Bei einer korrekten Antwortreaktion wurde zum nächsten Stimulusobjekt übergegangen, ohne explizit auf die Richtigkeit der Antwortreaktion hinzuweisen. Im vorgelagerten Übungsdurchlauf bearbeiteten die Versuchspersonen für 30 s eine Buchstabensequenz aus der Schwierigkeitsstufe K2.

Matrizenaufgabe

Matrizen werden zur Messung des schlussfolgernden Denkens eingesetzt. Das Ziel dabei ist es, logische Zusammenhänge zwischen den abstrakten Objekten in einer Matrix zu entdecken. Für die Matrizenaufgabe (M) der Studie III wurden mehrere Matrizen konzipiert, die je nach Schwierigkeitsstufe (M1 bis M4) unterschiedlich komplex waren. Die Komplexität der Matrizen ergab sich aus der Anzahl der Objekte innerhalb eines Matrizenelements und aus der Anzahl der Regeln, denen die Matrizenelemente folgten. Einfache Matrizen beinhalteten nur wenige Objekte und Regeln (siehe Abbildung 8.4, oben),

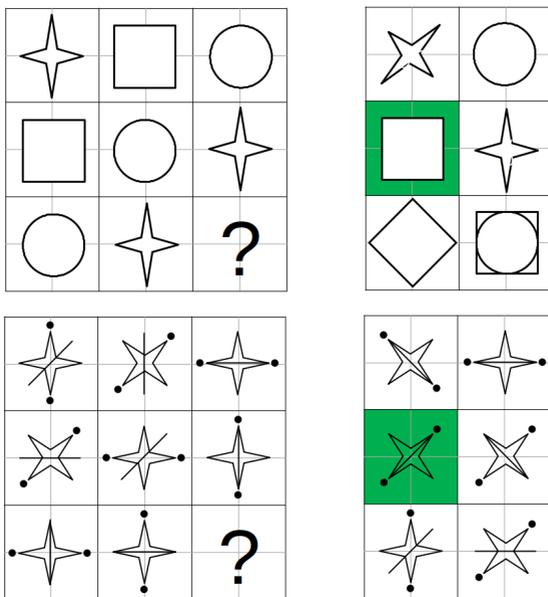


Abbildung 8.4: Leichte Matrize (oben): ein Objekt pro Matrizelement und eine einzige Regel (jedes Objekt ist ein Mal in einer Zeile oder Spalte vertreten). Schwere Matrize (unten): mehrere Objekten pro Matrizelement (Stern mit Punkten und Gerade) und separaten Objektregeln (Stern mit Punkten drehen sich im Uhrzeigersinn und die Gerade dreht sich gegen den Uhrzeigersinn). Die Lösung ist im rechten Auswahlbereich in Grün markiert.

wohingegen komplexe Matrizen mehrere Objekte und Regeln enthielten (siehe Abbildung 8.4, unten).

Das Ziel in der Matrizenaufgabe war die Identifikation des fehlenden Matrizelements aus einer Reihe möglicher Antwortelemente. Abbildung 8.5 zeigt den Ablauf der Matrizenaufgabe im Experiment. Alle Matrizen bestanden aus einer $3 \cdot 3$ Matrix. Von den insgesamt 9 Elementen fehlte ein Element. Das fehlende Element war stets unten rechts in der Matrix zu finde und wurde mit einem „?“ symbolisiert. Dieses musste aus einer $2 \cdot 3$ Matrix ermittelt werden. Eine Antwort konnte durch das Klicken auf das ausgewählte Antwortelement und der anschließenden Bestätigung mit der Eingabetaste abgegeben werden.

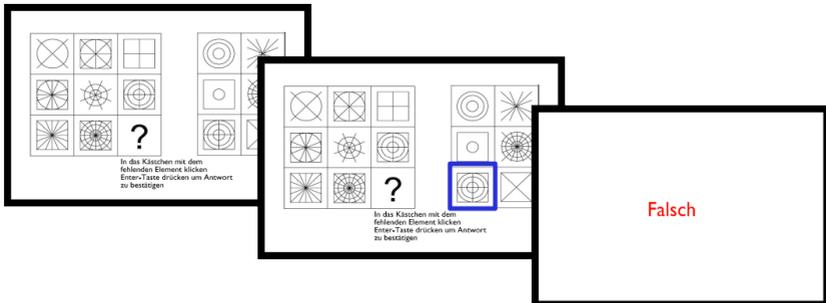


Abbildung 8.5: Dargestellt sind Ausschnitte aus einer schweren Matrizenaufgabe: Präsentation der Matrize (links), Auswahl eines Antworelements (Mitte) und Leistungsrückmeldung bei inkorrekt Lösungswahl (rechts).

Ein ausgewähltes Antworelement wurde mit einem blauen Rahmen markiert. Nach der anschließenden Bestätigung durch die Eingabetaste erschien für 0.5 s die Leistungsrückmeldung „Richtig“ bei korrekter und „Falsch“ bei falscher Antwort auf dem Bildschirm. Jede Matrize wurde für 7 s eingeblendet. Die Zeit, die zur Lösung einer Matrix zur Verfügung stand, betrug 20 s in den ersten drei Schwierigkeitsstufen (M1 bis M3) und 10 s in der vierten Schwierigkeitsstufe (M4). Erfolgte innerhalb der Antwortzeit von 10 s bzw. 20 s keine Eingabe, ging der Test automatisch zur nächsten Matrize über. Jeder Durchgang dauerte 60 s.

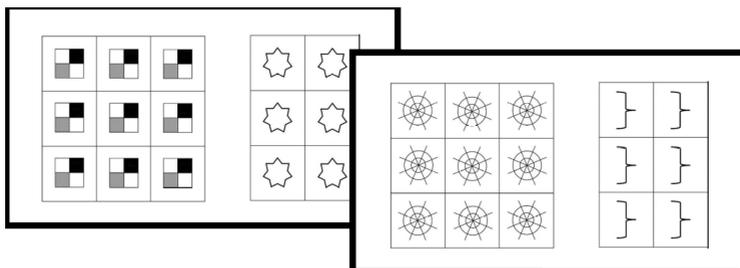


Abbildung 8.6: Beispiel für Matrizen aus der Ruhebedingung (M1).

Die erste Schwierigkeitsstufe (M1) wurde als Ruhebedingung konzipiert. Hier mussten keine Matrizen gelöst werden mussten. Stattdessen wurden nacheinander verschiedene Matrizen mit identischen Elementen präsentiert (siehe Abbildung 8.6). Die zweiten Schwierigkeitsstufe (M2) enthielt einfachen Matrizen mit wenigen Objekten und Regeln. In der dritten und vierten Schwierigkeitsstufe (M3 und M4) kamen schwere Matrizen (mehrere Objekte und Regeln) zum Einsatz. Zudem wurde in der Bedingung M4 die Antwortzeit auf 10 s verkürzt. In einem vorgelagerten Übungsdurchgang wurden den Versuchspersonen drei Übungsmatrizen in aufsteigender Schwierigkeit (M2 bis M4) vorgelegt.

Stroopaufgabe

Das Ziel einer Stroopaufgabe ist es, automatisierte Handlungsimpulse aktiv zu kontrollieren. Folglich werden Strooppaufgaben zur Messung der kognitiven Inhibition – also der Unterdrückung automatischer Reaktionsmechanismen – eingesetzt (z. B. Noreen & MacLeod, 2015). Aber auch zur experimentellen Induktion der mentalen Beanspruchung ist diese geeignet (z. B. Janssen et al., 2019; Vanitha, Suresh, Chandrasekar & Punita, 2017). In einer klassischen Stroopaufgabe werden Personen dazu aufgefordert, die Schriftfarbe eines präsentierten Farbwortes zu benennen, ohne sich vom Wortinhalt ablenken zu lassen (Stroop, 1935). Sind die Schriftfarbe und das Farbwort identisch (siehe Abbildung 8.7, links) ist dies mühelos möglich. Wenn das Farbwort und die Schriftfarbe jedoch inkongruent sind (siehe Abbildung 8.7, rechts), kommt es zur sogenannten Stroop-Interferenz. Diese ist anhand einer erhöhten Reaktionszeit und einer steigenden Antwortfehlerrate erkennbar (Dyer & Severance, 1973; MacLeod, 1991).³

³ Dieser Effekt kommt dadurch zustanden, dass die eher ungewohnte Reaktion, nämlich das Erkennen der Schriftfarbe, sich gegen die trainierte und dadurch dominante Reaktion des Lesens durchsetzen muss. Präsentiert man den Personen eine Reihe farbiger „X“ anstatt der Farbwörter, verschwindet die Stroop-Interferenz.

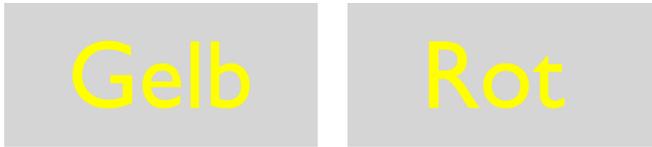


Abbildung 8.7: Links: Übereinstimmung zwischen Farbwort und Schriftfarbe (Kongruenz). Rechts: Keine Übereinstimmung zwischen Farbwort und Schriftfarbe (Inkongruenz).

Im Gegensatz zur klassischen Stroopaufgabe musste in der Stroopaufgabe (S) der Studie III die Schriftfarbe nicht benannt, sondern durch einen Mausklick in das entsprechende Antwortfeld eingegeben werden. In Abbildung 8.8 ist zu sehen, dass das Antwortfeld aus Antwortelementen mit den entsprechenden Farbwörtern in schwarzer Schriftfarbe bestand. Je nach Schwierigkeitsstufe wurden vier oder sechs Farbwörter bzw. Schriftfarben verwendet: Gelb, Rot, Blau und Grün für die ersten drei Schwierigkeitsstufen (S1 bis S3) und zusätzlich Grau und Lila für die höchste Schwierigkeitsstufe (S4). Innerhalb einer Experimentalbedingung wurde die Reihenfolge der Farbwörter randomisiert. In den ersten drei Schwierigkeitsstufen (S1 bis S3) wurde jedes Farbwort für 2 s eingeblendet. Die Präsentationsdauer für die Schwierigkeitsstufe S4 betrug 1.4 s. Innerhalb dieser Zeit musste die Antwort erfolgen, bevor der Test automatisch

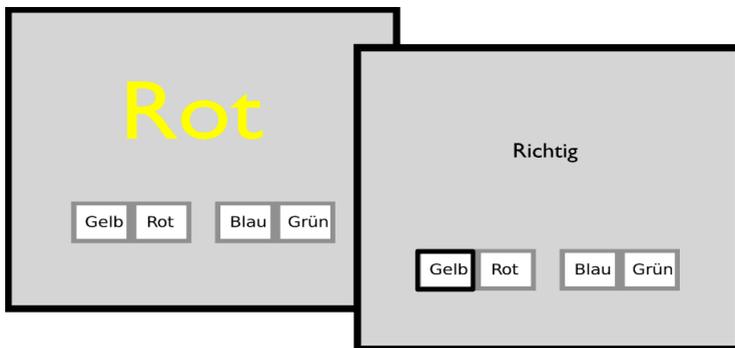


Abbildung 8.8: Dargestellt sind Ausschnitte aus der Stroopaufgabe mit der Schwierigkeitsstufe S3: Präsentation des inkongruenten Farbworts mit Antwortfeld (links) und Leistungsrückmeldung bei korrekter Reaktion (rechts).

zum nächsten Farbwort übergang. Nach jeder Antwort erschien für 0.5 s die Leistungsrückmeldung „Richtig“ bei korrekter und „Falsch“ bei inkorrektter Antwortreaktion.

Die Schwierigkeitsstufe S1 stellte die Ruhebedingung der Stroopaufgabe dar. Entsprechend wurde in dieser Bedingung kein Antwortfeld und keine Leistungsrückmeldung angezeigt. Alle in der Bedingung S1 präsentierten Farbwörter waren kongruent zur Schriftfarbe. In der einfachen Schwierigkeitsstufe (S2) wurden ebenfalls kongruentes Stimulusmaterial für jeweils 2 s dargeboten. Jedoch wurde hier eine Antwortreaktion verlangt. In S3 wurde die Schwierigkeitsstufe durch die Präsentation von inkongruenten Farbwörtern erhöht. Auch hier wurde jedes Farbwort für 2 s eingeblendet. Die schwerste Schwierigkeitsstufe (S4) wurde ebenfalls mit inkongruentem Stimulusmaterial durchgeführt. Zusätzlich wurde die Gesamtanzahl an Farbwörtern von vier auf sechs angehoben und die Präsentationsdauer auf 1.4 s pro Farbwort verkürzt. Vor Beginn der Stroopaufgabe konnten alle Versuchspersonen einen Übungsdurchgang absolvieren. Hier wurden jeweils vier kongruente und vier inkongruente Farbwörter mit einer Präsentationsdauer von 2 s pro Farbwort bearbeitet.

Messung der mentalen Beanspruchung

In Studie III wurde die subjektive mentale Beanspruchung mit dem NASA-RTLX (ohne die Skala körperliche Anforderung) im Anschluss an jede Bedingung erfasst (siehe Unterkapitel 5.4.1). Auch wurde eine Leistungsmessung für jede Bedingung durchgeführt. Da die Leistungsmessung nicht im Fokus der Analyse der Studie III steht, ist diese im Anhang A.8.4 beschrieben.

8.2.2 Versuchsdesign

Dem Experiment der Studie III liegt ein quasiexperimentelles Design mit Messwiederholung zugrunde. Es umfasst 20 Bedingungen, bestehend aus fünf

Aufgabentypen (B, G, K, M, S) mit jeweils vier Schwierigkeitsstufen (1 bis 4). Die zu einem Aufgabentyp gehörenden Experimentalbedingungen wurden hintereinander durchgeführt, wobei die Reihenfolge der Schwierigkeitsstufen interindividuell randomisiert wurde. Ebenfalls randomisiert wurde die Reihenfolge der Aufgabentypen pro Versuchsperson. Folglich bearbeitete jede Versuchsperson die gleichen Bedingungen, jedoch in einer unterschiedlichen Reihenfolge.

Im Gegensatz zu Studie I und II wird in Studie III nur eine Kriteriumsvariable – nämlich die subjektive mentale Beanspruchung gemessen mit dem NASA-RTLX – zur Extraktion physiologischer Profile mentaler Beanspruchung verwendet. Die NASA-RTLX-Werte bilden die Kriteriumsmatrix im Rahmen der NPLS-Regression. Die Prädiktormatrix wird jeweils durch die kardialen und elektrodermalen Parameter repräsentiert.

8.2.3 Versuchsablauf

Der Versuch fand im Labor des Instituts für Arbeitswissenschaft und Betriebsorganisation (ifab) des KIT statt. Vor Beginn informierte die Versuchsleitung alle Versuchspersonen mündlich und schriftlich über den Zweck der Studie sowie über die Erfassung der Daten. Danach wurde eine Einverständniserklärung zur Teilnahme am Experiment unterzeichnet. Im Anschluss brachte die Versuchsleitung die Elektroden zur Messung der kardialen und elektrodermalen Aktivität an der Versuchsperson an und verband diese mit der Messsensorik. Vor dem Start des Experiments wurden alle Sensoren auf ihre Funktionstüchtigkeit sowie Signalplausibilität von der Versuchsleitung geprüft. Daran anschließend instruierte die Versuchsleitung die Versuchsperson über den Versuchsablauf, die Experimentalaufgaben sowie den NASA-RTLX-Fragebogen. Alle Versuchspersonen wurden gebeten, sich im Experiment möglichst wenig zu bewegen. Zur Steigerung der Leistungsmotivation wurde den Versuchspersonen mitgeteilt, dass die zu bearbeitenden Aufgaben unter anderem bei der Intelligenzdiagnostik eingesetzt werden. Nach Klärung offener Fragen, wurde der Versuch gestartet.

Dieser fand automatisiert am Versuchsrechner statt, um Versuchsleitereffekte zu vermeiden. Als Experimentalsoftware wurde die Blender Game Engine Version 2.78 (Blender Online Community, 2015) verwendet. Die Instruktion zu jedem Aufgabentyp und jedem Schwierigkeitsgrad wurde der Versuchsperson auf dem Bildschirm präsentiert. Am Anfang jedes neuen Aufgabentyps durchlief die Versuchsperson einen Übungsdurchgang, um sich an die Experimentalaufgabe zu gewöhnen und mögliche Fragen zu klären. Anschließend folgten die zum Aufgabentyp gehörenden vier Experimentalbedingungen in randomisierter Reihenfolge. Jede Experimentalbedingung hatte eine Dauer von 60 s. Nach jeder Experimentalbedingung beantwortete die Versuchsperson den NASA-RTLX-Fragebogen. Die fünf Aufgabentypen wurden in randomisierter Reihenfolge bearbeitet. Zwischen den Aufgabentypen konnten die Versuchspersonen nach Bedarf eine Pause machen. Nach Beendigung des Versuchs beantwortete die Versuchsperson einen Abschlussfragebogen (Kontrollvariablen) und legte alle Sensoren ab. Anschließend wurde der Versuchsperson für die Teilnahme gedankt und sie wurde von der Versuchsleitung verabschiedet. Der gesamte Versuch dauerte ca. 90 min.

8.2.4 Abschlussbefragung und Kontrollvariablen

Zur Erfassung der soziodemografischen Daten wurde am Ende des Experiments ein Abschlussfragebogen an die Versuchspersonen ausgehändigt. Die Versuchspersonen wurden gebeten, ihr Alter, Geschlecht und ihre beruflichen Hintergründe anzugeben. Als Kontrollvariablen wurden die Temperatur und die relative Luftfeuchtigkeit zu Beginn und am Ende des Experiments durch die Versuchsleitung gemessen.

8.2.5 Manipulationskontrolle

Zur validen Beantwortung der Forschungsfragen muss vorab statistisch geprüft werden, ob die experimentelle Manipulation der mentalen Beanspruchung über

die verschiedenen Schwierigkeitsstufen der Aufgaben gelungen ist. Zu diesem Zweck werden Hypothesen (H) für die Kontrollfragestellung KF3 im folgenden Abschnitt formuliert.

Hypothesen der Manipulationskontrolle

Für die subjektive mentale Beanspruchung (NASA-RTLX-Werte) wird eine Steigung mit zunehmender Aufgabenschwierigkeit erwartet.

H3.1 Je höher die Aufgabenschwierigkeit, desto höher sind die NASA-RTLX-Werte.

Zudem wird untersucht, ob der Aufgabentyp einen Effekt auf die NASA-RTLX-Werte hat. Idealerweise sollten keine Unterschiede zwischen den verschiedenen Aufgabentypen vorliegen, sodass jeder Aufgabentyp im Mittel eine gleich hohe subjektive mentale Beanspruchung hervorruft. Diese Annahme wird mit der folgenden Hypothese getestet.

H3.2 Es gibt keine signifikanten Unterschiede zwischen den Aufgabentypen in den NASA-RTLX-Werten.

Ob die Steigung der mentalen Beanspruchung (bei steigenden Schwierigkeitsstufen) sich zwischen den Aufgabentypen unterscheidet, wird mit Hilfe der Interaktion Aufgabenschwierigkeit \times Aufgabentyp geprüft. Da keine Annahmen zur Art der Interaktionen gemacht werden können, wird eine Unterfragestellung (UF) zur Kontrollfragestellung KF3 formuliert:

UF3.1 Gibt es einen signifikanten Interaktionseffekt zwischen der Aufgabenschwierigkeit und dem Aufgabentyp auf die NASA-RTLX-Werte?

Auswertemethoden der Manipulationskontrolle

Im Rahmen der Manipulationskontrolle wird mit Hilfe der MEA geprüft, ob die Schwierigkeitsstufen einen bedeutsamen Effekt auf die NASA-RTLX-Werte haben. Aufgrund der nominalen Faktorstufen der Schwierigkeit kommt eine Effektkodierung zum Einsatz. Um einen guten Überblick mit möglichst wenigen Dummyvariablen zu erreichen, wird die Rückwärts-Differenz-Kodierung angewandt. Dabei wird der Kriteriumsmittelwert der ersten Faktorstufe mit dem Kriteriumsmittelwert der zweiten Faktorstufe verglichen. Anschließend wird der Kriteriumsmittelwert der zweiten Faktorstufe mit dem der dritten Faktorstufe verglichen usw. Insgesamt ergeben sich für die vier Faktorstufen der Aufgabenschwierigkeit drei Dummyvariablen. Zusätzlich wird auf Unterschiede zwischen den fünf Faktorstufen des Aufgabentyps (B, G, K, M und S) geprüft. Hierfür wird die einfache Effektkodierung angewandt. Dabei wird der Kriteriumsmittelwert der jeweiligen Faktorstufe (z. B. G) mit dem Gesamtmittelwert des Kriteriums⁴ (Mittelwert aus B, G, K, M und S) verglichen. Insgesamt ergeben sich hier fünf Dummyvariablen, jeweils eine für jeden Aufgabentyp.

Zur Prüfung der Interaktion Aufgabenschwierigkeit \times Aufgabentyp werden aus den Dummyvariablen der Schwierigkeit und des Aufgabentyps Interaktionsvariablen gebildet. Insgesamt ergeben sich $3 \cdot 5$ Zweifachinteraktionen. Aufgrund der Vielzahl an Dummyvariablen wird eine Bonferroni-Korrektur zur Vermeidung der Alphafehlerkumulierung durchgeführt. Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei $\alpha = .002$.

8.2.6 Stichprobe

Am Experiment nahmen 56 Versuchspersonen freiwillig teil. Diese Stichprobe bestand zu 87 % aus Studenten und zu 11 % aus Berufstätigen. Eine Per-

⁴ Technisch gesehen ist es nicht der Mittelwert des Kriteriums, sondern der Mittelwert der Kriteriumsmittelwerte der einzelnen Faktorstufen.

son machte keine Angaben. 68 % waren männlichen und 32 % weiblichen Geschlechts. Der Altersdurchschnitt lag bei 24.1 ($SD = 3.6$) Jahren.

Bei den physiologischen Daten wurden Personen aufgrund von Sensorausfällen aus der weiteren Analyse ausgeschlossen. Die Analyse der kardialen Aktivität wurde mit 54 Personen durchgeführt. Zwei Personen mussten aufgrund eines Sensorausfalls aus der Analyse ausgeschlossen werden. Bei der Analyse der elektrodermalen Aktivität wurden drei Personen aufgrund von Sensorausfällen ausgeschlossen.

8.3 Ergebnisse

Die folgenden Unterkapitel berichten die Ergebnisse der Forschungsfragen. Zunächst wird in Abschnitt 8.3.1 geprüft, ob die Induktion der mentalen Beanspruchung im Experiment gelungen ist. Anschließend folgt die Extraktion der kardialen Profile mentaler Beanspruchung in Abschnitt 8.3.2 und der elektrodermalen Profile mentaler Beanspruchung in Abschnitt 8.3.3. Zu diesem Zweck werden NPLS-Regressionen der zweidimensionalen Kriteriumsmatrix (Bedingung \times Person) auf die dreidimensionale Prädiktormatrix (Bedingung \times Person \times Parameter) berechnet. Zum Schluss wird im Sinne der Strukturvalidität geprüft, ob die extrahierten kardialen Profile (Abschnitt 8.3.4) und elektrodermalen Profile (Abschnitt 8.3.5) mentaler Beanspruchung identisch zu den Profilen mentaler Beanspruchung aus Studie I und II sind. Die deskriptive Statistik kardialer und elektrodermalen Parameter ist aus Platzgründen im Anhang A.8.1 und A.8.2 beschrieben.

Tabelle 8.1: Deskriptive Statistiken der Kontrollvariablen

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
Anfangstemperatur	56	24.95	1.31	22	28	0.62	0.90
Endtemperatur	54	25.43	1.16	23	28	0.85	0.44
Luftfeuchtigkeit (Beginn)	56	51.30	11.04	23	67	-0.72	-0.13
Luftfeuchtigkeit (Ende)	54	52.00	8.80	30	67	-0.63	0.02

8.3.1 Kontrollfragestellung KF3: Ist die experimentelle Induktion mentaler Beanspruchung in Studie III gelungen?

In diesem Abschnitt werden die Kontrollvariablen und das NASA-RTLX analysiert.

Analyse der Kontrollvariablen

Die deskriptiven Statistiken der Kontrollvariablen sind in Tabelle 8.1 dargestellt. Für die Raumtemperatur sowie für die relative Luftfeuchtigkeit fehlen jeweils zwei Messpunkte für das Ende des Experiments. Ein t-Test für abhängige Stichproben⁵ deckt auf, dass es einen signifikanten Anstieg in der Raumtemperatur ($t(53) = 4.99, p < .001, d = 0.67$), jedoch nicht in der relativen Luftfeuchtigkeit ($t(48) = 0.82, p = .417$) zum Ende des Experiments gibt. Die Temperatur steigt im Mittel um 0.40 °C an.

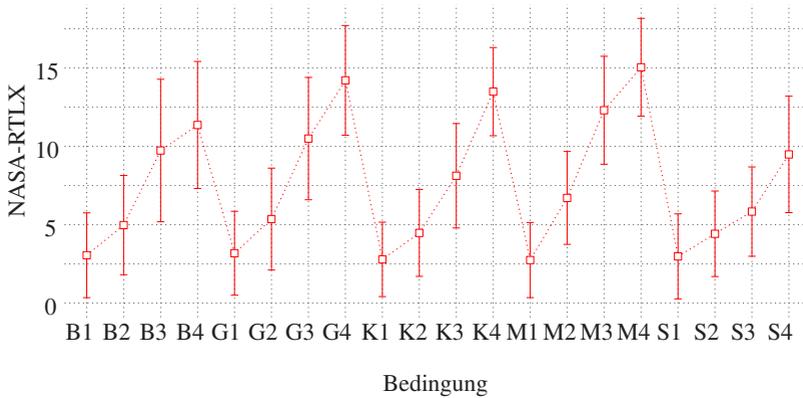


Abbildung 8.9: Mittelwert und Standardabweichung des NASA-RTLX in den Versuchsbedingungen.

Analyse des NASA-RTLX-RTLX

Die NASA-RTLX-Werte für die 20 Experimentalbedingungen sind in Abbildung 8.9 dargestellt. Eine Tabelle mit den deskriptiven Statistiken der NASA-RTLX-Werte findet sich im Anhang A-25. Ausreißer ($\pm 3 SD$) oder fehlende Datenpunkte liegen nicht vor. Deskriptiv ist bereits ein erwartungskonformer Trend im Sinne einer Zunahme der subjektiven mentalen Beanspruchung mit zunehmender Schwierigkeitsstufe innerhalb der fünf Aufgabentypen zu erkennen. Dies wird mit Hilfe der MEA inferenzstatistisch geprüft. Diese deckt die Hypothesen H3.1, H3.2 und die Unterfragestellung UF3.1 der Manipulationskontrolle ab (siehe Abschnitt 8.2.5).

Die Ergebnisse des Modellvergleichs in Tabelle 8.2 zeigen, dass sowohl die beiden Haupteffekte (Aufgabenschwierigkeit und Aufgabentyp) als auch die Interaktion (Aufgabenschwierigkeit \times Aufgabentyp) eine signifikante Reduktion der Devianz und des AIC bewirken. Hierbei klärt der Haupteffekt der

⁵ Signifikanzniveau von 5 % (zweiseitig).

Tabelle 8.2: Modellfit-Indizes für die geschachtelten Modelle des NASA-RTLX

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	6794.97	6788.97			0.00
+ Schwierigkeit (M_S)	6	5741.09	5729.09	1059.87***	3	53.15
+ Aufgabe (M_A)	10	5552.60	5532.60	196.50***	4	58.40
+ Schwierigkeit \times Aufgabe ($M_{S \times A}$)	22	5407.31	5363.31	169.29***	12	62.20

*** $p < .001$, ** $p < .01$, * $p < .05$.

Aufgabenschwierigkeit 53.2 % der Varianz auf, während der Haupteffekt des Aufgabentyps nur 5.3 % und der Interaktionseffekt nur 3.8 % der Varianz aufklären.

Die Regressionsgewichte der einzelnen Faktorstufen sind in Tabelle 8.3 für das Interaktionsmodell $M_{S \times A}$ dargestellt. Hier ist eine signifikante Steigung der subjektiven mentalen Beanspruchung für jede höhere Schwierigkeitsstufe vorzufinden (Teststärken der Effekte bei 100.0 %). Die Regressionsgewichte des Aufgabentyps geben Auskunft darüber, welcher Aufgabentyp vom Durchschnitt aller Aufgabentypen abweicht. Hier zeigt sich, dass die NASA-RTLX-Werte in der Stroopaufgabe signifikant geringer ausfallen. Signifikant höhere NASA-RTLX-Werte haben hingegen die Gedächtnis- und Matrizenaufgaben. Die Teststärken für die bedeutsamen Effekte des Aufgabentyps liegen zwischen 99.8 % und 100.0 %. Keine bedeutsamen Abweichungen vom Durchschnitt der Aufgabentypen zeigen die Balltracking- und Konzentrationsaufgaben.

Die 15 Interaktionsvariablen in Tabelle 8.3 zeigen auf, bei welchem Aufgabentyp die Steigung zwischen zwei Schwierigkeitsstufen geringer oder höher ausfällt als die durchschnittliche Steigung über alle Aufgabentypen hinweg. Die Interaktionseffekte sind deskriptiv bereits in Abbildung 8.9 anhand der unterschiedlichen Steigungen zwischen den Schwierigkeitsstufen eines Auf-

Tabelle 8.3: Parameter für das Modell ($M_{S \times A}$) des NASA-RTLX

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	7.54***	0.27	27.55	56	1.97
Schwierigkeit: 2 – 1	2.24***	0.21	10.66	1064	
Schwierigkeit: 3 – 2	4.11***	0.21	19.57	1064	
Schwierigkeit: 4 – 3	3.42***	0.21	16.28	1064	
Aufgabe: B	-0.26	0.15	-1.73	1064	
Aufgabe: G	0.77***	0.15	5.20	1064	
Aufgabe: K	-0.32	0.15	-2.15	1064	
Aufgabe: M	1.66***	0.15	11.18	1064	
Aufgabe: S	-1.86***	0.15	-12.51	1064	
(Schw.: 2 – 1)×(Auf.: B)	-0.31	0.42	-0.75	1064	
(Schw.: 3 – 2)×(Auf.: B)	0.65	0.42	1.54	1064	
(Schw.: 4 – 3)×(Auf.: B)	-1.79***	0.42	-4.26	1064	
(Schw.: 2 – 1)×(Auf.: G)	-0.06	0.42	-0.14	1064	
(Schw.: 3 – 2)×(Auf.: G)	1.03	0.42	2.44	1064	
(Schw.: 4 – 3)×(Auf.: G)	0.30	0.42	0.70	1064	
(Schw.: 2 – 1)×(Auf.: K)	-0.55	0.42	-1.31	1064	
(Schw.: 3 – 2)×(Auf.: K)	-0.46	0.42	-1.10	1064	
(Schw.: 4 – 3)×(Auf.: K)	1.95***	0.42	4.64	1064	
(Schw.: 2 – 1)×(Auf.: M)	1.73***	0.42	4.11	1064	
(Schw.: 3 – 2)×(Auf.: M)	1.48**	0.42	3.53	1064	
(Schw.: 4 – 3)×(Auf.: M)	-0.68	0.42	-1.62	1064	
(Schw.: 2 – 1)×(Auf.: S)	-0.80	0.42	-1.90	1064	
(Schw.: 3 – 2)×(Auf.: S)	-2.69***	0.42	-6.41	1064	
(Schw.: 4 – 3)×(Auf.: S)	0.23	0.42	0.54	1064	

*** $p < .0001$, ** $p < .001$, * $p < .002$.

gabentyps erkennbar. Ein signifikant geringere Steigung besteht zwischen der Schwierigkeitsstufe B3 und B4 der Balltrackingaufgabe sowie zwischen der Schwierigkeitsstufen S2 und S3 der Stroopaufgabe. Im Gegensatz dazu fällt bei der Konzentrationsaufgabe die Steigung zwischen K3 und K4 signifikant höher aus als die durchschnittliche Steigung über alle Aufgabentypen hinweg. Das Gleiche trifft auf die Steigungen zwischen M1, M2 und M3 der Matrizenaufgabe zu. Die Teststärken für die bedeutsamen Interaktionseffekte liegen zwischen 93.7 % und 100.0 %. Alle übrigen Experimentalbedingungen zeigen keine bedeutsamen Abweichungen zur durchschnittlichen Steigung über alle Aufgabentypen hinweg.

8.3.2 Forschungsfrage F1.1: Können aus den kardialen Einzelparametern kardiale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

Zunächst wird die Anzahl der Komponenten bestimmt sowie die Komponentenmodelle Kreuzvalidiert. Gemäß den Ergebnissen der Kreuzvalidierung in Tabelle 8.4 bietet nur die erste Komponente einen bedeutsamen Beitrag zur Vorhersagegüte der NASA-RTLX-Werte.⁶ Im Originalmodell erklärt diese Komponente 8.9 % der Varianz in den kardialen Parametern und sagt 48.6 % der Varianz in den NASA-RTLX-Werten vorher. In der Kreuzvalidierten Vorhersage verringert sich der Anteil erklärter Varianz im NASA-RTLX auf 29.5 %. Die Faktorladungen der Komponente 1 sind in Abbildung 8.10 dargestellt.

Für die Interpretation der Faktorladungen wird zunächst die Bedingungsdimension analysiert. Diese ist in Abbildung 8.10 im mittleren Diagramm abgebildet. Die Tabelle der Bedingungs-Faktorladungen ist im Anhang A-29 zu finden.

⁶ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Prädiktormatrix beträgt: $M[99\% \text{ CI}] = 0.23[0.09, 0.44]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{pls}}^2$ der Kriteriumsmatrix beträgt: $M[99\% \text{ CI}] = -2.77[-7.76, -0.08]$.

Tabelle 8.4: NPLS-Regression des NASA-RTLX auf kardiale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR^2_{PLS}) und das kreuzvalidierte Modell ($\text{val}\Delta R^2_{\text{PLS}}$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	ΔR^2_{PLS}	$\text{val}\Delta R^2_{\text{PLS}}$	valSQR
0					1580.33
1	8.90	7.43 [†]	48.63	29.47 [†]	723.64***
2	4.15	1.83 [†]	17.36	2.27 [†]	700.30
3	4.09	1.54 [†]	9.37	1.93 [†]	680.52
4	5.10	3.14 [†]	6.87	2.95 [†]	650.26
5	3.66	1.20 [†]	3.26	-0.44	654.82

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $\text{val}\Delta R^2 > 0.44$, der Kriteriumsmatrix: [†] $\text{val}\Delta R^2 > 0.0$.

