

Möglichkeiten eines Assessments psycho-physischer Belastbarkeit im Kontext der Personalauswahl

Gregor Hackfort

Vollständiger Abdruck der von der Fakultät für Humanwissenschaften
der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktors der Philosophie

genehmigten Dissertation.

Gutachter:

1. Prof. Dr. Andreas Schlattmann
2. Prof. Dr. Billy Sperlich

Die Dissertation wurde am 02.06.2014 bei der Universität der Bundeswehr München
eingereicht und durch die Fakultät für Humanwissenschaften am 26.07.2014 angenommen.
Die mündliche Prüfung fand am 28.07.2014 statt.

Für meine Eltern

Danksagung

An erster Stelle möchte ich meinen Doktorvater Prof. Dr. Andreas Schlattmann nennen, der mir mit seinen Ratschlägen wichtige Impulse für die Erstellung dieser Arbeit gegeben hat und mir von Anfang an mit anregenden Gesprächen und konstruktiven Diskussionen zur Seite stand und dem, aufgrund der Unterstützung während des gesamten Promotionsvorhabens, mein herzlicher Dank gilt.

Ebenso danke ich Prof. Dr. Billy Sperlich, der in der „zweiten Hälfte“ des Dissertationsprojekts eine Professur an der Universität Würzburg annahm und trotz der Distanz weiterhin für Fragen und fachliche Gespräche stets ein offenes Ohr hatte und hinsichtlich methodischer und trainingswissenschaftlicher Fragestellungen entscheidende Hinweise geben konnte, die zur Gestaltung dieser Arbeit wesentlich beigetragen haben.

Daneben möchte ich mich bei Herrn Dr. Ingo Seidelmeier bedanken, dessen Tür für mich stets offen stand und der mir mit seinen konstruktiven Anmerkungen bei der Fertigstellung dieser Arbeit eine erhebliche Unterstützung war.

Ein weiterer Dank gilt den Studierenden der Sportwissenschaft an der Universität der Bundeswehr München, die mich im Rahmen dieses Promotionsprojektes bei der Datenerhebung tatkräftig unterstützt haben. An dieser Stelle sind zu nennen: Florian Holzhüter, Christoph Sauer, Fabian Hauser, Dennis Chmiel, Marvin Heuer und Jan Landhuis.

In diesem Zuge möchte ich allen Versuchsteilnehmern einen besonderen Dank aussprechen, die bereitwillig an ihre physischen Leistungsgrenzen gegangen sind und sich Zeit für die teilweise umfangreichen Untersuchungen genommen haben. Ohne Sie wäre diese Arbeit nicht zu realisieren gewesen.

Inhaltsverzeichnis

1	Einführung	1
1.1	Problem- und Zielstellung	1
1.2	Relevanz des Forschungsvorhabens	4
1.3	Anfänge der Eignungsdiagnostik und Personalauswahl	10
2	Theoretisches Rahmenkonzept	17
2.1	Psycho-physische Belastbarkeit im situativen Kontext	17
2.2	Theoretische Grundlagen der Belastung, Belastbarkeit und Beanspruchung	22
2.3	Psycho-physisches Belastbarkeitskonzept	36
2.4	Ausdauer als Form physischer Beanspruchung	39
2.5	Psychische Belastbarkeit als ein Persönlichkeitskonstrukt	42
3	Bezugspunkte und Stand der Forschung	51
3.1	Persönlichkeit und Sport	52
3.2	Psychische Gesundheit und Sport	65
3.3	Stress und Sport	70
3.4	Stress, Emotionale Stabilität und Neurotizismus	78
4	Empirisch-methodischer Ansatz	83
4.1	Methodische Ausgangspunkte für die Verfahrensentwicklungen	85
4.2	Konzeption und Orientierung für die Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente	88
4.3	Operationalisierung des Persönlichkeitskonstrukts psychische Belastbarkeit	90
4.3.1	Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung	90
4.3.2	Verfahrensökonomische Überlegungen zur Erfassung der psychischen Belastbarkeit	94
4.3.3	Ansatz zur Validierung des Selbsteinschätzungsverfahrens	99

5	Verfahrensentwicklungen und Untersuchungen zur psycho-physischen Belastbarkeit	103
5.1	Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente	103
5.1.1	Explorationsstudie zur Kriteriendefinition des Testverfahrens zur physischen Belastbarkeit	111
5.1.2	Laborstudie zur Überprüfung der Gütekriterien des 3000-Meter-Fahrradergometertests	122
5.1.2.1	Untersuchungsdesign und -methodik	131
5.1.2.2	Beschreibung der Stichprobe	144
5.1.2.3	Untersuchungsdurchführung	145
5.1.2.4	Ergebnisse der Laborstudie	150
	5.1.2.4.1 Ergebnisdarstellung der Test-Retest-Reliabilität des 3000-m-FET	150
	5.1.2.4.2 Ergebnisdarstellung der Validitätsüberprüfung des 3000-m-FET	161
5.2	Verfahrensentwicklung zur Erfassung der psychischen Belastbarkeitskomponente	178
5.2.1	Überprüfung der Kurz- und Langform-Skalen zur psychischen Belastbarkeit	178
5.2.1.1	Korrelationsanalysen zur Kurz- und Langform der psychischen Belastbarkeitsskalen	180
5.2.1.2	Korrelationsanalysen zum State- und Trait-Aspekt mittels der Kurzform-Skalen	182
5.2.2	Analysen zur Überprüfung der konvergenten und divergenten Validität	184
5.2.2.1	Psychische Belastbarkeit und Neurotizismus (NEO-FFI)	185
5.2.2.2	Psychische Belastbarkeit und Selbstwirksamkeit	186
5.3	Untersuchungen zur psycho-physischen Belastbarkeit	188
5.3.1	Korrelationsanalytische Befunde	192
5.3.1.1	Ausdauerleistung und psychische Belastbarkeit	192
5.3.1.2	Ausdauerleistung und Neurotizismus	200
5.3.2	Regressionsanalytische Befunde	202
5.3.3	Extremgruppenvergleiche und varianzanalytische Befunde	205

6	Diskussion	225
7	Schlusskommentar und Ausblick	245
8	Zusammenfassung	253
9	Literaturverzeichnis	259
	Anhang	277
	Anhang A: Korrelationstabellen	278
	Anhang B: Fragebögen psychische Belastbarkeit (Kurzform-Skalen)	282
	Anhang C: Fragebogen zur Konstruktvalidierung	286
	Anhang D: E-Mail zur Probandengewinnung	287
	Anhang E: Deskriptive Statistik der einzelnen Items der psychischen Belastbarkeitsskala	288
	Anhang F: Häufigkeitsverteilung der 3000-m-FET-Zeiten	289
	Anhang G: Deskriptive Statistik des maximalen Atemminutenvolumens (VE_{\max})	290
	Anhang H: Häufigkeitsverteilungen der Itemscores der Persönlichkeitsskalen	291

Abbildungsverzeichnis

Abbildung 1	Triadische Konzeption der für diese Arbeit relevanten Wissenschaftsbereiche.	3
Abbildung 2	Aufbau der Arbeit.	15
Abbildung 3	Kategorien der sportlichen Belastbarkeit (nach Fröhner, 1993; 2007; Martin & Nicolaus, 1998; aus Schnabel, Harre & Krug, 2008, S. 244).	26
Abbildung 4	Beziehung zwischen Belastung, psycho-physischer Belastbarkeit und Beanspruchung (vgl. Klimmer, 1978).	31
Abbildung 5	Wechselwirkungszusammenhang zwischen Belastung, Beanspruchung und Bewältigungsressourcen.	34
Abbildung 6	Strukturelle Ableitung der psycho-physischen Belastbarkeit aus der Belastungs-Beanspruchungsdifferenzierung.	35
Abbildung 7	Übersicht über die Konstrukte des psycho-physischen Belastbarkeitskonzepts.	37
Abbildung 8	Erweiterte Konzeptualisierung psycho-physischer Belastbarkeit.	38
Abbildung 9	Untersuchungsergebnisse im EPI bei verschiedenen Gruppen männlicher Athleten (nach Singer, 2000, S. 306).	56
Abbildung 10	Ursachen-Wirkungshypothese (nach Schlicht, 1995, S. 33).	67
Abbildung 11	Psycho-physisches Belastbarkeitskonzept.	87
Abbildung 12	Dimensionen-/Faktorenmodell des BIP (nach Hossiep & Paschen, 1998, S. 17).	92
Abbildung 13	Ableitung der Kurzsкала aus den Skalen „Emotionale Stabilität“ und „Belastbarkeit“ des BIP.	98

Abbildung 14	Darstellung des Bildschirms (Screenshot) der Tacx-Trainer-Software während der Testphase des 3000-m-FET.	108
Abbildung 15	3000-m-FET zur Feststellung der Basisfitness mit den Instrumenten der Firma Tacx.	109
Abbildung 16	Sporteinheiten der Pbn pro Woche (N = 50).	114
Abbildung 17	Altersverteilung der Teilnehmer der Studie zur Testentwicklung (N = 50).	115
Abbildung 18	Zusammenhang des 1000-m-Laufs mit dem 6:30-Min.-FET (N = 50).	117
Abbildung 19	Screenshot des B.I.T. Soft-Testsystems.	118
Abbildung 20	ERGO-FIT Display.	119
Abbildung 21	Beziehungen zwischen den Hauptgütekriterien (nach Lienert & Raatz, 1994, S. 13).	131
Abbildung 22	Laboruntersuchung mit Spiroergometrie zur Validierung des 3000-m-FET.	132
Abbildung 23	Darstellung des Bildschirms während der Testphase eines Rampentests auf dem Fahrradergometer.	133
Abbildung 24	Untersuchungsdesign zur Prüfung der Gütekriterien des 3000-m-FET.	135
Abbildung 25	Sicht auf den Bildschirm aus der Perspektive des Pbn und des Testleiters während der Testung.	148
Abbildung 26	Absolute $VO_{2peak/max}$ -Mittelwerte in l/Min. der Pbn für alle vier Ausdauerests (N = 145).	165
Abbildung 27	Einordnung der durchschnittlichen rel. VO_{2max} (ml/kg/Min.) des 3000-m-FET, differenziert nach Geschlecht (weibl.: n = 26; männl.: n = 119) in Anlehnung an Care Fusion (2013).	169

XIII

Abbildung 28	Hf _{max} -Werte in Abhängigkeit vom Alter eingeteilt nach Lebensdekaden (N = 145).	171
Abbildung 29	Darstellung der einzelnen Studien, deren Stichprobengröße, dabei eingesetzte Verfahren und korrespondierende Indikatoren.	191
Abbildung 30	Altersverteilung der Gesamtstichprobe in Gruppen nach Lebensdekaden (N = 424).	193
Abbildung 31	Zusammenhang zwischen objektiv erfasster (Zeit 3000-m-FET) physischer und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit (N = 424).	195
Abbildung 32	Zusammenhang zwischen objektiv erfasster physischer (Zeit 1000-m-Lauf) und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit (N = 145).	196
Abbildung 33	Zusammenhang zwischen objektiv erfasster physischer (VO _{2max}) und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit (N = 145).	199
Abbildung 34	Vergleich der LG (aufgeteilt nach den Zeiten des 3000-m-FET) hinsichtlich der Langform-Skala psychische Belastbarkeit (N = 424).	209
Abbildung 35	Profildiagramm der LG (eingeteilt nach der Berechnung: M ± SD der 3000-Zeiten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; N = 424).	213
Abbildung 36	Profildiagramm der LG (eingeteilt nach Drittelung der 3000-Zeiten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; N = 424).	216
Abbildung 37	Profildiagramm der LG (eingeteilt nach den VO _{2max} -Werten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Kurzform-Skala; N = 145).	219
Abbildung 38	Profildiagramm der LG (Faktor L) und Altersgruppen (Faktor A)	221

(jeweils eingeteilt nach dem Drittelungsverfahren) in Abhängigkeit der psychischen Belastbarkeit (Kurzform-Skala; N = 424).

Abbildung 39	Profildiagramm der LG (Faktor L) und Altersgruppen (Faktor A) (jeweils eingeteilt nach dem Drittelungsverfahren) in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; N = 424).	222
Abbildung 40	Berufsbedingte Eignung in Abhängigkeit des Ausprägungsgrades der psycho-physischen Belastbarkeit.	243
Abbildung B1	Instruktion der Langform-Skala.	282
Abbildung B2	Kurzform der psychischen Belastbarkeitsskala (Trait).	283
Abbildung B3	Kurzform der psychischen Belastbarkeitsskala (State, vorher).	284
Abbildung B4	Kurzform der psychischen Belastbarkeitsskala (State, nachher).	285
Abbildung C1	Skala zur allgemeinen Selbstwirksamkeit (nach Jerusalem & Schwarzer, 1999).	286
Abbildung D1	Wortlaut der E-Mail an die Mitarbeiter und Studenten der Universität der Bundeswehr München.	287
Abbildung F2	Häufigkeitsverteilung der 3000-m-FET-Zeiten (N = 424).	289
Abbildung H1	Häufigkeitsverteilung der Itemscores der psychischen Belastbarkeitsskala (Kurzform; N = 424).	291
Abbildung H2	Häufigkeitsverteilung der Itemscores der psychischen Belastbarkeitsskala (Langform; N = 424).	291
Abbildung H3	Häufigkeitsverteilung der Itemscores der Neurotizismus-Skala (N = 155).	292
Abbildung H4	Häufigkeitsverteilung der Itemscores der Skala zur allgemeinen Selbstwirksamkeit (N = 225).	292

Tabellenverzeichnis

Tabelle 1	Beispiele für Beanspruchungsfolgen bzw. Stressreaktionen (nach Greif, 1991; Kaufmann, Pornschlegel & Udris, 1982; Kaluza, 1996; aus Udris & Frese, 1999, S. 432).	24
Tabelle 2	Emotionskategorien (nach Schmidt-Atzert & Ströhm, 1983, S. 136).	80
Tabelle 3	Ergebnisse der Explorationsstudie (Teil I) zur Konzeption und Evaluierung eines Fahrradergometertests für die Personalgewinnung der Bundeswehr. Vergleich eines 1000-m-Laufs mit einem 6-minütigen Fahrradergometertest (N = 8).	113
Tabelle 4	Personen- und Leistungsdaten (N = 50).	116
Tabelle 5	Borg-Skala.	142
Tabelle 6	Messgenauigkeiten der Spiroergometrie (nach CareFusion, 2013).	149
Tabelle 7	Deskriptive Statistik der Testdauer im Vor- und Nachtest in Sekunden (n = 29).	151
Tabelle 8	Korrelationen der Testdauer zwischen Vor- und Nachtest (n = 29).	152
Tabelle 9	Maximale Sauerstoffaufnahme-Werte (VO_{2max}) zwischen Vor- und Nachtest in l/Min. (n = 29).	153
Tabelle 10	Korrelationen der maximalen Sauerstoffaufnahme-Werte (VO_{2max}) zwischen Vor- und Nachtest (n = 29).	154
Tabelle 11	Maximale Herzfrequenzen zwischen Vor- und Nachtest in Schlägen/Min. (n = 29).	155
Tabelle 12	Korrelationen der maximalen Herzfrequenzwerte (Hf_{max}) zwischen Vor- und Nachtest (n = 29).	156
Tabelle 13	Laktatkonzentrationen im Blut nach Beendigung der Testphase zwischen Vor- und Nachtest in mmol/l (n = 29).	157

Tabelle 14	Korrelationen der maximalen Laktatwerte zwischen Vor- und Nachtest (n = 29).	157
Tabelle 15	Werte der Borg-Skala zwischen den Vor- und Nachtests (n = 29).	158
Tabelle 16	Korrelationen der Borg-Werte zwischen den Vor- und Nachtests (n = 29).	159
Tabelle 17	T-Test für abhängige Stichproben der erhobenen Parameter bei dem 3000-m-FET (n = 29).	160
Tabelle 18	Benötigte Zeiten für die verschiedenen Ausdauertests in Sek. (N = 145).	162
Tabelle 19	Korrelationen der benötigten Zeiten (N = 145).	163
Tabelle 20	Maximale absolute Sauerstoffaufnahme-Werte in l/Min. (N = 145).	164
Tabelle 21	Korrelationen der absoluten maximalen Sauerstoffaufnahme-Werte (N = 145).	165
Tabelle 22	Die relativen VO_{2max} -Werte der einzelnen Ausdauertests in ml/kg/Min. (N = 145).	167
Tabelle 23	Durchschnitts- und Normwerte der VO_{2max} -Werte für unterschiedliche Leistungsniveaus (mod. nach Zintl, 1994, S. 56).	167
Tabelle 24	Korrelationen der relativen VO_{2max} -Werte (N = 145).	168
Tabelle 25	Deskriptive Statistik der relativen VO_{2max} -Werte (ml/kg/Min.) nach geschlechterspezifischer Unterscheidung (weibl.: n = 26; männl.: n = 119).	168
Tabelle 26	Maximale Herzfrequenzwerte in Schlägen/Min. (N = 145).	170
Tabelle 27	Korrelationen zwischen den Hf_{max} -Werten bei verschiedenen Ausdauertests (N = 145).	172
Tabelle 28	Maximale Laktatkonzentrationen in mmol/l (N = 145).	172

Tabelle 29	Korrelationen der maximalen Laktatkonzentrationen bei den verschiedenen Ausdauertests (N = 145).	173
Tabelle 30	Angegebene Werte der Borg-Skala als Ausbelastungskriterium (N = 145).	174
Tabelle 31	Korrelationen der Werte der Borg-Skala (N = 145).	174
Tabelle 32	Deskriptive Statistik der RQ-Werte bei den verschiedenen Ausdauertests (N = 145).	175
Tabelle 33	Durchschnittliche Leistungen (in Watt) bei den beiden Fahrradergometertests (N = 145).	177
Tabelle 34	Deskriptive Darstellung der Itemscores beider Skalen (Lang- und Kurzform) zur Erfassung der psychischen Belastbarkeit (N = 424).	181
Tabelle 35	Korrelationsanalyse der beiden Skalen (Kurz- und Langform) zur Erfassung der psychischen Belastbarkeit (N = 424).	182
Tabelle 36	Deskriptive Darstellung der Itemscores der beiden State-Kurzform Skalen zur Erfassung der psychischen Belastbarkeit (n = 275).	183
Tabelle 37	Korrelationsanalyse der beiden eingesetzten Kurzform-Skalen (State und Trait; n = 275).	183
Tabelle 38	Deskriptive Darstellung der Itemscores der Neurotizismus-Skala des NEO-FFI zur Erfassung der emotionalen Labilität (n = 155).	185
Tabelle 39	Korrelationsanalyse zur divergenten Validitätsüberprüfung (n = 155).	186
Tabelle 40	Deskriptive Darstellung der Itemscores der Skala zur allgemeinen Selbstwirksamkeitserwartung (n = 225).	186
Tabelle 41	Korrelationsanalyse zur konvergenten Validitätsüberprüfung (n = 225).	187
Tabelle 42	Zusammenhang zwischen dem physischen Indikator Zeit (3000-m-FET) und der psychischen Belastbarkeit (N = 424).	194

XVIII

Tabelle 43	Korrelationsanalyse des Parameters Zeit und der psychischen Belastbarkeit (N = 145).	197
Tabelle 44	Korrelationsanalyse der Parameter durchschnittliche Wattleistung (3000-m-FET) bzw. maximale Wattleistung (Rampentest F.-Ergom.) und der psychischen Belastbarkeit (N = 145).	197
Tabelle 45	Korrelationsanalyse der Parameter $VO_{2peak/max}$ und der psychischen Belastbarkeit (N = 145).	198
Tabelle 46	Korrelationsanalyse der physischen Parameter und der emotionalen Labilität, erfasst mit der Skala Neurotizismus des NEO-FFI (N = 145).	201
Tabelle 47	Regressionsanalyse hinsichtlich der Langform-Skala der psychischen Belastbarkeit (N = 145).	204
Tabelle 48	Regressionsanalyse hinsichtlich der Kurzform-Skala der psychischen Belastbarkeit (N = 145).	205
Tabelle 49	Vergleich der LG I und III (Zeit: 3000-m-FET) im Hinblick auf die psychische Belastbarkeit (n = 92), berechnet mit dem t-Test für unabhängige Stichproben.	206
Tabelle 50	Vergleich der LG I und III (VO_{2peak} : Rampentest F.-Ergom.; in l/Min.) im Hinblick auf die psychische Belastbarkeit (n = 48), berechnet mit dem t-Test für unabhängige Stichproben.	207
Tabelle 51	Vergleich der drei LG (Zeit: 3000-m-FET) im Hinblick auf die Lang und Kurzform-Skala der psychischen Belastbarkeit (N = 424), berechnet mit Hilfe der einfaktoriellen Varianzanalyse (ANOVA).	208
Tabelle 52	Zweifaktorielle Varianzanalyse der Faktoren L (LG) und G (Geschlecht) hinsichtlich beider Skalen der psychischen Belastbarkeit (N = 424).	210
Tabelle 53	Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem $M \pm SD$ der 3000-m-FET-Zeiten	212

(N = 424).

Tabelle 54	Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach der Drittelung der 3000-m-FET-Zeiten (N = 424).	215
Tabelle 55	LG der Frauen und Männer bestimmt durch die VO_{2max} -Werte (l/Min.) des Rampentests auf dem Laufband (N = 145).	217
Tabelle 56	Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem $M \pm SD$ der VO_{2max} -Werte des Rampentests auf dem Laufband (N = 145).	218
Tabelle 57	Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Alter (A) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem Drittelungsverfahren für die Faktoren L und A (N = 424).	220
Tabelle A1	Interkorrelationen der physischen Belastbarkeitsindikatoren (N = 145).	278
Tabelle A2	Interkorrelationen der Skalen des BIP (nach Hossiep & Paschen, 1998, S. 31).	279
Tabelle A3	Korrelationen der VE_{max} -Werte zwischen Vor- und Nachtest (n = 29).	280
Tabelle A4	Korrelationen zwischen den physischen Belastbarkeitsindikatoren und den Itemscores der Skala zur allgemeinen Selbstwirksamkeit (N = 145).	280
Tabelle A5	Korrelationen zwischen der 3000-m-FET-Zeit, gemessen mit dem Ergometer der Firma Ergo Fit und den Itemscores der Skala zur psychischen Belastbarkeit (Lang- und Kurzform; n = 82).	281
Tabelle A6	Korrelationen zwischen der 3000-m-FET-Zeit, gemessen mit den Tacx-Trainern und den Itemscores der Skala zur psychischen Belastbarkeit (Lang- und Kurzform; n = 342).	281

Tabelle A7	Korrelationen der VE_{\max} -Werte zwischen den verschiedenen Ausdauer-testverfahren (N = 145).	281
Tabelle E1	Itemstatistik: Skala zur psychischen Belastbarkeit (N = 424).	288
Tabelle F1	Tabellarische Darstellung der Häufigkeitsverteilung der Zeiten des 3000-m-FET (N = 424).	289
Tabelle G1	VE_{\max} -Werte in l/Min. zwischen Vor- und Nachtest (n = 29).	290
Tabelle G2	Deskriptive Statistik des VE_{\max} in l/Min. (N = 145).	290

1 EINFÜHRUNG

Mit der vorliegenden Arbeit werden drei zentrale Anliegen verfolgt, die zunächst zur Orientierung charakterisiert werden. Bevor auf die Relevanz der dokumentierten Forschung eingegangen wird, ist die Problem- und Zielstellung im Verbund erläutert. Der historische Hintergrund, wie in Kapitel 1.3 dargestellt, soll einen kurzen Überblick über die Anfänge und Entwicklungen der Eignungsdiagnostik bzw. der Personalauswahl liefern. Der Gang der Arbeit wird anschließend in einer zusammenfassenden Übersicht veranschaulicht.

1.1 Problem- und Zielstellung

Im Kontext der Personalauswahl und Eignungsfeststellung interessieren im Bewerbungsverfahren neben dem fachlichen Know-how und den jobspezifischen Kompetenzen des Bewerbers selbstverständlich auch die Ausprägungen bestimmter berufsrelevanter Persönlichkeitsmerkmale. Diese Faktoren werden je nach Art des Berufes und in Abhängigkeit von der angestrebten Position unterschiedlich stark beachtet und mit verschiedensten Methoden versucht zu überprüfen. Dazu werden z. B. interviewgeführte Gespräche, improvisierte Testpräsentationen, Rollenspiele, Logik-Tests oder Fallstudien eingesetzt (siehe hierzu die Literatur zu Eignungsdiagnostik und/oder Assessment Center, u. a. Fay, 2002; Fisseni & Preusser, 2007; Sarges, 2001). Für die Hinweisgewinnung auf die Eignung eines Bewerbers gibt es unterschiedliche Wege und Herangehensweisen, die mit Hilfe von diversen Instrumenten und spezifischen Verfahren sowie kombinierten multiplen Verfahren (Assessment Center; siehe u. a. Höft, 2007) mehr oder weniger aussagekräftige Informationen über den Bewerber liefern, welche den Entscheidungsträger bei der Auswahl unterstützen. Nach der Identifikation einer Person, die den Anforderungen des Arbeitsplatzes gerecht wird, besteht seitens des Arbeitgebers meist nicht nur der Wunsch, dass die an den neuen Mitarbeiter herangetragenen Anforderungen erwartungsgemäß erfüllt werden, sondern darüber hinaus auch die Erwartung einen in psychischer und physischer Hinsicht stabilen Mitarbeiter gewonnen zu haben, der den berufsbedingten Belastungen standzuhalten vermag.

In Tätigkeitsfeldern, bei denen sowohl hohe psychische als auch physische Anforderungen gegeben sind oder aufgrund der gegenseitigen Beeinflussung entweder psychische oder physische Besonderheiten hinzukommen, sollten im Rahmen der Personalauswahl beide Aspekte

gesondert und ihre Interaktion Berücksichtigung finden. Insofern erscheint es angebracht, bei der Entwicklung von entsprechenden Verfahren zur Personalauswahl auf methodische Bestände verschiedener Disziplinen zurückzugreifen und ggf. unterschiedliche Verfahrensansätze zu kombinieren oder gar zu integrieren. Der Kombination physischer und psychischer Aspekte bei der Personalauswahl ist man sich z. B. im militärischen Bereich durchaus bewusst und es wird dazu von der allgemeinen Annahme ausgegangen, dass körperliche Leistungsfähigkeit auch psychische bzw. mentale Aspekte stärkt (s. u.). Dennoch sind die Wechselwirkungen von physischen und psychischen Faktoren im Hinblick auf die berufliche Eignung noch kaum beachtet, was wohl auf fehlende Verfahren bzw. Verfahrensverbände zurückzuführen ist. Die sich daraus ergebende Problemstellung und in Anbetracht der in diesem Bezug bisher noch unzureichend berücksichtigt gebliebenen Wechselwirkungen psychischer und physischer Komponenten, wie sie geradezu exemplarisch hinsichtlich der Belastbarkeit auftritt, soll im Rahmen dieser Arbeit aufgegriffen werden. Darüber hinaus soll eine Möglichkeit zur Operationalisierung des Konstrukts der psycho-physischen Belastbarkeit für Maßnahmen des Personalmanagements entwickelt werden. Somit ergibt sich eine triadische Konzeption, die auf Erkenntnissen der Wissenschaftsbereiche Sportwissenschaft (insbesondere trainingswissenschaftlichen Aspekten), Arbeitswissenschaft (speziell Arbeits- und Organisationspsychologie) und der Psychologie (besonders der Persönlichkeitspsychologie) fußt und auf den für diese Arbeit maßgeblichen Kenntnisstand dieser Wissenschaftsbereiche basiert (siehe Abbildung 1).

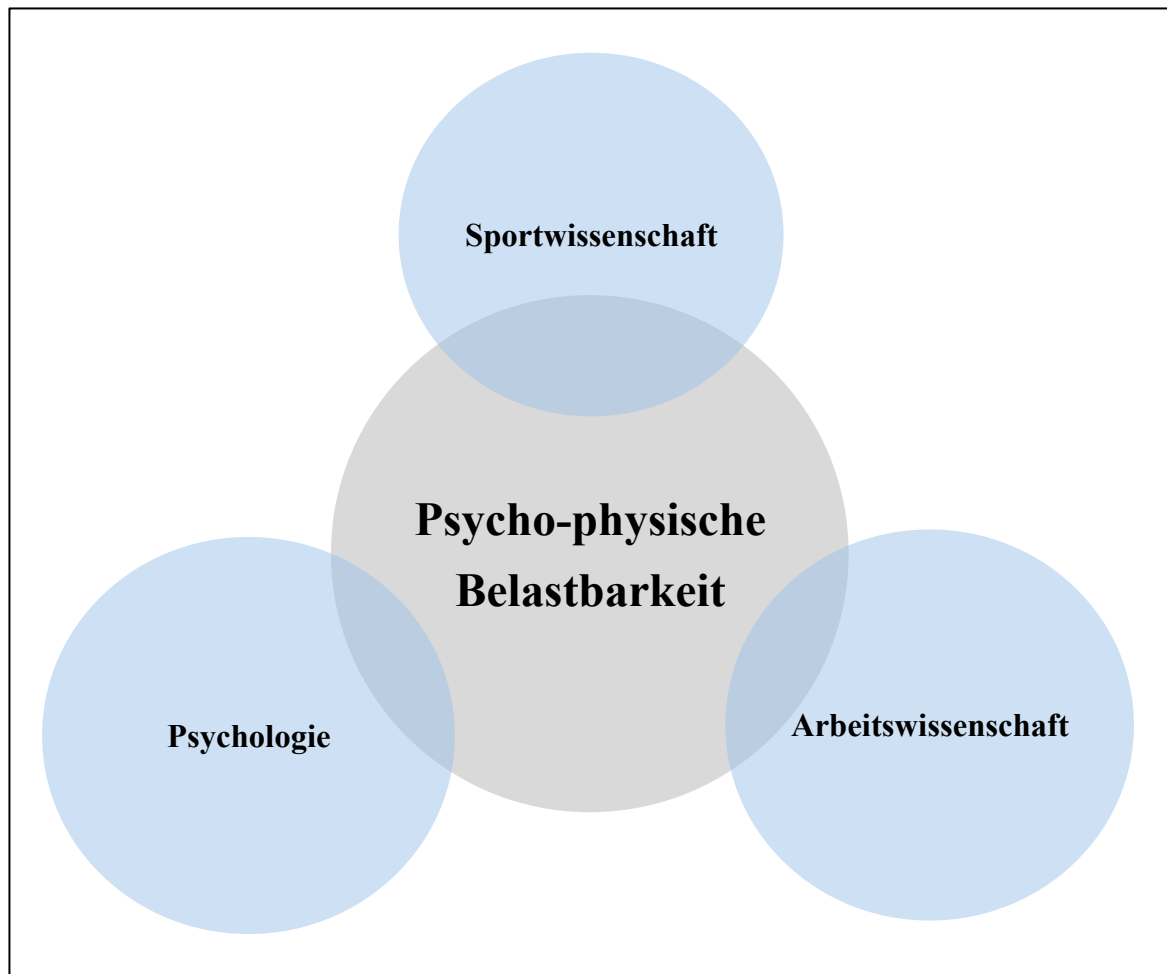


Abbildung 1. Triadische Konzeption der für diese Arbeit relevanten Wissenschaftsbereiche.

Die vorliegende Arbeit befasst sich dementsprechend mit der psycho-physischen Belastbarkeit, die in den Kontext der Personalauswahl gestellt wird. Durch die hierzu entwickelnden Verfahren zur Erfassung der physischen als auch der psychischen Belastbarkeit sollen Möglichkeiten aufgezeigt werden, die beiden Komponenten zeiteffizient im Sinne eines Screenings zu erfassen und darüber hinaus Unterschieden zwischen Gruppen mit differenziertem körperlichen Leistungsniveau nachzugehen. In Folge dieser Anliegen ergeben sich drei wesentliche Aufgabenstellungen für diese Arbeit: (1) Es soll ein Verfahren entwickelt werden, womit die physische Belastbarkeit mit speziellem Bezug auf die Ausdauerleistungsfähigkeit operationalisiert und erfasst werden kann. (2) Zusätzlich wird aus einem gängigen und bereits bestehendem Instrument zur Überprüfung von berufsbezogenen Persönlichkeitseigenschaften ein Verfahren abgeleitet und zur Operationalisierung der psychischen Belastbarkeit spezifiziert und eine daraus zeitökonomische Variante konzipiert. Im Anschluss daran werden die entwickelten Verfahren erprobt und hinsichtlich der Gütekriterien überprüft. (3) Mit den so definierten physischen und psychischen Belastbarkeitsindikatoren werden nachfolgend

psycho-physische Zusammenhänge analysiert. Die konzipierten Methoden sollen Möglichkeiten bieten, einerseits - im Anwendungsbezug - psycho-physische Belastbarkeitsaspekte im Kontext von Personalauswahlprozessen ökonomisch zu erfassen und andererseits - im Forschungsbezug - psycho-physischen Zusammenhängen differenziert nachgehen zu können.

1.2 Relevanz des Forschungsvorhabens

In der Arbeitswelt rücken immer mehr psychosoziale Belastungen (Stressfaktoren) in den Fokus der Arbeitgeber und Arbeitnehmervertreter. Insbesondere diesbezüglich einschlägige Erkrankungen, wie psychosomatische Störungen, Depressionen und Burn-Out-Syndrome schränken die Produktivität am Arbeitsplatz deutlich ein, führen zu betriebs- und volkswirtschaftlichen Schäden und verursachen dem Gesundheitssystem immense Versorgungskosten. Politik und Wirtschaft erkennen zunehmend die Relevanz von nicht nur physisch, sondern auch psychisch gesunden Mitarbeitern als wirtschaftlichen Faktor (siehe Statistisches Bundesamt, 2010). Gleichwohl bleibt es größtenteils unbemerkt, dass es lange vor der Manifestation auftretender Krankheitsbilder zu einer verminderten Belastbarkeit und nachlassenden Leistungsfähigkeit, dem sogenannten „Präsentismus“ (siehe Leyk, Franke, Hofmann, Klein, Weller, Hackfort, Löllgen & Piekarski, 2013, S. 163), kommt.

Die Bedeutsamkeit und Aktualität des Themas ergibt sich - was die praktische Relevanz angeht - daraus, dass gegenwärtig die Rate psychischer Erkrankungen deutlich zunimmt, insbesondere Burnout und Depressivität. Krankheitsstatistiken dokumentieren einen jährlichen Anstieg dieser Erkrankungen. Von 2011 auf 2012 stieg der Anteil von 12,1 auf 13,4 Prozent am Gesamtkrankenstand. In den zurückliegenden 15 Jahren hat sich der Anteil dieser Krankheitsgruppe am Krankenstand mehr als verdoppelt und macht inzwischen ein siebtel der Gesamtkennzahl für Erkrankungszeiten bei Berufstätigen aus (DAK Gesundheitsreport, 2012). Dies ist längst nicht mehr als ein Randphänomen einzuordnen, sondern macht seit geraumer Zeit nachdrücklich auf die wachsenden Diskrepanzen zwischen Leistungsanforderungen und der tatsächlichen Belastbarkeit der Arbeitnehmer aufmerksam (vgl. Lohman-Haislah, 2012). Der allgemeinen Annahme der stetigen Zunahme psychischer Erkrankungen steht jedoch die Überlegung entgegen, dass die Zahl derer gestiegen ist, die bereit sind, mit ihren psychischen Gesundheitsproblemen fachliche Hilfe aufzusuchen, was dann zu einer Feststellung und Registrierung einer beträchtlich höheren Anzahl von Patienten in den vergangenen Jahren geführt hat und eben zu dem Eindruck der Zunahme psychischer Gesundheitsprobleme führt

(Breucker, 2013). Die Symptome von Burnout umfassen u. a. eine starke Minderung der Leistungsfähigkeit und damit einhergehende körperliche, sowie kognitive und emotionale Erschöpfungszustände (siehe Maslach, 1993). Diese werden damit in Verbindung gebracht, dass sich die Wahrscheinlichkeit bzw. die Tendenz zu psychosomatischen Erkrankungen, Depressionen, erhöhter Aggressivität und Suchtgefahr erhöht (hat). Demzufolge ist Burnout meist ein Resultat beruflicher Überlastung, d. h. die Folge von externen Bedingungen, die den Organismus aus dem Gleichgewicht bringen. Je nach individueller Leistungsvoraussetzung reagiert die Person unterschiedlich auf eine bestimmte Belastung (auf äußere Einflussfaktoren), was als individuell erlebte (innere) Beanspruchung zu verstehen ist. Demzufolge liegt dem Burnout ein zu hohes und langandauerndes Beanspruchungserleben zugrunde, was sich in einem anhaltenden Stresszustand widerspiegelt. In diesem Stresszustand adaptiert sich das Individuum nicht an die äußeren Einflüsse (Belastung), sondern ist ihnen in diesem Sinne nicht gewachsen und die Überbeanspruchung kommt auf emotionaler und/oder körperlicher Ebene zum Ausdruck (vgl. Freudenberger, 1980; Perlman & Hartman, 1982).

In diesem Zusammenhang stellt sich die Frage, inwieweit sich Personen mit einer gut ausgeprägten Ausdauerleistungsfähigkeit im Sinne einer körperlichen Belastbarkeit von weniger physisch belastbaren Individuen im Hinblick auf die psychische Widerstandsfähigkeit (Resilienz) bzw. Belastbarkeit unterscheiden. Durch das zunehmende Auftreten von Burnout-Phänomenen und psychischen Erschöpfungszuständen speziell durch berufsbedingte Stresssituationen sind Überlegungen und Untersuchungen zu psycho-physischen Zusammenhängen angebracht, die sich speziell mit Belastbarkeitsaspekten befassen. Informationen und Erkenntnisse über Effekte, Zusammenhänge und Unterschiede zwischen physischen und psychischen Aspekten der Belastbarkeit könnten hier neben der betrieblichen Gesundheitsförderung v. a. auch für den Personalauswahlprozess Handlungsempfehlungen abgeleitet werden, die für die Praxis von Relevanz sein dürften.

Bevor im theoretischen Teil auf das Belastungskonzept spezifischer und detaillierter eingegangen wird, sei hier darauf hingewiesen, dass als Belastungen objektive, von außen auf den Menschen einwirkende Faktoren bezeichnet werden. Die Auswirkungen von Belastungen auf den Menschen, die auf Grund differierender Fähigkeiten und Eigenschaften unterschiedlich sind, werden als Beanspruchung bezeichnet (Rohmert & Rutenfranz, 1975, S. 9). Die objektiven Einflussgrößen wirken von außen in Form von Arbeitsanforderungen (sog. Kontext-Faktoren) und Faktoren aus der Arbeitsumwelt (sog. Kontext-Faktoren) auf die Person ein.

Solche Belastungen resultieren in Stress. Stress bedeutet allgemein ein Ungleichgewicht (siehe Semmer, 1988) und wurde ursprünglich mit der Verformung eines Materials in Verbindung gebracht. Nitsch (1981, S. 40) verweist zusammenfassend darauf, dass Stress „in allgemeinsten Sicht (. . .) etwas mit der Anpassung von Lebewesen an ihre jeweilige Umwelt zu tun hat: Stress ist mit Situationen verbunden, in denen sich ein Anpassungsproblem stellt.“ Da die Auswirkungen von Arbeitsbelastungen auf das Individuum als Beanspruchung verstanden werden kann (siehe auch Rutenfranz, 1972), wird davon ausgegangen, dass Stress als eine Form der Beanspruchung Anpassungsleistungen erfordert, die bei zu langer Dauer oder zu hoher Intensität Folgen für die physische wie die psychische Gesundheit haben (Semmer, 1988).

Psychische Erkrankungen haben Herz-Kreislauf-Probleme oder Gelenkerkrankungen in der Fehlzeitenstatistik und Inzidenz überholt und sind inzwischen wesentlich verantwortlich für enorme Steigerungen der Kosten im Gesundheitssystem (DAK Gesundheitsreport, 2013). Dies führt konsequenterweise bei denjenigen, die sich in der Forschung wie in der Praxis mit solchen Problemen beschäftigen, zu Überlegungen möglicher Maßnahmen zur Prävention und auch Reduktion von Risiken (Fehlzeiten, Ausfall, Kosten). D. h., auch bereits bei der Personalauswahl könnten Überlegungen in den Fokus rücken oder nach Möglichkeiten Ausschau gehalten werden, wie man Risiken der Fehltauswahl reduzieren kann oder positiv ausgedrückt: die Wahrscheinlichkeit einer nach Kriterien der Belastbarkeit etc. getroffenen guten bzw. passenden Selektion erhöhen kann. Im Zuge dieser Überlegungen kommt der psycho-physischen Stabilität bzw. Belastbarkeit, die auf Faktoren der psychischen Stabilität und Belastbarkeit einerseits, sowie der physischen Fitness andererseits beruht, spezielle Bedeutung zu. Unter dem Begriff Fitness wird umgangssprachlich im Allgemeinen die körperliche Leistungsfähigkeit und/oder Gesundheit verstanden und vielfach als Synonym für einen „umfassenden Gesundheitsbegriff“ verwendet (Bös, 1998, S. 188). In Gänze umfasst der Fitnessbegriff jedoch verschiedene Aspekte, darunter auch psychische sowie soziale Merkmale einer Person bzw. die Fähigkeit, Tauglichkeit oder Eignung, sich in Abhängigkeit einer bestimmten Aufgabe und von gegebenen Umweltbedingungen anzupassen (siehe u. a. Hollmann, 1990; Bloss, 1992). Es kann davon ausgegangen werden, dass ein weiterer wesentlicher Faktor der Fitness auch in der Erholung und Regeneration bzw. im Regenerationsvermögen einer Person besteht, das sowohl in sportlicher als auch in beruflicher Hinsicht ein bedeutender Faktor für die Belastbarkeit darstellt. Es erscheint die Annahme sinnvoll, dass physisch fittere Personen über eine bessere bzw. schnellere Regenerationsfähigkeit verfügen und sich nach Belastungen so-

wohl auf physischer als auch auf psychischer Ebene schneller regenerieren. Regeneration und Erholung haben im beruflichen und sportlichen Bereich eine protektive Bedeutung für die Gesundheit und können als stressausgleichende Ressource gesehen werden (Allmer, 1996, S. 75).

Die körperliche Leistungsfähigkeit ist ein Teil der Fitness und kann als „physical fitness“ gekennzeichnet werden, welche nach Clarke (1976, S. 174) die konditionelle Leistungsfähigkeit mit einer deutlichen Akzentuierung der Ausdauerleistungsfähigkeit beschreibt (vgl. Bös, 1998). Hollmann definiert Fitness als einen „Zustand guter psychophysischer Bereitschaft zum Lösen einer bestimmten Aufgabe“ (1990, S. 8). Bloss (1992) bezieht sich bei seiner Fitnessdefinition zudem auf die Leistungsfähigkeit, Belastbarkeit und Anpassungsfähigkeit für eine bestimmte Aufgabe. Mit Verweis auf Skarkey (1990, S. 1) werden zwei Komponenten herausgestellt, die der Begriff Fitness umfasst, „one is objective and physiological; the other subjective, emotional and psychological“ (Bös, 1998, S. 189). Bei Betrachtung des Fitnessbegriffs wird deutlich, dass v. a. von einer physischen und psychischen Leistungsfähigkeit bzw. Belastbarkeit ausgegangen wird, welche als Voraussetzungen für eine Anpassung an bestimmte Umweltgegebenheiten oder zur Bewältigung bestimmter Aufgaben gesehen werden können. Die Zusammenhänge dieser Faktoren haben für eignungsdiagnostische Überlegungen in verschiedenen Bereichen Bedeutung und bilden ein Potential für die Optimierung entsprechender Entscheidungsprozesse, die auf eine spezifische Entwicklung in der Diagnostik für die berufsbezogene Personalauswahl verweisen. Ganz besonders in solchen Handlungsfeldern mit hohen psycho-physischen Anforderungen, wie dies beispielhaft im militärischen Feld, etwa in der Bundeswehr und auch in Berufsfeldern wie Feuerwehr als auch Polizei der Fall ist, wo es in Kampfeinsätzen sogar um die Konfrontation mit lebensbedrohlichen Gefahren geht. Schuler und Funke (1991, S. 5) definieren Eignungsdiagnostik als:

„die Wissenschaft und Praxis der Zuordnung von Personen und Tätigkeiten. Die wissenschaftliche Seite ist die der Entwicklung von Theorien über den Zusammenhang von Fähigkeiten und Leistungen sowie Methoden, diese Merkmale zu messen und zu prognostizieren. Aufgabe der Eignungsdiagnostik in der Praxis ist es, sich dieser Methoden als Bestimmungs- und Entscheidungshilfen in einer Weise zu bedienen, die das Umfeld der wirtschaftlichen und sozialen Bedingungen berücksichtigt. Bei angemessener Anwendung liefert die berufsbezogene Eignungsdiagnostik nicht nur Entscheidungshilfen für die Organisation, die Mitarbeiter sucht und zu deren Weiterentwicklung sie beitragen möchte, sondern auch für den einzelnen Menschen, der sich für einen Beruf, einen Betrieb oder eine Umschulung zu entscheiden hat. Beide Seiten – Mensch und Organisation – profitieren von einem fähigkeits- und interessenentsprechenden beruflichen Einsatz, dessen Nutzen sich nicht nur in ökonomischen Größen bemessen lässt, sondern auch im Beitrag zur Zufriedenheit und Gesundheit, der dadurch zustande kommt, dass Über- und Unterforderung vermieden werden.“

Der Zusammenhang von Fähigkeiten und Anforderungen, sowie das Messen der korrespondierenden Merkmale, nimmt in der Personalauswahl eine besondere Stellung ein. Hierbei sollten vor dem Hintergrund der oben angeführten Krankheitsentwicklungen in der Arbeitswelt psychischen und physischen Faktoren sowie deren Zusammenhängen in der Berufseignungsdiagnostik zunehmend Aufmerksamkeit gewidmet werden, da diese grundlegende Bedeutung für die Gesundheit und das Wohlbefinden der Arbeitnehmer haben. Die in dieser Arbeit exemplarisch mit Bezug auf das militärische Handlungsfeld als relevant herausgestellten psycho-physischen Faktoren und deren Zusammenhänge können u. U. in diagnostischen Prozessen in anderen Handlungsfeldern einbezogen und darauf aufbauend Interventionsmaßnahmen deduziert werden. Wie in der Historie der Eignungsdiagnostik geschehen, könnten diese auf nicht-militärische Bereiche übertragen bzw. spezifiziert werden und lassen sich evtl. für das jeweilige Verfahren modifizieren, sodass sich diese Maßnahmen in dem jeweiligen Bereich im Kontext der Eignungsfeststellung bewähren.

Insbesondere in Folge der in den letzten Jahren wissenschaftlich aufgedeckten Problematik, dass im militärischen Bereich einsatzbedingte psychische Störungen wie posttraumatische Belastungsstörungen unerkannt bleiben und aufgrund dessen der Bedarf an Screeningverfahren zur Früherkennung zunimmt, ergibt sich für die Problemstellung dieser Arbeit eine spezielle Relevanz (siehe Wittchen, Schönfeld, Kirschbaum, Trautmann, Thurau, Siegert, Höfler, Hauffa & Zimmermann, 2013). Anhand des hier ausgearbeiteten Konzepts zur Feststellung der psycho-physischen Belastbarkeit im Kontext der Personalauswahl, soll ein Beitrag zur Erprobung und Weiterentwicklung weiterer Verfahrensüberlegungen für die Eignungsdiagnostik geleistet werden. Soldatische Tätigkeiten, ob im alltäglichen Dienstbetrieb oder verstärkt bei Einsätzen im Ausland, setzen sowohl in psychischer als auch physischer Hinsicht

eine hohe Belastbarkeit voraus. Das Vorhandensein einer gewissen physischen Robustheit im Sinne der körperlichen Belastbarkeit wird sogar als Grundeigenschaft für Soldaten bezeichnet (Bundeswehr, 2011). Letztendlich sind Soldaten, wie dies auch in anderen Berufsfeldern der Fall ist, der Gefahr ausgesetzt, durch einen hohen bis überwiegenden Anteil an sitzenden Tätigkeiten diese Eigenschaft (körperliche Belastbarkeit) zu verlieren bzw. auch unter Berücksichtigung des Alterns abzubauen. Eine derart „schleichende“ Abnahme der Leistungsfähigkeit und Belastbarkeit wird bereits mehrere Jahre vor der Manifestation einer Erkrankung erkennbar und im Vergleich zu gesunden Arbeitnehmern im Hinblick auf Produktivität und steigende Kosten deutlich (siehe hierzu z. B. Uhle & Treier, 2011; Leyk, 2012).

Wie bereits erwähnt, scheinen psychische Erkrankungen, die v. a. schwerer zu diagnostizieren sind, den Herz-Kreislauf-Erkrankungen den „Rang abzulaufen“. Verdeutlicht wird dieses Phänomen durch die Ergebnisse einer Untersuchung zu psychischen Störungen (Wittchen et al. 2013; zit. nach Süddeutscher Zeitung, Ausgabe 273, 26.11.2013), wo darauf hingewiesen wird, dass etwa jeder fünfte Soldat bereits mit einer manifesten, aber zumeist nicht erkannten, psychischen Störung in den Einsatz gehe, dies aber sogar noch unter dem gesellschaftlichen Durchschnitt liege. Weiterhin wird angeführt, dass bei dieser Gruppe der unerkannt vorbelasteten Soldaten ein vier- bis sechsfach höheres Risiko vorliege, aus dem Einsatz mit einer neuen Erkrankung zurückzukehren. Demzufolge dürfte verstärkt die Erprobung und Entwicklung eines Screening- und/oder Monitoringverfahrens als vielversprechend und praxisrelevant gesehen werden. Im weiteren Verlauf soll, wie bereits erwähnt, solch ein Beitrag zur Verfahrensentwicklung und -erprobung geleistet werden. Mit Erprobung ist in diesem Kontext v. a. das Aufzeigen einer Möglichkeit zur Operationalisierung der psycho-physischen Belastbarkeit zu verstehen und der damit verbundenen Überprüfung eines Zusammenhangs der herangezogenen Parameter.

Neben der anwendungsbezogenen Relevanz ist das Forschungsprojekt wesentlich von dem erkenntnisleitenden Interesse der weiteren Aufklärung psycho-physischer Zusammenhänge und gruppenbezogener Unterschiede getragen. Das „psycho-physische Problem“ hat grundlegende Bedeutung für die Psychologie. Es wird in diesem Projekt auf die psycho-physische Belastbarkeit konkretisiert und im theoretischen Teil spezifisch thematisiert.

1.3 Anfänge der Eignungsdiagnostik und Personalauswahl

Ein allgemeiner Bezug und Hintergrund dieser Arbeit ist im weiteren Bereich der Eignungsdiagnostik und im Hinblick eines Screenings im Zuge der Personalauswahl zu sehen. Die Feststellung der Eignung und Passung von Personen für eine Tätigkeit oder Aufgabe geht über 3000 Jahre zurück. Die Ursprünge der Eignungsdiagnostik lassen sich in etwa auf 2200 v. Chr. datieren, „when the Chinese emperor examined his officials every third year to determine their fitness for continuing in office“ (DuBois, 1970, S. 3), wobei sich Staatsdiener einer systematischen Überprüfung der Eignung für „Managementaufgaben“ im Kaiserreich unterziehen mussten (vgl. Schuler, 2000). Inwiefern hier der Begriff der Fitness als Form eines physischen Auswahlverfahrens gesehen werden kann oder hiermit eine andere Form der Fitness gemeint ist, wird nicht konkretisiert. Interessant ist auch, dass in diesem Zusammenhang der Fitnessbegriff auftaucht – offensichtlich nicht auf physische Fitness beschränkt oder reduziert und im Hinblick auf Passung zu verstehen ist, was der Wortbedeutung direkt entspricht. Mit dem heutzutage in der Alltagssprache üblichen Gebrauch des Fitnessbegriffs ist zumeist die physische Komponente gemeint und es kann aber angenommen werden, dass bei der ersten berufsbezogenen Eignungsdiagnostik nicht ausschließlich körperliche Leistungsaspekte erfasst wurden.

Moser und Schuler (2013) referieren in ihrem historischen Überblick der Management Diagnostik über eine „Testbatterie“, bei der sich chinesische Staatsdiener in den Disziplinen Bogenschießen, Reiten und Arithmetik behaupten mussten. Demnach spielten bereits bei dem ersten überlieferten eignungsdiagnostischen Verfahren neben geistigen auch körperliche Fertigkeiten eine entscheidende Rolle. Der Forscher Sir Francis Galton interessierte sich nachweislich (siehe DuBois, 1970) als Erster für Messungen und die Zusammenhänge physischer und psychischer Merkmale und untersuchte erstmals die Ähnlichkeiten von physischen und mentalen Charakteristika bei erfolgreichen Generationen. Dabei entdeckte er die Notwendigkeit einer Methode, die es ermöglichte, den Zusammenhang zwischen zwei Variablen zu beschreiben und legte den Grundstein für die Entwicklung von statistischen Verfahren wie die Korrelationsanalyse (siehe DuBois, 1970, Barthel, 1988). Galton eröffnete in Verbindung mit der International Health Exhibition ein Labor, wo er sowohl physische Leistungsparameter (Sauerstoffaufnahmekapazität, Krafttests, etc.) erfasste als auch die Versuchspersonen auf ihre geistigen Fähigkeiten hin (Gedächtnisleistung, usw.) testen lies (siehe DuBois, 1970), um unter anderem Zusammenhänge zwischen der körperlichen Konstitution und bestimmten

Verhaltensweisen sowie Charaktereigenschaften der Teilnehmer zu analysieren. Der amerikanische Psychologe James McKeen Cattell, ein Assistent, der in Galtons Labor mitarbeitete, brachte Galtons Ansatz in die Vereinigten Staaten und gründete im Jahre 1891 an der Columbia University in New York ein psychologisches Labor. Dort entwickelte er eine Testbatterie, welche Messungen von mentalen, sensorischen und physischen sowie motorischen Fähigkeiten beinhaltete (siehe DuBois, 1970).

Die erste umfangreiche systematische psychologische Eignungsdiagnostik fand im Rahmen eines Personalauswahlprogramms der deutschen Armee für Spezialkräfte und Offiziere statt. Die Niederlage im ersten Weltkrieg wurde mitunter der schlechten Leistung der deutschen Soldaten zugeschrieben, bspw. auch aufgrund von mangelndem emotionalen Anpassungsvermögen. „Simoneit attributed the disappointing performance of German soldiers to (a) poor emotional adjustment to life at the battle front“ (Wiggins, 1973, S. 516). Die deutsche Armee sah die Notwendigkeit eines neuen Personalauswahlverfahrens auf Grundlage von psychologischen Testungen im Sinne eines Assessment Centers, bei dem jedes Jahr in 15 psychologischen Laboren der Armee die Anwärter auf verschiedene Parameter hin untersucht wurden. Das besondere bei dieser Eignungsdiagnostik war die Observierung und Eignungsfeststellung nach ganzheitlichen Gesichtspunkten. D. h. die Soldaten wurden auf ihre Handlungsreaktion in militärspezifischen Situationen im Hinblick auf Führung, Handeln in Stresssituationen, auf das Durchhalte- bzw. Ausdauervermögen, etc. von Offizieren und Militärpsychologen untersucht. Mit dieser Herangehensweise wurden die Fähigkeiten des Kandidaten nicht mehr isoliert betrachtet, sondern im Kontext und in Relation zueinander gesehen und eine Verbindung der erhobenen Parameter bzw. beobachteten Fähigkeiten hergestellt (siehe Wiggins, 1973).

„The candidate was no longer viewed as the sum of his isolated abilities (e.g., knowledge of the classics), but rather as a 'total personality' whose potentialities must be globally assessed in realistic situations that simulate the press of the criterion environments in which he is expected to perform.“ (Wiggins, 1973, S. 517)

Nach dem zweiten Weltkrieg begannen erstmalig Unternehmen diese militärspezifischen Personalauswahlverfahren und -prozesse zu übertragen und in der Wirtschaft und Industrie zu implementieren. Diese multiplen Verfahren zur Offiziersauswahl in der Weimarer Republik können als Vorläufer der heutigen Assessment Center zur Führungskräfteauswahl gesehen werden (vgl. Plate, 2006).

Bereits in den 50-er Jahren wurde am Institute of Personality Assessment and Research (IPAR) an der University of California in Berkeley mit 343 Offizieren der amerikanischen Air Force eine empirische Studie zur Eignungsfeststellung durchgeführt. Die Besonderheit bei dieser Studie war die Eignungsfeststellung anhand von mehreren Kriterien und deren Zusammenhänge für die Evaluierung eines erfolgreichen Offiziers, darunter sowohl Kriterien physischer Widerstandsfähigkeit und Robustheit, als auch Kriterien und Merkmale die Persönlichkeit betreffend:

„Health and vitality. Consistent good health and stability of physiological functioning. Absence of minor as well as major illness. Resistance to disease. Ability to withstand stress and endure hardship (. . .). Vigor, robustness, stamina, sense of physical well-being.“ (MacKinnon, 1958, S. 1 f.)

Das herauszustellende Ziel dieser Studie bestand darin, Zusammenhänge zwischen diesen verschiedenen Eignungskriterien zu identifizieren, um die besten Kandidaten für Führungspositionen ausfindig zu machen. Wiggins (1970, S. 540) definiert nach McKinnon (1958, S. 1) die beiden Leitfragen der Studie:

„(1) What patterns of ability, motivation and interest are associated with outstanding effectiveness of Air Force officers in their military assignments?’ and (2) ‘By what psychological tests and procedures or by what combinations of such techniques can the distinguishing characteristics and the potential promise of the effective officer be identified and measured?’“

Es kann somit konstatiert werden, dass die Anfänge der Eignungsdiagnostik, speziell die Kombination eignungsdiagnostischer Instrumente und Verfahren in Form des Assessment Centers, in einem engen Verbund mit der Personalauswahl im militärischen Bereich stehen. Nach Hull (1928, S. 5) sinnierte bereits Plato im alten Griechenland über die Auswahl von Soldaten durch Tests zum Erinnerungsvermögen, zur Widerstandskraft und u. a. zur Müdigkeit, welche die Eignung für soldatische Tätigkeiten feststellen sollten (vgl. Barthel, 1989). Die Testung der Eignungsfeststellung von Soldaten ging den Tests und anderen systematischen Verfahren zur Bewerberauswahl im industriellen und wirtschaftlichen Bereich voran (Barthel, 1989). Im zweiten Weltkrieg erreichte die Testung ihren Höhepunkt als 14 Millionen amerikanische Soldaten getestet wurden. Dies war nach Thorndike (1949) die Grundlage für die Entwicklung „brauchbarer Kriterien zur Messung des Berufserfolgs - nicht zuletzt, um damit auch den Nutzen von Auswahlmethoden unter Beweis stellen zu können“ (Barthel, 1989, S. 28). Insofern kann der militärische Bereich als Grundlage für die Entwicklung, Stan-

dardisierung und Verbreitung von Selektions- und Personalauswahlverfahren in anderen Bereichen verstanden werden. Nicht zuletzt befindet sich der Vorläufer des heutigen Assessment-Centers, als Diagnostikum für die Auswahl der passenden Mitarbeiter und Manager, in der Offiziersauswahl der Reichswehr in der Weimarer Republik. Hier wurden mehrere Bewerber innerhalb von zwei bis zweieinhalb Tagen „verschiedenen diagnostischen Prozeduren ausgesetzt“ (Moser & Schuler, 2013, S. 37), die von mehreren supervisierenden Personen begutachtet und anschließend ergebnisbasierend beraten wurden. Diese Ansätze sollen aufgegriffen und fortgeführt werden.

Für diese Arbeit ergibt sich folgender Aufbau (siehe Abbildung 2): Nach dem allgemeinen Problemaufriss, welcher vor dem Hintergrund aktueller Entwicklungen in der Arbeitswelt dargestellt wurde und der Charakterisierung der drei wesentlichen Anliegen dieser Arbeit, folgte die Erläuterung der Relevanz dieses Forschungsvorhabens. Da die diagnostischen Verfahren zur Untersuchung der psycho-physischen Belastbarkeit weitestgehend exemplarisch im militärischen Kontext entwickelt werden und sich daraus Implikationen für andere Bereiche - bspw. für Wirtschaft und Industrie - ergeben können, wie es sich in der Vergangenheit bereits in Form des Assessment Centers gezeigt hat, wurde im Rahmen der Einführung auch auf den historischen Hintergrund der Eignungsdiagnostik und Personalauswahl kurz eingegangen. Im Anschluss an die Einführung wird das theoretische Rahmenkonzept für diese Arbeit erarbeitet. Hierbei sind zunächst die Begriffe Belastung, Beanspruchung und (psycho-physische) Belastbarkeit und deren Zusammenhänge insbesondere vor dem Hintergrund arbeits- und sportwissenschaftlicher Überlegungen zu klären, bevor das hier konzipierte psycho-physische Belastbarkeitskonzept vorgestellt wird. Daran anschließend wird auf die beiden Konstrukte für die Erfassung der beiden hier relevanten Merkmale eingegangen. Nach der Vorstellung des theoretischen Rahmenkonzepts wird auf die für das Forschungsvorhaben bedeutenden Bereiche Bezug genommen und dazu wichtige Arbeiten und Beiträge gesichtet. Aufbauend auf den vorherigen Kapiteln folgt nun der empirisch-methodische Ansatz für die Verfahrensentwicklung zur Erfassung der beiden hier interessierenden Merkmale und die späteren Analysen zur psycho-physischen Belastbarkeit. Dieser Teil umfasst u. a. die zentralen Fragestellungen und Hypothesen, sowie die grundsätzlichen methodischen Herangehensweisen, mit denen die anschließenden Untersuchungen realisiert werden sollen. Den Hauptteil dieser Arbeit stellt Kapitel 5 dar.

Die Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente beinhalten in einem ersten Schritt die Explorationsstudie zur Kriteriendefinition des Testverfahrens und als nächsten Schritt die Studie zur Überprüfung der Gütekriterien des neu konzipierten Ausdauertests. Die Laborstudie soll neben der Erkenntnis über die Validität und Reliabilität des Verfahrens auch weitere Indikatoren für die physische Belastbarkeit liefern, die in die Untersuchungen der psycho-physischen Belastbarkeit (Kapitel 5.3) miteinbezogen werden. Dieser Laborstudie kommt somit aus zweierlei Gesichtspunkten eine dementsprechende Gewichtung in dieser Arbeit zu Teil. Neben der Verfahrensentwicklung der physischen wird in Kapitel 5.2 die Untersuchungen zur Verfahrensentwicklung der psychischen Belastbarkeitskomponente beschrieben. Hierbei werden ebenfalls Instrumente und Methoden eingesetzt, die für die darauf folgende Untersuchung der psycho-physischen Belastbarkeit Berücksichtigung finden und mit in die Analysen einbezogen werden. Nach den Untersuchungen und Analysen sowie der Darstellung der Ergebnisse, werden die Befunde in Kapitel 6 interpretiert und diskutiert. In Anbetracht der Befunde und deren Interpretation sowie dem Rückbezug auf die theoretischen Ausführungen, folgt der Schlusskommentar und ein Ausblick. Zuletzt werden die wesentlichen Inhalte in Kapitel 8 zusammengefasst.

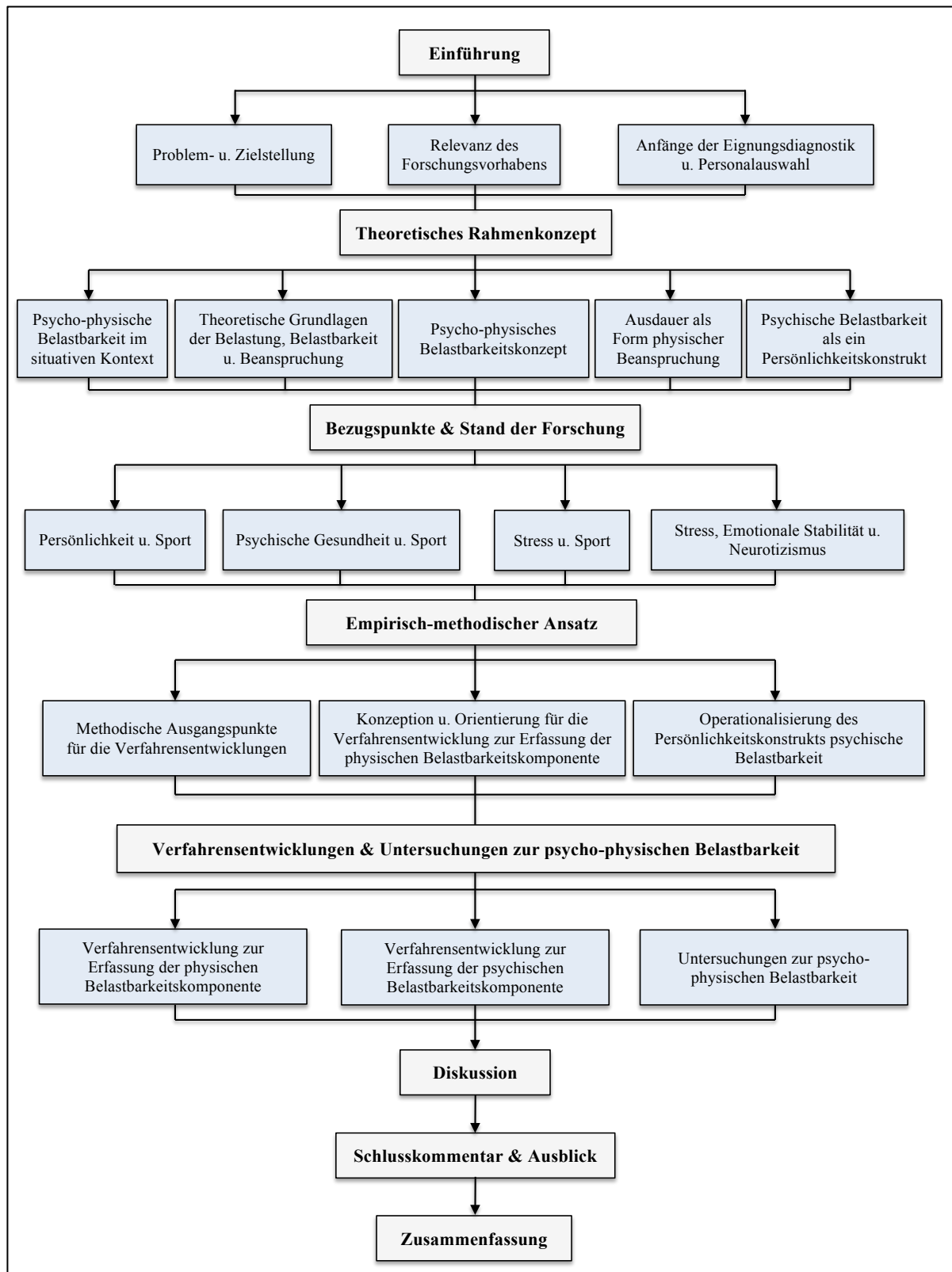


Abbildung 2. Aufbau der Arbeit.

2 THEORETISCHES RAHMENKONZEPT

In diesem Abschnitt der Arbeit wird im ersten Schritt der psycho-physische Zusammenhang mit spezieller Berücksichtigung der psycho-physischen-Belastbarkeit erläutert und in einen theoretischen Bezugsrahmen eingeordnet. Das Konstrukt der psycho-physischen Belastbarkeit wird dazu in den Kontext des Belastungs-Beanspruchungs-Konzepts gestellt und sein Erklärungspotenzial diskutiert. In Kapitel 2.3 wird dann das psycho-physische Belastbarkeitskonzept, welches den Kern dieser Arbeit darstellt, genauer beschrieben und erläutert. Inhaltlich geht es bei der Vorstellung der Konzeptualisierung vornehmlich darum, die theoretischen Grundlagen dafür aufzuzeigen, wie die im Fokus stehenden Konstrukte anschließend operationalisiert werden und auf welchem Wege die Zusammenhänge der psychischen und physischen Komponenten ausgelotet werden, um so eine Grundlage für das Verständnis der empirischen Befunde zu bilden.

2.1 Psycho-physische Belastbarkeit im situativen Kontext

„Die Einsätze prägen auch die Anforderungen an die Soldaten der Bundeswehr, (. . .). Sie müssen fest damit rechnen, dass Ihre körperliche und psychische Belastbarkeit im Einsatz gefordert wird. Dafür brauchen Sie eine gefestigte und belastbare Persönlichkeit. Sie müssen unübersichtliche und gefährliche Situationen bestehen.“ (De Maizière, 2012, Abs. 20).

Die Aussage des ehemaligen Bundesministers der Verteidigung verweist auf die Relevanz sowohl der psychischen als auch der physischen Belastbarkeitskomponente im Kontext des Soldatenberufs sowie auf die Brauchbarkeit einer stabilen Persönlichkeit im Sinne einer psychischen Belastbarkeit. Soldaten können während Einsätzen, Ausbildungs- oder Weiterbildungsmaßnahmen hohen Belastungen, sowohl physischer als auch psychischer Art ausgesetzt sein. Kritische Situationen und psychotraumatische Ereignisse wie Unfälle, Schusswaffengebräuche, tätliche Angriffe oder Anschläge, unter zum Teil extremen Umwelt- und Arbeitsbedingungen, als auch das Transportieren von Lasten sind starke psycho-physische Belastungen (Stressoren) für das Individuum und können ausgeprägte Stressreaktionen hervorrufen. Entscheidende Faktoren in solchen schwierigen Handlungssituationen sind hohe Belastbarkeit, Durchhaltevermögen und Leistungsfähigkeit der Soldaten (Leyk et al., 2013). Diese Situationen hoher Belastung können über körperliche Konsequenzen hinaus auch emotionale Folgen

haben und unter Umständen in Abhängigkeit von der individuellen psychischen Stabilität des Soldaten zu Belastungsstörungen führen.

Die Begriffe Stress und Belastung sind aus wissenschaftlicher Sicht nicht als Synonym zu verstehen. „In jüngerer Zeit“, so Udris und Frese (1999, S. 429), „hat sich ein Minimalkonsens entwickelt, der die angelsächsische ‚Stresstradition‘ und die deutschsprachige Belastungs- und Beanspruchungstradition zusammenführt.“ Stress ist als eine Form der Überbeanspruchung zu sehen. Den Begriff „Belastung“ kann man weiter differenzieren in psychische und physische Belastung. In dieser Arbeit und in diesem Kontext kommt sowohl der Ausdauerbelastung auf der physischen Seite und der psychischen Belastbarkeit, auf der anderen Seite eine besondere Bedeutung zu. Aus trainingswissenschaftlicher Sicht erhöht sich mit zunehmender Ausdauer nicht nur die physische, sondern auch die psychische Widerstandsfähigkeit (Belastbarkeit) gegenüber leistungsmindernden Ermüdungssymptomen (Schnabel, Harre & Krug, 2008). Die Weltgesundheitsorganisation definiert (siehe ICD-10; F43.0) eine akute Belastungsstörung als "eine vorübergehende Störung, die sich bei einem psychisch nicht manifest gestörten Menschen als Reaktion auf eine außergewöhnliche physische oder psychische Belastung entwickelt.“

Analog der gesundheitlichen Entwicklung der Gesellschaft belegen internationale Studien und ein NATO-Bericht, dass die Leistungsfähigkeit auch bei den Soldatinnen und Soldaten tendenziell abzunehmen scheint, obwohl die Annahme verbreitet ist, dass körperliche Betätigungen zur alltäglichen Aufgabe der Soldaten gehören (Leyk et al., 2013). Den Trend der abnehmenden Leistungsfähigkeit zeigen z. B. Längsschnittuntersuchungen von norwegischen Wehrpflichtigen über einen Zeitraum von zwei Dekaden, bei denen die maximale Sauerstoffaufnahme als Indikator für die Ausdauerleistungsfähigkeit um acht Prozent gesunken ist und demgegenüber der Body-Mass-Index sowie die Körpermasse um sechs bzw. sieben Prozent zugenommen haben (Leyk et al., 2013). Die US-Streitkräfte haben indes auf den Rückgang der gesundheitlichen Entwicklungen reagiert und verfolgen seit 2010 einen Präventionsansatz namens „Total Force Fitness for the 21st Century“, der neben körperlichen u. a. auch psychische Komponenten der alltäglichen Leistungsanforderungen und Belastungen der Soldaten einbezieht und fördert (Leyk et al., 2013, S. 165). So gesehen erscheint ein Interesse an psycho-physischen Zusammenhängen, insbesondere das Erforschen von physischer Fitness in Verbindung mit psychischer Stabilität bzw. Belastbarkeit insbesondere im Soldatenberuf speziell relevant und indiziert zu sein.

In einer Untersuchung über die Bedeutung und Effektivität der Widerstandsfähigkeit („hardiness“) und des Durchhaltevermögens („grit“) zur Voraussage der Leistungsfähigkeit („performance“) von US Marine Kadetten im ersten Ausbildungsjahr von Maddi et al. (2012, S. 26) weisen die Autoren auf ein validiertes psychisches Widerstands- bzw. Belastbarkeits-training hin („HardiTraining“), welches möglicherweise zu einer Leistungssteigerung führt. Neben der Trainierbarkeit der physischen Komponente wird von Seiten der Autoren auch eine Beeinflussbarkeit des Persönlichkeitsfaktors hardiness durch Training und somit der psychischen Komponente angenommen. Insbesondere in Stresssituationen oder in gesundheitsgefährdeten Umgebungen, wie in Kampfeinsätzen spielen Persönlichkeitsfaktoren eine entscheidende Rolle für die Vorhersagbarkeit der Leistung.

„There is increasing evidence that personality factors play a key role in determining human performance in a variety of contexts. (. . .) This may be particularly true in settings that are highly stressfull, challenging, or that even pose a threat to physical well-being (. . .).“ (Maddi et al., 2012, S. 20)

Der Persönlichkeitsfaktor hardiness hat sich nach den Autoren (Maddi, 2002; Maddi, Harvey, Khoshaba, Fazel & Resurreccion, 2009a; Maddi, Harvey, Khoshaba, Fazel & Resurreccion, 2009b; Maddi, Harvey, Khoshaba, Lu, Persico & Brow, 2006; Maddi, Harvey, Resurreccion, Giatras & Raganold, 2007; Maddi & Hess, 1992; Maddi & Kobasa, 1984) als gesundheitsfördernder und leistungssteigernder Faktor in belastenden Situationen erwiesen: „Hardiness has been shown to enhance performance and health despite stress in a variety of situations and occupational settings, such as business, sports, firefighting, and college.“ (Maddi et al., 2012, S. 21) Die Befunde der Untersuchung von Maddi et al. (2012) zeigen eine signifikante Vorhersagegenauigkeit der beiden Persönlichkeitsfaktoren hardiness und grit im Hinblick auf die Retention (Mitarbeiterbindung) der Soldaten. Auch wenn die Interpretation von hardiness als gesundheitsförderlicher oder leistungsförderlicher Faktor sehr weit geht und Korrelationsanalysen diese Zusammenhänge nicht eindeutig hergeben, lässt sich nachvollziehen, dass die Autoren in diesem Kontext empfehlen, den Persönlichkeitsfaktor hardiness bei Nachwuchskräften zu messen und als mögliches Selektionskriterium für die Personalauswahl heranzuziehen.

Ergänzend dazu wird in der Arbeitspsychologie auf den Persönlichkeitsaspekt „negative affectivity“ (siehe Udris & Frese, 1999, S. 433) hingewiesen. Er soll in diesem Zusammenhang eine wichtige Rolle spielen, denn bei der Tendenz eines Individuums, Dinge vorwiegend

negativ zu beurteilen, korrelieren bei diesem Merkmal die Faktoren Stressoren und Stressauswirkung stark (vgl. Spector, Zapf, Chen & Frese, 1998). Folgen von zu starker Belastung oder Überbeanspruchung sind bekannte Phänomene, z. B. das Burnout („Ausbrennen und Ausgebranntsein“; siehe Kern, 1997) und das Phänomen der „Inneren Kündigung“, mit dem ein abnehmendes Engagement bis hin zur Leistungsverweigerung einhergeht (vgl. Faller, 1993; Krystek, Becherer & Deichmann, 1995).

Neben der Bundeswehr stehen auch die Wirtschaft und andere Bereiche zukünftig aufgrund der demografischen Entwicklung, dem Anstieg des Bildungsniveaus, des Gesundheitszustands in der Bevölkerung, sowie des gesellschaftlichen Wertewandels vor Herausforderungen bei der Personalgewinnung (siehe dazu National Audit Office, 2006). Laut dem Bundesministerium der Verteidigung (2010) kann die Bundeswehr den Bedarf an Soldatinnen und Soldaten nicht decken. So seien seit Jahren 7000 Stellen nicht besetzt. Auf eine Anfrage der Bundeszentrale für politische Bildung seitens des Bundesministeriums der Verteidigung (2001) wurde konstatiert, dass in Folge der Strukturreform der Bundeswehr (die Aussetzung der Wehrpflicht zum 01. Juli 2011) die Gesamtstärke auf 210550 Soldaten reduziert wurde. Die Bundeswehr ist seitdem eine Freiwilligenarmee mit insgesamt 188000 Berufs- und Zeitsoldaten und 22550 freiwillig länger Wehrdienst Leistenden (FWDL). Es besteht also ein Ergänzungsbedarf von 18400 Soldaten auf Zeit und 4000 FWDL. Die Sollstruktur soll im Rahmen der Strukturreform weiter reduziert werden (Bundeszentrale für politische Bildung, 2011). Bestehen bleibt weiterhin die Herausforderung der Personalgewinnung für die Bundeswehr.

Die Hauptzielgruppe sind ungediente Frauen und Männer mit deutscher Staatsbürgerschaft zwischen 17 und 24 Jahren, die sich in der Orientierungsphase für Ausbildung und Beruf befinden. Dabei wird vor allem gut ausgebildetes Personal mit einer hohen physischen und psychischen Belastbarkeit nachgefragt. Das definierte soldatische Anforderungsprofil des Generalinspektors sieht vor, hervorragend qualifizierte und motivierte Soldaten mit einem hohen Maß an Sozialkompetenz, Führungsstärke und Flexibilität, die in einem komplexen Umfeld, schnell, vorausschauend und mit der nötigen Umsicht agieren können. Abgesehen von der fortschreitenden Technologisierung militärischer Aufgaben erfordert eine Vielzahl der Tätigkeiten eines Soldaten ein hohes Maß an konditioneller Leistungsfähigkeit. Insbesondere die Gewichtsbelastung durch Ausrüstung und das Heben und Tragen von schweren Lasten im Einsatz bei hoher Belastungsdauer und extremen Bedingungen setzen gute konditionelle Fä-

higkeiten (physische Fitness) wie Ausdauer, Kraft, Schnelligkeit, Koordination und Beweglichkeit, sowie auch eine gewisse psychische Stabilität voraus (Witzki, Sievert, Gorges, Glitz, Kreim & Leyk, 2010).

Angesichts der hohen physischen und psychischen Anforderungen im Soldatenberuf und empirischer Analysen zu Gesundheitsrisiken und Defiziten in der Bevölkerung hat der Sanitätsdienst der Bundeswehr die physische Leistungsfähigkeit von Bewerbern analysiert. Dabei wurde festgestellt, dass zwischen den Jahren 2000 und 2004 insgesamt 39 Prozent der weiblichen und 37 Prozent der männlichen Bewerber die Minimalanforderungen der physischen Leistungsfähigkeit, erfasst durch den zu der Zeit eingesetzten Physical Fitness Test (PFT), nicht erfüllten. Bei den Untersuchungen war besonders auffällig, dass es bei den männlichen Bewerbern ab dem Jahr 2001 zu einem signifikanten Anstieg der PFT-Durchfallquoten kam. Daher wurden die physischen Leistungsanforderungen im Annahmeverfahren weiter gesenkt (Leyk, Rhode, Gorges, Ridder, Wunderlich, Dinklage, Sievert, Rütter & Essfeld, 2006; vgl. Bundesministerium der Verteidigung, Hrsg., Bericht des Generalinspektors der Bundeswehr zum Prüfauftrag aus der Kabinettsklausur vom 7. Juni 2010, S. 25)

In der „Weisung zur Ausbildung und zum Erhalt der Individuellen Grundfertigkeiten“ (Weisung IGF) der Bundeswehr kommt die Relevanz der Anforderungen an die Soldaten hinsichtlich der psychischen und physischen Konstitution bzw. Leistungsfähigkeit zum Ausdruck:

„Sie fordert von den militärischen Angehörigen der Streitkräfte ein hohes Maß an physischer und psychischer Robustheit sowie persönlicher Flexibilität, um die besonderen Anforderungen des Soldatenberufs und der soldatischen Gemeinschaft anzunehmen und zu verinnerlichen [und] darüber hinaus muss jeder Soldat / jede Soldatin eine ausreichende physische und psychische Leistungsfähigkeit besitzen, um die Anforderungen des soldatischen Dienstes insbesondere mit Blick auf die Einsätze jederzeit erfüllen zu können.“ (IGF, 2009, S. 4)

Die „Weisung zur Ausbildung und zum Erhalt der Individuellen Grundfertigkeiten“ der Bundeswehr hebt die besondere Bedeutung sowohl physischer als auch psychischer Leistungsfähigkeit im Soldatenberuf hervor und betont die Notwendigkeit zur Ausbildung beider Faktoren, nämlich physischer und psychischer Robustheit, die im Kontext dieser Arbeit im Vordergrund stehen und wobei von einer Abhängigkeit bzw. Wechselbeziehung der beiden Faktoren ausgegangen wird. Die Bundeswehr und somit der militärische Bereich stellt somit konkrete Anforderungen an seine Mitarbeiter im Hinblick auf beide Faktoren. Mit dieser Arbeit soll auf

die Relevanz dieser Faktoren hinsichtlich berufsbezogener Diagnostik nicht nur für den militärischen, sondern auch im Kontext anderer Bereiche aufmerksam gemacht werden. Dies gilt v. a. vor dem Hintergrund von Berufsgruppen und Jobs, die nicht nur eine hohe psychische Belastbarkeit, sondern auch eine eventuell damit zusammenhängende körperliche Leistungsfähigkeit bzw. Robustheit erfordern.

2.2 Theoretische Grundlagen der Belastung, Belastbarkeit und Beanspruchung

Das Belastungs-Beanspruchungskonzept stellt sowohl im Rahmen der Sportwissenschaft, speziell in den Teildisziplinen Trainingswissenschaft (u. a. Schnabel, Harre & Krug, 2008) und Sportpsychologie (u. a. Nitsch & Udris, 1976), als auch in der Arbeitswissenschaft (z. B. Rohmert, 1983), hier insbesondere in der Arbeitspsychologie (u. a. Kleinbeck & Rutenfranz, 1987) und Arbeitsmedizin (z. B. Rohmert & Rutenfranz, 1975) sowie bei entsprechenden sozialwissenschaftlichen (u. a. Richter & Hacker, 1998) Fragestellungen ein grundlegendes Konzept dar, wenn es z. B. um die Erklärung möglicher Risikokonstellationen an den Schnittstellen von Mensch-Tätigkeit-Gesundheit geht. Mit dem Rückgriff auf das Konzept geht häufig das Anliegen und das Bemühen einher, humane Phänomene insbesondere im Kontext der Arbeit in einen (möglichst kausalen) Erklärungszusammenhang zu bringen.

Zur Koordination und Integration unterschiedlicher arbeitswissenschaftlicher Perspektiven auf das Arbeitsleben, dessen Analyse, Beschreibung, Erklärung und daraus abgeleiteten Vorschlägen zur Optimierung werden unter Belastungen die äußeren Merkmale bzw. die Anforderungen der Arbeitssituation, wie die Arbeitsaufgabe, Arbeitsbedingungen, Zeitdruck, hohe Verantwortung, wechselnde Technologie sowie die soziale Interaktion mit den betrieblichen Akteuren verstanden (vgl. Kuhn, 2007, S. 27). Als Beanspruchungen, die in verschiedenen Formen (z. B. Monotonie, Stress, psychische Sättigung) auftreten können, werden die Reaktionen (körperlich-physiologisch, psycho-somatisch sowie erlebens- und verhaltensmäßig) gekennzeichnet. Die jeweilige Beanspruchung wird dabei allerdings nicht nur als eine Funktion der Belastung gesehen, sondern es werden weitere Einflussfaktoren, sozusagen moderierende oder modifizierende Einflussgrößen (intervenierende Variablen), veranschlagt. Dazu wird z. B. auf individuelle Eigenschaften und Fähigkeiten oder den Gewöhnungsgrad an die Arbeitsanforderungen (vgl. Schlick, Bruder & Luczak, 2010, S. 38 f.) sowie die psychophysische Belastbarkeit hingewiesen.

Oppolzer (1989) verweist darauf, dass die jeweilig eintretende Beanspruchung bzw. die Beanspruchungsfolgen je nach Art und Grad sowie je nach den Möglichkeiten ihrer Bewältigung bzw. Bewältigbarkeit und im Hinblick auf persönliche, soziale und ökonomische Konsequenzen sehr unterschiedlich ausfallen können. Dementsprechend ist davon auszugehen, dass Über-, Fehl- und Unterbeanspruchungen zu Beeinträchtigungen des Wohlbefindens, der Gesundheit und wesentlich auch der Leistungsfähigkeit führen, insbesondere wenn die Grenzen der individuellen Reproduktions- und Anpassungsfähigkeit überschritten werden (vgl. Oppolzer, 1989, S. 25).

In der folgenden Tabelle 1 ist systematisch zusammengefasst, wie sich Fehlbeanspruchungen einerseits kurzfristig und andererseits langfristig oder sogar dauerhaft auf körperlicher, psychischer und sozialer Ebene darstellen:

Tabelle 1

Beispiele für Beanspruchungsfolgen bzw. Stressreaktionen (nach Greif, 1991; Kaufmann, Pornschlegel & Udris, 1982; Kaluza, 1996; aus Udris & Frese, 1999, S. 432).

	kurzfristig, aktuelle Reaktionen	mittel- bis langfristige chronische Reaktionen
physiologisch, somatisch	<ul style="list-style-type: none"> ▪ erhöhte Herzfrequenz ▪ erhöhter Blutdruck ▪ Ausschüttung von Cortisol und Adrenalin („Streßhormone“) 	<ul style="list-style-type: none"> ▪ Psychosomatische Beschwerden und Erkrankungen ▪ Unzufriedenheit ▪ Resignation ▪ Depressivität ▪ Burnout
psychisch, kognitiv- emotional	<ul style="list-style-type: none"> ▪ Anspannung, Nervosität, innere Unruhe ▪ Frustration ▪ Ärger ▪ Ermüdungs-, Monotonie-, Sättigungsgefühl 	
Verhalten, individuell	<ul style="list-style-type: none"> ▪ Leistungsschwankungen ▪ Nachlassen der Konzentration ▪ Fehlhandlungen ▪ schlechte sensumotorische Koordination ▪ Hastigkeit und Ungeduld 	<ul style="list-style-type: none"> ▪ vermehrter Nikotin-, Alkohol-, Tablettenkonsum ▪ Fehlzeiten (Krankheitstage) ▪ innere Kündigung
Verhalten, sozial	<ul style="list-style-type: none"> ▪ erhöhte Reizbarkeit ▪ Konflikte ▪ Mobbing ▪ Streit ▪ Aggressionen gegen andere ▪ Rückzug (Isolierung) innerhalb und außerhalb der Arbeit 	

Physische wie psychische Fehlbelastungen stehen den Bedingungen für ein physisches, psychisches und soziales Wohlbefinden entgegen. Konsequenzen schlagen sich speziell in Befindlichkeitsstörungen nieder, die sich bei längerer Dauer oder wiederholtem Auftreten zu Gesundheitsbeschwerden bis hin zu Erkrankungen mit langfristigen Folgen entwickeln kön-

nen. Von daher ist es aus arbeitswissenschaftlicher Sicht relevant aufzuklären, inwiefern arbeitsbedingte Gesundheitsgefahren mit dem Risiko verbunden sind, zu korrespondierenden arbeitsbedingten Gesundheitsbeeinträchtigungen zu führen (vgl. Kuhn, 2007, S. 30).

Beispielhaft kann hoher Zeit- und Leistungsdruck zu spezifischen Beanspruchungsfolgen wie Anspannung, Nervosität, Frustration und Ärger führen und sich darüber in Befindlichkeitsstörungen und Unzufriedenheit manifestieren (siehe Tabelle 1). Zu Gesundheitsbeeinträchtigungen können beispielsweise auch Belastungen führen, wie sie durch die Arbeitszeit bzw. Arbeitszeitgestaltung etwa in Form von Wechselschicht, Nachtarbeit oder Wochenendarbeit ergeben. Damit sind häufig Beanspruchungsfolgen wie erhöhte Herzfrequenz, erhöhter Blutdruck und Ausschüttung von Stresshormonen verbunden, die sich zu gravierenden Folgen für das Herz-Kreislauf-System bis hin zur Herzinfarktgefährdung entwickeln.

Mit der Diskussion der aufgezeigten Zusammenhänge bekommt der Begriff Belastbarkeit Relevanz, und zwar sowohl hinsichtlich des physischen wie psychischen Aspekts. Im Folgenden soll zunächst auf die Bedeutung und Begrifflichkeit der Belastbarkeit eingegangen werden um zu einem besseren Verständnis im Gang der weiteren Erläuterungen in dieser Arbeit beizutragen. Aus trainingswissenschaftlicher Perspektive beschreiben Schnabel, Harre und Krug (2008, S. 243) Belastbarkeit als

„die Fähigkeit des Organismus und des psychischen Funktionssystems, Beanspruchungen ohne Störung der Gesundheit zu tolerieren. (. . .) Belastbarkeit bedeutet, dass physische und psychische Belastungen ohne Störungen der Gesundheit oder Trainierbarkeit verarbeitet werden und sich Zustandsgrößen von Geweben und Systemen des Organismus nach Belastungen wiederherstellen können.“

In Abbildung 3 sind die Kategorien der Belastbarkeit, so wie sie von den Autoren angesprochen werden, anschaulich und übersichtlich dargestellt. Hinsichtlich der Belastungsart und -qualität, der Belastungsquantität und -intensität, sowie im Hinblick auf häufige Beeinträchtigungen der Belastbarkeit bzw. Gesundheit lassen sich nach Fröhner (1993, 2007) sowie nach Martin und Nicolaus (1998) folgende Kategorien unterscheiden:

- die allgemein-organismische Belastbarkeit
- die mechanische Belastbarkeit
- die Belastbarkeit der leistungsbestimmenden Systeme
- die psychische Belastbarkeit

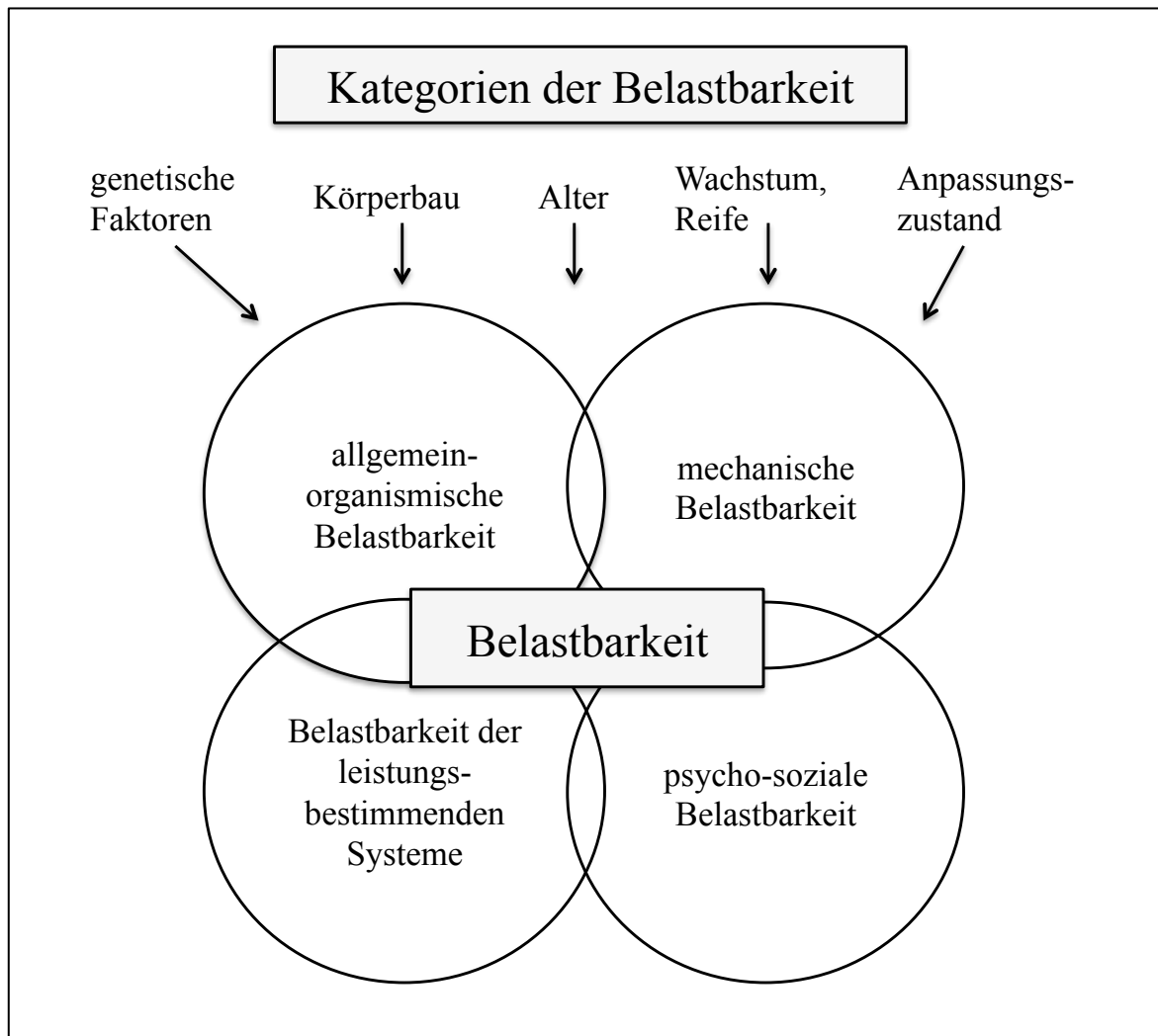


Abbildung 3. Kategorien der sportlichen Belastbarkeit (nach Fröhner, 1993; 2007; Martin & Nicolaus, 1998; aus Schnabel, Harre & Krug, 2008, S. 244).

Die Differenzierung der Belastbarkeit, so wie sie Schnabel et al. (2008) vornehmen, lässt folgende Kritikpunkte aufkommen: Die psychischen Aspekte sind gegenüber den anderen Kategorien unterrepräsentiert und unzureichend differenziert. Zudem wird die psychische Belastbarkeit nicht gesondert, sondern nur in Verbund mit der sozialen Belastbarkeit gesehen. Des Weiteren ist eine genaue Differenzierung der beiden Kategorien „allgemein-organismische Belastbarkeit“ und „Belastbarkeit der leistungsbestimmenden Systeme“ nicht auszumachen bzw. nicht eindeutig und aufgrund dessen nicht nachvollziehbar. Gleichwohl lässt sich das Bemühen erkennen, Belastbarkeit differenziert in Teilkomponenten (Kategorien) zu strukturieren und dabei auch den psychischen Bereich zu berücksichtigen. Dies bedarf allerdings weiterer Ausarbeitung. Auch dazu sollen die folgenden Ausführungen beitragen.

In der Arbeitswissenschaft werden Belastungen als Reize (Einflüsse) gesehen, die während der Arbeitstätigkeit auf das Individuum einwirken, wie bspw. Umweltbedingungen in denen die Person tätig ist, die Art und Menge von Informationsangeboten, oder Arbeitsgegenstände und Arbeitsmittel. Die durch Belastungen ausgelösten Reaktionen werden als Beanspruchung verstanden (Sonntag, 2007). Auch bei diesem klassischen Belastungs-Beanspruchungsmodell der Ergonomie wird ein psychophysiologischer Zusammenhang angenommen. Es wird davon ausgegangen, dass sich durch Belastungen (Reize unterschiedlicher Art) ausgelöste Reaktionen in Form von beobachtbaren Verhaltensänderungen und Veränderungen körperlicher Zustände (z. B. Erhöhung der Herzfrequenz, Atemfrequenz, verstärkte Schweißsekretion etc.) zeigen können (Sonntag, 2007).

Auch hier wird der psycho-physische Zusammenhang noch eher implizit denn systematisch explizit verdeutlicht. Im Weiteren wird die, nicht nur aber besonders für das Anliegen dieser Arbeit, grundlegende Bedeutung herauszuarbeiten sein. Im Rahmen dieser Arbeit kann und soll das allerdings nicht im Sinne eines philosophischen Diskurses aufgearbeitet werden (siehe dazu z. B. Bunge, 1980), sondern wird vielmehr im Sinne eines pragmatischen Bezugs angegangen, wie er z. B. in sport- und arbeitswissenschaftlichen Zusammenhängen im Hinblick auf die Belastungs-Beanspruchungsdynamik relevant ist.

Im Sport versteht man nach Kunath (1975, S. 123) unter psychischer Belastbarkeit die Fähigkeit eines Sportlers, „mit hohen Anforderungen bzw. starken Reizen des Trainings und Wettkampfes fertig zu werden, ohne daß das psycho-physische Gleichgewicht erst gestört wird, d. h. leistungsstabil zu bleiben.“ Die psychische Belastbarkeit ist insbesondere im beruflichen Kontext maßgebend für die Leistung der Mitarbeiter. Hohe bzw. zu starke Anforderungen oder Situationen führen zu Beanspruchungen, die über die Kapazitäten des Individuums hinausgehen und zu psychischen Belastungen führen, die sich auf Gesundheit und Wohlbefinden auswirken (vgl. Udris & Frese, 1999). Da psychische Belastungen im Sinne einer Überbeanspruchung körperliche Veränderungen induzieren, stellt sich die Frage, inwieweit die physische Verfassung eines Individuums dazu beiträgt, hohen Anforderungen und Beanspruchungen im Berufsleben standzuhalten, Resistent gegen bestimmte psychische Belastungsreize zu sein bzw. diese als nicht oder weniger belastend zu erleben und so leistungsfähig zu bleiben.

Lässt sich im Anschluss an diese Überlegungen evtl. davon ausgehen, dass die physische Leistungsfähigkeit eine Rolle für Steuerungs- und Regulationsvorgänge für die Verarbeitung und Beantwortung von psychischen Belastungsreizen spielt? Bereits Kunath (1975, S. 124) weist darauf hin, dass

„die psychische Belastbarkeit (. . .) vom Ausbildungsgrad und der aktuellen Funktionstüchtigkeit des psycho-physischen Systems oder seiner Teilsysteme, ihrer gegenseitigen Beeinflussbarkeit und Kompensationsfähigkeit, ihren strukturell-dynamischen Verbindungen sowie dem Niveau der Einzeleigenschaften [abhängig ist].“

Jedoch merkt Kunath (1975, S. 124) kritisch an, dass die spezifische psychische Belastbarkeit eines Sportlers auf Grundlage der antrainierten physischen Konstitution bzw. der erworbenen psycho-physischen Struktur nicht ohne weiteres „auf die Bewältigung hoher Anforderungen in anderen oder gar allen Lebensbereichen übertragbar ist.“ Allerdings erscheint die Annahme zumindest plausibel, dass je ähnlicher die Anforderungen oder Aufgabenstrukturen der verschiedenen Tätigkeiten sind, desto besser bzw. eher davon ausgegangen werden kann, dass sich die im Sport oder durch den Sport erworbene psychische Belastbarkeit transferieren lässt (Kunath, 1975).

An dieser Stelle scheint eine weitergehende Differenzierung der Begriffe Belastung, Belastbarkeit und Beanspruchung unter Berücksichtigung von Leistung bzw. Leistungsfähigkeit angebracht, um die damit verbundenen psychophysiologischen Zusammenhänge präziser zu beschreiben. Ein entsprechender Vorschlag wird von Ulmer (2013) angeboten. Die Belastung wird in dem Kontext in Verbindung mit der an das Individuum gestellten Aufgabe gesehen, die Leistung in physischer und psychischer (untergliedert in emotional, mental und kreativ) Hinsicht ist die Erfüllung der vorgegebenen Aufgabe (Belastung). Die Beanspruchung ist die individuelle Reaktion (physisch und psychisch) beim Erbringen der Leistung und die Leistungsfähigkeit kann als die Fähigkeit zur Erfüllung der Aufgabe betrachtet werden. Die Belastbarkeit ist nach Ulmer (2013) als reduzierte Leistungsfähigkeit beschrieben. Damit ist auf die individuellen, situativen Gegebenheiten verwiesen, die den Abruf bzw. Einsatz der potentiellen Fähigkeiten beeinflussen. Dies bedeutet also, dass eine Person mit potentiell hoher Leistungsfähigkeit durch situative Bedingungen, z. B. die aktuelle Ermüdung oder sonstige Vorbeanspruchungen nicht weiter belastbar ist und ihre spezifische Leistungsfähigkeit in einer bestimmten Situation nicht auszuschöpfen vermag. In verschiedenen Tätigkeitsfeldern der Arbeit besteht oftmals Unkenntnis über die individuelle Leistungsfähigkeit von Mitarbeitern und somit eine gewisse Gefahr, Mitarbeiter einer inadäquaten Belastung auszusetzen. Um

gesundheitlich nachteilige Konsequenzen zu vermeiden sollte darauf geachtet werden, dass die Arbeitsbelastungen weder zu stark über- noch unterfordern. Ferner ist ohne den Willen sich anzustrengen, auch keine Leistungserbringung möglich. Dieser Wille ist von psychophysiologischen und auch von sozialen Einflüssen abhängig (Ulmer, 2013).

„Leisten oberhalb der individuellen Dauerleistungsgrenze geht mit physischer Ermüdung oder gar Erschöpfung einher, die damit verbundenen Empfindungen bzw. Mißempfindungen bleiben nicht ohne Auswirkungen auf die Bereitschaft zum Sich-Anstrengen bzw. Sich-weiter-Anstrengen. Ansonsten gilt wie bei der physiologischen Beanspruchung: Je geringer die Leistungsfähigkeit, desto mehr muß sich ein Mensch bei gleicher Leistung anstrengen.“ (Ulmer, 2013, S. 83)

Das „Sich-Anstrengen“, so wie es Ulmer (2013) nennt, ist eher als Leistungsbereitschaft zu verstehen, wobei das Motivationsproblem diesbezüglich einen wichtigen Aspekt darstellt. In diesem Kontext treten auch die Begriffe Tüchtigkeit und Tauglichkeit häufig auf (siehe Münzberger, 2008). Das Anliegen dieses Projektes ist allerdings vorrangig im Anwendungsbezug zu sehen und nicht der Klärung aller möglichen nomologischen Verbindungen, die in Bezug auf die Belastungs-Beanspruchungs-Konzeption auftreten.

Mit Blick auf die Belastungs-Beanspruchungsdynamik und die Entwicklung bzw. Ausbildung einer diesbezüglichen Belastbarkeit ist aus handlungstheoretischer Sicht davon auszugehen, dass sich die Belastbarkeit auf solche Anforderungen bezieht, die bei der Ausübung der jeweiligen Tätigkeit bzw. Handlung im Sinne von Regulationserfordernissen der psychischen Funktion gefordert sind (vgl. dazu Hacker, 1978, speziell seine Ausführungen zur psychischen Struktur und Regulation von Arbeitstätigkeiten, Kapitel 3 und zur Anforderungsbewältigung und Veränderungen der psychischen Struktur und Regulation von Arbeitstätigkeiten, Kapitel 8). Dies bedeutet wenn z. B. spezifische kognitive Anforderungen vorliegen, werden entsprechende psychische Funktionen gefordert, trainiert und ausgebildet.¹

In dem hier vorgestellten Belastungs-Beanspruchungskonzept wird von psycho-physischen Zusammenhängen ausgegangen, die sich insofern auf die Entwicklung des psycho-physischen Belastbarkeitskonzepts abbilden lassen, als ein funktionaler Zusammenhang zwischen Aus-

¹ Von solchen spezifischen (nicht allgemeinen) Zusammenhängen ist sowohl im Hinblick auf psychische und physische Funktionen auszugehen. Ein einfaches Beispiel soll dies verdeutlichen: Wenn jemand die Ausdauer durch gezielte sportliche Tätigkeiten wie bspw. regelmäßiges Joggen trainiert und so seine entsprechende Belastbarkeit steigert, wird er bei langen Märschen Vorteile haben, was hingegen die Kraftleistung anbelangt z. B. Heben oder Transportieren einer schweren Last, aber nicht.

dauertraining, der sich dadurch ergebenden Verbesserung des Herz-Kreislauf-Systems und damit verbundener Auswirkungen auf die psychische Belastbarkeit angenommen wird. Hierzu sei auch auf die enge Verbindung zwischen emotionalen Zuständen, Stress und physiologischen Prozessen hingewiesen, auf die noch näher eingegangen wird. Darüberhinaus spricht die Befundlage aus empirischen Studien dafür, dass bestimmte sportliche Betätigungen, speziell aerobes Ausdauertraining, mit einer Verbesserung der maximalen Sauerstoffaufnahme einhergeht und sich bei Interventionsmaßnahmen gegen Depressions- oder Burn-Out-Erkrankungen als effektiv erwies. Dies verdeutlichen Forschungsarbeiten (siehe dazu auch Kapitel 3.2 Sport und psychische Gesundheit) von Martinsen, Medhus und Sandvik (1985), Martinsen (1987), Mutrie (1988, 2000), Sexton, Mære und Dahl (1989), McCann und Holmes (1984), Rueter, Mutrie und Harris (1982), Sime (1987), Doyne, Chambless und Beutler (1983). Martinsen (2002, S. 209) fasst zusammen, dass die Befunde der Studien zu sportlicher Betätigung, insbesondere aerobes Ausdauertraining, alle in die gleiche Richtung deuten: „Aerobic exercise is more effective than no treatment, and not significantly different from other forms of treatment, including various forms of psychotherapy.“

Martinsen (2002) merkt hingegen kritisch an, dass eine höhere aerobe Fitness kein entscheidendes Kriterium ist und Patienten ohne physiologische Vorteile ähnliche psychologische Effekte aufweisen, wie Personen mit einer ausgeprägten Fitness. Diese Schlussfolgerung steht jedoch ambivalent zu den Befunden einer Studie von Martinsen (1987), bei der sich Ergebnisse zeigen, die auf einen Zusammenhang zwischen einer verbesserten Fitness, speziell einer Erhöhung der maximalen Sauerstoffaufnahme, und einer Reduktion der Scores in verschiedenen Depressionsinventaren hinweisen (mit einem signifikanten $p = .05$ Korrelationskoeffizienten speziell bei Männern in Höhe von $r = .40$). Daneben verweist Fuchs (2003) auf eine Untersuchung von Mutrie (2000), die zusammenfassend feststellt, dass die bisherigen Meta-Analysen einen Zusammenhang sportlicher Aktivität und klinischer Depression von Effektstärken zwischen .53 und .72 dokumentieren. Darüber hinaus konkludiert Mutrie (2000),

„dass aerobe und anaerobe Sportprogramme zur Verringerung der klinischen Depression beitragen können und dass dieser antidepressive Effekt von der Größenordnung her vergleichbar ist mit den Wirkungen, die sich mit psychotherapeutischen Behandlungsverfahren erzielen lassen. (. . .) Insgesamt aber zeigen Mutries Ergebnisse, dass Sport und Bewegung eine wichtige Rolle bei der Prävention und Behandlung depressiver Störungen spielen können – zumindest als flankierende Maßnahme.“ (Fuchs, 2003, S. 90)

Die hierbei angesprochenen psycho-physischen Zusammenhänge werden im Kontext der Belastungs-Beanspruchungsdynamik mit Hilfe des folgenden Variablenansatzes erläutert.

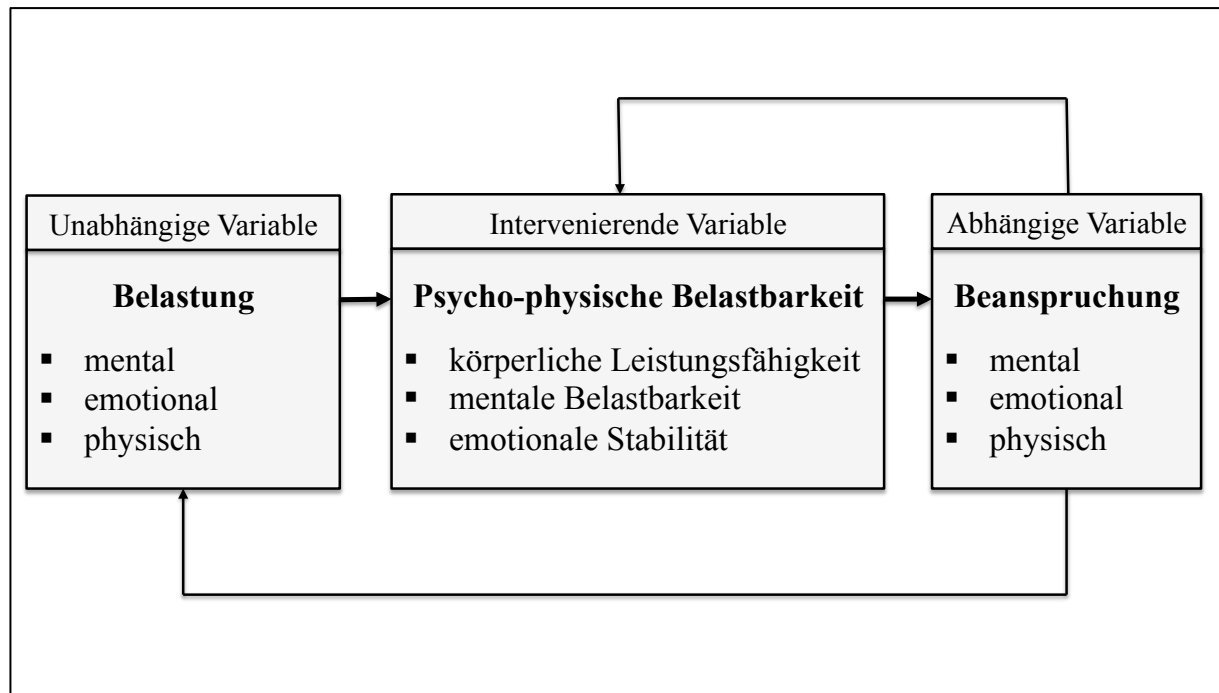


Abbildung 4. Beziehung zwischen Belastung, psycho-physischer Belastbarkeit und Beanspruchung (vgl. Klimmer, 1978).

Abbildung 4 ist im Anschluss und in Erweiterung der Darstellung von Klimmer (1978) konzipiert, um die Beziehung zwischen Belastung, psycho-physischer Belastbarkeit und Beanspruchung mit den hier wesentlichen Komponenten zu verdeutlichen. Die Belastung (unabhängige Variable) stellt den Einflussfaktor dar, der in Abhängigkeit von Qualität und Quantität bei einer Person auf die psycho-physische Belastbarkeit (intervenierende Variable) trifft und je nach individuellen Voraussetzungen (personenspezifischen Merkmalen und Ausprägung der psycho-physischen Belastbarkeit) eine bestimmte Beanspruchung (abhängige Variable) in Form und Intensität von Reaktionen auslöst. Belastungen sind zunächst externe Faktoren, die dann in Abhängigkeit der psycho-physischen Belastbarkeit zu internen Beanspruchungen führen, die wiederum zu Verhaltenskonsequenzen wie externen Reaktionen führen können. Die Beanspruchungsreaktion selbst kann dann wiederum als interne Belastung wirken, wenn bspw. die Wahrnehmung und Bewältigung der Beanspruchungsreaktion selbst zum Problem wird. Im Soldatenberuf würden z. B. bei Kampfhandlungen sowohl psychische als auch physische Einflussfaktoren auftreten, die es zu bewältigen gilt. Insbesondere bei Auslandseinsätzen kommt es zu stark belastenden sich wiederholenden Ereignissen in hohem Ausmaß. Bei

lang andauernden Gefechten unter sehr heißen und trockenen Bedingungen mit unzureichender Flüssigkeitszufuhr hoch konzentriert zu bleiben, erfordert in physischer und psychischer Hinsicht eine hohe Belastbarkeit (körperliche Ausdauerleistungsfähigkeit, mentale Belastbarkeit und emotionale Stabilität). Die jeweils gegebene psycho-physischen Belastbarkeit hat Einfluss auf die Entstehung einer Beanspruchung bei dem sich in einer Kampfeshandlung befindenden Soldaten, d. h. auf die Art und Stärke der internen Reaktionen, die durch die Belastung ausgelöst werden. Die durch diese Reaktionen geprägte Beanspruchung wirkt rückkoppelnd auf die Belastbarkeit, entweder im Sinne einer Einschränkung oder bei Wiederholung und längerfristig durchaus auch im Sinne eines Trainings bzw. einer Verbesserung/Steigerung. Die psycho-physische Belastbarkeit, die sich durch wiederholte Exposition bzw. Konfrontation der Person mit belastenden Situationen verändert, geht auch einher mit einer veränderten Einschätzung der Belastung und führt schließlich zu einer veränderten Beanspruchung und einer anderen Wahrnehmung sowie einem veränderten Erleben der Beanspruchung.

Von diesem Annahmenverbund ausgehend kann zur Beschreibung der Beanspruchung die Konzeption des Verhältnisses von Anforderungen und personalen Leistungsvoraussetzungen (Person-Environment-Fit-Ansatz) als relevant herangezogen werden, sodass eine erweiterte Erklärungsmöglichkeit für interindividuellen Unterschiede bei Reaktionen von Personen auf ähnliche oder identische physische und psycho-soziale Anforderungen eröffnet wird. Muster solcher subjektiver Reaktionen auf Beanspruchungen können als Ermüdungserscheinungen, Monotonie, psychische Sättigung, Stress, Burnout sowie in Form von psychosomatischen Erkrankungen auftreten (vgl. Richter & Hacker, 1998).

Bei einem Definitionsansatz von Weber (1985, S. 21) wird Beanspruchung als ein Kontinuum mit fließenden Grenzen zwischen psychischer und physischer Beanspruchung betrachtet, die als eine von externen Belastungsparametern abhängige Variable gefasst werden kann. Nach Weber (1985) kann demnach eine Trennung bzw. ein isoliertes Auftreten rein psychischer und rein körperlicher Beanspruchung in realen Arbeitssituationen keine diskutabile Möglichkeit darstellen. Die hierbei auftretende Wechselbeziehung zwischen psychischen und physischen Beanspruchungen wie sie Weber (1985) beschreibt, ist dabei als sinnvoll zu erachten, da Beanspruchungen in beiderlei Hinsicht als nicht unabhängig gesehen werden können. Gleichwohl hat auch eine Unterscheidung und differenzierte Betrachtung der Beanspruchung Berechtigung und macht Sinn, wenn es gilt die Art der Beanspruchung zu spezifizieren, um z.

B. mit Interventionen gezielt und ggf. schwerpunktmäßig im physischen oder psychischen Bereich anzusetzen.

„Ausgehend von aktivierungstheoretischen Modellvorstellungen eines Kräftespeichers, der durch physische und psychische Aktivitäten zunehmend erschöpft wird und sich durch Ruhepausen regeneriert, ist der Grad der Beanspruchung zu einem bestimmten Zeitpunkt mit dem zur Bewältigung der aktuellen Reizsituation notwendigen Aufwand gleichzusetzen. Die Höhe dieses Aufwandes ist mit situations- und personenspezifischer Differenzierung erlebbar und ergibt sich unter ergonomischen Gesichtspunkten aus

- den von der Qualität und Quantität der Belastungssituation abhängigen Graden der gesamten Aktivierung aller betroffenen menschlichen Leistungsfunktionen aus den Bereichen der
 - Informationsaufnahme,
 - Informationsverarbeitung,
 - emotionaler Qualitäten,
 - motorischen Arbeitshandlungen
- in Relation zu dem aktuellen Erschöpfungsgrad des Kräftespeichers, der das Ausmaß der Ermüdung anzeigt.“ (Weber, 1985, S. 22)

In unmittelbarer Nähe zu dieser Vorstellung liegt auch eine sportmedizinische bzw. biowissenschaftliche Konzeptualisierung, die Belastung als äußere Einflussgröße bzw. Einwirkung auf ein System (Organismus, Mensch) und Beanspruchung als Reaktion des Systems dieser externen Wirkungsmechanismen (siehe z. B. Ulmer, 2001) beschreibt.