Die aus der NASA-RTLX-Matrix extrahierte Komponente 1 (gestrichelte rote Linie in Abbildung 8.10, Mitte) stimmt in den Bedingungs-Faktorladungen mit den Mittelwerten des NASA-RTLX (gestrichelte dunkelgraue Linie) überein. Somit kann die NASA-RTLX-Komponente 1 als ein Repräsentant der subjektiven mentalen Beanspruchung interpretiert werden. Die NASA-RTLX-Komponente 1 kann mit $B = 0.50$, $t(1078) = 11.49$, $p < .001$ durch die kardiale Komponente 1 (durchgezogene rote Linie) in der Bedingungsdimension vorhergesagt werden. Für beide Komponenten liegen die niedrigsten Schwierigkeitsstufen im unterdurchschnittlichen und die höchsten Schwierigkeitsstufen im überdurchschnittlichen Bereich der subjektiven mentalen Beanspruchung bzw. der kardialen Aktivität. Es liegen aber auch Unterschiede zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente 1 und der kardialen Komponente 1 vor. Bedeutsame Faktorladungsunterschiede ($|a_{\Delta}| \geq .20$) finden sich in den Bedingungen B2, K4, M4 und S3. Zudem zeigt sich ein marginaler Unterschied ($|a_{\Delta}| \geq .15$) in der Bedingung B3. Diese Unterschiede deuten darauf hin, dass der in Abschnitt 8.3.1 beschriebenen Haupteffekt des Aufgabentyps sowie der Interaktionseffekt (Aufgabenschwierig-

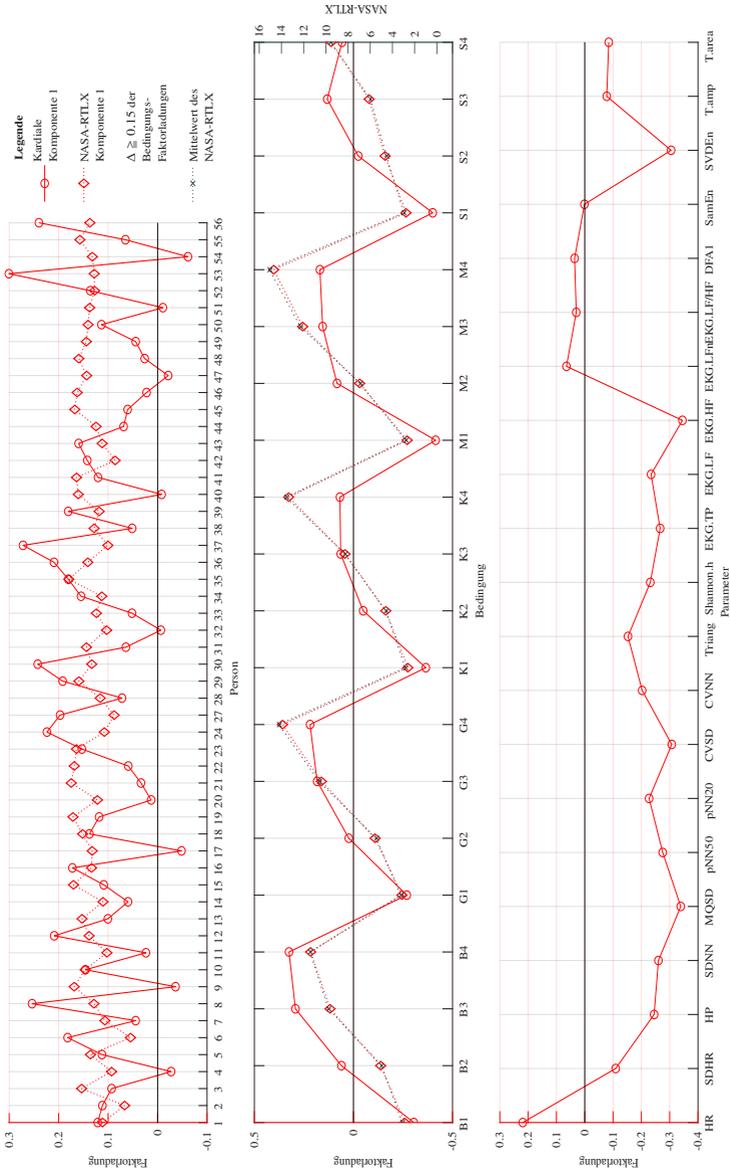


Abbildung 8.10: Faktorladungen α für das NASA-RTLX und die kardiale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.

keit \times Aufgabentyp) auf die subjektive mentale Beanspruchung nicht durch die kardiale Komponente 1 abgebildet wird. In Abbildung 8.10 (Mitte) sind zudem unterschiedliche Zusammenhangsmuster zwischen den Schwierigkeitsstufen und den Bedingungs-Faktorladungen erkennbar: Während die NASA-RTLX-Komponente 1 einen linearen Anstieg der Bedingungs-Faktorladungen mit zunehmender Aufgabenschwierigkeit beschreibt, zeigt die kardiale Komponente 1 einen abflachenden Anstieg bzw. ein exponentielles Wachstum mit Schranke (degressiv).

In Abbildung 8.10 (unten) ist der Zusammenhang der Parameter mit der kardialen Komponente zu sehen. Den höchsten Zusammenhang zeigen hier die Parameter MQSD, CVSD, EKG.HF und SVDEn mit $|a| \geq .30$. Marginale Faktorladungen ($|a| \geq .20$) ergeben sich für die Parameter HR, HP, SDNN, pNN50, pNN20, CVNN, Shannon.h, EKG.TP und EKG.LF. An den Faktorladungen der Personen (Abbildung 8.10, oben) ist zu erkennen, dass für die NASA-RTLX-Komponente 1 (gestrichelte rote Linie) geringe und für die kardiale Komponente 1 (durchgezogene rote Linie) hohe interindividuelle Unterschiede vorliegen.

8.3.3 Forschungsfrage F2.1: Können aus den elektrodermalen Einzelparametern elektrodermale Profile abgeleitet werden, welche die subjektive mentale Beanspruchung zuverlässig repräsentieren?

Das Ergebnis der Kreuzvalidierung ist für die Regression vom NASA-RTLX auf die elektrodermalen Parameter in Tabelle 8.5 dargestellt.⁷ Hier zeigt sich ein bedeutsamer Zuwachs an Vorhersagegüte in der Prädiktor- und Kriteriumsmatrix nur für die Komponente 1. Sie erklärt im Originalmodell lediglich 7.6 % der

⁷ Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Prädiktor-matrix beträgt: $M [99 \% CI] = 0.32 [0.12, 0.59]$. Der mittlere zufällige Zuwachs $\text{rand}\Delta R_{\text{PLS}}^2$ der Kriteriumsmatrix beträgt: $M [99 \% CI] = -2.96 [-7.85, -0.40]$.

Tabelle 8.5: NPLS-Regression des NASA-RTLX auf elektrodermale Parameter. Dargestellt ist der Zuwachs an erklärter Varianz durch die Komponenten 1 bis 5 jeweils für das Originalmodell (ΔR_{PLS}^2) und das kreuzvalidierte Modell ($val\Delta R_{PLS}^2$)

Komponente	Prädiktormatrix		Kriteriumsmatrix		
	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	ΔR_{PLS}^2	$val\Delta R_{PLS}^2$	valSQR
0					1054.00
1	7.64	6.47 [†]	64.59	37.73 [†]	627.10***
2	0.88	-2.78	9.13	3.05 [†]	596.41***
3	1.38	-0.12	4.73	0.50 [†]	591.37
4	-0.03	0.09	2.77	-0.17	593.05
5	1.68	-0.98	2.76	-0.29	595.97

Bedeutsamer Abfall der valSQR: *** $p < .001$, ** $p < .01$, * $p < .05$.

Bedeutsamer Zuwachs der Prädiktormatrix: [†] $val\Delta R^2 > 0.59$, der Kriteriumsmatrix: [†] $val\Delta R^2 > 0.0$.

Varianz in den elektrodermalen Parametern, kann aber 64.6 % der Varianz im NASA-RTLX vorhersagen. In der kreuzvalidierten Vorhersage verringert sich jedoch der Anteil erklärter Varianz auf 37.7 %. Die Ladungen der Komponente 1 sind in Abbildung 8.11 dargestellt.

Zur Interpretation wird die Dimension der Bedingungen (Abbildung 8.11, Mitte) betrachtet. Eine tabellarische Übersicht der Bedingungs-Faktorladungen ist im Anhang A-30 zu finden. Die extrahierte NASA-RTLX-Komponente 1 ist in ihren Bedingungs-Faktorladungen (siehe Abbildung 8.11, Mitte, gestrichelte rote Linie) mit dem Verlauf der NASA-RTLX-Mittelwerte (gestrichelte dunkelgraue Linie) identisch und repräsentiert somit die subjektive mentalen Beanspruchung. In der Bedingungsdimension kovariiert die NASA-RTLX-Komponente 1 mit der elektrodermalen Komponente 1 (durchgezogene rote Linie). Dabei sagt die elektrodermale Komponente 1 die NASA-RTLX-Komponente 1 mit $B = 0.81$, $t(1058) = 22.05$, $p < .001$ vorher. In Abbildung 8.11 (Mitte) ist zu sehen, dass mit einer steigenden Aufgabenschwierigkeit nicht nur eine höhere subjektive

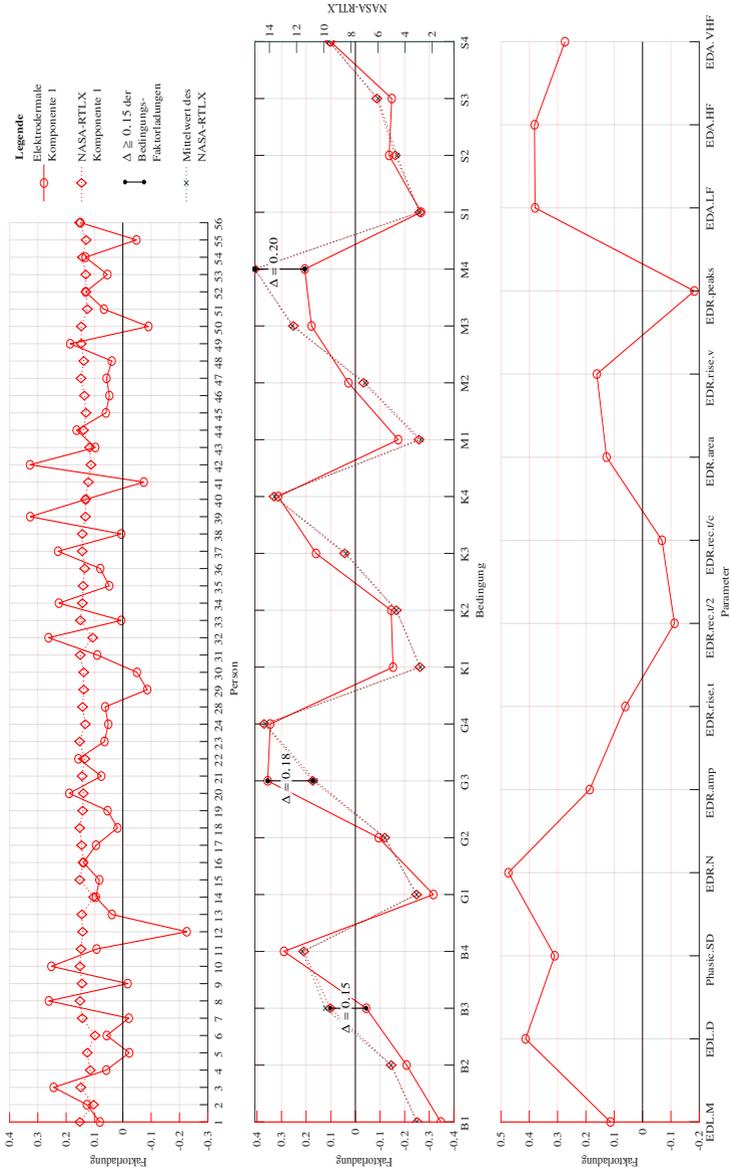


Abbildung 8.11: Faktorladungen α für das NASA-RTLX und die elektrodermale Aktivität jeweils für Personen (oben), Bedingungen (Mitte) und Parameter (unten). Zum Vergleich werden die Mittelwerte des NASA-RTLX in den Bedingungen (Mitte) angegeben.

mentale Beanspruchung, sondern auch eine höhere elektrodermale Aktivität einhergeht.

Für die drei Bedingungen M4, B3 und G3 sind jedoch Abweichungen zwischen den Bedingungs-Faktorladungen der Komponenten zu erkennen. In der Bedingung M4 liegt die elektrodermale Faktorladung um .20 über der Faktorladung der NASA-RTLX-Komponente. Somit ist der Anstieg zwischen der dritten und vierten Schwierigkeitsstufe der Matrizenaufgabe für die elektrodermale Komponente steiler als für die NASA-RTLX-Komponente. Des Weiteren zeigt die Bedingung B3 für die NASA-RTLX-Komponente eine marginal höhere und die Bedingung G3 eine marginal geringere Faktorladung als die elektrodermale Komponente. Letztere Differenz zeigt auf, dass zwischen G3 und G4 nicht die erwartete Steigung, sondern sogar ein minimaler Abfall von $a = .36$ auf $a = .35$ in der elektrodermalen Aktivität vorliegt. Insgesamt ist für die elektrodermale Komponente 1 ein eigenes Interaktionsmuster aus Aufgabenschwierigkeit und Aufgabentyp zu vermuten. Dieses ist somit divergent zu dem vorgefundenen Interaktionsmuster auf die subjektive mentale Beanspruchung (siehe Abschnitt 8.3.1).

Die Faktorladungen der elektrodermalen Parameter (Abbildung 8.11, unten) definieren das elektrodermale Profil. Auf diesem laden bedeutsam die Parameter EDL.D und EDR.N mit $a > .40$, gefolgt von EDA.LF, EDA.HF und Phasic.SD mit $a > .30$. Der Parameter EDA.VHF zeigt eine marginale Faktorladung ($a = .27$). Die Faktorladungen der Personen sind in Abbildung 8.11 (oben) zu sehen. Die Person-Faktorladungen der NASA-RTLX-Komponente (gestrichelte rote Linie) weisen dabei kaum interindividuelle Unterschiede auf, wohingegen die elektrodermalen Person-Faktorladungen (durchgezogene rote Linie) hohe interindividuelle Unterschiede zeigen.

8.3.4 Forschungsfrage F7.1: Sind die aufgabenübergreifenden kardialen Profile mentaler Beanspruchung der Studie III kongruent zu den kardialen Profilen mentaler Beanspruchung der Studien I und II?

Die kardialen Profile des NASA-RTLX aus Studie I und II wurden bereits als identische Repräsentanten der mentalen Beanspruchung identifiziert (siehe Abschnitt 6.4.2). Mit diesen beiden Profilen wird das in Studie III extrahierte kardiale Profil des NASA-RTLX auf die Kongruenz hin untersucht. Des Weiteren wird die Kongruenz mit dem kardialen Profil der Trefferquote aus Studie I bestimmt. Dieses Profil wurde in Studie I als ein Mischprofil der mentalen und emotionalen Beanspruchung interpretiert.

In Tabelle 8.6 sind die Parameter-Faktorladungen der Studie I bis III aufgeführt. Bereits hier ist die Ähnlichkeit in der Ladungsstruktur der kardialen Profile aus den Studien I bis III erkennbar. Zur statistischen Prüfung werden die Kongruenzkoeffizienten herangezogen. Eine Identität der Parameter-Faktorladungen ergibt sich sowohl zwischen den kardialen Profilen des NASA-RTLX aus Studie I und III ($C = .98$) als auch zwischen den kardialen Profilen des NASA-RTLX aus Studie II und III ($C = 1.0$). Dabei liegen die a_{Δ} der Parameter bei maximal $.07$ zwischen Studie III und I und bei maximal $.05$ zwischen Studie III und II (siehe Tabelle 8.7).

Der Kongruenzkoeffizient zwischen dem kardialen Profil des NASA-RTLX aus Studie III und dem kardialen Profil der Trefferquote aus Studie I deckt eine Ähnlichkeit der Profile auf Ebene der Parameter auf, $C = .93$. Tabelle 8.7 (letzte Spalte) belegt ausschließlich für den Parameter HP einen marginalen Faktorladungsunterschied von $.15$. Für die restlichen Parameter werden keine bedeutsamen oder marginalen Faktorladungsunterschiede sichtbar. Nichtsdestoweniger sind die Faktorladungsunterschiede zwischen dem Profil des NASA-RTLX aus Studie III und dem kardialen Profil der Trefferquote aus

Tabelle 8.6: Parameter-Faktorladungen a für die kardialen Komponenten des NASA-RTLX (aus Studie I, II und III) und der Trefferquote (aus Studie I)

Parameter	NASA-RTLX			Trefferquote
	Studie III	Studie II	Studie I	Studie I
HR	.22	.21	.28	.09
SDHR	-.11	-.12	-.09	-.22
HP	-.24	-.24	-.30	-.09
SDNN	-.26	-.27	-.26	-.31
MQSD	-.34	-.35	-.32	-.31
pNN50	-.28	-.32	-.26	-.27
pNN20	-.23	-.23	-.30	-.20
CVSD	-.31	-.32	-.28	-.31
CVNN	-.20	-.21	-.19	-.28
Triang	-.15	-.14	-.15	-.24
Shannon.h	-.23	-.22	-.20	-.29
EKG.TP	-.27	-.25	-.25	-.23
EKG.LF	-.23	-.18	-.23	-.23
EKG.HF	-.35	-.32	-.30	-.25
EKG.LFn	.06	.05	.10	-.06
EKG.LF/HF	.03	.03	.10	-.09
DFA1	.04	.08	.10	-.02
SamEn	.00	.00	-.04	.03
SVDEn	-.31	-.31	-.28	-.29
T.amp	-.08	-.08	-.03	-.14
T.area	-.08	-.09	-.04	-.18

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

Studie I im Durchschnitt um .40 bis .60 höher als die zwischen den kardialen Profilen des NASA-RTLX der Studien I bis III (vgl. Tabelle 8.7, Spalte 2 und 3).

Tabelle 8.7: Differenz zwischen den Parameter-Faktorladungen der kardialen Komponenten des NASA-RTLX (Na) aus Studie I, II und III sowie der kardialen Komponente der Trefferquote (Tq) aus Studie I

Parameter	Na (III) – Na (I)	Na (III) – Na (II)	Na (III) – Tq (I)
HR	–.06	.01	.13
SDHR	–.02	.01	.11
HP	.06	–.01	–.15
SDNN	.00	.01	.05
MQSD	–.02	.01	–.03
pNN50	–.01	.04	–.01
pNN20	.07	.00	–.03
CVSD	–.02	.01	.00
CVNN	–.02	.01	.08
Triang	.00	–.01	.08
Shannon.h	–.03	–.01	.06
EKG.TP	–.02	–.01	–.04
EKG.LF	.00	–.05	.00
EKG.HF	–.04	–.02	–.10
EKG.LFn	–.03	.01	.13
EKG.LF/HF	–.07	.00	.12
DFA1	–.06	–.04	.06
SamEn	.04	.00	–.03
SVDEn	–.02	.01	–.01
T.amp	–.05	.01	.06
T.area	–.05	.00	.10

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, **marginal** $|a_{\Delta}| \geq .15$, **nicht bedeutsam** $|a_{\Delta}| < .15$.

8.3.5 Forschungsfrage F7.2: Sind die aufgabenübergreifenden elektrodermalen Profile mentaler Beanspruchung der Studie III kongruent zu den elektrodermalen Profilen mentaler Beanspruchung der Studien I und II?

Die Kongruenz des in Studie III extrahierten elektrodermalen Profils wird mit den Repräsentanten der mentalen Beanspruchung aus Studie I und II berechnet. Zu diesen zählen die elektrodermale Profil des NASA-RTLX der Studien I und

Tabelle 8.8: Differenz zwischen den Parameter-Faktorladungen der elektrodermalen Profile des NASA-RTLX (Na) aus Studie III, II und I und der Valenz (Va) aus Studie I

Parameter	Na (III) – Na (I)	Na (III) – Na (II)	Na (III) – Va (I)
EDL.M	–.25	–.05	.00
EDL.D	.19	.15	.22
Phasic.SD	.06	–.05	.00
EDR.N	.04	.14	.40
EDR.amp	.03	–.13	–.23
EDR.rise.t	.29	.00	.01
EDR.rec.t/2	.22	.04	–.11
EDR.rec.t/c	.26	.12	–.07
EDR.area	.07	–.09	–.16
EDR.rise.v	–.09	–.13	–.26
EDR.peaks	.18	.03	–.15
EDA.LF	.26	.07	–.05
EDA.HF	.17	–.01	–.03
EDA.VHF	.10	–.01	.02

Differenz der Ladungen ist **bedeutsam** $|a_{\Delta}| \geq .20$, marginal $|a_{\Delta}| \geq .15$, nicht bedeutsam $|a_{\Delta}| < .15$.

Tabelle 8.9: Parameter-Faktorladungen α für die elektrodermalen Komponenten des NASA-RTLX (aus Studie I, II und III) und der Valenz (aus Studie I)

Parameter	NASA-RTLX			Valenz
	Studie III	Studie II	Studie I	Studie I
EDL.M	.11	.16	.36	.11
EDL.D	.41	.26	.22	.19
Phasic.SD	.31	.36	.25	.31
EDR.N	.47	.33	.43	.07
EDR.amp	.19	.32	.16	.42
EDR.rise.t	.06	.06	-.23	.05
EDR.rec.t/2	-.11	-.15	-.33	.00
EDR.rec.t/c	-.07	-.19	-.33	.00
EDR.area	.13	.22	.06	.29
EDR.rise.v	.16	.29	.25	.42
EDR.peaks	-.18	-.21	-.36	-.03
EDA.LF	.38	.31	.12	.43
EDA.HF	.38	.39	.21	.41
EDA.VHF	.27	.28	.17	.25

Faktorladung ist **bedeutsam** $|a| \geq .30$, **marginal** $|a| \geq .20$, nicht bedeutsam $|a| < .20$.

II.⁸ Das Profil des NASA-RTLX aus Studie III ist dabei als verschieden zu dem Profil des NASA-RTLX aus Studie I zu interpretieren, $C = .77$. Tabelle 8.8 (zweite Spalte) deckt bedeutsame Unterschiede für fünf und marginale Unterschiede für weitere drei Parameter auf. Auch die Parameter-Faktorladungen in Tabelle 8.9 zeigen eindeutige Unterschiede in der Ladungsstruktur dieser Profile. Hierbei laden die Parameter EDL.D, Phasic.SD, EDA.LF und EDA.HF hoch auf dem elektrodermalen Profil des NASA-RTLX aus Studie III, jedoch nur marginal oder unbedeutend gering auf dem Profil des NASA-RTLX aus Studie I.

⁸ Wie bereits in Abschnitt 7.4.2 diskutiert, zeigen die elektrodermalen Profile des NASA-RTLX der Studien I und II keine Ähnlichkeit untereinander.

Im Gegensatz dazu laden die Parameter EDL.M, EDR.rec.t/2, EDR.rec.t/c und EDR.peaks bedeutsam auf dem Profil des NASA-RTLX aus Studie I, jedoch gering auf dem Profil des NASA-RTLX aus Studie III. Lediglich der Parameter EDR.N hängt bedeutsam mit dem elektrodermalen Profil aus Studie I und III zusammen.

Das elektrodermale Profil des NASA-RTLX aus Studie II ist als ähnlich zum Profil des NASA-RTLX aus Studie III zu interpretieren: Die Ähnlichkeit der Parameter-Faktorladungen liegt dabei mit $C = .94$ knapp unterhalb der Identitätsschwelle. In Tabelle 8.8 (dritte Spalte) ist alleinig für den Parameter EDL.D ein marginaler Unterschied zu sehen. Die hohe Ähnlichkeit ist an der Ladungsstruktur in Tabelle 8.9 erkennbar: Die Parameter Phasic.SD, EDR.N, EDA.LF und EDA.HF zeigen übereinstimmend bedeutsame Faktorladungen in beiden Studien. Keine Übereinstimmung findet sich für die Parameter EDL.D, EDR.amp, EDR.area, EDR.rise.v und EDR.peaks. Diese laden entweder auf dem Profil der Studie II oder auf dem Profil der Studie III bedeutsam bzw. marginal (vgl. Spalte 2 und 3 in Tabelle 8.9).

In Studie II wurde eine Ähnlichkeit zwischen dem elektrodermalen Profil mentaler Beanspruchung (Profil des NASA-RTLX) und dem Profil emotionale Beanspruchung aus Studie I (Profil der Valenz) festgestellt (siehe Abschnitt 7.3.7). Aus diesem Grund wird auch das elektrodermale Profil der Studie III auf eine Kongruenz mit dem Profil der emotionalen Beanspruchung (aus Studie I) geprüft. Der Kongruenzkoeffizient der Parameter-Faktorladungen zeigt mit $C = .80$ auf, dass das elektrodermale Profil des NASA-RTLX der Studie III keine Ähnlichkeit zum elektrodermalen Profil der Valenz aus Studie I besitzt. Beide unterscheiden sich in vier Parametern (EDL.D, EDR.N, EDR.amp und EDR.rise.v) bedeutsam voneinander (siehe Tabelle 8.8 und 8.9, letzte Spalte).

8.4 Diskussion

Dieses Kapitel beinhaltet die Zusammenfassung und Diskussion der Befunde der Studie III. Abschnitt 8.4.1 beinhaltet die Zusammenfassung und Interpretation der Ergebnisse zur Manipulationskontrolle. Darauf folgend werden die extrahierten Profile kardialer Aktivität (Abschnitt 8.4.2) und elektrodermalen Aktivität (Abschnitt 8.4.3) zusammengetragen und deren Implikationen diskutiert. In Abschnitt 8.4.4 werden die Einschränkungen der Studie III dargelegt.

8.4.1 Zusammenfassung und Interpretation der Manipulationskontrolle

Die Ergebnisse belegen eine erwartungskonforme und bedeutsame Veränderung der subjektiven mentalen Beanspruchung durch die Schwierigkeitsstufen der Aufgaben: Je schwerer die Aufgabe, desto höher war die wahrgenommene mentale Beanspruchung. Folglich ist die Hypothese H3.1 zu bestätigen. Im Gegensatz dazu muss die Hypothese H3.2 verworfen werden: Denn entgegen der Erwartung zeigten sich bedeutsame Unterschiede in der subjektiven mentalen Beanspruchung zwischen den Aufgabentypen. Zudem wurden bei der Analyse der Unterfragestellung UF3.1 aufgabenspezifische Steigungen in der mentalen Beanspruchung zwischen den Schwierigkeitsstufen der Aufgabe gefunden (bedeutsame Interaktion Aufgabenschwierigkeit \times Aufgabentyp). Von den insgesamt 15 Steigungen zwischen den benachbarten Schwierigkeitsstufen zeigten fünf eine bedeutsame Abweichung zur durchschnittlichen Steigung über alle Aufgabentypen hinweg (siehe Abschnitt 8.3.1). Allein in der Gedächtnisaufgabe fanden sich keine aufgabenspezifischen Steigungen in der mentalen Beanspruchung. Hier stieg die mentale Beanspruchung mit jeder weiteren Schwierigkeitsstufe um den gleichen Betrag, wie die durchschnittliche Steigung über alle Aufgabentypen hinweg.

Eine Erklärung für den nicht erwartungskonformen Haupteffekt (des Aufgabentyps) und den Interaktionseffekt ist die Operationalisierung der fünf Bereiche der Kognition. Dabei unterscheiden sich die Aufgabentypen nicht nur in der Art des Wissensabrufs, sondern auch in der Art der Lösungseingabe und der Anzahl präsentierter Aufgaben pro Durchgang. Diese Störfaktoren könnten die Wahrnehmung der mentalen Beanspruchung beeinflusst und zu den vorgefundenen Effekten geführt haben. Diese nicht erwartungskonformen Effekte sind für die Analyse und Interpretation der Forschungsfragen F1.1 und F2.1 nicht hinderlich, solange der Hauptanteil der Varianz auf die mentalen Belastungsfaktoren zurückzuführen ist und die konfundierenden Varianzquellen bekannt sind. Insgesamt erklärt der Haupteffekt des Aufgabentyps 3.8 % und der Interaktionseffekt 5.3 % der Varianz in der subjektiven mentalen Beanspruchung. Beide Effekte spielen somit im Vergleich zum Haupteffekt der Aufgabenschwierigkeit mit 52.2 % Varianzaufklärung eine eher untergeordnete Rolle.

Zusammenfassend ist in Hinblick auf die Fragestellung KF3 eine gelungene Induktion der subjektiven mentalen Beanspruchung durch die mentalen Belastungsfaktoren aus den fünf kognitiven Bereichen festzustellen.

8.4.2 Zusammenfassung und Interpretation der kardialen Profile

In Hinblick auf die Forschungsfrage F1.1 konnte in Studie III ein kardiales Profil des NASA-RTLX extrahiert werden. Dieses repräsentiert die subjektive mentale Beanspruchung über verschiedene Bereiche der Kognition hinweg. Im Originalmodell erklärt es mit $B = 0.50$ bis zu 48.6 % der Varianz im NASA-RTLX. Allerdings nimmt die Vorhersagegüte für Personen, die dem Modell nicht bekannt sind um 19.2 % ab. Auch wenn der verbleibende Anteil von 48.6 % erklärter Varianz weiterhin bedeutsam und beachtlich ist, verdeutlicht die hohe Reduktion die Relevanz der personenspezifischen Unterschiede für die

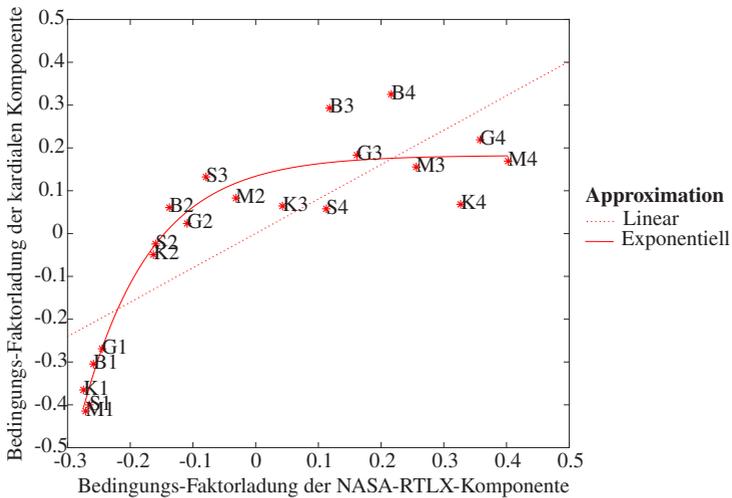


Abbildung 8.12: Zusammenhang zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente und der kardialen Komponente der Studie III.

Vorhersagegüte der subjektiven mentalen Beanspruchung durch das kardiale Profil. Den stärksten Zusammenhang mit dem kardialen Profil haben die Parameter MQSD, CVSD, EKG.HF und SVDEn mit $|a| > .30$. Diese nehmen konform zu den bisherigen Forschungsbefunden bei Zunahme der mentalen Beanspruchung ab (vgl. Abschnitt 3.2.4). Eine endgültige Bewertung der kardialen Einzelparameter erfolgt am Ende der Arbeit in Abschnitt 9.1.2.

Ein Zusatzbefund der Studie III ist der (degressiv) exponentielle Zusammenhang zwischen der kardialen Komponente und der subjektiven mentalen Beanspruchung in der Bedingungsdimension. Zur Veranschaulichung sind in Abbildung 8.12 die Bedingungs-Faktorladungen der kardialen Komponente gegen die Faktorladungen der NASA-RTLX-Komponente aufgetragen. Hierbei ergibt sich eine höhere Passung für die (degressiv) exponentielle Wachstumskurve (durchgezogene rote Linie) im Vergleich zur linearen Wachstumskurve

(gestrichelte rote Linie): Die durchschnittlichen Abweichungen⁹ sind für die lineare Approximation doppelt so hoch (RMSE = 0.133) wie für die exponentielle Approximation (RMSE = 0.065). In den Studien I und II konnte der exponentielle Zusammenhang nicht aufgedeckt werden, da hier nur zwei Schwierigkeitsstufen umgesetzt wurden. Das (degressiv) exponentielle Zusammenhangsmuster zeigt auf, dass die kardiale Aktivität (des kardialen Profils) mit steigender subjektiver mentaler Beanspruchung zwar zunimmt, jedoch dieser Anstieg nach oben hin abflacht. Dies gilt für alle fünf Bereiche der Kognition.

Das in Studie III extrahierte aufgabenübergreifende Profil mentaler Beanspruchung ist identisch zu denjenigen kardialen Profilen, welche in Studie I und II als Repräsentanten mentaler Beanspruchung identifiziert wurden (Profile des NASA-RTLX). Die kardialen Profile des NASA-RTLX aus Studie I und II wurden bereits im Rahmen der Studie II (vgl. Unterkapitel 7.4.2) als identische Repräsentanten der mentalen Beanspruchung interpretiert. Folglich sind die kardialen Profile des NASA-RTLX aus der Studie I, II und III zu einem Profil der mentalen Beanspruchung zusammenzufassen. Darüber hinaus zeigt das kardiale Profil mentaler Beanspruchung aus Studie III eine Ähnlichkeit zu demjenigen kardialen Profil, welches in Studie I als Mischprofil der mentalen und emotionalen Beanspruchung interpretiert wurde (Profil der Trefferquote). Ein marginaler Unterschied zwischen den beiden Profilen findet sich nur für den Parameter HP, welcher bereits in Studie I zur Differenzierung mentaler und emotionaler Beanspruchung diente (siehe Abschnitt 6.3.10). Dieser Befund ist konsistent zu den Ergebnissen aus Studie II (vgl. Unterkapitel 7.4.2).

Insgesamt ist die Forschungsfrage F7.1 zu bejahen und somit als weiterer Beleg für die Strukturvalidität des kardialen Profils mentaler Beanspruchung zu werten. Die Replizierbarkeit des kardialen Profils mentaler Beanspruchung bestätigt, dass die in Studie I und II gefundenen kardialen Profile mentaler Beanspruchung keine Zufallsbefunde waren. Zudem belegen die Befunde der Studie III die

⁹ Die mittlere Abweichung der Approximationskurve wird als Wurzel der mittleren Fehlerquadratsumme (engl.: Root Mean Square Error; RMSE) berechnet.

Generalisierbarkeit des kardialen Profils der mentalen Beanspruchung über verschiedene Bereiche der Kognition.

8.4.3 Zusammenfassung und Interpretation der elektrodermalen Profile

Im Rahmen der Forschungsfrage F2.1 wurde ein elektrodermales Profil als Repräsentant der subjektiven mentalen Beanspruchung identifiziert. In Abbildung 8.13 ist zu sehen, dass das elektrodermale Profil die verschiedenen Schwierigkeitsstufen über die fünf Aufgabentypen hinweg zuverlässig abbildet ($B = 0.81$). Im Gegensatz zum kardialen Profil mentaler Beanspruchung ist der Zusammenhang zwischen der elektrodermalen Komponente und der NASA-RTLX-Komponente in der Bedingungsdimension linear (vgl. Abbildung 8.13). Das elektrodermale Profil mentaler Beanspruchung klärt im Originalmodell der Studie III bis zu 64.6 % der Varianz in der subjektiven mentalen Beanspruchung auf. Allerdings reduziert sich der Anteil erklärter Varianz um 26.9 % (auf 37.7 %) im kreuzvalidierten Modell. Dies verdeutlicht die Relevanz der personenspezifischen Unterschiede für eine zuverlässige Vorhersage der mentalen Beanspruchung durch das elektrodermale Profil. Auf diesem Profil laden die Parameter Phasic.SD, EDL.D, EDR.N, EDA.LF und EDA.HF am höchsten. Eine endgültige Bewertung der elektrodermalen Einzelparameter erfolgt am Ende der Arbeit in Unterkapitel 9.1.4.

In Bezug auf Forschungsfrage F7.2 wurde eine sehr hohe Ähnlichkeit, jedoch keine Identität, zum elektrodermalen Profil des NASA-RTLX aus Studie II festgestellt. Dieses Profil wurde in Studie II als ein Repräsentant der mentalen Beanspruchung interpretiert. Eine mögliche Erklärung für den fehlenden Identitätsnachweis könnte darin liegen, dass das elektrodermale Profil des NASA-RTLX aus Studie II die mentale Beanspruchung eines spezifischen Aufgabentyps (nämlich der Multiplikationsaufgabe) repräsentiert. Demgegenüber bildet das elektrodermale Profil aus Studie III die subjektive mentale Beanspru-

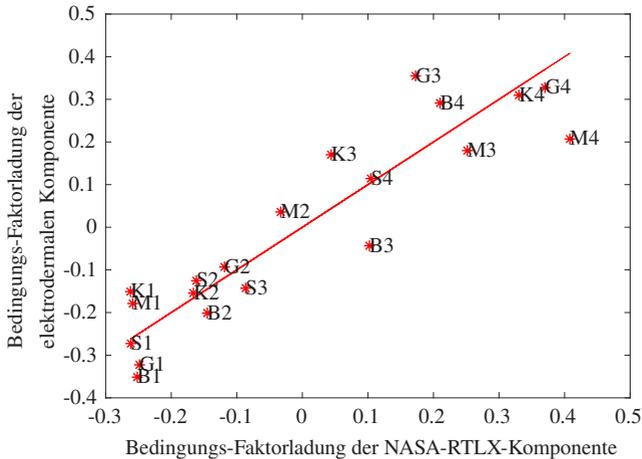


Abbildung 8.13: Zusammenhang zwischen den Bedingungs-Faktorladungen der NASA-RTLX-Komponente und der elektrodermalen Komponente der Studie III.

chung über verschiedene Aufgabentypen hinweg ab. Eine weitere Ursache für die fehlende Identität könnte darin begründet sein, dass die Konfundierung mit der emotionalen Beanspruchung in Studie III geringer ausfällt als in Studie II. Immerhin weist das elektrodermale Profil des NASA-RTLX aus Studie II, jedoch nicht das Profil aus Studie III, eine Ähnlichkeit zum elektrodermalen Profil der emotionalen Beanspruchung (Profil der Valenz aus Studie I) auf.