Die aktivierungstheoretisch begründete und biowissenschaftlich akzentuierte Sichtweise wird erweitert durch den und im sozialwissenschaftlichen Kontext, in dem Anforderungen, die durch Umweltbedingungen, Aufgaben oder einer Kombination von Umweltbedingungen und Aufgaben gegeben sind, auch in psychischen Wirkzusammenhängen gesehen werden. Danach führen Anforderungen über individuelle Einschätzungen oder subjektive Bewertungen (siehe hierzu z. B. das transaktionale Stresskonzept von Lazarus & Launier, 1981; Lazarus & Folkman, 1984) und in Abhängigkeit von der individuellen (psycho-physischen) Belastbarkeit zu Beanspruchungen. Bereits in der Genese von Anforderungen sind danach psychische Prozesse (speziell Einschätzungen) wesentlich. Das Beanspruchungserleben kann diesem Verständnis folgend im Sinne einer Rückkopplung selbst wieder zu einer Belastung werden. Dies ist von der Art (Qualität) und dem Aufwand (Quantität) der Bewältigung bzw. des Bewältigungshandelns (siehe Schönplflug, 1983) abhängig. Entsprechend der oben darge-

stellten Grundkonzeption lässt sich dieser Zusammenhang wie in der folgenden Abbildung 5 veranschaulicht verdeutlichen.

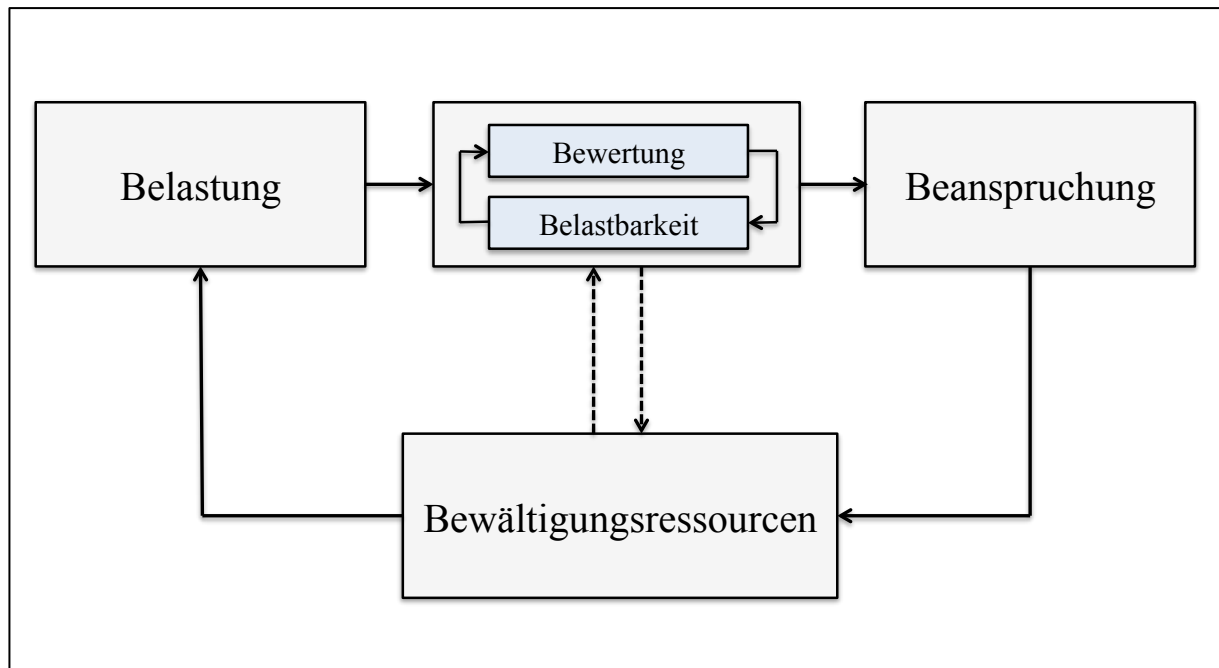


Abbildung 5. Wechselwirkungszusammenhang zwischen Belastung, Beanspruchung und Bewältigungsressourcen.

Während in der Stress- und Bewältigungs-/Coping-Forschung der Umgang (Bewältigungshandeln) mit der spezifischen Beanspruchung „Stress“ untersucht wird, interessiert hier speziell die personale Voraussetzung, die psychisch-physische Belastbarkeit, die einerseits die Wahrnehmung und Bewertung von Belastungen und andererseits das Bewältigungshandeln beeinflusst. Fasst man z. B. Training im Sinne der sportlichen Trainings als eine präventive Betätigung zum Aufbau von personalen Bewältigungsressourcen auf, so ist von einem Wechselwirkungszusammenhang (Interaktion) zwischen Bewältigung und Belastbarkeit auszugehen. Hierzu soll erneut ein erläuterndes tätigkeitsbezogenes Beispiel folgen:

Bei der Tätigkeit als Unternehmensberater gehören hohe Belastungen im Sinne von Zeitdruck, insbesondere enge Zeitfenster für die Anfertigung von Präsentationen, Aufarbeitung und Analysen von Daten, sowie Belastungen in Form von häufigen Reisen zu Terminen, Leistungserbringung unter erschwerten Bedingungen (Zeitzonewechsel, Jetlag, unzureichend Schlaf und ungewohnten Arbeitsbedingungen), sich oftmals wechselnde Arbeitsbedingungen und Arbeitsteams und teilweise eigenwillig und ablehnende Klienten zu den bekannten Anforderungen. Die Beratertätigkeit erfordert demnach eine hohe Stressresistenz, Ausdauer und

eine hohe Frustrationstoleranz (vgl. Plate, 2006), sowie eine gewisse emotionale Stabilität und Belastbarkeit um mit Drucksituationen gelassen umzugehen und handlungsfähig zu bleiben. Diese branchentypischen Umweltbedingungen und Aufgaben, wie bspw. die zur Verfügung stehende Zeit für ein Projekt, Dauer der Reise zu einem Kunden, Komplexität der auszuführenden Dienstleistung und die dem Berater zur Verfügung stehenden Ressourcen im Rahmen eines Projekts, stellen für den Berater Anforderungen (Belastungen) dar, die indirekt über vorhergehende Bewertungen bzw. Einschätzungen dieser Bedingungen und Aufgaben zu einer bestimmten Beanspruchung führen. Das Ausmaß der Beanspruchung (Reaktionen des Organismus, Stress) hängt dabei von der individuellen Wahrnehmung und Bewertung der Anforderungen ab, die je nach dem Grad der Ausprägung der individuellen psycho-physischen Belastbarkeit in Qualität und Quantität unterschiedlich ausfällt. Eine individuell stark ausgeprägte psycho-physische Belastbarkeit kann also auch als personale Bewältigungsressource aufgefasst werden, die sich in Abhängigkeit von der Ausprägung unterschiedlich im Beanspruchungserleben niederschlägt.

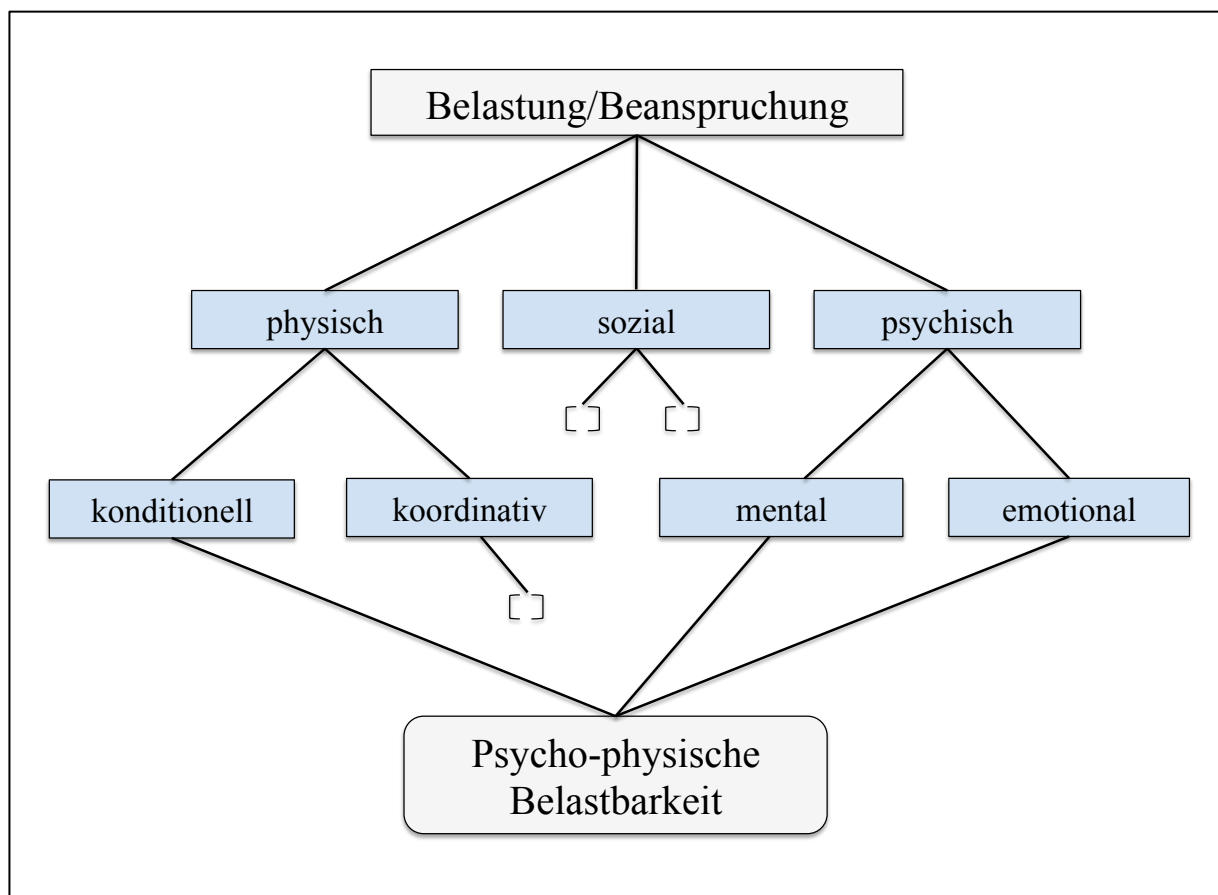


Abbildung 6. Strukturelle Ableitung der psycho-physischen Belastbarkeit aus der Belastungsbeanspruchungsdifferenzierung.

Zur weiteren Differenzierung und Strukturierung lassen sich sowohl für die Belastung als auch für die Beanspruchung „top-down“ betrachtet drei Bereiche voneinander abheben (vgl. Abbildung 6). Belastung und Beanspruchung können demnach grundsätzlich in physische, soziale und psychische Bereiche unterteilt werden. Der physische Bereich kann wiederum in konditionelle und koordinative Formen der Belastung bzw. der Beanspruchung differenziert werden. Daneben lassen sich psychische Belastungen und Beanspruchungen in mentaler und emotionaler Hinsicht aufteilen. Die psycho-physische Belastbarkeit setzt sich letztendlich aus emotionalen und mentalen Aspekten seitens der psychischen Belastbarkeit, sowie aus konditionellen Aspekten, hier speziell die Ausdauerleistungsfähigkeit auf Seiten der physischen Belastbarkeit, zusammen. Im hier verfolgten Anliegen zu vernachlässigen und auszuklammern sind auf der physischen Seite die koordinativen Aspekte, die im Sinne einer Belastbarkeit und speziell im relevanten Kontext dieser Arbeit, der Personalauswahl, eine untergeordnete Rolle spielen. Im Rahmen der konditionellen Belastbarkeit nimmt der Bereich der Ausdauerleistungsfähigkeit eine besondere Stellung hinsichtlich des psycho-physischen Zusammenhangs ein. Bei der Konzeptualisierung der psycho-physischen Belastbarkeit bleiben soziale Aspekte genauso ausgeklammert wie koordinative Aspekte bei der physischen Belastbarkeit. Dies dient in erster Linie einer Komplexitätsreduktion zu Gunsten der inhaltlichen und methodischen Bewältigung der Problem- und Zielstellung dieser Arbeit.

Das Ausmaß der Beanspruchung bzw. das Beanspruchungserlebnis ist im Wesentlichen von psychischen Merkmalen insbesondere spezifischer Persönlichkeitsmerkmale sowie von physischen Faktoren, speziell der körperlichen Leistungsfähigkeit (physischer Fitness) abhängig. Wie im nachfolgenden Kapitel erläutert, lassen sich die hier relevanten personenspezifischen Merkmalskonfigurationen zu bestimmten Konstrukten integrieren, die sich sowohl aus physischen als auch psychischen Komponenten, wie z. B. Ausdauerleistungsfähigkeit oder emotionale Stabilität und Belastbarkeit, zusammensetzen.

2.3 Psycho-physisches Belastbarkeitskonzept

Ein Kardinalanliegen bzw. die zentrale Herausforderung dieser Arbeit besteht in der Entwicklung einer passenden Konzeptualisierung in Verbindung mit der Operationalisierung der psycho-physischen Belastbarkeit. Dazu erscheint es notwendig, das Konstrukt bzw. die Teilstrukturen (Komponenten) in einen systematischen Zusammenhang zu bringen, der es er-

laubt, die einzelnen Komponenten in ihrer Abhängigkeit zu verstehen und die Verbindung nachvollziehen zu können. Die nachstehende Abbildung 7 veranschaulicht diese Konzeptualisierung. Den Ausgangspunkt bilden die beiden Komponenten physische und psychische Belastbarkeit. Die Konstrukte repräsentieren die spezifischen Komponenten, für die jeweils (dies wird im methodischen Teil konkretisiert) Indikatoren bestimmt werden: Zum einen auf der physischen Seite die objektive Ausdauer, zum anderen - auf der psychischen Seite - die kombinierten Persönlichkeitsmerkmale Emotionale Stabilität und Belastbarkeit.

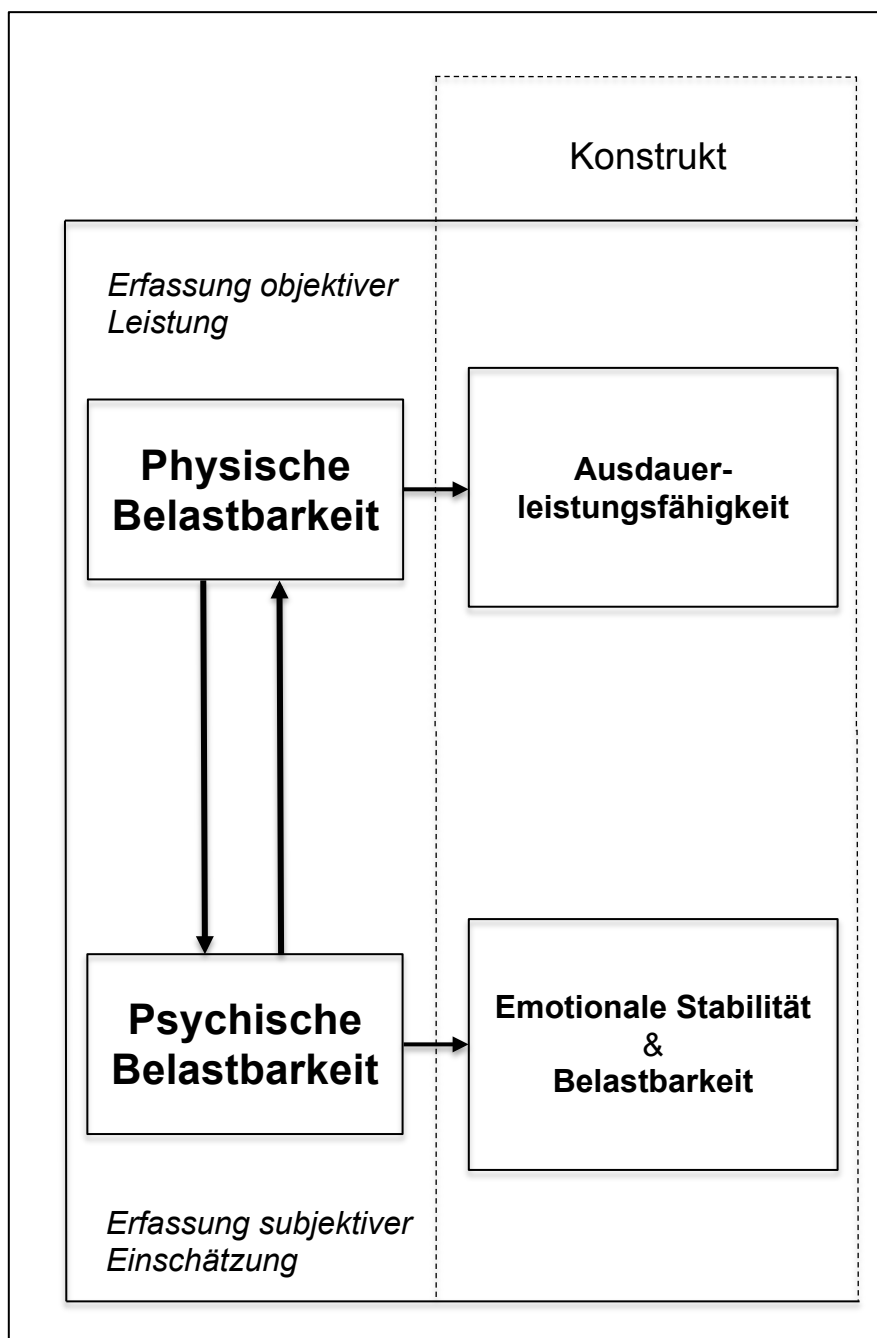


Abbildung 7. Übersicht über die Konstrukte des psycho-physischen Belastbarkeitskonzepts.

Für die Konzeptualisierung ließen sich zum einen andere Akzentuierungen denken, zum anderen wären Erweiterungen möglich, die sich sowohl auf physische Komponenten als auch auf psychische Komponenten beziehen. Physische Komponenten könnten andere Ausdauerformen oder -arten (bspw. Langzeitausdauer) sein, die auch mit anderen Belastungsprotokollen bzw. -schemata erfasst werden könnten oder es könnten andere Facetten der physischen Leistungsfähigkeit berücksichtigt werden, wie z. B. Kraft-, Beweglichkeits-, Koordinations- und Schnelligkeitstests. Genauso kommen für die psychische Komponente andere Facetten und entsprechend andere Verfahren oder Inventare in Frage, die ebenfalls Aufschluss über Persönlichkeitsdimensionen hinsichtlich der psychischen Belastbarkeit geben können. Dies wären z. B. Merkmale wie „Resilienz“, „Hardiness“ und/oder „Mental Toughness“. Wobei jede Facette, das gilt für den physischen wie für den psychischen Bereich, sich durch bestimmte Spezifika, Vor- und Nachteile insbesondere in der Operationalisierung unterscheiden. Zusammenfassend lässt sich die Konzeptualisierung psycho-physischer Belastbarkeit anhand der folgenden Abbildung 8 verdeutlichen:

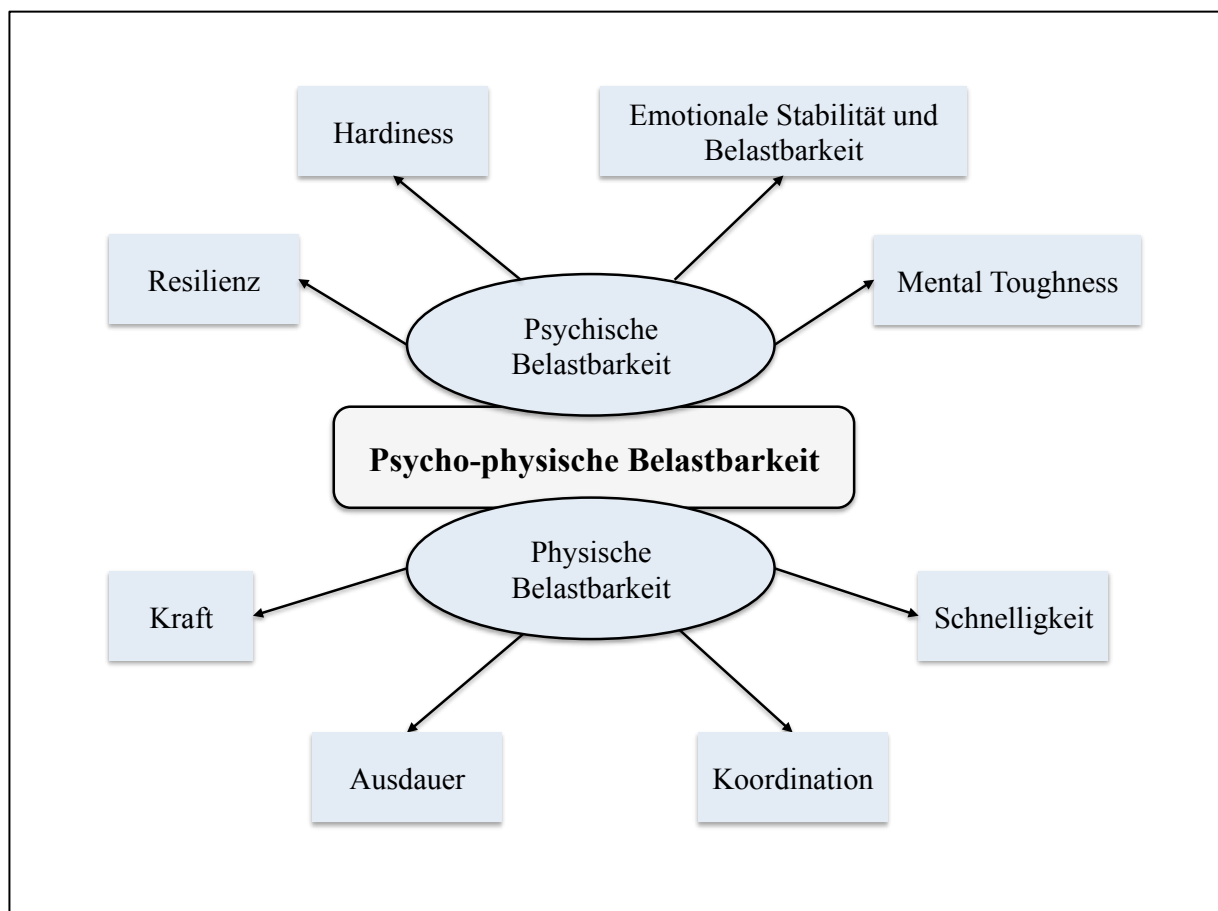


Abbildung 8. Erweiterte Konzeptualisierung psycho-physischer Belastbarkeit.

Die hier und weiter oben angeführten Konstrukte „Resilienz“ (resilience, resiliency; siehe Maddi & Khoshaba, 2005), „Hardiness“ (Maddi, 2006) und „Mental Toughness“ (Jones, Hanton & Connaughton, 2002), die sich als Facetten der psychischen Belastbarkeit verstehen lassen, sollen hier nicht eingehender erläutert werden. Sie werden momentan insbesondere in der sportpsychologischen Literatur diskutiert (siehe u. a. Crust & Clough, 2011) und gewinnen aktuell auch in der Auseinandersetzung mit dem Thema der (psychischen) Belastbarkeit bzw. Resilienz im Arbeitsleben, speziell auch hinsichtlich der Fitness und psychischen Stärke von Managern, zunehmend an Bedeutung (siehe z. B. Leitl, 2014). Im folgenden wird deutlich, warum speziell die Ausdauer in dieser Arbeit als physische Komponente für die Ausarbeitung des psycho-physischen Belastbarkeitskonzepts gewählt wurde.

2.4 Ausdauer als Form physischer Beanspruchung

Die Ausdauer gehört bekanntermaßen zu den Hauptfaktoren sportlicher Leistung und kann als „Widerstandsfähigkeit gegenüber Ermüdung, die bei sportlichen Belastungen ermüdungsbedingte Leistungsverluste mindert“, definiert werden (Schnabel et al. 2008, S. 180). Die Ausdauerleistungsfähigkeit gewährleistet bei sportlichen Betätigungen mit länger andauernden Belastungen möglichst lange handlungsfähig zu bleiben und bei längeren Belastungsdauer und -anforderungen Beeinträchtigungen der beanspruchten Regulationssysteme standzuhalten. Hierbei können psychische Merkmale, wie bspw. das Durchhaltevermögen, zu leistungslimitierenden Faktoren werden, die letztendlich notwendig sind, die körperlichen Ressourcen bei fortschreitender Belastung auszuschöpfen und abzurufen. Hierzu konstatieren Schnabel et al. (2008, S. 180), dass die „Ausdauer (. . .) die Leistungsfähigkeit und Belastbarkeit bei der Bewältigung jeglicher Anforderungen des täglichen Lebens [stabilisiert]. Die mit der Ausbildung der Ausdauer entstehenden Anpassungen im Herz-Kreislauf-, Atmungs- und Stoffwechselsystem erhöhen die Gesundheit und verbessern die Lebensqualität.“ Demzufolge wird von einer positiven Auswirkung der Ausdauerleistungsfähigkeit in anderen Lebensbereichen im Sinne der Belastbarkeit und im Hinblick auf das Bewältigen verschiedener Anforderungen generell ausgegangen (siehe hierzu auch Kapitel 4.2).

Die Ausdauer kann in verschiedenen Arten, je nach Dauer der Belastung, eingeteilt werden. Ein Anliegen dieser Arbeit ist es, einen Fahrradergometertest für die Überprüfung der Ausdauerleistungsfähigkeit zu entwickeln, der im Rahmen des BFT für die Personalgewinnung

der Bundeswehr eingesetzt werden soll. Das dafür zu konzipierende Belastungsschema ist vorwiegend an die Vorgaben der Bundeswehr gekoppelt, die u. a. eine maximale Testdauer für die Feststellung der Ausdauerleistung von 10 Minuten beinhalten. Aufgrund der Vorgaben in Bezug auf die Testkriterien kann die zu überprüfende Art der Ausdauer des neu zu konzipierenden Tests als Mittelzeitausdauer definiert werden, die im Allgemeinen als Belastung in Zeitbereichen zwischen zwei bis zehn Minuten gesehen wird, wobei die kürzeren Belastungen eher die anaerobe und die längeren Belastungen eher die aerobe Energiebereitstellung betrifft (Schnabel et al., 2008; Graf, Hartmann, Platen, Rost & Schänzer, 2001). Hierzu sei angemerkt, dass die Differenzierung der allgemeinen aeroben dynamischen Ausdauer u. a. von Hollmann und Strüder anders vorgenommen wird. Die Autoren definieren eine Belastungsdauer von 3 bis 10 Min. als „allgemeine aerobe Kurzeitausdauer“ (2009, S. 433). Hollmann und Strüder (2009, S. 433) erklären hierzu, dass „mit der allgemeinen aeroben Ausdauer vornehmlich die Kurzeitausdauer gemeint [ist]. Pädagogen bezeichnen sie auch als *>>Organ-kraft<<*. Als entscheidend leistungsbegrenzendes Kriterium steht im Mittelpunkt die kardiopulmonale Kapazität. Ihr Bruttokriterium stellt die maximale Sauerstoffaufnahme/min dar.“ Für die genauen Vorgänge und Prozesse des Energiestoffwechsels bzw. der Energiebereitstellung sei auf die ausführlichen Darstellungen von bspw. Heck (1990), Graf et al. (2001) oder de Marées (2002) verwiesen.

Die Ausdauerleistungsfähigkeit kann mit verschiedenen Verfahren erfasst werden. Die Testverfahren können anhand von verschiedenen Kriterien und mit Hilfe von unterschiedlichen Gerätschaften oder Aufgabenstellungen die Ausdauer einer Person messen. Unter Berücksichtigung der vorhandenen Ressourcen, wie die Möglichkeit, Ergometer, Laufbänder o. ä. einzusetzen und welche Rahmenbedingungen vorzufinden sind (Labor, Laufbahn, Sporthalle, etc.), ist des Weiteren festzulegen, welche Art der Ausdauer durch die jeweilige Testmethode speziell gemessen werden soll. Bestimmte Tests erfassen ganz bestimmte Arten der Ausdauer, d. h. wenn das primäre Anliegen die Feststellung der Ausdauerleistungsfähigkeit in einer bestimmten Sportart ist, liegt es nahe, dass Testdesign so zu gestalten, dass die sportartspezifische Ausdauer gemessen wird. Dementsprechend bedarf es dann einer spezifischen Belastungsform bzw. eines für diese Sportart speziellen Testverfahrens. Die allgemeine Ausdauerleistungsfähigkeit, also die sportartunspezifische Ausdauer, kann mit dem Laufband oder dem Fahrradergometer im Labor getestet werden. Nach Heck (1990) hat der Labortest gegenüber dem Feldtest den Vorteil, dass spirographische Parameter wie die maximale Sauerstoffaufnahme, der maximale Sauerstoffpuls, der maximale Respiratorischer Quotient, oder das ma-

ximale Atemäquivalent besser erfasst werden können. Außerdem gewährleistet der Labortest eine „Reproduzierbarkeit der Belastungsbedingungen bzgl. des Belastungsgeräts (Fahrradergometer, Laufbandergometer) und gleichbleibende Umgebungsbedingungen“ (Heck, 1990, S. 102). Folgende Beurteilungskriterien sind nach Heck (1990) wesentliche Indikatoren für die Leistungsbeurteilung bei dem Fahrradergometer- und Laufbandtest:

- „maximal erreichte Belastungsstufe,
- maximale Sauerstoffaufnahme (VO_2),
- maximaler Sauerstoffpuls (O_2 -Puls),
- PWC170
- Schwellen“

Nach Heck (1990, S. 118) „setzen die ersten beiden Kriterien die Ausbelastung der Testperson voraus, während die letzten beiden im submaximalen Belastungsbereich bestimmt werden.“ Da eine subjektive Aussage über den Belastungszustand während oder nach dem Test oft nicht ausreicht, um die tatsächliche Ausbelastung festzustellen, sind folgende objektive Kriterien der Ausbelastung von besonderer Relevanz (Heck, 1990, S. 118):

- „maximale Herzfrequenz
- maximales Atemäquivalent ($\text{A}\ddot{\text{A}}$),
- maximaler Respiratorischer Quotient (RQ),
- maximaler Laktatwert,
- Levelling-off der Sauerstoffaufnahme“

Heck (1990, S. 118) fügt hinzu, dass „zur Beurteilung der Leistungsfähigkeit die maximal erreichte Belastung das am einfachsten zu messende leistungsdiagnostische Kriterium [ist], da man zur Bestimmung nur das Fahrrad- bzw. Laufbandergometer benötigt. Gegenüber der maximalen Sauerstoffaufnahme hat die erzielte maximale Belastungsstufe jedoch den Nachteil, dass sie stark von der Belastungsanstiegsgeschwindigkeit und der Motivation abhängt.“

Die Autoren Graf et al. (2001) weisen darauf hin, dass der Fahrradergometertest einem bestimmten Testmanual und Testschema bzw. Belastungsprotokoll folgen sollte, um die Ergebnisse der Testpersonen interpretieren und vergleichen zu können. Das Testmanual ist eine zumeist standardisierte Instruktion oder Anweisung zur Durchführung und umfasst Aspekte

der Vorbereitung für die Ergometrie. Die Testperson sollte nach Graf et al. (2001) zur Vermeidung hypoglykämischer Zustände (Unterzuckerung) darauf hingewiesen werden, dass sie nicht nüchtern, sondern nach einer leichten Mahlzeit den Test durchführen sollte. Bei der Einstellung der Sattelhöhe ist darauf zu achten, „dass bei gestrecktem Bein die Fußsohle bequem auf dem durchgetretenen Pedal ruht. Bei zu hoher oder zu niedriger Sattelleinstellung beeinflussen die ungünstigen Hebelverhältnisse das Ergebnis negativ“ (Graf et al., 2001, S. 52). Bei der Fahrradergometrie können verschiedene Belastungsstufen in Watt bestimmt werden. Die Belastungsintensität sollte bei der Eingangsbelastung so gewählt sein, dass die Testperson die Gelegenheit hat, sich mit dem Ergometer vertraut zu machen und sich bei 70 Watt ca. ein bis zwei Minuten einzufahren (siehe Hollman-Venrath Schema; siehe Graf et al., 2001, S. 53). Anschließend beginnt ohne Pause der Test bei einer erhöhten Wattzahl, die entweder konstant gehalten wird mit dem Ziel in einer bestimmten Zeit eine vorgegebene Strecke zurückzulegen, oder sich je nach Testmethode minütlich um eine festgesetzte Wattstufe bei einer empfohlenen Drehzahl von 60 bis 80 Umdrehungen pro Minute steigert. Das Testdesign des Ausdauer-tests mit dem Fahrradergometer wurde mit Hilfe einer Studie an der Universität der Bundeswehr in München im Rahmen dieses Dissertationsprojekts entwickelt und durch Untersuchungen validiert und auf Reliabilität geprüft (siehe Kapitel 5.1.1).

2.5 Psychische Belastbarkeit als ein Persönlichkeitskonstrukt

Dem Anliegen der Arbeit entsprechend, eine ökonomische und valide Einschätzung der körperlichen und psychischer Belastbarkeit zur Auswahl geeigneter Personen für Tätigkeitsbereiche vornehmen zu können, die mit hohen psycho-physischen Anforderungen verbunden sind, wurde folgende Überlegung ausgearbeitet:

In Stresssituationen ist die psycho-physische Belastbarkeit eine Voraussetzung, effektiv handlungsfähig zu sein und zu bleiben, sodass die Anforderungen effizient bewältigt werden können. Belastbarkeit wird dabei als positives Pendant zur Stressanfälligkeit gesehen (siehe Nitsch, 1982). Wesentliche Komponenten der psycho-physischen Belastbarkeit sind einerseits die aerobe Ausdauer und andererseits die emotionale Stabilität (siehe z. B. Hossiep & Paschen, 1998). Auch die selbsteingeschätzte Belastbarkeit mit starkem Bezug zur - ebenfalls per Selbsteinschätzung erhobenen - physischen Robustheit erscheint in diesem Zusammenhang relevant (siehe Hossiep & Paschen, 1998, S. 24). Die drei Faktoren „aerobe Ausdauer“,

die über verschiedene Ausdauer-Testverfahren (anhand von objektiven Leistungsparametern operationalisiert), sowie Emotionale Stabilität (per Selbsteinschätzung erfasst) und Belastbarkeit (per Selbsteinschätzung erfasst) können als eine fundamentale Trias der Belastbarkeit konzipiert werden. Bereits Hossiep und Paschen (1998) hatten bei der Entwicklung und Überprüfung des BIP darauf hingewiesen, dass die beiden Selbsteinschätzungen der Emotionalen Stabilität und Belastbarkeit einen gemeinsamen Stabilitätsfaktor bilden können. Dies wird hier durch die objektive Erfassung der Ausdauer ergänzt (dualer methodischer Ansatz: objektiver Test in Verbindung mit der Selbsteinschätzung), da gerade dieser Faktor, der auf die Funktionalität des Herz-Kreislauf-Systems verweist, mit Stress und Stresstoleranz, respektive Belastbarkeit, in Verbindung gebracht wird (siehe z. B. Klaperski, von Dawans, Heinrichs & Fuchs, 2013).

Der BIP stellt in erster Linie ein Instrument für die Berufseignungsdiagnostik dar. Die durch den BIP und die Überprüfung der Ausdauerleistungsfähigkeit durch verschiedene Ausdauer-Tests gewonnenen Parameter werden miteinander in Zusammenhang gebracht. Die Art und Ausprägung der so konzipierten Merkmalskombination kann der Statusdiagnostik zugeordnet werden. Somit wird in dieser Arbeit dem Ist-Zustand der untersuchten Personen im Hinblick auf die physische und psychische Komponente in Form der Ausdauerleistungsfähigkeit als Indikator für die körperliche Fitness und der eigenen Persönlichkeitseinschätzung des Faktors psychische Belastbarkeit besondere Bedeutung beigemessen. Dieses Vorgehen entspricht also der Statusdiagnostik und ist im Sinne einer Querschnittanalyse angelegt, während eine Prozessdiagnostik im Sinne einer Längsschnittanalyse anzulegen wäre. Für die Überprüfung der physischen Komponente bzw. der körperlichen Fitness und Leistungsfähigkeit kommen im Rahmen der Untersuchung verschiedene Ausdauer-Testverfahren zum Einsatz. Das hauptsächlich eingesetzte Testverfahren zur Überprüfung der physischen Leistungsfähigkeit ist ein 3000-Meter-Fahrradergometertest (3000-m-FET; Streckentest mit einem rektangulärem Belastungsmodus), der zu Personalgewinnungszwecken der Bundeswehr entwickelt und erprobt wird (siehe Kapitel 5.1) und deswegen für eine breit angelegte Datenerhebung auch hinsichtlich der Evaluierung der Persönlichkeitseigenschaften (Fragebogen BIP, etc.) als sehr geeignet erscheint, um beide Komponenten des psycho-physischen-Belastbarkeit-Konzepts (siehe Kapitel 5.2) zu erfassen.

Bereits Stern (1935), der zwischen Rüstungs- und Richtungsdispositionen unterscheidet, verknüpfte psychische und physische Merkmale im Sinne einer psychischen und physischen

„Vitalkapazität“: Bei den Rüstungsdispositionen bezog er sich auf mentale Personenmerkmale (Intelligenz, Temperament) und auf körperliche Merkmale wie Körpergröße und Gewicht (vgl. Kieschke, 2003).

Persönlichkeitstests können auch „gewinnbringend für die Personalauswahl und –platzierung eingesetzt werden“ (Mussel, 2012, S. 446). Speziell im Hinblick auf den hier interessierenden (und auch unter den sog. Big Five aufgeführten) Faktor „Neurotizismus“ konnten Zusammenhänge mit der Intelligenz (Größenordnung: $r = -.15$) nachgewiesen werden (siehe Reeve & Bonaccio, 2008). „Eine plausible Erklärung für diese Befunde ist, dass neurotische Personen ängstlicher sind und Prüfungsangst die Leistung in Intelligenztests verringert.“ (Mussel, 2012, S. 442) Allerdings erscheint es auch naheliegend anzunehmen, dass „möglicherweise (. . .) Prozesse des Arbeitsgedächtnisses bei neurotischen Personen durch aufgabenirrelevante Informationen beeinträchtigt werden“ (Mussel, 2012, S. 442), wie es besonders die Forschung zur Leistungs- und Prüfungsangst verdeutlicht (siehe z. B. Sarason, 1984).

Der Persönlichkeitsbegriff kann sowohl als populär als auch in der wissenschaftlichen Psychologie etabliert gekennzeichnet werden und umfasst „die gesamte psychische Eigenart eines Menschen“ (Meili & Steingrüber, 1978, S. 142). Eysenck hat bereits auf Grundlage des von ihm entwickelten „Maudsley Medical Questionnaire (MMQ)“ und des „Maudsley Personality Inventory (MPI)“ 1964 das „Eysenck Personality Inventory (EPI)“ entwickelt, das die Faktoren Introversion und Extraversion als Hauptmerkmale durch ein Fragebogenverfahren erfasst. Die faktorenanalytische Forschung hat, so Meili und Steingrüber (1978), zur Konstruktion weiterer Persönlichkeitsfragebogen beigetragen, mit denen gewisse Teilaspekte der Persönlichkeit eruiert werden können. Kieschke (2003, S. 36) fasst in seiner Dissertation die Befunde zusammen und konstatiert, dass Persönlichkeitseigenschaften in der Kindheits- und Jugendphase „einer zunehmenden Konsolidierung unterworfen“ sind „und erreichen bei jungen Erwachsenen schließlich ein über viele Jahre stabiles Niveau, bevor sie im höheren Alter – z. B. aufgrund physiologischer Abbauvorgänge – neuerlich an Beständigkeit einbüßen“ (vgl. dazu auch Roberts & DelVecchio, 2000). Kieschke (2003) weiter: „Vermutlich beruht dieser ‚Aushärtungseffekt‘ u. a. auf Nachhaltigkeit von Prozessen der Selbstbildkonturierung bis in die Adoleszenz hinein.“ (S. 36).

Faktoren wie Emotionalität, emotionale Stabilität/Labilität, Neurotizismus u. ä. sind in allen gängigen Instrumenten der Persönlichkeitsdiagnostik seit der Entwicklung des MMPI (Minnesota Multiphasic Personality Inventory von Hathaway & McKinley, 1943) enthalten, das als Ausgang und Grundlage der Persönlichkeitsdiagnostik, insbesondere per Fragebogen, angesehen werden kann. Im deutschen Sprachraum ist besonders das Freiburger Persönlichkeitsinventar (FPI von Fahrenberg & Selg, 1970) zu nennen, oder die deutsche Version des NEO-FFI (siehe Costa & McCrae, 1992) und dem Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP von Hossiep & Paschen, 1998).

Aufgrund des persönlichkeitsrelevanten Faktors „psychische Belastbarkeit“, der aus den beiden im BIP beinhalteten Dimensionen „Emotionale Stabilität“ und „Belastbarkeit“ konzeptualisiert wurde, kann von einem abgeleiteten Persönlichkeitskonstrukt gesprochen werden. Meili und Steingrüber (1978, S. 153) bemerken, dass „erklärende Eigenschaften den Status von Konstrukten haben“. Die Autoren weiter: „Dimensionen und Faktoren können als Konstrukte aufgefaßt werden, sofern sie in erklärendem Sinne gebraucht werden. Es fehlt ihnen aber noch weitgehend die Einordnung in ein theoretisches System.“ (Meili & Steingrüber, 1978, S. 154) Insofern ist spezifischer von einem explikativen im Gegensatz zu einem deskriptiven Konstrukt, d. h. von einem erklärenden Konstrukt auszugehen, welches allerdings noch zu operationalisieren ist. Deskriptive Konstrukte, wie z. B. Aggression, Ärger oder Angst im Sinne von emotionalen Zuständen, beschreiben einen Zustand bzw. ein Phänomen. Im folgenden wird das systematische Vorgehen im Bemühen um eine angemessene Operationalisierung des hier relevanten Persönlichkeitskonstrukts „Emotionale Stabilität“ anhand einer Studie von Westhoff, Steinborn und Schurz (2013) aufgezeigt. Die Autoren bemühen sich speziell um eine begriffssystematische Klärung dieses Konstrukts und der Differenzierung von Facetten.

Westhoff et al. (2013) entwickelten ein methodisch stringent aufgebautes Kategoriensystem für die beiden Persönlichkeitseigenschaften bzw. die beiden Konstrukte Gewissenhaftigkeit und Emotionale Stabilität. Die beiden Hauptdimensionen aus dem Fünf-Faktoren-Modell (FFM) von Costa und McCrae (1985) wurden durch einen Vergleich der Items von gängigen Fragebogenverfahren zur Erfassung dieser Persönlichkeitskonstrukte operationalisiert. Die Zusammenstellung der Facetten bei den verschiedenen Autoren wurden auf Gemeinsamkeiten und Unterschiede hin analysiert und dabei aus dem Datenpool diejenigen Facetten herausgestellt, die in der Mehrheit oder in allen Konzeptionen enthalten bzw. extrahierbar waren

(Westhoff et al., 2013). Die am häufigsten vorgefundenen Facetten wurden als der Kern des Konstrukts angesehen. Der Sinn solch einer Analyse und Strukturierung liegt „besonders in der praktischen diagnostischen Tätigkeit“, da sie „für die Planung, Durchführung und Auswertung diagnostischer Strategien und Prozeduren ausgesprochen nützlich [sind].“ (Westhoff, et al., 2013, S. 204) Das Kategoriensystem bietet im personaldiagnostischen Kontext eine Hilfestellung bei der Beratung und Intervention (Westhoff et al., 2013). Damit dem Diagnostiker in der Praxis eine bestmögliche Unterstützung zukommt, sollte er mit dem Inhalt des Kategoriensystems vertraut sein (Westhoff et al., 2013). In einer von Westhoff betreuten Diplomarbeit von Frau Liebert konnten zwei Kategoriensysteme für die Dimensionen Emotionale Belastbarkeit und Umgang mit emotionalen Belastungen entwickelt werden, die den derzeitigen Wissensstand über die Facettierung und Operationalisierung dieser Dimensionen wiedergeben (siehe Liebert, 2006).

Das von Liebert (2006) entwickelte Kategoriensystem für Emotionale Belastbarkeit basiert in erster Linie auf der Sichtung der Fachliteratur und dem Zusammentragen der Operationalisierungen der Facetten sowie einer Analyse, wie häufig eine Facette des Konstrukts in der gesichteten Literatur vorkam und in welchem Ausmaß der Facette Akzeptanz von Seiten der Forschergruppen beigemessen wurde. Zusätzlich wurden die Faktorenanalysen zu den Persönlichkeitsdimensionen als ein weiterer Ansatzpunkt zur Facettenbildung herangezogen und aus bewährten und standardisierten Testverfahren, wie bspw. dem NEO-Persönlichkeitsinventar (NEO-PI-R; Ostendorf & Angleitner, 2004), die bereits bestehenden Operationalisierungen in Form von Items gesichtet (vgl. Liebert, 2006). Hierzu konstatiert Westhoff (2013, S. 201), „ein Verständnis des Konstrukts erhofft man sich in der Literatur von Korrelationen zwischen verschiedenen Fragebogen, die dasselbe oder ein verwandtes Konstrukt erfassen“. Ein hier von Westhoff (2013) genannter Ansatz ist die Erstellung einer Multitrait-multimethod-Matrix (MTMM) nach Campbell und Fiske (1959). Westhoff (2013, S. 201) weiter:

„Die Korrelationen zwischen den verschiedenen Operationalisierungen müssten sich dabei in der Höhe der Ähnlichkeit der Konstrukte und der Methoden zu ihrer Erhebung entsprechend ordnen. Theoretisch und empirisch nicht verwandte Konstrukte sollten dabei also nicht oder niedrig mit den verschiedenen Operationalisierungen des fraglichen Konstrukts korrelieren, während gleiche oder verwandte Konstrukte jedoch miteinander korrelieren.“

Westhoff (2013) beschreibt in diesem Zusammenhang einen weiteren Zugang für einen besseren Überblick der mit unterschiedlichen Fragebogen erfassten Konstrukte, dieser besteht

„in der zumeist faktorenanalytischen Weiterverarbeitung von Korrelationen zwischen den unterschiedlichen Operationalisierungen desselben und verwandter beziehungsweise nicht verwandter Konstrukte (Kenny & Kashy, 1992). Aus der Faktorenstruktur kann man dann vorher aufgestellte Zusammenhangshypothesen zu den Konstrukten prüfen. Besonders wertvoll sind Korrelationen von Fragebogen mit Kriterien, die das Konstrukt oder bestimmte Aspekte davon anders operationalisieren, zum Beispiel Tests oder berufliche Leistungen oder Leistungsbeurteilungen durch Experten, denn diese sagen etwas über die Kriteriumsvalidität aus (Costa, McCrae & Dye, 1991).“ (Westhoff, 2013, S. 201)

Die Arbeit von Liebert (2006) zum Kategoriensystem zur emotionalen Belastbarkeit enthält auf Grundlage des beschriebenen methodischen Vorgehens insgesamt zehn Facetten mit den jeweiligen Operationalisierungen. Die zehn Facetten des Kategoriensystems emotionale Belastbarkeit, darunter geringe Ängstlichkeit, Gelassenheit, geringe Neigung zu Depression, soziales Selbstbewusstsein, leistungsbezogene Selbstsicherheit, Zufriedenheit, Robustheit, Impulskontrolle, emotionale Stabilität und körperliches Wohlbefinden, wurden einer empirischen Überprüfung auf die Wirksamkeit hin untersucht, indem einer Stichprobe von Psychologiestudenten ein transkribierter Interviewauszug vorgelegt wurde, den die Vpn hinsichtlich der Facetten des Konstrukts emotionale Belastbarkeit analysieren sollten. Die Sätze des sechsseitigen Interviewauszugs waren durchnummeriert und sollten den passenden Facetten des Konstrukts, welche auf einem Untersuchungsbogen abgebildet waren, zugeordnet werden. Hierbei wurden zwei Gruppen miteinander verglichen. Die Versuchsgruppe bekam für die Analyse des Interviews ein Kategoriensystem zur emotionalen Belastbarkeit als Auswertungshilfe an die Hand. Die Kontrollgruppe hingegen hatte für die Analyse des Interviews lediglich den Untersuchungsbogen mit der Facettierung von emotionaler Belastbarkeit zur Verfügung (siehe Liebert, 2006). Bei der Analyse des schriftlich vorliegenden Interviews mussten die Pbn die Sätze erkennen, „die eine Aussage zur emotionalen Belastbarkeit machten. Der zweite Schritt bestand in der Zuordnung dieser Sätze zu den passenden Facetten des Kategoriensystems.“ (Liebert, 2006, S. 40)

Westhoff (2005) definiert emotionale Belastbarkeit als „die stark genetisch festgelegte Neigung, gefühlsmäßige Belastungen eher ‚wegstecken‘ zu können, ein ‚dickeres Fell‘ zu haben oder andererseits eher empfindlich zu reagieren und ‚zart besaitet‘ zu sein“ Westhoff (2005, S. 130) und beschreibt im Anschluss an Westhoff und Kluck (2003), mit welchem Grad an

Robustheit oder Empfindlichkeit der Mensch im Allgemeinen reagiert. Sie wird auch als Synonym für emotionale Stabilität gesehen (Liebert, 2006). „Von der emotionalen Belastbarkeit, die in hohem Maße genetisch festgelegt ist, ist der Umgang mit Belastungen abzugrenzen, der im Laufe des Lebens erlernt wird.“ (Westhoff & Kluck, 2003, S. 29) Das Antonym emotionale Labilität/Neurotizismus „bezeichnet die gefühlsmäßige Empfindlichkeit eines Menschen“ (Westhoff & Kluck, 2003, S. 29) und gibt Aufschluss über die emotionale Belastbarkeit bzw. Nichtbelastbarkeit eines Menschen (vgl. Liebert, 2006).

Mit anderer Akzentuierung, nämlich mit Bezug auf persönlichkeitspsychologische Konstrukte wird insbesondere im Kontext eignungsdiagnostischer Anliegen versucht der Einsicht Rechnung zu tragen, dass von einer Dynamik auszugehen ist. So verweist Kieschke (2003) darauf, dass zunehmend das Modell „der richtige Mensch am richtigen Platz“ (S. 41) mit Bemühungen flankiert wird, Lern- und Entwicklungsprozesse bei betrieblichen Entscheidungsprozessen besser als bisher geschehen zu berücksichtigen (siehe hierzu auch den Ansatz und das Konzept des „person-environment-fit“ nach French, Rogers und Cobb, 1974). Mit diesen Hinweisen ist auch eine Erweiterung des Spektrums der methodischen Herangehensweisen angesprochen.

„Das reicht von komplexeren Auswahlverfahren, Bedarfsanalysen und Optimierungsangeboten (Assessment Center, Verfeinerung instruktionspsychologischer Designs für Qualifizierungsmaßnahmen etc.; vgl. Drenth, Therry & de Wolf, 1998; Sarges, 1995) bis hin zu meßtheoretischen Erwägungen (z. B. Ausweitung der Lerntestidee auf Fragestellungen der Personaldiagnostik; vgl. Guthke, 1991). Zudem könnte die prädiktive Validität vieler Persönlichkeitsfragebogen hartnäckige Methodenkritiker dezent über die Nützlichkeit der Erhebungsinstrumente belehren (vgl. Ones, Viswesvaran & Schmidt, 1993; Robertson & Kinder, 1993; Schuler, 1996). Barrick und Mount (1991) etwa berichten in ihrer Metaanalyse, „daß allein schon der Big-Five-Faktor ‘Gewissenhaftigkeit’, erfaßt per Selbsteinschätzung, mit dem Vorgesetztenurteil zum beruflichen Erfolg der Aspiranten zu immerhin .23 korreliert“ (Kieschke, 2003, S. 41).

Der Gedanke der Erweiterung der methodischen Herangehensweise in derartigen Entscheidungsprozessen ist auch in dieser Arbeit ein Leitgedanke, der systematisch verfolgt wird und der in einen Vorschlag zur Operationalisierung psycho-physischer Belastbarkeit mündet.

Nerdinger, Blickle und Scharper (2008, S. 301) verweisen zudem darauf, dass

„Judge, Higgins, Thoresen und Barrick (1999) Langzeitstudien ausgewertet [haben], bei denen in Kalifornien Persönlichkeitsmerkmale von Personen im Alter zwischen 12 und 14 Jahren erfasst und dann mit dem Berufserfolg ca. 30-35 Jahre später, also im Alter zwischen 41 und 50 Jahren in Beziehung gesetzt wurden. Dabei zeigte sich zum einen eine relativ hohe Stabilität der Persönlichkeitsmerkmale (die Durchschnittskorrelation betrug $r = .43$) und zum anderen, dass zwei Persönlichkeitsmerkmale besonders bedeutsam für den objektiven Berufserfolg - erfasst über Einkommen und sozialen Status - waren, nämlich Gewissenhaftigkeit und Neurotizismus. Das Persönlichkeitsmerkmal Gewissenhaftigkeit sagte den Berufserfolg positiv vorher, das Persönlichkeitsmerkmal Neurotizismus dagegen negativ. Diese Befunde decken sich auch mit einer Metaanalyse zum Zusammenhang von Persönlichkeitsmerkmalen und beruflichen Erfolg (Barrick, Mount & Judge, 2001). Das Persönlichkeitsmerkmal Neurotizismus ist der negative Pol der Dimension der emotionalen Stabilität. Je höher die emotionale Stabilität ist, desto besser kann sie mit Belastungssituationen umgehen und konstruktive Bewältigungsstrategien dafür finden. Barrick und Kollegen fanden auch folgenden Zusammenhang: Je höher die emotionale Stabilität und die Gewissenhaftigkeit von Personen ausgeprägt sind, desto besser integrieren diese sich in Arbeitsteams und arbeiten dort konstruktiver mit. Judge, Bono, Ilies und Gerhardt (2002) konnten im Rahmen einer Metaanalyse außerdem aufzeigen, dass das Persönlichkeitsmerkmal der Extraversion sowohl mit der spontanen Anerkennung als Führungspersönlichkeit (>> leadership emergence <<) im jeweiligen sozialen Kontext (Wirtschaft, Militär, Universität) als auch mit dem Führungserfolg in konsistent positiver Beziehung steht.“

Aus diesen Arbeiten geht zusammenfassend eindrücklich hervor, dass Neurotizismus bzw. dessen Gegenpol die emotionale Stabilität, gerade hinsichtlich beruflichen Erfolgs und dementsprechend auch bereits bei Prozessen der Personalgewinnung und -auswahl eine besondere Bedeutung zukommt. Dies ist teilweise auch in vorliegenden Verfahren zur berufsbezogenen Persönlichkeitsdiagnostik berücksichtigt und wird hier bei der Konzeptualisierung und Entwicklung eines Verfahrens zur Diagnostik der psychischen Komponente der psychophysischen Belastbarkeit aufgegriffen (siehe Kapitel 4.3.1). Mit diesen den theoretischen Teil abschließenden Hinweisen ist bereits auf die Komplexität der zu berücksichtigenden Faktoren (Kapitel 3) für eine angemessene empirische Herangehensweise (Kapitel 4) und auch die angezeigten Verfahrensentwicklungen (siehe Kapitel 5) aufmerksam gemacht. In der nun folgenden Darstellung des Stands der Forschung werden die Faktoren Persönlichkeit, psychische Gesundheit, Stress und Neurotizismus bzw. Emotionale Stabilität als Bezugspunkte in Verbindung mit Sport aufgearbeitet.

3 BEZUGSPUNKTE UND STAND DER FORSCHUNG

Bei den Literaturrecherchen und der Aufarbeitung des Forschungsstandes für diese Arbeit soll zunächst ein Überblick über die einzelnen, hier speziell relevanten Begrifflichkeiten geschaffen werden und es geht darum, Studien und Untersuchungen über Zusammenhänge mit relevanten Konstrukten zu sichten. Suchkriterien wie „physische Leistungsfähigkeit“, „körperliche Fitness“, „physische Fitness“, „Ausdauerleistungsfähigkeit“ und „Sport“ wurden im Verbund mit den Kategorien „psychische Stabilität“, „mentale Stärke“, „psychisches Wohlbefinden/Gesundheit“, „(mentale) Belastbarkeit“, „psychische Belastbarkeit“, „Neurotizismus“ und „Stress(-resistenz)“ in diversen Datenbanken, Katalogen und Suchmaschinen erkundet und gesichtet. Vor allem die uneinheitliche Definition von „emotionaler Stabilität“, oder auch „Ich-Stärke“ und „Zufriedenheit“ bzw. die enorme Weitläufigkeit der Begriffe erschwerte die Recherche nach Arbeiten von unterschiedlicher wissenschaftlicher Provenienz (Sportwissenschaft, Arbeitswissenschaft, Psychologie, Soziologie) und mit einem Bezug zum Thema bzw. Schlüsselkonzept „psycho-physische Belastbarkeit“. Von besonderer Relevanz war es auch, möglichen Hinweisen in der Literatur auf eventuelle Zusammenhänge zwischen physischer Fitness und Persönlichkeitseigenschaften nachzugehen, insbesondere hinsichtlich der Faktoren „Belastbarkeit“ und „Emotionale Stabilität“. Betrachtet man speziell das Forschungsfeld „Sport und Persönlichkeit“, was sich besonderer sportwissenschaftlicher (speziell sportpädagogischer und sportpsychologischer) Beliebtheit in den 70er und 80er Jahren erfreute, so ist in diesem Forschungsbereich nicht gerade ein Mangel an Untersuchungen und Beiträgen feststellbar. Allerdings sind die Befunde und Ergebnisse vieler Untersuchungen zum Zusammenhang von sportlicher Betätigung und speziellen Persönlichkeitsmerkmalen kritisch zu sehen, da häufig methodische Defizite zu konstatieren sind, worauf im einzelnen zurückzukommen sein wird.

Darüber hinaus ist im Hinblick auf die Forschung zum Zusammenhang von Sport bzw. körperlicher Betätigung und Persönlichkeitseigenschaften eine mehrperspektivische Komplexität zu beachten. Diese bezieht sich zum einen auf die Art (Sportart) und das Niveau (Freizeit-, Breiten-, Leistungssport) der sportlichen Betätigung, zum anderen auf die Auswahl der jeweiligen Eigenschaft(en) und deren Operationalisierung bzw. die Art und Weise (Methode) der Erfassung der Eigenschaft(en). Zudem ist die methodische Herangehensweise (Art und Auswahl der Fragebogen, Skalen, Tests etc.), die zur Persönlichkeitserfassung Anwendung finden, überaus heterogen. Insofern ist festzuhalten, dass sich eine äußerst diversifizierte Sachla-

ge und komplexe Herausforderung zur Feststellung von Zusammenhängen sowie für Vergleiche zwischen Studien und deren Befunden mit dieser Arbeit ergibt. Dazu äußerte sich auch bereits Wankel (1985), der insbesondere die Theorielosigkeit etlicher Studien zum einen und die eingesetzten Messinstrumente zum anderen kritisiert und der ebenfalls eine daraus resultierende Problematik der Vergleichbarkeit erkannte und diese als abwegig einschätzte (vgl. auch Pahmeier, 1994).

Da mit diesem Beitrag auch die Intention verfolgt wird, der praktischen Relevanz und dem Anwendungsbezug Rechnung zu tragen, sind noch weitere Praxisfelder von Interesse, speziell die Bezugfelder der Arbeits- und Organisationspsychologie. Persönlichkeitsmerkmale und deren Erfassung, die Themen Belastung, Beanspruchung und Stress am Arbeitsplatz und in Organisationen, Präventions- und Interventionsmöglichkeiten, sowie die Eignungsdiagnostik stellen wesentliche Bezugspunkte dar. Des Weiteren besteht die Herausforderung nicht nur darin, für diese Arbeit wichtige Studien und Forschungsarbeiten ausfindig zu machen und so Möglichkeiten einer Operationalisierung des Konstrukts "psycho-physische Belastbarkeit" auszuloten, sondern auch einen angemessenen theoretischen Rahmen zu entwickeln, in dem die Thematik einzuordnen ist. Die wesentlichen Befunde, die für diese Arbeit eine besondere Relevanz darstellen, seien im weiteren Verlauf insbesondere im Hinblick auf die Forschungsschwerpunkte Sport und psychische Gesundheit, Sport bzw. körperliche Aktivität und Persönlichkeitsaspekte, die Auswirkung bzw. die präventive Wirkung des Sports/Fitnesszustands auf das Stressempfinden und der Zusammenhang von emotionaler Stabilität/Labilität und Stress resümiert.

3.1 Persönlichkeit und Sport

Im Hinblick auf den Zusammenhang von Persönlichkeit und Sport gilt es in diesem Abschnitt primär Studien vorzustellen, die sich auf die hier relevanten Persönlichkeitsmerkmale beziehen oder allgemeine Hinweise liefern, die zur Differenzierung des Zusammenhangs zwischen Sport und Persönlichkeit beitragen. Von einer umfassenden Aufarbeitung des Themas Sport und Persönlichkeit in seiner Gänze soll aufgrund der gegebenen Spezifität in dieser Arbeit abgesehen werden. Den gesichteten Beiträgen und Arbeiten ist insgesamt zu entnehmen, dass sich die Zusammenhänge insbesondere dann verdeutlichen, wenn die Strukturähnlichkeit in den Handlungsbezügen (z. B. Anforderungen an psychische Funktionen) berücksichtigt wer-

den. In diesem Sinne spezifiziert und konkretisiert sich der hier interessierende Zusammenhang auf psycho-physische Anforderungen und die diesbezügliche Belastbarkeit. Dies wird im Rahmen des Belastungs-Beanspruchungskonzepts aufgegriffen (siehe Kapitel 2.2) und weiter ausgeführt.

Die Persönlichkeitsforschung im Sport wurde bereits früh in der ersten Hälfte des 20. Jahrhunderts in einen wissenschaftlich sportbezogenen Kontext gebracht. Der Höhepunkt der Untersuchungen liegt allerdings in den 60er bis Anfang der 80er Jahre. In den vergangenen 20 Jahren sind nur sporadisch Studien und Untersuchungen in diesem Bereich vorgenommen worden (Allen, Greenless & Jones, 2013). Die frühen Untersuchungen beschäftigten sich insbesondere mit der Thematik der „Sportlerpersönlichkeit“ und hatten vorwiegend Persönlichkeitsunterschiede zwischen Sportlern und Nichtsportlern zum Forschungsgegenstand (vgl. Allen et al., 2013; Fisher, 1984; Singer, 2000). Die zugrunde liegenden Arbeiten in diesem Bereich waren überwiegend ohne theoretischen Bezugsrahmen und die dabei verwendete Methodik beinhaltete verschiedene persönlichkeitsdiagnostische Verfahren, wie bspw. das von Eysenck und Eysenck (1975) entwickelte Eysenck-Personality Questionnaire (EPQ), das Eysenck-Personality-Inventory (EPI, Eysenck, 1964), den Sixteen Personality Factor Questionnaire (16 P. F.-Test) von Cattell (1946), das Maudsley Personality Questionnaire (Eysenck, 1952), das Maudsley Personality Inventory (MPI) von Eysenck (1959) und das Freiburger Persönlichkeits Inventar (FPI) von Fahrenberg, Selg und Hampel (1970). Zu den herangezogenen Stichproben gehörten Sportler und Nicht-Sportler, Olympiateilnehmer, Wettkampfsportler, Teilnehmer an Sportprogrammen verglichen mit Nicht-Teilnehmern, Sportler verschiedener Sportarten, wie z. B. Mannschaftssportler und Individualsportler, Spitzenathleten und Mitglieder von Universitäts- und Schulmannschaften.

„Die Skala der eingesetzten persönlichkeitsdiagnostischen Verfahren reicht vom Rating (Einschätzung) von Persönlichkeitsdimensionen durch Fremdpersonen über verschiedenartigste Fragebogen, zumeist sogenannten Persönlichkeitsinventaren, bis hin zu verschiedenen sog. projektiven Verfahren und klinischen Fallanalysen.“ (Singer, 2000, S. 299)

Nach Singers Analyse der vorliegenden Literatur und der zugrunde liegenden empirischen Untersuchungen, kann resümiert werden, dass keine Aussage über eine spezifische Sportlerpersönlichkeit getroffen werden kann. Singer und Haase (1975, S. 28) konstatierten bereits, dass die Ergebnisse, „so inkonsistent und widersprüchlich sind, daß sich – überspitzt formuliert – kaum eine positive oder negative Aussage über die Persönlichkeit des Sportlers denken

läßt, die nicht durch die eine oder andere Untersuchung zu stützen wäre“. Vanek und Hosek (1977) setzten in einer Studie, bei der über 800 tschechische Spitzenathleten- und athletinnen (146 weiblich, 678 männlich) aus 27 Sportarten teilnahmen, mehrere persönlichkeitsdiagnostische Testverfahren (EPI, PI-MAS, Raven-Test und 16 P.F.) ein. Mit der Manifest-Anxiety-Scale (MAS) weisen Hosek und Vanek bei den Sportlern Angstwerte auf, die über dem Durchschnitt liegen. Durch das Eysenck-Personality-Inventory (EPI) konnte eine etwas höhere Ausprägung der Extravertiertheit, sowie ein geringerer Neurotizismuswert im Vergleich mit der Normpopulation ermittelt werden, jedoch gehen die Autoren davon aus, „daß die Gesamtheit der Sportler sich im EPI wahrscheinlich nicht signifikant von der übrigen Bevölkerung unterscheiden wird“ (Vanek & Hosek, 1977, S. 88). Auf Grundlage dieser Ergebnisse schlussfolgern Hosek und Vanek (1977), dass ein hypothetisches Persönlichkeitsprofil des Sportlers äußerst spekulativ sei und stützen somit die Aussage von Singer (2000), dass von einer typischen (Spitzen-)Sportlerpersönlichkeit nicht gesprochen werden kann. Dies kann in erster Linie auf inkonsistente Ergebnisse, aufgrund der unterschiedlichen methodischen Vorgehensweisen, zurückgeführt werden (siehe Martens, 1975; Morgan, 1980; Sack, 1982; Singer, 2000; Vealey, 1989). Diese inkonsistenten Ergebnisse sind jedoch, so Singer (2000), auch bei methodisch vergleichbaren Untersuchungsvorgehen nicht auszuschließen. Speziell im Hinblick auf den Persönlichkeitsfaktor Neurotizismus wird eine geringere Ausprägung bei Sportlern vermutet bzw. wird bei Sportlern von der Tendenz ausgegangen, emotional stabiler zu sein als Nicht-Sportler. Singer (2000, S. 303 ff.) stellt dazu eine inkonsistente Befundlage fest, interpretiert diese allerdings differenzierend so:

„Da Neurotizismus mit Reizbarkeit, emotionaler Labilität, innerer Spannung und Ängstlichkeit verbunden ist und Angst als Triebreiz von denjenigen Reizen, die für eine optimale Erfüllung der jeweiligen Aufgabe verarbeitet werden müssen, ablenken kann, gehen die Autoren davon aus, daß zumindest herausragende Sportler eher niedrigere Neurotizismuswerte aufweisen. Allerdings ist hier die Situation insofern kompliziert, als zwischen Angst und Leistung auch eine kurvilineare Beziehung möglich ist, d. h., zu niedrige Werte gehen mit geringer Motivation einher, und zu hohe Werte führen zur Ablenkung von der eigentlichen Aufgabe. Genauere Vorhersagen sind also schwierig.“

Die am zahlreichsten eingesetzten Verfahren bei der Persönlichkeitsforschung von Sportlern stellen die drei von Eysenck entwickelten Skalen Maudesley-Personailty-Inventory (MPI), Eysenck-Personailty-Questionnaire (EPQ) und das Eysenck-Personality-Inventory (EPI) dar, anhand derer die Persönlichkeitstendenzen von (Spitzen- und Freizeit-)Sportlern durch die Ausprägung der Dimensionen (MPI: Extraversion und Neurotizismus; EPI: Extraversion,

Neurotizismus und Lügenskala; EPQ: Extraversion, Neurotizismus und Psychotizismus) dargestellt wurden. In seinem Überblicksbeitrag berücksichtigt Singer (2000) einige dieser Forschungsarbeiten mit dem MPI, EPI und dem EPQ in Bezug auf die Sportlerpersönlichkeit. Hinzu kommen noch Arbeiten von Knapp (1965), der mit dem MPI 46 Wimbledon Tennisspieler (28 weiblich, 18 männlich) untersuchte und keine signifikanten Unterschiede in Hinblick auf die Extravertiertheit und den Neurotizismus im Vergleich zur englischen Normalpopulation nachweisen konnte.

Morgan (1968) untersuchte 23 Ringer im Rahmen der Amateurweltmeisterschaft mit dem EPI und fand ebenfalls keine signifikanten Unterschiede bei den Dimensionen Extraversion und Neurotizismus im Vergleich zu erwachsenen und studentischen Normgruppen. Dahingegen konnten Stone (1961), der mit dem MPI 50 Athleten auf nationalem britischen Niveau untersuchte, sowie Dimsdale (1968), der anhand des EPI 33 Athleten und Teilnehmer eines nationalen walisischen Lehrgangs befragte, beide erhöhte Neurotizismuswerte feststellen. Morgan und Costill (1972) untersuchten mit dem EPI neun amerikanische Marathonläufer und konnten genauso wie Hagberg, Mullin, Bahrke und Limburg (1979) bei einer ähnlichen Studie mit dem EPI bei neun Spitzenradrennfahrer aus den USA erhöhte Introversions- und durchschnittliche Neurotizismuswerte nachweisen (siehe Abbildung 9). Sack (1975) konnte mit einer Stichprobe von 364 Mittel- und Langstreckenläufern und einer Kontrollgruppe von 291 Jugendlichen durch den FPI und einer Eigenschaftswörterliste herausfinden, dass die Athleten emotional stabiler, psychisch robuster und leistungsorientierter sind im Vergleich zur herangezogenen Kontrollgruppe, wenn auch insgesamt gesehen hinsichtlich der Persönlichkeitsunterschiede nur geringfügige Gefälle festzustellen waren. In einer Untersuchung mit dem 16 P.F.-Test von Hendry (1968) stellte sich heraus, dass die Stichprobe von Schwimmerinnen und Schwimmern (N = 126) zwischen 14 und 22 Jahren im Vergleich zur Normalpopulation erhöhte Angstwerte aufwiesen, jedoch keine Unterschiede in der Extraversion feststellen waren. Nach Kirkcaldy (1982) konnte in einer Untersuchung mit dem EPQ festgestellt werden, dass Athletinnen und Athleten (unterschiedlicher Leistungsstufen und Sportarten) in Bezug auf die Normierungsgruppe höhere Extraversions-, Psychotizismus- und niedrigere Neurotizismuswerte aufweisen. Diese Ergebnisse stützen die Folgerungen von Eysenck et al. (1982), dass Sportler im Gegensatz zu Nicht-Sportlern niedrigere Neurotizismuswerte, sowie höhere Werte auf dem Psychotizismusfaktor verzeichnen und extravertierter sind.

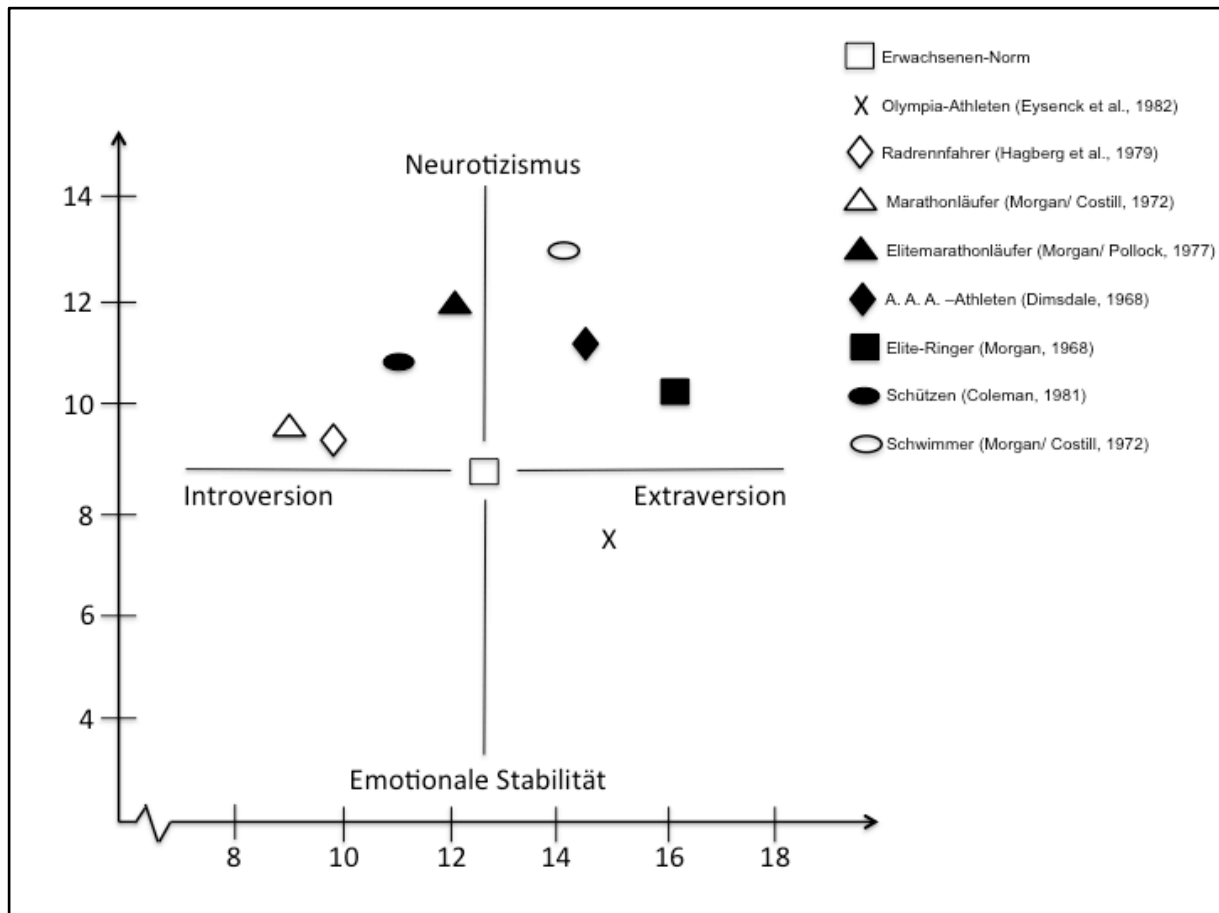


Abbildung 9. Untersuchungsergebnisse im EPI bei verschiedenen Gruppen männlicher Athleten (nach Singer, 2000, S. 306).

Als Fazit stellt Singer (2000, S. 308) in seinem Übersichtsbeitrag fest, dass

„die Suche nach ‚der‘ Persönlichkeit des Sportlers bzw. der Sportlerin – operationalisiert über relativ generelle Persönlichkeitseigenschaften – (. . .) vergebliche Liebesmüh` [ist] und (. . .) so nicht fortgesetzt werden [sollte]. Sport ist ein so komplexes und inhomogenes Phänomen und wird in so vielfältigen Realisierungsformen (Sportarten, privat und organisiert betriebener Sport, Breiten- und Hochleistungssport etc.) betrieben, daß es offenbar keine durchgehenden Bedingungen und Anforderungsstrukturen gibt. Von daher ist auch nicht zu erwarten, daß, der‘ Sport nur Menschen mit bestimmten Persönlichkeitseigenschaften anzieht (Selbstselektion der SportlerInnen), auswählt (Fremdselektion etwa über Talentauswahl und Förderung), an sich bindet (Dropout) und/oder bestimmte Persönlichkeitseigenschaften durchgehend verstärkt bzw. herausbildet (Sozialisation). Die Inkonsistenz der Ergebnisse liegt daher nicht nur in dem unterschiedlichen methodischen vorgehen (unterschiedliche Persönlichkeitstests, unterschiedliche SportlerInnen-Stichproben und Vergleichsgruppen) der diesbezüglichen Untersuchungen begründet, sondern zu einem wesentlichen Teil auch in der Sache selbst.“

Den Forschungsstand der sportbezogenen Persönlichkeitsforschung wird auch von Rushall (1972) annähernd identisch (nach Martens, 1975, S. 15) zusammengefasst:

„After all the years of research no clear findings are available. Physical education and sports personality researchers are not yet off the ground. It is evident that the investigative process must be restored in new directions, utilizing new techniques and designs adopting theoretically sound bases for each work, and avoiding all the errors in the past.“

In einem Beitrag von Schwinger, Olbricht und Stiensmeier-Pelster (2013) sind Wettkampf- (n = 149) und Freizeitschwimmer (n = 79) anhand der Persönlichkeitsfaktoren des Big Five Modells (Extraversion, Neurotizismus, Offenheit, Gewissenhaftigkeit und Verträglichkeit) verglichen und auf Zusammenhänge mit der Bestzeit in ihrer jeweiligen Schwimmdisziplin hin untersucht worden. Der Zusammenhang der Big Five Persönlichkeitseigenschaften mit der sportlichen Leistung wird in dieser Studie über eine Kausalkette bzw. Hierarchiestufen hergestellt, wobei die Persönlichkeitsfaktoren des Big Five Modells als dispositionale Traits (stabile Persönlichkeitsmerkmale) die erste Stufe darstellen, die jeweils bestimmte habituelle Merkmale (Gewohnheiten, die je nach Situation und Anreiz veränderbar sind) wie motivationale Zielorientierungen auf der zweiten Stufe und spezifische Verhaltensweisen, hier speziell Self-Handicapping und Ausdauer, auf der dritten und vierten Stufe der Hierarchieebene beeinflussen und in der sportlichen Leistung der Schwimmer resultieren. Mit der Ausdauer ist in dieser Studie nicht die körperliche Ausdauerleistungsfähigkeit, sondern das psychische Durchhaltevermögen und die mentale Belastbarkeit gemeint, die - wie auch alle anderen Faktoren in der Untersuchung - von Schwinger et al. (2013) über subjektive Selbsteinschätzung durch verschiedene Fragebogen erfasst wurden (NEO-FFI, Skalen zur Erfassung der Lern- und Leistungsmotivation, Academic Self-Handicapping Scale, einer Subskala aus dem Sportbezogenen Leistungsmotivationstest, sowie der Angabe der jeweiligen Bestleistung in der Paradedisziplin Schwimmen).

Für die beiden Teilstichproben Wettkampf- und Freizeitschwimmer wurde zur Überprüfung des hierarchisch aufgebauten Modells separat jeweils ein Pfadmodell erstellt, in dem die „Einflüsse der psychologischen Prädiktoren auf die Bestzeit im Schwimmen modelliert wurden“ (Schwinger et al., 2013, S. 88). Insgesamt konnten signifikante Zusammenhänge bei den Wettkampfschwimmern zwischen Vermeidungs-Leistungzielorientierung und Self-Handicapping mit schlechteren Schwimmzeiten, sowie signifikante Korrelationen in Höhe von $r = -.32$ ($p \leq .01$) zwischen der Ausdauer und der Schwimmbestzeit festgestellt werden.

Des Weiteren konnten die Autoren zwischen den Annäherungs-Leistungsziel orientierten Personen und der Bestzeit einen Korrelationskoeffizienten in Höhe von $r = -.26$ ($p \leq .05$) berechnen. Durch die Pfadanalyse konnte u. a. festgestellt werden, dass der Faktor Neurotizismus die Verfolgung von Vermeidungs-Leistungszielen signifikant positiv beeinflusst ($r = .37$; $p \leq .01$) bei Wettkampfschwimmern und zu $r = .21$ ($p \leq .05$) bei Freizeitschwimmern. Weiterhin konnte die „spezifische Verhaltensweise Ausdauer“ (Schwinger, et al., 2013; S. 88) bei beiden Gruppen durch die Prädiktoren Annäherungs-Leistungsziele und negativ durch Vermeidungs-Leistungsziele sowie den Persönlichkeitsfaktor Neurotizismus vorhergesagt werden. Die Korrelationskoeffizienten liegen dabei von $r = -.23$ ($p \leq .01$) bis $r = -.41$ ($p \leq .01$) in einem signifikanten Bereich. Der Ansatz der Pfadanalyse zur Überprüfung von direkten und indirekten Zusammenhängen zwischen Persönlichkeitsfaktoren und sportlicher Leistung über Variablen in der Kausalkette, stellt ein sinnvolles Vorgehen dar, welches als sehr nützlich für ähnliche Untersuchungen zum Zusammenhang von Persönlichkeitseigenschaften und körperlicher Leistungsfähigkeit erscheint. Auch in dieser Arbeit liegt die Annahme zugrunde, dass zwischen sportlicher Leistung und Persönlichkeitsfaktoren wie Emotionale Stabilität und Belastbarkeit (= „Psychische Belastbarkeit“) bzw. zwischen der physischen Leistung und einer der fundamentalen Persönlichkeitseigenschaften des Big Five Modells, nämlich Neurotizismus, Zusammenhänge bestehen, worüber in Kapitel 5.3.1.2 detailliert berichtet wird.

McKelvie, Lemieux und Stout (2002) untersuchten verschiedene Gruppen von Sportlern einer Universität (Athleten von Kontakt-Sportarten, $n = 46$ und Athleten von Nicht-Kontakt-Sportarten, $n = 40$) sowie einer Kontrollgruppe von Personen, die keiner professionellen Universitätsmannschaft angehörten und keiner sportlichen Betätigung nachgingen ($n = 86$) mit dem Eysenck Personality Inventory im Hinblick auf die Persönlichkeitseigenschaften Extraversion und Neurotizismus. Bei den varianzanalytischen Berechnungen zeigte sich, dass die Athleten wesentlich geringere Itemscores bei der Neurotizismus-Skala aufweisen, als die Nicht-Sportler ($F = 7.63$, $p \leq .01$) und sich demnach emotional stabiler zeigen als die Kontrollgruppe. Zwischen den Athleten der Kontakt- bzw. Nicht-Kontakt-Sportarten können die Autoren keinen Unterschied feststellen.

In einer Untersuchung „über Persönlichkeitsunterschiede in Abhängigkeit von Niveau und Art des sportlichen Leistungsvermögens“ von Wolfgang Conrad (ohne Jahresangabe) vom Psychologischen Institut der Universität Mannheim wurde den folgenden Hypothesen nachgegangen:

„H0: Es bestehen keine Persönlichkeitsunterschiede zwischen Gruppen mit unterschiedlichem sportlichen Leistungsniveau

H1: Zwischen solchen Gruppen bestehen Unterschiede im Hinblick auf folgende Persönlichkeitsdimensionen:

- Ausdauer/Belastbarkeit
- Allgemeine Leistungsmotivation
- Ängstlichkeit
- Geltungsstreben/Narzismus
- Aktivität
- Wettkampfmotivation
- Toleranz
- Paranoia
- Dominanzstreben
- Neurotizismus
- Selbstkritik
- Extraversion
- Aggressivität
- Maskulinität
- Soziale Einstellung
- Depressivität“

Conrad (ohne Jahresangabe) operationalisierte das Leistungsniveau, indem er die Teilnehmer (N = 808), erfasst durch einen Fragebogen, in drei Gruppen einteilte („Leistungssportler“, „Freizeitsportler“ und „Nicht-Sportler“). Die Gruppe der befragten Leistungssportler bestand ausschließlich aus Leichtathleten, die der Bestenliste [des Deutschen Leichtathletik Verbands] in ihrer Disziplin, angehörten. Als Freizeitsportler (n = 262) galten alle Adressaten verschiedener Landessportbünde, die die „Normen zur Erlangung des deutschen Sportabzeichens erfüllten“ (o. J., S. 154). Die Stichprobe der Nicht-Sportler (n = 194) ergab sich aus Personen, die bei der Beantwortung des Fragebogens angaben, sich im vorausgegangenen Kalenderjahr nicht sportlich betätigt zu haben. Durch den Fragebogen wurden darüber hinaus die unabhängigen Variablen „sportliches Leistungsniveau“, „Geschlecht“, „Lebensalter“ und „Beruf“ erfasst, wobei die Variable Beruf in Akademiker (B1) und Nicht-Akademiker (B2) unterteilt wurde, um so durch Varianzanalysen eventuell weitere Einflussfaktoren auf die Persönlichkeitsmerkmale im Verlauf der Untersuchung aufzudecken.

Zur Erfassung der 16 Persönlichkeitsfaktoren wurde auf Grundlage verschiedener Skalen eine Fragebogenbatterie mit insgesamt 394 Persönlichkeits-Items konstruiert. Mit Hilfe verschiedener Prozeduren bzw. Varianten der Faktorenanalyse wurde eine optimale bzw. passen-

de Faktorenstruktur (signifikante Einfachstruktur) berechnet. Danach ergaben sich drei gut zu interpretierende Faktoren („Emotionale Labilität“, „Leistungsorientierung“ und „Selbstbehauptung“).

Durch eine Vierweg-Varianzanalyse konnte der Autor feststellen, dass sich das sportliche Leistungsniveau signifikant auf die beiden Persönlichkeitsfaktoren „Emotionale Labilität“ sowie „Leistungsorientierung“ auswirkt. Conrad (o. J., S. 169) fasst dazu zusammen:

„Emotional stabilisierende Effekte durch die Ausübung von Leistungssport, besonders jedoch durch den in der Regel spielnäheren Freizeitsport, liegen zum Beispiel ebenso auf der Hand wie die Annahme, daß psychisch anfälligeren Persönlichkeiten von vornherein weniger Bedürfnisse nach sportlicher Betätigung verspüren. Nimmt man an, daß diese beiden Tendenzen zusammenwirken, dann liegt eine Verstärkung im Sinne des empirisch nachgewiesenen signifikanten Haupteffekts nahe.“

Hinsichtlich der Dimension „Selbstbehauptung“ wurde kein Unterschied zwischen den Gruppen in Abhängigkeit des sportlichen Leistungsniveaus festgestellt. Darüber hinaus konnten wesentliche Befunde im Hinblick auf die parallel erfassten unabhängigen Variablen und den Persönlichkeitsdimensionen ausgemacht werden. So zeigte sich, dargestellt durch die orthogonal abgebildeten Werte der Skala „Emotionale Labilität“ und „Leistungsniveau“, ein bedeutender Unterschied zwischen den Geschlechtern und Unterschiede in Abhängigkeit vom Alter. Außerdem zeigte sich zwischen den beiden varianzanalytischen Faktoren Disziplin, erfasst durch die Gruppe der Leistungssportler, und Geschlecht ein signifikanter Interaktionseffekt ($F = 10.63$; $p \leq .01$), sowie signifikante Haupteffekte dieser beiden Faktoren im Hinblick auf „Emotionale Labilität“ (Conrad, o. J.).

Young und Ismail (1977) setzten zur Bestimmung des Persönlichkeitsprofils von regelmäßig körperlich aktiven Personen, sowie von Einsteigern zu Beginn eines Sportprogramms und bei Aussteigern im Verlauf der Studie das 16 PF-Modell von Cattell (1946, 1993) ein. Bei den regelmäßig Aktiven konnte ein größeres Selbstvertrauen und größere emotionale Stabilität nachgewiesen werden (vgl. Pahmeier, 1994).

Fuchs (1997) weist auf die Tatsache hin, dass sich Athleten offensichtlich in körperlicher Hinsicht von anderen Menschen unterscheiden und es somit plausibel erscheint, dass es auch im Hinblick auf die Ebene psychischer Merkmale charakteristische Unterschiede geben müsste und stellt demzufolge die altbekannte Frage erneut, ob es eine typische Sportlerpersönlich-

keit oder eine Persönlichkeit in Abhängigkeit der Ausübung einer bestimmten Sportart gibt. Der Einfluss sportlicher Aktivität auf die Entwicklung und Entfaltung der Persönlichkeit wird im weiteren Sinne kaum bezweifelt. Betrachtet man jedoch Überblicksartikel bzw. zusammenfassende Forschungsarbeiten, die schwerpunktmäßig in den 60er bis Anfang der 80er Jahre zum Zusammenhang sportlicher Aktivität und Persönlichkeit auf empirischer Basis verfasst wurden, so findet man kein einheitliches Bild z. T. sogar widersprüchliche Befunde (siehe Vanek & Hosek, 1973; Sack, 1975; Singer & Haase, 1975; Mummendey 1983, 1989). Auch im Bereich der sportwissenschaftlichen Persönlichkeitsforschung muss, wie Fuchs (1993) es erläutert, im Hinblick auf die Persönlichkeit als abhängige und unabhängige Variable unterschieden werden. Bei der Persönlichkeit als abhängige Variable „geht es um die Frage, ob und ggf. wie sich die Persönlichkeit eines Menschen infolge des Sporttreibens verändert (Sozialisationshypothese).“ (Fuchs, 1993, S. 9) Hingegen geht es bei der Persönlichkeit als unabhängige Variable „um das Problem, ob bestimmte Persönlichkeitsmerkmale dazu prädestinieren, überhaupt sportlich aktiv zu sein, bestimmte Sportarten auszuwählen oder sportlichen Erfolg zu haben (Selektionshypothese)“ (Fuchs, 1993, S.9). In dieser Arbeit steht nicht die sportliche Aktivität, vielmehr die körperliche Leistungsfähigkeit bzw. der Ist-Zustand der physischen Fitness im Vordergrund, der durch die Ausübung sportlicher Aktivitäten beeinflusst werden kann und in Verbindung mit der Erfassung personaler Dispositionen (bestimmter Persönlichkeitseigenschaften) Hinweise auf eine evtl. berufliche Eignung bzw. eine Hilfestellung im Kontext des Selektionsprozesses bei der Personalauswahl geben soll. Die Genese personaler Dispositionen und ein möglicher Einfluss sportlicher Aktivität darauf bleibt hier ausgeklammert.

Speziell im Hinblick auf den hier relevanten Faktor „Emotionale Stabilität“ scheint ein Hinweis von Eysenck und Mitarbeitern (1982) jedoch interessant und erwähnenswert, der sich auf den Gegenpol von psychischer Stabilität bezieht, nämlich auf den Faktor „Neurotizismus“ (psychische Labilität). Sie fassen die Befundlage dahingehend zusammen, dass Sportler, insbesondere Top-Athleten, tendenziell niedrigere Neurotizismus-Werte aufweisen und weniger durch Angst beeinträchtigt werden als Nichtsportler bzw. unsportliche Personen (Eysenck, Nias & Cox, 1982).

Sack (1982) referiert und interpretiert 28 Untersuchungen zum Zusammenhang von sportlicher Aktivität und Persönlichkeitseigenschaften und resümiert, dass Jugendliche zwischen 12 und 22 Jahren, die sich sportlich betätigen „etwas geselliger, unbekümmerter (Extraversion),

weniger ängstlich (Neurotizismus) und dominanter“ (1982, S. 116) sind als Nichtsporttreibende und stellt zusammenfassend fest, dass „die Varianzaufklärung im Persönlichkeitsbereich erheblich ansteigt, je größer die Differenzen der Population hinsichtlich ihres Sportengagements werden“ (1982, S. 142).

Egloff und Gruhn (1998) untersuchten eine Stichprobe von 86 Ausdauersportlern (58 männlich, 28 weiblich; $M = 41,4$ Jahre; $SD = 9,9$ Jahre; Spannweite: 27-64 Jahre) anhand der Skalen Extraversion und Neurotizismus („Emotionalität“) des Freiburger Persönlichkeitsinventars (FPI-R) und eines Fragebogens zur Erfassung der Trainingsintensität und der Bestzeiten in der jeweiligen Ausdauerdisziplin (Laufen oder Triathlon). Ein wesentliches Anliegen der Studie von Egloff und Gruhn besteht in der Analyse von Zusammenhängen zwischen den beiden Persönlichkeitsdimensionen Extraversion und Neurotizismus mit Gründen für die Zuwendung zum Sport und deren Veränderungen durch das Ausüben von Sport, in diesem Fall speziell Ausdauersport. Da nicht alle Teilnehmer der Untersuchung Wettkämpfe bestritten, wurden sie nicht nach dem Niveau der sportlichen Aktivität durch die Bestzeit, sondern über den wöchentlichen Trainingsumfang kategorisiert.

Die Stichprobe bestand sowohl aus wettkampforientierten Ausdauerathleten nationalen Niveaus als auch aus Hobby-Läufern. Es kamen spezielle Skalen zur Erfassung von Gründen zum Beginn von Ausdauersport und zur wahrgenommenen Veränderung aufgrund der Ausübung von Ausdauersport zum Einsatz, die aus Skalen vorheriger Studien zusammengestellt wurden. Die beiden Skalen bestanden aus 39 und 36 Items, bei denen sich mittels einer Faktorenanalyse jeweils vier Faktoren identifizieren ließen. Gründe, mit Ausdauersport zu beginnen: (1) „Umgang mit negativen Affekten“ (bspw. „Depression“, „Ärger“, „Angst“), (2) „Erholung“ (z. B. „Ausgleich zu Alltäglichen Belastungen“), (3) „Leistung“ (z. B. „Leistung erbringen“) und (4) „Einfluß anderer und Aussehen“ (bspw. „Aussehen verbessern“). Die vier Faktoren zur wahrgenommenen Veränderung durch Ausdauersport wurden als (1) „Entspannung“ („Umgang mit Streß“), (2) „Zielerreichung/Erfolg“ („Gelegenheit mich richtig zu verausgaben“ und „Erfolgserlebnisse“), (3) „physische Gesundheit“ („Bemühungen für meine Gesundheit“) und (4) „Geselligkeit und Selbstkonzept“ (z. B. „neue Bekanntschaften“ und „Selbstbewusstsein“) charakterisiert. Mit insgesamt acht Regressionsanalysen sind die Effekte der Persönlichkeitsdimensionen Neurotizismus und Extraversion sowohl hinsichtlich der Gründe, mit Ausdauersport zu beginnen als auch im Hinblick auf die Veränderungen durch Ausdauersport untersucht worden. Zudem kamen die Prädiktoren Trainingsintensität, Alter

und Geschlecht als Kontrollvariablen zum Einsatz. Im Rahmen dieser Arbeit sind speziell die Befunde mit Bezug auf das Persönlichkeitsmerkmal Neurotizismus von Interesse, da dieses Konstrukt ein komplementäres Merkmal zur emotionalen Stabilität darstellt und in dieser Arbeit zur Analyse psycho-physischer Zusammenhänge herangezogen werden soll. Dazu ließen sich durch die von Egloff und Gruhn (1998) durchgeführte Studie folgende Hinweise gewinnen: „Personen mit hohen Werten in Neurotizismus suchen in der sportlichen Aktivität vor allem ‚Umgang mit negativen Affekten‘ und ‚Erholung‘. Als Veränderung werden Verbesserungen in ‚Entspannung‘ registriert“ (Egloff & Gruhn, 1998, S. 51). Die Autoren weiter:

„Neurotizismus sollte dagegen mit Maßnahmen zur Bewältigung von Zuständen emotionaler Labilität korrespondieren. Diese Hypothese konnten für die Dimension Neurotizismus in vollem Umfang bestätigt werden: Umgang mit negativen Affekten und Erholung als Ursache sowie Entspannung als Folge sportlicher Aktivität waren positiv mit Neurotizismus assoziiert. Personen mit hohen Werten in Neurotizismus, die durch mangelnde emotionale Stabilität charakterisiert sind, welche sich u. a. in einer erhöhten Anfälligkeit für negative Emotionen wie Gereiztheit oder Angst manifestiert, versuchen also, durch Ausdauersport diese Zustände zu bewältigen bzw. deren Auftreten vorzubeugen. Somit konnte die Annahme der ‚Passung‘ von Persönlichkeitsdispositionen auf der einen Seite und Motiven für die Aufnahme der sportlichen Tätigkeit bzw. wahrgenommenen Veränderung durch die sportliche Aktivität auf der anderen Seite durch dieses Ergebnis gut demonstriert werden.“ (Egloff & Gruhn, 1998, S. 52)

Die Autoren merken an, dass der Berücksichtigung des Faktors „physische Ausdauer“ nicht durch eine subjektive Selbsteinschätzung genüge getan wird, sondern um diesen besser zu überprüfen, „Datenquellen objektiver Art“ (Egloff & Gruhn, 1998, S. 53) wie bspw. Fitness-tests herangezogen werden müssten. In der hier durchgeführten Studie ist dies berücksichtigt und zur Erfassung der physischen Komponente auf objektive Daten aus einem spezifischen Ausdauerstest (siehe Kapitel 4.2 und 5.1) zurückgegriffen worden, die mit Selbstberichtsdaten aus Fragebogen zu spezifischen Persönlichkeitsmerkmalen in Zusammenhang gebracht werden (siehe dazu Kapitel 4.3 und 5.2).

In einer Studie von Piedmont, Hill und Blanco (1997) wurden 79 weibliche Fußballspielerinnen der höchsten amerikanischen Universitätsliga mit einer speziellen Version des Fünf-Faktoren-Inventars (FFI) von McCrae und Costa (1985, 1987) auf ihre Persönlichkeitseigenschaften hin untersucht und diese auf Zusammenhänge hinsichtlich der Prognostizierbarkeit der fußballerischen Leistungsfähigkeit analysiert. Ein vorrangiges Anliegen war es herauszufinden, inwieweit eine eventuelle Prognosefähigkeit bzw. Vorhersagbarkeit der fünf Persönlichkeitsdimensionen des FFI auf die tatsächliche sportliche Leistung festzustellen ist. Die

Leistungsfähigkeit der Athletinnen definierte sich zum einen über die in einer Saison gesammelten Spielstatistiken und zum anderen über die Bewertungen durch die Trainer. Die Autoren machen selbst auf die methodischen Schwachpunkte ihrer Herangehensweise aufmerksam indem sie konstatieren, dass Spielerstatistiken sich alleine aufgrund der unterschiedlichen Spielpositionen unterscheiden und die individuelle Leistung innerhalb einer Mannschaft des Weiteren stark von den Mitspielern abhängt. Somit sind Vergleiche mit anderen Sportlern aus anderen Sportarten nur schwer oder kaum zu realisieren. Diese Studie sollten Piedmonts (1995) Befunde einer früheren Untersuchung stützen bzw. untermauern, in der sich signifikante Zusammenhänge zwischen den Persönlichkeitsmerkmalen des FFI und der Prognostizierbarkeit der akademischen Leistung, ebenfalls bei weiblichen Personen, feststellen ließen. Die Persönlichkeitseigenschaften wurden durch die „Bipolar Adjective Scale“ mit insgesamt 80 Items, die die fünf Faktoren des FFI repräsentierten, von McCrae und Costa (1985, 1987) erfasst. Die Beurteilungen durch die Trainer bezogen sich auf fünf Dimensionen wie bspw. die „sportliche Leistung“ („athletic ability“), „Teamfähigkeit“ („team playerness“) und „Arbeitsethik“ („work ethic“) der Spielerinnen. Die Leistungsindizes wurden im Zeitraum einer Saison gesammelt und beinhalteten z. B. die Anzahl der geschossenen Tore, gespielte Pässe, Anzahl der absolvierten Spiele, usw.

Die Ergebnisse weisen mittlere und hoch signifikante Zusammenhänge zwischen den subjektiv eingeschätzten Persönlichkeitsdimensionen, insbesondere der Dimensionen Neurotizismus (bis zu $r = -.45$; $p \leq .001$ für „Game performance“; $r = -.40$; $p \leq .001$ für „Athletic ability“), sowie Gewissenhaftigkeit (bis zu $r = .39$; $p \leq .001$ für „work ethic“) und den Leistungsbewertungen der Trainer auf. Die Spielerstatistiken wiesen dagegen keine oder lediglich sehr geringe Zusammenhänge, bis auf Gewissenhaftigkeit und „Games played“ ($r = .47$; $p \leq .001$), mit den Persönlichkeitseigenschaften der Spielerinnen auf. In Folge dieser Befunde von Piedmont et al. (1997) kann ein signifikanter Zusammenhang des auch in dieser Arbeit relevanten Persönlichkeitsmerkmals Neurotizismus und der Leistungsfähigkeit („athletic performance“) der Fußballspielerinnen konstatiert werden. Die Autoren merken an, dass auf Grundlage der Ergebnisse die Variable „Spielerstatistik“ als ungeeignet erscheint. Wie bereits angemerkt sind die aus der Spielerstatistik entnommenen Leistungsfaktoren, v. a. vor dem Hintergrund der verschiedenen Spielpositionen und der daraus resultierenden differierenden Statistiken, für die eigentliche Leistungsfähigkeit nicht aussagekräftig. Wie auch in anderen Untersuchungen haben Piedmont et al. (1997) mit ihrer Studie nachgewiesen, dass der Faktor Neurotizismus in einem nicht unbedeutenden Zusammenhang mit der Leistungsfähigkeit steht

und gehen sogar soweit zu konstatieren: „The combination of low neuroticism and high conscientiousness represents the personality profile of the prototypical achiever: emotionally stable, capable, with a heightened sense of competence and drive to succeed.“ (Piedmont et al., 1997, S. 776) Mit der Variable „Spielerstatistiken“ machen die Autoren einen Versuch die subjektive Selbsteinschätzung mit (neben den Trainerbewertungen) objektiven Kriterien in Zusammenhang zu bringen. Obgleich diese Herangehensweise von Piedmont et al. (1997) bis auf die referierten kaum signifikante Resultate geliefert hat, wird der Ansatz subjektive Einschätzungen mit objektiven Merkmalen zu untersuchen in der vorliegenden Arbeit in anderer Form aufgegriffen.

Bevor spezifischer Konsequenzen der vorgestellten Forschung aufgezeigt werden, kann als vorläufiges Resumée Folgendes festgehalten werden: Der Komplexität des Zusammenhangs zwischen Sport und Persönlichkeit bzw. speziellen Persönlichkeitsmerkmalen wie Neurotizismus und emotionaler Stabilität ist bisher nicht hinreichend Rechnung getragen. Dies liegt insbesondere daran, dass sich z.B. Geschlechts- und Altersgruppen einerseits und Tätigkeitsmerkmale wie Sportart, Leistungsniveau, Dauer und Intensität der Betätigung sowie Umweltgegebenheiten (sozial wie ökologisch) im Verbund kaum umfassend berücksichtigen bzw. kontrollieren lassen. Für dieses Forschungsprojekt bedeutet dies, dass die notwendigerweise spezifische und eingeschränkte Herangehensweise bei der Untersuchung der Fragestellung bewusst zu wählen und bei der Auslegung der Befunde zu berücksichtigen ist.

3.2 Psychische Gesundheit und Sport

Die Wirksamkeit von körperlicher Aktivität auf psychische Gesundheit, bzw. psychische Belastbarkeit und emotionale Stabilität wird in der Wissenschaft genauso unterschiedlich diskutiert wie der Zusammenhang zwischen physischer und psychischer oder mentaler Fitness. Während einige Autoren, darunter Morgan und Goldston (1987), konstatieren, Fitnesstraining steigere das allgemeine Wohlbefinden, reduziere Angst und Depressionen, wie auch Neurotizismus und Ängstlichkeit, ziehen andere Autoren, wie Brown (1990), eine generelle Wirksamkeit von sportlicher Betätigung auf psychische Faktoren in Zweifel. Ein möglicher Grund für die inkonsistente Beurteilung der Auswirkung von Fitnesstraining auf die psychische Gesundheit ist nach Schlicht (1995) neben methodisch defizitären Studien (vgl. Folkins & Sime, 1981; Steptoe, 1994) insbesondere auch in der Definition des Begriffs „psychische Gesund-

heit" zu suchen. Unter dem Begriff werden Intelligenz- und Gedächtnisprozesse, Selbstachtung, Stressbewältigung und Stressvulnerabilität, Stimmung und Wohlbefinden, sowie Angst und Depressivität subsumiert (Schlicht, 1995). Somit stellt der Begriff „psychische Gesundheit" bzw. „psychische Stabilität" eher ein summarisches Konstrukt dar mit einer Definitionsvielfalt bzw. Vielfalt an Kompositionen, weshalb ein klares Verständnis nicht gewährleistet ist. In dieser Situation kann demzufolge auch von einem Mangel hinsichtlich der Operationalisierung der psychischen Gesundheit gesprochen werden. Schlicht (1990, S. 14) führt weiter an:

“Psychische Gesundheit wird zumeist auf Stimmungszustände eingegrenzt und über globale, unstandardisierte Befindensbeschreibungen operationalisiert. Objektive Testverfahren und Verhaltensmaße werden nur selten als abhängige Variablen definiert. Schwer wiegt das Versäumnis, die körperliche Leistungsfähigkeit im Anschluss an das Training zu überprüfen. Durch diesen Mangel wird die Entscheidung beeinträchtigt, ob veränderte Variablenausprägungen der psychischen Gesundheit auf der bloßen Teilnahme am Treatment beruhen oder mit veränderten physiologischen Prozessen kovariieren.”

Die systematische Eignungsfeststellung bzw. die Personalauswahl durch bestimmte Verfahren oder per Erfassung spezifischer Persönlichkeitskonstrukte dient nicht nur der Leistungsoptimierung, sondern auch der Erhaltung psychischer und physischer Gesundheit, dem allgemeinen Wohlbefinden und der Zufriedenheit, an denen nicht nur die Arbeitnehmer sondern auch die Arbeitgeber stark interessiert sein sollten (vgl. Schuler, 2000).

Schlicht (1990) fasst zusammen, dass aufgrund methodischer Mängel der Studien keine längerfristig stabilen Befindensverbesserungen nach sportlicher Betätigung konsistent nachgewiesen werden konnten. Die meisten Studien - weisen der Kritik dieses Autors folgend - eine mangelnde interne Validität durch konfundierte Treatmentbedingungen, d. h. eine Vermischung der experimentellen Maßnahmen (Ausdauertraining und zusätzliche psychotherapeutische Interventionen) auf. Andere Autoren verneinen nach Schlicht (1990) zwar, dass eine konsistente Befundlage vorliegt, behaupten aber auf Grundlage ihrer Untersuchungen erkannt zu haben, dass aerobes Ausdauertraining insbesondere die emotionale Komponente der Zustandsangst (Aufgeregtheit) beeinflusst, wohingegen die kognitive Komponente (Besorgtheit) nicht beeinflusst wird (Van Andel & Austin, 1984; Kerr & Vlaswinkel, 1990; Brown, 1990). Eine von Schlicht (1993) durchgeführte Untersuchung führte zu inkonsistenten Befunden, stützt im Ergebnis allerdings die Hypothese eines generell angstmindernden Effekts sportlichen Trainings. Eine Wirkung zeigt sich speziell bei Personen mittleren Lebensalters (zwi-

schen 30 und 50 Jahren), die ihre aerobe Ausdauer trainieren und ihre Angstgefühle um eine 16 Prozent höhere Wahrscheinlichkeit reduzieren als sportlich inaktive Personen.

Anhand der von Schlicht (1995) betrachteten Studien und deren Befunde, wird die körperliche Fitness bzw. sportliches Training als unabhängige Variable und die psychische Gesundheit als abhängige Variable gesehen. Er verweist darauf, dass bei den meisten Arbeiten nicht hinreichend erkennbar ist, ob zwischen den beiden Variablen ein direkter oder indirekter Zusammenhang besteht. In der folgenden Abbildung 10 wird die angemahnte Differenzierung der Wirkungszusammenhänge bzw. -wege im Sinne direkter und indirekter Wirkung veranschaulicht.

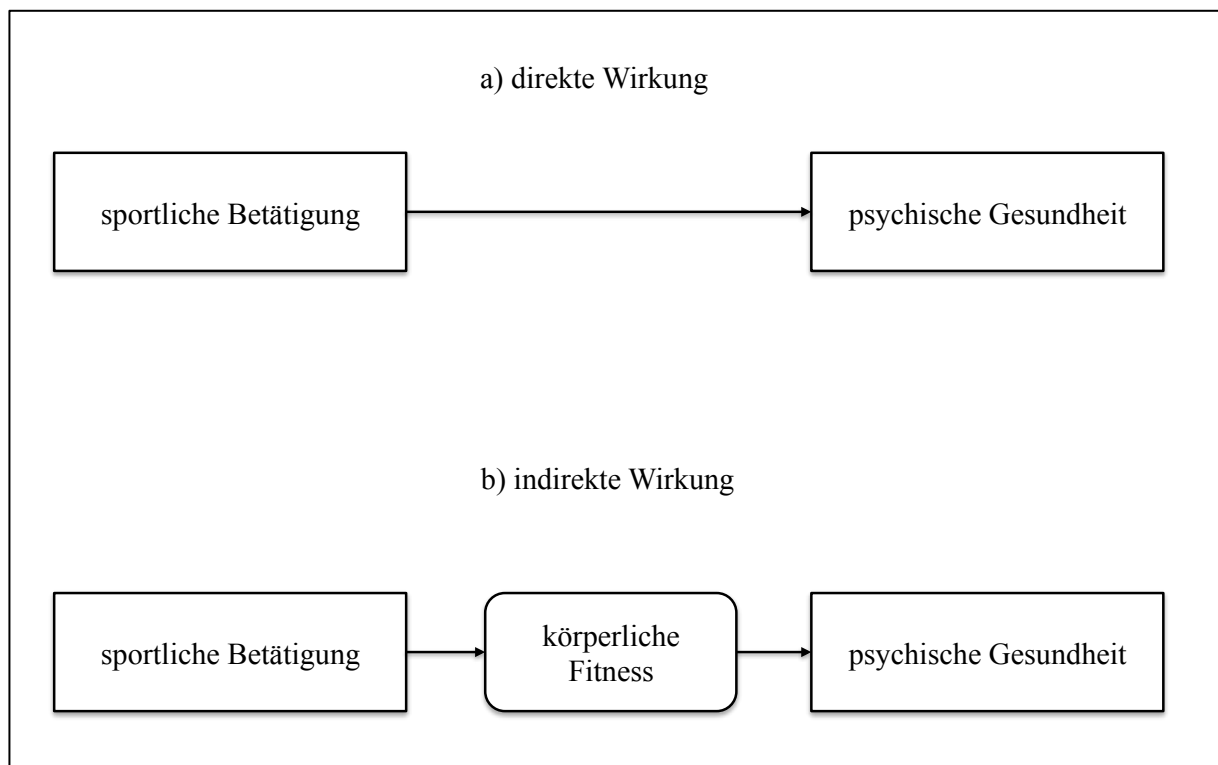


Abbildung 10. Ursachen-Wirkungshypothese (nach Schlicht, 1995, S. 33).

Zusammenfassend formuliert Schlicht (1995, S. 32): „Als globales Ziel ihrer Studien definieren die meisten Autoren die Analyse des Einflusses einer körperlich-sportlichen Betätigung (exercise) und der körperlichen Leistungsfähigkeit (physical fitness) auf die psychische Gesundheit (mental health).“ So allgemein gehalten können die inkonsistenten Ergebnisse nicht verwundern. Spezifischere Zusammenhänge (welche physischen und psychischen Aspekte, Komponenten sind jeweils angesprochen?) und differenziertere Ansätze (Design und Methoden) sind erforderlich, um einen Erkenntnisgewinn zu erzielen und praktisch nutzbares Wis-

sen zu generieren. Außerdem verweist Ulmer (2001, S. 1) auf den Zusammenhang zwischen Leistungsfähigkeit, Belastbarkeit und Gesundheit derart hin, dass die Leistungsfähigkeit anzeigt,

„inwieweit ein Mensch in der Lage ist, auf eine Anforderung (Belastung) adäquat zu reagieren, abhängig besonders von Eignung einschließlich 'Einübungszustand'. Hingegen zeigt die Belastbarkeit an, inwieweit ein Mensch ohne gesundheitliche Risiken belastet werden kann, speziell in Abhängigkeit von seinem Gesundheitszustand. Gesundheit und Leistungsfähigkeit korrelieren nur sehr bedingt miteinander, insbesondere ist eine überdurchschnittliche Leistungsfähigkeit kein Indiz für Gesundheit (vergl. Mit Leistungssportlern) und umgekehrt.“

Plante und Rodin (1990) sichteteten im Rahmen ihrer Untersuchung zum Zusammenhang zwischen physischer Fitness und psychischer Gesundheit („physical fitness and enhanced psychological health“) den Forschungsstand bis Ende der achtziger Jahre und schenkten 41 Studien und wissenschaftlichen Beiträgen der ausgemachten Literatur nähere Aufmerksamkeit. Die ausgewählte Literatur konnte in vier Bereiche der avisierten psychischen Befindlichkeit unterteilt werden, nämlich dem psychischen Wohlbefinden und der allgemeinen Stimmungslage („psychological well-being and mood“), dem Selbstkonzept („personality and self-concept“), der physischen Stressreagibilität („physiological stress responsivity“) und der Kognition („cognition“). Bei Studien zur Untersuchung zum psychischen Wohlbefinden und der Stimmungslage wurde unterschieden zwischen Studien, die sich im Hinblick auf die Langzeit- und Kurzeiteffekte von körperlichem Training konzentrierten. Nach Plante und Rodin (1990) kann resümierend konstatiert werden, dass regelmäßiges körperliches Training das psychische Wohlbefinden und die allgemeine Stimmungslage positiv beeinflusst und Angst, Depression und Stress reduziert („in summary, empirical research conducted since 1980 suggests that exercise improves mood and well-being and reduces anxiety, depression, and stress“; Plante & Rodin, 1990, S. 9). Diese Befunde werden auch von Martinsen (2002) gestützt, der die Ergebnisse mehrerer Studien zusammenfasst, die sich mit dem Zusammenhang von körperlichen Training, insbesondere aerobes Ausdauertraining und Depression beschäftigen. Martinsen (2002, S. 209) kommt angesichts der Befundlage dieser und auch auf Grundlage eigener Untersuchungen zu dem Schluss, „the results in these studies all point in the same direction: Aerobic exercise is more effective than no treatment, and not significantly different from other forms of treatment, including various forms of psychotherapy.“

Bezüglich des Selbstkonzepts konnte anhand der vorliegenden Literatur festgehalten werden, dass regelmäßiges körperliches Training die Selbsteinschätzung, das Selbstwertgefühl

und das Selbstbewusstsein verbessert. Die Autoren halten gleichzeitig auch fest, dass solche Ergebnisse kritisch zu betrachten sind, da Langzeiteffekte und Kurzeiteffekte, sowie kumulative Effekte von sportlicher Betätigung und deren Zusammenhang zu Persönlichkeitsaspekten und der Selbsteinschätzung bisher noch nicht adäquat untersucht worden sind. In Bezug auf die Stressbewältigung überprüften die Autoren 21 Studien, die nach 1980 veröffentlicht wurden. Es handelte sich dabei zum einen um experimentelle Laboruntersuchungen, bei denen die Stressbewältigung (Reaktion auf oder Regeneration von Stress) unter Berücksichtigung von physiologischen Indikatoren wie Herzfrequenzrate, Blutdruck, Hautleitfähigkeit und Hormonraten (Katecholamine: Adrenalin, Noradrenalin und Dopamin) erfasst wurden. Die Versuchspersonen sollten bei diesen Untersuchungen ein regelmäßiges körperliches Training über einen bestimmten Zeitraum absolvieren oder es erfolgte die Zuweisung einer inaktiven Kontrollgruppe zugeordnet. Zum anderen ordneten Plante und Rodin (1990) weitere Studien der Kategorie Korrelationsdesign zu, bei denen die Testpersonen in fit oder unfit in Bezug auf ihre aerobe Ausdauerleistungsfähigkeit eingeteilt wurden und in einem Labortest mit kognitiven Aufgabestellungen als Stressor konfrontiert. Hierbei wurden ebenfalls Variablen wie Herzfrequenz, Blutdruck, und Hormonwerte als psychophysische Indikatoren für Stress herangezogen. Obwohl eine beträchtliche Anzahl von Studien zu physischer Fitness und dessen Zusammenhang mit Stress vorliegen, hat die Forschung laut Plante und Rodin (1990) keine übereinstimmenden Ergebnisse vorzuweisen, die zeigen, dass körperliche Fitness physische Reaktionen auf Stressoren beeinflusst. Der stichhaltigste Nachweis der betrachteten Studien zeigt, so die Autoren, dass körperliches Training und ein damit einhergehender guter Fitnesszustand die physische Stressregeneration wahrscheinlich beeinflusst („The most impressive evidence favors the notion that exercise might influence physiological stress recovery“; Plante & Rodin, 1990, S. 14).