Ebenfalls im Rahmen der Forschungsfrage F7.2 wurde die Kongruenz mit demjenigen elektrodermalen Profil untersucht, welches in Studie I als ein Repräsentant der mentalen Beanspruchung interpretiert wurde (Profil des NASA-RTLX). Anzumerken ist, dass bereits das elektrodermale Profil der mentalen Beanspruchung aus Studie II keine Ähnlichkeit zu diesem elektrodermalen Profil der mentalen Beanspruchung aus Studie I aufwies (siehe Abschnitt 7.4.3). Offen blieb bisher die Frage, welches der beiden Profilen tatsächlich die mentale Beanspruchung repräsentiert. Mit den vorliegenden Ergebnissen ist ersichtlich, dass das elektrodermale Profil mentaler Beanspruchung aus Studie I weder in Studie II noch in Studie III repliziert werden konnte. Dieser Befund er-

härtert die Annahme einer Konfundierung in Studie I. Vor allem der große Beanspruchungsunterschied zwischen den zwei Aufgabentypen (Balltracking und IAPS) der Studie I könnte die Ladungsstruktur des extrahierten Profils des NASA-RTLX mit beeinflusst haben. In Abschnitt 6.4.4 wurde bereits diskutiert, dass dieser Beanspruchungsunterschied potentiell konfundiert ist mit dem Grad an Bewegung, der Reihenfolge der Versuchsblöcke oder dem Typ der Aufgabe. Aber auch die Verwendung unterschiedlicher Messsensoren könnte zur Extraktion unterschiedlicher Profile zwischen den Studien geführt haben. Da das elektrodermale Profil mentaler Beanspruchung aus Studie I nicht repliziert werden konnte, ist die Interpretation dieses Profils als Repräsentant der mentalen Beanspruchung anzuzweifeln.

8.4.4 Limitationen

Im Folgenden werden die Einschränkungen der Ergebnisinterpretation aus Studie III diskutiert.

Einschränkungen der Messdaten

Wie bereits für die Studien I und II ergeben sich auch für die Studie III Einschränkungen durch die ungeklärte Reliabilität kardialer Ultra-Kurzzeitmaße sowie durch die fehlende Kontrolle der Atemaktivität. Beide Limitationen wurden bereits in Unterkapitel 6.4.4 diskutiert und sind entsprechend auf die Ergebnisse der Studie III übertragbar. Eine ebenfalls in Studie I und II beschriebene Einschränkung ist der potentielle Einfluss der Raumtemperatur auf die elektrodermale Messung. Für die Studie III liegt zwar ein geringer, jedoch bedeutsamer Anstieg für die Raumtemperatur um durchschnittlich 0.48 °C vor. Auch wenn die Bedingungen in ihrer Reihenfolge randomisiert wurden, um Reihenfolgeeffekte zu vermeiden, kann eine Konfundierung durch die Raumtemperatur nicht gänzlich ausgeschlossen werden.

Eine weitere potentielle Einschränkung ist die Verwendung unterschiedlicher Messsensoren bei der Erhebung kardialer als auch elektrodermalen Aktivität in der Studie I. In Studie II und III wurde die kardiale Aktivität mit dem varioport-B Biosignal Rekorder und in Studie I mit dem Schimmer3 ECG Unit gemessen. Die Sensoren unterscheiden sich sowohl in den Ableitungen am Brustkorb als auch in den Abtastraten (siehe Unterkapitel 5.1.1). Die Ähnlichkeit der kardialen Profile zwischen den Studien belegt, dass die Sensorunterschiede zumindest keinen bedeutenden Einfluss auf die extrahierten kardialen Profile hatten. Im Gegensatz dazu kann der Einfluss der unterschiedlichen Sensoren auf die extrahierten elektrodermalen Profile nicht ausgeschlossen werden. In Studie II und III wurde der varioport-B Biosignal Rekorder (Abtastrate 500 Hz) und in Studie I die Shimmer3 GSR+ Unit (Abtastrate 51.2 Hz) verwendet (siehe Unterkapitel 5.2.1). Zudem wurde in Studie II und III der Hautleitwert (unter Verwendung von Elektroden mit einer isotonischen Paste) und in Studie I der Hautwiderstand (unter Verwendung von Trockenelektroden) gemessen. Die fehlende Kongruenz der elektrodermalen Profile des NASA-RTLX aus Studie II und III zu dem aus Studie I könnte auf die Verwendung verschiedener Sensoren zurückzuführen sein.

Einschränkungen in der Konstruktrennung

Da eine psychometrische Trennung emotionaler und mentaler Beanspruchung kein Ziel der Studie III darstellte, wurde im Versuchsdesign nur eine reine mentale Belastungsinduktion umgesetzt. Somit können die Ergebnisse der Studie III keine Auskunft darüber geben, ob die extrahierten physiologischen Profile eine reine mentale Beanspruchung abbilden oder Einflüsse der emotionalen Beanspruchung beinhalten.

Eine weitere Einschränkung der Studie III besteht darin, dass die Unterschiede zwischen den Schwierigkeitsstufen in den verschiedenen Aufgabentypen nicht äquidistant sind. Dies ist durch den bedeutsamen Interaktionseffekt Aufgabenschwierigkeit \times Aufgabentyp auf die subjektive mentale Beanspruchung

belegt. Um den Interaktionseffekt aus Aufgabenschwierigkeit und Aufgabentyp auf die Extraktion physiologischer Profile zu ergründen, sind weitere Studien notwendig.

Einschränkungen in der externen Validität

Die Ergebnisse der Studie III sind auf den Laborkontext und die verwendeten Aufgabentypen der Studie beschränkt. Zudem handelt es sich um eine anfängliche Stichprobe, die nicht repräsentativ für die Gesamtpopulation ist. Beide Limitationen wurden bereits in vorangegangenen Kapiteln diskutiert (siehe Abschnitt 6.4.4) und sind auf die vorliegende Studie III übertragbar.

9 Psychophysiologische Profile in der beanspruchungssensitiven Mensch-Maschine-Schnittstelle

Dieses Kapitel fasst die wichtigsten Erkenntnisse der Arbeit zusammen, indem die physiologischen Profile aus den vorangegangenen Studien zusammengetragen und bewertet werden (Abschnitt 9.1). Anschließend erfolgt in Abschnitt 9.2 eine Einordnung der Profile in bestehende theoretische Erklärungsansätze. Zum Schluss werden Möglichkeiten der Anwendung in der Mensch-Maschine-Schnittstelle diskutiert (Abschnitt 9.3) und ein Ausblick für zukünftige Forschungsbemühungen gegeben (Abschnitt 9.4). Die Arbeit endet mit dem Fazit in Abschnitt 9.5.

9.1 Psychometrische Bewertung der extrahierten Profile

Die Studien aus den vorangegangenen Kapiteln dienen der Neustrukturierung des Zusammenhangs zwischen den physiologischen Einzelparametern und der mentalen Beanspruchung. In Anlehnung an das Rahmenmodell von Cacioppo et al. (Cacioppo & Tassinari, 1990; Cacioppo et al., 2000, 2007) wurden für die Neustrukturierung die physiologischen Parameter Φ mit Hilfe der NPLS-Regression in physiologische Profile höherer Ordnung Φ' umgewandelt. Auf diese Weise konnte zumindest ein Viele-zu-eins-Zusammenhang hergestellt werden. Ein Viele-zu-eins-Zusammenhang ist nach Cacioppo et al. (2000, 2007)

eine Mindestanforderung, um von den gemessenen physiologischen Messgrößen auf die mentale Beanspruchung schließen zu können (siehe Kapitel 4.1). Um eine Orientierung zu geben, welche der extrahierten Profile einen validen Rückschluss auf die mentale Beanspruchung erlauben, wird die Güte der physiologischen Profile bewertet. Zu diesem Zweck werden die Ergebnisse aus den drei Studien verglichen und zusammengefasst. Gleichzeitig werden die Ergebnisse in die Taxonomie psychophysiologischer Zusammenhänge (siehe Unterkapitel 4.1.3) von Cacioppo et al. (2017) eingeordnet. Zur Bewertung der Güte physiologischer Profile werden folgende Gütekriterien herangezogen:

1. Strukturvalidität: Hohe Kongruenz der physiologischen Profile mentaler Beanspruchung zwischen den Studien I, II und III.
2. Generalisierbarkeit im engeren Sinne: Gültigkeit der physiologischen Profile mentaler Beanspruchung über verschiedene mentale Belastungsfaktoren hinweg.
3. Sensitivität (Kriteriumsvalidität): Vorhersagbarkeit der mentalen Beanspruchung durch das physiologische Profil.
4. Spezifität (Diagnostizität) im weiteren Sinne: Abgrenzung der physiologischen Profile mentaler Beanspruchung zu den physiologischen Profilen der emotionalen Beanspruchung.

Die Generalisierbarkeit im weiteren Sinne – d. h. eine Übertragbarkeit der gefundenen physiologischen Profile auf die Gesamtheit aller möglichen Situationen außerhalb des Laborkontextes (externe Validität) – ist für keines der Studienergebnisse gegeben und wird daher nicht als Bewertungskriterium herangezogen.

Neben der Bewertung der physiologischen Profile wird zusätzlich eine Bewertung der physiologischen Einzelparameter vorgenommen. Auf diese Weise können relevante und redundante Parameter identifiziert werden. Ein Parameter gilt als relevant, wenn dieser zur Sensitivität (bedeutsame Faktorladung) und optimalerweise zur Spezifität (profilspezifische Faktorladung) des Profils

beiträgt. Im Gegensatz zur Relevanz ist die Redundanz der Einzelparameter nicht eindeutig quantifizierbar. Ein Hinweis auf Redundanz besteht, wenn inhaltlich ähnliche Parameter auf demselben physiologischen Profil bedeutsam oder zumindest marginal laden. Es ist anzumerken, dass Redundanz im Rahmen eines Profils erwünscht sein kann, da dadurch zufällige Fehlerquellen eliminiert und die Reliabilität gesteigert werden kann (siehe Kapitel 4.2).

9.1.1 Bewertung kardialer Profile

Zur Bewertung der extrahierten kardialen Profile der Studien I, II und III sind diese mitsamt der statistischen Kennwerte in Tabelle 9.1 zusammengefasst. Die Bewertung erfolgt anhand der zuvor beschriebenen Gütekriterien.

Strukturvalidität

Als Repräsentanten der mentalen Beanspruchung wurden die kardialen Profile des NASA-RTLX aus Studie I, II und III sowie das kardiale Profil der Trefferquote aus Studie II identifiziert. Gemäß dem Kongruenzkoeffizienten der Parameter-Faktorladungen (siehe Tabelle 9.1) sind alle vier Profile als identisch zu interpretieren und somit zu einem kardialen Profil mentaler Beanspruchung zusammenzufassen. Die Replizierbarkeit der Ladungsstruktur über die verschiedenen Studien hinweg belegt eine hinreichende Strukturvalidität für das kardiale Profil mentaler Beanspruchung. Ein nonkonformes Ergebnis liegt lediglich für das kardiale Profil der Trefferquote aus Studie I vor. Dieses zeigt nur ein ähnliches, jedoch kein identisches Ladungsmuster zu den restlichen Profilen mentaler Beanspruchung und wurde als Mischprofil emotionaler und mentaler Beanspruchung gedeutet. Die Reflektion emotionaler Beanspruchung durch das kardiale Profil der Trefferquote in Studie I ist darauf zurückzuführen, dass die gemessene Trefferquote der Studie I stärker durch die emotionalen als durch die mentalen Belastungsfaktoren beeinflusst wurde. Als Ursache hierfür kommt zum einen die gleichzeitige Induktion mentaler und emotionaler Belastungsfak-

Tabelle 9.1: Kardiale Profile des NASA-RTLX (Na), der Trefferquote (Tq), der Erregung (Er) und der Valenz (Va) aus den Studien I bis III zugeordnet zur Art der Beanspruchung. Zu sehen sind die Parameter-Faktorladungen α , deren Kongruenzkoeffizient C , die Bestimmtheitsmaße $\text{val}R_{\text{PLS}}^2$ und $\text{val}R_{\text{PLS}}^2$ sowie das Regressionsgewicht B .

Parameter	Mental			Emotional		Mischform	
	$C > .97$			Na (I)	Va (I)	$C = .95$	
	Na (III)	Na (II)	Tq (II)			Tq (I)	Er (I)
HR	.22	.21	.19	.28	.02	.09	.06
SDHR	-.11	-.12	-.13	-.09	-.18	-.22	-.28
HP	-.24	-.24	-.21	-.30	-.05	-.09	-.07
SDNN	-.26	-.27	-.28	-.26	-.22	-.31	-.31
MQSD	-.34	-.35	-.34	-.32	-.41	-.31	-.34
pNN50	-.28	-.32	-.31	-.26	-.28	-.27	-.21
pNN20	-.23	-.23	-.22	-.30	-.16	-.20	-.14
CVSD	-.31	-.32	-.32	-.28	-.40	-.31	-.31
CVNN	-.20	-.21	-.23	-.19	-.22	-.28	-.34
Triang	-.15	-.14	-.14	-.15	-.18	-.24	-.20
Shannon.h	-.23	-.22	-.23	-.20	-.14	-.29	-.28
EKG.TP	-.27	-.25	-.27	-.25	-.21	-.23	-.27
EKG.LF	-.23	-.18	-.20	-.23	-.19	-.23	-.26
EKG.HF	-.35	-.32	-.34	-.30	-.32	-.25	-.21
EKG.LFn	.06	.05	.05	.10	.12	-.06	-.05
EKG.LF/HF	.03	.03	.02	.10	.12	-.09	-.10
DFA1	.04	.08	.07	.10	.12	-.02	-.05
SamEn	.00	.00	.00	-.04	.06	.03	.11
SVDEn	-.31	-.31	-.31	-.28	-.39	-.29	-.31
T.amp	-.08	-.08	-.06	-.03	.02	-.14	.01
T.area	-.08	-.09	-.06	-.04	-.01	-.18	.03
Prädiktormatrix							
R_{PLS}^2	8.90	6.86	6.66	9.01	4.19	11.49	4.03
$\text{val}R_{\text{PLS}}^2$	7.43	5.34	5.11	6.48	1.88	10.03	1.52
B	0.50	0.71	-0.63	0.37	-0.59	-0.42	0.49
Kriteriumsmatrix							
R_{PLS}^2	48.63	82.37	63.6	73.17	33.20	40.58	31.00
$\text{val}R_{\text{PLS}}^2$	29.47	55.81	40.36	56.78	25.31	19.51	14.40

Faktorladung ist **bedeutsam** $|\alpha| \geq .30$, **marginal** $|\alpha| \geq .20$, nicht bedeutsam $|\alpha| < .20$.

toren im Versuchsdesign in Frage. Zum anderen könnte die Operationalisierung der leistungsbezogenen mentalen Beanspruchung als Sekundäraufgabenmaß zu diesem Befund geführt haben (siehe Abschnitt 6.4.4). Im Gegensatz zu Studie I wurde in Studie II die Trefferquote als Primäraufgabenmaß gemessen. Zudem wurde in Studie II keine emotionale Belastung im Versuch induziert. Hier ist das Profil der Trefferquote identisch zu den restlichen kardialen Profilen mentaler Beanspruchung.

Generalisierbarkeit

Die Gültigkeit des kardialen Profils mentaler Beanspruchung konnte sowohl über die verschiedenen Studien als auch über die verschiedene Aufgabentypen der Studie III hinweg nachgewiesen werden. Somit kann die subjektive und leistungsbezogene mentale Beanspruchung über die verschiedenen Bereiche der Kognition durch das kardiale Profil mentaler Beanspruchung abgebildet werden.

Eine Generalisierbarkeit des Profils auf Situationen mit unterschiedlichen körperlichen Anforderung ist in zukünftigen Forschungsbestrebungen zu untersuchen. Der vermutete Störeinfluss der körperlichen Anforderung auf die Extraktion des kardialen Profils mentaler Beanspruchung in Studie I konnte durch die Ergebnisse der Studien II und III widerlegt werden: Hier zeigten sich trotz unterschiedlicher Bewegungsanforderungen der Aufgaben identische Profile mentaler Beanspruchung. Zwar wurde die physische Belastung nicht explizit über die Aufgaben hinweg variiert, jedoch erforderten die jeweiligen Aufgaben der Studien ein unterschiedliches Ausmaß an Bewegung zur Lösung der Aufgabe (vgl. Balltrackingaufgabe in Studie I mit Kopfrechenaufgabe in Studie II). Nichtsdestoweniger befanden sich alle Bewegungsanforderungen im unteren physischen Beanspruchungsbereich. Bei allen Aufgaben musste nur ein Arm für die Eingabe der Aufgabenlösung bewegt werden. Ob sich das kardiale Profil mentaler Beanspruchung auch auf Situationen mit einer mittleren oder

hohen körperlichen Belastung generalisieren lässt, sollte in zukünftigen Studien geprüft werden.

Sensitivität

Über alle Studien hinweg zeigt sich, dass nur ein geringer Varianzanteil in der kardialen Aktivität mit der mentalen Beanspruchung zusammenhängt. In der kardialen Prädiktormatrix werden nämlich weniger als 10 % der Varianz durch die Komponente aufgedeckt (siehe R^2_{PLS} der Prädiktormatrix in Tabelle 9.1). Über 90 % der Varianz in den kardialen Parametern ist somit auf unbekannte Einflüsse zurückzuführen. Nichtsdestoweniger kann die kardiale Komponente zwischen 48.6 % und 82.4 % der Varianz in der leistungsbezogenen und subjektiven mentalen Beanspruchung erklären (siehe R^2_{PLS} der Kriteriumsmatrix in Tabelle 9.1). Bei der Bewertung der Sensitivität ist zu beachten, dass die Vorhersage der subjektiven und leistungsbezogenen mentalen Beanspruchung für unbekannte Personen mit nur 29.5 % bis 56.8 % Varianzaufklärung deutlich schlechter ausfällt als für Personen, deren personenspezifischen Faktorladungen im Vorhersagemodell bekannt ist (vgl. R^2_{PLS} mit $\text{val}R^2_{\text{PLS}}$ der Kriteriumsmatrix in Tabelle 9.1). Für eine bessere Vorhersagegüte ist vor allem die Kenntnis über die personenspezifische Höhe der kardialen Aktivität erforderlich, da diese sich zwischen den Personen stark unterscheidet. Die personenspezifische Leistung bzw. die personenspezifische Wahrnehmung der mentalen Beanspruchung spielt dabei nur eine untergeordnete Rolle. Die Sensitivität bei der Vorhersage unbekannter Personen ist zwar verringert, dennoch ist ein substanzieller Varianzanteil von durchschnittlich¹ 45.6 % (vgl. 66.9 % im Originalmodell) in der mentalen Beanspruchung durch das kardiale Profil der mentalen Beanspruchung prognostizierbar. Der niedrigste Varianzanteil von 29.5 % (bzw. 48.6 % im Originalmodell) liegt für die Studie III vor. Dies ist nicht zuletzt darauf zurückzuführen, dass ein (degressiv) exponentieller Zusammenhang zwischen

¹ In die Berechnung des Durchschnitts gehen die $\text{val}R^2_{\text{PLS}}$ (bzw. R^2_{PLS}) der kardialen Profile des NASA-RTLX aus Studie I bis III und des kardialen Profils der Trefferquote aus Studie II ein.

der kardialen Aktivität und der subjektiven mentalen Beanspruchung gefunden wurde. In den Studien I und II konnte ein solches Zusammenhangsmuster nicht identifiziert werden, da die Studien nur zwei Schwierigkeitsstufen zur Operationalisierung mentaler Beanspruchung verwendeten. Zusammenfassend zeigt das kardiale Profil der mentalen Beanspruchung eine zufriedenstellende Sensitivität für die mentale Beanspruchung mit einem (degressiv) exponentiellen Zusammenhangsmuster.

Spezifität

Die Spezifität wird in Hinblick auf die Abgrenzung zur emotionalen Beanspruchung bewertet. Für den Überblick ist das Strukturmodell der kardialen Profile in Abbildung 9.1 dargestellt. Als ein Repräsentant der emotionalen Beanspruchung wurde das kardiale Profil der Valenz aus Studie I interpretiert.² Die Sensitivität des Profils emotionaler Beanspruchung ist mit einer Varianzaufklärung von 25.3 % (kreuzvalidiertes Modell) bzw. 33.2 % (Originalmodell) als zufriedenstellend zu bewerten.³ Das kardiale Profil der emotionalen Beanspruchung ist mit einem Kongruenzkoeffizienten von .89 als ähnlich zum kardialen Profil der mentalen Beanspruchung zu interpretieren. Diese Ähnlichkeit ist in Abbildung 9.1 anhand der fehlenden Einfachstruktur der Parameter-Faktorladungen zu erkennen: Nur die Parameter HR und HP zeigen eine profilspezifische Faktorladung auf dem kardialen Profil mentaler Beanspruchung und einen Faktorladungsunterschied $\geq .20$ zum kardialen Profil emotionaler Beanspruchung. Die restlichen Parameter erlauben keine bedeutsame oder marginale Differenzierung der Profile. So laden alle bedeutsamen oder marginalen Parameter des emotionalen Beanspruchungsprofils gleichzeitig auf dem Profil der mentalen

² Das kardiale Profil der Erregung (Studie I) wurde analog zum kardialen Profil der Trefferquote (Studie I) als Mischprofil mentaler und emotionaler Beanspruchung interpretiert. In Bezug auf ihre Kongruenz liegen die Profile an der Identitätsschwelle von $C = .95$. Da es sich um Mischprofile handelt, werden diese nicht weiter fokussiert.

³ Aussagen zur Generalisierbarkeit oder Strukturvalidität des kardialen Profils der Valenz können nicht gemacht werden, da dieses Profil nur in Studie I extrahiert wurde.

Beanspruchung. Die fehlende Diagnostizität des kardialen Profils emotionaler Beanspruchung führt dazu, dass die subjektive mentale Beanspruchung auch durch das kardiale Profil der emotionalen Beanspruchung mit 22.5 % Varianzaufklärung (im Originalmodell der Studie I) bedeutsam vorhergesagt werden kann (siehe Abbildung 9.1). Umgekehrt ist es nicht möglich, die emotionale Beanspruchung (bzw. die subjektive Valenz der Studie I) durch das kardiale Profil der mentalen Beanspruchung bedeutsam vorherzusagen.⁴ Dies ist ein weiterer Beleg für die Spezifität des kardialen Profils mentaler Beanspruchung.

Einordnung in die Taxonomie psychophysiologischer Zusammenhänge

Insgesamt ist es in der vorliegenden Arbeit erfolgreich gelungen, die kardialen Einzelparameter Φ durch eine Neustrukturierung zu kardialen Profilen Φ' umzuwandeln (siehe Abbildung 9.1). Zusammenfassend zeigt das kardiale Profil der mentalen Beanspruchung eine hinreichende Strukturvalidität, Sensitivität, Generalisierbarkeit und eine ausreichende Spezifität. Allerdings ist eine Generalisierung des Profils mentaler Beanspruchung auf Situationen außerhalb des Labors nicht gegeben. Zudem ist die Spezifität des Profils mentaler Beanspruchung eingeschränkt, da eine Abgrenzung zum Profil emotionaler Beanspruchung lediglich durch zwei marginal bedeutsame Parameter gewährleistet wird. Ordnet man das in Abbildung 9.1 dargestellte Zusammenhangsmuster in die Taxonomie psychophysiologischer Zusammenhänge (Cacioppo & Tassinari, 1990; Cacioppo et al., 2017; zsf. Unterkapitel 4.1.3) ein, dann ist das kardiale Profil der mentalen Beanspruchung mindestens als eine Auswirkung (situationsspezifischer Viele-zu-eins-Zusammenhang) und höchstens als ein Marker (situationsspezifischer Eins-zu-eins-Zusammenhang) einzuordnen.

⁴ Die gekreuzten Regressionskoeffizienten beziehen sich auf die Ergebnisse der Studie I (siehe Abschnitt 6.3.10).

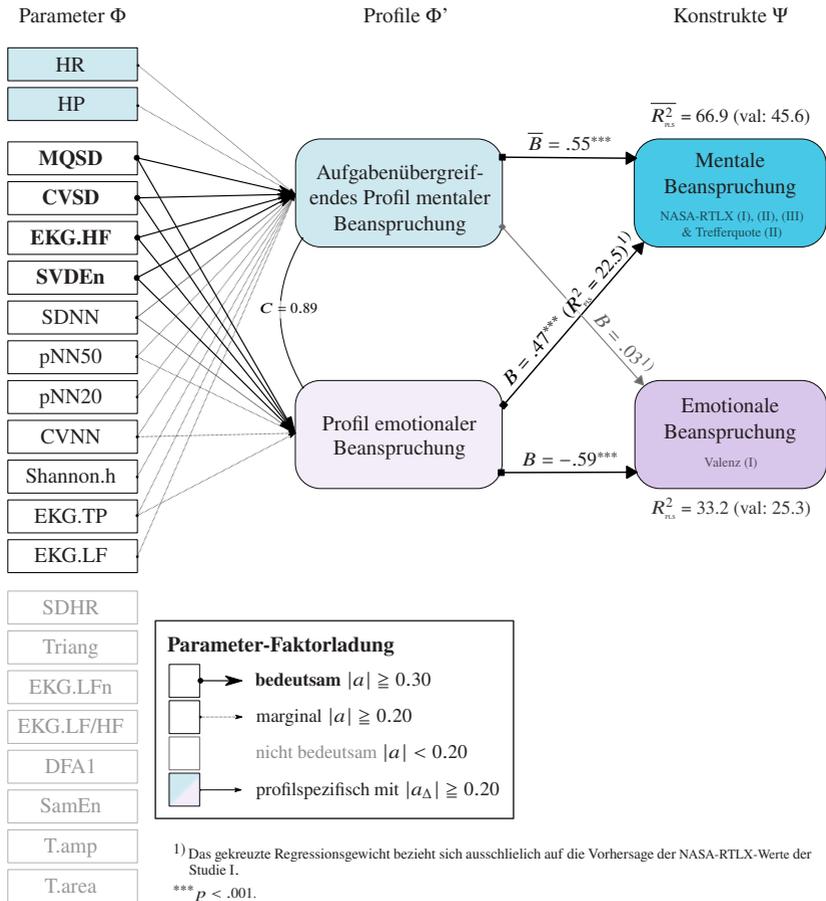


Abbildung 9.1: Strukturmodell für den Zusammenhang der kardialen Profile mit den Konstrukten der mentalen Beanspruchung (hellblau; repräsentiert durch das NASA-RTLX der Studien I bis III und durch die Trefferquote der Studie II) und emotionalen Beanspruchung (lila; repräsentiert durch die Emotionsdimension Valenz der Studie I). Die Höhe des Zusammenhangs ist am Regressionsgewicht B bzw. dessen Durchschnitt \overline{B} und am Anteil erklärter Varianz $R_{pl,s}^2$ bzw. dessen Durchschnitt \overline{R}_{ns}^2 ablesbar (val bezieht sich auf das kreuzvalidierte $R_{m,s}^2$). Die Profile sind mit den für sie bedeutsamen oder marginalen Parametern verbunden. Profilspezifische Parameter sind farblich dargestellt. C beschreibt die Kongruenz der Profile.

9.1.2 Bewertung kardialer Einzelparameter

Von den insgesamt 21 kardialen Parametern sind 12 Parameter relevant für die Vorhersage der mentalen Beanspruchung. Diese sind in Tabelle 9.1 und Abbildung 9.1 zusammengefasst. Von den 12 relevanten Parametern zeigen allerdings nur zwei Parameter eine Diagnostizität zwischen mentaler und emotionaler Beanspruchung. Für die Bewertung werden die Einzelparameter in relevante-unspezifische, relevant-spezifische und irrelevante Parameter unterteilt. Diese werden im Folgenden beschrieben.

Relevant-unspezifische Parameter

Das kardiale Profil der mentalen Beanspruchung zeigt bedeutsame Faktorladungen ($|a| \geq .30$) auf den Parametern MQSD, CVSD, EKG.HF und SVDEn (siehe Tabelle 9.1). Alle vier Parameter besitzen keine Profilspezifität und können somit nicht zwischen mentaler und emotionaler Beanspruchung differenzieren. Sie leisten jedoch den höchsten Beitrag zur Sensitivität der kardialen Profile mentaler und emotionaler Beanspruchung. Diese vier Parameter sind Maße der Herzratenvariabilität (HRV). Sie bilden die HRV jedoch in unterschiedlichen Domänen ab. So zählen die Parameter MQSD und CVSD zu den Maßen aus dem Zeitbereich, der Parameter EKG.HF zu den Maßen aus dem Frequenzbereich und der Parameter SVDEn zu den nichtlinearen Maßen.

Die Parameter MQSD und CVSD sind als redundant zu bewerten. Der Parameter CVSD wird aus dem Verhältnis der MQSD zu HP berechnet. Folglich liegt die Vermutung nahe, dass dieser Parameter vom Zusammenhang des MQSD mit dem kardialen Profil der mentalen Beanspruchung profitiert. Der Parameter MQSD wird als das mittlere Quadrat sukzessiver NN-Intervalle berechnet und berücksichtigt dadurch schnelle Veränderungen der Herzperiode stärker als andere zeitbezogene Maße der HRV. Auch der Parameter EKG.HF misst die schnellen Veränderungen der Herzperiode. Dieser gilt als ein Repräsentant der parasympathischen Kontrolle (Fahrenberg, 2001; Shaffer & Ginsberg, 2017) und

korrespondiert mit der respiratorischen Sinusarrhythmie. Die Ergebnisse der vorliegenden Studien zeigen auf, dass unter den sensitivsten HRV-Parametern vor allem diejenigen vertreten sind, welche die schnellen Anteile der HRV und somit verstärkt die parasymphatische Kontrolle repräsentieren. Gemäß den Ergebnissen nimmt diese mit Zunahme der mentalen und emotionalen Beanspruchung ab. Das Zusammenhangsmuster ist kongruent zu den bisherigen Ergebnissen aus der Literatur (z. B. Backs et al., 1999; Salahuddin et al., 2007; Sauvet et al., 2009; F. Schneider et al., 2017).

Der vierte bedeutsame Parameter (SVDEn) gibt die Regelmäßigkeit der HRV – durch die Zerlegung der HRV in Vektoren – an. Die Ergebnisse der vorliegenden Studien zeigen auf, dass die Regelmäßigkeit der HRV mit Zunahme der mentalen Beanspruchung steigt. Folglich werden bei einer höheren mentalen Beanspruchung weniger Vektoren benötigt, um die HRV zu erklären. Soweit bekannt, sind die vorliegenden Untersuchungen die Ersten, die einen Zusammenhang zwischen dem Parameter SVDEn und der mentalen bzw. emotionalen Beanspruchung belegen.

Weitere relevante Parameter mit einer marginaler Faktorladung von $|a| \geq .20$ sind SDNN, pNN50, pNN20 und CVNN (vgl. Tabelle 9.1). Sie alle zählen zu den HRV-Maßen aus dem Zeitbereich. Eine inhaltliche Ähnlichkeit zum Parameter MQSD ist für SDNN gegeben. SDNN berechnet aus den NN-Intervallen die Standardabweichung – wohingegen beim MQSD das mittlere Quadrat ermittelt wird. Auch der Parameter CVNN zeigt eine gewisse Redundanz, da dieser (analog zu CVSD) das Verhältnis zwischen SDNN und der HP abbildet. Aufgrund der Redundanz bei gleichzeitig geringerem Beitrag zur Sensitivität und Spezifität spielen die Parameter SDNN und CVNN nur eine zweitrangige Rolle für das kardiale Profil mentaler Beanspruchung.

Unter den Parametern mit marginal bedeutsamer Faktorladung finden sich Vertreter der geometrischen Maße (Shannon.h) und frequenzbezogenen Maße (EKG.TP und EKG.LF) der HRV. Der Parameter EKG.TP umfasst sowohl den Frequenzbereich der HF- als auch der LF-Komponente und bildet dahingehend

eine Redundanz zu beiden Parametern. Der Parameter EKG.LF zeigt im Gegensatz zur EKG.HF nur eine marginale Faktorladung auf dem Profil mentaler Beanspruchung. Dabei galt die LF-Komponente lange Zeit als ein sensibles Korrelat mentaler Beanspruchung. Diese Annahme ist aus heutiger Sicht anzuzweifeln (vgl. Fahrenberg, 2001) und wird durch die vorliegenden Befunde bestätigt. Allerdings könnte die Ursache für die marginale Faktorladung der EKG.LF auch in einer potentiellen verminderten Reliabilität liegen, da diese aus einer Ultra-Kurzzeitmessung berechnet wurde.

Bei allen bisher genannten Parametern findet sich eine hohe Überlappung in der Ladungsstruktur zwischen dem kardialen Profil mentaler Beanspruchung und dem kardialen Profil emotionaler Beanspruchung. Folglich sind die bisher genannten Parameter relevant für eine hohe Sensitivität der kardialen Profile, jedoch nicht für deren Diagnostizität untereinander. Diese Befunde bestätigen die bisherige Feststellung einer mangelnden Diagnostizität der HRV-Maße durch Manzey (1998). Gemäß Manzey (1998) gelten die HRV-Parameter als unspezifische Beanspruchungsindikatoren sowohl für die mentale, emotionale als auch körperliche Beanspruchung.

Relevant-spezifische Parameter

Eine profilspezifische Ladungsstruktur liegt lediglich für die marginal bedeutsamen Parameter HR und HP vor. Beide Parameter laden bedeutsam auf dem Profil mentaler Beanspruchung und unbedeutsam auf dem Profil emotionaler Beanspruchung. Kongruent zu den bisherigen Forschungsbefunden (z. B. Backs et al., 1999; Heine et al., 2017; Lenneman & Backs, 2007; Mukherjee et al., 2011; Moya-Albiol et al., 2001; Salahuddin et al., 2007) nimmt die Herzrate (HR) mit steigender mentalen Beanspruchung zu und dessen reziproker Kennwert, das Schlagintervall (HP), ab. Allerdings steht die vorgefundene Spezifität der HR und HP im Widerspruch zu den Schlussfolgerungen von Manzey (1998). Dieser geht von einer mangelnden Diagnostizität für diese Parameter aus. Manzey (1998) stellt in seiner Übersichtsarbeit fest, dass die

Parameter HR und HP nicht nur eine Reaktion auf mentale, sondern auch auf emotionale oder körperliche Belastungsfaktoren zeigen. Die von Manzey festgestellte, fehlende Diagnostizität könnte auf die Konfundierung zwischen mentaler, emotionaler und körperlicher Beanspruchung in den Studien aus der einschlägigen Literatur zurückzuführen sein. So stellt auch Manzey (1998) final fest, dass eine solche Konfundierung sich nur in kontrollierten Laborstudien ausschalten ließe.

Die Parameter HR und HP sind zwar durch Quotientenbildung ineinander überführbar, jedoch nicht redundant. Aufgrund der Nichtlinearität dieser Transformation ist es nach Fahrenberg (2001) nicht irrelevant, welches der beiden Maße bei der Analyse verwendet wird. Bspw. wird der HP eine biometrische Überlegenheit sowie eine höhere Eignung bei der Untersuchung kleiner Zeitintervalle zugesprochen (Fahrenberg, 2001). Diese Überlegenheit der HP (im Vergleich zur HR) zeigt sich auch in dessen höheren Faktorladung auf dem Profil mentaler Beanspruchung (vgl. Tabelle 9.1).

Irrelevante Parameter

Irrelevante Parameter für das Profil mentaler oder emotionaler Beanspruchung sind die Parameter SDHR, Triang, EKG.LFn, EKG.LF/HF, DFA1, SamEn, T.amp und T.area. Sie alle tragen weder zur Sensitivität noch zur Diagnostizität der Profile bei. Eine mögliche Ursache hierfür könnte die Berechnung dieser Parameter als Ultra-Kurzzeitmaß sein. So ist bspw. die Reliabilität für die geometrische Kenngröße Triang in einem Zeitbereich von 60 s nicht nachgewiesen (Nussinovitch et al., 2011; Salahuddin et al., 2007; Shaffer et al., 2016; zsf. Shaffer & Ginsberg, 2017). Auch der Parameter EKG.LF/HF – der lange Zeit als Indikator für das Gleichgewicht sympathischer und parasympathischer Einflüsse galt (z. B. Malik et al., 1996; Pagani et al., 1986; Shaffer & Ginsberg, 2017) – basierte in seiner ursprünglichen Berechnung auf einer Langzeitmessung. Zwar berichten einige Autoren (z. B. Salahuddin et al., 2007; Shaffer et al., 2016)

kürzere Zeitintervalle (von 20 s bis 180 s), jedoch fehlt ein fundierter Nachweis für die Reliabilität dieser Zeitintervalle.

Für die nichtlinearen Maße DFA1 und SampEn als auch für die morphologischen Maße T.amp und T.area besteht bis heute kein Konsens darüber, ob diese Parameter die mentale Beanspruchung valide abbilden können (vgl. Furedy et al., 1996; F. Schneider et al., 2017; Voss et al., 2009). Die vorliegenden Ergebnisse legen nahe, dass diese Parameter irrelevant für die Vorhersage der mentalen Beanspruchung sind.

9.1.3 Bewertung elektrodermalen Profile

Die extrahierten elektrodermalen Profile der Studien I, II und III sind zusammen mit den statistischen Kennwerten in Tabelle 9.2 zusammengefasst. Die Bewertung erfolgt anhand der zuvor definierten Gütekriterien.

Strukturvalidität

In den Studien I bis III konnten mehrere elektrodermale Profile mentaler Beanspruchung identifiziert werden: Das elektrodermale Profil des NASA-RTLX aus Studie I, II und III sowie das elektrodermale Profil der Trefferquote aus Studie II. Allerdings weichen die Parameter-Ladungsmuster der Profile des NASA-RTLX zwischen den Studien I, II und III voneinander ab. Aufgrund dieser fehlenden Replizierbarkeit der Profile über die Studien hinweg ist keine hinreichende Strukturvalidität gegeben. Von allen elektrodermalen Profilen mentaler Beanspruchung wird dem Profil aus Studie III die höchste Augen-scheinvalidität zugeschrieben. Der Grund hierfür liegt in den potentiellen Konfundierungsfaktoren der Studien I und II und wird im Folgenden erläutert.

Das kardiale Profil des NASA-RTLX aus Studie I wird als Repräsentant mentaler Beanspruchung verworfen. Zum einen ist dieses Profil mit $C \leq .82$ unterschiedlich zu den übrigen Profilen mentaler Beanspruchung. Zum anderen

Tabelle 9.2: Elektrodermale Profile des NASA-RTLX (Na), der Trefferquote (Tq), der Erregung (Er) und der Valenz (Va) aus den Studien I bis III zugeordnet zur Art der Beanspruchung. Zu sehen sind die Parameter-Faktorladungen α , deren Kongruenzkoeffizient C , die Bestimmtheitsmaße $\text{val}R_{\text{PLS}}^2$ und $\text{val}R_{\text{PLS}}^2$ sowie das Regressionsgewicht B .

Parameter	Mental			Emotional			-
	$C \geq .94$			$C \geq .97$			
	Na (III)	Na (II)	Tq (II)	Tq (I)	Er (I)	Va (I)	Na (I)
EDL.M	.11	.16	.17	.18	.12	.11	.36
EDL.D	.41	.26	.27	.28	.15	.19	.22
Phasic.SD	.31	.36	.37	.20	.29	.31	.25
EDR.N	.47	.33	.34	.11	.12	.07	.43
EDR.amp	.19	.32	.31	.37	.42	.42	.16
EDR.rise.t	.06	.06	.03	.08	-.02	.05	-.23
EDR.rec.t/2	-.11	-.15	-.13	-.01	-.07	.00	-.33
EDR.rec.t/c	-.07	-.19	-.18	-.03	-.07	.00	-.33
EDR.area	.13	.22	.21	.31	.31	.29	.06
EDR.rise.v	.16	.29	.29	.37	.44	.42	.25
EDR.peaks	-.18	-.21	-.20	-.02	-.06	-.03	-.36
EDA.LF	.38	.31	.31	.40	.40	.43	.12
EDA.HF	.38	.39	.40	.43	.39	.41	.21
EDA.VHF	.27	.28	.28	.33	.25	.25	.17
Prädiktormatrix							
R_{PLS}^2	7.64	6.85	6.79	8.61	8.91	7.91	11.67
$\text{val}R_{\text{PLS}}^2$	6.47	5.22	5.12	6.38	5.94	6.08	10.15
B	0.81	0.92	-0.81	-0.54	0.45	-0.53	0.67
Kriteriumsmatrix							
R_{PLS}^2	64.59	82.52	63.36	41.91	27.98	33.51	80.97
$\text{val}R_{\text{PLS}}^2$	37.73	51.2	37.17	10.32	8.16	4.18	52.5
Faktorladung ist bedeutsam $ a \geq .30$, marginal $ a \geq .20$, nicht bedeutsam $ a < .20$.							

wurde eine Reihe von Störfaktoren in Studie I identifiziert, die zur Extraktion eines abweichenden Profils beigetragen haben könnten. So ist der große Beanspruchungsunterschied in Studie I mit dem Grad an Bewegung, der Reihenfolge der Versuchsblöcke und dem Aufgabentyp (Balltrackingaufgabe versus IAPS-Bildersets) konfundiert. Zudem wurde in Studie I eine andere Messsensorik genutzt. Auch zwischen den Profilen mentaler Beanspruchung der Studien II und III liegt eine unzureichende Strukturvalidität vor. Das elektrodermale Profil des NASA-RTLX aus Studie III zeigt zwar mit $C = .94$ eine sehr hohe Ähnlichkeit zum elektrodermalen Profil der mentalen Beanspruchung aus Studie II, jedoch nicht die notwendige Identität. Das Profil der mentalen Beanspruchung aus Studie II wird durch die identischen Profile des NASA-RTLX und der Trefferquote ($C = 1.0$) repräsentiert. Eine mögliche Ursache für die mangelnde Strukturvalidität zwischen Studie II und III könnte darin liegen, dass das elektrodermale Profil mentaler Beanspruchung aus Studie II durch die emotionale Beanspruchung oder durch den spezifischen Aufgabentyp der Studie verzerrt wird. Diese Einschränkungen werden im Rahmen der Gütekriterien Generalisierbarkeit und Spezifität diskutiert. Im Gegensatz zum Profil des NASA-RTLX aus Studie I wird das Profil mentaler Beanspruchung aus Studie II weiterhin als ein Repräsentant mentaler Beanspruchung, jedoch mit den genannten Einschränkungen, interpretiert. Da das Profil des NASA-RTLX aus Studie I als Repräsentant der mentalen Beanspruchung ausgeschlossen wurde, wird im Folgenden auf eine Diskussion der Sensitivität, Generalisierbarkeit und Spezifität dieses Profils verzichtet.

Sensitivität

Wie bereits bei den kardialen Parametern, hängt auch bei den elektrodermalen Parametern nur ein geringer Varianzanteil mit der mentalen Beanspruchung zusammen. In der kardialen Prädiktormatrix werden weniger als 10 % der Varianz durch die elektrodermale Komponente aufgedeckt (siehe R_{pls}^2 der Prädiktormatrix in Tabelle 9.2). Über 90 % der Varianz sind somit auf unbekannte

Einflüsse zurückzuführen. Nichtsdestoweniger kann die elektrodermale Komponente 64.6 % (Originalmodell der Studie III) der Varianz in der mentalen Beanspruchung vorhersagen.

In Bezug auf die Sensitivität der subjektiven mentalen Beanspruchung schneidet das Profil des NASA-RTLX aus Studie II mit 82.5 % Varianzaufklärung (Originalmodell) am besten ab (vgl. Tabelle 9.2). Der geringere Anteil erklärter Varianz in Studie III (64.6 % Originalmodell) lässt sich durch die höhere Anzahl an Varianzquellen in der subjektiven mentalen Beanspruchung dieser Studie begründen. In Studie III wird die subjektive mentale Beanspruchung durch vier Schwierigkeitsstufen, fünf Aufgabentypen und deren Interaktion beeinflusst. Das elektrodermale Profil mentaler Beanspruchung aus Studie III bildet nur die Schwierigkeitsstufen erwartungskonform ab. Dies resultiert in dem geringeren Anteil erklärter Varianz der Studie III. Im Gegensatz dazu wird das NASA-RTLX in der Studie II lediglich durch zwei Schwierigkeitsstufen der Aufgabe moduliert. Die Trefferquote der Studie II wird ebenfalls durch eine zusätzliche Varianzquelle beeinflusst. Diese besteht aus der Interaktion zwischen Aufgabenschwierigkeit, visueller Komplexität und Bildmaterial. Da dieser Interaktionseffekt durch das elektrodermale Profil mentaler Beanspruchung aus Studie II nicht abgebildet wird, ist auch der Anteil erklärter Varianz für das elektrodermale Profil für die Trefferquote geringer (63.4 % Originalmodell) als für das NASA-RTLX (82.5 % Originalmodell).

Für die Profile mentaler Beanspruchung aus Studie II und III gilt, dass die Vorhersage der mentalen Beanspruchung für unbekannte Personen deutlich schlechter ausfällt (37.2 % bis 51.2 % Varianzaufklärung) als für Personen, die dem Vorhersagemodell bekannt sind (63.4 % bis 82.5 % Varianzaufklärung). Für eine bessere Vorhersage ist vor allem die Kenntnis über die personenspezifische Höhe der elektrodermalen Aktivität entscheidend.

Generalisierbarkeit

Das elektrodermale Profil aus Studie II wurde lediglich innerhalb eines spezifischen Aufgabentyps (Rechenaufgaben) extrahiert. Eine Konfundierung dieses Profils mit den spezifischen Aufgabenanforderungen der Studie II kann hier nicht ausgeschlossen werden. Sie könnte sogar eine Ursache für die mangelnde Strukturvalidität zwischen den Profilen der Studie II und III sein. Im Gegensatz dazu bildet das elektrodermale Profil aus Studie III die subjektive mentale Beanspruchung über verschiedene Aufgabentypen hinweg ab. Dies belegt, dass dieses Profil auf verschiedene Bereiche der Kognition angewendet werden kann. Wie bereits bei den kardialen Profilen wurde die Generalisierbarkeit des elektrodermalen Profils nur für kognitive Anforderungen mit geringen physischen Anforderungen geprüft.

Spezifität

Die Spezifität wird in Hinblick auf die Abgrenzung zur emotionalen Beanspruchung bewertet. Für den Überblick ist das Strukturmodell der kardialen Profile in Abbildung 9.2 dargestellt. Hier wird die mentale Beanspruchung durch das elektrodermale Profil des NASA-RTLX aus Studie III repräsentiert. Die emotionale Beanspruchung wird durch die identischen Profile der Valenz und Erregung aus Studie I repräsentiert ($C = .99$).⁵ Im kreuzvalidierten Modell ist die Sensitivität des Profils emotionaler Beanspruchung mit einer Varianzaufklärung von 4.2 % in der Valenz und 10.3 % in der Erregung als gering zu bewerten.⁶

⁵ Da die Profile emotionaler Beanspruchung nur in Studie I ermittelt wurden, kann keine Aussage bezüglich der Strukturvalidität und Generalisierbarkeit über Aufgaben und Situationen hinweg gemacht werden. Das elektrodermale Profil der Trefferquote ist ebenfalls identisch zu den Profilen der Erregung und der Valenz. Dieses wird bei der Bewertung als Repräsentant mentaler Beanspruchung ausgeklammert, da die Trefferquote kein Kriterium der emotionalen Beanspruchung, sondern der mentalen Beanspruchung darstellt.

⁶ Sind die Personen im Vorhersagemodell bekannt, erhöht sich der Anteil erklärter Varianz in der emotionalen Beanspruchung auf durchschnittlich 30.8 %. Diese Ergebnisse verdeutlichen, dass

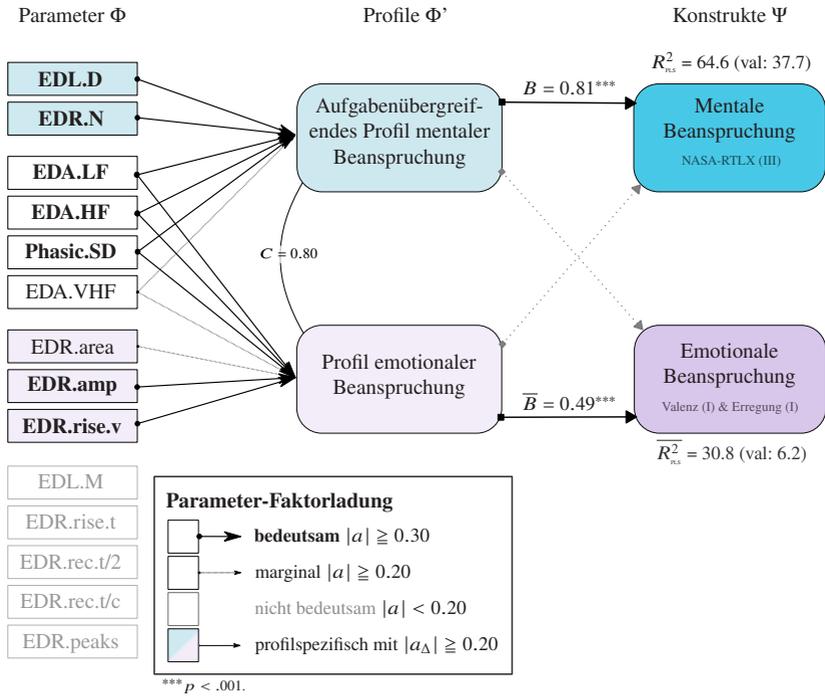


Abbildung 9.2: Strukturmodell für den Zusammenhang der elektrodermalen Profile mit den Konstrukten der mentalen Beanspruchung (hellblau; repräsentiert durch das NASA-RTLX der Studie III) und der emotionalen Beanspruchung (lila; repräsentiert durch die Emotionsdimensionen Erregung und (invertierte) Valenz der Studie I). Die Höhe des Zusammenhangs ist am Regressionsgewicht B bzw. dessen Durchschnitt \bar{B} und am Anteil erklärter Varianz R^2_{pls} bzw. dessen Durchschnitt \bar{R}^2_{ns} ablesbar (val bezieht sich auf das kreuzvalidierte R^2_{pls}). Die gekreuzten Regressionsgewichte (grau gestrichelte Pfeile) sind nicht bekannt. Die Profile sind mit den für sie bedeutsamen oder marginalen Parametern verbunden. Profilspezifische Parameter sind farblich dargestellt. C beschreibt die Kongruenz der Profile.

Das elektrodermale Profil der mentalen Beanspruchung aus Studie II besitzt keine Diagnostizität zum elektrodermalen Profil der emotionalen Beanspruchung

die Kenntnis der personenspezifischen Unterschiede für eine zufriedenstellende Vorhersage der emotionalen Beanspruchung unabdingbar ist. Im Gegensatz zum elektrodermalen Profil mentaler Beanspruchung, trägt beim Profil emotionaler Beanspruchung nicht nur die personenspezifische Höhe der elektrodermalen Aktivität, sondern auch die personenspezifische Bewertung der Erregung und Valenz zur Vorhersagegüte bei.

(aus Studie I). Beide Profile sind als ähnlich zu interpretieren ($C = .89$). Diese Ähnlichkeit könnte darauf zurückzuführen sein, dass die aufgabenspezifischen Belastungsfaktoren in Studie II mit der emotionalen Beanspruchung konfundiert sind. Im Gegensatz dazu wurden in der Studie III mehrere Aufgabentypen operationalisiert. Da eine Konfundierung mit der emotionalen Beanspruchung nicht für alle Aufgabentypen zu erwarten ist, mittelt sich dieser Störeinfluss eher aus. Entsprechend unterscheidet sich das elektrodermale Profil aus Studie III zum Profil emotionaler Beanspruchung (aus Studie I) mit $C = .80$. In Abbildung 9.2 ist zu erkennen, dass EDL.D und EDR.N profilspezifische und bedeutsame Parameter des Profils mentaler Beanspruchung darstellen. Demgegenüber besitzen die Parameter EDR.area, EDR.amp und EDR.rise.v eine profilspezifische und bedeutsame Faktorladung auf dem elektrodermalen Profil der emotionalen Beanspruchung. Die restlichen Parameter erlauben keine Differenzierung der Profile und laden entweder auf beiden oder auf keinem Profil. Ob eine Vorhersage der emotionalen Beanspruchung durch das elektrodermale Profil der mentalen Beanspruchung (und umgekehrt) möglich ist, kann bedingt durch die Unterschiede zwischen den Studien I und III nicht beantwortet werden.

Einordnung in die Taxonomie psychophysiologischer Zusammenhänge

In der vorliegenden Arbeit konnten die elektrodermalen Einzelparameter Φ erfolgreich zu elektrodermalen Profilen Φ' umgewandelt werden (vgl. Abbildung 9.2). Dadurch konnte der ursprüngliche Viele-zu-viele-Zusammenhang zwischen den elektrodermalen Parametern und der mentalen Beanspruchung zu einem Eins-zu-eins-Zusammenhang⁷ umgewandelt werden. Dabei ist das elektrodermale Profil mentaler Beanspruchung aus Studie III den übrigen Profilen überlegen und daher als der beste Repräsentant der mentalen Beanspruchung zu bewerten. Insgesamt zeigt das elektrodermale Profil mentaler Beanspruchung

⁷ Im Rahmen der in dieser Arbeit untersuchten Konstrukte.

aus Studie III eine hinreichende Sensitivität, Generalisierbarkeit und Diagnostizität zur emotionalen Befindlichkeit. Lediglich die Strukturvalidität bedarf noch weiterer Schärfung. Die schwache Strukturvalidität über die unterschiedlichen Studien hinweg zeigt die Notwendigkeit, die Wirkung der gefundenen Störfaktoren – Bewegung, Aufgabentyp und emotionale Beanspruchung – auf die Extraktion elektrodermalen Profil zu untersuchen. Nach der Taxonomie psychophysiologischer Zusammenhänge (Cacioppo & Tassinary, 1990; Cacioppo et al., 2017) ist das elektrodermale Profil mentaler Beanspruchung (aus Studie III) mindestens als ein Marker (situationsspezifischer Eins-zu-eins-Zusammenhang) der mentalen Beanspruchung einzuordnen.

9.1.4 Bewertung elektrodermalen Einzelparameter

Aus dem bisherigen Forschungsstand wurde eine geringe Diagnostizität elektrodermalen Einzelparameter gefolgert. Im Gegensatz dazu zeigen die Befunde der vorliegenden Arbeit, dass elektrodermale Parameter im Verbund eines Profils Spezifität besitzen: Die Parameter EDL.D und EDR.N laden spezifisch und bedeutsam auf dem Profil mentaler Beanspruchung, die Parameter EDR.area, EDR.amp und EDR.rise.v hingegen spezifisch und bedeutsam auf dem Profil emotionaler Beanspruchung (siehe Abbildung 9.2). Dadurch bestätigt sich die Auffassung von Dawson et al. (2017), dass elektrodermale Parameter unabhängige Informationsquellen darstellen. Für die Bewertung werden die Einzelparameter in relevant-unspezifische, relevant-spezifische und irrelevante Parameter unterteilt. Diese werden im Folgenden beschrieben.

Relevant-spezifische Parameter

Zu den profilspezifischen Parametern der mentalen Beanspruchung zählen der Parameter EDR.N, definiert als Anzahl der NS-EDRs, und EDL.D, definiert als die Veränderung des tonischen Hautleitwerts. Beide Parameter zeigen einen Anstieg mit steigender mentaler Beanspruchung. Diese Ergebnisse sind konsistent

mit den bisherigen Befunden. In der Literatur wird eine Zunahme in der Anzahl der NS-EDRs sowie ein bedeutsamer Anstieg des tonischen Hautleitwerts mit steigender mentaler Beanspruchung berichtet (z. B. Chanel et al., 2011; Collet et al., 2014; Engström et al., 2005; Kohlisch & Schaefer, 1996; Mehler et al., 2012; Moya-Albiol et al., 2001; Nourbakhsh et al., 2017; Setz et al., 2010; Visnovcova et al., 2016; Widyanti et al., 2017). Für das Profil mentaler Beanspruchung ist jedoch nicht der Anstieg des tonischen Hautleitwerts zwischen den Aufgaben (repräsentiert durch EDL.M), sondern innerhalb der Aufgaben (repräsentiert durch EDL.D) relevant. Die fehlende Sensitivität des Parameters EDL.M könnte darauf zurückzuführen sein, dass die versuchsbedingten Pausen zwischen den einzelnen Schwierigkeitsstufen der Experimentalaufgaben zu einer Rückbildung des mittleren tonischen Hautleitwerts geführt haben. Die Parameter EDL.D und EDR.N zählen zu den Maßen der tonischen Veränderung. Sie sind allerdings nicht redundant zueinander: EDL.D repräsentiert die tonischen Veränderungen abgeleitet aus der tonischen EDA-Komponente. EDR.N repräsentiert hingegen die tonischen Veränderungen abgeleitet aus der phasischen EDA-Komponente.

Zu den profilspezifischen Parametern der emotionalen Beanspruchung zählen EDR.area, EDR.amp und EDR.rise.v. Sie alle nehmen mit Zunahme der emotionalen Beanspruchung (Steigerung der Erregung oder Verringerung der Valenz) zu. Diese Ergebnisse sind konform mit den einschlägigen Befunden aus der Literatur (siehe Nourbakhsh et al., 2017; Traxel, 1960). Die Parameter EDR.area und EDR.rise.v sind dahingehend redundant zum Parameter EDR.amp, dass beide Parameter die EDR.amp zur Berechnung ihres Kennwerts – unter Berücksichtigung der Zeit – nutzen (siehe Unterkapitel 5.2.2). Somit wird das Profil der emotionalen Beanspruchung vor allem durch diejenigen Parameter repräsentiert, welche aus der Amplitude der NS-EDR abgeleitet werden.

Relevant-unspezifische Parameter

Die Parameter EDA.LF und EDA.HF und EDA.VHF laden bedeutsam oder zumindest marginal auf dem Profil mentaler als auch emotionaler Beanspru-

chung. Folglich sind diese nur für eine höhere Sensitivität, jedoch nicht für die Diagnostizität, der Profile zu berücksichtigen. Alle drei Parameter bilden Maße aus dem Frequenzbereich ab. Sie nehmen konform zu den Befunden aus der Literatur mit zunehmender Beanspruchung zu (vgl. Posada-Quintero et al., 2016a, 2016b). Im Gegensatz zu den frequenzbezogenen Maßen der kardialen Aktivität sind die frequenzbezogenen Maße der elektrodermalen Aktivität keine etablierten Parameter. Folglich gibt es für diese keine eindeutigen Richtlinien, welche Frequenzbereiche sinnvoll unterteilt werden können. Bspw. kombinieren Posada-Quintero et al. (2016a) die LF-Komponente mit Frequenzbereichen aus der HF-Komponente (0.08 Hz bis 0.25 Hz) und berechnen daraus ein Maß der sympathischen EDA-Aktivität. Aus diesem Grund kann für die vorliegenden Studien eine Redundanz der frequenzbezogenen Parameter nicht ausgeschlossen werden.

Ein weiterer relevanter, jedoch unspezifischer Parameter ist Phasic.SD. Dieser ist ebenfalls kein etabliertes Maß mentaler oder emotionaler Beanspruchung. Phasic.SD ist definiert als die Standardabweichung der phasischen EDA-Komponente. Diese nimmt mit steigender mentaler aber auch emotionaler Beanspruchung zu. Soweit bekannt, sind die Ergebnisse der vorliegenden Arbeit die Ersten, die einen Zusammenhang dieses Parameters mit der mentalen und emotionalen Beanspruchung belegen. Für Phasic.SD besteht eine potentielle Redundanz zu denjenigen Parametern, welche auf Basis der NS-EDRs berechnet werden (EDR.N, EDR.area, EDR.amp, EDR.rise.v). Denn sowohl Phasic.SD als auch die NS-EDRs werden auf Basis der phasischen EDA-Komponente ermittelt. Mit steigender NS-EDR-Anzahl und NS-EDR-Amplitude steigt auch die Standardabweichung der phasischen EDA-Komponente.

Irrelevante Parameter

Zu den irrelevanten Parametern zählen die zeitbezogenen Kenngrößen für den Anstieg und Abfall der NS-EDRs: EDR.rise.t, EDR.rec.t/2 und EDR.rec.t/c. Diese Parameter besitzen weder bedeutsame noch profilspezifische Faktorla-

dungen auf dem Profil mentaler Beanspruchung oder dem Profil emotionaler Beanspruchung. Gleiches gilt für den Parameter EDR.peaks, welcher die mittlere Zeit zwischen den NS-EDRs abbildet. Es ist allerdings anzumerken, dass all diese zeitbezogenen Parameter bedeutsame Faktorladungen auf dem elektrodermalen Profil des NASA-RTLX aus Studie I besitzen (siehe Tabelle 9.1). Durch die vorliegenden Störgrößen in Studie I ist es allerdings nicht möglich, dieses Profil valide einem Konstrukt zuzuordnen. Jedoch scheinen genau diese zeitbezogenen Parameter bedeutsam auf diesem unbekanntem Konstrukt zu laden. Die Faktorladungen dieser zeitbezogenen Parameter indizieren, dass die Zeiten für den Anstieg, den Abfall und den Abstand der EDRs verringert wird, je höher das unbekannte Konstrukt ist. Ein weiterer Parameter, der nur auf dem elektrodermalen Profil des NASA-RTLX aus Studie I lädt, ist der mittlere Hautleitwert der tonischen EDA-Komponente. Dieser nimmt mit Zunahme des unbekanntem Konstrukts aus Studie I zu.

9.2 Theoretische Verortung der extrahierten Profile

Basierend auf dem Rahmenmodell von Cacioppo et al. (Cacioppo & Tassinary, 1990; Cacioppo et al., 2017) konnten in dieser Arbeit die physiologischen Einzelparameter zu bedeutsamen Profilen mentaler und emotionaler Beanspruchung umstrukturiert werden. Das angewandte Rahmenmodell bietet jedoch keine Erklärung dafür, welche psychophysiologischen Wirkmechanismen für die Ausbildung der physiologischen Profile verantwortlich sind. Um die gefundenen Profile theoretisch erklären zu können, werden die bisherigen Erklärungsansätze mentaler Beanspruchung diskutiert.

Eine erste Erklärung für die Ausbildung physiologischer Profile bieten die Steuermechanismen des autonomen Nervensystems – bestehend aus dem Sympathikus und Parasympathikus. Dieser Ansatz ist allerdings nicht dazu geeignet, um die Wirkmechanismen der elektrodermalen Profile aufzuklären. Schließlich wird die EDA rein sympathisch gesteuert. Der Ansatz der autonomen Steuerung

eignet sich vor allem für die kardialen Profile und wird auch von Backs (1995, 1998) im Rahmen seiner psychophysiologischen Studien herangezogen. In diesen Studien extrahiert Backs aus unterschiedlichen kardialen Parametern (des EKG und der Impedanzkardiographie) zwei Komponenten, die er als Repräsentanten der sympathischen und parasympathischen Aktivität interpretiert. Dabei ordnet Backs (1998) den Parameter HP der sympathischen Aktivität und die Parameter EKG.HF und EKG.LF der parasympathischen Aktivität zu. Vergleicht man die Ergebnisse dieser Arbeit mit denen von Backs (1998), so ergibt sich ein wesentlicher Unterschied⁸: In den vorliegenden Studien konnte stets nur eine bedeutsame kardiale Komponente aus den kardialen Parametern extrahiert werden, während sich bei Backs zwei Komponenten zeigen. Dieser Umstand ist auf die unterschiedlichen Extraktionsmethoden zurückzuführen. Backs (1995, 1998) wendet in seinen Studien die PCA an. In der vorliegenden Arbeit wird hingegen die NPLS-Regression angewandt. Bei der NPLS-Regression wird die gemessene mentale Beanspruchung als abhängige Variable zur Komponentenextraktion herangezogen. Das Ziel dabei ist es, den Zusammenhang der Komponente mit der mentalen Beanspruchung zu maximieren. Dass auf diese Weise valide Repräsentanten sympathischer oder parasympathischer Kontrolle extrahiert werden, ist unwahrscheinlich. Naheliegender ist die Extraktion eines Profils, welches simultan die sympathische und parasympathische Kontrolle abbildet. Diese Annahme wird dadurch gestützt, dass das kardiale Profil mentaler Beanspruchung simultan durch den von Backs (1998) definierten parasympathischen Parameter EKG.HF und den sympathischen Parameter HP repräsentiert wird.

Moderne Ansätze der autonomen Steuerung, wie die allodynamische Steuerung, sind für die Erklärung der Profile den zuvor beschriebenen Wirkmechanismen überlegen. Bspw. könnten die verschiedenen Formen der allodynamischen

⁸ Backs (1998) unterscheidet in seiner Studie zwischen Parametern, die um personenspezifische Effekte bereinigt wurden und Parametern, die nicht bereinigt wurden. Die Ergebnisse der vorliegenden Arbeit werden nur mit den Befunden bereinigter Parametern verglichen, da diese der angewandten Zentrierung und Normierung der NPLS-Regression am ähnlichsten sind.

Steuerung hinter der Ausprägung der physiologischen Profile stehen (siehe Berntson et al., 1991, 1994, 2007; zsf. Kapitel 3.1). Für das kardiale Profil der mentalen Beanspruchung wäre der autonome Antagonismus (Aktivierung des Sympathikus bei gleichzeitiger Deaktivierung des Parasympathikus) anzunehmen, da dieses Profil gleichzeitig durch den – von Bacs (1998) definierten – sympathischen Parameter HP und den parasympathischen Parameter EKG.HF repräsentiert wird. Das kardiale Profil der emotionalen Beanspruchung wäre hingegen der autonomen Entkopplung (alleinige Aktivierung des Parasympathikus) zuzuschreiben, da dieses Profil nur einen Zusammenhang mit dem parasympathischen Parameter EKG.HF zeigt.

Weitere Erklärungsansätze finden sich in den aktivierungstheoretischen Modellen (siehe Unterkapitel 2.2.2). Bspw. ist das Konzept der zentralen Aktivierung von Bartenwerfer (1969) auf das kardiale Profil emotionaler Beanspruchung anwendbar. Dieses Profil besitzt keine Diagnostizität und sagt sowohl die emotionale als auch die mentale Beanspruchung bedeutsam vorher. Die fehlende Spezifität könnte ein Hinweis darauf sein, dass dieses Profil die zentrale (unspezifische) Erregung repräsentiert. Der eindimensionale Ansatz von Bartenwerfer (1969) bietet jedoch wenig Erklärungsgrundlage für die Wirkmechanismen der anderen physiologischen Profile. Aus diesem Grund werden mehrdimensionale Aktivierungsmodelle als Erklärung herangezogen. Relevant ist hierbei das 4-Arousal-Modell von Bacs und Boucsein (2009) und Boucsein (2012, S. 353–362). In diesem Modell nehmen die Forscher eine Zuordnung der physiologischen Kenngrößen zu den vier Aktivierungssystemen des Modells vor. Diese Kenngrößen fungieren als Repräsentanten des jeweiligen Aktivierungssystems (siehe Tabelle 2.1 in Unterkapitel 2.2.2). Diese physiologischen Repräsentanten können dazu genutzt werden, um die extrahierten Profile dieser Arbeit zu den vier Aktivierungssystemen zuzuweisen. Eine solche Herangehensweise führt allerdings zu unplausiblen Ergebnissen: Bspw. schreiben die Forscher einen Anstieg der HR den Aktivierungssystemen *Vorbereitende-Aktivierung* und *allgemeine Erregung* zu. Im Gegensatz dazu gilt in der vorliegenden Arbeit ein Anstieg der HR als spezifisch für das Profil der mentalen Bean-

spruchung. Plausibel wäre eigentlich eine Zuordnung des Profils mentaler Beanspruchung – und somit des Parameters HR – zum *Anstrengungs-System*. Ein ähnliches Ergebnis zeigt sich für die elektrodermalen Profile: Hier würde das elektrodermale Profil mentaler Beanspruchung dem *Affekt-Erregungs-System* zugeordnet werden, da beide durch den Anstieg der NS-EDR-Anzahl repräsentiert werden. Das elektrodermale Profil der emotionalen Beanspruchung wäre hingegen dem *Vorbereitenden-Aktivierungs-System* zuzuordnen, da beide durch die NS-EDR-Amplitude vertreten werden. Bei einer plausiblen Zuordnung der physiologischen Profile zu den Aktivierungssystemen wäre zu erwarten, dass die Profile mentaler Beanspruchung im *Anstrengungs-System* und die Profile emotionaler Beanspruchung im *Affekt-Erregungs-System* verortet sind. Die Ursache für die fehlende Plausibilität ist vor allem in der mangelnden Validität der Parameterzuordnung von Backs und Boucsein (2009) und Boucsein (2012, S. 353-362) zu vermuten. Die Forscher nehmen eine oberflächliche Zuordnung weniger Parameter und lediglich auf Basis einer Literaturrecherche vor. Überlappungen der Parameter für verschiedene Aktivierungssysteme werden kaum berücksichtigt. Darüber hinaus sind Messgrößen, wie die HRV, nicht spezifiziert. Dadurch ist es nicht ersichtlich, welche Parameter konkret unter die HRV fallen. Eine empirische Validierung der von Backs und Boucsein (2009) und Boucsein (2012, S. 353–362) vorgenommenen Parameterzuordnung ist nicht bekannt. Alles in allem ist die Parameterzuweisung zu den Aktivierungssystemen als Grundwahrheit anzuzweifeln. Nichtsdestoweniger bieten die vier Aktivierungssysteme einen vielversprechenden Erklärungsansatz für potentielle Wirkmechanismen hinter den extrahierten Profilen. Zukünftige Forschungsarbeiten sollten sich der Forschungsfrage widmen, ob die vier Aktivierungssysteme durch jeweilige physiologische Profile repräsentiert werden können. Auf diese Weise kann eine empirische Basis für die Zuordnung der Parameter zu den Aktivierungssystemen geschaffen werden.

9.3 Anwendungsmöglichkeiten der physiologischen Profile in der Mensch-Maschine-Schnittstelle

Die Entwicklung technischer Systeme hin zu intelligenten Unterstützungssystemen ist an die Voraussetzung geknüpft, dass die definierten Nutzerzustände durch die Schnittstelle gemessen und korrekt interpretiert werden. Übertragen auf die beanspruchungssensitive Mensch-Maschine-Schnittstelle bedeutet dies, dass die mentale Beanspruchung valide aus der gemessenen physiologischen Aktivität vorherzusagen ist. Vielversprechende multivariate Methoden zur Prädiktion der mentalen Beanspruchung wurden bereits in Abschnitt 3.4.1 aus dem Bereich des Psychophysiological Computing vorgestellt. Hierbei werden überwiegend Klassifikationsverfahren angewandt und der Fokus auf die Maximierung der Sensitivität gelegt. Zur Maximierung der Sensitivität greifen viele Forscher und Entwickler auf personenspezifischen Vorhersagemodelle zurück. Dabei wird die mentale Beanspruchung mit einer personenspezifischen Parameterkombination vorhergesagt, ohne die Personenspezifität theoretisch zu begründen. Als Forschungslücke des Psychophysiological Computing wurde bereits der fehlende Validitätsnachweis der Klassifikationsergebnisse diskutiert (siehe Abschnitt 3.4.2). Für einen Validitätsnachweis müsste zum einen sichergestellt werden, dass sich dieselbe Parameterkombination über verschiedene Studien und Aufgaben hinweg zeigt. Zum anderen wäre zu belegen, dass die vorgefundene Parameterkombination tatsächlich mit der mentalen Beanspruchung zusammenhängt und sich nicht lediglich aus einer Scheinkorrelation ergibt. Für individuelle Vorhersagemodelle ist so ein Validitätsnachweis mit einem unverhältnismäßigen Aufwand verbunden, da der Nachweis für jede einzelne Person und dessen Parameterkombination zu erbringen ist.