Methodische Mängel sehen Plante und Rodin (1990) auch bei Studien, die sich mit physischer Fitness und dessen Einfluss auf kognitive Fähigkeiten befassen. Es besteht bei solchen Studien stets auch das Problem von motivationalen Aspekten bei der Durchführung von kognitiven Tests. In der Arbeit von Dustman, Ruhling, Russel, Shearer, Bonekat, Shigeoka, Wood und Bradford (1984) zeigt sich bspw. eine Verbesserung bei neuropsychologischen Tests durch aerobes Fitnessstraining bei älteren Personen, die ein 48-wöchiges Trainingsprogramm absolvierten. Dagegen können Blumenthal und Madden (1988), die die Auswirkung eines 12-wöchigen Fitnessprogramms auf die Gedächtnisleistung untersuchten, keinen Zusammenhang feststellen (Plante & Rodin, 1990). Plante und Rodin (1990) fassen zusammen,

dass die näher betrachteten Studien, trotz methodischer Defizite und Untersuchungsmängel darauf hinweisen, dass körperliches Training das Wohlbefinden und das Selbstkonzept sowie Selbstwertgefühl positiv beeinflussen. Des Weiteren kann angenommen werden, dass Angst und Depressionen durch regelmäßiges Training reduziert werden können und arbeitsrelevante Verhaltensweisen verbessert werden. „In addition, our review suggests that exercise is likely to decrease mild anxiety, depression, and stress, and may improve certain work-relevant behaviors“ (Plante & Rodin, 1990, S. 21).

Gleicherweise fasst Bamberg (2004) zusammen, dass Ergebnisse zum Zusammenhang zwischen Stress und Gesundheit heterogen sind. Nach Zapf und Semmer (2004) haben sich in Untersuchungen Korrelationen um $r = .50$ gezeigt. Demgegenüber liegen aber auch Ergebnisse aus anderen Studien vor, die wesentlich geringere oder nicht signifikante Werte darlegen.

Zu den risikomindernden Maßnahmen gehören nach Bamberg (2007) personen- und verhaltensbezogene Interventionen und die Entwicklung von personenbezogenen Ressourcen, sowie die Unterstützung individueller Stressbewältigung. Hierzu gehören Sport-, Fitness- und Ernährungsprogramme (Bamberg, 2007). Von Busch (2004) konnten in diesem Zusammenhang Effektstärken bis zu $r = .70$ nachgewiesen werden. Hierzu verweisen Frey, Kaminski, Streicher und Niesta (2007, S. 638) auf medizinische Untersuchungen, die zeigen, „dass sich schon moderate körperliche Betätigung begünstigend auf die allgemeine berufliche Leistung auswirkt“ (Pronk, Martinson, Kessler, Beck, Simon & Wang, 2004).

3.3 Stress und Sport

Sportliche Aktivität bzw. körperliches Training kann man in Bezug auf Stressmanagement (siehe Berger, Pargman & Weinberg, 2007) als ein Mittel sehen, das wissenschaftlich nachgewiesen äußerst effektiv ist, um Stresseffekte abzubauen oder zu kompensieren und gleichzeitig andere gesundheitliche Vorteile (siehe z. B. American College of Sports Medicine Position Stand, 1998) mit sich bringt. Hinsichtlich des Stressmanagements kann festgehalten werden, dass durch sportliche Betätigung wie bspw. Ausdauertraining, körperliche Faktoren wie zu hoher Blutdruck oder die Ruheherzfrequenz und stressbedingte Muskelspannungen gesenkt werden können. Es besteht eindeutig ein diesbezüglich spezifischer Zusammenhang zwischen dem, was umgangssprachlich als Körper und Geist angesprochen wird. Auch Un-

ternehmen und Konzerne haben bereits zum Großteil erkannt, nicht nur die Leistungsfähigkeit der Führungskräfte, sondern auch die ihrer Mitarbeiter durch Fitness,- Sport- und Mentalprogramme zu fördern und sicher zu stellen. Nach Berger et al. (2007) wirkt sich physische Fitness in vielen Aspekten positiv auf die Psyche aus, insbesondere auf die mit der sportlichen Betätigung einhergehende Hormonausschüttung und dem damit verbundenen „besserem Gefühl“ und positiver Stimmungslage. Die sog. Stresshormone (Adrenalin und Nor-Adrenalin) sind Indikatoren für den engen funktionalen Zusammenhang zwischen physiologischen und psychischen Faktoren. Konkrete Beispiele für eine durch Sport bedingte positivere Stimmungslage und Stressminderung kann neben einer entspannteren Muskulatur auch eine bessere Ausgeglichenheit oder verminderte Müdigkeit und mehr Vitalität (mehr Energie) sein (Berger, 1994). Adäquate sportliche Betätigung, insbesondere aerobes Ausdauertraining, hat neben dem Abbau von Aggressionen, Wut, Ärger und Feindseligkeit, oder dem Gefühl von Schlappeheit weitere positive körperliche Auswirkungen wie ein gesteigertes Herzschlagvolumen und damit einen geringeren Ruhepuls, sowie einen verminderten systolischen und diastolischen Blutdruck (American College of Sports Medicine, 2006).

Des Weiteren stützen Untersuchungen die Annahme, dass Personen, die regelmäßig Sport treiben eine positivere Einstellung zu ihrem Körper und damit ein besseres Selbstverständnis und Selbstbewusstsein entwickeln als Nichtsporttreibende. Dies ist begründet durch die Möglichkeit einer trainierten Person, sich körperlich anspruchsvollen Herausforderungen zu stellen und sie zu meistern, sowie durch ein ebenfalls geringeres Anstrengungsempfinden und mehr Spaß an körperlichen Aktivitäten zu haben. Nach Hong, Farag, Nelesen, Ziegler und Millis (2004), Plante, Coscarelli und Ford (2001) und der Autorengruppe Spalding, Lyon, Steel und Hatfield (2004) kann konstatiert werden, dass trotz der Komplexität des Zusammenhangs von Ausdauertraining und Stressbewältigung, der Frage nach der richtigen Trainingsintensität, der Berücksichtigung individueller Charaktereigenschaften und dem Grad des Stressors regelmäßiges Training grundsätzlich eine stressmindernde Wirkung hat. Dagegen zeigen Studien wie die von Claytor (1991), dass sowohl körperlich trainierte Männer als auch männliche Personen, die ausschließlich sitzenden Tätigkeiten nachgehen, gleichermaßen mit Herzkreislaufanzeichen- und dem Nervensystem auf Stressoren ansprechen.

Einerseits deuten Studien, wie die von Claytor (1991) darauf hin, dass keine Unterschiede von trainierten und untrainierten Personen vorhanden sind. Andererseits konnte in weiteren Studien gezeigt werden, dass sich bedeutsame Zusammenhänge zwischen gut trainierten Per-

sonen und einer besseren Stressresistenz andeuten (Holmes & Roth, 1985; Hong et al., 2004; Keller & Seraganian, 1984; Sinyor, Golden, Steinert & Seraganian, 1986). In einer spezifischen Situation konnte bei Frauen, die einen psychologischen Test unter Stressbedingungen zu lösen hatten, nachgewiesen werden, dass diejenigen von ihnen, die fitter waren während des Tests eine geringere Herzfrequenz aufwiesen, als die unfitteren Frauen (Holmes & Roth, 1985; Long, 1991). Darüber hinaus stellten Sinyor, Schwartz, Peronnet, Brisson und Seraganian (1983) in einer Studie mit körperlich fitten Männern fest, dass diese während einer Laboruntersuchung in einer Stresssituation höhere Blutwerte an Noradrenalin und Prolaktin hatten, was auf eine starke Stressreaktion hinwies. Es konnte in der gleichen Untersuchung nachgewiesen werden, dass die gut trainierten Männer schneller wieder zum Ausgangspunkt ihres Befindens zurückkehrten als dies bei den weniger trainierten bzw. untrainierten Probanden² der Fall war. In einer Studie von Hong et al. (2004) wurden insgesamt 48 Personen, davon die Hälfte der Versuchsgruppe trainierte und die andere Hälfte untrainierte Personen, sowohl einem psychologischen Test (psychischer Stressor) in Form einer 15-minütigen Sprachaufgabe, als auch einem physischen Test in Form eines 15-18 Minuten (Min.) andauernden Stufentests mit dem Fahrradergometer bei einer maximalen Herzfrequenz von 75 Prozent (physischer Stressor) unterzogen. Bei dem psychischen Test konnte bei den fitteren Versuchspersonen eine höhere Konzentration an Lymphozyten, insbesondere der NK-Zellen (natürliche Killerzellen), die zu dem adaptiven Immunsystem gehören, damit für die spezifische Abwehr und für die Erkennung von Fremdstoffen (Viren und Bakterien) und deren Entfernung zuständig sind, als bei den unfitten Testpersonen nachgewiesen werden. Hingegen wurden bei dem physischen Test keine Unterschiede von Hormonwerten im Blut festgestellt. Diese Befunde lassen darauf schließen, dass physische Fitness die Immunantwort auf psychische Stressoren beeinflusst, jedoch nicht bezüglich physischer Stressoren.

An dieser Stelle soll kritisch angemerkt werden, dass sich bei der Studie von Hong et al. (2004) die Probanden hinsichtlich ihres körperlichen Fitnesszustands selbst einschätzten und keine objektive Beurteilung des Fitnesslevels der Untersuchung voranging. Relativ unbeachtet sind die Dauer der Stressreaktion und diesbezügliche Unterschiede zwischen trainierten und untrainierten Personen. Aufgrund kardiorespiratorischer Unterschiede (die Herztätigkeit und Atmung betreffend) bei fitten und unfitten Personen und der kürzeren Regenerationszeiten bei fitten Individuen, kann dies auch für Stressreaktionen und die Rückkehr zu einem Zustand der

² Der besseren Lesbarkeit halber werden nur dann beide Geschlechtsformen aufgeführt, wenn damit zusätzliche Informationen vermittelt werden.

Ausgeglichenheit angenommen werden. Unter Berücksichtigung negativer Auswirkungen auf gesundheitliche Aspekte ist eine möglichst schnelle Rückkehr in den Zustand der Ausgeglichenheit nach einer Stresssituation erstrebenswert. Eine kürzere Regenerationszeit in Form einer zügigeren Rückkehr zum Ruhepuls indiziert eine kürzere Stressreaktion bei trainierten Personen (Sinyor et al., 1983). Eine schnellere Regenerationsfähigkeit von ausdauertrainierten Personen indiziert eine effektivere Bewältigung und bessere Belastbarkeit bei emotionalem Stress als dies bei weniger fitten Personen der Fall ist (siehe Berger et al., 2007, S. 158).

In einer Studie über den Zusammenhang der aeroben Ausdauer und der Dauer der Stress- bzw. Belastungsreaktion untersuchten Keller und Seraganian (1984) Männer und Frauen, die nach dem Zufallsprinzip in eine Gruppe eingeteilt wurden, die vier Mal in der Woche ein 30-minütiges Ausdauerprogramm für insgesamt zehn Wochen vollzog und in zwei weitere Gruppen, die sitzenden Aktivitäten wie meditative Yogaübungen und Zuhören von Musik im gleichem Zeitrahmen durchführten. Die Gruppe, die über zehn Wochen regelmäßig ein meditatives Yogaprogramm ausübte, wurde als Kontrollgruppe eingesetzt. Dies geschah vor dem Hintergrund der weit verbreiteten Annahme, dass Meditationsübungen stressmindernd wirken. In den Wochen zwei, sechs und zehn wurden alle Pbn in einem Labor stressvollen Aufgaben unterzogen. Es zeigte sich eine schnellere elektrodermale Regeneration, d. h. schnellere Erholung gemessen durch den Hautwiderstand (GSR = galvanic skin response), bei der Ausdauergruppe im Gegensatz zu den anderen beiden Gruppen, die die sitzenden Aktivitäten ausübten. Somit konnte Ausdauertraining als eine Stressmanagement-Technik ausgemacht werden und die Annahme aufgestellt werden, dass aerob trainierte Personen eine effizientere emotionale Stressbewältigung als Personen mit einem geringeren aeroben Fitnesslevel haben.

Sinyor et al. (1983) konnten ebenfalls einen Unterschied zwischen hochtrainierten Männern im Alter von 20 bis 30 Jahren und weniger fitten Personen gleichen Alters in deren Stressbewältigungsprofil feststellen. Dieser drückte sich wie bei Hong et al. (2004) unter anderem in einer schnelleren Reduktion der Herzfrequenz nach der Stresssituation sowie einem verminderten Angstzustand aus. In Anlehnung an Berger et al. (2007) wurden in einer Untersuchung von Long (1991) der Blutdruck und die Blutwerte an Adrenalin, Noradrenalin und Kortisol gemessen. Frauen im Alter von 18 bis 48 Jahren wurden auf Grundlage der Ergebnisse eines submaximalen Fahrradergometertests in zwei Gruppen eingeteilt (hohe Fitness und geringe Fitness). Nach dem Ausdauerstest mussten alle Pbn einen Rechentest und ein „Elektrokardiogramm Quiz“ innerhalb von zehn Minuten absolvieren. Die Stressreaktion anhand der oben

genannten Faktoren wurde jeweils nach dem Ausdauerstest sowie nach den insgesamt 10-minütigen mentalen Belastungstests und in den Regenerationsphasen gemessen. Bei den physisch fitteren Frauen erholte sich die Herzfrequenzrate schneller, außerdem sank der Noradrenalinwert (Stresshormon bzw. Indikator für Stress) im Blut rapider ab als bei den Untrainierten.

In Anlehnung an Dienstbier (1991) reduziert physische Fitness und regelmäßiges und moderates Sporttreiben die Gefahr, in stressvollen Lebensphasen Erkrankungen zu erleiden. Vor allem in den Phasen, die als besonders stressvoll erscheinen, wie beispielsweise Prüfungsphasen in der Schule oder der Universität, mindert Sport die Stressreaktion und durch die vermehrte Ausschüttung an NK-Zellen wird das Immunsystem gestärkt. Im Gegensatz dazu können starke sportliche Aktivitäten und Belastungen die über eine moderate Belastung hinausgehen das Immunsystem in einer stressvollen Phase noch stärker beeinträchtigen und schwächen. Regelmäßige moderate aerobe Ausdauerbelastungen sind zeitweilige Stressoren, die nicht nur zu einer positiven Anpassung der spezifischen Ausdauer führen, sondern auch zu einer erhöhten Belastbarkeit in anderen Bereichen beitragen (vgl. Berger et al., 2007). Außerdem gewährleistet eine bessere aerobe Fitness auch eine bessere Bereitstellung von Katecholaminen (Adrenalin und Noradrenalin), die zu einer besseren Stressbewältigung beitragen. Durch das vermehrte Aufkommen von Adrenalin und Noradrenalin im Körper durch Ausdauertraining kommt es zu einer längerfristigen Leistungsbereitschaft und dies macht den Körper durch Aufbau und Bereitstellung neuer Reserven leistungs- und widerstandsfähiger und führt zu einer verbesserten Stressresistenz, Belastbarkeit, Stressbewältigung, einem stärkeren Immunsystem und einer besseren emotionalen Stabilität (vgl. Dienstbier, 1991).

In einer Untersuchung von Reiner, Niermann, Krapf und Woll (2013) wurde geprüft, inwiefern Sport und körperlicher Aktivität einen Puffereffekt für bestimmte negative gesundheitliche Auswirkungen von Stress (Herzbeschwerden, Gliederschmerzen, Magenschmerzen, Erschöpfung, Beschwerdedruck und physische sowie psychische Beschwerden) haben. Im ersten Schritt erfassten Reiner et al. (2013) im Rahmen einer kommunalen Gesundheitsuntersuchung von insgesamt 210 Männern und 243 Frauen im Alter zwischen 25 und 79 Jahren ($M = 53,56$ Jahre; $SD = 11,31$ Jahre) das Gesundheits- und Sportverhalten, die körperliche Aktivität, die Beschwerdewahrnehmung, das Stressempfinden und die Konstrukte Selbstwirksamkeit und Selbstbestimmung mit Hilfe von spezifischen Fragebogeninstrumenten. Die Angaben zur sportlichen Aktivität (Sportart, Dauer und Häufigkeit der Trainingseinheiten,

Summe der ausgeübten Sportarten) und zur allgemeinen körperlichen Aktivität wie z. B. Gartenarbeit, regelmäßigen Geh- und Radstrecken, etc. wurden durch ein bestimmtes Verfahren operationalisiert, indem die sportlichen und allgemeinen körperlich angegebenen Tätigkeiten in einem Energieindex zu einem Energieverbrauch (kcal/Woche) umgerechnet worden sind. Der wahrgenommene Stress ist anhand des Summenscores (modifiziert auf neun anstatt zehn Items) der Perceived Stress Scale (PSS) von Cohen, Karmack und Mermelstein (1983) gemessen worden. Die wahrgenommenen Beschwerden wurden mit dem Gießener Beschwerdebogen (GGB-24) nach Brähler, Hintz und Scheer (2008) erfasst. Darauffolgend führten die Autoren mit den z-transformierten Werten Regressionsanalysen durch, wobei auch Geschlechts- und Altersunterschiede berücksichtigt wurden. In einem ersten Schritt prüften Reiner et al. (2013) die Haupteffekte der Variablen Sport, körperliche Aktivität und Stress im Hinblick auf die erlebten Beschwerden (abhängige Variable).

Im Folgenden nahmen die Autoren jeweils den Interaktionsterm von Stress*Sport und Stress*körperliche Aktivität mit in die statistische Berechnungen hinzu, um den postulierten Puffereffekt zu untersuchen. In dem Zusammenhangsmodell von Reiner et al. (2013) stellt Stress den Prädiktor und die Beschwerdewahrnehmung das Kriterium dar, Sport und körperliche Aktivität sind die Moderatorvariablen. Die Studie von Reiner et al (2013) zeigen folgende Ergebnisse: Grundsätzlich geht ein erhöhtes Maß an erlebten Stress mit mehr wahrgenommenen Beschwerden, sowohl in psychischer als auch in physischer Hinsicht einher. Personen mit erhöhtem Stresserleben weisen erwartungskonform ein höheres Ausmaß an Beschwerden auf ($r = .37; p < .01$). Ein signifikant negativer Zusammenhang ($r = -.16; p < .01$) zeigt sich ebenfalls, wie von den Autoren erwartet, zwischen dem Energieverbrauch durch sportliche Betätigungen und den berichteten Beschwerden. Dagegen zeigt sich bei der körperlichen Aktivität und den wahrgenommenen Beschwerden kein Zusammenhang. Die Befunde von Reiner et al. (2013) deuten in die Richtung, dass nicht nur der Energieverbrauch alleine maßgebend für eine Minderung der Beschwerdewahrnehmung ist, sondern dass v. a. die Art des Energieverbrauchs (Sport vs. allgemeine körperliche Aktivität) im Hinblick auf die Wahrnehmung der stressassoziierten Beschwerden eine entscheidende Rolle spielt. Anhand der signifikanten ($p \leq .01$) Befunde der Regressionsanalysen, konstatieren die Autoren, dass insbesondere die Moderatorvariable Sport einen Einfluss auf den Zusammenhang zwischen erlebten Stress und wahrgenommenen Beschwerden (psychisch und physisch) nehmen kann bzw. einen Puffereffekt bewirkt, jedoch so nicht für alle Formen von Beschwerden im gleichen Maße wirksam ist. Bei den Beschwerdearten Magenbeschwerden und Herzbeschwerden blieben hingegen

signifikante Interaktionsterme aus. Puffereffekte von alltäglichen körperlichen Betätigungen konnten ebenfalls nicht nachgewiesen werden. Für ihre Befunde liefern Reiner et al. (2013) u. a. folgenden Erklärungsansatz, der insbesondere für den thematischen Kontext dieser Arbeit als interessant erachtet werden kann:

„durch die [sportliche] Aktivität bzw. die damit einhergehende körperliche Fitness, [könnten] körperliche Ressourcen gestärkt werden, was beispielsweise positiv auf stressbedingte Rücken- oder Nackenschmerzen wirken dürfte. Zum anderen könnte die mit sportlicher Aktivität verbrachte Zeit als eine mentale Auszeit vom Arbeitsalltag und Verpflichtungen betrachtet werden und so psychische Ressourcen schützen und/oder stärken (Sonntag, 2012). Hierüber lassen sich auch positive Effekte auf psychische Beschwerden erklären. Dieser Effekt zeigt sich in der vorliegenden Studie in Bezug auf die psychische Beschwerdeart Erschöpfung, hier kann sportliche Aktivität als Puffer wirken.“ (Reiner et al., 2013, S. 272)

Auf Grundlage der Befunde und deren Interpretation von Reiner et al. (2013) ist somit dem Faktor Sport und darüber dem Fitnesszustand eine puffernde Wirkung in Bezug auf psychische Erschöpfungszustände zuzuschreiben. Das Stressempfinden und die damit einhergehenden wahrgenommenen psychischen Beschwerden, wie der Zustand des „erschöpft seins“, kann demnach durch sportliche Betätigung und die damit verbesserte Fitness im Sinne einer höheren Belastbarkeit „abgepuffert“ werden. Abgesehen von den Befunden weisen die Autoren kritisch auf die Erfassung der Daten hin. Insbesondere die Operationalisierung der Angaben zur sportlichen Aktivität und alltäglichen körperlichen Aktivität wirft das Problem auf, dass die Befragten eine tägliche körperliche Betätigung, wie bspw. Radfahren zur Arbeitsstelle als sportliche Betätigung deklarierten, anstatt es zur körperlichen Aktivität zuzuordnen. Die Autoren merken an, dass eine andere Herangehensweise hinsichtlich der Messmethodik, wie z. B. ein Fitnesstest als objektive Untersuchungsmethode zur Erfassung der physischen Leistungsfähigkeit, eine gute bzw. bessere Alternative darstellen würde, so wie es in dieser Arbeit umgesetzt wird.

Wenn so der Eindruck vermittelt und explizit ausgeführt wird, dass Sport einen Puffereffekt zwischen Stresserleben und Beschwerden haben kann, bleibt anzumerken, dass solche ‚Kann-Aussagen‘ wissenschaftlich wenig aufschlussreich und tragfähig sind. Die Zusammenhänge der berichteten Befunde (siehe die Regressionsanalysen) erweisen sich teilweise als signifikant, bleiben jedoch in ihrer durchgehend geringen Ausprägung kaum entsprechend interpretiert und diskutiert. Es fehlt auch ein konzeptueller Rahmen zur differenzierten Ableitung und Begründung von Hypothesen und zur angemessenen Interpretation der Befunde. Dies wird

den Autoren am Ende (siehe „Ausblick“, S. 274) wohl auch selbst deutlich, wenn sie erkennen, dass „die Betrachtung von Stress und Sport in nachfolgenden Studien auf einer fundierten theoretischen Grundlage beruhen [sollte]“. Weiterhin bleibt kritisch festzustellen, dass bei sorgfältiger Betrachtung der Befunde Unstimmigkeiten in der Darstellung und Interpretation auffallen. So stimmen Angaben in Tabellen und im Text nicht überein (siehe z. B. Tabelle 2 zu Beschwerdedruck und Aktivität: -0.14^{**} , im Text auf Seite 270: -0.16) und es wird falsch interpretiert (vgl. Tabelle 1 und Text auf Seite 268), wenn es um den Vergleich von psychischen (und nicht, wie geschrieben, physischen) Beschwerden/Erschöpfung bei älteren und jüngeren Männern geht. Insgesamt wird deutlich, dass die Bearbeitung dieser Fragestellung eine anspruchsvolle Herausforderung darstellt. Dies bezieht sich auf die notwendige theoretische Auseinandersetzung und Grundlegung wie auf die Methodik bis hin zur statistischen Analyse der Daten und deren differenzierte und detaillierte Betrachtung und Auslegung.

Die aufgeführten Studien und Erläuterungen verdeutlichen zusammenfassend den engen psycho-physischen Zusammenhang und funktionalen Verbund speziell mit Bezug auf Stress-situationen. Auch wenn sich aufgrund der unterschiedlichen Untersuchungsanliegen und -anlagen und verschiedenen Methoden die Befundlage als komplex darstellt, ist eine Tendenz deutlich erkennbar: Für körperlich fittere Personen zeigt sich eine günstigere Konstellation für das Stressmanagement. Inwieweit dies auf die Auswirkungen sportlicher Aktivitäten und gesundheitsförderliches Verhalten oder auf konstitutionelle Faktoren zurückzuführen ist, oder auf eine diesbezügliche Wechselwirkung bedarf weiterer Klärung. Im Rahmen dieser Arbeit interessieren dazu speziell die Persönlichkeitsfaktoren Neurotizismus bzw. die emotionale Stabilität und (mentale) Belastbarkeit.

3.4 Stress, Emotionale Stabilität und Neurotizismus

Neurotizismus bzw. die als emotionale Labilität bezeichnete starke Ausprägung des einen Pols des bipolar (emotionale Stabilität vs. emotionale Labilität) angelegten Merkmals, stellt ein Persönlichkeitsmerkmal dar, das gekennzeichnet ist durch Ängstlichkeit, häufige körperliche Beschwerden, Stressanfälligkeit, ein öfter eintretendes Gefühl entmutigt und/oder überfordert zu sein, Depressivität, geringem Selbstwert, Verletzlichkeit, Verlegenheit, Reizbarkeit, Impulsivität und Nervosität. Der Gegenpol Emotionale Stabilität ist ein Persönlichkeitskonstrukt, das demgegenüber durch die Fähigkeit geprägt ist, Misserfolge und Rückschläge rasch zu überwinden, eigene negative Gefühle kontrollieren zu können und durch eine allgemeine Ausgeglichenheit, ein hohes Maß an Gelassenheit und die Fähigkeit, sich nach persönlichen Niederlagen erneut motivieren zu können (vgl. Bamberg & Keller, 2013; Rammsayer & Weber, 2010; Hossiep & Paschen, 1998). Bamberg und Keller (2013) verweisen in einem Beitrag zur Stressbewältigung auf eine Reihe von Untersuchungen zu organisationalem Verhalten, nach denen Zusammenhänge zwischen emotionaler Stabilität und Leistung und zwischen emotionaler Stabilität und Arbeitszufriedenheit nachgewiesen werden konnten (Judge, Thoresen, Bono & Patton, 2001; Judge & Bono, 2009; Bruk-Lee, Khoury, Nixon, Goh & Spector, 2009). Darüber hinaus beziehen sich die Autoren auf Metanalysen von Judge, Klinger, Simon & Yang (2008), die sich mit den Big Five-Persönlichkeitsfaktoren befassen und das Merkmal Neurotizismus inkludieren. Hier nach haben die Variablen emotionale Stabilität bzw. emotionale Labilität einen Bezug „zu nahezu allen Indikatoren organisationalen Verhaltens“, u. a. zur Arbeitsleistung und zur Arbeitsmotivation (Bamberg & Keller, 2013, S. 281).

Der Begriff Emotionale Stabilität umfasst im Bezugsrahmen der berufsbezogenen Persönlichkeitsforschung, wie bereits erwähnt, die Eigenschaften, Rückschläge oder Niederlagen schnell verkraften zu können, auch in Stresssituationen gelassen zu bleiben, weniger ängstlich zu sein etc. Was allerdings unter „Emotionen“ genau verstanden wird, soll zunächst aufgrund der zentralen Bedeutung des Wortes im Kontext des Begriffs Emotionale Stabilität differenzierter beleuchtet werden.

Im Alltagsverständnis glaubt jeder zu wissen, was mit „Emotion“ gemeint ist (Schlattmann & Hackfort, 1988). Im Allgemeinen werden Emotionen auch als „Gefühle“ verstanden, die sich in einem bestimmten Sozialverhalten, durch physiologische Veränderungen und kognitive Prozesse widerspiegeln. Aus wissenschaftlicher Sicht wurden verschiedene Definitionsver-

suche aus unterschiedlichen theoretischen Perspektiven angeboten, allerdings konnte keine allgemein akzeptierte oder gar verbindliche Definition geleistet werden. Immerhin lassen sich aber Gemeinsamkeiten in den Konzeptualisierungen erkennen, die im Sinne eines kleinsten gemeinsamen Nenners wie folgt festzuhalten sind: Für Emotionen sind drei Komponenten grundlegend, nämlich (1) die kognitive, (2) die physiologische und (3) die Verhaltenskomponente (siehe u. a. Izard, 1981, S. 20; Schlattmann & Hackfort 1988, S. 4; Lang, 1993). In Bezug auf die Auswirkungen von Emotionen auf den Menschen beschreibt Izard (1981, S. 26) auch einen psycho-physischen Zusammenhang mit Hinweis auf Simonov (1975) derart, dass mit Emotionen Veränderungen

„in der elektrischen Aktivität des Gehirns, im Kreislauf- und im Atmungssystem [auftreten]. Bei starkem Zorn oder großer Furcht kann die Herzfrequenz sogar um 40-60 Schläge pro Minute ansteigen (Ruslova u. a., 1975). Diese dramatischen Veränderungen in den Körperfunktionen bei intensiven Emotionen deuten darauf hin, daß praktisch alle neurophysiologischen Systeme und Subsysteme des Körpers mehr oder weniger stark an emotionalen Zuständen beteiligt sind.“

Auch Alpers, Mühlberger und Pauli (2009, S. 413) beschreiben physiologische Komponenten von Emotionen, wie kardiovaskuläre Aktivitäten (z. B. Herzschlagrate; Blutdruck), elektrodermale Aktivitäten (Hautleitfähigkeit), Respiratorischer Aktivitäten (Atemfrequenz; Hyperventilation) und somatische Aktivitäten, wie z. B. elektromuskuläre Aktivitäten, die häufig als Parameter in der psychophysiologischen Forschung herangezogen werden.

Die Kopplung von Emotionen und körperlichen Reaktionen bzw. die von Izard (1981) beschriebene Auswirkung von Emotionen auf physische Veränderungen führt im Umkehrschluss zu der Überlegung, inwiefern Individuen mit einem leistungsfähigerem bzw. belastbarerem physischem Vermögen, insbesondere der Ausdauerleistungsfähigkeit, eine größere Emotionale Stabilität besitzen. Mit Emotionaler Stabilität ist in diesem Sinne die Fähigkeit zur Kontrolle der eigenen Emotionen bzw. der emotionalen Reaktionen gemeint. Dies bezieht sich also auf eine allgemeine Emotionalität in Abhebung und Abgrenzung zu einzelnen Emotionen, für die er das Konzept differentieller bzw. fundamentaler Emotionen postuliert.

Nach Izard (1981) lassen sich zehn „fundamentale“ Emotionen differenzieren: Ärger, Ekel, Freude, Furcht, Interesse; Scham, Schuld, Traurigkeit, Überraschung und Verachtung. Plutchik (1980) unterscheidet in seinem psychoevolutionären Ansatz Primär- und Sekundäremotionen, die sekundären sind Kombinationen/Zusammensetzungen aus den primären Emotio-

nen: Ärger, Aufnahmebereitschaft, Ekel, Erwartung, Freude, Furcht, Traurigkeit und Überraschung. Schlattmann und Hackfort (1988) fassen zusammen, dass

„in einer Untersuchung zur Struktur des Wortfeldes ‚Emotionen‘ [von Schmidt-Atzert und Ströhm (1983) insgesamt 14 Emotionskategorien identifiziert werden konnten.] (. . .) Studentische Versuchspersonen hatten in dieser Studie 56 Emotionskategorien in Stapel von jeweils ähnlichen Emotionen einzusortieren. (. . .) Betrachtet man diese Ergebnisse, dann entsteht der Eindruck, als gäbe es eindeutig mehr negative als positive Emotionen. D.h. auch, die Wahrscheinlichkeit für das Erleben von negativen Emotionen ist höher. Sinnvoll scheint es aber auch, der Annahme nachzugehen, daß positive emotionale Erlebnisse eine geringere Differenzierbarkeit aufweisen“ (Schlattmann, 1988, S. 6).

Tabelle 2

Emotionskategorien (nach Schmidt-Atzert & Ströhm, 1983, S. 136).

Bezeichnung des Clusters	Zum Cluster gehörende Emotionen
1) Abneigung	Abneigung, Verachtung, Ekel, Widerwille
2) Ärger	Ärger, Aggressionslust, Groll, Wut, Hass, Trotz
3) Neid	Eifersucht, Neid, Misstrauen, Schadenfreude
4) Frustration	Frustration, Verstimmtheit
5) Leere	Langeweile, Leere, Unlust
6) Angst	Angst, Entsetzen, Panik, Schreck, Verzweiflung
7) Sexuelle Erregung	Erregung, Leidenschaft, Spannung
8) Unruhe	Ungeduld, Unruhe
9) Traurigkeit	Heimweh, Sehnsucht, Kummer, Traurigkeit, Sorge
10) Scham	Reue, Schuld, Scham, Verlegenheit
11) Freude	Begeisterung, Lust, Freude, Glück, Zufriedenheit, Erleichterung
12) Stolz	Stolz, Triumphgefühl
13) Zuneigung	Dankbarkeit, Verehrung, Wohlwollen, Zuneigung, Zutrauen, Mitgefühl, Rührung
14) Überraschung	Überraschung, Erstaunen, Verwunderung

Weiterhin lassen sich Emotionen abgrenzen in emotionale Episoden, wie dies in Tabelle 2 repräsentiert ist (z. B. Angst) und in emotionale Dispositionen (z. B. Ängstlichkeit), die sich in ihrer Ausprägung unterscheiden, d. h. hinsichtlich der Qualität differenzieren und in der Intensität unterschiedlich stark auftreten. Der Emotionsbegriff umfasst, wie man der Tabelle entnehmen kann, über 50 emotionale Phänomene. Diese sind für die berufsbezogene Emotionale Stabilität bzw. für die Persönlichkeitsdimension Neurotizismus nicht alle relevant und finden auch in den Skalen zu den jeweiligen Persönlichkeitsfaktoren nicht alle Beachtung. Persönlichkeitsdimensionen, die auf Emotionen, speziell die Emotionale Stabilität und Labilität abzielen, berücksichtigen spezifische Emotionen. Bei der Emotionalen Stabilität handelt es sich um ein Konstrukt, welches zur Aufschlüsselung eines bestimmten Persönlichkeitsmerkmals konstruiert wurde. Der Begriff „Emotional“ verweist auf eine Spezifikation dieser Persönlichkeitseigenschaft, womit nicht das gesamte Emotionspektrum abgedeckt wird. Welche Kategorien und Facetten dieses Konstrukt beinhaltet, ist im Kapitel 2.5 bereits näher beleuchtet worden.

In diesem Kapitel wurde zum einen auf die Emotionale Stabilität bzw. auf das komplementäre Merkmal Emotionale Labilität in Verbindung mit Sport eingegangen und in diesem Zusammenhang auch kurz auf die einschlägige Emotionsforschung Bezug genommen. Aufgrund des allgemein vermuteten Zusammenhangs von Sport und Stress (siehe hierzu die o. a. Literatur), sowie Stress und psychischer Belastbarkeit bzw. emotionaler Stabilität/Neurotizismus sind hier relevante Studien und Beiträge zur Stressforschung gesichtet worden. Da auch Aspekte der psychischen Gesundheit in Verbindung mit sportlicher Aktivität für diese Arbeit von Bedeutung sind, wurden dazu Untersuchungen analysiert, die sich mit diesen Zusammenhängen befassen.

Je breiter man die Suche nach Studien zum Zusammenhang zwischen Sport, Persönlichkeit, Stress und psychischer Gesundheit anlegt, desto heterogener stellt sich die Befundlage dar. Schränkt man dieses Forschungsspektrum ein und spezifiziert die Suche auf Zusammenhänge zwischen sportlicher Aktivität bzw. Leistungsfähigkeit und Neurotizismus bzw. emotionaler Stabilität, so sind überwiegend bestätigende Ergebnisse zu finden (siehe u. a. Cattell, 1946, 1993; Egloff & Gruhn, 1998; Eysenck et al., 1982; Kirkcaldy, 1982; McKelvie et al., 2002; Piedmont et al., 1997; Sack, 1975, 1982; Schwinger et al., 2013; Vanek & Hosek, 1977).

4 EMPIRISCH-METHODISCHER ANSATZ

In diesem Kapitel werden die methodischen Grundlagen und Orientierungen für die nachfolgenden empirischen Studien dargestellt. Die spezifischen methodischen Angaben zum Untersuchungsdesign und zum Untersuchungsablauf der einzelnen Studien sind unter den jeweiligen Untersuchungsberichten (siehe Kapitel 5.1, 5.1, 5.3) aufgeführt. Entsprechend der Problem- und Zielstellung geht es in dieser Arbeit um Verfahrensentwicklungen und darauf aufbauend um die Überprüfung ausgewählter berufsbezogener psycho-physischer Zusammenhänge bzw. im Hinblick auf körperlich leistungsdifferenzierte Gruppenunterschiede. Dazu ergeben sich folgende übergeordnete Fragestellungen und Hypothesen:

- (1) Lässt sich ein Verfahren zur Messung der Ausdauerleistungsfähigkeit konzipieren, welches für die Personalgewinnung (speziell in der Bundeswehr) besser (insbesondere hinsichtlich der Justiziabilität) geeignet ist als der bisher verwendete 1000-Meter-Lauf?
- (2) Lassen sich für die Erfassung der psychischen Belastbarkeit zwei Skalen eines gängigen berufsbezogenen Persönlichkeitsinventars kombinieren und zur zeiteffizienten Erfassung dieses Merkmals ökonomisieren?
- (3) Lassen sich auf der Grundlage der zu entwickelnden Verfahren und den daraus gewonnenen Parametern die hier speziell interessierenden psycho-physischen Zusammenhänge einerseits und gruppenbezogene Unterschiede andererseits nachweisen?

Mit Bezug auf die theoretischen Erläuterungen ergeben sich dazu folgende Arbeitshypothesen:

(H1) Es bestehen Zusammenhänge zwischen der Ausdauerleistungsfähigkeit als physische Belastbarkeitskomponente und der psychischen Belastbarkeit.

Hierzu werden Korrelationsanalysen durchgeführt.

(H2) Es lässt sich ein Einfluss der physischen Belastbarkeit auf die psychische Belastbarkeit anhand spezifischer Indikatoren nachweisen.

Hierzu werden Regressionsanalysen durchgeführt.

(H3) Es bestehen Unterschiede zwischen Gruppen mit unterschiedlichen physischen Leistungsniveaus im Hinblick auf die psychische Belastbarkeit.

Hierzu werden Varianzanalysen durchgeführt.

Im Rahmen der Verfahrensentwicklungen zur Erfassung sowohl der physischen als auch der psychischen Belastbarkeit wurden bei verschiedenen Gelegenheiten Daten zu beiden Aspekten der Belastbarkeit erhoben, welche dann auch zur Analyse von psycho-physischen Zusammenhängen und Gruppenunterschieden herangezogen wurden. In den Studien zur Verfahrensentwicklung wurden insbesondere für die Validierung weitere Methoden eingesetzt. Die verschiedenen relevanten Datensätze gehen in die Analysen zur psycho-physischen Belastbarkeit ein. Eine Übersicht zu den angelegten Studien, den eingesetzten Verfahren sowie den daraus gewonnenen Indikatoren zur psycho-physischen Belastbarkeit bietet die Abbildung 29. Die einzelnen Studien entsprechen den Gelegenheiten der verschiedenen Datenerhebungen und gehen in die Kapitel zur Beantwortung der o. a. Fragestellungen ein. Dazu werden die Daten aus den verschiedenen Studien für die jeweilige Analyse zusammengefasst. Der Bericht weicht dementsprechend bis auf die Darstellung der Laborstudie (Kapitel 5.1.2) diesbezüglich vom klassischen Vorgehen für separat durchgeführte Studien ab.

Bevor das methodische Vorgehen dargestellt wird, soll zur Orientierung vorab (Kapitel 4.1) der Übergang von der Konzeptualisierung (Kapitel 2.3) zur Operationalisierung als Ausgangspunkt für die daran anschließenden Studien (Kapitel 5) erläutert werden. Dann wird die methodische Herangehensweise zur Konzeption des Verfahrens zur Erfassung der physischen Belastbarkeitskomponente (Kapitel 4.2) beschrieben. Anschließend wird das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung vorgestellt (Kapitel 4.3.1). Dieses Instrument beinhaltet die beiden Skalen, welche die Grundlage des Verfahrens zur Erfassung der psychischen Belastbarkeitskomponente darstellen. Es bildet zudem die Grundlage für den Ansatz der Verfahrensökonomisierung, auf die in Abschnitt 4.3.2 eingegangen wird.

4.1 Methodische Ausgangspunkte für die Verfahrensentwicklungen

Das Forschungsvorhaben bezieht sich zunächst auf Verfahrensentwicklungen zur Erfassung einerseits der physischen Belastbarkeitskomponente und andererseits der psychischen Belastbarkeitskomponente, um darüber das entwickelte Konzept zur psycho-physischen Belastbarkeit zu operationalisieren. Darauf folgend lässt sich dann mit diesen Verfahren die hier herausgestellte Fragestellung eines psycho-physischen Zusammenhangs, spezifiziert auf die berufsbezogene Belastbarkeit, angehen. Die dazu o.a. allgemeine Hypothese wird weiter unten in dem Kapitel spezifiziert.

Für die Operationalisierung des vorgestellten psycho-physischen Belastbarkeitskonzepts (Abbildung 11) ist auf die Abgrenzung folgender wichtiger Begrifflichkeiten hinzuweisen: Als Parameter sind Kennwerte bezeichnet, z. B. kennzeichnende Werte für eine Stichprobe (etwa Mittelwert und Standardabweichung) oder auch ein Funktionssystem (Herz-Kreislauf-Prozesse). Indikatoren sind in dieser Konzeptualisierung als Anzeichen für einen Zustand zu sehen. Dies kann bspw. eine bestimmte Pulsfrequenz als Indikator für die Herz-Kreislauf-Aktivität oder eine Belastung/Beanspruchung sein, sowie in diesem Fall die maximale Sauerstoffaufnahme (VO_{2max}) als Indikator für die allgemeine Ausdauerleistungsfähigkeit gesehen wird. Die physische Belastbarkeit wird dabei anhand von objektiven Kriterien erfasst. Durch die enge Verbindung dieser Arbeit zu einem Projekt, in dem es um die Entwicklung eines alternativen Verfahrens zum 1000-Meter-Lauf (1000-m-Lauf) im Rahmen des BFT geht, ergab sich die Aufgabe, die physische Konstitution bzw. die körperliche Leistungsfähigkeit durch einen neu konstruierten Ausdauerstest zu erfassen. Die Ausdauer, als eine der motorischen Hauptbeanspruchungsformen, stellt somit die Komponente (Teilkonstrukt) der körperlichen Leistungsfähigkeit im Zusammenhang der Konzeptualisierung psycho-physischer Belastbarkeit dar. Die Ausdauerleistungsfähigkeit wird über spezifische Testverfahren und im Rahmen dieser Arbeit über einen eigens entwickelten 3000-Meter-Fahrradergometertest (3000-m-FET) operationalisiert. Diesen galt es daraufhin zu überprüfen, inwiefern er sich als aussagekräftiger und valider Test zur Überprüfung der Ausdauerleistungsfähigkeit erweist (siehe dazu Kapitel 5.1.2).

Die dabei ermittelte Zeit, die durchschnittliche Leistung in Watt, sowie die absoluten und relativen VO_{2max} -Werte stellen die kennzeichnenden Parameter der physischen Belastbarkeit dar. Die zusätzlich erhobenen Parameter, wie die maximalen Herzfrequenzen, die Respiratori-

scher Quotienten (RQ), die Laktat-Werte sowie die Werte der Borg-Skala, dienen als Ausbelastungskriterien. Die subjektiv eingeschätzten Merkmale „Emotionale Stabilität“ und „Belastbarkeit“ machen in diesem Rahmenkonzept die beiden Komponenten der psychischen Belastbarkeit aus. Weiterhin sind die Persönlichkeitsdimensionen Neurotizismus und Selbstwirksamkeit Konstrukte, welche die psychische Komponente der Belastbarkeit repräsentieren. Diese Konstrukte zur Kennzeichnung der herangezogenen Persönlichkeitseigenschaften werden über bestimmte Skalen aus ausgewählten Persönlichkeitsinventaren operationalisiert und darüber hinaus untereinander auf Zusammenhänge hin untersucht, um so die divergente oder konvergente Validität der Skalen zu überprüfen. Aus den für diese Arbeit wesentlichen Skalen, Emotionale Stabilität und Belastbarkeit, wurde eine eigens entwickelte Kurzsкала eingesetzt, die zur Ökonomisierung des Selbsteinschätzungsverfahrens beitragen soll (siehe Kapitel 4.3.2). Die Itemwerte bzw. die Mittelwerte der einzelnen Itemscores der Persönlichkeitsskalen werden als Kennwerte für die psychische Belastbarkeit herangezogen. Die durch die Ausdauertests gewonnenen Daten werden mit den Mittelwerten der Itemscores der Persönlichkeitsskalen korreliert und auf Zusammenhänge hin untersucht. Infolgedessen kommt es zu einer Kopplung der subjektiv eingeschätzten Kennwerte der Persönlichkeitsskalen mit den objektiv ermittelten Leistungsdaten durch die Ausdauertests, die zusammen die Ausprägung der psycho-physischen Belastbarkeit charakterisieren.

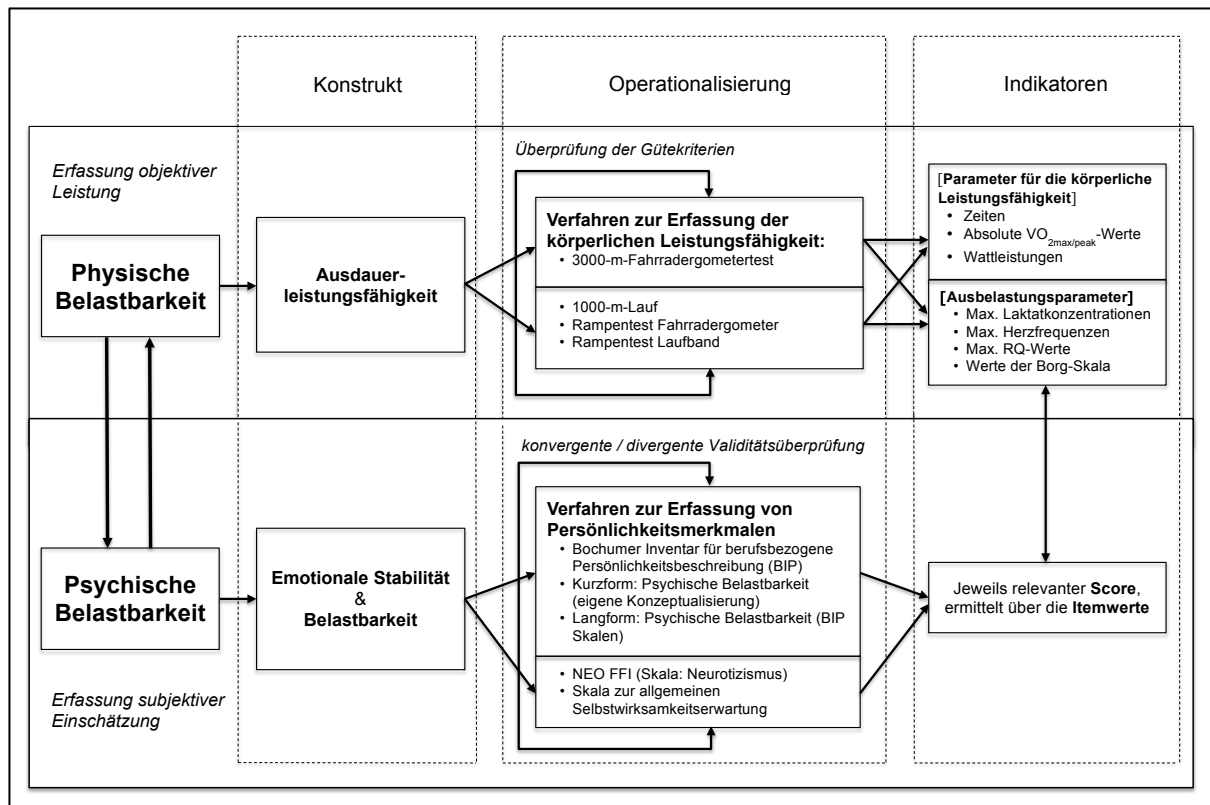


Abbildung 11. Psycho-physisches Belastbarkeitskonzept.

Da für diese Arbeit die berufliche Eignung im Fokus steht, ist für die Erfassung der psychischen Komponente des psycho-physischen Belastbarkeits-Konzepts ein ganz bestimmtes Instrument ausgewählt und eingesetzt worden, welches sich für die Feststellung der berufsbezogenen Eignungsfeststellung seit längerem in verschiedenen Bereichen bewährt hat. Aufgrund der Relevanz dieses Instruments soll im Folgenden dem Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung besondere Aufmerksamkeit zuteil werden.

4.2 Konzeption und Orientierung für die Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente

Mit Hilfe von einer Explorationsstudie, bei der die Untersuchungsteilnehmer sowohl einen Fahrradergometertest als auch den im Rahmen des Basis-Fitness-Tests (BFT) bisher durchgeführten 1000-m-Lauf absolvierten, sollte herausgefunden werden, welche Kriterien (Distanz in Metern und Wattzahl) für das Testdesign eines Fahrradergometertests zur Erfassung der physischen Belastbarkeit für die Personalgewinnung der Bundeswehr festgelegt werden sollte. Ein Anliegen der Testkonzeption und Kriteriendefinition war es, einen vom Belastungsumfang mit dem 1000-m-Lauf vergleichbaren bzw. strukturähnlichen Test zu entwickeln. Der zu konzipierende und zu evaluierende ausdauerorientierte Fitnessstest mit dem Fahrradergometer sollte zum einen justizierbar und somit unter standardisierten Bedingungen für alle Bewerber gleich sein, und zum anderen eine gewisse Strukturähnlichkeit zum 1000-m-Lauf v. a. in Hinblick auf die zeitliche Durchführbarkeit sowie eine Einordnung der erbrachten Leistung in eine Normwerttabelle ermöglichen.

Das Motiv dieser Strukturähnlichkeit ist zum einen der zeitliche Aufwand, der im Rahmen des Eignungsfeststellungsverfahrens für die Erfassung der gesamten physischen Konstitution des Bewerbers vorgesehen ist und im Speziellen in der Feststellung der Ausdauerkomponente im Rahmen des Basis-Fitness-Tests zu sehen. Betrachtet man die Kriterien der Leistungsanforderungen des 1000-m-Laufs, ist ein maximaler Zeitaufwand von 6:30 Minuten für die Testung pro Pb einzukalkulieren, in der der Bewerber die 3000 Meter Strecke bewältigen soll. Bei dem Lauf wurde je nach benötigter Zeit für die Distanz von 1000 Metern der Bewerber in eine Normwerttabelle eingeordnet und aus dieser eine Punktzahl entnommen. Diese Punktzahl wird mit den Punkten der anderen beiden BFT-Disziplinen zu einer Gesamtpunktzahl addiert. Der im Rahmen dieser Studie zu entwickelnde Fahrradergometertest sollte vom zeitlichen Aspekt her gleichermaßen durchführbar sein und die Streckenvorgabe hinsichtlich der Leistungsanforderung dem bisherigen 1000-m-Lauf entsprechen.

Die Explorationsstudie zur Kriteriendefinition kann in zwei Teile gegliedert werden. Die Stichprobe der Explorationsstudie bestand aus 8 Versuchspersonen davon ausschließlich Studierende der Universität der Bundeswehr München. Um die Strukturähnlichkeit der beiden Ausdauerstests zu gewährleisten und den Widerstand bzw. die Wattzahl als Anforderungskriterien für das spätere Testdesign festzulegen, wurde zunächst eine längsschnittlich angelegte

Untersuchung als Feldtest durchgeführt. Zum Zeitpunkt t1 haben die Versuchspersonen den Basis-Fitness-Test (BFT) mit dem 1000-m-Lauf durchgeführt und 2 Wochen später (Zeitpunkt: t2) den BFT mit dem neuen Fahrradergometertest absolviert. Die Testpersonen hatten am Messzeitpunkt t2 die Aufgabe innerhalb von sechs Minuten bei verschiedenen Wattstufen eine maximale Strecke auf dem Fahrradergometer zurückzulegen und sich subjektiv auszubelasten. Anschließend wurden die Leistungen des 1000-m-Laufs (Zeit) und des Fahrradergometertests (Strecke) miteinander verglichen (siehe Kapitel 5.1.1). Bei der Explorationsstudie II zur Evaluierung der vorzugebenden Wattstufe bewältigten 50 Vpn bei einer konstanten Wattstufe auf dem Fahrradergometer, die auf Grundlage der Ergebnisse der Explorationsstudie I festgelegt wurde, einen 6:30-Minütigen Fahrradergometertest. Außerdem absolvierten die Pbn ebenfalls den 1000-m-Lauf. Somit war es möglich die Ergebnisse des bisher im Rahmen des BFT durchgeführten 1000-m-Laufs mit dem auf Grundlage der Explorationsstudie I festgelegten Testkriterien des neu zu konzipierenden Fahrradergometertests zu vergleichen, eine strukturähnliche Streckenlänge zu bestimmen und über eventuelle weitere Modifikationen zu entscheiden.

Die Pbn bekamen bei beiden Teilen der explorativen Studie zur Definition der Testkriterien direkt im Anschluss an die Tests die Borg-Skala vorgelegt, um somit ihr Belastungsempfinden einzuschätzen (siehe Borg, 1970). Die Herzfrequenz und die Borg-Skala können zur Beurteilung der Ausbelastung bei der Fahrradergometrie eingesetzt werden (vgl. Graf et al., 2001).

Eine Problematik bei der Testkonzeption bestand darin, dass die seitens der Bundeswehr angeschafften Ergometer aufgrund der Drehkurbelunabhängigkeit sowie der dazugehörigen Software keinen großen Spielraum für die Testkonzeption zuließen. Außerdem wurden von der Bundeswehr Anforderungen in Form eines zeitlichen Rahmens für die Testdurchführung gestellt, die sich auf maximal zehn Minuten beliefen. In erster Linie sollten mit Hilfe des zu konzipierenden Tests die zu erfüllenden Minimalanforderungen im Hinblick auf die Ausdauerleistungsfähigkeit der Anwärter und Anwärterinnen der Bundeswehr getestet werden.

Aufgrund dieser Restriktionen sowohl seitens der gewünschten Vorgaben der Bundeswehr, als auch der Limitierungen der Ergometer und der entsprechenden Software, wurde im Laufe des Projekts auf die sogenannten Hometrainer der Firma Tacx mit der dazugehörigen Software zurückgegriffen. Des Weiteren wurden für die Umsetzung des Tests Fahrräder der Firma

Cannondale (Modell CAAD 10) in Verbindung mit den Tacx-Trainern eingesetzt, sodass die Testdurchführung nach den definierten Testkriterien weitestgehend realisiert werden konnte.

Im Anschluss an die Definition der Testkriterien und der Lösungsfindung in Bezug auf die Gerätschaften bzw. der Kombination der Messinstrumente wurde eine Laborstudie zur trainingswissenschaftlichen Überprüfung der Gütekriterien des konzipierten Tests durchgeführt. Hierbei wurde das neu konzipierte Testverfahren zur Erfassung der physischen Belastbarkeit im Sinne der Ausdauerleistungsfähigkeit mit dem bisher bei der Bundeswehr eingesetzten 1000-m-Lauf zur Erfassung der Ausdauerleistung verglichen. Außerdem kamen für die Vergleichsstudie auch zwei Rampentests (Laufband und Fahrradergometer) zum Einsatz. Das Untersuchungsdesign der Laborstudie ist in Kapitel 5.1.2.1 dargestellt und erläutert.

4.3 Operationalisierung des Persönlichkeitskonstrukts psychische Belastbarkeit

Mit Hilfe des Bochumer Inventars zur berufsbezogenen Persönlichkeitsbeschreibung, kurz „BIP“ genannt, von Hossiep und Paschen (1998) als einem inzwischen bewährtem Instrument zur Erfassung von zentralen und speziell für die Personalsichtung relevanten Persönlichkeitseigenschaften wird in dieser Arbeit die Operationalisierung der psychischen Komponente des Konstrukts der psycho-physischen Belastbarkeit verfolgt.

4.3.1 Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung

Das BIP gilt als ein etablierter und im deutschsprachigen Raum als eines der gängigsten Instrumente zur Erfassung der berufsrelevanten Persönlichkeitseigenschaften. Entwickelt von Hossiep und Paschen ab Anfang der 90er Jahre und 1998 erstmals veröffentlicht, stellt das BIP eine Methode dar, welche bis heute wegen der wissenschaftlichen Fundiertheit und der ausgereiften Handhabung in der Praxis in verschiedenen Berufsfeldern und bei der Karriereberatung zum Einsatz kommt.

Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung umfasst insgesamt 14 Faktoren, die auf Grundlage von in der Literatur gesichteten Hinweisen als berufserfolgsrelevante Persönlichkeitsdimensionen ausgesucht und „auf der Basis vorliegender Validitätshinweise zu Persönlichkeitskonstrukten und mit Blick auf Anforderungen der diagnostischen Praxis“ (Hossiep und Paschen, 1998, S. 16) zusammengestellt worden sind. Der Skalen- und Testfragenkonzeption des BIP liegen einschlägige und ausgereifte Verfahren und Theorien, wie das Big-Five-Globalfaktorenmodell der Persönlichkeit von Costa und McCrae (1992), das Konzept der Handlungs- und Lageorientierung nach Kuhl (2000), oder die Motivationstheorie von McClelland (1987) zugrunde (vgl. Hossiep & Bräutigam, 2006). Um ergänzende Hinweise auf erfolgsrelevante Dimensionen zu bekommen, haben die Autoren des weiteren Validitätsstudien zu bisher durchgeführten Persönlichkeitstests analysiert und zusätzlich Expertenbefragungen mit Personalverantwortlichen und Personalmanagern hinsichtlich deren Einschätzung wichtiger Persönlichkeitsattribute durchgeführt. Bei der Vorstellung des BIP ist es essentiell wichtig darauf hinzuweisen, dass der BIP, so Hossiep und Paschen (1998, S. 11), zum Ziel hat relevante Beschreibungsdimensionen aus dem Berufsleben anhand des Selbstbildes des Bewerber standardisiert zu erfassen und nicht, die menschlichen Persönlichkeitseigenschaften in Gänze abzubilden, sondern lediglich die relevanten Facetten für den beruflichen Kontext zu eruieren (vgl. Hossiep & Bräutigam, 2006).

Die 14 Dimensionen (Persönlichkeitseigenschaften, -merkmale) des BIP können, wie es in der nachfolgenden Abbildung 12 verdeutlicht ist, in die Kategorien „Psychische Konstitution“, „Arbeitsverhalten“, „Soziale Kompetenzen“ und „Berufliche Orientierung“ eingeteilt werden.

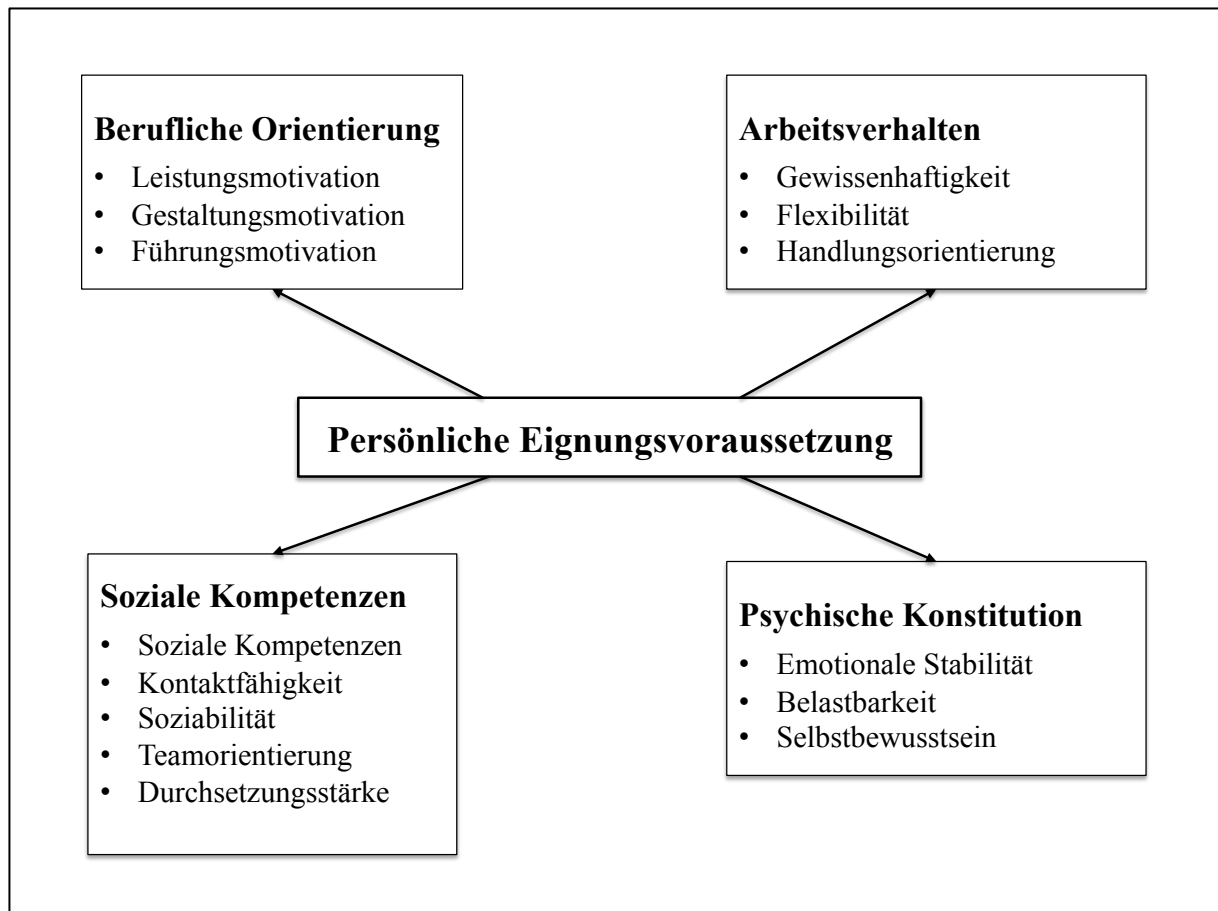


Abbildung 12. Dimensionen-/Faktorenmodell des BIP (nach Hossiep & Paschen, 1998, S. 17).

Der Itempool besteht aus 196 Items, die den oben genannten Dimensionen zugeordnet werden. Hinzu kommen 14 Zusatzindizes, welche eher für ein retrospektives Interview geeignet sind. Aufgrund bisher unzureichender Validität dieser Indizes fließen diese nicht in die Skalen mit ein und werden aufgrund dessen in dieser Arbeit auch nicht weiter berücksichtigt. Die 196 Items verteilen sich mit einer Anzahl zwischen 12 und 16 Items auf die 14 Dimensionen. Wie in Kapitel 5.3 angeführt, ist es ein wesentliches Anliegen dieser Arbeit eine ökonomische und valide Einschätzung der körperlichen und psychischen Belastbarkeit vornehmen zu können, um Zusammenhänge zwischen physischer Fitness und den psychischen Faktoren Emotionale Stabilität und Belastbarkeit auszuloten. Hierbei dienen insbesondere zwei Dimensionen und deren Skalen des BIP als Grundlage für die Erfassung der beiden, für das Anliegen dieser Arbeit speziell relevanten, psychischen Faktoren:

- Belastbarkeit
- Emotionale Stabilität

Hossiep und Paschen (1998) betonten bereits die hohe Interkorrelation dieser beiden Skalen ($r = .68$) und erwähnten die Möglichkeit der Verschmelzung der beiden Faktoren zu einem möglichen neuen Konstrukt:

„Während bei der Emotionalen Stabilität v. a. der gefühlsmäßige Umgang mit schwierigen Situationen erfasst wird, zielen die Fragen dieser Skala eher auf physische Merkmale ab. Die beiden Skalen sind hoch interkorreliert, denn häufig geht physisches Unwohlsein mit körperlichen Beschwerden einher. Dennoch eignen sich beide Dimensionen durch die unterschiedlichen Iteminhalte zur Exploration verschiedener Aspekte eines denkbaren Konstrukts ‘Allgemeine Stabilität‘.“ (Hossiep & Paschen, 1998, S. 51)

Die Idee der Konstruktion eines neuen Faktors, die durch die hohe Interkorrelation der beiden Skalen begründet ist, soll in dieser Arbeit aufgegriffen und umgesetzt werden. Zum einen wird aus den beiden Skalen der Dimensionen Emotionale Stabilität und Belastbarkeit eine neue Dimension „Psychische Belastbarkeit“ gebildet, die sich inhaltlich bezüglich der Items nicht von den zugrunde liegenden Skalen unterscheidet. Hossiep und Paschen (2012) bezeichnen diesen Faktor als „Allgemeine Stabilität“ und beschreiben dessen Eigenschaft mit beruflichen Belastungen wie Misserfolgen, Überforderung und Stress umzugehen. Im Hinblick auf die Emotionale Stabilität besteht eine inhaltliche Übereinstimmung mit dem entgegengesetzten Faktor emotionale Labilität bzw. Neurotizismus des FFM von Costa und McCrae (1989, 1992).

Neben der hohen Interkorrelation der Skalen Emotionale Stabilität und Belastbarkeit, die in der Korrelationsmatrix den höchsten Korrelationskoeffizienten bilden (siehe Anhang A2), besteht darüber hinaus eine hohe interne Konsistenz der beiden Skalen (Cronbach's alpha: Emotionale Stabilität: $r = .89$; Belastbarkeit: $r = .92$), was für eine inhaltliche Homogenität spricht und dafür, dass die über die beiden Dimensionen erfassten Facetten mit zufriedenstellender Genauigkeit gemessen werden (Hossiep & Paschen, 1998, S. 25). Betrachtet man in diesem Zuge die interne Konsistenz der integrierten Skala „Psychische Belastbarkeit“, die sich aus den beiden genannten Skalen zusammensetzt, ergibt sich ebenfalls ein Ergebnis, dass für eine inhaltlich homogene Skala spricht (Cronbach's alpha: $r = .90$; 29 Items; $N = 424$). Insofern kann von einem hohen Niveau hinsichtlich der internen Konsistenz der psychischen Belastbarkeitsskala ausgegangen werden. Für einen genaueren Überblick der Itemstatistik sei auf Anhang E1 verwiesen.

4.3.2 Verfahrensökonomische Überlegungen zur Erfassung der psychischen Belastbarkeit

Hinsichtlich der Erfassung der psychischen Komponente geht es speziell darum, die beiden hier interessierenden und relevanten Persönlichkeitsfaktoren (Emotionale Stabilität und Belastbarkeit) zeiteffizient zu messen. Die Konzeption und Erprobung der Modifikation des Fragebogens liegt in erster Linie der Gedanke und das Anliegen der Ökonomisierung der Datenerhebung zugrunde. Bei der Ökonomie bzw. bei dem Bemühen um Ökonomisierung des Verfahrens stehen einerseits Material- und Zeitkosten bei der Durchführung und andererseits der Aufwand bei der Auswertung im Vordergrund.

„Ein Auswahlinstrument ist dann ökonomisch, wenn die Durchführbarkeit kurz ist, wenn kein aufwändiges Material benötigt wird, wenn die Handhabung einfach ist, wenn mehrere oder viele Personen gleichzeitig untersucht werden können und wenn die Befunde schnell und bequem auswertbar sind.“ (Nerdinger, Blicke & Schaper, 2008, S.259)

Die Nettodurchführungszeit des Tests ist vor allem durch die Anzahl der Items gekennzeichnet. Die Langform, bestehend aus den zwei Dimensionen des BIP „Emotionale Stabilität“ und „Belastbarkeit“ beinhalten insgesamt 29 Items, welche eine Bearbeitungszeit von durchschnittlich 8 Minuten benötigen. Dagegen weist die Kurzskala, mit durchschnittlich deutlich weniger als einer Minute Bearbeitungszeit, eine wesentlich kürzere Dauer zur Bearbeitung auf und ist somit höchst zeiteffizient. Neben der Itemanzahl ist die Verständlichkeit ein weiterer wichtiger Faktor, der zur Ökonomisierung des Fragebogenverfahrens beiträgt. Dazu wird eine Idee aufgegriffen, die sich als summative Einschätzung kennzeichnen lässt. Sie wird der sog. additiven Einschätzung gegenübergestellt, die über - wie die Bezeichnung nahelegt - eine Addition der Antworten über alle Items erfolgt. Die summative Einschätzung bezieht sich auf die gemeinsame Einschätzung von für den Faktor über Faktorenanalysen eruierte Markieritems, sodass für einen Faktor auch nur eine Einschätzung bzw. Antwort erfolgt, was das Vorgehen äußerst ökonomisch und zeiteffizient gestaltet. Die Testgütekriterien dürfen selbstverständlich nicht darunter leiden. Dies ist für die Reliabilität durchaus kritisch zu sehen, da die Reliabilität im Allgemeinen bekanntlich mit der Anzahl an Items steigt. Die Validität sollte anhand der Korrelation zwischen additiver Messung (Langform-Skala) und summativer Messung (Kurzform-Skala) belegbar sein. In diesem Bezug sei darauf verwiesen, dass aufgrund der einfacher nachvollziehbaren Vergleichbarkeit zwischen den beiden Skalenformen und auch anderen Skalen zum Zweck der Überprüfung der Konstruktvalidität auf die Rohwerte zurückgegriffen

wird und nicht auf die (empirisch) ermittelten skalenbezogenen Werte (Mittelwerte der empirisch ermittelten Verteilung auf die Antwortkategorien).

Überlegungen zur Ökonomisierung des Fragebogenverfahrens erschienen aufgrund zeitlicher Aspekte hinsichtlich der Datenerhebung angebracht. Die Erhebung der Persönlichkeitseigenschaften sollte in Verbindung mit der Erfassung der physischen Komponente (u. a. Fahrradergometer Testung) realisiert werden. Insofern spielte der Faktor Zeit für die Untersuchung eine wesentliche Rolle. Da für die Datenerhebung möglichst viele Personen an einem Untersuchungstermin getestet werden sollten und nur ein begrenzter Zeitrahmen für die Testung eines Pb angesetzt war, kamen Überlegungen zur Ökonomisierung des Verfahrens auf. Infolgedessen wurde eine „Kurzskala“ mit einer positiv und einer negativ gepolten „Wolke“ (symbolhaft in einer Wolke zusammengefasste Adjektive) aus den beiden für diese Arbeit relevanten Skalen Emotionale Stabilität und Belastbarkeit entwickelt (siehe Abbildung 13). Bei diesem Vorgehen wurde eine systematische Berücksichtigung positiver und negativer Formulierungen vorgenommen, um so evtl. Antworttendenzen zu begegnen, oder auch um inhaltlich den Interpretationsspielraum einzuschränken und die Klarheit des Faktors zu erhöhen. Wie bereits erwähnt wurden aufgrund der hohen Interkorrelation die beiden Skalen zu einer Skala „Psychische Belastbarkeit“ zusammengefasst. Als Grundlage für die Erfassung der psychischen Belastbarkeitsskala wird das BIP herangezogen.

Bei der Erhebung der Daten zur psycho-physischen Belastbarkeit wird die Skala der Psychischen Belastbarkeit in Form eines Fragebogens für diese Arbeit zum Einsatz gebracht (BIP-Skalen: Emotionale Stabilität und Belastbarkeit) und stellt die „Langform“ des Befragungsinstruments dar. Parallel zu der Langform wird den Versuchspersonen während der Untersuchungen auch eine eigens dazu konzipierte „Kurzform“ dieser Skala vorgelegt. Diese Kurzform(-Skala) basiert auf Markieritems, die aus vorliegenden Faktorenanalysen identifiziert werden konnten. Es scheint in dieser Hinsicht aufgrund der Untersuchungen von Rösler, Baumann und Marake (1980) sinnvoll und zweckmäßig, anstelle der additiven (über die Einzelitems), eine summative (auf die Faktoren bezogene) Einschätzung vorzunehmen, die vor allem eine Ökonomisierung des Erhebungsvorgangs durch eine verkürzte Datenerhebungszeit gewährleistet. Bei dem summativen Verfahren, werden die Items mit der höchsten Ladung bei der Faktorenanalyse (Markieritems) oder mit den besten Trennschärfeindizes herangezogen, um die beiden in dieser Arbeit relevanten Dimensionen zu kennzeichnen und summativ einschätzen zu lassen. Hackfort und Schlattmann (1995) erläutern ein solches Vorgehen am Bei-

spiel ihrer Stimmungs- und Befindensskalen und berichten über Studien, nach denen sich dieser Verfahrensansatz bei den Skalen bewährt hat. Im Rahmen der Untersuchungen der psycho-physischen Belastbarkeit soll hier ein zweigleisiges Vorgehen angewendet werden um herauszufinden, ob sich die lange und kurze Variante der Befragung in für den Einsatz und die Messung wichtigen Gesichtspunkten unterscheidet oder hinsichtlich der Zielstellung ebenbürtig ist. Rösler et al. (1980) wie Hackfort und Schlattmann (1995) konnten bereits durch hohe Korrelationen und hoch signifikante Befunde bei dem Vergleich der beiden Ansätze nachweisen, dass durch die Kurzsкала im wesentlichen die gleichen angestrebten Informationen gewonnen werden konnten. Unter vorangehender Betrachtung der Testfragen der beiden Skalen werden für die einzelnen Fragen prägnante, maßgeblich kennzeichnende Adjektive (Markieritems) definiert bzw. ausgewählt, die dann zusammengefasst in einer Wolke zur Repräsentation des Faktors dargestellt werden. In Abbildung 13 ist gekennzeichnet, welche Markieritems für die Bildung der Wolken herangezogen worden sind und welche Adjektive sich aus diesen Items für die „Wolkenbildung“ ableiten ließen.

Im Folgenden wird beispielhaft das Vorgehen gekennzeichnet, wie sich die Itemkennwerte der Testfragen des BIP zu einer summativen Einschätzung (Kurzform der Befragung) zusammenfassen und präsentieren lassen. Um die Kurzform des Fragebogens zu bilden, wurde eine Itemselektion und Transformation (von der Fragesatzform zur Adjektivform) durchgeführt. Auf Grundlage der Testfragenbatterie der beiden Skalen „Emotionale Stabilität“ und „Belastbarkeit“ wurden die Itemkennwerte mit den höchsten Ladungen ausgewählt. Da, wie bereits beschrieben, diese Skalen eine hohe Interkorrelation aufweisen, wurden jeweils von den Items mit den höchsten Ladungen beider Faktoren, die sowohl in Richtung einer hohen Ausprägung des Konstrukts formuliert sind (positiv gepolt) als auch die in Richtung einer niedrigen Ausprägung formuliert sind (negativ gepolte Items), Adjektive abgeleitet, welche die Items mit der höchsten Ladung am besten repräsentieren. Die Messlatte für die Ladung der Items, die für die Kurzform des Fragebogens in Frage kamen, wurden bei der Skala der Emotionalen Stabilität auf $rf \geq .65$ gesetzt (oberes Drittel in der Ladung) und bei der Skala der Belastbarkeit auf $rf > .70$. Anschließend wurde eine Wolke gebildet, die vier Adjektive umfasst („instabil“, „nervös“, „reizbar“ und „entmutigt“), welche die negativ gepolten Items mit der höchsten Ladung beider Skalen repräsentiert und eine weitere Wolke, die vier Adjektive beinhaltet („robust“, „willensstark“, „gelassen“ und „standhaft“), die die positiv gepolten Items mit der höchsten Ladung repräsentiert.

Das Item Nr. 183 bezieht sich auf die Skala Emotionale Stabilität „Mich wirft so leicht nichts aus der Bahn“ und weist mit einer Ladung von $rf = .73$ eine besonders hohe Ladung auf diesem Faktor auf. Das Item kann mit dem Adjektiv „robust“ gekennzeichnet werden. Das Item 19, „Ich bleibe gelassen, auch wenn vieles gleichzeitig auf mich einströmt“, sowie das Item 65 der Skala Belastbarkeit, „Auch wenn ich sehr hart arbeiten muss, bleibe ich gelassen“, sind ebenfalls positiv gepolt und haben mit einer Ladung von $rf = .71$ bzw. $rf = .75$ eine ähnlich hohe Ladung vorzuweisen. Eine naheliegende Kennzeichnung ausgedrückt in einem für die Kurzskala passendem Adjektiv wäre hier der in beiden Items vorhandene Begriff „gelassen“. Für das Item fünf „Wenn ich extrem hart arbeiten muss, gerate ich bei zusätzlichen Schwierigkeiten aus dem Gleichgewicht“ (Polung: negativ; $rf = .77$) der Skala Belastbarkeit des BIP wurde hier das Adjektiv „instabil“ herangezogen. Das Adjektiv „reizbar“ ließ sich aus dem negativ gepolten Item Nr. 156 ($rf = .67$) und aus dem Item 127 ($rf = .72$) der Skala Belastbarkeit herleiten „Wenn ich sehr viel zu tun habe, wirke ich auf andere gereizt“ und „Wenn ich unter starkem Druck stehe, reagiere ich gereizt.“ Für die beiden anderen Stichwörter „nervös“ und „entmutigt“ der negativen Wolke der aus den beiden BIP-Skalen entwickelten Kurzform wurden aus den Items 126 ($rf = .67$) „Ich fühle mich manchmal ziemlich entmutigt“ der Skala der emotionalen Stabilität und des Items 141 ($rf = .79$) „Bei gleichzeitigen Anforderungen von mehreren Seiten werde ich nervös“ der Belastbarkeitsskala herangezogen. Für die positive Wolke der Kurzskala wurden die beiden Adjektive „standhaft“ und „willensstark“ ausgewählt und aus den Items 169 ($rf = .73$) der Emotionalen Stabilitätsskala „Mich wirft so leicht nichts aus der Bahn“ und Nr. 49 ($rf = .71$) „Ich verkrafte lang andauernde, hohe Belastungen besser als andere“ (Skala: Belastbarkeit) hergeleitet. Fasst man diese vier Adjektive, die für die Faktoren charakteristisch sind, zusammen, kann eine summative Einschätzung so abgegeben werden, wie es im Folgenden veranschaulicht ist (Abbildung 13). Die Antwortmöglichkeiten waren genauso strukturiert wie bei der Originalform des BIP und wurden in sechs Antwortkategorien unterteilt, von „trifft voll zu“ bis „trifft überhaupt nicht zu“, sodass die „goldene Mitte“ ausgeschlossen war und eine Antworttendenz bei der Selbsteinschätzung der Teilnehmer ersichtlich wurde.

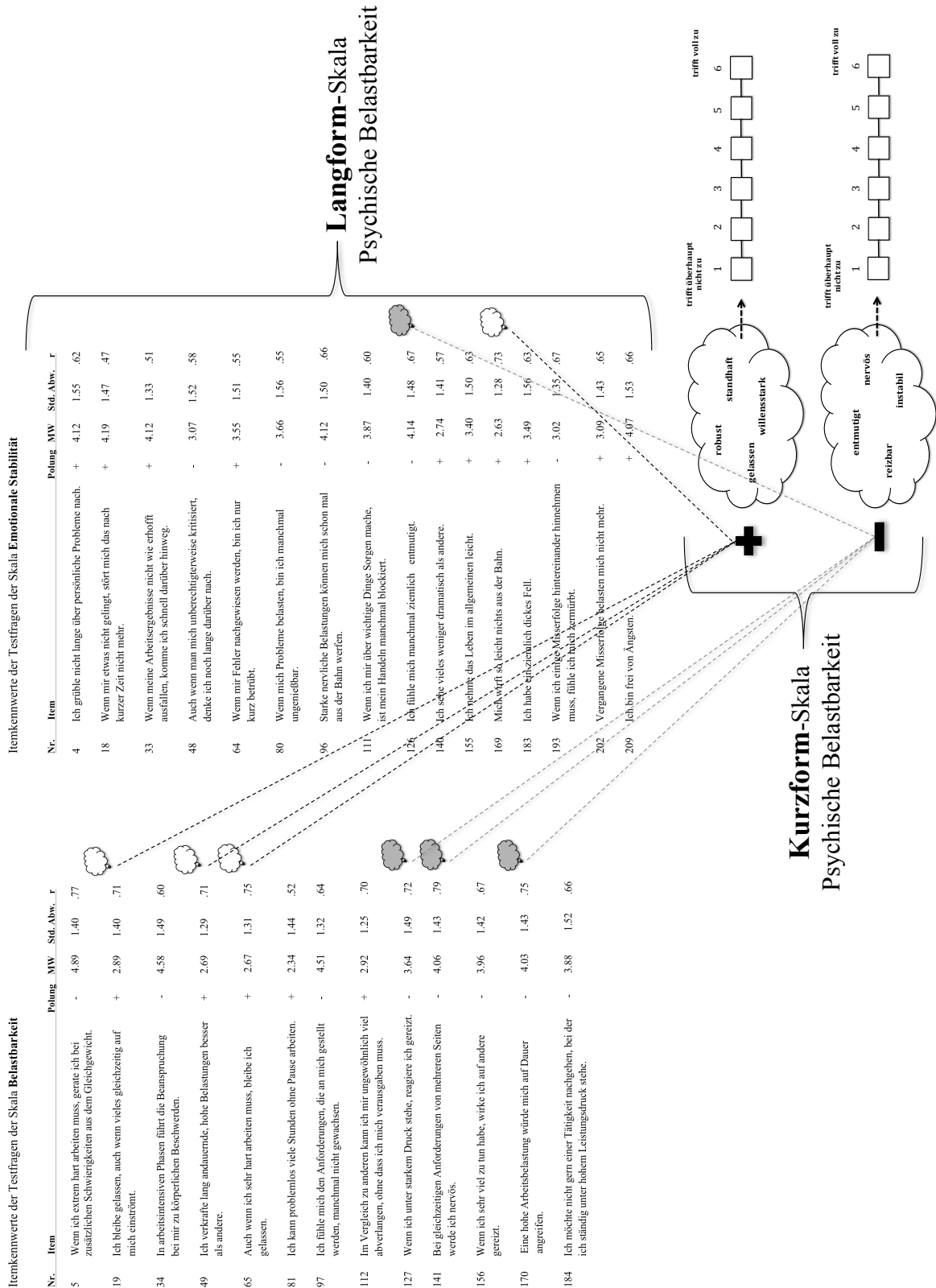


Abbildung 13. Ableitung der Kurzskala aus den Skalen „Emotionale Stabilität“ und „Belastbarkeit“ des BIP.

4.3.3 Ansatz zur Validierung des Selbsteinschätzungsverfahrens

Für die Validierung des Persönlichkeitskonstrukts psychische Belastbarkeit bzw. des Instruments zur Erfassung der psychischen Belastbarkeit, bestehend aus den beiden Skalen Emotionale Stabilität und Belastbarkeit des BIP, kamen im Rahmen von bestimmten Teiluntersuchungen weitere Persönlichkeitsskalen zum Einsatz, um die Konstruktvalidität der Skalen (Lang- und Kurzform) empirisch zu belegen. Aufgrund der beschriebenen hohen Interkorrelation der beiden Skalen aus dem BIP (Emotionale Stabilität und Belastbarkeit) wurden diese zu einer Skala „Psychische Belastbarkeit“ zusammengefasst und daraus eine selbstkonzipierte Kurzskala entwickelt (siehe Kapitel 4.3.2).

Für die Konstruktvalidierung beider Skalenformen der psychische Belastbarkeit wurden bei verschiedenen Testungen im Rahmen dieser Untersuchung weitere Skalen ausgewählt und den Pbn vorgelegt, um somit zum einen die Konvergenz und zum anderen die Divergenz des Konstrukts psychische Belastbarkeit (Lang- und Kurzskala) zu überprüfen. Eine konvergente Validität liegt vor, wenn ein bestimmtes Persönlichkeitskonstrukt anhand von verschiedenen Messmethoden bzw. Verfahren (Skalen) gemessen wird, die das gleiche Konstrukt erfassen und die Ergebnisse dieser Messmethoden in hohem Maße übereinstimmen. Darüber hinaus liegt eine konvergente Validität vor, wenn im nomologischen Netzwerk naheliegende Konstrukte bzw. deren Operationalisierung korrelieren. Hingegen liegt eine divergente Validität vor, wenn das zu erfassende Konstrukt und das dafür herangezogene Instrument etwas anderes misst, als andere Konstrukte und deren Messinstrumente (Operationalisierung), d. h. die Korrelationen der für die divergente Validitätsüberprüfung herangezogenen Skalen sollten im Vergleich zu dem Zielkonstrukt gegenläufig sein. Für die divergente Validitätsüberprüfung wurde die Dimension Neurotizismus aus dem NEO-Fünf-Faktoren-Inventar (NEO-FFI), nach Borkenau und Ostendorf (2008) eingesetzt und die Skala zur allgemeinen Selbstwirksamkeitserwartung nach Schwarzer und Jerusalem (1999) für die konvergente Validitätsüberprüfung herangezogen.

Das NEO-FFI wurde in Anlehnung an Costa und McCraes amerikanischer Originalversion (1989) entwickelt. Die Skala Neurotizismus zielt auf die Erfassung des Persönlichkeitsmerkmals „emotionale Labilität“ ab. Mit den beiden Skalen psychische Belastbarkeit und Neurotizismus werden gegensätzliche Merkmale gemessen. Bei dem ersten Untersuchungstermin, einer trainingswissenschaftlichen Leistungsdiagnostik (siehe Kapitel 5.1.2) mit dem Fahrrad-

ergometer, kamen die beiden Originalskalen des BIP (Konstrukt: Psychische Belastbarkeit) sowie die eigens konstruierte Kurzskala zum Einsatz. Der zweite Termin umfasste eine Leistungsdiagnostik via Laufband, ebenfalls mit Spiroergometrie und Laktatabnahme und einer Persönlichkeitsskala, basierend auf dem NEO-FFI, die das Konstrukt Neurotizismus anhand von 12 Items mit einer fünfstufigen Antwortmöglichkeit erfasst. Die Anwendung des Neurotizismus-Fragebogens hat den Zweck einer Konstruktvalidierung (divergenter Aspekt) der beiden psychischen Belastbarkeitsskalen. Durch die Bearbeitung beider Instrumente, die gegensätzliche Konstrukte erfassen, soll die Überprüfung einer divergenten Validierung erfolgen, welche dann als erfolgreich gelten kann, wenn die beiden Instrumente, die unterschiedliches messen, Emotionale Stabilität (BIP) und Emotionale Labilität (NEO-FFI), idealtypisch gegenläufig bzw. negativ miteinander korrelieren (Nerdinger, Blickle & Schaper, 2008).

Darüber hinaus wurde überprüft, inwieweit die einzelnen Items der Skalen mit den Items der Skalen korreliert, die das gleiche Konstrukt messen (BIP-Skalen) und negativ mit den Items korreliert, die ein anderes bzw. gegensätzliches Konstrukt, d. h. hier speziell emotionale Labilität (Neurotizismus), messen. „Ist dies für alle Items eines Instrumentes der Fall, spricht man von einer faktoriellen Validierung im Rahmen der Konstruktvalidierung“ (Nerdinger, Blickle & Schaper, 2008, S. 255). Zudem verweist Wenzel (2008, S. 59) darauf: „Neben Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit zählt Neurotizismus als Gegenpol der emotionalen Stabilität, zu den fünf die Persönlichkeit beschreibenden Faktoren. Mit Neurotizismus assoziierte Adjektive sind nach John (1990, S. 80, übersetzt von Amelang & Bartussek, 2001, S. 366) ‚gespannt‘, ‚ängstlich‘, ‚nervös‘, ‚launisch‘, ‚besorgt‘, ‚empfindlich‘, ‚reizbar‘, ‚furchtsam‘, ‚selbst bemitleidend‘, ‚instabil‘, ‚mutlos‘, ‚verzagt‘ und ‚emotional‘. Mit emotionaler Stabilität verbundene Adjektive sind ‚stabil‘, ‚ruhig‘, ‚zufrieden‘, ‚ausgeglichener‘ und ‚entspannt.“ Mit dem Vergleich der Itemscores der Personen, die sowohl die psychische Belastbarkeits-Skala und die Neurotizismus-Skala beantworteten, wird tendenziell ein hoher negativer Zusammenhang erwartet, der insofern auf eine Divergenz der beiden Konstrukte schließen lassen würde.

Zusätzlich kam im Rahmen einer weiteren Teilstudie die Skala zur allgemeinen Selbstwirksamkeitserwartung neben der Lang- und Kurzskala zu Erfassung der Psychischen Belastbarkeit zum Einsatz. Die eindimensionale Skala der Selbstwirksamkeit beinhaltet zehn Items, die gleichsinnig gepolt sind und vierstufig beantwortet werden (Schwarzer & Jerusalem, 1999). „Jedes Item bringt eine internal-stabile Attribution der Erfolgserwartung zum Ausdruck. (. . .)

Der individuelle Testwert ergibt sich durch das Aufsummieren aller zehn Antworten, so dass ein Score zwischen 10 und 40 resultieren muss.“ (Schwarzer & Jerusalem, 1999, S. 2) Schwarzer und Jerusalem (1999) fügen an, dass die Skala die selbsteingeschätzte Überzeugung misst, kritische Anforderungssituationen aus eigener Kraft erfolgreich bewältigen zu können. Untersuchungen von u. a. Schumacher, Leppert, Gunzelmann, Strauß und Brähler (2004) liefern Grund zur Annahme, dass Personen, die sich als allgemein selbstwirksam betrachten, sich selbst auch als psychisch belastbarer bzw. psychisch widerstandsfähiger einschätzen. Die Autoren berichten im Rahmen einer Studie über die Überprüfung der Konstruktvalidität (konvergent) der Resilienz-Skala (gekürzte Version: RS-11) zur Erfassung der psychischen Widerstandsfähigkeit mit der Skala zur allgemeinen Selbstwirksamkeitserwartung (SWE) anhand einer bivariaten Korrelationsberechnung, wobei eine signifikant positive Korrelation i. H. v. $r = .70$ ($p < .001$) nachgewiesen werden konnte. Wie von den Autoren erwartet, stellen die beiden Merkmale „eng verwandte Konstrukte [dar]. Beide werden als personale Ressourcen und als Protektivfaktoren angesichts von Belastungen und riskanten Lebensbedingungen betrachtet“ (Schumacher et al., 2004, S. 11). Das Konstrukt Selbstwirksamkeit (self-efficacy) ist eine spezielle Art bzw. Form des Selbstvertrauens, nämlich ein aufgabenbezogenes Selbstvertrauen; genauer ist es eine Einschätzung hinsichtlich des Vertrauens in die eigenen Fähigkeiten eine Aufgabe zu bewältigen. Auch im Sinne einer Überzeugung mit den eigenen Fähigkeiten eine Aufgabe zu vollbringen, deswegen auch Selbstwirksamkeitsüberzeugung genannt (vgl. Morris, 1995).

5 VERFAHRENTWICKLUNGEN UND UNTERSUCHUNGEN ZUR PSYCHO-PHYSISCHEN BELASTBARKEIT

In diesem Teil der Arbeit geht es um die empirischen Studien zu den im Kapitel 4 formulierten Fragestellungen. Im ersten Schritt (Kapitel 5.1) wird die Verfahrensentwicklung zur Erfassung der physischen Belastbarkeit für die Personalauswahl (bei der Bundeswehr) beschrieben und die Befunde der entsprechenden Studien zur Entwicklung und Validierung eines speziellen Ausdauertests im Kontext der Personalgewinnung dargestellt. Im Anschluss (Kapitel 5.2) wird das konzipierte Verfahren zur Erfassung der (berufsbezogenen) psychischen Belastbarkeitskomponente (Lang- und Kurzskala) mit Hilfe der in Kapitel 4.3.2 vorgestellten Skalen validiert und die Ergebnisse aus den dazu durchgeführten Untersuchungen beschrieben. Im dritten Abschnitt dieses Teils (Kapitel 5.3) wird die Verknüpfung der beiden entwickelten Verfahren zum einen zur Überprüfung psycho-physischer Zusammenhänge und zum anderen von Gruppenunterschieden als eine Möglichkeit des Assessments psycho-physischer Belastbarkeit im Rahmen von Personalauswahlprozessen.

5.1 Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente

Im Rahmen eines Drittmittel-Projekts zur Evaluierung und Validierung eines Ergometertests für die Personalgewinnung in der Bundeswehr, wird unter Mitwirkung des Autors ein bisher durchgeführter 1000-m-Lauf als Teil des Basis-Fitness-Tests (BFT; siehe Weisung IGF des Generalinspektors der Bundeswehr, 2009) durch einen hinsichtlich der Beanspruchung der Ausdauerleistungsfähigkeit gleichzusetzenden Test mit dem Fahrradergometer ersetzt. Der Grund dafür liegt in den Durchführungsbedingungen, die beim Test mit einem Fahrradergometer leichter konstant gehalten werden können. Der BFT ist zum einen Teil des eignungsdiagnostischen Verfahrens bei Berufssoldaten und Offiziersanwärtern und zum anderen stellt es ein Monitoringtool zur Kontrolle der physischen Verfassung der Soldaten und Soldatinnen dar. Mit dem BFT soll die allgemeine körperliche Leistungsfähigkeit und körperliche Belastbarkeit des zukünftigen Personals der Bundeswehr erfasst werden sowie eine jährliche Überprüfung der Basisfitness aller im Dienstverhältnis stehenden Personen durchgeführt werden. Ziel des BFT ist „einerseits valide und effiziente Leistungstests zur Ermittlung der körperlichen Leistungsfähigkeit, die zeitökonomisch und ohne größeren apparativen Aufwand die

Überprüfung von grundlegenden physischen Fähigkeiten ermöglichen“ (Leyk, Witzki, Gorges, Rhode, Lison, Rondé, Wömpener, Schlattmann, Dobmeier, Rütter & Wunderlich, 2010, S. 279). Der BFT wurde im Jahre 2010 flächendeckend in der Bundeswehr eingeführt, um effiziente und valide Aussagen bzw. Daten über die Ausdauer-, Schnelligkeits- und Kraftleistungen des Personals der Bundeswehr, sowie deren Bewerberinnen und Bewerber, treffen und erfassen zu können. Die Daten dienen des Weiteren als präventives Screening-Verfahren, „mit dem frühzeitig negative Veränderungen sichtbar und auf die Bedeutung von gesundheitsförderlichem Verhalten aufmerksam gemacht werden kann“ (Leyk et al., 2013, S. 164; siehe auch Leyk et al., 2010). Der Einsatz des Fahrradergometers zur Ermittlung der Ausdauerbelastbarkeit und die damit einhergehende Substitution des 1000-m-Laufs erfolgt aufgrund evtl. auftretender Probleme bei der notwendig sicherzustellenden Gleichheit der Durchführungsbedingungen beim 1000-m-Lauf. So soll in Zukunft bei der Rekrutierung von Berufssoldaten eine einheitliche Testbedingung gesichert werden können, um somit auch eine Vergleichbarkeit der Leistungen und Bewertungen zu gewährleisten. Das Ergometerverfahren, welches anstelle des 1000-m-Laufs zukünftig zum Einsatz kommt, hat den Zweck einer bestmöglichen Standardisierung, Reproduzierbarkeit und Praktikabilität über die Aussage der Ausdauerleistungsfähigkeit, sowie einer ganzjährigen witterungsunabhängigen Durchführbarkeit unter optimierten verletzungsvermeidenden Bedingungen. Die Testung mit dem Ergometer bietet darüber hinaus, durch unterschiedliche Gegebenheiten der Bundeswehr Karrierecenter, infrastrukturell (eine Laufbahn ist nur teilweise vorhanden) und logistisch (Gerätschaften sind transportierbar und auch einsetzbar auf Berufsmessen) weitere Vorteile.

Im Rahmen der Entwicklung des Ergometertests wird auch eine Studie durchgeführt, in der auf der Basis des bisherigen 1000-m-Laufs für den Fahrradergometertest eine vergleichbare Belastungsintensität ermittelt werden soll. Der 1000-m-Lauf als Bestandteil des BFT wurde unter verschiedensten Bedingungen aufgrund unterschiedlicher infrastruktureller Voraussetzungen und jahreszeitlich bedingter inkonstanter Witterungsverhältnisse durchgeführt. So wurde er teilweise auf Außensportanlagen mit oder ohne 400-m-Laufbahn, in der Halle mit unterschiedlichen Raumausmaßen oder sogar in Anlagen durchgeführt, die eigentlich nicht (z. B. Bodenbeschaffenheit) für sportliche Zwecke vorgesehen sind. Dadurch bedingt mangelte es dem so durchgeführten 1000-m-Lauf an den o. g. Kriterien und einer standardisierten Durchführbarkeit. Der zu entwickelnde Ergometertest soll im Hinblick auf diese Mängel Abhilfe schaffen, trotzdem in einem zeitlich effizienten Rahmen bleiben und nach Vorgabe der Bundeswehr in maximal zehn Minuten abgeschlossen sein. Die Normwerttabelle des 1000-m-

Laufs sieht eine Bewältigung der zu laufenden Kilometerdistanz in maximal 6:30 Minuten vor. Diese zeitliche Vorgabe soll auch für den neuen Fahrradergometertest eine Zeitzielgröße darstellen. Die zurückzulegende Distanz für den zukünftigen Ergometertest wird auf Grundlage der Ergebnisse der vergleichenden Beanspruchungsstudie (Explorationsstudie, siehe Kapitel 5.1.1) definiert und für die darauffolgenden Validierungsstudien des Ergometerverfahrens festgelegt (Kapitel 5.1.2).

Der Anlass und Zweck des Projekts sind dadurch gegeben, dass die Personalgewinnungsorganisation der Bundeswehr Bewerberinnen und Bewerber im Rahmen der Eignungsfeststellung auf ihre körperliche Leistungsfähigkeit testet. Hierzu wurde bisher der PFT (Physical Fitness Test) angewandt, der für die Zentren für Nachwuchsgewinnung in der Teildisziplin Cooper-Test (12-Min.-Lauf) aus infrastrukturellen Gründen mittels einer Ergometertestung abgewandelt wurde. In Zukunft soll der BFT diese Funktion wahrnehmen. Aufgrund weiterhin bestehender infrastruktureller Einschränkungen (kein Sportplatz bzw. große Laufsporthalle verfügbar) kann in den zukünftigen Karrierecentern der Bundeswehr jedoch keine Testung des originären und im Vergleich zum Cooper-Test deutlich schneller (höhere Verletzungsgefahr durch wesentlich höhere Fliehkräfte in den Kurven) zu absolvierenden 1000-m-Laufs des BFT (Basis-Fitness-Test) Truppe vollzogen werden. Für die Personalgewinnungsorganisation (PersGOrg) muss das Testverfahren justiziabel, vergleichbar, ganzjährig witterungsunabhängig und logistisch durchführbar sein.

Diese Vorgaben sind für die Verfahrensentwicklung und –anwendung, wie sie im Rahmen dieser Arbeit erfolgt, maßgebend. Zur Einordnung und zum besseren Verständnis ist auf die Phasen des Projekts „BFT PersG“ zur Entwicklung und Validierung eines neuen Fahrradergometertests kurz hinzuweisen: Auf der Grundlage von Voruntersuchungen (siehe Kapitel 5.1.1) und den o. a. zeitlichen Vorgaben zur Testdurchführung legte man sich auf eine Distanzvorgabe von 3000 Meter für den Fahrradergometertest und einen Zeitrahmen zum bewältigen dieser Distanz, identisch der Zeitvorgabe des 1000-m-Laufs, von 6:30 Min., fest. Zur Entwicklung und Evaluierung wurden seitens der Bundeswehr mehrere Fahrradergometer des Typs CYCLE 457 MED zur Verfügung gestellt, die anfangs der Untersuchungen im Dezember 2012 bereits als zukünftige einzusetzenden Fahrradergometer in den KarriereCentern der Bundeswehr vorgesehen waren.

Die erste Datenerhebung, sowohl der Einsatz der Fragebögen (Lang- und Kurzsкала zur psychischen Belastbarkeit; siehe Kapitel 4.3.2), als auch die Erprobung des Fahrradergometertests, fand auf der Fitnessfachmesse FIBO (11.04.2013 bis 14.04.2013) in Köln statt. Die Pbn wurden mit Hilfe einer Studentengruppe in die Testung eingewiesen und während der Testung die zurückgelegte Distanz und die benötigte Zeit für die 3000 Meter von den angehenden Sportwissenschaftlern erfasst. Hierbei zeigte sich im Zusammenhang mit dem Ergometer die Problematik der Umdrehungszahl deutlich, welches das Testdesign im weiteren Verlauf der Studie erschwerte und im Ganzen in Frage stellte. Die fitteren bzw. die schwereren männlichen Versuchspersonen fingen mit zunehmender Trittfrequenz bei den gleichbleibenden 130 Watt an, kaum noch das Gefühl zu haben, Widerstand bei der Bewältigung der Strecke zu spüren und fingen nach mehrfacher Rückmeldung an, „durchzutreten“. Insgesamt zeigten sich folgende Probleme mit diesem Ergometer: Funktionalität der Software (das Testdesign ließ sich nicht programmieren), Drehkurbelunabhängigkeit des Ergometers und eine unpräzise Distanzanzeige auf dem Display.

Im Gegensatz zu männlichen Pbn mit einem höheren Körpergewicht, waren Frauen mit einem niedrigen Körpergewicht bei einer Wattzahl von 130 Watt klar überfordert und aufgrund der limitierenden Funktionalität des Ergometers benachteiligt. Je nach Körperkonstitution der Versuchsperson, bzw. der körperlichen Voraussetzung wurden mit dem Test durch die konstant gehaltene Wattzahl unterschiedliche Ausdauerarten getestet. Demnach ist die erbrachte Leistung der vom Körpergewicht her leichten Probandinnen eher einer Kraftausdauerleistung zuzuordnen. Die Leistung der fitteren Teilnehmer oder schwereren männlichen Pbn ist, aufgrund der hohen Trittfrequenz und des kaum gespürten Widerstands, tendenziell dem einer Schnelligkeitsausdauerleistung zuzuordnen. Die Ursache dieser Problematik ist, wie bereits erwähnt auf die „Drehkurbelunabhängigkeit“ zurückzuführen.

Darüber hinaus zeigte sich ein Problem hinsichtlich der Messungengenauigkeit bei der Feststellung der tatsächlichen Distanz. Die Versuchsleitern konnten beobachten, dass sich der Unterschied der zurückgelgten Strecke in einem Bereich von 100 bis 400 Meter bewegte. Jede Versuchsperson sollte sich vor dem Start des 3000-Meter-Tests erst etwa 1-2 Min. an das Fahrradergometer gewöhnen und eine „Warm-up-Phase“ durchlaufen bevor der Test begann. Da keine Voreinstellung des Testdesign möglich war, mussten die Versuchsleiter im Anschluss an die Warm-up-Phase die Einstellung für den Test stets manuell ändern und konnten erst nachdem die Zeit für den Test bereits lief, die vorgegebene Wattstufe (130 Watt) von Null

nach oben regulieren. Während die Versuchsleiter die Einstellung der Wattstufe vornahmen, befanden sich die Pbn bereits in der Testphase und erhöhten die Trittfrequenz, um die vorgegebene Distanz in einer möglichst kurzen Zeit zurückzulegen. Die Wattstufenregulierung konnte erst nach dem Start der Zeit vorgenommen werden, eine Programmierung der Wattstufe vor dem Teststart war nicht möglich. Somit war in der Anfangsphase des Tests eine Messgenauigkeit gegeben, die sich in Abhängigkeit von der Dauer der Wattregulierung bzw. Einstellung des Widerstands des Versuchsleiters auf die zurückgelegte Distanz und somit auf das Ergebnis des Pbn auswirkte.

Die zweite Gelegenheit der Datensammlung fand auf dem Gesundheitstag der Universität der Bundeswehr in München (23.05.2013) statt. Nach der ersten Testungen auf der FIBO und den dabei erkannten Faktoren, die zur Schwierigkeiten bei der Umsetzung des Testdesigns führten, wurden daraufhin zwei i-Flow-Trainer der Firma Tacx von dem Modell T2250 mit der dazugehörigen Software eingesetzt (siehe Abbildung 14). Dabei handelt es sich um ein aus dem Radsport bekanntes Trainingsgerät für den Indoorbereich, bei dem man bspw. mit einem Rennrad ein Straßentraining simulieren kann, das die Strecke relativ genau misst und auch für wissenschaftliche Zwecke geeignet ist. Auf dem Gesundheitstag wurde ein Stand mit jeweils zwei Tacx-Trainern mit jeweils einem Rennrad der Marke Cannondale (Modell: CAAD 10; siehe Abbildung 15) aufgebaut und beide Trainer mit jeweils einem Laptop und der zugehörigen Tacx-Trainer-Software („4 Basic“) über ein USB-Kabel verbunden. Es bestand die Möglichkeit, über die Software das Testdesign als Trainingsprogramm anzulegen und die Distanzvorgabe von 3000 Meter sowie das Streckenprofil (0 % Steigung) zu programmieren, sodass der Test automatisch nach einem fünf-sekündigen Countdown gestartet und nach Erreichen der 3000 Meter beendet wurde. Über die Anlage eines Pbn-Profiles mit Angabe über Körpergröße und Gewicht konnte die maximale und die Durchschnittsleistung des Pbn von der Software berechnet werden. Zudem wurden über die Software die genaue Distanz, die maximale und durchschnittliche Trittfrequenz, Watt und Geschwindigkeit automatisch gemessen und konnten als zusätzliche Parameter erfasst werden. Bei den ersten beiden umfangreichere Datensammlung auf der FIBO und dem Gesundheitstag an der Universität der Bundeswehr München konnten insgesamt 47 Teilnehmer gewonnen werden, die in die Gesamtstichprobe (siehe Kapitel 5.3) einfließen.



Abbildung 14. Darstellung des Bildschirms (Screenshot) der Tacx-Trainer-Software während der Testphase des 3000-m-FET.

Zu Beginn sehen die Pbn auf dem Bildschirm einen Countdown von fünf Sekunden. Nach Ablauf des Countdowns ertönt ein Startsignal und es erscheint ein visuelles Startzeichen auf dem Bildschirm des Laptops. Der Pbn beginnt nun mit dem Treten des Pedals und hat die Option während des Testvorgangs den Gang am Lenker des Rads frei zu wählen und den Gang bzw. den Widerstand beliebig zu verstellen. Somit kann der Pbn eine individuelle optimale Leistung erreichen unabhängig von Gewicht und Konstitution. Sobald die 3000 Meter erreicht sind, ist der Test beendet und die Daten werden von der Software automatisch aufgezeichnet und gespeichert.



Abbildung 15. 3000-m-FET zur Feststellung der Basisfitness mit den Instrumenten der Firma Tacx.

Der Einsatz der Tacx-Trainer mit den Rennrädern erwies sich als sehr fruchtbar für weitere Untersuchungen im Rahmen des Projekts BFT-PersG und als äußerst nützlich für die Evaluierung des 3000-m-FET.

In einer Teilstudie dieser Arbeit wurde im Rahmen einer Laboruntersuchung der 3000-m-FET und drei weitere Ausdauertests (siehe Kapitel 5.1.2; Abbildung 24) mit dem Einsatz einer Spiroergometrie, durch die Abnahme des Blutlaktats, die Feststellung der maximalen Herzfrequenz und der Abfrage des subjektiven Belastungsempfindens trainingswissenschaftlich analysiert und auf Zusammenhänge hin untersucht. Die Teilnehmer dieser Studie absolvierten zum ersten Termin den 3000-m-FET und nach einer Pause im Anschluss einen Rampentest, der den Zweck hatte, dass sich die Versuchspersonen bis an ihre Leistungsgrenze ausbelasten konnten, um so die maximalen Atemgaswerte, speziell die maximale Sauerstoffaufnahme (VO_{2max}) und Blutlaktatwerte als Indikatoren der körperlichen Leistungsfähigkeit bzw. der Ausbelastung zu ermitteln und diese nachfolgend auf Zusammenhänge hin zu unter-

suchen. Bei dem zweiten Termin legten die Versuchspersonen den 1000-m-Lauf auf dem Laufband ab und bewältigten nach einer Erholungspause einen Rampentest auf dem Laufband. Für die beiden Fahrradergometertests (3000 Meter und Rampentest) wurden wie bei dem Gesundheitstag die Rennräder mit dem Tacx-Trainer eingesetzt, da sich der Einsatz dieser Gerätschaft als sehr nützlich und aufgrund von positiven Feedback der Teilnehmer am Gesundheitstag als besonders tauglich erwies. Die dazugehörige Software ermöglichte den Teilnehmern außerdem während der Testung ein Überblick über die zurückgelegte Strecke, die Zeitangabe, die Geschwindigkeit, die Trittfrequenz und die Leistung (siehe Abbildung 14), sodass die Pbn ihre Leistung im Laufe der Testung beobachten konnten. Dies wirkte sich v. a. positiv auf die Motivation sich auszubelasten und letztendlich auf die Leistung der Teilnehmer aus. Neben der Spiroergometrie und der Laktatabnahme wurde der Puls zum Zeitpunkt des Testendes festgehalten und die Pbn nach ihrem Anstrengungsempfinden anhand der Borg-Skala direkt nach dem Test befragt. Diese Studie stellt im Rahmen dieser Arbeit einen wesentlichen Kernpunkt dar, der zu einer fundierten trainingswissenschaftlichen Validierung des neu konzipierten Ausdauertests für die Personalgewinnung beiträgt. Das Kapitel 5.1.2 befasst sich im Detail mit der Laborstudie zur Validitäts- und Reliabilitätsüberprüfung des 3000-m-FET.

Parallel zu den Laboruntersuchungen am Department für Sportwissenschaft an der Universität der Bundeswehr München fanden im Zeitraum vom 15.05.2013 bis 15.12.2013 Testungen am KarriereCenter der Bundeswehr in Wilhelmshaven statt. Die Testungen beinhalteten ausschließlich den 3000-m-FET mit der Erfassung der Zeit bzw. der Testdauer und den Einsatz der Skalen zur psychischen Belastbarkeit (Lang- und Kurzform). Die Studie wurde vom Personal des KarriereCenters Wilhelmshaven begleitet und die Ergebnisse sowohl schriftlich als auch über eine Excel-Tabelle festgehalten und dokumentiert.

Des Weiteren fand am Standort der Bundeswehr in Feldafing eine zusätzliche Datenerhebung vom 08.01.2014 bis 06.02.2014 unter Einsatz der Tacx-Trainer und der oben beschriebenen Software sowie des Cannondale Fahrrads statt. Hierbei wurden Soldaten der Bundeswehr getestet und zur Beurteilung der physischen Fitness in Form der Ausdauerleistungsfähigkeit die Zeit für die Bewältigung der 3000-Meter-Strecke herangezogen. Im Zuge dieser Gelegenheit kamen wie bei den anderen Datenerhebungen die Fragebögen zur psychischen Belastbarkeit und darüber hinaus die Skala zur allgemeinen Selbstwirksamkeit (siehe Kapitel 4.3.3) zum Einsatz. Für die Untersuchungen der psycho-physischen Belastbarkeit bzw. zur

Beantwortung der Hypothesen H1, H2 und H3 (siehe Kapitel 4) wurden die Daten zu einer Gesamtstichprobe zusammengefügt. Eine genaue Darstellung der einzelnen Studien und Datenerhebungen, sowie die Beschreibung der Stichprobe kann Kapitel 5.3 entnommen werden.

5.1.1 Explorationsstudie zur Kriteriendefinition des Testverfahrens zur physischen Belastbarkeit

Der Laborstudie zur Überprüfung der Gütekriterien des Fahrradergometertests ging eine explorative Untersuchung voraus, bei der der angemessene Widerstand für die Hauptuntersuchung zum neu zu konzipierenden Fahrradergometertest eruiert werden sollte. Hierzu wurde ein Vergleich anhand von acht Versuchsteilnehmern zwischen der benötigten Zeit beim 1000-m-Lauf und der zurückgelegten Strecke bei einem 6-minütigen Fahrradergometertest durchgeführt. Die Pbn sollten eine möglichst weite Strecke während der sechs Minuten bei einem konstanten Widerstand (Wattstufe) zurücklegen, d. h. während der sechs Minuten Belastung wurde eine Wattstufe gefahren, die bis zum Ende für den jeweiligen Pbn gleich blieb. Während des Lauf- und des Fahrradergometertests wurde der Puls mit Pulsfrequenzuhren der Firma Polar erfasst. Der Lauftest wurde auf einer 400-Meter-Leichtathletikbahn durchgeführt. Nach abgelaufener Zeit wurde die Distanz schriftlich festgehalten. Während des Laufs wurde an einer bestimmten Stelle der 400-m-Bahn die Pulsfrequenz von den Pbn per Zuruf ohne Unterbrechung der Belastung an die Versuchsleiter weitergegeben. Bei dem Fahrradergometertest waren die Herzfrequenzgurte an die Ergometer gekoppelt, welche wiederum mit einem Laptop verbunden waren. Eine extra für die Ergometer der Firma ERGO-Fit programmierte Software von B.I.T. Soft (Health Center Software 1.4.42) erfasste Herzfrequenzdaten und die Trittfrequenz. Die gefahrene Distanz der Testpersonen wurde auf dem Display des Ergometers während des Tests verfolgt und nach Ablauf der Zeit dem Pbn die Testbeendigung signalisiert und schriftlich festgehalten.

Ein weiteres Ziel der Voruntersuchungen war es, die Software in Zusammenarbeit mit der Herstellerfirma für das spätere Testdesign so zu modifizieren bzw. anzupassen, dass der Test nach der erreichten Distanz automatisch zu Ende ist und direkt in die „cool down-Phase“ übergeht, ohne dass der Versuchsleiter die Distanzangabe auf dem Ergometer-Display manuell abzulesen hat. In diesem Teil der Vorstudie wurden die acht Untersuchungsteilnehmer mit verschiedenen Wattstufen zwischen 110 Watt und 200 Watt getestet. Tabelle 3 zeigt den Ab-

bruch des Tests der Pb Nr. 6 mit einer Wattzahl von 200 Watt. Demnach kann die Belastungsanforderung in diesem Fall als zu hoch angesehen werden. Des Weiteren fiel auf, dass die Teilnehmer bei den niedrigeren Wattstufen hohe durchschnittliche Trittfrequenzen (127, 129, 139, 142, 147) aufwiesen. Aus entsprechender Literatur (siehe z. B. Graf et al., 2001) geht hervor, dass 80 Umdrehungen pro Minute als sinnvolle Richtlinie gelten. Dennoch hatte der selbe Untersuchungsteilnehmer Schwierigkeiten, bei einer reduzierten Wattzahl von 150 Watt die Trittfrequenz (\approx 80 Umdrehungen pro Min.) konstant zu halten. Es erschien sinnvoll, auf Grundlage der Befunde der explorativen Studie zur Eruierung eines angemessenen Widerstandes für die weitere Untersuchung einen Widerstand von 130 Watt und eine zu fahrende Strecke von 3000 Meter zu wählen, die die Testpersonen innerhalb eines Zeitrahmens von 6:30 Min., wie auch beim 1000-m-Lauf bewältigen sollen.