Die oben aufgeführten Lücken können mit der mehrdimensionalen Analyse-methode sowie den daraus abgeleiteten Profilen dieser Arbeit geschlossen werden. Durch die Profilextraktion wurde eine Grundlage für die Validität

physiologischer Profile geschaffen. Zudem konnten in der vorliegenden Arbeit relevante Parameter zur Vorhersage der mentalen Beanspruchung identifiziert werden. Ein weiterer Vorteil ist, dass die mehrdimensionale Methode nicht nur personenübergreifende sondern auch personenspezifische Informationen für die Vorhersage berücksichtigt. Im Gegensatz zu den personenspezifischen Vorhersagemodellen sind die physiologischen Profile der dreidimensionalen Methode flexibel: Dadurch kann die mentale Beanspruchung unbekannter Personen oder unbekannter Situationen in einem Mensch-Maschine-System vorhergesagt werden. Die mentale Beanspruchung einer unbekannt Person wird durch

$$\hat{y} = \mathbf{X} \cdot (\mathbf{T} \otimes \mathbf{Z})^T \quad (9.1)$$

geschätzt. Dabei steht \hat{y} für den vorhergesagten Beanspruchungswert der Person, \mathbf{Z} für die Ladungsmatrix der Parameter, \mathbf{T} für die Ladungsmatrix der Situationen bzw. Bedingungen und \mathbf{X} für die entfaltete Matrix der gemessenen Parameter über verschiedene Situationen hinweg. Dasselbe Vorhersagemodell kann dazu verwendet werden, um die mentale Beanspruchung einer unbekannt Situation – unter Zuhilfenahme der Parameter-Faktorladungen \mathbf{Z} und Person-Faktorladungen \mathbf{W} – vorherzusagen. Dies beschreibt die Gleichung

$$\hat{y} = \mathbf{X} \cdot (\mathbf{W} \otimes \mathbf{Z})^T \quad (9.2)$$

mit \mathbf{X} als die entfaltete Matrix der gemessenen Parameter über verschiedene Personen hinweg. Durch das Regressionsgewicht B der Bedingungsdimension wird die Vorhersage der mentalen Beanspruchung auf

$$\hat{y} = B \cdot (\mathbf{X} \cdot (\mathbf{W} \otimes \mathbf{Z})^T) \quad (9.3)$$

erweitert. Die Schätzung unbekannter Situationen in Gleichung 9.2 und 9.3 impliziert, dass die Person-Faktorladungen bekannt sind. In einem Mensch-Maschine-System würde das eine initiale Kalibrierung des maschinellen Systems auf die Personen erfordern, um die jeweilige Person-Faktorladungen zu

berechnen und im System zu hinterlegen. Für die Kalibrierung kann eine Standardaufgabe – bspw. mit zwei Schwierigkeitsstufen – genutzt werden. Die Kalibrierung eines beanspruchungssensitiven Mensch-Maschine-Systems auf die jeweilige Person wird bspw. in der Forschungsarbeit von M. Schneider (2017) umgesetzt. Allerdings verwendet M. Schneider (2017) personenspezifische Vorhersagemodelle (bzw. personenspezifische Parameter) zur Schätzung der mentalen Beanspruchung. Die mehrdimensionale Methode der vorliegenden Arbeit eignet sich gleichfalls, um eine Kalibrierung der Person vorzunehmen. Gleichzeitig bietet die hier aufgeführte mehrdimensionale Methode einen Mehrwert, indem sie die Validitätseinschränkungen der individuellen Vorhersagemodelle überwindet.

Die Ermittlung der mentalen Beanspruchung mit Hilfe physiologischer Profile kann in diversen Unterstützungssystemen gewinnbringend eingesetzt werden. Ein Anwendungsbeispiel aus dem Arbeitskontext zeigt das Forschungsprojekt KonsensOP (Kontextsensitive Assistenz im aufmerksamen OP; Stiefelhagen, Martinez, Wolf, Philipp & Marzi, 2019; Peinsipp-Byma & Voit, 2019; Schwarz, Mack, Reisig, Gundling & Schuette, 2019). Das Ziel dieses Projekts besteht in der Entwicklung eines Assistenzsystems für den Operationsaal. Dabei wird der Operateur bei schwierigen Arbeitsschritten oder auftretenden Komplikationen durch die Präsentation von Hinweisen, Anleitungen oder Informationen aus der Operationsplanung auf dem Bildschirm unterstützt. Da die Durchführung einer Operation bereits eine anspruchsvolle Aufgabe darstellt, würde die Bedienung eines Assistenzsystems das Operationsteam zusätzlich belasten. Im Gegensatz dazu würde ein dauerhaftes Unterstützungsangebot das Operationsteam unnötig ablenken. Die Erfassung der mentalen Beanspruchung durch die physiologische Mensch-Maschine-Schnittstelle des Assistenzsystems ermöglicht es, dem Operateur genau dann eine Unterstützung bereitzustellen, wenn eine zunehmende Auslastung der kognitiven Ressourcen gemessen wird. In solchen Anwendungsszenarien ist es wichtig, dass die Beanspruchungsvorhersage über verschiedene Arbeitsaufgaben hinweg Gültigkeit besitzt und von anderen Nutzerzuständen, darunter der emotionalen Beanspruchung, abgegrenzt ist. Die

extrahierten physiologischen Profile können für diesen Zweck eingesetzt werden. Die Unterscheidung mentaler und emotionaler Beanspruchung ist wichtig, da ansonsten irrelevante Emotionen eine unangemessene Unterstützungsreaktion auslösen oder eine notwendige Unterstützung verhindern würden. Im Idealfall ist die Unterstützung an die Art der Beanspruchung (mental oder emotional) anzupassen. Bspw. könnte eine unerwartete Komplikation sowohl eine negative Emotion als auch eine hohe mentale Beanspruchung auslösen. Demgegenüber ist das alleinige Auftreten einer negativen Emotion auf aufgabenunabhängige Ursachen (z. B. zwischenmenschliche Konflikte) zurückzuführen.

Da die meisten Operationen einem standardisierten Ablauf folgen, können die gemessenen Daten vorangegangener Operationen dazu genutzt werden, um die Vorhersage der mentalen Beanspruchung in zukünftigen Operationen zu verbessern. Das zugrundeliegende Prinzip ist mit einer automatischen Kalibrierung des Mensch-Maschine-Systems vergleichbar. Hierbei können sowohl die Personen als auch die Situationen kalibriert werden, indem sie mit bereits erfassten Personen und Situationen in Zusammenhang gesetzt werden. Das Grundprinzip eines solchen Ansatzes wird im Folgenden skizziert. Beim ersten Einsatz des Assistenzsystems ist davon auszugehen, dass personenspezifische oder situationsspezifische Daten (also \mathbf{W} oder \mathbf{T}) nicht vorhanden sind. In diesem Fall stehen für die Vorhersage einer unbekannt Person in einer unbekannt Situation nur die Parameter-Ladungsmatrix \mathbf{Z} und die gemessenen Parameterwerte \mathbf{X} zur Verfügung. Der vorhergesagte Beanspruchungswert \hat{y} ergibt sich in diesem Fall aus:

$$\hat{y} = \mathbf{X} \cdot \mathbf{Z}^T. \quad (9.4)$$

Unter der Annahme, dass mehr als eine Operation mit einem solchen Unterstützungssystem durchgeführt wurden, können die vorangegangenen physiologischen Daten der Operateure zur Verbesserung der Vorhersage genutzt werden. Dafür ist es zunächst notwendig, die gemessenen Daten den jeweiligen Arbeitsschritten zuzuweisen. Dadurch wird eine dreidimensionale Datenstruktur

(Person \times Situation \times Parameter) generiert.⁹ Mit der parallelen Faktorenanalyse (PARAFAC; siehe Abschnitt 5.6.2) können die Ladungsmatrizen der Personen (**W**) und Situationen (**T**) gewonnen werden.¹⁰ Die Person- und Situations-Faktorladungen erlauben es nun, die mentale Beanspruchung neuer Personen oder Situationen entsprechend der Gleichungen 9.1 bis 9.3 zu bestimmen. Würde man also einen bereits im System gespeicherten Arbeitsschritt in einer Operation mit einer neuen Person durchführen, dann wäre eine präzisere Schätzung der mentalen Beanspruchung durch die Gewichtung mit der situationsspezifischen Faktorladung möglich. Eine solche Verbesserung wird beispielhaft in Anhang A.9 veranschaulicht. Dasselbe Prinzip gilt für die Personen. Dieser Ansatz verdeutlicht, wie ein Mensch-Maschine-System auch ohne eine manuelle Kalibrierung auskommt. Ein solches System würde sich durch einen anwachsenden Datensatz selbst kalibrieren und verbessern.

Weitere Anwendungsbereiche für eine beanspruchungssensitive Mensch-Maschine-Schnittstelle auf Basis physiologischer Profile sind Tätigkeiten, bei denen eine kontinuierliche Überwachung von technischen Systemen durch den Mensch notwendig ist. Hierzu zählen u. a. Arbeitssysteme von Piloten, Fluglotsen, Astronauten oder Leitwarteoperatoren. Die Messung der mentalen Beanspruchung könnte hier Unterstützung in hoch beanspruchenden Situationen bereitstellen und dadurch zur Vermeidung von menschlichen Fehlern führen. Im Bereich der computerbasierten Lernprogramme könnte eine beanspruchungssensitive Mensch-Maschine-Schnittstelle dazu genutzt werden, um den Lerninhalt an die Auslastung der kognitiven Informationsverarbeitung der Person anzupassen. Aber auch hilfsbedürftige Menschen (bspw.

⁹ Die Personen sind in diesem Fall die bisher erfassten Operateure. Die Situationen werden durch die Arbeitsschritte abgebildet. Die Parameter werden durch das Mensch-Maschine-System erfasst.

¹⁰ Mit Hilfe der PARAFAC wird die dreidimensionale Datenmatrix in Komponenten und deren Faktorladungen zerlegt. Hierbei ist es wichtig, die Faktorladungen der Parameter – auf die in dieser Arbeit definierten Ladungswerte des mentalen Beanspruchungsprofils – zu fixieren. Die Fixierung stellt sicher, dass die Komponente (bzw. Ladungsstruktur) der mentalen Beanspruchung extrahiert wird.

Demenzkranke) könnten durch ein beanspruchungssensitives Assistenzsystem an Selbstständigkeit und somit Lebensqualität gewinnen.

9.4 Ausblick

Aus den erarbeiteten Forschungsergebnissen dieser Arbeit können mehrere Ideen und Fragen für zukünftige Forschungsbestrebungen abgeleitet werden. Diese werden im Folgenden beleuchtet. Forschungsbedarf ergibt sich in Bezug auf die Frage, ob sich eine hohe physische Aktivität auf die Validität der kardialen und elektrodermalen Profile auswirkt. Einflüsse hoher körperlicher Aktivität auf kardiale Parameter, wie den MQSD, werden bspw. von Wagner (2013) berichtet. Da die physiologischen Profile dieser Arbeit im Bereich geringer bis mittlerer körperlicher Aktivität abgeleitet wurden, ist die Gültigkeit der Profile bei hoher körperlicher Aktivität in zukünftigen Forschungsarbeiten zu klären.

In zukünftigen Arbeiten können neue physiologische Parameter sowie neue psychologische Konstrukte (bspw. körperliche Beanspruchung) zur Profilextraktion genutzt werden. Dadurch kann die erarbeitete Übersicht relevanter und irrelevanter Parameter erweitert werden. Es ist anzumerken, dass die in dieser Arbeit vorgenommene Einteilung in relevante und irrelevante Parameter ausschließlich für die Vorhersage mentaler und emotionaler Beanspruchung gilt. Für andere psychologische Konstrukte (bspw. körperliche Beanspruchung) können diese scheinbar irrelevanten Parameter wieder relevant sein. Ein Beispiel hierfür ist das elektrodermale Profil des NASA-RTLX der Studie I. Dieses Profil wurde als Repräsentant eines nicht spezifizierten Störfaktors der Studie I vermutet. Auf diesem laden diejenigen Parameter hoch, welche für die mentale und emotionale Beanspruchung irrelevant sind. Welches Konstrukt das Profil aus Studie I repräsentiert, ist zukünftig zu untersuchen. Potentielle Kandidaten zur Erklärung dieses Profils sind die körperliche Beanspruchung, der zeitliche Verlauf oder die Art der mentalen Belastung.

Die extrahierten Profile dieser Arbeit ermöglichen es, die gemessenen physiologischen Kenngrößen als Indikatoren mentaler Beanspruchung psychometrisch zu validieren. Bspw. ist es durch die physiologischen Profile nun möglich, die Konstruktvalidität statistisch, z. B. mit Hilfe einer MTMM-Analyse, zu bestimmen. Forschungsbestrebungen zur Konstruktvalidierung der Profile würden eine objektive Bewertung bereitstellen, wie gut physiologische Profile mentaler Beanspruchung mit anderen Maßen mentaler Beanspruchung übereinstimmen bzw. zu Maßen konkurrierender Konstrukte abgegrenzt werden können.

Ein weiterer Forschungsansatz wäre die Kombination physiologischer Profile unterschiedlicher Konstrukte mit dem Ziel, die Vorhersagegüte der mentalen Beanspruchung zu steigern. So können bspw. das kardiale und elektrodermale Profil gemeinsam zur Prädiktion mentaler Beanspruchung genutzt werden. Von Interesse ist hier vor allem die Frage, ob die beiden Profile unabhängige oder überlappende Anteile in der mentalen Beanspruchung erklären (Multikollinearität). Bei der Kombination physiologischer Profile können auch Profile weiterer physiologischer Daten, wie der okulomotorischen Aktivität, herangezogen werden. Bspw. könnten aus den okulomotorischen Daten Profile der Blickstrategien extrahiert und zu bestimmten Aufgabentypen zugeordnet werden. Da der Aufgabentyp einen signifikanten Effekt auf die mentale Beanspruchung zeigt, könnte das Profil der Blickstrategie als zusätzlicher Prädiktor der mentalen Beanspruchung dienen. Auf diese Weise können auch weitere physiologische Profile – z. B. ein Profil der körperlichen Beanspruchung mit Hilfe von Bewegungsdaten – ermittelt und in einem Mensch-Maschine-System als Prädiktor mentaler Beanspruchung genutzt werden.

Ein zusätzlicher Erkenntnisgewinn wäre durch die Erforschung interindividueller Unterschiede der physiologischen Aktivität mit Hilfe physiologischer Profile zu erwarten. In der vorliegenden Arbeit werden die Profile so extrahiert, dass der Zusammenhang zwischen der physiologischen Aktivität und der mentalen Beanspruchung auf Ebene der intraindividuellen Unterschiede (Bedingungsdimension) maximiert wird. Es ist aber auch möglich, die physiologischen Profile derart zu extrahieren, dass der Zusammenhang zwischen

der physiologischen Prädiktormatrix und mentalen Beanspruchung für die Personendimension maximiert wird. Ein solches Vorgehen macht vor allem dann Sinn, wenn Dispositionen der Personen – also interindividuelle Unterschiede – den Forschungsfokus bilden. Auf diese Weise lassen sich Person- und Persönlichkeitseigenschaften identifizieren, welche für die Unterschiede in der physiologischen Aktivität verantwortlich sind. Diese könnten wiederum Aufschluss darüber geben, weshalb der Zusammenhang zwischen der physiologischen Aktivität und der mentalen Beanspruchung für manche Personengruppen hoch, niedrig oder sogar entgegengesetzt ist. Einen vielversprechenden Ansatz bietet die Personeneigenschaft der sog. elektrodermalen Labilität (Lacey & Lacey, 1958; Mundy-Castle & McKiever, 1953). Als elektrodermal labil gelten Personen mit einer hohen Anzahl an NS-SCRs oder einer langsamen SCR-Habituation. Forschungsarbeiten von Koelega (1990) und Munro, Dawson, Schell und Sakai (1987) zeigen auf, dass elektrodermal labile Personen bei Aufgaben besser abschneiden als elektrodermal stabile Personen. Informationen über solche physiologische Dispositionen würden zu einer präziseren Prädiktion der mentalen Beanspruchung bei den entsprechenden Personengruppen führen.

9.5 Fazit

Der vorliegenden Arbeit ist es gelungen, kardiale und elektrodermale Einzelparameter zu kardialen und elektrodermalen Profilen mentaler Beanspruchung zusammenzufassen. Dadurch konnte die Voraussetzung geschaffen werden, um von der gemessenen physiologischen Aktivität auf die mentale Beanspruchung zu schließen. Damit leistet diese Arbeit einen wichtigen Beitrag zur validen Vorhersage der mentalen Beanspruchung in einem nutzerzentrierten Mensch-Maschine-System. Die abgeleiteten Profile mentaler Beanspruchung sind sensitiv und spezifisch für die mentale Beanspruchung und können über verschiedene mentale Belastungsfaktoren generalisiert werden. Darüber hinaus wurde eine Übersicht erstellt, die relevante, spezifische und redundante Parameter für die Vorhersage der mentalen Beanspruchung beschreibt. Die

angewandte mehrdimensionale Methode und die extrahierten Profile ebnen den Weg für zukünftige Forschungsbestrebungen und die Anwendung in einer beanspruchungssensitiven Mensch-Maschine-Schnittstelle.

Literaturverzeichnis

- Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (pls regression). *Wiley Interdisciplinary Reviews: Computational Statistics*, 2 (1), 97–106. doi: 10.1002/wics.51
- Afzal, S. & Robinson, P. (2011). Designing for automatic affect inference in learning environments. *Journal of Educational Technology & Society*, 14 (4), 21–34.
- Andersson, C. A. & Bro, R. (2000). The n-way toolbox for matlab. *Chemometrics and Intelligent Laboratory Systems*, 52 (1), 1–4. doi: 10.1016/S0169-7439(00)00071-X
- Annala, P. A., Yli-Hankala, A. M. & Lindgren, L. (1994). The effect of atropine on the T-wave amplitude of ecg during isoflurane anaesthesia. *International journal of clinical monitoring and computing*, 11 (1), 43–47.
- Aust, F. & Barth, M. (2018). *papaja: Prepare apa journal articles with r markdown*. Zugriff am 01.10.2018 auf <https://github.com/crsh/papaja>
- Backs, R. W. (1995). Going beyond heart rate: autonomic space and cardiovascular assessment of mental workload. *The International journal of aviation psychology*, 5 (1), 25–48. doi: 10.1207/s15327108ijap0501_3
- Backs, R. W. (1998). A comparison of factor analytic methods of obtaining cardiovascular autonomic components for the assessment of mental workload. *Ergonomics*, 41 (5), 733–745. doi: 10.1080/001401398186883
- Backs, R. W. (2000). Application of psychophysiological models to mental workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44 (21), 3-464-3-467. doi: 10.1177/154193120004402123

- Backs, R. W. & Boucsein, W. (2009). Psychophysiology in digital human modeling. In V. G. Duffy (Hrsg.), *Handbook of digital human modeling* (S. 16-1–16-14). Boca Raton, Fla.: CRC Press.
- Backs, R. W., Lenneman, J. K. & Sicard, J. L. (1999). The use of autonomic components to improve cardiovascular assessment of mental workload in flight. *The International journal of aviation psychology*, 9 (1), 33–47. doi: 10.1207/s15327108ijap0901_3
- Backs, R. W., Rohdy, J. & Barnard, J. (2005). Cardiac control during dual-task performance of visual or auditory monitoring with visual-manual tracking. *Psychologia*, 48 (2), 66–83. doi: 10.2117/psychoc.2005.66
- Backs, R. W. & Seljos, K. A. (1994). Metabolic and cardiorespiratory measures of mental effort: the effects of level of difficulty in a working memory task. *International Journal of Psychophysiology*, 16 (1), 57–68. doi: 10.1016/0167-8760(94)90042-6
- Baek, H. J., Cho, C.-H., Cho, J. & Woo, J.-M. (2015). Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability. *Telemedicine journal and e-health: the official journal of the American Telemedicine Association*, 21 (5), 404–414. doi: 10.1089/tmj.2014.0104
- Baltaci, S. & Gokcay, D. (2016). Stress detection in human-computer interaction: Fusion of pupil dilation and facial temperature features. *International Journal of Human-Computer Interaction*, 32 (12), 956–966. doi: 10.1080/10447318.2016.1220069
- Bartenwerfer, H. (1969). Einige praktische Konsequenzen aus der Aktivierungstheorie. *Zeitschrift für experimentelle und angewandte Psychologie*, 16 (2), 195–222.
- Bartoń, K. (2019). *Mumin: Multi-model inference*. Zugriff am 23.09.2020 auf <https://CRAN.R-project.org/package=MumIn>
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67 (1), 1–48. doi: 10.18637/jss.v067.i01

- Becker Meditec. (2016, 06). Gebrauchsanweisung varioport [Software-Handbuch]. Karlsruhe: Becker Meditec.
- Beckers, F., Ramaekers, D. & Aubert, A. E. (2001). Approximate entropy of heart rate variability: Validation of methods and application in heart failure. *Cardiovascular Engineering: An International Journal*, 1 (4), 177–182.
- Beckmann, C. F. & Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject fmri analysis. *NeuroImage*, 25 (1), 294–311. doi: 10.1016/j.neuroimage.2004.10.043
- Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G. & Sleight, P. (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *Journal of the American College of Cardiology*, 35 (6), 1462–1469. doi: 10.1016/S0735-1097(00)00595-7
- Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . Stone, P. H. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34 (6), 623–648.
- Berntson, G. G. & Cacioppo, J. T. (2002). Psychophysiology. In H. D'Haenen, J. A. den Boer & P. Willner (Hrsg.), *Biological psychiatry* (S. 1–16). Chichester: John Wiley & Sons, Ltd.
- Berntson, G. G., Cacioppo, J. T. & Grossman, P. (2007). Whither vagal tone. *Biological psychology*, 74 (2), 295–300. doi: 10.1016/j.biopsycho.2006.08.006
- Berntson, G. G., Cacioppo, J. T. & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review*, 98 (4), 459–487. doi: 10.1037/0033-295X.98.4.459
- Berntson, G. G., Cacioppo, J. T., Quigley, K. S. & Fabro, V. T. (1994). Autonomic space and psychophysiological response. *Psychophysiology*, 31 (1), 44–61.

- Berntson, G. G., Norman, G. J., Hawkley, L. C. & Cacioppo, J. T. (2008). Cardiac autonomic balance versus cardiac regulatory capacity. *Psychophysiology*, 45 (4), 643–652. doi: 10.1111/j.1469-8986.2008.00652.x
- Berntson, G. G., Quigley, K. S., Norman, G. J. & Lozano, D. L. (2017). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 183–216). Cambridge, United Kingdom: Cambridge University Press.
- Betella, A. & Verschure, P. F. M. J. (2016, 02). The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLOS ONE*, 11 (2), 1-11. doi: 10.1371/journal.pone.0148037
- Blascovich, J. J., Vanman, E., Berry Mendes, W. & Dickerson, S. S. (2011). *Social psychophysiology for social and personality psychology*. London: SAGE Publications.
- Blender Online Community. (2015). Blender - a 3d modelling and rendering package [Software-Handbuch]. Stichting Blender Foundation, Amsterdam. Zugriff am 18.04.2017 auf <http://www.blender.org>
- Boonnithi, S. & Phongsuphap, S. (2011). Comparison of heart rate variability measures for mental stress detection. *Computing in Cardiology*, 38, 85-88.
- Boucsein, W. (2001). Physiologische Grundlagen und Meßmethoden der dermalen Aktivität. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich C. Theorie und Forschung. Grundlagen und Methoden der Psychophysilogie* (Bd. 4, S. 551–623). Göttingen: Hogrefe, Verlag für Psychologie.
- Boucsein, W. (2012). *Electrodermal activity*. Boston, MA: Springer Science+Business Media LLC. doi: 10.1007/978-1-4614-1126-0
- Bower, G. H. (1981). Mood and memory. *American Psychologist* (36), 129–148.
- Bradley, M. M., Codispoti, M., Cuthbert, B. N. & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1 (3), 276–298. doi: 10.1037/1528-3542.1.3.276

- Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25 (1), 49–59. doi: 10.1016/0005-7916(94)90063-9
- Brickenkamp, R. (2002). *Test d2: Aufmerksamkeits-Belastungs-Test: Manual* (9., überarbeitete und neu normierte Aufl.). Göttingen: Hogrefe, Verlag für Psychologie.
- Bro, R. (1997). Parafac. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems*, 38 (2), 149 - 171. doi: [https://doi.org/10.1016/S0169-7439\(97\)00032-4](https://doi.org/10.1016/S0169-7439(97)00032-4)
- Byers, J. C., Bittner, A. C. & Hill, S. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Hrsg.), *Advances in industrial ergonomics and safety* (S. 481–485). London, New York, Philadelphia: Taylor & Francis.
- Cacioppo, J. T., Berntson, G. G. & Klein, D. J. (1992). What is emotion? The role of somatovisceral "illusions". *Review of Personality and Social Psychology* (14), 63–98.
- Cacioppo, J. T. & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45 (1), 16–28. doi: 10.1037//0003-066X.45.1.16
- Cacioppo, J. T., Tassinary, L. G. & Berntson, G. G. (2000). Psychophysiological science. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 3–23). New York, NY, US: Cambridge University Press.
- Cacioppo, J. T., Tassinary, L. G. & Berntson, G. G. (2007). Psychophysiological science: Interdisciplinary approaches to classic questions about the mind. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 1–16). Cambridge, England: Cambridge University Press.
- Cacioppo, J. T., Tassinary, L. G. & Berntson, G. G. (2017). Strong inference in psychophysiological science. In J. T. Cacioppo, L. G. Tassinary

- & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 3–15). Cambridge, United Kingdom: Cambridge University Press.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56* (2), 81–105. doi: 10.1037/h0046016
- Cannon, W. B. (1927). The James-Lange theory of emotion: A critical examination and an alternative theory. *American Journal of Psychology* (39), 10–124.
- Carreiras, C., Alves, A. P., Lourenço, A., Canento, F., Silva, H. & Fred, A. (2015). *BioSPPy: Biosignal processing in Python*. Zugriff am 15.02.2019 auf <https://github.com/PIA-Group/BioSPPy>
- Castaldo, R., Montesinos, L., Wan, T. S., Serban, A., Massaro, S., Pecchia, L., . . . Hyttinen, J. (2018). Heart rate variability analysis and performance during a repeated mental workload task. In *EMBECE & NBC 2017* (S. 69–72). Singapore: Springer Singapore. doi: 10.1007/978-981-10-5122-7_18
- CEN Europäisches Komitee für Normung. (2011). *Ergonomie - Genereller Ansatz, Prinzipien und Konzepte* (Bd. 13.180) (Nr. EN ISO 26800).
- Chanel, G., Kronegg, J., Grandjean, D. & Pun, T. (2006). Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals. In B. Günsel, A. Jain, A. Tekalp & B. Sankur (Hrsg.), *Multimedia content representation, classification and security* (Bd. 4105, S. 530–537). Springer Berlin Heidelberg. doi: 10.1007/11848035_70
- Chanel, G., Rebetez, C., Bétrancourt, M. & Pun, T. (2011). Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *41* (6), 1052–1063. doi: 10.1109/TSMCA.2011.2116000
- Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S. Z., Khawaji, A. & Conway, D. (2016). *Robust multimodal cognitive load measurement* (1st Aufl.). Springer Publishing Company, Incorporated. doi: 10.1007/978-3-319-31700-7
- Chen, S. & Epps, J. (2013). Automatic classification of eye activity for cognitive load measurement with emotion interference. *Computer methods and*

- programs in biomedicine*, 110 (2), 111–124. doi: 10.1016/j.cmpb.2012.10.021
- Collet, C., Salvia, E. & Petit-Bou langer, C. (2014). Measuring workload with electrodermal activity during common braking actions. *Ergonomics*, 57 (6), 886–896. doi: 10.1080/00140139.2014.899627
- Collins, R. (2020). Clinician cognitive overload and its implications for nurse leaders. *Nurse Leader*, 18 (1), 44–47. doi: <https://doi.org/10.1016/j.mnl.2019.11.007>
- Dawson, M. E., Schell, A. M. & Filion, D. L. (2017). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 217–243). Cambridge, United Kingdom: Cambridge University Press.
- DIN Deutsches Institut für Normung e. V. (2000). *Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung; Teil 1: Allgemeines und Begriffe* (Bd. 01.040.13; 13.180) (Nr. EN ISO 10075-1). Berlin: Beuth Verlag GmbH.
- DIN Deutsches Institut für Normung e. V. (2004). *Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung; Teil 3: Grundsätze und Anforderungen an Verfahren zur Messung und Erfassung psychischer Arbeitsbelastung* (Bd. 13.180) (Nr. EN ISO 10075-3). Berlin: Beuth Verlag GmbH.
- Ditton, H. (1998). *Mehrebenenanalyse. Grundlagen und Anwendungen des hierarchisch linearen Modells*. Weinheim: Juventa.
- Dix, A., Finlay, J., Abowd, G. D. & Beale, R. (2004). *Human-computer interaction* (3. Aufl.). Harlow, England: Pearson/Prentice-Hall.
- Duschek, S., Muckenthaler, M., Werner, N. & del Paso, G. A. R. (2009). Relationships between features of autonomic cardiovascular control and cognitive performance. *Biological psychology*, 81 (2), 110–117.
- Dyer, F. N. & Severance, L. J. (1973). Stroop interference with successive presentations of separate incongruent words and colors. *Journal of Experimental Psychology*, 98 (2), 438. doi: 10.1037/h0034353

- Edelberg, R. (1972). Electrodermal recovery rate, goal-orientation, and aversion. *Psychophysiology*, 9 (5), 512–520. doi: 10.1111/j.1469-8986.1972.tb01805.x
- Edelmann, W. (2000). *Lernpsychologie* (6. Auflage Aufl.). Beltz, PVU.
- Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Hrsg.), *Handbook of cognition and emotion* (S. 45–60). Chichester, England: Wiley.
- Engström, J., Johansson, E. & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8 (2), 97–120. doi: 10.1016/j.trf.2005.04.012
- Eysenck, H. J. (1965). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychological Bulletin*, 63 (4), 258–270. doi: 10.1037/h0021921
- Fahrenberg, J. (1983). Psychophysiologische Methodik. In K.-J. Groffmann & C. F. Graumann (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich B: Methodologie und Methoden. Psychologische Diagnostik* (Bd. 4, S. 1–192). Göttingen: Hogrefe, Verlag für Psychologie.
- Fahrenberg, J. (2001). Physiologische Grundlagen und Meßmethoden der Herz-Kreislaufaktivität. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Serie Biologische Psychologie. Bereich Psychophysiology. Grundlagen und Methoden der Psychophysiology* (Bd. 1, S. 317–483). Göttingen: Hogrefe, Verlag für Psychologie.
- Fairclough, S. H. & Houston, K. (2004). A metabolic measure of mental effort. *Biological psychology*, 66 (2), 177–190.
- Fairclough, S. H. & Venables, L. (2006). Prediction of subjective states from psychophysiology: a multivariate approach. *Biological psychology*, 71 (1), 100–110. doi: 10.1016/j.biopsycho.2005.03.007
- Feigenspan, A. (2017). *Prinzipien der Physiologie: Grundlegende Mechanismen und evolutionäre Strategien*. Berlin: Springer Spektrum. doi: 10.1007/978-3-662-54117-3

- Furedy, J. J., Szabo, A. & Peronnet, F. (1996). Effects of psychological and physiological challenges on heart rate, t-wave amplitude, and pulse-transit time. *International Journal of Psychophysiology*, 22 (3), 173–183.
- Gaillard, A. W. (1993). Comparing the concepts of mental load and stress. *Ergonomics*, 36 (9), 991–1005. doi: 10.1080/00140139308967972
- Galy, E., Cariou, M. & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, 83 (3), 269–275. doi: 10.1016/j.ijpsycho.2011.09.023
- Galy, E., Paxion, J. & Berthelon, C. (2018). Measuring mental workload with the NASA-TLX needs to examine each dimension rather than relying on the global score: An example with driving. *Ergonomics*, 61 (4), 517–527. doi: 10.1080/00140139.2017.1369583
- Gauthier, G., Vercher, J., Ivaldi, F. M. & Marchetti, E. (2004). Oculo-manual tracking of visual targets: Control learning, coordination control and coordination model. *Experimental Brain Research*, 73, 127-137.
- Ghaderyan, P. & Abbasi, A. (2016). An efficient automatic workload estimation method based on electrodermal activity using pattern classifier combinations. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology*, 110, 91–101. doi: 10.1016/j.ijpsycho.2016.10.013
- Goldberger, A. L. (1991). Is the normal heartbeat chaotic or homeostatic? *News in physiological sciences: an international journal of physiology produced jointly by the International Union of Physiological Sciences and the American Physiological Society*, 6, 87–91. doi: 10.1152/physiologyonline.1991.6.2.87
- Grassmann, M., Vlemincx, E., Leupoldt, A., Mittelstädt, J. & Van den Bergh, O. (2016). Respiratory changes in response to cognitive load: A systematic review. *Neural Plasticity*, 2016, 1-16. doi: 10.1155/2016/8146809
- Greco, A., Valenza, G., Lanata, A., Scilingo, E. P. & Citi, L. (2016). cvxeda: A convex optimization approach to electrodermal activity processing. *IEEE transactions on bio-medical engineering*, 63 (4), 797–804. doi: 10.1109/TBME.2015.2474131

- Greco, A., Valenza, G., Nardelli, M., Bianchi, M., Citi, L. & Scilingo, E. P. (2017). Force–velocity assessment of caress-like stimuli through the electrodermal activity processing: Advantages of a convex optimization approach. *IEEE Transactions on Human-Machine Systems*, *47* (1), 91-100. doi: 10.1109/THMS.2016.2586478
- Green, P. & MacLeod, C. J. (2016). simr: an r package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, *7* (4), 493–498. doi: 10.1111/2041-210X.12504
- Gupta, R., Mitra, M. & Bera, J. (2014). *ECG acquisition and automated remote processing*. New Delhi: Springer.
- Haapalainen, E., Kim, S., Forlizzi, J. F. & Dey, A. K. (2010). Psycho-physiological measures for assessing cognitive load. In *Proceedings of the 12th acm international conference on ubiquitous computing* (S. 301–310). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/1864349.1864395
- Hair, J. F., Tatham, R. L., Anderson, R. E. & Black, W. (1998). *Multivariate data analysis* (5. Aufl.). London: Prentice Hall.
- Hancock, P. A. & Meshkati, N. (1988). *Human mental workload* (1. Aufl.). Amsterdam: North-Holland.
- Harshman, R. A. & Lundy, M. E. (1994). Parafac: Parallel factor analysis. *Computational Statistics & Data Analysis*, *18* (1), 39 - 72. doi: [https://doi.org/10.1016/0167-9473\(94\)90132-5](https://doi.org/10.1016/0167-9473(94)90132-5)
- 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.), *Advances in psychology: Human mental workload* (Bd. 52, S. 139–183). Amsterdam: North-Holland. doi: 10.1016/S0166-4115(08)62386-9
- Healey, J. A. & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6* (2), 156–166. doi: 10.1109/TITS.2005.848368
- Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O. & Deml, B. (2017). Electrocardiographic features for the measurement of drivers'

- mental workload. *Applied Ergonomics*, 61, 31-43. doi: <https://doi.org/10.1016/j.apergo.2016.12.015>
- Hendy, K. C., Hamilton, K. M. & Landry, L. N. (1993). Measuring subjective workload: When is one scale better than many? *Human factors*, 35 (4), 579–601. doi: 10.1177/001872089303500401
- Heslegrave, R. J. & Furedy, J. J. (1979). Sensitivities of HR and Tx -wave amplitude for detecting cognitive and anticipatory stress. *Physiology & Behavior*, 22 (1), 17–23.
- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2. Aufl.). New York: Routledge.
- Huang, W.-H. D. & Tettegah, S. (2010). Cognitive load and empathy in serious games. In R. van Eck (Hrsg.), *Gaming and cognition* (S. 137–151). Hershey, Pa: IGI Global. doi: 10.4018/978-1-61520-717-6.ch006
- Hussain, M. S., Calvo, R. A. & Chen, F. (2014). Automatic cognitive load detection from face, physiology, task performance and fusion during affective interference. *Interacting with Computers*, 26 (3), 256–268.
- Ibrahim, N., Shiratuddin, M. F. & Wong, K. (2015). Instruments for measuring the influence of visual persuasion: Validity and reliability tests. *European Journal of Social Sciences Education and Research*, 4. doi: 10.26417/ejser.v4i1.p25-37
- Immordino-Yang, M. H. & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*, 1 (1), 3–10. doi: 10.1111/j.1751-228x.2007.00004.x
- James, W. (1884). What is emotion? *Mind* (9), 188–205.
- Janssen, S., Heijs, J. J. A., van der Meijs, W., Nonnekes, J., Bittner, M., Dorresteijn, L. D. A., . . . Heida, T. (2019). Validation of the auditory Stroop task to increase cognitive load in walking tasks in healthy elderly and persons with parkinson's disease. *PLOS ONE*, 14 (8), 1-14. doi: 10.1371/journal.pone.0220735
- Jennings, J. R. & Allen, B. (2017). Methodology. In J. T. Cacioppo, L. G. Tassinari & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 583–611). Cambridge, United Kingdom: Cambridge University Press.

- Johannsen, G. (1993). *Mensch-Maschine-Systeme*. Berlin: Springer. doi: 10.1007/978-3-642-46785-1
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Kapoor, A., Burleson, W. & Picard, R. W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65 (8), 724–736. doi: 10.1016/j.ijhcs.2007.02.003
- Kim, J., André, E., Rehm, M., Vogt, T. & Wagner, J. (2005). Integrating information from speech and physiological signals to achieve emotional sensitivity. In *INTERSPEECH 2005 - eurospeech, 9th european conference on speech communication and technology* (S. 809–812). Lisbon, Portugal: ISCA. doi: 10.21437/Interspeech.2005-380
- Kim, K. H., Bang, S. W. & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing*, 42 (3), 419–427. doi: 10.1007/BF02344719
- Kirchner, J.-H. (1986). Belastungen und Beanspruchungen - Einige begriffliche Klärungen zum Belastungs-Beanspruchungs-Konzept. In R. Hackstein, F.-J. Heeg & F. Below (Hrsg.), *Arbeitsorganisation und neue Technologien* (S. 553–569). Berlin: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-93339-4_23
- Kirchner, S. (2015). Konturen der digitalen Arbeitswelt. Eine Untersuchung der Einflussfaktoren beruflicher Computer- und Internetnutzung und der Zusammenhänge zu Arbeitsqualität. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 67, 763–791.
- Kline, K. P., Ginsburg, G. P. & Johnston. (1998). T-wave amplitude: Relationships to phasic RSA and heart period changes. *International Journal of Psychophysiology*, 29 (3), 291–301.
- Koelega, H. S. (1990). Vigilance performance: A review of electrodermal predictors. *Perceptual and Motor Skills*, 70 (3), 1011-1029. doi: 10.2466/pms.1990.70.3.1011
- Koenig, A., Novak, D., Omlin, X., Pulfer, M., Perreault, E., Zimmerli, L., . . . Riener, R. (2011). Real-time closed-loop control of cognitive load in neurological patients during robot-assisted gait training. *IEEE transac-*

- tions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society, 19* (4), 453–464. doi: 10.1109/TNSRE.2011.2160460
- Kohlisch, O. & Schaefer, F. (1996). Physiological changes during computer tasks: Responses to mental load or to motor demands? *Ergonomics*, 39 (2), 213–224. doi: 10.1080/00140139608964452
- Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Hrsg.), *Multiple task performance* (S. 279–328). London: Taylor & Francis.
- Kron, A., Pilkiw, M., Banaei, J., Goldstein, A. & Anderson, A. K. (2015). Are valence and arousal separable in emotional experience? *Emotion*, 15 (1), 35–44. doi: 10.1037/a0038474
- Kroonenberg, P. M. (1983). *Three-mode principal component analysis: Theory and applications*. Leiden: DWO Press.
- Kroonenberg, P. M. (2008). *Applied multiway data analysis*. Hoboken N.J.: Wiley-Interscience.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82 (13). doi: 10.18637/jss.v082.i13
- Lacey, J. I. & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *The American Journal of Psychology*, 71 (1), 50. doi: 10.2307/1419197
- Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Shlien (Hrsg.), *Research in psychotherapy* (S. 90–102). Washington, DC, US: American Psychological Association. doi: 10.1037/10546-004
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson & T. A. Williams (Hrsg.), *Technology in mental health care delivery systems* (S. 119–137). Norwood, NJ: Ablex.

- Lang, P. J., Bradley, M. M. & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. technical report A-8*. University of Florida, Gainesville, FL.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY, US: Oxford University Press.
- Lazarus, R. S., Averill, J. R. & Opton, E. M., JR. (1970). Toward a cognitive theory of emotion. In M. B. Arnold (Hrsg.), *Feelings and emotion* (S. 207–232). New York: Academic Press.
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer Pub. Co.
- Lenneman, J. K. & Backs, R. W. (2003). The evolution of autonomic space as a method of mental workload assessment for driving. In *Proceedings of the 2nd international driving symposium on human factors in driver assessment, training and vehicle design: Driving assessment 2003* (S. 125–129). Iowa City, Iowa: University of Iowa. doi: 10.17077/drivingassessment.1106
- Lenneman, J. K. & Backs, R. W. (2007). Diagnosticity of cardiac modes of autonomic control elicited by simulated driving and verbal working memory dual-tasks. In D. Harris (Hrsg.), *Engineering psychology and cognitive ergonomics* (S. 541–550). Berlin: Springer.
- Li, S.-y., Yang, M., Li, C.-c. & Cai, P. (2008). Analysis of heart rate variability based on singular value decomposition entropy. *Journal of Shanghai University (English Edition)*, 12 (5), 433–437. doi: 10.1007/s11741-008-0511-3
- Liao, W., Zhang, W., Zhu, Z., Ji, Q. & Gray, W. D. (2006). Toward a decision-theoretic framework for affect recognition and user assistance. *International Journal of Human-Computer Studies*, 64 (9), 847–873. doi: 10.1016/j.ijhcs.2006.04.001
- Lippman, N., Stein, K. M. & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology-Heart and Circulatory Physiology*, 267 (1), H411–H418.

- Lorenzo-Seva, U. & Ten Berge, J. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, 2 (2), 57–64. doi: 10.1027/1614-2241.2.2.57
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109 (2), 163–203. doi: 10.1037/0033-2909.109.2.163
- Makowski, D. (2016). *Neurokit: A python toolbox for statistics and neurophysiological signal processing*. Paris, France: Memory and Cognition Lab. Zugriff am 23.04.2018 auf <https://pypi.org/project/neurokit>
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J. & Schwartz, P. J. (1996, 03). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17 (3), 354–381. doi: 10.1093/oxfordjournals.eurheartj.a014868
- Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich C: Theorie und Forschung. Biologische Psychologie. Ergebnisse und Anwendungen der Psychophysiologie* (Bd. 5, S. 799–864). Göttingen: Hogrefe, Verlag für Psychologie.
- Martin, J., Schneider, F., Kowalewskij, A., Jordan, D., Hapfelmeier, A., Kochs, E. F., . . . Schulz, C. M. (2016). Linear and non-linear heart rate metrics for the assessment of anaesthetists' workload during general anaesthesia. *BJA: British Journal of Anaesthesia*, 117 (6), 767–774.
- Mathôt, S., Schreij, D. & Theeuwes, J. (2011). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods*, 44, 314–24. doi: 10.3758/s13428-011-0168-7
- MATLAB. (2018). *Matlab:2018*. Natick, Massachusetts: The MathWorks Inc.
- 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. doi: 10.1177/0018720814539505
- Maurissen, J. P. & Vidmar, T. J. (2017). Repeated-measure analyses: Which one? A survey of statistical models and recommendations for reporting.

- Neurotoxicology and Teratology*, 59, 78–84. doi: 10.1016/j.ntt.2016.10.003
- Mehler, B., Reimer, B. & Coughlin, J. F. (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: An on-road study across three age groups. *Human factors*, 54 (3), 396–412. doi: 10.1177/0018720812442086
- Merenda, P. F. (1997). A guide to the proper use of factor analysis in the conduct and reporting of research: Pitfalls to avoid. *Measurement and Evaluation in Counseling and Development*, 30 (3), 156–164. doi: 10.1080/07481756.1997.12068936
- Meyer, S.-C., Tisch, A. & Hünefeld, L. (2019). Arbeitsintensivierung und Handlungsspielraum in digitalisierten Arbeitswelten - Herausforderung für das Wohlbefinden von Beschäftigten? *Industrielle Beziehungen. Zeitschrift für Arbeit, Organisation und Management*, 26 (2-2019), 207–231. doi: 10.3224/indbez.v26i2.06
- Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63 (2), 81–97.
- Miwakeichi, F., Martinez-Montes, E., Valdes-Sosa, P. A., Nishiyama, N., Mizuhara, H. & Yamaguchi, Y. (2004). Decomposing EEG data into space-time-frequency components using parallel factor analysis. *NeuroImage*, 22 (3), 1035–1045. doi: 10.1016/j.neuroimage.2004.03.039
- Morup, M., Hansen, L. K., Herrmann, C. S., Parnas, J. & Arnfred, S. M. (2006). Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG. *NeuroImage*, 29 (3), 938–947. doi: 10.1016/j.neuroimage.2005.08.005
- Moya-Albiol, L., Salvador, A., Costa, R., Martínez-Sanchis, S., González-Bono, E., Ricarte, J. & Arnedo, M. (2001). Psychophysiological responses to the Stroop task after a maximal cycle ergometry in elite sportsmen and physically active subjects. *International Journal of Psychophysiology*, 40 (1), 47–59. doi: 10.1016/S0167-8760(00)00125-2

- Mukherjee, S., Yadav, R., Yung, I., Zajdel, D. P. & Oken, B. S. (2011). Sensitivity to mental effort and test-retest reliability of heart rate variability measures in healthy seniors. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, 122 (10), 2059–2066. doi: 10.1016/j.clinph.2011.02.032
- Mulder, G., Mulder, L. J. M., Meijman, T. F., Veldman, J. B. P. & van Roon, A. M. (2000). A psychophysiological approach to working conditions. In R. W. Backs & W. Boucsein (Hrsg.), *Engineering psychophysiology* (S. 139–160). Mahwah NJ: Lawrence Erlbaum.
- Mundy-Castle, A. C. & McKiever, B. L. (1953). The psychophysiological significance of the galvanic skin response. *Journal of Experimental Psychology*, 46 (1), 15–24. doi: 10.1037/h0060100
- Munro, L. L., Dawson, M. E., Schell, A. M. & Sakai, L. M. (1987). Electrodermal lability and rapid vigilance decrement in a degraded stimulus continuous performance task. *Journal of Psychophysiology*, 1 (3), 249–257.
- Myrtek, M., Deutschmann-Janicke, E., Strohmaier, H., Zimmermann, W., Lawrenz, S., Brügger, G. & Müller, W. (1994). Physical, mental, emotional, and subjective workload components in train drivers. *Ergonomics*, 37 (7), 1195–1203.
- Nezlek, J. B., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung. *Psychologische Rundschau*, 57 (4), 213–223. doi: 10.1026/0033-3042.57.4.213
- Nickel, P. & Nachreiner, F. (2003). Sensitivity and diagnosticity of the 0.1-Hz component of heart rate variability as an indicator of mental workload. *Human factors*, 45 (4), 575–590. doi: 10.1518/hfes.45.4.575.27094
- Noreen, S. & MacLeod, M. D. (2015). What do we really know about cognitive inhibition? Task demands and inhibitory effects across a range of memory and behavioural tasks. *PLOS ONE*, 10 (8), 1–21. doi: 10.1371/journal.pone.0134951
- Nourbakhsh, N., Chen, F., Wang, Y. & Calvo, R. (2017, 09). Detecting users' cognitive load by galvanic skin response with affective interference.

- ACM Transactions on Interactive Intelligent Systems*, 7, 1–20. doi: 10.1145/2960413
- Novak, D., Mihelj, M. & Munih, M. (2012). A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting with Computers*, 24 (3), 154–172. doi: 10.1016/j.intcom.2012.04.003
- Nussinovitch, U., Elishkevitz, K. P., Katz, K., Nussinovitch, M., Segev, S., Volovitz, B. & Nussinovitch, N. (2011). Reliability of ultra-short ECG indices for heart rate variability. *Annals of Noninvasive Electrocardiology*, 16 (2), 117–122. doi: 10.1111/j.1542-474X.2011.00417.x
- Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human factors*, 33 (1), 17–33. doi: 10.1177/001872089103300102
- Ochs, M., Pelachaud, C. & Sadek, D. (2007). An empathic rational dialog agent. In A. Paiva, R. Prada & R. W. Picard (Hrsg.), *Affective computing and intelligent interaction* (S. 338–349). Berlin: Springer.
- O'Donnell, R. D. & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman & J. P. Thomas (Hrsg.), *Handbook of perception and human performance. Volume 2. Cognitive processes and performance* (S. 1–49). Oxford: John Wiley & Sons, Inc.
- Ortony, A. & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97 (3), 315–331. doi: 10.1037/0033-295x.97.3.315
- Pagani, M., Lombardi, F., Guzzetti, S., Rimoldi, O., Furlan, R., Pizzinelli, P., . . . Piccaluga, E. (1986). Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circulation Research*, 59 (2), 178–193.
- Pecchia, L., Castaldo, R., Montesinos, L. & Melillo, P. (2018). Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations. *Healthcare technology letters*, 5 (3), 94–100. doi: 10.1049/htl.2017.0090

- Peinsipp-Byma, E. & Voit, M. (2019). *KonsensOP - Kontextsensitive Assistenz im aufmerksamen OP: Schlussbericht, Kontextsensitive Assistenz im aufmerksamen OP* (Bericht). Karlsruhe: Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung. doi: 10.2314/KXP:1677281243
- Peng, C. K., Havlin, S., Stanley, H. E. & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5 (1), 82–87. doi: 10.1063/1.166141
- Petermann, F. & Wechsler, D. (2012). *Wechsler adult intelligence scale - fourth edition*. Frankfurt, M.: Pearson.
- Peterson, R. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters*, 11 (3), 261–275. doi: 10.1023/A:1008191211004
- Pfetsch, J. M. (2019). *Richtlinien zur Manuskriptgestaltung* (5., aktualisierte Aufl.). Göttingen: Hogrefe, Verlag für Psychologie.
- Picard, R. W., Vyzas, E. & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 (10), 1175–1191. doi: 10.1109/34.954607
- Plutchik, R. (1980). *Emotion. A psychoevolutionary synthesis*. New York: Harper & Row.
- Porges, S. W. (1992). Autonomic regulation and attention. In B. A. Campbell, H. Hayne & R. Richardson (Hrsg.), *Attention and information processing in infants and adults*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., Aljama-Corrales, T., Charleston-Villalobos, S. & Chon, K. H. (2016a). Power spectral density analysis of electrodermal activity for sympathetic function assessment. *Annals of biomedical engineering*, 44 (10), 3124–3135. doi: 10.1007/s10439-016-1606-6
- Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, Á. D. & Chon, K. H. (2016b). Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity. *American journal*

- of physiology. Regulatory, integrative and comparative physiology*, 311 (3), R582–R591. doi: 10.1152/ajpregu.00180.2016
- Pribram, K. H. & McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review*, 82 (2), 116–149. doi: 10.1037/h0076780
- Quené, H. & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43 (1), 103–121. doi: 10.1016/j.specom.2004.02.004
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Zugriff am 05.01.2019 auf <http://www.R-project.org/>
- Randall, T. D. (1995). An introduction to partial least squares regression. In *Proceedings of the twentieth annual sas users group international conference* (Bd. 20, S. 1250-1257). Cary, NC: SAS Institute Inc.
- Renaud, P. & Blondin, J.-P. (1997). The stress of Stroop performance: Physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27 (2), 87–97. doi: 10.1016/S0167-8760(97)00049-4
- Ribback, S. (2003). *Psychophysiologische Untersuchung mentaler Beanspruchung in simulierten Mensch-Maschine-Interaktionen* (Dissertation). Universität Potsdam, Potsdam.
- Richter, M. & Slade, K. (2017). Interpretation of physiological indicators of motivation: Caveats and recommendations. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology*, 119, 4–10. doi: 10.1016/j.ijpsycho.2017.04.007
- Rieger, A., Stoll, R., Kreuzfeld, S., Behrens, K. & Weippert, M. (2014). Heart rate and heart rate variability as indirect markers of surgeons' intraoperative stress. *International archives of occupational and environmental health*, 87 (2), 165–174. doi: 10.1007/s00420-013-0847-z
- Rohmert, W. (1984). Das Belastungs-Beanspruchungs-Konzept. *Zeitschrift für arbeitswissenschaft*, 38 (4), 193–200.

- Rolls, E. T. (2000). Précis of the brain and emotion. *Behavioral and Brain Sciences*, 23 (2), 177–191. doi: 10.1017/s0140525x00002429
- Rossion, B. & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33 (2), 217-236. doi: 10.1068/p5117
- Rubio, S., Diaz, E., Martin, J. & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, 53 (1), 61–86. doi: 10.1111/j.1464-0597.2004.00161.x
- Rüchel, J. (2013). *Multiparametrische Ermittlung der Beanspruchung bei gezielter emotionaler und mentaler Beanspruchung* (Dissertation). Universitätsmedizin Rostock, Rostock.
- Salahuddin, L., Cho, J., Jeong, M. G. & Kim, D. (2007). Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In *2007 29th annual international conference of the IEEE engineering in medicine and biology society* (S. 4656-4659). Lyon, France: IEEE. doi: 10.1109/iembs.2007.4353378
- Salo, M. A., Huikuri, H. V. & Seppänen, T. (2001). Ectopic beats in heart rate variability analysis: Effects of editing on time and frequency domain measures. *Annals of Noninvasive Electrocardiology*, 6 (1), 5–17. doi: 10.1111/j.1542-474X.2001.tb00080.x
- Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C. & Bigdeli, A. (2008). “How do you know that I don't understand?” A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24 (4), 1342–1363. doi: 10.1016/j.chb.2007.07.008
- Sarter, M., Berntson, G. G. & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist*, 51 (1), 13–21. doi: 10.1037/0003-066X.51.1.13
- Sauvet, F., Jouanin, J. C., Langrume, C., van Beers, P., Papelier, Y. & Dussault, C. (2009). Heart rate variability in novice pilots during and after a multi-

- leg cross-country flight. *Aviation, space, and environmental medicine*, 80 (10), 862–869.
- Schachter, S. & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional states. *Psychology Review* (69), 379–399.
- Scher, H., Furedy, J. J. & Heslegrave, R. J. (1984). Phasic T-wave amplitude and heart rate changes as indices of mental effort and task incentive. *Psychophysiology*, 21 (3), 326–333.
- Schlick, C. M., Bruder, R. & Luczak, H. (2010). *Arbeitswissenschaft*. Berlin: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-78333-6
- Schmidt, R. F., Thews, G. & Lang, F. (2000). *Physiologie des Menschen* (28., korrigierte und aktualisierte Aufl.). Berlin: Springer. doi: 10.1007/978-3-662-09346-7
- Schmitz, S. K., Hasselbach, P. P., Ebisch, B., Klein, A., Pipa, G. & Galuske, R. A. W. (2014). Application of parallel factor analysis (PARAFAC) to electrophysiological data. *Frontiers in neuroinformatics*, 8, 84. doi: 10.3389/fninf.2014.00084
- Schneider, F., Martin, J., Hapfelmeier, A., Jordan, D., Schneider, G. & Schulz, C. M. (2017). The validity of linear and non-linear heart rate metrics as workload indicators of emergency physicians. *PLOS ONE*, 12 (11), 1–12. doi: 10.1371/journal.pone.0188635
- Schneider, M. (2017). *Blickbasierte Beanspruchungsmessung: Entwicklung und Evaluation eines Kalibrierungssystems zur individuellen Bewertung der mentalen Beanspruchung in der Mensch-Technik-Interaktion* (Dissertation). Karlsruher Institut für Technologie, Karlsruhe.
- Schwarz, M., Mack, M., Reisig, G., Gundling, R. & Schuette, A. (2019). *Kontextsensitive Assistenz im aufmerksamen OP, Teilvorhaben: Workflowanalyse von Operationen und Entwicklung von chirurgischen Instrumenten mit integrierter Sensorik: Verbundprojekt: KonsensOP: Schlussbericht, Workflowanalyse von Operationen und Entwicklung von chirurgischen Instrumenten mit integrierter Sensorik, KonsensOP* (Bericht). Mannheim: Orthopädisch-Unfallchirurgisches Zentrum, Sektion Experimentelle Orthopädie und Unfallchirurgie. doi: 10.2314/KXP:1678802093

- Sepehr, M. (1988). *NASA Task Load Index: deutsche Version* (Bericht). Berlin: Technische Universität Berlin, Institut für Fahrzeugtechnik.
- Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G. & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable EDA device. *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society*, 14 (2), 410–417. doi: 10.1109/TITB.2009.2036164
- Shaffer, F. & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in public health*, 5, 258. doi: 10.3389/fpubh.2017.00258
- Shaffer, F., Shearman, S. & Meehan, Z. M. (2016). The promise of ultra-short-term (UST) heart rate variability measurements. *Biofeedback*, 44 (4), 229–233. doi: 10.5298/1081-5937-44.3.09
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27 (3), 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Shao, S., Wang, T., Song, C., Su, Y., Chen, X. & Zhao, H. (2020). The novel indices of short-time heart rate variability for prediction of cardiovascular and cerebrovascular events. *Journal of Medical Imaging and Health Informatics*, 10 (3), 769–774. doi: 10.1166/jmih.2020.2931
- Shen, L., Wang, M. & Shen, R. (2009). Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment. *J. Educ. Technol. Soc.*, 12, 176–189.
- Sheridan, T. B. (1981). Understanding human error and aiding human diagnostic behaviour in nuclear power plants. In J. Rasmussen & W. B. Rouse (Hrsg.), *Human detection and diagnosis of system failures* (S. 19–35). Boston, MA: Springer US. doi: 10.1007/978-1-4615-9230-3_3
- Shimmer. (2016a). ECG User Guide Revision 1.11 [Software-Handbuch]. Dublin: Shimmer Realtime TechnologiesLtd. Zugriff am 15.09.2018 auf http://www.shimmersensing.com/images/uploads/docs/ECG_User_Guide_Rev1.11.pdf

- Shimmer. (2016b). GSR+ User Guide Revision 1.10 [Software-Handbuch]. Dublin: Shimmer Realtime TechnologiesLtd. Zugriff am 15.09.2018 auf http://www.shimmersensing.com/images/uploads/docs/GSR_User_Guide_rev1.10.pdf
- Silvestrini, N. & Gendolla, G. H. E. (2007). Mood effects on autonomic activity in mood regulation. *Psychophysiology*, 44 (4), 650–659. doi: 10.1111/j.1469-8986.2007.00532.x
- Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis*. New York: Oxford University Press. doi: 10.1093/acprof:oso/9780195152968.001.0001
- Smith, A.-L., Owen, H. & Reynolds, K. J. (2013). Heart rate variability indices for very short-term (30 beat) analysis. Part 1: Survey and toolbox. *Journal of clinical monitoring and computing*, 27 (5), 569–576. doi: 10.1007/s10877-013-9471-4
- Snodgrass, J. G. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6 (2), 174–215. doi: 10.1037/0278-7393.6.2.174
- Sokolowski, K. (2002). Kapitel 2c: Emotion. In J. Müsseler & W. Prinz (Hrsg.), *Allgemeine Psychologie* (S. 337–384). Heidelberg: Spektrum Akademischer Verlag.
- Stanton, N. (2005). *Handbook of human factors and ergonomics methods*. Boca Raton: CRC Press.
- Stanton, N. (2013). *Human factors methods: A practical guide for engineering and design* (2. Aufl.). Farnham: Ashgate.
- Stemmler, G. (1989). The autonomic differentiation of emotions revisited: Convergent and discriminant validation. *Psychophysiology*, 26 (6), 617–632. doi: 10.1111/j.1469-8986.1989.tb03163.x
- Stemmler, G. (2001). Grundlagen psychophysiologischer Methodik. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich C. Theorie und Forschung. Grundlagen und Methoden der Psychophysiologie* (Bd. 4, S. 1–84). Göttingen: Hogrefe, Verlag für Psychologie.

- Sterling, P. & Eyer, J. (1988). Allostasis: A new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Hrsg.), *Handbook of life stress, cognition and health* (S. 629–649). Oxford, England: John Wiley & Sons.
- Stiefelhagen, R., Martinez, M., Wolf, E., Philipp, P. R. & Marzi, C. (2019). *Kon-sensOP - Kontextsensitive Assistenz im aufmerksamen OP: Schlussbericht* (Bericht). Karlsruhe: Computer Vision for Human-Computer Interaction Lab, Institut für Arbeitswissenschaft und Betriebsorganisation, Institut für Anthropomatik und Robotik, Karlsruher Institut für Technologie. doi: 10.2314/KXP:1677704829
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18 (6), 643–662. doi: 10.1037/h0054651
- Strube, M. J., Newman, L. C., Lord, A. N. & Nguyen, P. L. (2017). Psychometrics. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology* (S. 612–627). Cambridge, United Kingdom: Cambridge University Press.
- Stuiver, A., Brookhuis, K. A., de Waard, D. & Mulder, B. (2014). Short-term cardiovascular measures for driver support: Increasing sensitivity for detecting changes in mental workload. *International Journal of Psychophysiology*, 92 (1), 35–41. doi: 10.1016/j.ijpsycho.2014.01.010
- Stuiver, A. & Mulder, B. (2014). Cardiovascular state changes in simulated work environments. *Frontiers in Neuroscience*, 8. doi: 10.3389/fnins.2014.00399
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10 (3), 251–296. doi: 10.1023/A:1022193728205
- Timpe, K.-P. & Jürgensohn, T. (2002). Perspektiven der Mensch-Maschine-Systemtechnik. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Hrsg.), *Mensch-Maschine-Systemtechnik* (S. 337–347). Düsseldorf: Symposion.
- Traxel, W. (1960). Die möglichkeit einer objektiven messung der stärke von gefühlen. *Psychologische Forschung*, 26, 75-90.

- Tsang, P. S. & Wilson, G. B. (1997). Mental workload measurement and analysis. In *Handbook of human factors and ergonomics* (2. Aufl., S. 417–449). New York, NY: John Wiley & Sons, Ltd.
- Tulga, M. K. & Sheridan, T. B. (1980). Dynamic decisions and work load in multitask supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics*, 10 (5), 217–232. doi: 10.1109/TSMC.1980.4308481
- van der Voet, H. (1994). Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, 25 (2), 313–323. doi: [https://doi.org/10.1016/0169-7439\(94\)85050-X](https://doi.org/10.1016/0169-7439(94)85050-X)
- van Lien, R., Neijts, M., Willemsen, G. & de Geus, E. J. C. (2015). Ambulatory measurement of the ECG T-wave amplitude. *Psychophysiology*, 52 (2), 225–237.
- van Roon, A. M., Mulder, L. J. M., Althaus, M. & Mulder, G. (2004). Introducing a baroreflex model for studying cardiovascular effects of mental workload. *Psychophysiology*, 41 (6), 961–981.
- Vanitha, L., Suresh, G. R., Chandrasekar, M. & Punita, P. (2017). Development of four stress levels in group Stroop colour word test using HRV analysis. *Biomedical Research*, 28, 98–105.
- Visnovcova, Z., Mestanik, M., Gala, M., Mestanikova, A. & Tonhajzerova, I. (2016). The complexity of electrodermal activity is altered in mental cognitive stressors. *Computers in biology and medicine*, 79, 123–129. doi: 10.1016/j.combiomed.2016.10.014
- Voss, A., Kurths, J., Kleiner, H. J., Witt, A., Wessel, N., Saperin, P., . . . Dietz, R. (1996). The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. *Cardiovasc. Res.*, 31 (3), 419–433.
- Voss, A., Schroeder, R., Heitmann, A., Peters, A. & Perz, S. (2015). Short-term heart rate variability - influence of gender and age in healthy subjects. *PLOS ONE*, 10 (3), e0118308. doi: 10.1371/journal.pone.0118308
- Voss, A., Schulz, S., Schroeder, R., Baumert, M. & Caminal, P. (2009). Methods derived from nonlinear dynamics for analysing heart rate

- variability. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367 (1887), 277–296. doi: 10.1098/rsta.2008.0232
- Wagner, M. (2013). *Der Einfluss emotionaler, kognitiver und körperlicher Beanspruchung auf die Herzrate bei Jugendlichen: eine Laborstudie* (Dissertation). Universität Ulm, Ulm.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Hrsg.), *Varieties of attention* (S. 63–101). Orlando: Academic.
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3 (2), 159–177. doi: 10.1080/14639220210123806
- Wickens, C. D. (2008). Multiple resources and mental workload. *Human factors*, 50 (3), 449–455. doi: 10.1518/001872008X288394
- Wickens, C. D., Hollands, J. G., Banbury, S. & Parasuraman, R. (2013). *Engineering psychology and human performance* (4. internat. Aufl.). Boston, Mass.: Pearson.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer. Zugriff am 16.03.2018 auf <https://ggplot2.tidyverse.org>
- Widyanti, A., Hanna, Muslim, K. & Satalaksana, I. Z. (2017). The sensitivity of galvanic skin response for assessing mental workload in Indonesia. *Work*, 56 (1), 111–117. doi: 10.3233/WOR-162479
- Wierwille, W. W., Rahimi, M. & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human factors*, 27 (5), 489–502.
- Wilson, G. F. & Russell, C. A. (2003a). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human factors*, 45 (3), 381–389. doi: 10.1518/hfes.45.3.381.27252
- Wilson, G. F. & Russell, C. A. (2003b). Real-time assessment of mental workload using psychophysiological measures and artificial neural net-

- works. *Human factors*, 45 (4), 635–643. doi: 10.1518/hfes.45.4.635.27088
- Wilson, G. F. & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Human factors*, 49 (6), 1005–1018. doi: 10.1518/001872007X249875
- Wirtz, M. A. & Strohmer, J. (2014). *Dorsch - Lexikon der Psychologie* (17., vollst. überarb. Aufl.). Bern: Huber.
- Wolf, E., Martinez, M., Roitberg, A., Stiefelhagen, R. & Deml, B. (2018). Estimating mental load in passive and active tasks from pupil and gaze changes using bayesian surprise. In *Proceedings of the workshop on modeling cognitive processes from multimodal data* (S. 1–6). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3279810.3279852
- Wundt, W. (1905). *Grundriß der Psychologie* (7. Aufl.). Leipzig: Engelmann.
- Xiao, Y. M., Wang, Z. M., Wang, M. Z. & Lan, Y. J. (2005). The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index [Abstract]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi = Chinese journal of industrial hygiene and occupational diseases*, 23 (3), 178–181.
- Xu, Q., Nwe, T. L. & Guan, C. (2015). Cluster-based analysis for personalized stress evaluation using physiological signals. *IEEE journal of biomedical and health informatics*, 19 (1), 275–281. doi: 10.1109/JBHI.2014.2311044
- Yeh, Y.-Y. & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human factors*, 30 (1), 111–120. doi: 10.1177/001872088803000110
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative and Neurological Psychology* (18), 459–482.

A Anhang

A.1 Theoretische Verortung des Konstrukts Emotionen

Ein Ansatz zur Definition von Emotionen geht auf Eysenck (1965), Lang (1968) und später auf Lazarus, Averill und Opton (1970) zurück. Diese trennen die Emotionen in ihre Komponenten, die sogenannten Reaktionstrias. Demnach manifestieren sich Emotionen in den drei Komponenten: (1) subjektives Erleben von Gefühlen und Kognition, (2) physiologische Antwort im Sinne einer vegetativen (z. B. Herzrate, Blutdruck) und neuronalen Veränderungen (z. B. elektrische Hirnaktivität) und (3) behaviorale Reaktion im Sinne eines expressiven (z. B. Mimik und Gestik) und instrumentellen Verhaltens (z. B. Fluchtverhalten). Ein Vorteil dieser Differenzierung ist der direkte Bezug der Konstruktebene zur Messebene. Innerhalb der drei Komponenten gibt es eine weitere Aufteilung in theoretische Modellansätze und Modelle. Aus diesem Grund soll in diesem Abschnitt lediglich die für diese Arbeit relevante Komponente des subjektiven Erlebens näher betrachtet werden.

Die Erfassung der Komponente *subjektives Erleben* erfolgt üblicherweise durch Selbstbeobachtung der erlebten Emotion in Form einer schriftlichen oder verbalen Befragung. Die am häufigsten verwendeten Methoden bilden die sog. Emotions-Adjektiv-Skalen, welche die Emotion mit Hilfe von Adjektiven erfassen. Auf theoretischer Ebene wird zwischen dem dimensional und kategorialen Ansatz sowie einer Mischform aus beiden unterschieden. Der dimensionale Ansatz geht auf Wundt (1905) zurück, der die Emotionen als einen mehr oder minder ausgeprägten Zustand auf den Dimensionen „Lust-Unlust“,

„Spannung-Lösung“ sowie „Erregung-Beruhigung“ beschrieb. Im Laufe der Jahre wurden diese Dimensionen modifiziert und erweitert. Bspw. nimmt Lang (1980) – neben den zwei orthogonalen Dimensionen Valenz und Erregung – die Dominanz im Sinne einer subjektiv empfundenen Situationskontrolle als dritte Emotionsdimension an. Ob die wenigen Emotionsdimensionen es vermögen, die gesamte Bandbreite einer emotionalen Empfindung abzubilden ist fraglich und zählt zu den Kritikpunkten dieses Ansatzes. Im Gegensatz zu den dimensional Ansätzen geht der kategoriale Ansatz von disjunkten Grundemotionen – auch bekannt als Basisemotionen – aus. Am bekanntesten sind die von Ekman (1999) postulierten sechs Basisemotionen Freude, Traurigkeit, Ekel, Furcht, Wut und Überraschung. Kritisiert wird dieser Ansatz unter anderem wegen den unklaren Abgrenzungskriterien für eine Basisemotion sowie der stark schwankenden Anzahl an Emotionen (Ortony & Turner, 1990). Als Mischform dimensionaler und kategorialer Ansätze gilt der Ansatz von Plutchik (1980). Nach Plutchik (1980) setzen sich Emotionen analog eines Farbkreismodells aus Basisdimensionen zusammen. Dabei nimmt Plutchik die Existenz von acht Primäremotionen an, die gemäß ihrer Ähnlichkeit näher oder weiter voneinander entfernt sind. Aus der Mischung zweier Primäremotionen entstehen je nach Entfernung der Primäremotionen sogenannte primäre, sekundäre und tertiäre Dyaden (z. B. Scham aus den Emotionen Ekel und Furcht).

Mit dem Einzug der „kognitiven Wende“ rückte die Rolle der Kognition bei der Emotionsentstehung in den Vordergrund (vgl. Rolls, 2000). In Hinblick auf die Wirkreihenfolge kognitiver und physiologischer Vorgänge bei der Emotionsentstehung und beim Emotionserleben entstanden zwei theoretische Richtungen: Theorien, welche die kognitiven Anteile der Emotionen betonen und Theorien, welche den physiologischen Anteil betonen. Während kognitive Emotionstheorien versuchen, kognitive Muster bei Auslösereizen und -situationen ausfindig zu machen (z. B. Ich habe Angst, weil die Situation mich gefährdet), widmen sich die physiologischen Emotionstheorien den reinen körperlichen Empfindungen und biologischen Grundprozessen (z. B. Ich habe Angst, weil mein Herz rast). Letztere Theorien gehen auf James (1884) zurück, der die Entstehung von

Emotionen auf die Wahrnehmung körperlicher Veränderungen zurückführte. Dass dieses nicht ausreichte, um die Vielfalt an Emotionen zu unterscheiden, betonten Cannon (1927) und später auch Schachter und Singer (1962) in ihren kognitiven Emotionstheorien. Viele weitere Theorien folgten und erweiterten die vorangegangenen Modellannahmen durch neue Erkenntnisse. Heute wissen wir, dass sowohl die Kognition als auch die Physiologie maßgebend an der Emotionsentstehung beteiligt ist und diese begleitet (siehe Cacioppo, Berntson & Klein, 1992). Auch die kognitive Komponente der Emotionen kann subjektiv erfasst werden. Hierunter fallen unterschiedliche theoretische Strömungen, wie die Netzwerk-Theorien (z. B. Bower, 1981) und die Appraisal-Theorien (z. B. Lazarus, 1991).