Tabelle 3

Ergebnisse der Explorationsstudie (Teil I) zur Konzeption und Evaluierung eines Fahrradergometertests für die Personalgewinnung der Bundeswehr. Vergleich eines 1000-m-Laufs mit einem 6-minütigen Fahrradergometertest (N = 8).

Vpn	Test	Wattleistung	Zeit (in Min.)	Distanz (in Metern)	Max. Pulsfrequenz	Durchschn. Trittfrequenz
1	1000-m-Lauf		03:13		177	
	6-min. FET	160	06:00	4800	177	142
2	1000-m-Lauf		03:16		187	
	6-min. FET	160	06:00	4600	175	139
	6-min. FET	110	06:00	4800	182	147
3	1000-m-Lauf		03:49		177	
	6-min. FET	140	06:00	4000	157	125
4	1000-m-Lauf		04:38		181	
	6-min. FET	150	06:00	4200	156	119
5	1000-m-Lauf		05:51		179	
	6-min. FET	130	06:00	3200	170	99
6	1000-m-Lauf		05:51		162	
	6-min. FET	200	06:00	1000	166	78
	6-min. FET	150	06:00	2600	163	84
7	1000-m-Lauf		05:01		186	
	6-min. FET	130	06:00	4200	177	129
8	1000-m-Lauf		04:49		191	
	6-min. FET	140	06:00	4200	177	127

An dieser Stelle sei angemerkt, dass zu diesem Zeitpunkt der Testentwicklung davon auszugehen war, dass die Realisierung bzw. die Datenerhebung der gesamten Untersuchung mit den (drehkurbelunabhängigen) Ergometern der Firma ERGO-Fit, sowie der entsprechenden Software durchzuführen sei. Insofern bestand zu diesem Zeitpunkt der Testkonzeption die Herausforderung, das Kriterium der vorzugebenen Wattstufe im Rahmen des Streckentests

auszuloten und zu fixieren. Eine große Schwierigkeit bestand hierbei v. a. in den Vorgaben des Auftraggebers beiden Geschlechtern das gleiche Belastungsprotokoll vorzugeben und lediglich den weiblichen Personen, im nachhinein, bei der Einordnung in die Normwerttabelle einen Geschlechterbonus zu gewähren.

Nach Graf et al. (2001) berechnet sich die untere maximale Herzfrequenz (Schläge/Min.) für Fahrradergometriebelastungen mit der Faustformel 200 Schläge/Min. minus dem Lebensalter in Jahren. Bei Laufbelastungen liegt diese nach Rost um etwa 10 Schlägen/Min. höher (2001, S. 57). Graf et al. führen weiter an, dass die so festgestellte absolute Leistungsfähigkeit eines schweren und längeren Menschen größer ist, als die eines kleineren und leichteren Menschen und muss von daher auf die Körperdimensionen normiert werden. Die maximale Leistungsfähigkeit der Frau liegt insbesondere aufgrund der geringeren Muskelmasse im Mittel etwa 20 Prozent unter der des Mannes und kann nach Graf et al. (2001, S. 58) wie folgt beurteilt werden: „Die maximale Soll-Leistung beträgt für den Mann 3 Watt pro kg Körpergewicht minus 10 Prozent für jede Lebensdekade jenseits des 30. Lebensjahres. Die maximale Soll-Leistung beträgt für die Frau 2,5 Watt pro kg Körpergewicht minus 8 Prozent für jede Lebensdekade jenseits des 30. Lebensjahres.“

Insgesamt konnten für den zweiten Teil der Explorationsstudie der Testkriterien 50 Versuchsteilnehmer gewonnen werden. Alle waren zum Zeitpunkt der Untersuchung Studierende der Universität der Bundeswehr München. Der weibliche Anteil der getesteten Personen belief sich auf 22 Prozent ($n = 11$) und der männliche Anteil auf 78 Prozent ($n = 39$).

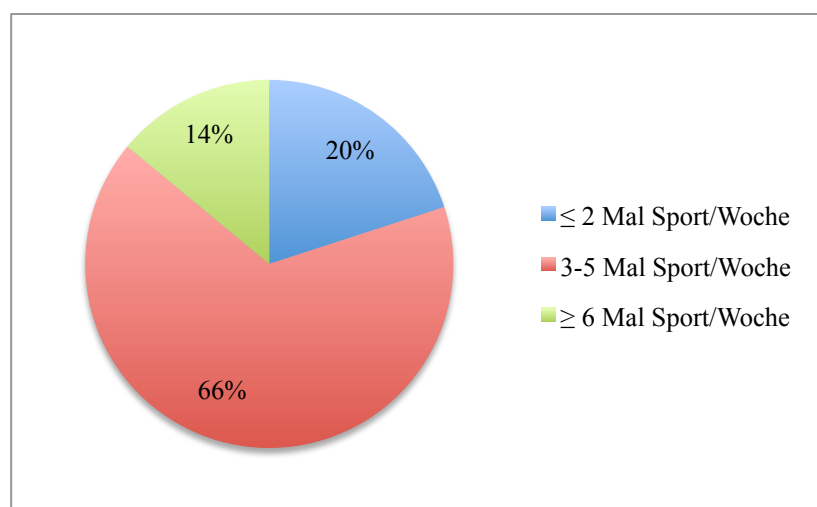


Abbildung 16. Sporteinheiten der Pbn pro Woche (N = 50).

Die Pbn haben zu 66 Prozent angegeben (siehe Abbildung 16), dass sie sich zwischen 3 und 5 Mal pro Woche sportlich betätigen, 20 Prozent machen ein bis zwei Mal die Woche oder gar keinen Sport. Da es sich zum größten Teil um Studierende der Sportwissenschaft handelte, gab es sogar einen Anteil an Studierenden, die mehr als sechs Sporteinheiten pro Woche (14 %) absolvieren und somit ausgesprochen sportlich aktiv sind.

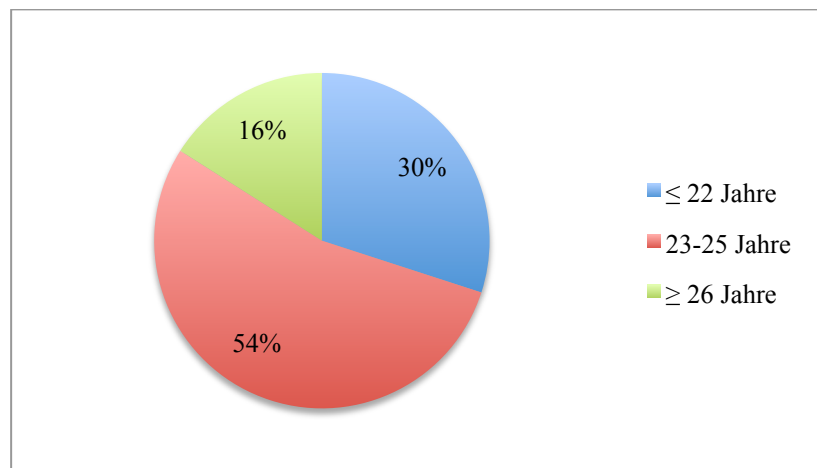


Abbildung 17. Altersverteilung der Teilnehmer der Studie zur Testentwicklung (N = 50).

Abbildung 17 zeigt, dass die Altersverteilung zu 54 Prozent (27 Pbn) zwischen 23 und 25 Jahren liegt. 30 Prozent der Pbn waren unter 22 Jahren (15 Pbn) und acht Pbn (16 %) waren mindestens 26 Jahre alt. Der älteste Teilnehmer der Studie war 27 Jahre und der jüngste war 18 Jahre alt.

Tabelle 4

Personen- und Leistungsdaten ($N = 50$).

Variable	Stichprobe									
	Gesamt ($N = 50$)		Frauen ($n = 11$)		Männer ($n = 39$)		Kürzeste Zeiten ($n = 10$)		Längste Zeiten ($n = 10$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Alter	23,6	1,9	23,6	2,0	23,8	1,6	23,1	2,4	23,9	2,0
Sport / Woche	1,9	0,6	1,9	0,6	1,9	0,5	2,2	0,8	1,7	0,5
Zeit (in Min.): 1000-m-Lauf	4:18	0:46	5:06	0:55	3:55	0:27	3:21	0:10	5:21	0:44
Werte der Borg-Skala	15,5	2,7	15,5	2,7	15,5	2,8	16,8	2,1	15,7	2,9
Strecke (in Metern): 6:30-Min.-FET	3890	764	3118	817	4107	596	4430	499	3060	684
Werte der Borg-Skala	16,1	2,1	16,5	2,5	16,0	2,0	16,7	1,6	16,6	2,5

Die Pbn absolvierten die beiden Ausdauerests im Rahmen des BFT an der Universität der Bundeswehr München. Neben der Ausdauerdisziplin, haben die Soldaten der Bundeswehr die Aufgabe zusätzlich einen Krafttest (Klimmhang) sowie einen Geschicklichkeits- und Schnelligkeitstest (Pendellauf) zu absolvieren. Da diese beiden Disziplinen für die Testentwicklung keine Relevanz darstellen, bleiben diese im weiteren Verlauf unberücksichtigt.

Tabelle 4 zeigt ein Durchschnittsalter von 23 Jahren und eine wöchentliche sportliche Betätigung zwischen 1,7 und 2,2. Die zehn Pbn, die bei dem 1000-m-Lauf am besten abschnitten, treiben mehr als zwei Mal pro Woche Sport.

Mittels der Borg-Skala konnte ein Unterschied beim subjektiven Anstrengungsempfinden zwischen den beiden ausdauerorientierten Fitnessstests festgestellt werden. Nach Beendigung des Ergometerests war das Anstrengungsempfinden der Frauen mit einem durchschnittlichen Wert der Borg-Skala von 16,5 etwas höher als das der Männer ($M = 16,0$). Am subjektiv angestrengtesten waren die besten zehn Pbn nach dem Lauf- und Fahrradergometerest mit Werten von 16,8 (1000-m-Lauf) und 16,7 (6:30-Min.-FET) auf der Borg-Skala. Die Pbn, die bei beiden Tests am schlechtesten abschnitten, fühlten sich mit einem Wert in Höhe von 16,6

anhand der Borg-Skala nahe zu ausbelastet (6:30-Min.-FET). Nach dem Laufen gaben die schlechtesten 10 Pbn im Mittel einen Wert hinsichtlich des Belastungsempfindens von 15,7 an. Beim Fahrradergometertest haben die schlechtesten 10 Pbn des 1000-m-Laufs im Durchschnitt insgesamt eine Strecke von 3060 ($SD = 684$) Meter innerhalb der Zeitvorgabe von 6:30 Min. zurückgelegt. Der Durchschnitt aller 50 Pbn lag bei einem Wert von 3890 Metern ($SD = 764$). Geht man von einer festgelegten Strecke von 3000 Metern als Zielvorgabe für den Fahrradergometertest aus, dann hätten die Pbn im Durchschnitt 5:16 Min. für die Strecke gebraucht. Beim 1000-m-Lauf haben die Pbn also durchschnittlich 58 Sek. weniger für die vorgegebene Strecke gebraucht (Mittelwert des 1000-m-Laufs: 4:18 Min.).

Zusammenhang der Zeiten beim 1000-m-Lauf und dem 6:30-Min.-FET

Abbildung 18 zeigt den Zusammenhang zwischen den beiden BFT-Ausdauerdisziplinen, die von den Pbn beide jeweils einmal absolviert worden sind.

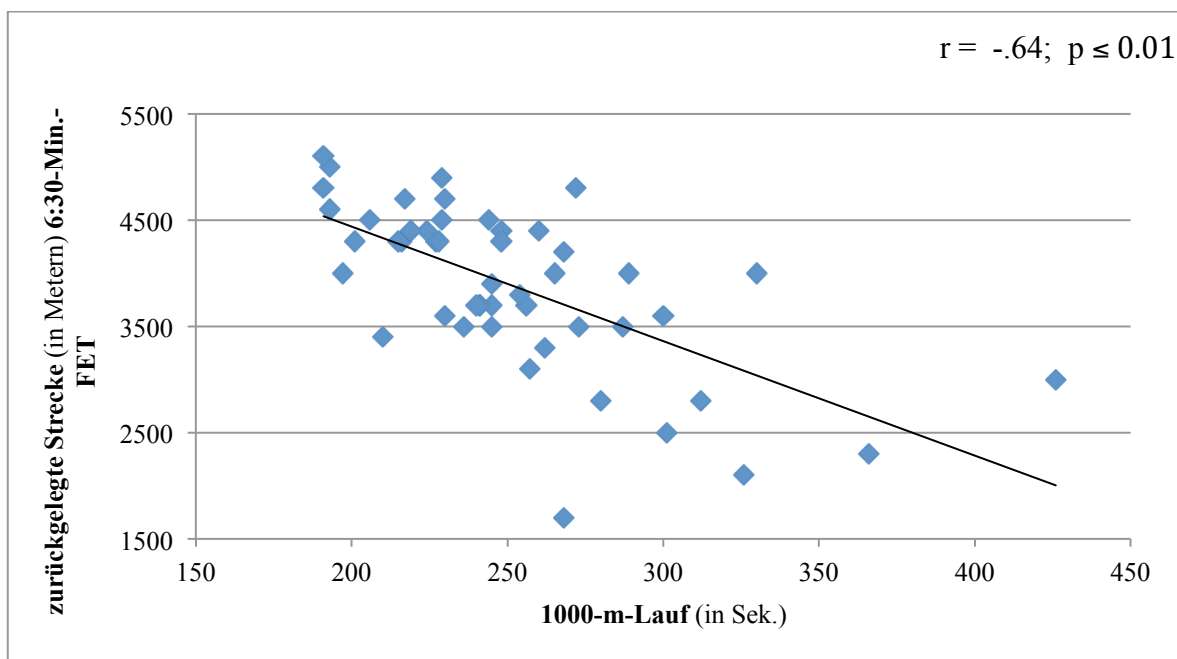


Abbildung 18. Zusammenhang des 1000-m-Laufs mit dem 6:30-Min.-FET (N = 50).

Die Berechnungen zeigen einen mittleren negativen Zusammenhang ($r = -.64$, $p \leq 0.01$) zwischen den Leistungen des 1000-m-Laufs und dem 6:30-Min.-FET (siehe Abbildung 18). Dadurch wird ersichtlich, dass die Leistung der Pbn (Zeit/Strecke) beim Lauftest mit der Leistung beim Fahrradergometer in etwa übereinstimmt.



Abbildung 19. Screenshot des B.I.T. Soft-Testsystems.

Die Abbildung 19 zeigt einen Ausschnitt aus der Ergometersoftware der Firma B.I.T. Soft und die Aufzeichnung der Trittfrequenz, sowie den Pulsverlauf eines Pbn während des Fahrradergometertests mit der Dauer von 6:30 Min. Des Weiteren lässt sich durch die Software die Pulskurve während des Testverlaufs entnehmen, sowie die durchschnittliche Trittfrequenz von 120 Umdr./Min. Die zurückgelegte Strecke wird nicht über die Software dargestellt. Der Pb sowie der Versuchsleiter können, wie in Abbildung 20, die Strecke am Display des Ergometers ablesen, wobei die Zahl hinter dem Komma die Hunderterschritte in Metern angibt.



Abbildung 20. ERGO-FIT Display.

Die Studie hat gezeigt, dass es einen mittleren negativen Zusammenhang ($r = -64$, $p \leq 0.01$) zwischen dem 1000-m-Lauf und dem zur Konzeption und Evaluation eingesetzten 6:30-Min.-FET im Hinblick auf die erbrachten Leistungen gibt. Der Tabelle 4 kann entnommen werden, dass die 50 Pbn im Durchschnitt 3890 Meter ($SD = 764$) in den vorgegebenen 6:30 Min. zurückgelegt haben. Für das Testdesign des neuen Fahrradergometertests im BFT liegt demnach eine Streckenvorgabe von 3000 Meter nahe. Dagegen spricht allerdings die Durchfallquote von 12,4 % der Versuchsteilnehmer. Besonders hervorzuheben ist dabei die Anzahl der Probandinnen, die die Distanz von 3000 Meter in der Zeit von 6:30 Min. nicht geschafft haben, während es beim 1000-m-Lauf nur eine Probandin gewesen ist, welche die zu laufende Strecke nicht in der Zeitvorgabe bewältigen konnte. Von den insgesamt zehn Teilnehmerinnen an der Studie, haben vier den Fahrradergometertest nicht bestanden. Dies kann allerdings auch an der vorgegebenen Wattzahl von 150 Watt gelegen haben, die auf Grund der Ergebnisse des ersten Teils der Explorationsstudie festgelegt wurde und evtl. zu hoch angesetzt ist. Speziell für weibliche Personen erscheint eine konstant zu bewältigende Wattstufe von 150 Watt zu hoch, was darüber hinaus auch das Anstrengungsempfinden der am schlechtesten abschneidenden Pbn, anhand der zum Einsatz gekommenen Borg-Skala, bestätigt. Die darunter fallenden zehn Pbn haben nach dem Fahrradergometertest eine Durchschnittsdistanz von 3060 Me-

tern ($SD = 684$) zurückgelegt und dabei im Mittel einen erhöhten Wert der Borg-Skala angegeben ($M = 16,6$; $SD = 2,5$). Im Hinblick auf das subjektive Belastungsempfinden im Anschluss der beiden Ausdauerdisziplinen gaben lediglich die zehn besten Versuchsteilnehmer mit $16,7$ ($SD = 1,6$) und $16,8$ ($SD = 2,1$; siehe Tabelle 4) höhere Werte anhand der Borg-Skala an. Darüber hinaus wurde erkennbar, dass die durchschnittlich berechnete Zeit für die 3000-Meter-Strecke mit 5:16 Min. fast eine Minute höher ist, als der Durchschnittswert beim 1000-m-Lauf. Dies lässt auf den ersten Blick entweder die vorgegebene Wattstufe von 150 Watt als zu schwer, oder die Wahl der zurückzulegenden Distanz (3000 Meter) als zu hoch erscheinen.

Für das finale Testdesign erscheint es sinnvoll, die Wattzahl für Frauen nach unten zu korrigieren. Dagegen spricht aber, dass der BFT, wie bereits erwähnt, einheitlich von Männern und Frauen unter gleichen Voraussetzungen durchgeführt werden sollte, um die Durchführung möglichst einfach und konsistent zu gestalten und anschließend die erbrachte Leistung in eine Normwerttabelle einzuordnen, welche zwischen Mann und Frau und Alter differenziert bzw. einen Frauenbonus gewährt. Alternativ wäre in Erwägung zu ziehen, die Wattzahl (Gang) bzw. den Widerstand durch die Pbn während der Testphase individuell bestimmen zu lassen, wobei sich über eine zu programmierende Formel von der entsprechenden Firma die Geschwindigkeit über die Wattstufe, welche den Gang eines drehkurbelabhängigen Rads simuliert und über die Trittfrequenz berechnet. Die Umsetzung dieser spezifischen Testkriterien bzw. die Erweiterungen der Einstellungsoptionen konnte mit den gegebenen Gerätschaften und der vorprogrammierten Ergometersoftware der Firma B.I.T. Soft nicht ermöglicht werden. Die Software bietet diverse voreingestellte Testdesigns, darunter einige für den medizinischen Bereich geeignete und bekannte Tests, die von dem Benutzer bzw. der benutzenden Institution nicht selbstständig modifiziert werden können. Somit stellte sich die Ergometersoftware als stark limitierender Faktor bezüglich der gewünschten Testkriterien und für die Umsetzung eines geeigneten Testdesigns unter den Vorgaben der Bundeswehr für den BFT heraus. Eine weitere Überlegung dazu, die Problematik des unterschiedlich geeigneten Widerstandes in den Griff zu bekommen, wäre die Wattzahl einmal für 120 Watt und 150 Watt von Pbn selbstständig zu Beginn des Tests (während der „Warm-up-Phase“) wählen zu lassen. Dadurch wäre die Möglichkeit gegeben die Pbn je nach Gewicht, Größe und Fitnesslevel den individuell am geeignetsten erachteten Widerstand zu bestimmen. Dabei könnte allerdings das Problem der Über- oder Unterschätzung der eigenen Leistungsfähigkeit auftreten. Der Pbn ist evtl. nicht dazu in der Lage, innerhalb der kurzen Warm-up-Phase den für sich angemessenen Widerstand (120 Watt oder 150 Watt) über die Distanz von 3000 Metern festzulegen. Eine

Variante, die bereits zu Beginn (s. o.) diskutiert wurde, ist den optimalen Widerstand über die Formel Watt/kg zu berechnen. Dieses Vorgehen würde dann von der gewünschten Strukturähnlichkeit zum bisher durchgeführten 1000-m-Lauf abweichen, da beim 1000-m-Lauf auch auf Grundlage der Körperkonstitution Testpersonen eine solche Differenzierung außer Acht gelassen wurde.

Aufgrund der bereits angesprochenen Softwareproblematik gab es bei der Untersuchung diverse Faktoren, die die Datenerhebung erschwerten. So war die Durchführung ohne Pulsgurt auf dem Fahrradergometer in Verbindung mit der Softwarenutzung und der damit verbundenen Aufzeichnung der Trittfrequenz nicht möglich. Der Widerstand in der Einfahrphase konnte nicht verändert werden und lag lediglich bei 25 Watt, was von den Pbn zu gering eingeschätzt wurde. Ein weiterer Nachteil ergab sich durch den ohne Pause kaum zu bewerkstellende Wechsel der Wattzahl zwischen der Warm-up bzw. „Vertrautmachungs-Phase“ mit dem Ergometer und der eigentlichen Testphase von 25 Watt auf die gewünschten 150 Watt. Erst wenn das Fahrradergometer den Puls des Pbn erkannt hat, kann das Programm auf die vorgesehenen 150 Watt erhöhen. Die Zeitdifferenz vom Startpunkt (Beginn der Zeitmessung) des Tests bis zur Erhöhung auf die 150 Wattleistung betrug mehrere Sekunden. Ebenso sind die angezeigten Meter im Programm nur anhand von 100 Meter Sprüngen im Display (siehe Abbildung 20) zu erkennen. Angesichts dieser Problematik und Ungenauigkeit bzw. Zeitdifferenz bis zur Erhöhung auf die vorgesehene Wattstufe, hat dies zu einer Messwertdifferenz bei jeder Testung zwischen 100 und 400 Metern geführt. Während der Testentwicklung wurde mit den Firmen Ergo Fit und B.I.T. Soft zusammengearbeitet, um die gewünschten Softwaremodifikationen auszuloten und zu realisieren, um die genannten Messdifferenzen in den Griff zu bekommen.

Festzuhalten bleiben folgende Hinweise und Konsequenzen: Eine erste konkrete Testkonzeption und -konfiguration konnte entwickelt und durch die Studie eine Strukturähnlichkeit zwischen den beiden Tests festgestellt werden. Es konnte insbesondere ein Zusammenhang in Hinblick auf die erzielten Leistungen zwischen dem neu konzipierten Fahrradergometertests und dem bisher im Rahmen des BFT durchgeführten 1000-m-Lauf nachgewiesen werden. Auf Grundlage dieser Befunde können weiterführenden Untersuchungen durchgeführt und mit dem hier konzipierten Fahrradergometertest weitere Daten erhoben werden um eventuelle Modifikationen am Testdesign vorzunehmen. Zunächst scheint es jedenfalls auf Grundlage der Ergebnisse legitim, den, unter Berücksichtigung der Vorgaben seitens der Bundeswehr,

neu konzipierten Test für die Personalrekrutierung einzusetzen und später für die Normierung relevante Datensätze zu generieren. Die vorgegebenen 3000 Meter bleiben zunächst, gleichermaßen wie die Zeitvorgabe von 6:30 Min., die zu erfüllenden Testkriterien für den weiter zu erprobenden Fahrradergometertest. Die diskutierte Wattzahl wird hingegen aufgrund der Befundlage für weitere Testungen mit dem ERGO-Fit-Ergometer auf 130 Watt reduziert und konnte durch den späteren Einsatz der Tacx-Trainer und die Möglichkeit einen individuellen Gang zu wählen, eliminiert werden. Im Anschluss an Folgeuntersuchungen soll dann eine Normwerttabelle für den Fahrradergometertest erarbeitet werden, die eine Einordnung der erbrachten Leistung der getesteten Personen ermöglicht. Die Normwerttabelle soll eine Einordnung des individuellen Testergebnisses ermöglichen und kann somit als Bezugssystem gesehen werden, woraus die prüfende Person relativ einfach eine Punktzahl für die erbrachte Leistung entnehmen kann, die dann mit der Punktzahl aus den Normwerttabellen der anderen Tests verrechnet wird und somit letztendlich einen Tauglichkeitsmaßstab darstellt (siehe Heck & Schulz, 1998). Im folgenden Kapitel wird der Fahrradergometertest mit den hier ausgearbeiteten Testkriterien anhand der klassischen Testkriterien evaluiert und mit anderen Ausdauertests verglichen.

5.1.2 Laborstudie zur Überprüfung der Gütekriterien des 3000-Meter-Fahrradergometertests

Der konzipierte 3000-Meter-Fahrradergometertest wurde in einer Laborstudie durch die Anwendung einer Leistungsdiagnostik mit Hilfe spiroergometrischer Analysen an $N = 145$ Pbn überprüft und auf seine Reliabilität und Validität hin getestet. Die Testkriterien bestehen darin, eine Strecke von 3000 Metern in kürzester Zeit und in maximal 6:30 Min. zu bewältigen. Für die Realisierung der definierten Testkriterien und der Testkonzeption wurden für die Validierungsstudie die Hometrainer der Firma Tacx mit der entsprechenden Software (Tacx-Trainer-Software, Basic 4) benutzt. Durch diese Gerätekonfiguration konnte ermöglicht werden, dass die Untersuchungsteilnehmer den Gang während der Testphase individuell wählen bzw. verstellen konnten.

Für die Überprüfung der Gütekriterien des neu entwickelten 3000-Meter-Fahrradergometertests, kurz 3000-m-FET, sind folgende Parameter, die Aufschluss über den Grad der

Belastung geben und die anschließend aufgelisteten Indikatoren für die Ausdauerleistungsfähigkeit von besonderer Relevanz:

Ausbelastungsparameter:

- Maximale Herzfrequenz
- Maximale Laktatkonzentration
- Werte der Borg-Skala
- Maximal erreichter Respiratorischer Quotient

Indikatoren für die Ausdauerleistungsfähigkeit bzw. für die allgemeine physische Fitness:

- Zeiten der Strecken- (3000-m-FET und 1000-m-Lauf) sowie der Rampentests (F.-Ergom. und Laufband), bzw. die Dauer der Tests
- Absolute und relative VO_{2max} bzw. VO_{2peak} -Werte
- Maximale Wattleistung des Rampentests (F.-Ergom.) und durchschnittliche Wattleistung des 3000-m-FET

Die Prämisse für die Aussagefähigkeit der Indikatoren der Ausdauerleistungsfähigkeit ist die Gewissheit, ob sich die Untersuchungsteilnehmer tatsächlich ausbelastet haben. Werden bestimmte Höhen bei den jeweiligen Ausbelastungsparametern erreicht, können die in dieser Arbeit relevanten Indikatoren für die Ausdauerleistungsfähigkeit bzw. für die physische Belastbarkeit zum einen für die Überprüfung der Gütekriterien sowie auch für die Untersuchung der psycho-physischen Zusammenhänge (Kapitel 5.3) herangezogen werden

Bevor das Design und die Methodik dargestellt werden, soll noch auf den durchaus komplexen Charakter der Studie, die grundlegenden Überlegungen und Annahmen eingegangen werden, sodass das Anliegen und der Ansatz, nämlich eine trainingswissenschaftlich fundierte Überprüfung des 3000-m-FET hinsichtlich der Validität und Reliabilität, deutlich wird. Mit der Überprüfung der Gütekriterien soll gewährleistet werden, dass die Erfassung der Ausdauerleistungsfähigkeit durch den 3000-m-FET für die Folgeuntersuchungen gegeben ist. Im Kontext der psycho-physischen Belastbarkeit stellt der 3000-m-FET das Tool zur physischen Leistungsbewertung dar. Ein Testverfahren, worüber die physische Komponente des psycho-physischen Belastbarkeits-Konzepts operationalisiert ist. Infolgedessen bedarf es einer Überprüfung der Testgütekriterien, die primär Aufschluss darüber geben soll, ob der 3000-m-FET die Ausdauerleistung im Vergleich zu anderen Tests adäquat widerspiegelt. Darüber hinaus

werden mit dieser Laborstudie auch weitere Parameter der körperlichen Leistungsfähigkeit erfasst und mit den Werten der Persönlichkeitsskalen korreliert, die Einflüsse dieser Parameter auf die psychische Belastbarkeit analysiert, sowie anhand der Leistungsindikatoren Gruppen definiert und diese dann auf Unterschiede hinsichtlich der psychischen Belastbarkeit überprüft. In dieser Laborstudie werden neben dem Leistungsparameter „Zeit“ aus anderen Datenerhebungen, auch weitere Indikatoren der körperlichen Leistungsfähigkeit (bspw. VO_{2max} , maximale Wattleistung) für die Untersuchungen der psycho-physischen Belastbarkeit herangezogen.

Für diese Laborstudie zur Überprüfung der Gütekriterien des neu konzipierten 3000-m-FET ergeben sich folgende Fragestellungen und Hypothesen:

1. Gibt es Zusammenhänge zwischen dem neu konzipierten 3000-m-FET und dem bisher im Rahmen der körperlichen Leistungsfeststellung der Bundeswehr eingesetzten 1000-m-Lauf, sowie anderen Ausdauerdiagnostiken in Form von Rampentests auf dem Fahrradergometer und auf dem Laufband?

Durch die Laborstudie soll anhand von Kennwerten, insbesondere der VO_{2max} , festgestellt werden, ob der neu konzipierte 3000-m-FET ein geeigneter Test für die Feststellung der Ausdauerleistungsfähigkeit im Vergleich zu anderen Ausdauer tests darstellt. Für diese Überprüfung wird der im Rahmen der Feststellung der Basis Fitness (Basis-Fitness-Test = BFT) in der Bundeswehr eingesetzte 1000-m-Lauf, sowie zwei Rampentests (Laufband und Fahrradergometer) herangezogen. Die Rampentests und die dafür verwendeten Belastungsprotokolle stellen ein geeignetes Testverfahren dar, um die Maximalwerte der Sauerstoffaufnahme, Herzfrequenz, usw. zu ermitteln. Die über die Spiroergometrie ermittelten Atemgase, sowie maximalen Laktat- und Herzfrequenzwerte und Werte der Borg-Skala der vier Ausdauer testverfahren, die im Rahmen der Laborstudie eingesetzt wurden, sollen auf Zusammenhänge hin überprüft werden. Dabei ist bei dem Wert VO_{2max} , der in dieser Untersuchung als wesentlicher leistungsdiagnostischer Parameter angesehen werden kann, eine hohe Korrelation zu erwarten. Außerdem ist bei der Zeit, die für die jeweiligen Tests benötigt wurde, ebenfalls eine hohe Korrelation zu erwarten. Bei der Überprüfung der Ausdauerleistungsfähigkeit sollte demnach eine gut ausdauertrainierte Person, die bei dem Fahrradergometertest die 3000 Meter schnell zurücklegt, auch die 1000 Meter auf dem Laufband verhältnismäßig schnell bewältigen und bei den Rampentests die Belastung relativ lange durchhalten, was sich somit in einer hohen

negativen Korrelation niederschlägt. Aufgrund der angeführten Überlegungen und Annahmen wird folgende Hypothese formuliert:

(H4) Die ermittelte VO_{2max} und die benötigte Zeit bei dem 3000-m-FET korrelieren in hohem Maße mit den VO_{2max} -Werten und Zeiten der anderen Ausdauerests dieser Studie.

Die zusätzlich erhobenen Werte der vier Tests werden, wie bereits erwähnt, ebenfalls verglichen und auf Zusammenhänge hin geprüft. Der Respiratorische Quotient (RQ), die Laktatwerte, die Herzfrequenz-Werte und der Wert der Borg-Skala, die am Ende jeder Testung erfasst wurden, werden hierbei als objektive und als subjektiver Ausbelastungsparameter herangezogen, um zu analysieren, ob und wie stark sich die Pbn ausbelastet haben.

Insgesamt werden von jedem der Testpersonen vier Tests absolviert:

- 3000-m-FET
- Rampentest-Fahrradergometer
- 1000-m-Lauf (Laufband)
- Rampentest-Laufband

Die oben genannten Kennwerte werden herangezogen, um speziell den 3000-m-FET mit jedem der anderen Testverfahren hinsichtlich der Validität zu vergleichen und v. a. eine Aussage darüber treffen zu können, inwieweit der neu konzipierte Test das Kriterium Ausdauerleistungsfähigkeit tatsächlich misst.

- 3000-m-FET - 1000-m-Lauf
- 3000-m-FET - Rampentest-Fahrradergometer
- 3000-m-FET - Rampentest-Laufband.

Aufgrund der unterschiedlichen Belastungsprotokolle und der Art der Belastung (Laufband und Fahrradergometer) sind hauptsächlich bei den maximalen Sauerstoffaufnahme-Werten und teilweise bei den benötigten Zeiten bzw. der Testdauer Korrelationen im hohen Bereich zu erwarten. Dies würde bedeuten, dass der neu konzipierte 3000-m-FET ein Testverfahren darstellt, welches die Ausdauerleistungsfähigkeit aussagekräftig erfasst. Bei den zusätzlich

erhobenen Parametern sind Zusammenhänge in nicht ganz so hohen Bereichen zu vermuten. Trotz der zu erwartenden Zusammenhänge zwischen den Tests werden aufgrund der unterschiedlichen Belastung signifikante Unterschiede bei den Mittelwerten angenommen. Diese Annahme basiert auf der Tatsache, dass bei Fahrradergometerbelastungen i. d. R. bis zu 10 Prozent niedrigere absolute VO_{2max} -Werte oder geringe maximale Herzfrequenzen aufgrund der unterschiedlichen Belastungsart und der beteiligten Muskelmasse erreicht werden und es, außer bei trainierten Radsportlern, nicht zu einer Ausbelastung des Herz-Kreislauf-Systems kommt (siehe hierzu bspw. auch Holmann & Strüder, 2009, S. 344 ff.; aus der Fünten et al., 2013).

Neben der Validitätsüberprüfung soll der 3000-m-FET im Rahmen dieser Studie auch auf seine Reliabilität hin untersucht werden. Dafür werden ebenfalls die über die Spiroergometrie generierten Parameter, sowie die Laktat-, Herzfrequenz- und Borg-Werte herangezogen. Hierfür absolvierten 29 Pbn die vier Tests zu einem späteren Zeitpunkt noch einmal. Dabei handelt es sich um einen Retest (siehe Lienert und Raatz, 1998, S.9). Hinsichtlich der Reliabilität ergibt sich folgende Fragestellung sowie Hypothese, der im weiteren Verlauf nachgegangen werden soll:

2. Gibt es im Hinblick der Maximalwerte (Zeit, VO_{2max} , Laktat, Herzfrequenz, Werte der Borg-Skala und Atemminutenvolumen) der vier Ausdauerests einen Zusammenhang bei erneuter Durchführung zu einem späteren Zeitpunkt?

(H5) Bei erneuter Durchführung der vier Ausdauerests zu einem späteren Zeitpunkt korrelieren die erhobenen Parameter (Zeit und VO_{2max}) in hohem Maße mit den Werten aus der ersten Testung.

Auch diese Hypothese soll durch die Betrachtung der Kennwerte wie Herzfrequenz, Atemminutenvolumen, Laktatkonzentration und Werte der Borg-Skala ergänzt werden. Es ergeben sich die folgenden Zusammenhangsuntersuchungen (Retest-Reliabilität):

- 3000-m-FET - 3000-m-FET
- 1000-m-Lauf - 1000-m-Lauf
- Rampentest Fahrradergometer - Rampentest Fahrradergometer
- Rampentest Laufband - Rampentest Laufband

Die im Folgenden abzuhandelnden Fragestellungen beziehen sich insbesondere auf die Gütekriterien, auf die zunächst noch kurz eingegangen wird.

Diagnostische Grundlagen

Die Laborstudie zur Überprüfung der Gütekriterien des 3000-m-FET ist ein wesentlicher Bestandteil dieser Arbeit. Das wesentliche Anliegen dieser Studie besteht darin, den entwickelten 3000-m-FET im Sinne einer Evaluation mit anderen Ausdauertests zu vergleichen. Zum einen gilt es im Rahmen eines Projekts zur Methodenentwicklung für die Personalauswahl in der Bundeswehr den dazu bisher eingesetzten 1000-m-Lauf durch einen neuen Fahrradergometertest zu ersetzen, der adäquate Messergebnisse liefert. Zum anderen sollte eine Überprüfung der kriterienbezogenen Validität des dafür konzipierten 3000-m-FET mit in der sportwissenschaftlich-trainingswissenschaftlichen Leistungsdiagnostik gängigen Tests, dies sind die sog. Rampentests, durchgeführt werden. Im Speziellen ist herauszustellen, wie stark der 3000-m-FET die Ausdauerleistung tatsächlich misst, inwiefern also dieses Verfahren das misst, was es messen soll. Die Ergebnisse dieser Untersuchung sollten die angemessene Erfassung der Daten der physischen Komponente mit dem 3000-m-FET absichern. Die Entwicklung und Validierung des 3000-m-FET ist die Grundlage für die Datengenerierung in weiteren Studien zur Erfassung der psycho-physischen Belastbarkeit. Ohne den Nachweis der kriterienbezogenen Validität wäre ein Einsatz dieses Ergometertests für eine breit angelegte Datensammlung ungeeignet und würde die methodische Herangehensweise zur Erfassung der physischen Komponente im Rahmen der psycho-physischen Belastbarkeit in Frage stellen. Aufgrund dessen soll der 3000-m-FET auf folgende Gütekriterien hin geprüft werden.

Objektivität

Die Objektivität eines Tests setzt voraus, dass die Ergebnisse des Testverfahrens unabhängig von dem Untersucher und/oder Auswerter sind. Objektivität wird also erreicht, wenn verschiedene Untersucher bei einem Test mit demselben Pbn zu gleichen Ergebnissen gelangen. In diesem Fall spricht man von einer interpersonellen Übereinstimmung der Untersucher (Lienert & Raatz, 1994, S. 7). Es können drei verschiedene Aspekte der Objektivität unterschied-

den werden. Die Durchführungsobjektivität beschreibt den Grad der Unabhängigkeit der Testergebnisse von zufälligen oder systematischen Verhaltensmustern der Untersucher (Lienert & Raatz, 1994, S. 8). Damit die Durchführungsobjektivität möglichst stark ausgeprägt ist, sollten Instruktionen wenn möglich schriftlich festgelegt und die Testsituation standardisiert sein. Das bedeutet, dass soziale Interaktionen zwischen Versuchsleiter und Pb auf ein Minimum reduziert sein sollten. Den Untersuchungsteilnehmern wurde seitens der Versuchsleiter nach einem festgelegten Schema eine Testinstruktion gegeben bzw. die Aufgabenstellung erläutert, während der Testung wurden dem Pb in regelmäßigen Abständen die zurückgelegte Distanz und die benötigte Zeit mündlich mitgeteilt und weitere Interaktionen mit dem Pbn vermieden. Heck und Schulz (1998, S. 14) erwähnen, dass gerade sogenannte Maximal-Tests, also Ausbelastungstest, wie sie in dieser Studie genutzt werden, „anfällig gegenüber motivationsbeeinflussenden Verhalten des Untersuchers [sind].“ Die Auswertungsobjektivität gewährleistet, dass die erhobenen Daten nach vorgegebenen Regeln ausgewertet werden. Insofern wird in dieser Arbeit bei den physischen Belastbarkeitsindikatoren auch von objektiv erhobenen Parametern gesprochen. Es werden unabhängig vom Auswerter immer die gleichen Auswertungsergebnisse erzielt (Lienert & Raatz, 1994, S. 8). Die computergestützte Auswertung garantiert einen hohen Grad an Auswertungsobjektivität, so wie es bei dem 3000-m-FET der Fall ist. Die Interpretationsobjektivität beschreibt den Grad der Unabhängigkeit der Interpretation des Testergebnisses von der interpretierenden Person (Lienert & Raatz, 1994, S. 8). Liefert die Auswertung eines Tests einen numerischen Wert, wie die maximale Sauerstoffaufnahme oder die Herzfrequenz, so ist die Interpretationsobjektivität nahezu vollkommen und lässt sich als trivial (Lienert & Raatz, 1994, S. 8) kennzeichnen.

Reliabilität

Eine Teil Anzahl ($n = 29$) der Versuchsteilnehmer konnte für die Überprüfung der Zuverlässigkeit und Messgenauigkeit der Ausdauerests bzw. für die Reliabilitätsprüfung gewonnen werden. Wie bereits erwähnt, nahmen diese Versuchspersonen jeweils an beiden Tests zwei Mal teil, verteilt auf insgesamt vier Termine, sodass ein gewisser zeitlicher Abstand auch unter dem Aspekt der Regeneration gewährt wurde. Nach Heck und Schulz (1998) ist die Reliabilität eines Ausdauerests insbesondere von der messmethodischen, sowie der individuellen biologischen Variabilität abhängig. Bei Ausbelastungstests kommt die (in diesem Zusammenhang) Störgröße der Leistungsbereitschaft bzw. der bereits erwähnte motivationale Aspekt

hinzu. „Die meßmethodische Variabilität setzt sich zusammen aus der Variation der Leistung der Belastungsgeräte und der Variation bei der Meßwertgewinnung. Bei der Laktatbestimmung wird neben dem apparativen Analysefehler die gesamte Kette der Blutabnahme-prozedur die Höhe des Analysefehlers mitbestimmen“ (Heck & Schulz, 1998, S. 16). Hier können bspw. zu langes Warten oder Komplikationen während der Blutabnahme über den Ohr-lappen des Versuchsteilnehmers bereits gravierende Auswirkungen auf die Messergebnisse haben und je nach Grad des Analysefehlers bringt dies eine Folgewirkung mit sich, die zu einer unzulänglichen Reliabilität des Tests führt (vgl. Heck & Schulz, 1998).

Validität

Lienert und Raatz, (1998) beschreiben einen Test als valide, wenn seine Ergebnisse einen fehlerfreien Rückschluss auf den Ausprägungsgrad des zu erfassenden Merkmals, hier die Ausdauerleistungsfähigkeit, zulassen. Nach den Autoren können drei Aspekte der Validität unterschieden werden. Die Inhaltsvalidität ist gegeben, wenn die Elemente des Tests so beschaffen sind, dass sie das Persönlichkeitsmerkmal am besten repräsentieren (Lienert & Raatz, 1998, S. 10). Lienert und Raatz (1998) beziehen sich aus einer psychologischen Perspektive zwar hauptsächlich auf Persönlichkeitsmerkmale, die Ausführungen lassen sich aber auf Tests auch in anderen Bereichen, z. B. sportwissenschaftliche oder trainingswissenschaftliche Tests zur Leistungsdiagnostik, anwenden. Ein anderer Aspekt wird als Konstruktvalidität bezeichnet. Diese gibt den Grad der Übereinstimmung der Testergebnisse mit der Wirklichkeit an. Dazu können die Messwerte mit Ergebnissen aus anderen Tests verglichen werden, die das gleiche Merkmal betreffen (Lienert & Raatz, 1998, S. 11)³. Unter dem vergleichbaren Aspekt der Konstruktvalidität wird in der vorliegenden Arbeit untersucht, ob ein 3000-m-FET zu gleichen Ergebnissen führt, wie z. B. der aktuell eingesetzte 1000-m-Lauf im Rahmen des BFT oder gängige Ausdauer-testverfahren wie Rampen oder Stufentests. Ergänzend dazu würde sich der dritte Aspekt der Validität, die Kriteriumsvalidität zeigen, wenn die erhobenen Daten mit Werten korrelieren, die gegenstandsbezogene Bereiche, in diesem Fall Ausdauerleistungen in konkreten Handlungsfeldern, betreffen (vgl. Lienert & Raatz, 1998, S. 11). Solche gegenstandsbezogenen Daten können beispielsweise Leistungen sein, die in Wettkämpfen erbracht wurden. Als eine in diesem Zusammenhang besonders relevante Über-

³ Im Zusammenhang der Konstruktvalidierung der hier relevanten Persönlichkeitsskalen wird auf die Aspekte der konvergenten und divergenten Konstruktvalidität in Kapitel 5.2.2 noch einmal gesondert und differenziert eingegangen.

prüfung der Validität bietet sich der Vergleich mit einem externen Kriterium an (kriteriumbezogene Validität),

„das in einem anderen Verfahren (Außenkriterium) gewonnen wurde und dessen Aussagefähigkeit als gesichert angenommen werden kann (äußere Validierung). Zum Beispiel kann die Durchschnittsgeschwindigkeit oder Laufzeit eines Marathonlaufs, die Zeit für die zu bewältigende Strecke, oder die relative Leistung (Watt/kg) als valides Außenkriterium angenommen werden, wenn man die allgemeine aerobe dynamische Ausdauer als die Ermüdungswiderstandsfähigkeit definiert“ (Heck & Schulz, 1998, S. 18).

Die Durchführung der Validitätsprüfung erfolgt dann (vgl. Heck & Schulz, 1998) mittels Korrelation der objektiven Kriteriumsleistungen (durchschnittliche Geschwindigkeit, durchschnittliche Leistung in Watt) mit Testwerten, wie bspw. der maximalen Sauerstoffaufnahme-kapazität (VO_{2max}). So wird es in der vorliegenden Studie realisiert.

Zwischen den Hauptgütekriterien bestehen Zusammenhänge, die von Lienert und Raatz (1998, S. 13) mit der nachfolgenden Abbildung 21 in einem Beziehungsgefüge dargestellt sind.

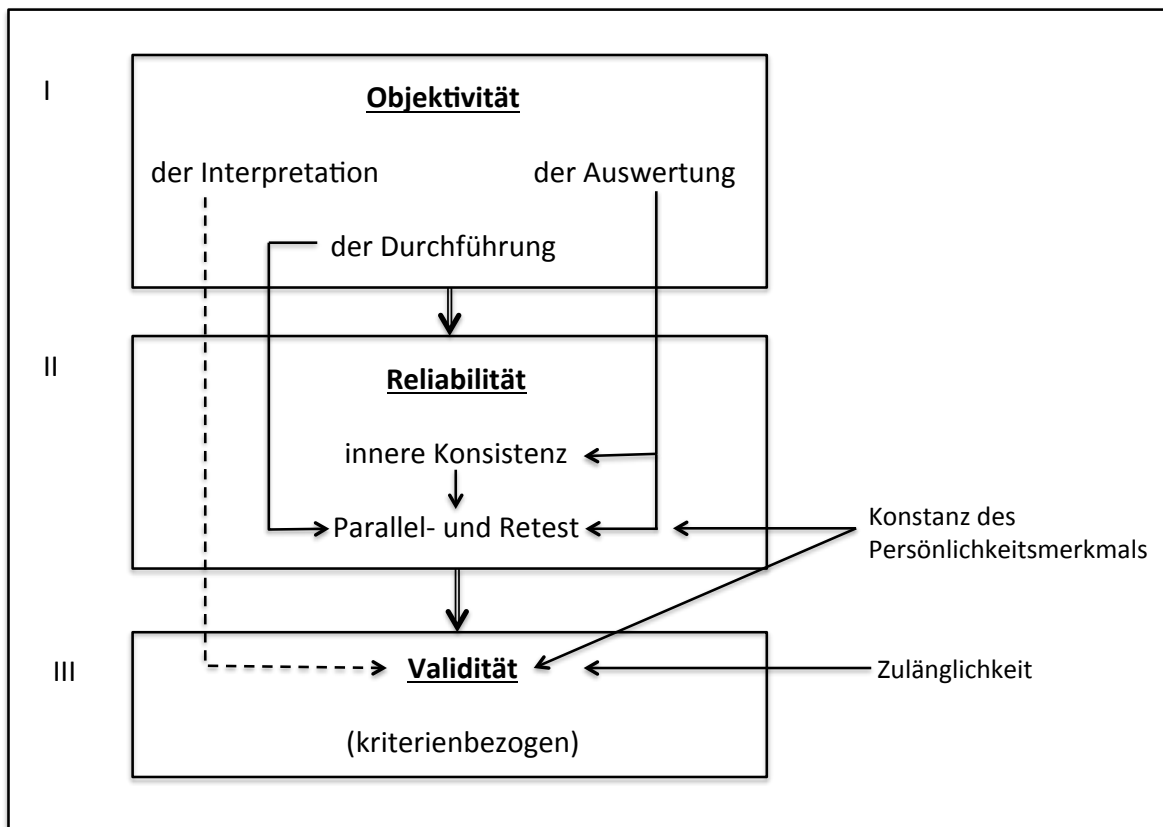


Abbildung 21. Beziehungen zwischen den Hauptgütekriterien (nach Lienert & Raatz, 1994, S. 13).

Die Beziehungen zwischen den Hauptgütekriterien lassen in diesem Gefüge eine gewisse Reihenfolge bzw. Wertigkeit erkennen. Danach kann ein Test nicht valider sein als reliabel und die Reliabilität kann nicht höher sein als die Objektivität. Demnach können bei einem Test mit hoher kriterienbezogener Validität die übrigen Gütekriterien gewissermaßen als erfüllt angesehen werden. Eine hohe kriterienbezogene Validität erlaubt ein Auslassen (vorsichtiger bzw. defensiver formuliert: ein Vernachlässigen) der Überprüfung der restlichen Gütekriterien (Lienert & Raatz, 1994, S. 13).

5.1.2.1 Untersuchungsdesign und -methodik

Die Untersuchung umfasste für die Mehrheit der Versuchsteilnehmer vier leistungsdiagnostische Ausdauer tests, die jeweils zu zwei Zeitpunkten (mindestens eine Woche Zwischenzeit), einmal auf dem Fahrradergometer und einmal auf dem Laufband stattfanden. Mit einer Spiroergometrie wurden die Atemgase gemessen und erfasst. Außerdem wurden die Pbn durch die Messung der Blutlaktatwerte, die Erfassung des Respiratorischen Quotienten und der maxi-

malen Herzfrequenzwerte, sowie durch die Abfrage des subjektiven Belastungsempfindens mit der Borg-Skala hinsichtlich des Ausbelastungszustands untersucht.

Bei dem ersten Termin wurde nach der Feststellung des Gewichts und der Körpergröße der 3000-m-FET durchgeführt. Hierbei hatte die Versuchsperson in einer Warm-up-Phase etwas Zeit, sich an das Rad zu gewöhnen und sich mit der Gangschaltung vertraut zu machen. Der Gang konnte während des Tests von dem Pbn frei gewählt und verändert werden. Nach einer angemessenen Pause von mehreren Minuten, bis der Pbn bzw. die Probandin ein Signal zum weitermachen gab, absolvierten die Teilnehmer im Anschluss nochmal einen Rampentest mit dem Ziel der subjektiven Ausbelastung. Dies hatte den Zweck einer zeiteffizienten Ermittlung der maximalen Leistungsfähigkeit bzw. des aktuellen Leistungsniveaus, sowie der Laktattoleranz des Pbn. Die Anfangsleistung betrug 130 Watt und wurde automatisch jede Minute um 30 Watt erhöht. In Abbildung 22 ist ein Versuchsteilnehmer während der Testphase abgebildet.



Abbildung 22. Laboruntersuchung mit Spiroergometrie zur Validierung des 3000-m-FET.

Über eine Motorbremse der Tacx-Trainer-Apparatur wurde der Widerstand während des Rampentests reguliert. Die Studienteilnehmer hatten die Möglichkeit, auf dem Bildschirm die eigene Wattleistung zu sehen und konnten so die Trittfrequenz anpassen, sobald die Wattleistung über oder unter der zu tretenden Leistung lag. Die Tacx Software gab für jede minutig vorgegebene Wattleistung einen Toleranzbereich an, der sich immer im Bereich plus/minus zehn Watt bewegte (grüner Bereich). Sobald die Pbn mehr als zehn Watt über oder unter der zu tretenden Wattleistung waren, wechselte die Wattzahl auf dem Bildschirm die Farbe von grün nach rot, somit wurde dem Pb signalisiert, dass die Trittfrequenz angepasst werden musste. Kam es vor, dass ein Pb zu häufig unter der zu bewältigenden Wattleistung bzw. unterhalb des Toleranzbereichs lag, wurde der Rampentest von der supervisierenden Person beendet, da somit offensichtlich die Wattleistung von dem Pbn nicht mehr bewältigt werden konnte. Der Pb sah, wie in Abbildung 23 zu erkennen ist, neben der Wattleistung auch noch andere Parameter, wie beispielsweise die zurückgelegte Strecke, die Zeitangabe, die Trittfrequenz etc. und hatte so die Möglichkeit, die gemessene Leistung während der Testphase zu beobachten.



Abbildung 23. Darstellung des Bildschirms während der Testphase eines Rampentests auf dem Fahrradergometer.

Beim zweiten Termin wurden die gleichen Parameter mit einer Testung auf dem Laufband (Modell: h/p/cosmos, quasar) erfasst. Zunächst führte die Versuchsperson einen 1000-m-Lauf durch. Der Test startete bei einer Geschwindigkeit von 9,3 km/h und konnte während der Testung von dem Pb frei gewählt werden. Die Anfangsgeschwindigkeit von 9,3 km/h beruht auf der Überlegung, dass mit dieser Geschwindigkeit die 1000 Meter genau in der maximalen Zeitvorgabe von 6:30 Min. bewältigt werden konnten. Nach einer Regenerationszeit von mehreren Minuten, die von dem Pbn individuell bestimmt werden konnte, folgte der Rampentest auf dem Laufband. Hierbei startete der Pb mit einer Anfangsleistung von acht km/h mit einem Anstieg der Geschwindigkeit um ein km/h pro Minute.

Ein Teil der Pbn ($n = 29$) stellte sich für die Reliabilitätsstudie zur Verfügung und wurde nach dem gleichen Schema an zwei weiteren Terminen getestet. Jeder Test umfasste die Aufzeichnung der spiroergometrischen Daten durch eine Spiroergometrie-Software der Firma Fusion Care, die Erfassung der Laktatwerte, die Dokumentation der Pulsfrequenzen nach den Belastungsphasen, die Abfrage der Anstrengungseinschätzung mit der Borg-Skala direkt nach Beendigung der Tests und bei den Fahrradergometertests zusätzlich die Berechnung der maximalen und durchschnittlichen Trittfrequenz, der maximalen und durchschnittlichen Geschwindigkeit, der maximalen und durchschnittlichen Wattzahl und der relativen Leistung (Watt/kg). Die Laktatabnahme erfolgte ebenfalls jeweils direkt nach der Belastung. Durch die Spiroergometrie konnten die Atemgase (VO_{2max} , VCO_{2max}) gemessen werden und der Respiratorische Quotient (Verhältnis von eingeatmetem Sauerstoff zu ausgeatmetem Kohlendioxid) ermittelt werden.

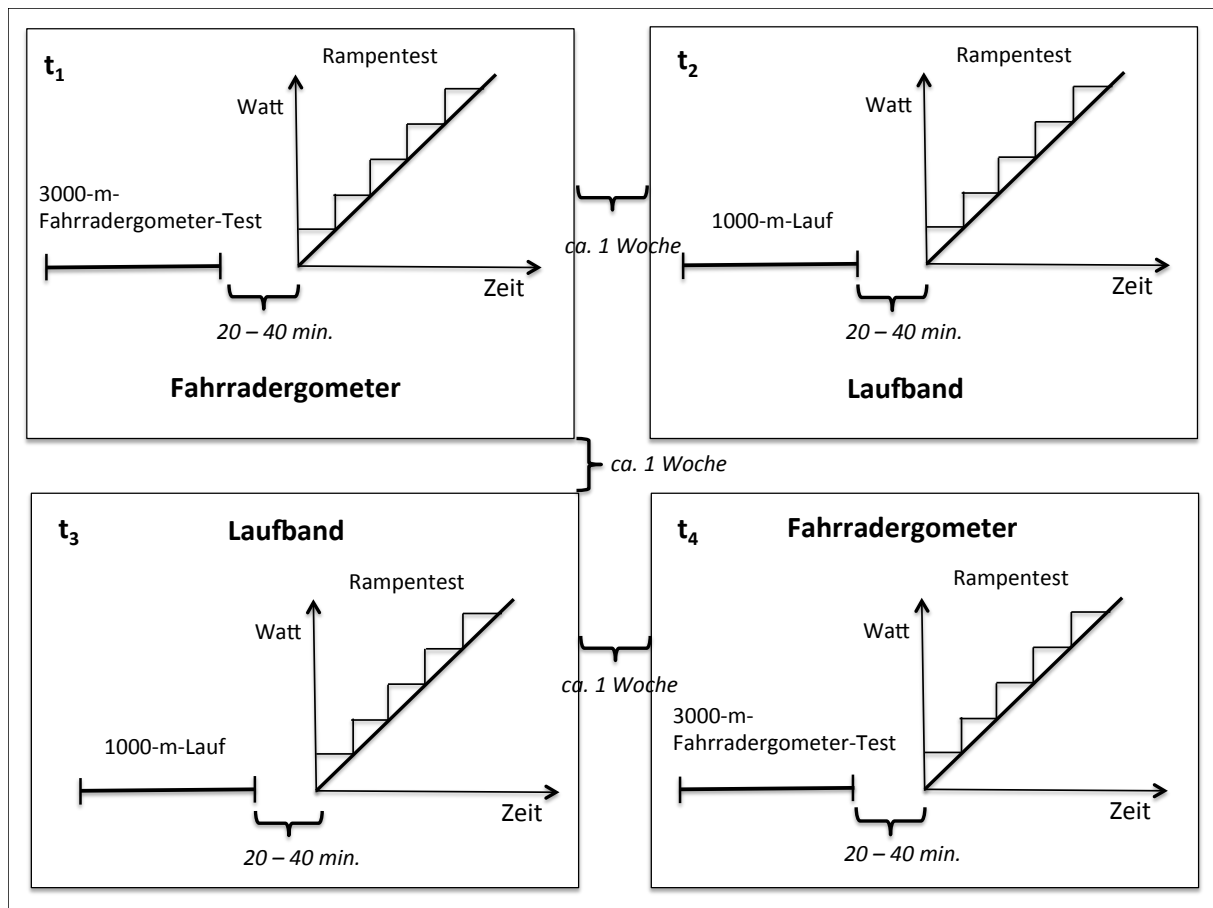


Abbildung 24. Untersuchungsdesign zur Prüfung der Gütekriterien des 3000-m-FET.