A.2 Güte psychophysiologischer Parameter

Um die Frage zu beantworten, wie gut sich die einzelnen physiologischen Parameter zur Erfassung der mentalen Beanspruchung eignen, bedienen sich unterschiedliche Forschungsbereiche festgelegter Bewertungskriterien. Subjektive, leistungsbezogene und physiologische Maße mentaler Beanspruchung werden üblicherweise mit Hilfe von Gütekriterien bewertet, die spezifisch für die Beanspruchungsmessung zusammengetragen wurden (Matthews et al., 2015). Sie werden als anwendungsorientierte Kriterien angesehen. Hierzu kann ein breites Spektrum an Kriterien aus der Literatur herangezogen werden (siehe EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Kramer, 1991; Manzey, 1998; O'Donnell & Eggemeier, 1986; Schlick et al., 2010; Tsang & Wilson, 1997): u. a. die Benutzerakzeptanz, die Gebrauchstauglichkeit (Effektivität, Effizienz und Zufriedenheit), die Breite der Anwendbarkeit oder die Praktikabilität, die Störanfälligkeit eines Verfahrens und die Intrusion bzw. Inferenzfreiheit als das Ausmaß, indem das Erhebungsverfahren die Bearbeitung einer Aufgabe stört. Besonders hervorzuheben sind die Gütekrите-

rien Sensitivität und Diagnostizität (bzw. Spezifität)¹ (EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Kramer, 1991; Manzey, 1998; O'Donnell & Eggemeier, 1986; Schlick et al., 2010; Tsang & Wilson, 1997). Auf letztere wird weiter unten vertieft eingegangen.

Seltener werden die allgemeineren, messtheoretischen Gütekriterien aus der Psychometrie zur Bewertung physiologischer Maße verwendet (Matthews et al., 2015). Zu diesen zählt u. a. die Objektivität, Reliabilität und Validität. Die Psychometrie erfüllt in der psychophysiologischen Forschung, ebenso wie in anderen Forschungsbereichen, einen fundamentalen Auftrag. Sie zeigt auf, wie wahre Werte aus den messfehlerbehafteten Beobachtungswerten identifiziert, geschätzt und maximiert werden können. Reliabilität und Validität legen den Grundstein, um von einem manifesten Messwert (hier: physiologischen Parameter) auf das dahinterliegende latente Konstrukt, wie die mentale Beanspruchung, schließen zu können. Es ist anzumerken, dass anwendungsorientierte und psychometrische Gütekriterien nicht als unabhängige Anforderungen, sondern als sich überschneidende und ergänzende Bewertungsmerkmale zu verstehen sind. Mithin können Sensitivität und Diagnostizität als ein Teilaspekt der Validität verstanden werden (vgl. Matthews et al., 2015). Im Folgenden werden die Kriterien Sensitivität, Diagnostizität, Reliabilität sowie Validität definiert und in Bezug zur physiologischen Erfassung mentalen Beanspruchung gesetzt.

A.2.1 Sensitivität

Sensitivität bezeichnet die Messempfindlichkeit des Messverfahrens zwischen unterschiedlichen Ausprägungsstärken mentaler Beanspruchung zu differenzieren (EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Manzey,

¹ Die Gütekriterien Sensitivität und Diagnostizität (bzw. Spezifität) sind keine klassischen Gütekriterien, wie man diese in der testtheoretischen Literatur findet, sondern Bewertungskriterien aus der psychophysiologischen Forschung. Beide Kriterien beziehen sich auf die analytische Sensitivität und Spezifität und sind nicht synonym mit der Spezifität und Sensitivität im Sinne der diagnostischen Entscheidungsanalyse (diagnostische Sensitivität und Spezifität) oder der Signalentdeckungstheorie zu verstehen.

1998; O'Donnell & Eggemeier, 1986; Kramer, 1991; Schlick et al., 2010 Tsang & Wilson, 1997). Zur Bestimmung der Sensitivität wird üblicherweise die Schwierigkeit bzw. Komplexität der Aufgaben variiert und untersucht, inwieweit das Messverfahren im Stande ist, die Nuancen abzubilden. Ein sensitives Erfassungsverfahren soll dabei nicht nur die Schwierigkeitsunterschiede zwischen formal ähnlichen Aufgaben, sondern auch zwischen formal unterschiedlichen Aufgaben wiedergeben. Neben dem Intensitätsaspekt mentaler Beanspruchung können auch weitere Sensitivitätsaspekte, wie bspw. die temporale Sensitivität (Kramer, 1991), betrachtet werden. Diese bezieht sich auf den Grad eines Messverfahrens zeitnah auf Veränderungen der mentalen Beanspruchung reagieren zu können.

A.2.2 Diagnostizität bzw. Spezifität

Diagnostizität bzw. analytische Spezifität beschreibt den Grad mit dem ein Messinstrument im Stande ist, zwischen den unterschiedlichen Quellen und Typen mentaler Beanspruchung zu unterscheiden (EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Kramer, 1991; Manzey, 1998; O'Donnell & Eggemeier, 1986; Tsang & Wilson, 1997). Diagnostizität bezieht sich im weiteren Sinne auf die Abgrenzung mentaler, emotionaler und körperlicher Beanspruchung und im engeren Sinne auf die Fähigkeit zur Differenzierung verschiedener Arten mentaler Beanspruchung (Manzey, 1998). Zu beachten ist das antagonistische Zusammenspiel zwischen Sensitivität und Diagnostizität: Eine sehr hohe Diagnostizität (im engeren Sinne) bedeutet, dass das Messverfahren selektiv auf bestimmte (z. B. reaktionsbezogene vs. perzeptive) Aspekte der Aufgabe reagiert, jedoch andere Aspekte, wie z. B. die Intensität der Reaktion nicht mit abbildet. Ob jetzt die Sensitivität oder Diagnostizität im Fokus der Erfassung mentaler Beanspruchung stehen soll, ist vom Untersuchungsgegenstand und -ziel abhängig. Im Kontext der Mensch-Maschine-Interaktion spielt Diagnostizität vor allem dann eine wichtige Rolle, wenn bspw. auf die Ursache einer Leistungseinbuße geschlossen werden soll, sodass eine entsprechende

Intervention abgeleitet werden kann (Nickel & Nachreiner, 2003; Tsang & Wilson, 1997).

A.2.3 Reliabilität

Reliabilität beschreibt die Zuverlässigkeit eines Verfahrens bei der Messung mentaler Beanspruchung. Unter Annahme der klassischen Testtheorie (KTT) kann die Stabilität (identische Ergebnisse zwischen wiederholten Messungen), die Homogenität (Einzelparаметer oder -items sind auf das gleiche Merkmal zurückzuführen) und die Konsistenz (identische Ergebnisse zwischen mind. zwei parallelen Formen oder Hälften eines Verfahrens) der Messergebnisse zur Bewertung der Zuverlässigkeit verwendet werden (EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Tsang & Wilson, 1997). Im Kontext psychophysiologischer Messungen hat die KTT zwei entscheidende Limitationen: Zum einen werden Fehler als eine reine Zufallsvariable und zum anderen als eine Einheit definiert (Vertiefung siehe Strube et al., 2017, S. 613–614). Diese Einschränkungen hat nach Strube et al. (2017) die Generalisierbarkeitstheorie (GT) nicht, da sie die Definition multipler Fehlerquellen zulässt und somit die Multideterminiertheit eines Messwerts berücksichtigt (Strube et al., 2017). Somit ist es möglich, verschiedenen Fehlerquellen bzw. Einflussfaktoren, angefangen von Zeit- oder Situationsvariablen bis hin zu Personenvariablen, voneinander zu trennen und abzubilden (zsf. Strube et al., 2017, S. 614–621). Diese Vorgehensweise greift bereits Teilaspekte der Validität auf, weshalb die GT auch als eine Verschmelzung von Reliabilität und Validität aufgefasst wird (Strube et al., 2017).

A.2.4 Validität

Validität bezieht sich auf die Gültigkeit eines Messverfahrens (EN ISO 10075-3, DIN Deutsches Institut für Normung e. V., 2004; Tsang & Wilson, 1997). Einfacher ausgedrückt, stellt die Validität die Angemessenheit des Labels, welches

einem Messverfahren zugeschrieben wird, dar (Strube et al., 2017). Die Validität ist dann erfüllt, wenn ein physiologisches Messverfahren das zu messende latente Konstrukt – in der vorliegenden Arbeit die mentale Beanspruchung – hinreichend genau misst (Wirtz & Strohmer, 2014). Schlussendlich hängt von der Validität die Qualität der Schlussfolgerungen ab, die aus den Ergebnissen gezogen werden können. Üblicherweise wird die Validität in unterschiedliche Formen unterteilt (z. B. Wirtz & Strohmer, 2014), die je nach Anwendungs- und Interpretationskontext unterschiedliche Schlussfolgerungen zulassen (zsf. Strube et al., 2017, S. 621–625).

Um die Validität nachzuweisen, kann untersucht werden, ob die mit einem bestimmten Verfahren gemessene mentale Beanspruchung ausschließlich auf die Variation der mentalen Beanspruchung zurückzuführen ist und nicht durch andere Faktoren, wie z. B. Emotionen, mitbestimmt wird. Dieses Beispiel erinnert an die oben beschriebenen Kriterien Sensitivität und Diagnostizität: Ein Maß reagiert auf Veränderungen der mentalen Beanspruchung (Sensitivität), jedoch nicht auf Veränderung anderer Konstrukte (Diagnostizität). Ob ein Nachweis der Sensitivität und Diagnostizität als hinreichende Validität zu werten ist, hängt allerdings vom Umfang der Analyse ab. Validität in seiner höchsten Form versteht sich nämlich als nomologisches Netzwerk (Konstruktvalidität) ähnlicher, aber auch artfremder Konstrukte. Zudem sollte die interne (latenten) Struktur des Messverfahrens die theoretischen Annahmen des Konstrukts widerspiegeln. Bei niedriger Validität ist zu prüfen, ob die Erfassungsmethode oder das Konstrukt selbst (z. B. die mentale Beanspruchung) revidiert werden muss. Muss z. B. das Konstrukt der mentalen Beanspruchung revidiert werden, so ist auch die inhaltliche Validität als die Voraussetzung der Konstruktvalidität zu prüfen. Diese Voraussetzung bringt folgende Prämisse mit sich, die im Rahmen der Forschung zur mentalen Beanspruchung nicht immer Beachtung findet: Zur Validierung sollte ein neues Verfahren in eine bereits bewährte Theorie eingebettet werden; und umgekehrt sollte eine neue Theorie nur mit bereits validierten Messverfahren überprüft werden. Ansonsten besteht die Gefahr einer Zirkularität in der Form, dass die Theorie als Grundlage

zur Validierung eines Messverfahrens dient, welches später die Vorhersagen derselben Theorie überprüft. Eine Vertiefung sowie Lösungsmöglichkeiten findet sich bei Richter und Slade (2017).

Formen der Validität

Die *inhaltliche Validität* beschäftigt sich mit der Frage, ob das Verfahren das Wesen bzw. den Inhaltsraum des Konstrukts (mentale Beanspruchung) adäquat widerspiegelt und somit als repräsentativ interpretiert werden darf (Wirtz & Strohmmer, 2014). Im Vordergrund steht die Überführbarkeit der theoretischen Annahmen zur mentalen Beanspruchung in das Messverfahren. Da die Inhaltliche Validität in der Regel nicht numerisch bestimmt, sondern theoriegeleitet begründet wird, ist der Nachweis für physiologische Messverfahren nicht trivial. Aber auch die unterschiedlichen theoretischen Modelle mentaler Beanspruchung stellen eine Herausforderung für den Nachweis inhaltlicher Validität dar.

Die *Konstruktvalidität* gibt an, inwieweit ein Messverfahren das Konstrukt der mentalen Beanspruchung so abbildet, dass es mit den vorliegenden Theorien und Definitionen der mentalen Beanspruchung übereinstimmt und somit theoriebasiert interpretiert werden kann (Wirtz & Strohmmer, 2014). Bspw. kann überprüft werden, inwieweit physiologische Maße in der Lage sind, die unterschiedlichen Ressourcen nach Wickens (2008) abzubilden. Zumeist wird versucht, das mit dem Verfahren zu erfassende Zielkonstrukt in ein logisches oder theoriebasiertes Netzwerk (nomologisches Netzwerk) von ähnlichen, aber auch „artfremden“ Konstrukten einzubinden. Diese Form der Validität setzt somit ein fundiertes Wissen über das vorliegende Konstrukt sowie die Überprüfbarkeit bestehender Theorien voraus. Zur empirischen Überprüfung der Konstruktvalidität bedient man sich der konvergenten und diskriminanten Validität. Ersteres ist gegeben, wenn ein bedeutsamer Zusammenhang zwischen den Messergebnissen des zu prüfenden Verfahrens und den Messergebnissen eines anderen (am besten bereits validierten) Messverfahrens vorliegt. Im Gegensatz dazu, wird bei der diskriminanten Validität eine geringe Ergebnisübereinstimmung mit Verfahren,

die ein anderes Konstrukt abbilden, erwartet (z. B. die statistische Abgrenzung mentaler und emotionaler Beanspruchung). Zur statistischen Überprüfung der Konstruktvalidität kann die Multitrait-Multimethod-Analyse (MTMM-Analyse; Campbell & Fiske, 1959) genutzt werden. Diese prüft gleichzeitige die konvergente und diskriminante Validität unter Berücksichtigung der Methodenvarianz. Zudem kann die Faktorenanalyse zur Prüfung der faktoriellen Validität verwendet werden. Die faktorielle Validität stellt zwar keine hinreichende Bedingung für die Konstruktvalidität dar, sie ermöglicht aber sowohl ein theoriegetriebenes-deduktives Vorgehen, als auch ein empirisch-exploratives Vorgehen und kann somit dazu beitragen, bestehende Ansätze der mentalen Beanspruchung durch Faktoren zu ergänzen oder zu korrigieren. Nach Tsang und Wilson (1997) ist die Bestimmung der Konstruktvalidität für Messmethoden mentaler Beanspruchung ein äußerst schwieriges Unterfangen, weshalb oft auf den Nachweis der Kriteriumsvalidität, zurückgegriffen wird (Tsang & Wilson, 1997).

Die *Kriteriumsvalidität* betrachtet den Zusammenhang zwischen Messergebnissen und einem oder mehreren Außenkriterien, von denen man annimmt, dass sie mit dem Zielkonstrukt zusammenhängen (Wirtz & Strohmeyer, 2014). Im Gegensatz zur Konstruktvalidität spielt die theoretische Grundlage des Konstrukts nur eine zweitrangige Rolle. Hierbei unterscheidet man zwischen der Übereinstimmungsvalidität, bei der das zu validierende Messinstrument und das Kriterium gleichzeitig erhoben werden, und der Vorhersagevalidität, bei der das Kriterium in der Zukunft erfasst wird (Wirtz & Strohmeyer, 2014). Bspw. kann ein Nachweis angestrebt werden, dass ein physiologisches Messverfahren erfolgreich die Leistung in einer Testaufgabe vorhersagt (analog zur konvergenten Konstruktvalidität²). Eine andere Möglichkeit ist die (quasi-) experimentelle Überprüfung von Vorhersagen durch die Variation der Ausführungsbedingungen hinsichtlich bekannter Wirkzusammenhänge, wie z. B. der Aufgabenschwierigkeit.

² Die Bestimmung der Konstruktvalidität alleine auf der Grundlage der Kriteriumsvalidität birgt die Gefahr in einem Zirkelschluss zu enden, da anstelle von nomologischen Modellen paarweise Vergleiche herangezogen werden (Richter & Slade, 2017).

A.3 Psychometrische Reliabilität und Validität des NASA-TLX

Xiao, Wang, Wang und Lan (2005) berichten eine zufriedenstellende Reliabilität mit Korrelationswerten zwischen .52 und .75 für die Retest-Reliabilität und größer .80 für die Split-Half-Reliabilität, sowie ein Cronbach's Alpha größer .80. Auch die konvergente sowie konkurrente Validität wurde in mehreren Studien bestätigt (Hart & Staveland, 1988; Rubio, Diaz, Martin & Puente, 2004s; zsf. Stanton, 2005, S. 305). In Bezug auf die Sensitivität zeigt sich das NASA-TLX anderen Messverfahren der mentalen Beanspruchung gleichwertig bis überlegen (Rubio et al., 2004). Zur Strukturvalidität des NASA-TLX finden sich nur vereinzelte Studien. Galy, Paxion und Berthelon (2018) identifizieren in ihrer Fahrsimulationsstudie vier verschiedene Dimensionen im NASA-TLX. Diese stimmen überwiegend mit den postulierten Faktorkategorien von Hart und Staveland (1988) überein. Die erste Dimension ist die Komplexität der Situation bzw. der Aufgabe. Diese ist äquivalent zu Faktorkategorie (a) bei Hart und Staveland (1988). Die Komplexität bestimmt die Ausprägung der mentalen, körperlichen und zeitlichen Anforderung im NASA-TLX. Die Skala Frustration wird durch die zweite Dimension, die Anspannung der Person, determiniert. Diese ist äquivalent zur Faktorkategorie (c) bei Hart und Staveland (1988). Im Gegensatz zu Hart und Staveland (1988), ordnen Galy et al. (2018) den Skalen Anstrengung und Ausführung der Aufgabe unterschiedliche Dimensionen zu. Die Anstrengung repräsentiert die Dimension der kognitiven Ressourcen und ergibt sich aus der Interaktion der ersten beiden Dimensionen. Die Skala Ausführung der Aufgabe wird hingegen als Korrelat des Situationsbewusstseins einer vierten Dimension zugeordnet (Galy et al., 2018).

A.4 Voraussetzungen der Mehrebenenanalyse

Zu den Strukturvoraussetzungen der MEA zählt ein gegebener linearer Zusammenhang zwischen UV und AV. Dieser ist für alle Zufallseffekte, d. h. für

jede Person sowie für die festen Effekte zwischen den Prädiktoren und den individuellen Regressionsparametern zu prüfen. Für kategoriale Prädiktoren (z. B. Dummyvariablen) entfällt diese Überprüfung (siehe Singer & Willett, 2003, S. 129). Weitere Annahmen sind die Normalverteilung und Homoskedaschizität der Residuen der festen und zufälligen Effekte (Ditton, 1998; Hox, 2010; Singer & Willett, 2003). Sollten die Residuen nicht normalverteilt sein, so ist es möglich ein Modell mit einer anderen statistische Verteilungsform (z. B. logistisch oder binomial) zu spezifizieren. Zuletzt ist die Voraussetzung der Unabhängigkeit der Prädiktoren von den Residuen – innerhalb der fixen und der zufälligen Effekte – zu prüfen (Ditton, 1998, S 118f).

A.5 Unterschiede zwischen einer PARAFAC und einer generalisierten PCA

Zum einen werden bei der PARAFAC die Komponenten parallel und nicht wie bei der PCA sequentiell extrahiert. Deshalb sind die extrahierten Faktoren der PARAFAC nicht zwangsläufig orthogonal zueinander. Zweitens gibt es bei der PARAFAC keine Rotationsfreiheit, sodass im Optimalfall die Komponentenlösung einzigartig ist.³ Drittens ist die PARAFAC aufgrund der geringeren Freiheitsgrade weniger flexibel. Allerdings ist die PARAFAC im Gegensatz zur PCA weniger anfällig für die Extraktion von Komponenten, die auf Rauschvarianz zurückzuführen sind. Folglich ist der Anteil der erklärten Varianz in einer PARAFAC zwar geringer als in einer – von drei auf zwei Dimensionen entfaltenen – PCA, dafür ist das Ergebnis jedoch angemessener (Bro, 1997). Eine höhere Angemessenheit hat das Potential einer höheren Vorhersagekraft.

³ Bei der PCA ist die Faktorenlösung nicht einzigartig, da es möglich ist, die Faktoren ohne einen Verlust an Varianzaufklärung zu rotieren.

A.6 Anhänge zu Studie I

A.6.1 Deskriptive Statistik der kardialen Parameter

Tabelle A-1: Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie I

Parameter	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
HR	422	80.45	13.78	51.77	132.15	0.65	0.46
SDHR	422	4.62	2.70	1.30	31.50	3.60	25.84
HP	422	771.22	127.73	456.22	$1.16 \cdot 10^3$	0.24	-0.21
SDNN	422	48.31	22.17	11.15	228.85	2.12	11.20
MQSD	422	35.23	22.97	5.41	195.98	2.50	9.91
pNN50	422	13.43	15.05	0.00	62.26	1.27	0.71
pNN20	422	42.73	21.82	0.00	86.21	-0.17	-0.88
CVSD	422	0.04	0.03	0.01	0.28	3.77	21.60
CVNN	422	0.06	0.03	0.02	0.31	2.83	17.45
Triang	422	$6.15 \cdot 10^3$	$1.91 \cdot 10^3$	$2.05 \cdot 10^3$	$1.97 \cdot 10^4$	1.19	5.51
Shannon.h	422	4.65	0.55	2.81	5.61	-0.65	0.53
EKG.TP	422	$5.46 \cdot 10^5$	$8.11 \cdot 10^5$	$2.04 \cdot 10^4$	$9.26 \cdot 10^6$	6.69	57.98
EKG.LF	422	$21.1 \cdot 10^5$	$3.16 \cdot 10^5$	$7.99 \cdot 10^3$	$3.64 \cdot 10^6$	6.26	51.82
EKG.HF	422	$1.29 \cdot 10^5$	$2.29 \cdot 10^5$	$3.38 \cdot 10^3$	$3.38 \cdot 10^6$	8.36	101.43
EKG.LFn	422	63.10	17.17	11.94	94.05	-0.48	-0.38
EKG.LF/HF	422	2.57	2.27	0.14	15.81	2.08	5.43
DFA1	422	1.21	0.28	0.37	1.77	-0.46	-0.23
SamEn	412	1.55	0.49	0.36	3.00	0.40	0.20
SVDEn	422	0.15	0.06	0.05	0.54	2.26	8.73
T.amp	407	26.13	69.65	0.11	647.70	5.45	39.26
T.area	407	611.20	$1.63 \cdot 10^3$	2.07	$1.46 \cdot 10^4$	5.23	35.12

Die deskriptiven Statistiken der kardialen Parameter sind in Tabelle A-1 für die 10 Bedingungen dargestellt. Sechs Personen hatten aufgrund fehlerhafter Aufnahmen oder Sensorausfällen fehlende Werte und wurden aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlten pro Parameter

durchschnittlich 4.52 % (Min: 4.1 %; Max: 7.5 %) der Messpunkte. Diese sind hauptsächlich auf den Ausschluss von Daten aufgrund von Artefakten im EKG-Signal zurückzuführen. Ausreißer ($\pm 3 SD$) gab es keine.

A.6.2 Deskriptive Statistik der elektrodermalen Parameter

Tabelle A-2: Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie I

Parameter	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
EDL.M	474	10.73	5.90	1.14	27.39	0.54	-0.33
EDL.D	463	0.15	0.80	-2.81	3.23	0.40	1.50
Phasic.SD	470	0.37	0.35	$2.31 \cdot 10^{-4}$	1.64	1.23	0.98
EDR.N	474	7.55	3.77	0.00	18.00	0.21	-0.17
EDR.amp	458	0.34	0.31	0.01	1.35	1.31	1.03
EDR.rise.t	454	2.29	0.71	1.15	5.55	1.07	1.31
EDR.rec.t/2	459	1.95	0.84	0.61	6.62	1.12	1.99
EDR.rec.t/c	461	2.43	1.39	0.71	20.55	5.36	62.10
EDR.area	454	59.95	53.13	1.38	387.05	1.61	3.83
EDR.rise.v	458	0.06	0.06	$8.61 \cdot 10^{-4}$	0.32	1.59	2.36
EDR.peaks	442	8.29	4.19	2.52	38.85	2.39	10.11
EDA.LF	462	$2.67 \cdot 10^{-2}$	$4.32 \cdot 10^{-2}$	$3.24 \cdot 10^{-6}$	$2.54 \cdot 10^{-1}$	2.60	7.40
EDA.HF	470	$1.68 \cdot 10^{-2}$	$3.54 \cdot 10^{-2}$	$1.49 \cdot 10^{-7}$	$2.55 \cdot 10^{-1}$	3.33	12.91
EDA.VHF	462	$3.85 \cdot 10^{-4}$	$9.86 \cdot 10^{-4}$	$7.75 \cdot 10^{-9}$	$7.72 \cdot 10^{-3}$	3.98	17.60

Die deskriptiven Statistiken der elektrodermalen Parameter sind in Tabelle A-2 für die 10 Bedingungen dargestellt. Zwei Personen hatten aufgrund fehlerhafter Aufnahmen oder Sensorausfällen fehlende Werte und wurden aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlten pro Parameter durchschnittlich 2.45 % der Messpunkte (Min: 1.04 %; Max: 7.08 %). Diese kommen durch Ausschluss von Ausreißern ($\pm 3 SD$; insgesamt 1,44 %) oder Artefakten im EDA-Rohsignal zustande.

A.6.3 Deskriptive Statistik der Kriteriumsvariablen

Tabelle A-3: Deskriptive Statistiken des NASA-RTLX jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B1	50	11.14	2.73	4.00	16.40	-0.05	-0.35
B2	50	15.64	3.11	7.20	20.00	-0.66	-0.36
B1+	50	9.64	3.17	1.00	16.40	-0.53	0.49
B2+	50	12.08	2.83	5.20	18.00	-0.29	-0.25
B1-	50	14.00	3.45	4.80	19.20	-0.27	-0.65
B2-	50	15.56	3.09	7.20	20.00	-0.70	-0.13
EnV+	50	2.66	1.61	1.00	7.40	1.04	0.24
EhV+	50	3.57	2.08	1.00	7.80	0.54	-0.83
EnV-	50	3.83	1.81	1.00	8.20	0.29	-0.53
EhV-	50	5.29	3.54	1.00	17.00	1.42	1.89
Alle	500	9.34	5.60	1.00	20.00	0.11	-1.23

Tabelle A-4: Deskriptive Statistiken der Trefferquote jeweils für die 6 Bedingungen der Balltracking-aufgabe sowie über alle Bedingungen

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B1	50	46.25	23.72	0.00	87.50	0.18	-0.87
B2	50	28.00	23.50	0.00	87.50	0.42	-0.76
B1+	50	40.25	28.60	0.00	100.00	0.50	-0.66
B2+	50	29.25	21.37	0.00	87.50	0.48	-0.12
B1-	50	18.25	20.24	0.00	62.50	0.86	-0.46
B2-	50	11.25	17.17	0.00	75.00	1.75	2.71
Alle	300	28.88	25.50	0.00	100.00	0.68	-0.30

Tabelle A-5: Deskriptive Statistiken der Erregung jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B1	50	2.28	0.88	1.00	5.00	0.49	0.34
B2	50	2.86	0.95	1.00	5.00	-0.15	-0.62
B1+	50	2.16	0.79	1.00	4.00	-0.04	-0.97
B2+	50	2.46	0.84	1.00	5.00	0.63	0.44
B1-	50	2.86	0.95	1.00	5.00	0.13	-0.45
B2-	50	2.92	0.83	1.00	5.00	0.35	0.05
EnV+	50	1.78	0.79	1.00	4.00	0.64	-0.44
EhV+	50	2.74	0.88	1.00	5.00	0.16	-0.32
EnV-	50	2.50	0.95	1.00	5.00	0.28	-0.40
EhV-	50	3.18	1.04	1.00	5.00	0.17	-0.95
Alle	500	2.57	0.97	1.00	5.00	0.29	-0.28

Tabelle A-6: Deskriptive Statistiken der Valenz jeweils für die 10 Bedingungen sowie über alle Bedingungen der Studie I

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B1	50	3.64	0.94	1.00	5.00	-0.40	-0.18
B2	50	3.24	1.06	2.00	5.00	0.33	-1.17
B1+	50	3.82	0.90	2.00	5.00	-0.32	-0.74
B2+	50	3.82	0.75	2.00	5.00	-0.57	0.27
B1-	50	2.84	0.98	1.00	5.00	0.31	-0.31
B2-	50	2.76	0.98	1.00	5.00	0.35	-0.63
EnV+	50	4.00	0.73	3.00	5.00	0.00	-1.15
EhV+	50	4.00	0.86	2.00	5.00	-0.38	-0.78
EnV-	50	3.02	0.96	1.00	5.00	0.37	-0.54
EhV-	50	2.50	1.15	1.00	5.00	0.40	-0.63
Alle	500	3.36	1.07	1.00	5.00	-0.19	-0.72

A.6.4 Bedingungs-Faktorladungen der NPLS-Regression

Tabelle A-7: Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter

Bedingung	Kardiale Komponente	NASA-RTLX-Komponente
S1	0.18	0.11
S2	0.24	0.42
S1+	0.22	0.02
S2+	0.35	0.18
S1-	0.21	0.30
S2-	0.31	0.40
EnV+	-0.35	-0.44
EhV+	-0.38	-0.37
EnV-	-0.41	-0.35
EhV-	-0.42	-0.27

Tabelle A-8: Bedingungs-Faktorladungen der Trefferquoten- und der kardialen Komponente aus der NPLS-Regression der Trefferquote auf kardiale Parameter

Bedingung	Kardiale Komponente	Trefferquoten-Komponente
S1	-0.51	0.59
S2	0.29	-0.20
S1+	-0.57	0.46
S2+	0.01	0.04
S1-	0.26	-0.36
S2-	0.51	-0.52

Tabelle A-9: Bedingungs-Faktorladungen der Erregungs- und der kardialen Komponente aus der NPLS-Regression der Erregung auf kardiale Parameter

Bedingung	Kardiale Komponente	Erregungs-Komponente
S1	-0.22	-0.22
S2	0.06	0.14
S1+	-0.09	-0.23
S2+	0.10	-0.02
S1-	0.01	0.23
S2-	0.21	0.24
EnV+	-0.83	-0.72
EhV+	0.05	0.08
EnV-	0.16	0.00
EhV-	0.42	0.49

Tabelle A-10: Bedingungs-Faktorladungen der Valenz- und der kardialen Komponente aus der NPLS-Regression der Valenz auf kardiale Parameter

Bedingung	Kardiale Komponente	Valenz-Komponente
S1	-0.17	0.13
S2	0.21	-0.17
S1+	-0.05	0.27
S2+	-0.01	0.20
S1-	0.10	-0.30
S2-	0.33	-0.35
EnV+	-0.68	0.45
EhV+	-0.41	0.43
EnV-	0.22	-0.17
EhV-	0.35	-0.47

Tabelle A-11: Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	NASA-RTLX-Komponente
S1	0.26	0.12
S2	0.37	0.40
S1+	0.06	0.02
S2+	0.26	0.18
S1-	0.29	0.31
S2-	0.23	0.41
EnV+	-0.41	-0.44
EhV+	-0.35	-0.37
EnV-	-0.41	-0.36
EhV-	-0.36	-0.27

Tabelle A-12: Bedingungs-Faktorladungen der Trefferquoten- und der elektrodermalen Komponente aus der NPLS-Regression der Trefferquote auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	Trefferquoten-Komponente
S1	-0.53	0.57
S2	0.05	-0.06
S1+	-0.37	0.41
S2+	-0.19	0.08
S1-	0.56	-0.46
S2-	0.48	-0.54

Tabelle A-13: Bedingungs-Faktorladungen der Erregungs- und der elektrodermalen Komponente aus der NPLS-Regression der Erregung auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	Erregungs-Komponente
S1	-0.30	-0.34
S2	-0.16	-0.08
S1+	-0.30	-0.33
S2+	-0.34	-0.22
S1-	0.08	0.15
S2-	-0.05	0.06
EnV+	-0.06	-0.35
EhV+	0.19	0.22
EnV-	0.19	0.21
EhV-	0.77	0.69

Tabelle A-14: Bedingungs-Faktorladungen der Valenz- und der elektrodermalen Komponente aus der NPLS-Regression der Valenz auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	Valenz-Komponente
S1	-0.25	0.22
S2	-0.06	-0.03
S1+	-0.30	0.32
S2+	-0.36	0.29
S1-	0.11	-0.21
S2-	0.06	-0.25
EnV+	-0.05	0.25
EhV+	-0.13	0.33
EnV-	0.21	-0.27
EhV-	0.80	-0.65

A.7 Anhänge zu Studie II

A.7.1 Deskriptive Statistik der kardialen Parameter

Tabelle A-15: Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie II

Parameter	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
HR	586	77.51	12.20	52.86	115.76	0.55	0.16
SDHR	586	4.15	2.04	1.05	17.80	1.77	6.74
HP	586	796.49	122.83	518.82	1.13·10 ³	0.29	-0.05
SDNN	586	48.05	22.42	13.41	159.44	1.54	3.61
MQSD	586	35.58	20.53	3.67	143.53	1.53	2.76
pNN50	586	15.20	16.77	0.00	66.04	1.23	0.50
pNN20	586	47.41	20.94	0.00	89.09	-0.10	-0.78
CVSD	586	0.04	0.02	0.01	0.18	1.76	4.64
CVNN	586	0.06	0.02	0.02	0.18	1.47	3.49
Triang	586	6.21·10 ³	2.12·10 ³	1.66·10 ³	1.88·10 ⁴	1.38	3.87
Shannon.h	586	4.67	0.56	2.81	5.61	-0.28	-0.38
EKG.TP	586	1.28·10 ⁶	1.41·10 ⁶	7.72·10 ⁴	1.23·10 ⁷	3.47	16.89
EKG.LF	586	4.82·10 ⁵	6.16·10 ⁵	2.25·10 ⁴	6.68·10 ⁶	4.79	32.83
EKG.HF	586	3.50·10 ⁵	4.24·10 ⁵	7.17·10 ³	2.82·10 ⁶	2.62	8.04
EKG.LFn	586	59.45	16.83	9.77	91.54	-0.56	-0.06
EKG.LF/HF	586	2.02	1.59	0.11	10.83	2.06	6.04
DFA1	586	1.15	0.29	0.22	1.85	-0.55	0.38
SamEn	578	1.63	0.49	0.34	3.58	0.91	1.57
SVDEn	586	0.14	0.05	0.03	0.40	1.20	1.97
T.amp	586	307.31	209.18	2.35	1.09·10 ³	1.05	1.50
T.area	586	2.02·10 ⁴	1.51·10 ⁴	87.93	7.16·10 ⁴	1.03	0.81

Die deskriptiven Statistiken der kardialen Parameter sind in Tabelle A-15 für die 12 Bedingungen dargestellt. Eine Personen wurden komplett aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlten pro Parameter durchschnittlich 0.4 % (Min: 0.3 %; Max: 1.7 %) der Daten,

welche hauptsächlich auf den Ausschluss aufgrund von hohen Artefakten im EKG-Signal zurückzuführen sind. Ausreißer ($\pm 3 SD$) gab es keine.

A.7.2 Deskriptive Statistik der elektrodermalen Parameter

Tabelle A-16: Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie II

Parameter	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
EDL.M	573	2.93	0.76	1.37	5.35	0.27	0.00
EDL.D	573	-0.04	0.10	-0.78	0.54	-0.52	8.54
Phasic.SD	573	0.03	0.03	$3.34 \cdot 10^{-6}$	0.24	3.05	12.38
EDR.N	573	4.31	3.36	0.00	14.00	0.49	-0.61
EDR.amp	482	0.04	0.03	0.01	0.26	2.52	7.85
EDR.rise.t	483	1.89	0.53	1.10	5.81	2.37	9.62
EDR.rec.t/2	482	2.43	2.55	0.87	40.10	10.16	125.90
EDR.rec.t/c	480	3.08	2.58	0.74	36.10	7.75	79.03
EDR.area	482	77.21	72.95	10.52	736.27	3.91	23.29
EDR.rise.v	483	0.09	0.07	0.01	0.57	2.71	9.34
EDR.peaks	419	11.70	7.87	2.75	56.40	2.76	9.81
EDA.LF	572	$3.50 \cdot 10^{-4}$	$1.24 \cdot 10^{-3}$	$4.08 \cdot 10^{-8}$	0.02	9.85	135.34
EDA.HF	573	$1.65 \cdot 10^{-4}$	$5.41 \cdot 10^{-4}$	$5.50 \cdot 10^{-8}$	0.01	6.32	46.33
EDA.VHF	573	$5.31 \cdot 10^{-6}$	$2.10 \cdot 10^{-5}$	$2.28 \cdot 10^{-8}$	$2.62 \cdot 10^{-4}$	7.63	68.45

Die deskriptiven Statistiken der elektrodermalen Parameter sind in Tabelle A-16 für die 12 Bedingungen dargestellt. Eine Person wurde aufgrund einer fehlerhaften Aufnahme aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlen pro Parameter durchschnittlich 9.13 % der Daten. Der hohe Anteil fehlender Datenpunkte ist darauf zurückzuführen, dass bei einem relativ hohen Anteil der Personen keine EDRs detektiert werden konnten. So haben Parameter, die auf Grundlage von EDRs berechnet werden einen Anteil fehlender Datenpunkte von 17.51 % bis 27.26 %, wohingegen bei den restlichen Parameter anteilhaft weniger als 0.70 % der Daten fehlen. Als Ausreißer ($\pm 3 SD$) wurden insgesamt nur 0.09 % der Daten ausgeschlossen.

A.7.3 Deskriptive Statistik der Kriteriumsvariablen

Tabelle A-17: Deskriptive Statistiken des NASA-RTLX jeweils für die 12 Bedingungen sowie über alle Bedingungen der Studie II

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
S1z-	50	4.32	3.37	1.00	14.80	1.44	1.52
S1b-	50	4.29	3.06	1.00	14.20	1.36	1.73
S1p-	50	4.53	3.46	1.00	15.00	1.47	1.63
S1z+	50	4.08	2.87	1.00	13.40	1.26	1.52
S1b+	50	4.82	2.93	1.00	12.00	0.59	-0.59
S1p+	50	4.06	2.98	1.00	12.40	1.07	0.47
S2z-	50	13.68	2.91	5.20	18.60	-0.46	-0.39
S2b-	50	14.00	3.01	5.60	19.00	-0.60	0.04
S2p-	50	13.54	3.59	4.80	19.20	-0.30	-0.69
S2z+	50	14.71	3.27	7.00	19.60	-0.44	-0.54
S2b+	50	13.70	3.51	5.00	19.20	-0.33	-0.75
S2p+	50	14.65	2.97	8.40	19.80	-0.25	-0.83
Alle	600	9.20	5.79	1.00	19.80	0.09	-1.40

Tabelle A-18: Deskriptive Statistiken der Trefferquote jeweils für die 12 Bedingungen sowie über alle Bedingungen der Studie II

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
S1z-	50	99.28	2.88	88.00	100.00	-3.59	11.15
S1b-	50	99.76	1.70	88.00	100.00	-6.65	43.12
S1p-	50	97.80	5.94	75.00	100.00	-2.67	6.45
S1z+	50	99.52	2.38	88.00	100.00	-4.55	19.13
S1b+	50	95.40	6.93	75.00	100.00	-1.19	0.53
S1p+	50	99.52	2.38	88.00	100.00	-4.55	19.13
S2z-	50	77.26	16.01	25.00	100.00	-0.98	1.38
S2b-	50	69.22	16.24	38.00	100.00	-0.17	-0.91
S2p-	50	76.96	14.27	50.00	100.00	-0.04	-0.77
S2z+	50	66.00	14.91	38.00	100.00	0.07	-0.73
S2b+	50	72.78	16.14	38.00	100.00	-0.51	-0.31
S2p+	50	69.20	9.47	50.00	88.00	-0.34	-0.25
Alle	600	85.22	17.39	25.00	100.00	-0.92	-0.16

A.7.4 Bedingungs-Faktorladungen der NPLS-Regression

Tabelle A-19: Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter

Bedingung	Kardiale Komponente	NASA-RTLX-Komponente
S1z-	-0.29	-0.29
S1b-	-0.23	-0.29
S1p-	-0.20	-0.27
S1z+	-0.34	-0.31
S1b+	-0.21	-0.26
S1p+	-0.43	-0.30
S2z-	0.28	0.27
S2b-	0.27	0.28
S2p-	0.37	0.26
S2z+	0.23	0.33
S2b+	0.29	0.27
S2p+	0.22	0.32

Tabelle A-20: Bedingungs-Faktorladungen der Trefferquoten- und der kardialen Komponente aus der NPLS-Regression der Trefferquote auf kardiale Parameter

Bedingung	Kardiale Komponente	Trefferquoten-Komponente
S1z-	-0.32	0.30
S1b-	-0.24	0.31
S1p-	-0.21	0.26
S1z+	-0.34	0.30
S1b+	-0.17	0.21
S1p+	-0.42	0.30
S2z-	0.22	-0.17
S2b-	0.28	-0.33
S2p-	0.35	-0.20
S2z+	0.29	-0.40
S2b+	0.28	-0.24
S2p+	0.24	-0.34

Tabelle A-21: Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	NASA-RTLX-Komponente
S1z-	-0.35	-0.29
S1b-	-0.16	-0.29
S1p-	-0.25	-0.27
S1z+	-0.28	-0.31
S1b+	-0.29	-0.26
S1p+	-0.38	-0.31
S2z-	0.19	0.27
S2b-	0.16	0.27
S2p-	0.28	0.25
S2z+	0.30	0.33
S2b+	0.38	0.28
S2p+	0.34	0.33

Tabelle A-22: Bedingungs-Faktorladungen der Trefferquoten- und der elektrodermalen Komponente aus der NPLS-Regression der Trefferquote auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	Trefferquoten-Komponente
S1z-	-0.34	0.30
S1b-	-0.16	0.31
S1p-	-0.26	0.26
S1z+	-0.28	0.30
S1b+	-0.28	0.21
S1p+	-0.37	0.30
S2z-	0.14	-0.15
S2b-	0.19	-0.31
S2p-	0.25	-0.19
S2z+	0.33	-0.42
S2b+	0.38	-0.26
S2p+	0.36	-0.35

A.8 Anhänge zu Studie III

A.8.1 Deskriptive Statistik der kardialen Parameter

Tabelle A-23: Deskriptive Statistiken der kardialen Parameter über alle Bedingungen der Studie III

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
HR	1071	72.01	10.21	46.75	107.58	0.13	-0.05
SDHR	1066	3.63	1.54	0.82	10.43	0.98	1.16
HP	1071	853.85	126.07	559.29	$1.29 \cdot 10^3$	0.66	0.23
SDNN	1069	50.56	21.54	13.41	181.78	1.64	4.52
MQSD	1070	40.16	21.95	8.00	142.18	1.83	4.19
pNN50	1071	19.90	18.58	0.00	83.67	1.09	0.43
pNN20	1071	54.14	19.15	1.12	97.78	-0.20	-0.34
CVSD	1070	0.05	0.02	0.01	0.13	1.24	2.37
CVNN	1068	0.06	0.02	0.01	0.18	1.08	2.40
Triang	1069	$6.24 \cdot 10^3$	$1.89 \cdot 10^3$	$2.11 \cdot 10^3$	$1.66 \cdot 10^4$	0.81	1.54
Shannon.h	1071	4.74	0.51	2.81	5.61	-0.39	-0.15
EKG.TP	1068	$1.34 \cdot 10^6$	$1.39 \cdot 10^6$	$8.01 \cdot 10^4$	$1.58 \cdot 10^7$	3.90	23.00
EKG.LF	1067	$4.95 \cdot 10^3$	$5.52 \cdot 10^5$	$2.47 \cdot 10^4$	$5.38 \cdot 10^6$	3.80	20.40
EKG.HF	1070	$3.44 \cdot 10^5$	$4.44 \cdot 10^5$	$5.23 \cdot 10^3$	$3.69 \cdot 10^6$	3.57	16.15
EKG.LFn	1071	60.53	18.60	8.23	97.51	-0.45	-0.43
EKG.LF/HF	1066	2.37	2.33	0.09	17.28	2.69	9.94
DFA1	1071	1.11	0.28	0.28	1.83	-0.31	-0.23
SamEn	1041	1.72	0.49	0.40	3.50	0.67	0.58
SVDEn	1070	0.15	0.05	0.04	0.33	0.78	1.07
T.amp	1003	467.83	185.28	48.50	$1.02 \cdot 10^3$	0.59	-0.02
T.area	1003	$3.19 \cdot 10^4$	$1.39 \cdot 10^4$	$1.84 \cdot 10^3$	$7.26 \cdot 10^4$	0.47	-0.23

Die deskriptiven Statistiken der kardialen Parameter sind in Tabelle A-23 für die 20 Bedingungen dargestellt. Zwei Personen wurden aufgrund fehlender Daten aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlten pro Parameter durchschnittlich 1.5 % (Min: 0.8 %; Max: 7.1 %) der

Daten, welche hauptsächlich auf den Ausschluss durch Artefakten im EKG-Signal zurückzuführen sind. Zusätzlich wurden 0.1 % der Daten als Ausreißer ($\pm 3 SD$) identifiziert und aus der weiteren Analyse ausgeschlossen.

A.8.2 Deskriptive Statistik der elektrodermalen Parameter

Tabelle A-24: Deskriptive Statistiken der elektrodermalen Parameter über alle Bedingungen der Studie III

Parameter	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
EDL.M	1060	1.97	0.58	0.63	3.84	-0.11	-0.38
EDL.D	1055	-0.04	0.07	-0.34	0.36	-0.35	2.99
Phasic.SD	1060	0.01	0.01	$2.89 \cdot 10^{-6}$	0.08	1.71	3.81
EDR.N	1059	2.85	2.92	0.00	19.00	1.27	1.69
EDR.amp	790	0.03	0.02	0.01	0.20	2.68	11.97
EDR.rise.t	789	1.78	0.50	0.87	4.36	1.58	3.43
EDR.rec.t/2	781	2.52	2.25	0.60	40.21	10.61	148.80
EDR.rec.t/c	784	3.41	3.35	0.69	42.50	7.74	73.84
EDR.area	787	56.25	47.58	8.75	553.32	4.00	27.28
EDR.rise.v	793	0.07	0.04	0.01	0.32	1.49	3.45
EDR.peaks	612	14.00	9.35	1.06	55.72	1.90	4.00
EDA.LF	1052	$8.21 \cdot 10^{-5}$	$1.65 \cdot 10^{-4}$	$3.90 \cdot 10^{-8}$	$1.80 \cdot 10^{-3}$	5.23	39.94
EDA.HF	1054	$3.75 \cdot 10^{-5}$	$7.29 \cdot 10^{-5}$	$6.52 \cdot 10^{-8}$	$7.57 \cdot 10^{-4}$	4.18	23.87
EDA.VHF	1054	$1.52 \cdot 10^{-6}$	$2.62 \cdot 10^{-6}$	$2.31 \cdot 10^{-8}$	$2.62 \cdot 10^{-5}$	3.92	21.58

Die deskriptiven Statistiken der elektrodermalen Parameter sind in Tabelle A-24 für die 20 Bedingungen dargestellt. Drei Personen wurde aufgrund von Sensorausfällen aus der weiteren Analyse ausgeschlossen. Bei der verbleibenden Stichprobe fehlen pro Parameter durchschnittlich 13.83 % der Daten. Der hohe Anteil fehlender Datenpunkte ist darauf zurückzuführen, dass bei einem relativ hohen Anteil der Personen keine EDRs detektiert werden konnten. So haben Parameter, die auf Grundlage von EDRs berechnet werden einen Anteil fehlender Datenpunkte von 25.18 % bis 42.26 %, wohingegen bei den restlichen

Parameter anteilhaft weniger als 0.80 % der Daten fehlen. Als Ausreißer ($\pm 3 SD$) wurden insgesamt nur 0.39 % der Daten ausgeschlossen.

A.8.3 Deskriptive Statistik der Kriteriumsvariable

Tabelle A-25: Deskriptive Statistiken des NASA-RTLX jeweils für die 20 Bedingungen sowie über alle Bedingungen der Studie III

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B1	56	3.05	2.71	1.00	13.20	1.82	3.28
B2	56	4.97	3.17	1.00	14.20	0.85	-0.11
B3	56	9.73	4.54	2.20	19.20	0.39	-1.02
B4	56	11.36	4.05	2.20	19.40	-0.05	-0.62
G1	56	3.18	2.68	1.00	11.60	1.45	1.48
G2	56	5.36	3.25	1.00	13.40	0.61	-0.60
G3	56	10.49	3.90	1.00	19.40	-0.30	0.02
G4	56	14.21	3.50	4.60	20.00	-0.54	-0.01
K1	56	2.79	2.37	1.00	12.00	2.01	4.15
K2	56	4.47	2.77	1.00	12.00	0.94	0.15
K3	56	8.12	3.33	1.40	15.60	0.08	-0.51
K4	56	13.49	2.81	6.00	18.80	-0.45	-0.10
M1	56	2.74	2.40	1.00	12.00	1.91	3.43
M2	56	6.71	2.96	1.00	12.40	-0.10	-0.75
M3	56	12.30	3.44	6.20	19.60	0.02	-0.83
M4	56	15.04	3.12	6.20	20.00	-0.47	0.06
S1	56	2.98	2.71	1.00	11.60	1.81	2.61
S2	56	4.42	2.73	1.00	11.40	0.68	-0.49
S3	56	5.84	2.85	1.00	12.00	0.33	-0.93
S4	56	9.48	3.71	1.60	17.80	-0.09	-0.74
Alle	1120	7.54	5.16	1.00	20.00	0.47	-0.88

A.8.4 Analyse der leistungsbezogenen mentalen Beanspruchung

Dieses Kapitel beschreibt die Messung und Auswertung der leistungsbezogenen mentalen Beanspruchung im Experiment. Dabei werden zunächst die Hypothesen postuliert sowie die Methode zur Analyse dargestellt. Anschließend werden die Ergebnisse in berichtet und interpretiert.

Messung der leistungsbezogenen mentalen Beanspruchung

Die Leistungsmessung wurde für jede Bedingung durchgeführt.⁴ Berechnet wurde hierfür der Anteil der korrekten Antworten bzw. Reaktionen an der Gesamtanzahl möglich Antworten bzw. Reaktionen pro Bedingung (Trefferquote). Da in den Ruhebedingungen keine Leistung erfasst wurde, liegt diese nur für die Schwierigkeitsstufen 2 bis 4 vor. Bedingt durch die qualitativen Unterschiede der Aufgabentypen setzt sich das Leistungsmaß für jeden Aufgabentyp unterschiedlich zusammen. In der Balltrackingaufgabe diente der finale Punktestand, also die relative Anzahl der Treffer (Quotient aus Anzahl tatsächlicher Treffer und Anzahl möglicher Gesamttreffer), als Leistungsmaß. In der Gedächtnisaufgabe wurde die Leistung als der Anteil korrekt erinnerter Zahlenfolgen gemessen. Bei der Konzentrationsaufgabe wurden alle korrekten Reaktionen (also Reaktion auf Zielreize und keine Reaktion auf Nicht-Zielreize) als Anteil an den Gesamtreizen erfasst. In der Matrizenaufgabe wurde der Anteil der gelösten Matrizen und in der Stroopaufgabe der Anteil der korrekten Reaktion auf die Schriftfarbe berechnet.

⁴ Aufgrund der qualitativen und quantitativen Unterschiede zwischen den Aufgabentypen, diente die Leistungsmessung in Studie III lediglich der Manipulationskontrolle, jedoch nicht der Extraktion physiologischer Profile mentaler Beanspruchung.

Hypothesen und Auswertemethoden

Für die leistungsbezogene mentale Beanspruchung (Trefferquote) wird eine Steigung mit zunehmender Aufgabenschwierigkeit erwartet.

H3.3 Je höher die Aufgabenschwierigkeit, desto geringer ist die Trefferquote.

Mit Hilfe der MEA wird geprüft, ob die Schwierigkeitsstufen einen bedeutenden Effekt auf die Trefferquote haben. Die nominalen Faktorstufen der Schwierigkeit werden mit der Rückwärts-Differenz-Codierung dummykodiert. Hierbei wird der Kriteriumsmittelwert der einen Faktorstufe mit dem Kriteriumsmittelwert der benachbarten Faktorstufe verglichen. Es ist zu beachten, dass die Trefferquote erst ab der 2-ten Schwierigkeitsstufe erfasst wurde, weshalb der Faktor Aufgabenschwierigkeit in diesem Fall nur drei (2 bis 4) Faktorstufen hat. Insgesamt ergeben sich daraus zwei Dummyvariablen.

Des Weiteren wird untersucht, ob der Aufgabentyp einen Effekt auf die Trefferquote hat. Für die Trefferquote ist ein solcher Effekt aufgrund der quantitativen als auch qualitativen Unterschiede zwischen den Aufgabentypen zu erwarten. Durch die Verwendung der Primäraufgabenleistung ist anzunehmen, dass die Trefferquote nicht nur die leistungsbezogene mentale Beanspruchung, sondern auch die allgemeine Lösbarkeit der verschiedenen Aufgabentypen wiedergibt. Über die Richtung der Unterschiede können allerdings keine Annahmen getroffen werden. Folgende Hypothese wird postuliert:

H3.4 Die Trefferquote unterscheidet sich signifikant zwischen den Aufgabentypen.

Die Unterschiede zwischen den fünf Faktorstufen des Faktors Aufgabentyp (B, G, K, M und S) wird ebenfalls innerhalb der MEA geprüft. Hierfür wird die einfache Effektkodierung angewandt. Insgesamt ergeben sich hier fünf Dummyvariablen (eine für jeden Aufgabentyp).

Zusätzlich wird mit der Interaktion Aufgabenschwierigkeit \times Aufgabentyp geprüft, ob die Steigung der mentalen Beanspruchung (bei steigender Aufgabenschwierigkeit) sich zwischen den Aufgabentypen unterscheidet. Da keine

Annahmen zur Art der Interaktionen (Aufgabenschwierigkeit \times Aufgabentyp) gemacht werden können, wird eine Unterfragestellungen (UF) formuliert:

UF3.2 Gibt es einen signifikanten Interaktionseffekt zwischen der Schwierigkeit und dem Typ der Aufgabe auf die Trefferquote.

Zur Prüfung der Interaktion Aufgabenschwierigkeit \times Aufgabentyp werden aus den Dummyvariablen der Schwierigkeit und des Aufgabentyps Interaktionsvariablen gebildet. Insgesamt ergeben sich hierfür $2 \cdot 5$ Zweifachinteraktionen. Aufgrund der Vielzahl an Dummyvariablen wird eine Bonferroni-Korrektur zur Vermeidung der Alphafehlerkumulierung durchgeführt. Das korrigierte Alphaniveau liegt für die Hypothesentests der Regressionsgewichte bei 0.0025.

Ergebnisse der Mehrebenenanalyse

Die deskriptiven Statistiken der Trefferquote sind in Tabelle A-26 und in Abbildung A-1 dargestellt. Ausreißer ($\pm 3 SD$) oder fehlende Daten liegen nicht vor. Deskriptiv ist zu erkennen, dass es innerhalb jedes Aufgabentyps zwar einen erwartungskonformen Abfall der Trefferquote bei steigender Aufgabenschwierigkeit gibt, jedoch jeder Aufgabentyp sein spezifisches Leistungsniveau und Abfallmuster zeigt. Dies spiegelt die bereits erwähnten qualitativen und quantitativen Unterschiede durch die aufgabenspezifische Leistungserfassung wieder. Ob sich bedeutsame Effekte finden, wird mit der MEA überprüft.

Der Modellvergleich in Tabelle A-27 belegt einen signifikanten Haupteffekt der Aufgabenschwierigkeit, des Aufgabentyps sowie einen bedeutsamen Interaktionseffekt. Die Haupteffekt der Aufgabenschwierigkeit und des Aufgabentyps erklären mit jeweils 41.1 % und 31.1 % den Hauptanteil der Varianz in der Trefferquote. In Tabelle A-28 ist zu erkennen, dass bei jeder Steigung der Schwierigkeitsstufe die Trefferquote signifikant abfällt (Teststärken bei 100.0 %). Die Regressionsgewichte der Aufgabentypen zeigen auf, dass jede der fünf Aufgabentypen vom Durchschnitt abweicht. Am schlechtesten fällt die Trefferquote dabei für die Balltracking- und die Matrizenaufgabe aus. Im Gegensatz dazu liegt die Trefferquote für die Gedächtnis-, Konzentrations-

Tabelle A-26: Deskriptive Statistiken der Trefferquote jeweils für die 15 Bedingungen sowie über alle Bedingungen der Studie III

Bedingung	<i>N</i>	<i>M</i>	<i>SD</i>	Min	Max	Schiefe	Kurtosis
B2	56	92.65	6.19	56.30	97.42	-3.86	18.88
B3	56	23.92	6.57	11.66	39.78	0.40	-0.19
B4	56	15.89	6.99	2.14	31.43	0.36	-0.56
G2	56	98.68	3.93	83.33	100.00	-2.88	7.01
G3	56	78.61	24.10	20.00	100.00	-1.01	-0.28
G4	56	55.27	26.07	7.14	100.00	-0.07	-1.15
K2	56	99.93	0.37	97.96	100.00	-4.88	22.28
K3	56	94.37	3.95	79.10	100.00	-1.51	3.06
K4	56	61.68	7.64	44.68	77.88	0.10	-0.26
M2	56	95.25	6.57	75.00	100.00	-1.19	0.54
M3	56	46.67	25.31	0.00	100.00	0.00	-0.55
M4	56	10.42	14.61	0.00	50.00	1.21	0.42
S2	56	99.57	1.44	92.86	100.00	-3.16	9.20
S3	56	96.64	7.16	60.00	100.00	-3.13	11.32
S4	56	77.67	14.16	38.10	100.00	-0.77	-0.07
Mittelwert	840	69.81	33.88	0.00	100.00	-0.77	-0.94

und Stroopaufgabe über dem Durchschnitt (Teststärken für die Effekte der Aufgabentypen liegen bei 100.0 %). Der Haupteffekt des Aufgabentyps kann mit der unterschiedlichen Lösbarkeit der Aufgaben erklärt werden. So wurde bei der Balltrackingaufgabe der Zielball durchschnittlich 44.2 % der Zeit getroffen. Im Gedächtnistest wurden im Durchschnitt 77.5 % der Zahlenreihen richtig erinnert und im Konzentrationstest 85.3 % richtige Antworten gegeben. Die Matrizenaufgaben wurden durchschnittlich 50.8 % richtig gelöst und bei der Stroop-Aufgabe wurde im Durchschnitt für 91.3 % der Stimuli die korrekte Farbe gewählt. Neben den Haupteffekten trägt auch die Interaktion aus Aufgabenschwierigkeit und Aufgabentyp mit 12.9 % einen wesentlichen Beitrag

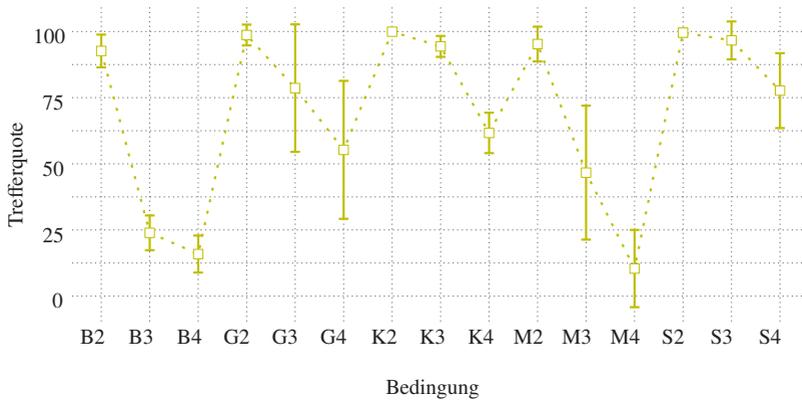


Abbildung A-1: Mittelwert und Standardabweichung der Trefferquote in den Versuchsbedingungen.

zur Varianzaufklärung in der Trefferquote (siehe Tabelle A-27). Anhand der Regressionsgewichte in Tabelle A-28 ist erkennbar, dass der Abfall der Trefferquote von einer Schwierigkeitsstufe zur nächsten für jede Aufgabe spezifisch ist. Ausnahmen bilden lediglich der Abfall der Schwierigkeitsstufen G3 zu G4 sowie S3 und S4 (keine signifikanten Regressionsgewichte; siehe Tabelle A-28). Diese unterscheiden sich nicht zum durchschnittlichen Trefferquotenrückgang der

Tabelle A-27: Modellfit-Indizes für die geschachtelten Modelle der Trefferquote

Modell	df	AIC	$-2LL$	χ^2_{diff}	df_{diff}	R^2_{GLMM}
Nullmodell	3	8307.03	8301.03			0.00
+ Schwierigkeit (M_S)	5	7867.49	7857.49	443.54***	2	41.05
+ Aufgabe (M_A)	9	7246.23	7228.23	629.26***	4	72.14
+ Schwierigkeit \times Aufgabe ($M_{S \times A}$)	17	6735.19	6701.19	527.04***	8	84.99

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Tabelle A-28: Parameter für das Modell ($M_{S \times A}$) der Trefferquote

	Feste Effekte				Zufallseffekte
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>SD</i>
(Ordinatenabschnitt)	69.81***	0.57	122.31	56	2.69
Schwierigkeit: 3 – 2	-29.18***	1.09	-26.87	784	
Schwierigkeit: 4 – 3	-23.85***	1.09	-21.97	784	
Aufgabe: B	-25.66***	0.89	-28.95	784	
Aufgabe: G	7.71***	0.89	8.69	784	
Aufgabe: K	15.51***	0.89	17.50	784	
Aufgabe: M	-19.03***	0.89	-21.47	784	
Aufgabe: S	21.48***	0.89	24.23	784	
(Schw.: 3 – 2)×(Auf.: B)	-39.56***	2.17	-18.22	784	
(Schw.: 4 – 3)×(Auf.: B)	15.83***	2.17	7.29	784	
(Schw.: 3 – 2)×(Auf.: G)	9.11***	2.17	4.19	784	
(Schw.: 4 – 3)×(Auf.: G)	0.52	2.17	0.24	784	
(Schw.: 3 – 2)×(Auf.: K)	23.61***	2.17	10.87	784	
(Schw.: 4 – 3)×(Auf.: K)	-8.84***	2.17	-4.07	784	
(Schw.: 3 – 2)×(Auf.: M)	-19.41***	2.17	-8.94	784	
(Schw.: 4 – 3)×(Auf.: M)	-12.40***	2.17	-5.71	784	
(Schw.: 3 – 2)×(Auf.: S)	26.25***	2.17	12.09	784	
(Schw.: 4 – 3)×(Auf.: S)	4.88	2.17	2.25	784	

*** $p < 0.0001$, ** $p < 0.001$, * $p < 0.0025$

Schwierigkeitsstufen 3 und 4 über alle Aufgabentypen hinweg. Die Teststärken für die bedeutsamen Interaktionseffekte liegen zwischen 98.6 % und 100.0 %.

Zusammenfassung und Interpretation der Ergebnisse

Die Ergebnisse belegen eine erwartungskonforme und bedeutsame Steigung der leistungsbezogenen mentalen Beanspruchung mit steigenden Schwierigkeits-

stufen der Aufgabe (Bestätigung der Hypothese H3.3). Auch die Hypothese H3.4 wird durch den signifikanten Haupteffekt des Aufgabentyps bestätigt. In Bezug auf die Unterfragestellung UF3.2 zeigen sich aufgabenspezifische Steigungen zwischen den Schwierigkeitsstufen der Aufgabe (bedeutsamer Interaktionseffekt). Diese aufgabenspezifischen Muster sind in der leistungsbezogenen mentalen Beanspruchung stärker ausgeprägt als in der subjektiven mentalen Beanspruchung, gemessen mit dem NASA-RTLX (vgl. Abschnitt 8.3.1). Bei der leistungsbezogenen mentalen Beanspruchung trägt der Haupteffekt des Aufgabentyps mit 31.1 % einen ähnlich hohen Beitrag zur Varianzaufklärung bei wie der Haupteffekt der Aufgabenschwierigkeit (41.0 %). Aber auch der Interaktionseffekt leistet mit 12.9 % einen wesentlichen Beitrag. Diese Effekte sind auf die charakteristische Leistungsmessung des jeweiligen Aufgabentyps zurückzuführen. Sie unterscheidet sich quantitativ (Anzahl präsentierter Aufgaben) als auch qualitativ (Art des Wissensabrufs und der Lösungseingabe) zwischen den Aufgabentypen. Die vorliegenden Daten lassen vermuten, dass die Leistungsmessung der Studie III durch konfundierende Faktoren, wie die Lösbarkeit des Aufgabentyps, die Anzahl präsentierter Aufgaben pro Aufgabentyp oder die aufgabenspezifische Antworteingabe, substantiell beeinflusst wird. Da die Trefferquotenvarianz wesentlich durch den Aufgabentyp und dessen Interaktion mit der Aufgabenschwierigkeit beeinflusst wird, würden die extrahierten physiologischen Profile nicht als valide Repräsentanten der leistungsbezogenen mentalen Beanspruchung interpretiert werden können. Aus diesem Grund wurde in Studie III keine Extraktion physiologischer Profile der leistungsbezogenen mentalen Beanspruchung vorgenommen.

A.8.5 Bedingungs-Faktorladungen der NPLS-Regression

Tabelle A-29: Bedingungs-Faktorladungen der NASA-RTLX- und der kardialen Komponente aus der NPLS-Regression des NASA-RTLX auf kardiale Parameter

Bedingung	Kardiale Komponente	NASA-RTLX-Komponente
B1	-0.30	-0.26
B2	0.06	-0.14
B3	0.29	0.12
B4	0.33	0.22
G1	-0.27	-0.24
G2	0.02	-0.11
G3	0.18	0.16
G4	0.22	0.36
K1	-0.37	-0.27
K2	-0.05	-0.16
K3	0.06	0.04
K4	0.07	0.33
M1	-0.41	-0.27
M2	0.08	-0.03
M3	0.16	0.26
M4	0.17	0.40
S1	-0.40	-0.27
S2	-0.02	-0.16
S3	0.13	-0.08
S4	0.06	0.11

Tabelle A-30: Bedingungs-Faktorladungen der NASA-RTLX- und der elektrodermalen Komponente aus der NPLS-Regression des NASA-RTLX auf elektrodermale Parameter

Bedingung	Elektrodermale Komponente	NASA-RTLX-Komponente
B1	-0.35	-0.25
B2	-0.21	-0.15
B3	-0.04	0.10
B4	0.29	0.21
G1	-0.32	-0.25
G2	-0.09	-0.12
G3	0.36	0.17
G4	0.35	0.37
K1	-0.15	-0.26
K2	-0.15	-0.17
K3	0.16	0.04
K4	0.31	0.33
M1	-0.17	-0.26
M2	0.03	-0.03
M3	0.18	0.25
M4	0.21	0.41
S1	-0.27	-0.26
S2	-0.14	-0.16
S3	-0.15	-0.09
S4	0.10	0.10

A.9 Vorhersage der mentalen Beanspruchung

Abbildung A-2 veranschaulicht beispielhaft, wie sich die Vorhersage der subjektiven mentalen Beanspruchung (NASA-RTLX) durch die Gewichtung der physiologischen Parameter mit den Faktorladungen verändert. Hierbei werden elektrodermale Daten von 10 Personen der Studie III verwendet. Zum Vergleich wird die Vorhersage der subjektiven mentalen Beanspruchung durch den Mittelwert aller elektrodermalen Parameter herangezogen (Abbildung A-2, A). In Abbildung A-2 sind die vorhergesagten NASA-RTLX-Werte der jeweiligen Bedingungen (linkes Diagramm) und der Zusammenhang zwischen den vorhergesagten und den tatsächlichen NASA-RTLX-Werten für jede einzelne Person (rechtes Diagramm) dargestellt. Die personenspezifischen Korrelationskoeffizienten sind in der Legende der Abbildung A-2 zu finden. Die Korrelationen zwischen den tatsächlichen und vorhergesagten NASA-RTLX-Werten erhöhen sich, wenn die subjektive mentale Beanspruchung mit Hilfe des elektrodermalen Profils mentaler Beanspruchung vorhergesagt wird (siehe Abbildung A-2, B). Dabei werden die elektrodermalen Parameterwerte mit den bedeutsamen Parameter-Faktorwerten gewichtet. Die Vorhersage kann weiter verbessert werden, indem zusätzlich die Bedingungs-Faktorladungen in die Berechnung einbezogen werden (Abbildung A-2, C). Bezogen auf das Beispiel aus Kapitel 9.3 bedeutet es, dass die mentale Beanspruchung neuer Personen für bereits bekannte Situationen (bzw. Operationsschritte) berechnet wird. In diesem Fall wäre das Mensch-Maschine-System bereits auf die Situationen kalibriert.

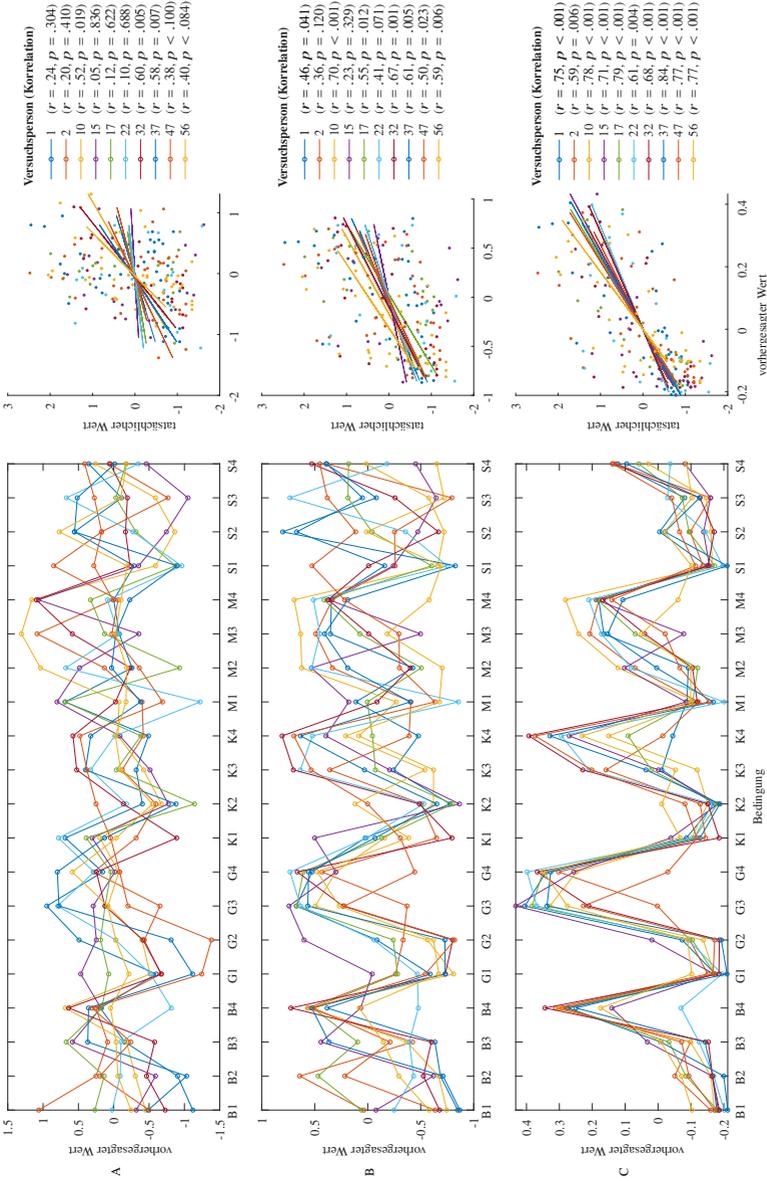


Abbildung A-2: Vorhersage personenspezifischer NASA-RTLX-Werte durch die elektrodermale Parameter für 10 Personen der Studie III.
 A: Vorhersage durch den Mittelwert der Parameter. B: Vorhersage durch das elektrodermale Profil (Parameter-Faktorladungen).
 C: Vorhersage durch das elektrodermale Profil und die Bedingungs-Faktorladungen.