In der Abbildung 24 ist das Untersuchungsdesign dargestellt. Jeder der Versuchsteilnehmer ($N = 145$) führte vier Tests durch, aufgeteilt auf zwei Testtage (t₁ und t₂). Bei dem ersten Termin absolvierten die Versuchspersonen die Tests auf dem Fahrradergometer und bei dem zweiten Termin auf dem Laufband. Für die Testung einer Person an einem Termin wurde in etwa eine Stunde eingeplant. Insgesamt konnten von den 145 Studienteilnehmern 29 für die Reliabilitätsstudie (Retest) gewonnen werden, die noch einmal zu zwei zusätzlichen Testtagen erschienen, um die Tests zu wiederholen. Der zeitliche Abstand zwischen den Tests belief sich auf eine Woche. Bei der Reliabilitätsüberprüfung (t₃ und t₄) wurde am dritten Termin die Laufbandtests und bei der vierten und letzten Testung die Fahrradergometertests durchgeführt. Diese Umkehrung der Reihenfolge ist vorgenommen worden, um dem Aspekt eventueller Leistungseinbußen bspw. aufgrund von Erschöpfung vorzubeugen und somit eventuelle Reihenfolgeeffekten auszubalancieren.

Im Hinblick auf die Validitätsprüfung durch diese Laborstudie konnte anhand der erfassten Parameter die Kriteriumsvariable Ausdauer in der dazu in der Sport- und Trainingswissenschaft üblichen Art und Weise herangezogen werden. Definiert man die allgemeine aerobe

dynamische Ausdauer als Ermüdungswiderstandsfähigkeit kann bspw. die Durchschnittsgeschwindigkeit, die Zeit für die zu bewältigende Strecke, oder die relative Leistung (Watt/kg) als valides Außenkriterium angenommen werden (Heck & Schulz, 1998). Die Validitätsprüfung kann mittels Korrelation der objektiven Kriteriumsleistungen (durchschnittliche Geschwindigkeit, durchschnittliche Leistung in Watt) mit Testwerten, wie bspw. der maximalen Sauerstoffaufnahmekapazität (VO_{2max}), vorgenommen werden (siehe hierzu auch Heck & Schulz, 1998). Für den Vergleich des bisherigen BFT-Ausdauertests, dem 1000-m-Lauf, und dem neuen Ausdauer test im Rahmen des BFT, des 3000-m-FET, kam ein Laufband zum Einsatz, womit eine Durchführung der gesamten Studie mit der stationären Spiroergometrie ermöglicht werden konnte und so gegenüber der Felduntersuchung auf der Laufbahn im Freien Unabhängigkeit von z. B. Witterungsbedingungen erreicht wurden. Die Schwierigkeit, dass unter Normalbedingungen auf der Laufbahn der Pbn i. d. R. vier bis sechs Sekunden benötigt, um seine volle Laufgeschwindigkeit für die Bewältigung der 1000 Meter zu erreichen, konnte eliminiert werden, indem die Anfangsgeschwindigkeit des Laufbandes durch die dazugehörige Software voreingestellt werden konnte und der Pbn nach etwa zwei bis drei Sekunden bereits eine Laufgeschwindigkeit von 9,3 km/h erreicht hat. Anschließend konnte der Pbn die Geschwindigkeit frei über das Display nach oben oder unten regulieren und eine für das eigene Leistungsniveau passende Geschwindigkeit anwählen.

In den Labortests zur Beurteilung der allgemeinen aeroben dynamischen Ausdauer als Indikator der physischen Fitness der Versuchspersonen sind sowohl für den Vergleich der Ausdauer tests (dem bisherigen 1000-m-Lauf, den beiden Rampentests und dem neuen 3000-m-FET), als auch für die Untersuchung zu den Zusammenhängen im Hinblick auf die psychische Belastbarkeit (Persönlichkeitsfaktoren: Emotionale Stabilität und Belastbarkeit bzw. Neurotizismus) von besonderer Relevanz. Nach Lienert und Raatz (1994; zit. nach Heck & Schulz, 1998, S. 13) versteht man unter einem Test „ein wissenschaftlich begründetes Verfahren zur Untersuchung von Persönlichkeitsmerkmalen mit dem Ziel, eine möglichst quantitative Aussage über die Ausprägung des Merkmals zu erhalten“. Aufgrund der Voraussetzung einer maximalen Belastung bzw. der Ausbelastung der Versuchsteilnehmer und um die maximale Leistungsstufe bestimmen zu können und eine möglichst genaue Erfassung der Merkmalsausprägung zu erreichen, wie bspw. die maximale Sauerstoffaufnahme, sind unter anderem auch motivationale Aspekte sowie das Verhalten des Untersuchers zu beachten (Heck & Schulz, 1998). Die leistungsdiagnostischen Untersuchungen wurden von verschiedenen Studierenden der Sportwissenschaft durchgeführt. Die Pbn unterlagen somit evtl. unterschiedlich motivati-

onsbeeinflussenden Verhaltensweisen seitens der Untersucher, was als Beeinträchtigung hinsichtlich der Durchführungsobjektivität zu werten ist. Allerdings schätzten die Pbn direkt nach den Tests ihr Belastungsempfinden anhand der Borg-Skala ein. Anhand dieser Daten konnte ein weiterer Indikator herangezogen werden, der eine Aussage über das individuell eingeschätzte Belastungsempfinden nach den Ausdauertests darstellt. Somit konnten Rückschlüsse darüber getroffen werden, inwiefern sich die Testpersonen nach den Tests auch tatsächlich ausbelastet gefühlt haben.

Eine Teil Anzahl ($n = 29$) der Versuchsteilnehmer konnte für die Überprüfung der Zuverlässigkeit und Messgenauigkeit der Ausdauertests bzw. für die Reliabilitätsprüfung gewonnen werden. Dazu nahmen, wie bereits erwähnt, diese Versuchspersonen jeweils an beiden Tests zwei Mal teil, verteilt auf insgesamt vier Termine, sodass ein gewisser zeitlicher Abstand auch unter dem Aspekt der Regeneration gewährt war. Nach Heck und Schulz (1998) ist die Reliabilität eines Ausdauertests insbesondere von der messmethodischen, sowie der individuellen biologischen Variabilität abhängig. Bei Ausbelastungstests kommt die Störgröße der Leistungsbereitschaft bzw. der bereits erwähnte motivationale Aspekt hinzu. „Die meßmethodische Variabilität setzt sich zusammen aus der Variation der Leistung der Belastungsgeräte und der Variation bei der Meßwertgewinnung. Bei der Laktatbestimmung wird neben dem apparativen Analysefehler die gesamte Kette der Blutabnahmeprozedur die Höhe des Analysefehlers mitbestimmen.“ (Heck & Schulz, 1998, S.16) Hierzu kann bspw. zu langes Warten genannt werden oder Komplikationen bei der Blutabnahme über den Ohrlappen des Versuchsteilnehmers, was dann durchaus gravierende Auswirkungen auf die Messergebnisse haben kann und je nach Grad des Analysefehlers bringt dies eine Folgewirkung mit sich, die zur Feststellung einer unzulänglichen Reliabilität des Tests führt. Die Interpretationsobjektivität müsste nach Heck und Schulz (1998) hoch sein, da die physiologischen Zusammenhänge zwischen dem Konstrukt Ausdauerleistungsfähigkeit und den Testvariablen VO_{2max} bzw. VO_{2peak} und Laktatwerte bekannt sind.

Zur Überprüfung der Gütekriterien des 3000-m-FET, wie er für den Einsatz im Rahmen der Personalgewinnung der Bundeswehr zum Einsatz kommen soll, wird also bei der Laborstudie insgesamt auf vier verschiedene Ausdauertests zurückgegriffen, welche alle in Verbindung mit einer Spiroergometrie durchgeführt wurden. Die Ergometrie stellt eine Messmethode dar, die eine exakte Dosierbarkeit und Reproduzierbarkeit der Belastung bzw. physischen Leistung erlaubt (Hollmann & Strüder, 2009). Ergometrische Verfahren, wie sie in dieser Studie

zum Einsatz kommen, messen „in erster Linie die Leistungsfähigkeit des Herz-Kreislauf-Systems und haben ihren Schwerpunkt daher in der Abschätzung der Ausdauer“ (aus der Fünter, Faude, Skorski & Meyer, 2013). Neben denjenigen spiroergometrischen Daten, die ebenfalls mit Hilfe des Spirometers erfasst werden, wie das Atemminutenvolumen (VE), die Atemgase VO_2 und CO_2 , wurden auch die Laktatkonzentration im Kapillarblut gemessen und die Herzfrequenz-Werte aufgezeichnet und darüber hinaus der Respiratorischer Quotient (Quotient aus der Kohlendioxidabgabe und Sauerstoffaufnahme; $RQ=VCO_2/VO_2$) und die relative maximale Sauerstoffaufnahmekapazität berechnet. Durch die bei der Leistungsdiagnostik herangezogenen Parameter und gewonnenen Daten konnte durch die Auswertung dieser Daten die Überprüfung der Gütekriterien aus sportwissenschaftlich-trainingswissenschaftlicher Perspektive geleistet werden. Dazu wurden auch die aus den jeweiligen Tests generierten spiroergometrischen Größen, sowie die Herzfrequenz- und Laktatwerte auf Zusammenhänge hin untersucht. Dies diente dem Zweck festzustellen, inwiefern die einzelnen Parameter der Tests miteinander übereinstimmen und ob Zusammenhänge zwischen den Tests bestehen. Das Hauptaugenmerk lag dabei auf der Validitäts- und Reliabilitätsüberprüfung des 3000-m-FET und auf der Überprüfung der angenommenen Strukturähnlichkeit zum 1000-m-Lauf.

Des Weiteren sollte evaluiert werden, ob der in dieser Arbeit entwickelte 3000-m-FET tatsächlich zur Feststellung der allgemeinen aeroben dynamischen Ausdauerleistungsfähigkeit geeignet ist und zumindest ähnliche Messergebnisse liefert wie andere (gängige) Ausdauer-tests in der sportwissenschaftlich-trainingswissenschaftlichen Leistungsdiagnostik. Aufgrund dieser Überlegungen kamen im Rahmen dieser Studie zwei weitere Rampentests, sowohl auf dem Fahrradergometer als auch auf dem Laufband, zum Einsatz. Dabei handelt es sich um ein Belastungsprotokoll, wobei die Pbn innerhalb einer Zeit von sieben bis 12 Minuten auf eine relativ hohe Leistung gebracht werden, wie es bei einem geeigneten Test zur Bestimmung der Leistung bei maximaler Sauerstoffaufnahme ($VO_{2max/peak}$) der Fall ist. Der VO_{2max} -Wert gilt in der einschlägigen Literatur als Bruttokriterium der allgemeinen Ausdauerleistungsfähigkeit. I. d. R. wird die VO_{2max} als ein Indikator für die Ausdauerleistungsfähigkeit und zur Beurteilung des entsprechenden physischen Zustands herangezogen. Die maximale Sauerstoffaufnahme ist demnach auch in dieser Arbeit neben der Überprüfung der Testgütekriterien und insbesondere im Kontext der psycho-physischen-Belastbarkeit eine wesentliche Kenngröße, aus der sich die körperliche Belastbarkeit bzw. der allgemeine körperliche Fitnesszustand ableiten lässt. Bei zunehmender Leistung steigt die Sauerstoffaufnahme bzw. die VO_{2max} , die

auch als maximale aerobe Kapazität bezeichnet wird „linear an und erreicht bei der symptomlimitierten Ausbelastung den maximalen Wert (. . .). Die VO_{2max} repräsentiert am zuverlässigsten die individuelle maximale Leistungsfähigkeit der Systemkette Atmung-Kreislauf-Muskelstoffwechsel“ (Haber, 2007, S. 142). Die maximale Sauerstoffaufnahme ist nicht als die maximale Aufnahme von Sauerstoff durch die Atmung zu verstehen, sondern die Verwertung des Sauerstoffs durch das Herz-Kreislauf-System, der durch die Atmung eingeatmet wird.

Die Rampentests wurden auch aus zeitökonomischen Gesichtspunkten gewählt, da sie nicht nur zu einer relativ schnellen subjektiven Ausbelastung der Versuchsperson führen, sondern auch aufgrund der kürzeren Zeitintervalle und des größeren Anstiegs der Wattleistung u. a. eine zeiteffiziente Feststellung des VO_{2max} -Wertes zulassen. Für die Ermittlung der VO_{2max} wird, wie es der Literatur zu entnehmen ist, nach wie vor als weitverbreitet eine Belastungsdauer von mindestens 7/8 bis 12 Minuten zur Messung der VO_{2max} empfohlen (siehe Buchfuhrer, Hansen, Robinson, Sue, Wassermann & Whipp, 1983). Hingegen merken Midgley, Bentley, Luttikholt, McNaughton und Millet (2008) in Bezug auf die Ausführungen und Studie von Buchfuhrer et al. (1983) an, dass aufgrund der homogenen getesteten Stichprobe eine solche Empfehlung der Belastungsdauer als durchaus kritisch zu sehen ist. Die Autoren fassen verschiedene Untersuchungen zusammen, bei denen sich andere Zeitspannen bzw. Belastungsprotokolle zur Feststellung der VO_{2max} als valide erwiesen und in Abhängigkeit von dem Gesundheits- und Fitnesszustand, sowie des Geschlechts und dem verwendeten Messinstrument (Fahrradergometer oder Laufband) zu sehen sei. Demnach können auch wesentlich kürzere oder längere Belastungszeiten sowohl für Fahrradergometer- als auch Laufbandtests für die Ermittlung der VO_{2max} eingesetzt werden.

Die körperliche Leistungsfähigkeit, speziell die allgemeine aerobe Ausdauerleistungsfähigkeit, ist an die Menge des zu Verfügung stehenden Sauerstoffs gekoppelt. Die physische Leistung ist abhängig von Stoffwechselprozessen, von den O_2 -Transportsystemen, dem Atmungs- und Herz-Kreislauf-System. Dabei wird der Sauerstoffaufnahme eine besondere Bedeutung beigemessen, d. h. der Fähigkeit, pro Zeiteinheit eine möglichst große Menge an Sauerstoff aufnehmen zu können. Somit lässt sich die Ausdauerleistungsfähigkeit wie bei einer Kette über das schwächste Element hinsichtlich der Gesamtbelastbarkeit limitieren (Heck, 1990). Bei der Feststellung und Messung der Leistungsfähigkeit von Teilsystemen lässt sich nach Heck (1990) dann auf die Gesamtleistung schließen.

„Es gilt somit, die limitierenden Indikatoren zu erfassen, die eine hohe Korrelation zur maximalen Energierate aufweisen und daher die Ausdauerleistungsfähigkeit im Leistungstest messen. Für leistungsdiagnostische Routineuntersuchungen sind jedoch nur solche Parameter geeignet, die ohne größere Belästigung des zu Untersuchenden einfach zu registrieren sind.“ (Heck, 1990, S. 74)

Dem entsprechend kommt dem Sauerstofftransport bzw. der Sauerstoffaufnahme unter Belastung als leistungsdiagnostischem Parameter des Energiestoffwechsels eine besondere Bedeutung zu. Die maximale Sauerstoffaufnahme ist ein objektives Kriterium zur Beurteilung der kardiopulmonalen (Herz-Lungen-Funktion) Leistungsfähigkeit, die in Milliliter (relative VO_{2max}) oder Liter pro Minute (absolute VO_{2max}) angegeben wird. Die maximale Sauerstoffaufnahme umfasst alle physiologischen Mechanismen, die an der Leistungserbringung beteiligt sind, und ist bei maximaler Ausbelastung bzw. Ausschöpfung aerober und anaerober Energiereserven ein guter Parameter der Ausdauerleistungsfähigkeit (vgl. Dieter, 2004). Zu den Mechanismen zählen die äußere Atmung, der Gasaustausch in der Lunge, das Herzzeitvolumen, der Sauerstofftransport im Blut und die O_2 -Aufnahme in die arbeitende Muskulatur (vgl. Meyer & Kindermann, 1999; Heck, 1990). Zu den limitierenden Faktoren für die maximale Sauerstoffaufnahme gehören nach Hollmann und Strüder (2009, S. 327):

„Interne Faktoren:

- Motivation
- Ventilation
- Distribution und
- Diffusion in der Lunge
- Herzzeitvolumen
- Blutverteilung
- Blutvolumen
- Total-Hämoglobingehalt
- Ernährungszustand

Externe Faktoren:

- Belastungsmodus
- Größe und Art der eingesetzten Muskulatur
- Körperposition
- Sauerstoffpartialdruck in der Einatemungsluft
- Klima“

Die maximale Leistung wird durch konstant ansteigende Belastungen erreicht. Nach Hollmann (1950) haben gesunde untrainierte Männer des dritten Lebensjahrzehnts eine maximale O₂-Aufnahme von 3300 plus/minus 200 ml/Min., Frauen liegen hingegen im Durchschnittbereich von 2000 plus/minus 200 ml/Min. (Hollmann & Strüder, 2009). Die relative maximale Sauerstoffaufnahme errechnet sich durch die Division der maximalen O₂-Aufnahme durch das Körpergewicht (ml/min/kg) und ist somit ein geeignetes Maß, um die individuelle Leistung zu bestimmen und mit anderen Personen vergleichen zu können. Hollmann und Strüder (2009, S. 321) nennen Vergleichswerte bei Männern von 40 bis 55 ml/kg*min⁻¹ und bei Frauen von 32 bis 38 ml/kg*min⁻¹.

$$\text{Rel. VO}_{2\text{max}} = \frac{\text{ml}}{\text{kg}} \cdot \text{min}^{-1}$$

Sobald die Sauerstoffaufnahme einer Person trotz weiterer ansteigender Belastung nicht weiter zunimmt und es zu einem Plateau hinsichtlich der Sauerstoffaufnahme kommt, stellt sich das sog. Leveling-off-Phänomen ein. Hierbei kann der Körper keinen weiteren Sauerstoff mehr verwerten, d. h. obwohl die Belastung zunimmt, bleibt die Sauerstoffaufnahme konstant.

Neben den erwähnten Parametern kam bei allen Ausdauer-tests nach sofortiger Beendigung der Belastung die Borg-Skala zum Einsatz, um so das subjektive belastungsempfinden nach den einzelnen Tests herauszufinden. Für diese Untersuchung wurde die Borg-Skala in nachstehende Variante verwendet (Tabelle 5):

Tabelle 5
Borg-Skala.

Borg-Skala	
6	
7	sehr, sehr leicht
8	
9	sehr leicht
10	
11	leicht
12	
13	etwas anstrengend
14	
15	anstrengend
16	
17	sehr schwer
18	
19	sehr, sehr schwer
20	

Die Testperson kann mittels der Borg-Skala (siehe Borg, 1970, 1978) von 7 (= sehr, sehr leicht) bis 19 (= sehr, sehr schwer) sein individuelles Belastungsempfinden angeben (Received Perception of Exertion, RPE). Der Hintergrund dieser Einteilung ist die Herzfrequenzzahl dividiert durch zehn. Ein Beanspruchungsempfinden von elf („leicht“) entspricht also etwa einer Herzfrequenz von 110 Schlägen pro Min., eine Angabe von 13 („etwas anstrengend“) würde einer Herzfrequenz von 130 Schlägen/Min. entsprechen. Borg (1961) entwickelte diese 15-stufige Skala so, dass sie der Herzfrequenzrate eines normal gesunden Mannes (60 bis 200 Schläge/Min.) gleicht. Die Vorhersagegenauigkeit der Linearität der Skala mit der Herzfrequenz ist jedoch nicht präzise genug (siehe Noble & Robertson, 1996) und sollte von daher lediglich als ein Anhaltspunkt bzw. als eine Orientierung zur Belastungseinschätzung gesehen werden. Die mangelnde Linearität wurde später von Borg (1978) verbessert, indem er zu den einzelnen Stufen der intervallskalierten Skala Adjektive hinzufügte, um so eine validere Aussagekraft über die angegebene Belastungsintensität zu erhalten. Der Ergänzung der Skala mit Adjektiven/Adverbien liegt das Anliegen der angemessenen quantitativen Semantik zugrunde. Die zugeordneten Begriffe weisen nach einer vorangegangenen Studie von Borg und Dahl-

strom (1962) eine gute intersubjektive Übereinstimmung und eine konstante semantisch widergegebene Intensität auf, d. h. die Abstufung zwischen den Begriffen „sehr leicht“ und „sehr sehr leicht“ ist gleich dem Unterschied zwischen „sehr schwer“ und „sehr sehr schwer“ (Borg & Lindblad, 1976) zu sehen.

Zusätzlich zur spiroergometrischen Messung wurden die Laktatkonzentrationen im arteriellen Blut über den Ohrlappen erfasst. Den Studienteilnehmern wurde vor und nach jeder Belastungsphase Blut aus dem Ohrlappen entnommen, und anschließend die Laktatwerte überprüft. Über die Milchsäurekonzentration im arteriellen Blut sollte die objektive Ausbelastung bzw. das Belastungsniveau der Pbn festgestellt werden. Milchsäure bzw. Laktat fällt bei dem anaerob-laktaziden Stoffwechselprozess der Glykolyse als Endprodukt an. Bei diesem Energiestoffwechselprozess wird durch Glukose oder Glykogen über chemische Reaktionen Laktat und Adenosintriphosphat (ATP) hergestellt. Das ATP ist eine chemische Verbindung, die zur Energiebereitstellung für die Muskelarbeit und andere energieerfordernde Prozesse dient, indem es zu Adenosindiphosphat (ADP) und Phosphat (P) abgebaut bzw. aufgespalten wird.

Das Laktat führt im Muskel zu einer Verschiebung des pH-Wertes hin zu einem sauren Milieu, so dass die Leistungsfähigkeit des Muskels herabgesetzt wird (Hollmann & Strüder, 2009). Laktat wird ständig und nicht nur während starker körperlicher Belastung gebildet und eliminiert. Abgebaut wird es vorwiegend in der arbeitenden Muskulatur, wie dem Herzmuskel, sowie der Leber und der Niere. Die Ruhewerte sind von der Ernährung und dem Trainingszustand abhängig und variieren zwischen 0,4 und 1,5 mmol pro Liter Blut (Pokan, Hofmann, Hörtnagel, Ledl-Kurkowski & Wonisch, 2004). Die Halbwertszeit, d. h. die Zeit in der sich die Laktatkonzentration im Blut halbiert beträgt etwa 20 Min. (Grosser, Starischka & Zimmermann, 2008, S. 122). Durch geringe körperliche Aktivität kann die Milchsäure schneller abgebaut werden. Mit zunehmender Belastungsintensität steigt die Laktatproduktion in Folge der vermehrten anaerob-laktaziden Energiebereitstellung an. Die Laktatkonzentration erreicht bei maximalen Belastungen einige Min. nach Belastungsende seine Höchstwerte (Hollmann & Strüder, 2009, S. 75). Diese können bei 20 bis 25 mmol/l liegen (Grosser et al., 2008, S. 123). Zur Messung des Laktats wird i. d. R. Blut am Ohrläppchen entnommen. Dabei werden 20 µl (Mikroliter) Blut in einer Glaskapillare aufgefangen und anschließend enzymatisch, elektrochemisch, reflexionsphotometrisch oder trockenchemisch analysiert. Für diese Studie werden die Proben enzymatisch ausgewertet. Allerdings sind bei der Abnahme am Ohrläppchen auch Messungenauigkeiten zu erwarten. Durch Quetschung des Ohres, unge-

naue Pipettierung, Vermischung mit Schweiß oder unterschiedlich lange Lagerung der Proben können Messfehler auftreten (Pokan et al., 2004, S. 104). Die Laktatdiagnostik ermöglicht, ebenso wie die Spiroergometrie, eine Aussage über die aerobe und anaerobe Leistungsfähigkeit des Untersuchten. Allerdings müssen dafür bestimmte Belastungsprotokolle mit mindestens 12 bis 16 Stufen konstant ansteigender Intensität durchlaufen werden, um eine differenzierte Aussage über die Leistungsfähigkeit zu erhalten (Pokan et al., 2004, S. 107). Obwohl, wie bereits erwähnt, das Laktat ein geeigneter Indikator der anaeroben Kapazität und indirekt auch ein Maß der Ausdauerleistungsfähigkeit darstellt (Heck, 1990), ist durch die Blutabnahme über den Ohrlappen nur eine recht ungenaue Feststellung der Laktatkonzentration möglich. Für die vorliegende Arbeit wird die Laktatkonzentration lediglich erhoben, um festzustellen, ob das von Hollmann und Strüder (2009, S. 348) beschriebene Ausbelastungskriterium (Konzentration von 8 bis 10 mmol/l) erreicht wurde.

5.1.2.2 Beschreibung der Stichprobe

Für die Studie wurden Mitarbeiter und Studierende der Universität der Bundeswehr München via Email angeschrieben. Es wurden zunächst alle Interessenten unabhängig vom Alter, Geschlecht und Fitnesslevel bzw. Leistungsniveau berücksichtigt, mit der Ausnahme von Personen, die gesundheitliche Beschwerden, wie bspw. Bluthochdruck oder die Einnahme von Beta-Blockern (Arzneimittel zur Senkung der Ruheherzfrequenz und des Blutdrucks) angaben.

Die Stichprobe zur Evaluierung des 3000-m-FET setzt sich zum größten Teil aus Soldaten und Studierenden der Bundeswehr Universität München zusammen ($\approx 70\%$). Von den insgesamt 145 getesteten Personen, die sich für die Studie zur Überprüfung der Gütekriterien nach einer Email an alle Angehörigen der Bundeswehr Universität freiwillig meldeten, konnten, wie bereits erwähnt, 29 Personen für die Reliabilitätsstudie gewonnen werden. Die weiteren Personen waren entweder Mitarbeiter der Universität oder einige wenige Personen darüber hinaus, die durch Mund zu Mund Propaganda von der Studie hörten und Interesse hatten, an einer Leistungsdiagnostik teilzunehmen. Die relativ homogene Gruppe von Studenten und Soldaten der Bundeswehr führte zu einem Altersdurchschnitt von 25,6 Jahren ($\pm 6,2$ Jahre). Das durchschnittliche Alter lag bei den Frauen ($n = 26$) mit 25,8 Jahren ($\pm 9,2$ Jahre) leicht über dem der Männer ($n = 119$) mit 25,6 Jahren ($\pm 5,5$ Jahre). Der jüngste Teilnehmer der Studie war zum Untersuchungszeitpunkt 17 Jahre alt, der älteste 60 Jahre. Die Männer

hatten im Mittel eine Körpergröße von 182 cm und ein Gewicht von 81,4 kg. Die Frauen waren im durchschnitt 170 cm groß und hatten ein Körpergewicht von 62,9 kg. Die männlichen Teilnehmer der Studie hatten einen durchschnittlichen Body-Mass-Index (BMI) von 24,6. Dies entspricht nach der Normwerttabelle der WHO (2013) dem Normalgewicht. Die weiblichen Pbn hatten im Durchschnitt einen BMI-Wert von 21,3, welcher sich ebenfalls im Intervall des Normalgewichts befindet. Alle Pbn erschienen mindestens zu zwei Testterminen und absolvierten somit einen kompletten Testdurchlauf. Keiner der Pbn brach den Test ab, oder sagte den zweiten Termin für die Testung auf dem Laufband ab. Somit konnten alle Pbn mit in die Datenanalyse miteinfließen. Die Untersuchungen der Teilstudie zur Überprüfung des 3000-m-FET, der im Rahmen von darauffolgenden Untersuchungen zur psycho-physischen Belastbarkeit die physische Komponente erfassen soll, wurden von Mai bis Ende Dezember 2013 im sportwissenschaftlichen Labor der Universität der Bundeswehr München durchgeführt.

5.1.2.3 Untersuchungsdurchführung

Im folgenden Abschnitt wird über das Vorgehen bei der Untersuchungsdurchführung berichtet. Es geht auch um die Standardisierung des 3000-m-FET. Zunächst werden die vorbereiteten Maßnahmen und anschließend der Testablauf erläutert.

Vorbereitung

Vor Beginn der eigentlichen Untersuchung wurde eine E-Mail verfasst (siehe Anhang D1) und an alle Studierenden und Mitarbeiter der Universität der Bundeswehr München geschickt. Darin wurde allen Adressaten eine kostenfreie Leistungsdiagnostik mit anschließender Beratung zur allgemeinen Fitness angeboten. Die Personen, die sich daraufhin freiwillig meldeten, wurden mit einer zweiten E-Mail über den Ablauf der Untersuchung informiert. Anschließend erfolgte dann die Terminabsprache ebenfalls per E-Mail mit den Interessenten.

Die Vorbereitung der Messungen an einem Testtag erforderte als erstes, dass die einzusetzenden Messsysteme gestartet und kalibriert wurden. Dazu gehören der spiroergometrische Messplatz und das Laktat-Analysegerät. In Abhängigkeit von Testtag und Messung war zu-

sätzlich entweder die zugehörige Software des Fahrradergometers der Firma Tacx an einem separatem Laptop, oder die Laufband-Software zu starten. Im Anschluss an die etwa 15 bis 20-minütige Warmlaufphase der Spiroergometrie erfolgte eine Kalibrierung der Messsensoren durch den Untersucher. Am ersten Testtag wurde nach einer Einweisung des Pbn über den Testhergang das Körpergewicht und die Körpergröße des Teilnehmers ermittelt. Vor dem 3000-m-FET füllten die Pbn zunächst die beiden Skalen des BIP „Emotionale Stabilität“ und „Belastbarkeit“, sowie die State- und Trait-Kurzformen dieser beiden Skalen aus. Nach dem Ausfüllen des Fragebogens wurde der Versuchsperson ein Pulsgurt angelegt und eine passende Maske gewählt. Der korrekte Sitz der Maske wurde durch den Versuchsleiter sichergestellt. Die ermittelten anthropometrischen Daten, Geburtsdatum und eine eindeutige Kennnummer wurden in das Programm „LabManager“ und in das Programm Tacx Trainer Software (TTS) eingegeben. Jeder Pbn erhielt ein eigenes Benutzerprofil, damit die Daten später korrekt zugeordnet werden konnten. Alle Daten wurden darüber hinaus zur Absicherung gegenüber Datenverlust handschriftlich festgehalten. Bei den Fahrradergometer-Untersuchungen wurde zunächst die Sattelhöhe des Fahrrads (Cannondale; Modell: CAAD 10) je nach Beinlänge des Pbn justiert. Die Teilnehmer durften selbst entscheiden, wie hoch der Sattel eingestellt werden sollte. Der Versuchsleiter achtete jedoch darauf, dass im tiefsten Punkt der Kurbeldrehung eine Kniegelenksstreckung von mehr als 160° vermieden wurde, um eine möglichst optimale Sitzposition sicher zu stellen. Anschließend wurde dem Pbn eine passende Gesichtsmaske angelegt und er darüber an die Spiroergometrie angeschlossen. Die Referenzphase wurde gestartet und währenddessen die Ruhewerte der Versuchsperson ermittelt.

Testablauf

Nach der Begrüßung und Erläuterung des physischen Tests füllten die Versuchsteilnehmer die Skalen Emotionale Stabilität und Belastbarkeit des BIP, sowie die konzipierten Kurzformen (State und Trait) aus bevor die Testung der physischen Fitness begann. Der 3000-m-FET konnte durch die TTS programmiert und gespeichert werden, sodass der Test für jeden Pbn direkt ausgewählt werden konnte, ohne neue Einstellungen vornehmen zu müssen. Vor jeder Testung wurde sichergestellt, dass der Pbn eine gute Sicht auf das Display des Laptops mit dem Testlayout hatte (Abbildung 25). Somit konnten die Versuchsteilnehmer während der Testung ihre Leistung in Watt/kg, Trittfrequenz, Geschwindigkeit und die bewältigte Strecke, sowie die Zeit mitverfolgen.

Sobald die Gesichtsmaske für die Messung der Atemgase angelegt und der Pbn die Sitzposition eingenommen und die Pedalen parallel zum Boden ausgerichtet hatte, wurde über das Programm ein fünf-sekündiger akustischer Countdown-Timer gestartet, der nach Ablauf den Testbeginn signalisierte. Der Pb begann direkt nach Ablauf des Timers die Tretkurbel zu beschleunigen und die 3000-Meter-Strecke zu bewältigen. Der Testleiter wechselte zeitgleich mit dem akustischen Signal auf dem Computer die Spiroergometrie-Software von der Referenz in die Testphase, sodass die spiroergometrischen Daten erfasst werden konnten. Sobald die Testperson die 3000 Meter erreicht hatte, wurde der Test beendet, die benötigte Zeit und die Parameter, wie die durchschnittliche Trittfrequenz, durchschnittliche relative Leistung, Geschwindigkeit usw., automatisch durch die TTS gespeichert. Dem Untersuchungsteilnehmer wurde direkt nach dem Erreichen der 3000 Meter die Gesichtsmaske entfernt, mit einer Glaskapillare Blut aus dem Ohrappen zur Laktatmessung entnommen und das Belastungsempfinden mittels der Borg-Skala abgefragt. Die maximalen Herzfrequenzwerte wurden mit einer Pulsuhr der Firma Polar erfasst und ebenfalls direkt nach Beendigung des Tests abgelesen und wie alle anderen relevanten Werte separat in einer Excel-Tabelle eingetragen. Im Nachgang des Fahrradergometertests beantwortete der Pbn die letzten Fragen des Fragebogens, bevor die spiroergometrischen Daten mit Hilfe der Auswertung der Spiroergometrie-Software dem Pb erläutert wurden.

Nach einer Pause von etwa 20 bis 45 Minuten begann der Rampentest. Dieser konnte wie auch der 3000-m-FET programmiert und vor jeder Testung über die TTS aufgerufen werden. Der Ablauf der Untersuchung, wie bspw. die Erfassung der maximalen Herzfrequenz und die Laktatkonzentration im Blut, liefen identisch zum 3000-m-FET ab. Nach Abbruch bzw. Beendigung des Tests durch die Testperson bei einer bestimmten Wattstufe wurden wieder alle Werte der Testung sowohl über die TTS gespeichert als auch in einer Excel-Tabelle festgehalten. Zum Abschluss des ersten Termins füllten die Testpersonen die State-Version (momentanes Befinden) der Kurzsкала zur psychischen Belastbarkeit aus.

Spiroergometrie

Bei dem eingesetzten Spiroergometrie-System handelt es sich um das System „Master Screen CPX“ der Firma CareFusion. Es ermöglicht die Untersuchung kardiopulmonaler Funktionen während eines Belastungstests. Dazu zählen die Sauerstoffaufnahme, die Kohlenstoffdioxidabgabe, der Respiratorische Quotient (RQ), das Atemminutenvolumen (VE), die Atemfrequenz, sowie der Atemreserve. Weiterhin können, in Verbindung mit einem Pulsmessgerät, auch Herzfrequenz und Sauerstoffpuls ermittelt werden. Die Berechnung erfolgt dabei Atemzug-für-Atemzug (Breath-by-Breath). Die hierbei gemessenen Parameter sind VO_2 , VCO_2 , VE und RQ. Die Messgenauigkeit wird durch die Firma CareFusion mit den folgenden Werten angegeben (CareFusion, 2013):

Tabelle 6

Messgenauigkeiten der Spiroergometrie (nach CareFusion, 2013).

Messgröße	Bereich	Abweichung
VO_2	0 – 7 l/min	3 % oder 0,05 l/min
VCO_2	0 – 7 l/min	3 % oder 0,05 l/min
VE	0 – 300 l/min	2 % oder 0,05 l/min
RQ	0,6 – 2,0	4 %

Zur Anlage des „MasterScreen CPX“ gehört ein Basismodul welches eine vollautomatische Kalibration ermöglicht. Für die Informationen zu den Kalibrationsvorgängen sei auf die Firma CareFusion verwiesen. Weiterhin besteht das System aus dem „CPX-Modul“, einer Maske, dem Volumensensor und einem Probenschlauch. Dem Pbn wird eine passende Maske aufgesetzt. An diese kann das Mundstück mit dem integrierten Volumensensor angeschlossen werden. So werden alle Atemgase erfasst, ohne dass die Untersuchungsperson übermäßig in ihrer Bewegungsfreiheit gestört wird. Der Probenschlauch soll die Atemgase trocknen, bevor sie analysiert werden. „Die Spiroergometrie erlaubt eine differenzierte Aussage über die Ausdauerleistungsfähigkeit, den Trainingszustand sowie den Muskelstoffwechsel. Dabei wird die Sauerstoffaufnahme und insbesondere der respiratorische Quotient gemessen“ (Graf et al., 2001).

5.1.2.4 Ergebnisse der Laborstudie

Der erste Teil der Ergebnisdarstellung zeigt die Werte der 29 Pbn, die für die Reliabilitätsüberprüfung sowohl die beiden Fahrradergometertests als auch die beiden Laufbandtests ein zweites Mal durchgeführt hatten. Im Anschluss werden die Werte der 145 Untersuchungsteilnehmer dargestellt, die zu den ersten beiden Testterminen erschienen sind und alle vier Ausdauer tests für die Überprüfung der Validität jeweils einmal absolviert hatten. Die Ergebnisse sind im Wesentlichen anhand der Mittelwerte (M), Standardabweichungen (SD), sowie der Minimum- und Maximum-Werte aller vier Tests aufgeführt. Außerdem werden die Ergebnisse der Tests auf Zusammenhänge hin überprüft. Des Weiteren wurde bei der Reliabilitäts- und Validitätsüberprüfung ein t-Test für abhängige Stichproben durchgeführt, um den Unterschied der Mittelwerte zwischen den Tests zu beleuchten (Tabelle 19). Die Ergebnisse wurden mit Hilfe des Statistik-Programms IBM SPSS[®] Statistics (Version 21) ausgewertet. Die Reihenfolge der Ergebnisse der einzelnen Parameter ist hierarchisch organisiert und ergibt sich aus der Bedeutsamkeit der Werte für diese Arbeit.

5.1.2.4.1 Ergebnisdarstellung der Test-Retest-Reliabilität des 3000-m-FET

Neben dem 3000-m-FET wurden mit einer Gruppe von Personen ($n = 29$), die sich für eine zweite Testung der vier Tests bereit erklärt hatten, auch die weiteren eingesetzten Verfahren (1000-m-Lauf, Rampentests) auf Reliabilität hin untersucht. Obwohl der Fokus auf dem 3000-m-FET hinsichtlich der Überprüfung der Gütekriterien liegt, sollen die Werte des Vor- und Nachtests der drei weiteren Verfahren ebenfalls aufgezeigt werden.

Zeiten

Der für diese Arbeit bedeutendste Parameter für die Ausdauerleistungsfähigkeit ist die benötigte Zeit für den 3000-m-FET. Die Zeit für die vorgegebene Strecke von 3000 Metern stellt in den Untersuchungen und bei der Datengenerierung außerhalb des Labors das alleinige Kriterium zur Beurteilung der körperlichen Leistungsfähigkeit bzw. der physischen Belastbarkeit dar. Aufgrund dessen ist insbesondere der Zusammenhang der Testdauer bei einer Messwiederholung des 3000-m-FET für die Überprüfung der Reliabilität von speziellem Interesse.

Tabelle 7

Deskriptive Statistik der Testdauer im Vor- und Nachtest in Sekunden (n = 29).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	226	174	456	426
<i>Max.</i>	306	309	844	883
<i>M</i>	255	243	617	613
<i>SD</i>	15	31	95	102
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	225	175	448	447
<i>Max.</i>	297	308	840	844
<i>M</i>	254	240	627	599
<i>SD</i>	15	28	91	98

Bis auf den Rampentest auf dem Laufband erreichten die Pbn bei den Nachtests im Vergleich zum ersten Testdurchlauf etwas bessere Ergebnisse. Die Durchschnittszeit verbesserte sich bei dem 3000-m-FET von 4:15 Min. (t1; *SD* = 15 Sekunden) auf 4:14 Min. (t2; *SD* = 15 Sekunden) lediglich um eine Sekunde (Sek.). Eine deutlichere Leistungsverbesserung erzielten die Pbn mit durchschnittlich 10 Sek. längerer Fahrzeit bei dem Rampentest auf dem Fahrradergometer, was einem höherem Leistungs- bzw. Belastungsniveau entspricht (t1: *M* = 10:17 Min., *SD* = 95 Sek.; t2: *M* = 10:27 Min., *SD* = 91 Sek.). Eine etwas schnellere Zeit und somit eine Verbesserung wurde bei dem 1000-m-Lauf erreicht. Die Testpersonen erzielten hier einen drei Sek. besseren Schnitt von t1 (*M* = 4:03 Min., *SD* = 31 Sek.) auf t2 (*M* = 4:00 Min., *SD* = 28 Sek.). Bei dem Rampentest auf dem Laufband konnte eine durchschnittliche Leistungseinbuße von 14 Sek. mit einer Zeit von 6:13 Min. (*SD* = 102 Sek.) des Vortests auf 5:59 Min. (*SD* = 98 Sek.) bei der zweiten Testung ermittelt werden (siehe Tabelle 7). Die Analyse von Zusammenhängen ist in Tabelle 8 zusammengefasst⁴:

⁴ In den Tabellen sind die Minimum-Werte mit „Min.“ abgekürzt, wohingegen im Text die Abkürzung Min. für Minuten steht.

Tabelle 8

Korrelationen der Testdauer zwischen Vor- und Nachtest (n = 29).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.923	.000
1000-m-Lauf	1000-m-Lauf	.913	.000
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.961	.000
Rampentest Laufband	Rampentest Laufband	.956	.000

Bei Betrachtung der Korrelationen aller vier Tests zu den Zeitpunkten t1 und t2 zeigen sich hohe und signifikante Zusammenhänge von $r > .90$. Der Zusammenhang des Vor- und Nachtests des 3000-m-FET liegt bei $r > .92$ ($p = .000$). Besonders hohe und nahezu vollkommene Zusammenhänge zeigten sich zwischen den Vor- und Nachtests bei den Rampentests.

Maximale Sauerstoffaufnahme (VO_{2max})

Die maximale Sauerstoffaufnahme gilt als klassischer Indikator der körperlichen Leistungsfähigkeit und wird in der Literatur als weitverbreitetes „Bruttokriterium“ der allgemeinen Ausdauerleistungsfähigkeit (kardiopulmonal-metabolischen Kapazität) bezeichnet. Die VO_{2max} integriert alle zur Leistungserbringung beteiligten Mechanismen und wird absolut in Litern (l) pro Min. angegeben (aus der Fünten et al., 2013; Schnabel et al., 2008; Meyer & Kindermann, 1999; Heck, 1990). Diese körperlichen Mechanismen umfassen die äußere Atmung, den O_2 -Transport im Blut und die O_2 -Aufnahme in der Arbeitsmuskulatur, den Gasaustausch in der Lunge, sowie das Herzminutenvolumen (HMV) und die der Belastung entsprechend angemessene Verteilung. Die Größe der eingesetzten Muskelmasse bei einer Belastung ist ein limitierender Faktor der VO_{2max} . Demnach ist die Art der Belastung ein wichtiger Bezug zur Interpretation für die VO_{2max} und darf aufgrund der verschiedenen Belastungsformen in dieser Untersuchung (Fahrradergometer und Laufband) nicht unberücksichtigt bleiben. Die eingesetzte Muskelmasse und die Anzahl der beteiligten Mitochondrien sind bei den Fahrradergometertests geringer als bei Laufbelastungen, bei denen das Körpergewicht gegen die Schwerkraft bewegt wird (Meyer & Kindermann, 1999). Demzufolge ist die erreichbare VO_{2max} auf

dem Laufband in der Regel höher als auf dem Fahrradergometer. Die auf dem Fahrradergometer erfasste höchste Sauerstoffaufnahme wird daher auch als VO_{2peak} bezeichnet (aus der Fünten et al., 2013). Eine Ausnahme von dieser Regel stellen radtrainierte Personen dar, die aufgrund der radsportspezifisch trainierten Muskulatur auch auf dem Fahrrad ihre tatsächliche VO_{2max} mobilisieren können (aus der Fünten et al., 2013; Meyer & Kindermann, 1999). Für eine interindividuelle Vergleichbarkeit ist die relative maximale Sauerstoffaufnahme besser geeignet, da die eingesetzte Muskelmasse, die Herzgröße und das Blutvolumen sich in Abhängigkeit vom Körpergewicht unterscheiden, sodass schwerere Menschen i. d. R. bei gleichen Fitnesszustand eine höhere VO_{2max} aufweisen als leichtere Personen (Meyer & Kindermann, 1999). „Normwerte der VO_{2max} (. . .) liegen bei männlichen untrainierten Erwachsenen bis 30 Jahren bei etwa $40-45 \text{ ml/min} \cdot \text{kg}^{-1}$ (bei Frauen bei $35-40 \text{ ml/min} \cdot \text{kg}^{-1}$). Analog zur maximalen Leistung nimmt die VO_{2max} ab dem 30. Lebensjahr jährlich um etwa 1 Prozent ab. Herzpatienten liegen bei $20-25 \text{ ml/min} \cdot \text{kg}^{-1}$ “ (aus der Fünten et al., 2013, S.175).

Tabelle 9

Maximale Sauerstoffaufnahme-Werte (VO_{2max}) zwischen Vor- und Nachtest in l/Min. ($n = 29$).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	2406	2937	3019	3050
<i>Max.</i>	5591	5548	4821	5496
<i>M</i>	3818	3942	3910	4185
<i>SD</i>	564,2	514,8	450,9	496,1
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	2565	2891	2763	3039
<i>Max.</i>	5601	5358	4991	5321
<i>M</i>	3854	3931	3860	4076
<i>SD</i>	590,9	500,1	525,3	514,1

Im Vergleich des Vor- und Nachtests zeigen die VO_{2max} -Werte bei dem 3000-m-FET nur eine geringfügige durchschnittliche Verbesserung von der ersten Testung 3818 l/Min. ($SD = 564,2 \text{ l/Min.}$) auf 3854 l/Min. ($SD = 590,9 \text{ l/Min.}$) zur zweiten Testung. Bei den drei anderen Tests ist eine minimale Verringerung der VO_{2max} -Werte von t1 auf t2 festzustellen. Trotz der im Durchschnitt leicht verminderten VO_{2max} -Werte bei t2 kann festgehalten werden,

dass die Werte der maximalen Sauerstoffaufnahme bei allen vier Ausdauer Tests bei Betrachtung des Mittelwertes sehr eng zusammenliegen. Wie Eingangs erläutert bedingen sich die Unterschiede der VO_{2max} -Werte zwischen den Tests durch die unterschiedliche Belastung und die Größe der beanspruchten Muskulatur. Die durchschnittlichen VO_{2max} -Werte der einzelnen Tests können Tabelle 9 entnommen werden.

Tabelle 10

Korrelationen der maximalen Sauerstoffaufnahme-Werte (VO_{2max}) zwischen Vor- und Nachtest ($n = 29$).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.930	.000
1000-m-Lauf	1000-m-Lauf	.913	.000
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.931	.000
Rampentest Laufband	Rampentest Laufband	.938	.000

Im Vergleich aller vier Tests bei dem Vor- und Nachtest zeigen sich wie auch bei der Zeit Korrelationskoeffizienten von stets $r > .90$. Somit sind auch bei der Untersuchung der VO_{2max} in Bezug auf die Reliabilität höchst signifikante Zusammenhänge ($p = .000$) zu vermerken (siehe Tabelle 10). Demgemäß kann auch in Betracht des Indikators der physischen Leistungsfähigkeit eine hohe Genauigkeit bei Messwiederholung festgestellt werden.

Maximale Herzfrequenzen

Ein weiterer Parameter, der zur Überprüfung der Gütekriterien herangezogen wurde, ist die maximale Herzfrequenz (Hf_{max}). Erfasst wurde diese bei den Laufbandtests über einen Pulsgurt in Verbindung mit einer dem Laufband zugehörigen Software, die während der Testphase bis zur Beendigung des Tests die Herzfrequenz in Sekundenabständen mitplottet und speichert. Bei den Fahrradergometer Tests geschah die Erfassung der maximalen Pulsfrequenz direkt nach Abbruch der Rampe bzw. nach Beendigung der 3000 Meter mit einem Pulsgurt und einer Uhr der Firma Polar. Die Ergebnisse der Hf_{max} des Vor- und Nachtests sind in Tabelle 11 dargestellt.

Tabelle 11

Maximale Herzfrequenzen zwischen Vor- und Nachtest in Schlägen/Min. (n=29).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	144	164	165	173
<i>Max.</i>	193	188	198	205
<i>M</i>	175	178	182	190
<i>SD</i>	9,3	6,5	7,9	7,8
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	146	167	167	173
<i>Max.</i>	191	193	191	200
<i>M</i>	176	180	182	188
<i>SD</i>	10,5	7,1	7,7	6,4

Tabelle 11 zeigt die Hf_{\max} -Werte aller vier Ausdauer-tests im Vergleich zwischen dem Vortest (t1) und Nachtest (t2). Die Hf_{\max} -Werte bei dem 3000-m-FET unterschieden sich im Mittel lediglich um einen Schlag/Min. von 175 ($SD = 9,3$ Schlägen/Min.) von t1 auf 176 ($SD = 10,5$ Schlägen/Min.) zu t2. Die durchschnittlichen Hf_{\max} -Werte des Rampentests auf dem Fahrrad-ergometer sind bei dem Vor- und Nachtest mit 182 Schlägen/Min. (t1: $SD = 7,9$ Schlägen/Min.; t2: $SD = 7,7$ Schlägen/Min.) identisch geblieben. Bei dem 1000-m-Lauf erhöhten sich die Hf_{\max} im Mittel von t1 auf t2 um zwei Schlägen/Min. von 178 ($SD = 6,5$ Schlägen/Min.) auf 180 ($SD = 7,1$ Schlägen/Min.). Bei dem Rampentest auf dem Laufband nahmen die Hf_{\max} -Werte um zwei Schlägen/Min. von 190 ($SD = 7,8$ Schlägen/Min.) auf 188 ($SD = 6,4$ Schlägen/Min.) leicht ab.

Tabelle 12

Korrelationen der maximalen Herzfrequenzwerte (Hf_{max}) zwischen Vor- und Nachtest ($n = 29$).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.807	.000
1000-m-Lauf	1000-m-Lauf	.687	.000
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.783	.000
Rampentest Laufband	Rampentest Laufband	.809	.000

Die nahe beieinander liegenden, teilweise identischen Hf_{max} -Werte, führen zu mittleren bis meist hohen und signifikanten Korrelationen. Wie man Tabelle 12 entnehmen kann, ist der Zusammenhang zwischen den Testungen zu den Zeitpunkten t1 und t2 bei dem 3000-m-FET ($r = .81$; $p = .000$) und bei dem Rampentest auf dem Laufband ($r = .81$; $p = .000$) am höchsten. Eine Ausnahme in der Höhe des Korrelationskoeffizienten war lediglich bei dem 1000-m-Lauf festzustellen. Hier konnte ein mittlerer und höchst signifikanter Zusammenhang zwischen dem Vor- und Nachtest in Bezug auf die maximale Herzfrequenz erreicht werden ($r = .69$; $p = .000$).

Laktatwerte

Die ermittelten Laktatkonzentrationen im arteriellen Blut geben Hinweise über die Ausbelastung der Pbn. Bei drei der vier Nachtests fallen die Konzentrationen im Vergleich zum Vortest geringer aus. Mit Ausnahme des 1000-m-Laufs, bei dem ein minimaler Anstieg des Blutlaktatwertes (t1: 7,2 mmol/l, $SD = 2,5$; t2: 7,3 mmol/l, $SD = 2,2$) zu verzeichnen ist. Die höchsten Werte werden beim 3000-m-FET erreicht. Im Vortest sind es 9,9 mmol/l ($SD = 3,2$) und im Nachtest bildete sich Laktat in Höhe von 9,5 mmol/l ($SD = 2,8$; siehe Tabelle 13).

Tabelle 13

Laktatkonzentrationen im Blut nach Beendigung der Testphase zwischen Vor- und Nachtest in mmol/l (n = 29).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	4	3	5	5
<i>Max.</i>	17	15	15	12
<i>M</i>	9,9	7,2	9,7	8,3
<i>SD</i>	3,2	2,5	2,0	1,9
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	5	4	5	5
<i>Max.</i>	15	14	13	11
<i>M</i>	9,5	7,3	9,5	8,0
<i>SD</i>	2,8	2,2	1,9	2,0

Die untenstehende Tabelle 14 zeigt die Zusammenhänge der maximalen Laktatwerte zwischen dem Vor- und Nachtest der jeweiligen Ausdauer Tests. Bis auf den Rampentest, der auf dem Laufband durchgeführt wurde, zeigen sich bei den drei anderen Tests höchst signifikante Zusammenhänge im hohen Bereich ($r > .70$; $p = .000$). Der höchste Zusammenhang zeigte sich bei dem 3000-m-FET mit einem Korrelationskoeffizienten in Höhe von $r = .88$ ($p = .000$).

Tabelle 14

Korrelationen der maximalen Laktatwerte zwischen Vor- und Nachtest (n = 29).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.875	.000
1000-m-Lauf	1000-m-Lauf	.806	.000
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.750	.000
Rampentest Laufband	Rampentest Laufband	.694	.000

Werte der Borg-Skala

Durch die Borg-Skala kann, wie erläutert, das subjektive Belastungsempfinden bestimmt werden. Direkt nachdem die Tests von den Pbn absolviert wurden, gaben die Pbn eine Einschätzung ein, wie stark sie sich von dem jeweiligen Test beansprucht gefühlt haben. Diese Einschätzung des Belastungsempfindens kam sowohl bei dem Vor- (t1) als auch bei dem Nachtest (t2) zum Einsatz, um die Werte der Borg-Skala vergleichen zu können und so zu überprüfen, inwiefern über den Grad des subjektiven Belastungsempfindens sich Rückschlüsse lässt, ob und wie stark sich die Testteilnehmer bei beiden Testdurchläufen ähnlich ausbelastet fühlten.

Tabelle 15

Werte der Borg-Skala zwischen den Vor- und Nachtests (n = 29).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	13	14	15	15
<i>Max.</i>	19	20	20	20
<i>M</i>	17	16	18	18
<i>SD</i>	1,4	1,5	1,4	1,3
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	15	15	15	16
<i>Max.</i>	19	19	20	20
<i>M</i>	17	17	18	18
<i>SD</i>	2,8	1,2	1,3	1,2

In Tabelle 15 sind zunächst die Daten aus den Vor- und Nachtests zusammenfassend dargestellt. Bei den Mittelwerten der Borg-Skala zeigen sich, bis auf die minimale Unterscheidung bei dem 1000-m-Lauf, identische Werte. Das Belastungsempfinden wird bei dem 3000-m-FET sowohl bei dem Vor- als auch bei dem Nachtest mit einem Wert von 17 ($SD = 1,4$ bzw. 2,8) auf der Borg-Skala angegeben. Während der Mittelwert konstant ist, fällt die Streuung zur zweiten Testung des 3000-m-FET größer aus.

Tabelle 16

Korrelationen der Borg-Werte zwischen den Vor- und Nachtests (n = 29).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.553	.002
1000-m-Lauf	1000-m-Lauf	.593	.001
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.836	.000
Rampentest Laufband	Rampentest Laufband	.768	.000

Bei den Korrelationen der Borg-Werte zeigen sich zwischen Vor- und Nachtest bei dem 3000-m-FET ($r = .55$, $p = .002$) und bei dem 1000-m-Lauf ($r = .59$, $p = .001$) hoch bzw. höchst signifikante Zusammenhänge im mittleren Bereich. Höhere Zusammenhänge zwischen t1 und t2 findet man bei beiden Rampentests mit $r > .70$ ($p = .000$; siehe Tabelle 16).

Atemminutenvolumen

Als Atemminutenvolumen (VE) bezeichnet man das Volumen an Atemluft, welches in einer Minute ein- und ausgeatmet (ventiliert) wird. Es ist das Produkt aus Atemzugvolumen und Atemfrequenz. Das maximale Atemminutenvolumen (VE_{max}) ist kein Parameter für die körperliche Leistungsfähigkeit und ist deswegen für den Vergleich der Reliabilitäts- und Validitätsüberprüfung im Rahmen der Spiroergometrie ein zusätzliches Kriterium. Da es jedoch sowohl für die körperliche Leistungsfähigkeit als auch für die Ausbelastung keinen Indikator darstellt, besitzt das VE_{max} für diese Arbeit eine untergeordnete Bedeutung. Für die deskriptive Statistik und die Korrelationen zwischen t1 und t2 sei auf Anhang G1 verwiesen. Die deskriptiven Befunde zeigen insgesamt eng zusammenliegende Mittelwerte des VE_{max} bei allen vier Ausdauer tests. Leichte Steigungen von der ersten auf die zweite Testung sind bei dem 3000-m-FET von 147,7 l/Min. (t1, $SD = 30,7$ l/Min.) auf 155,3 l/Min. (t2, $SD = 27,7$ l/Min.) sowie bei dem 1000-m-Lauf von 140,9 l/Min. (t1, $SD = 24,3$ l/Min.) auf 145,9 (t2, $SD = 21,7$ l/Min.) festzustellen. Somit ist die größte Differenz der VE-Mittelwerte beim 3000-m-FET mit einem Abfall von t1 auf t2 mit 7,7 l/Min. zu verzeichnen. Die beiden Rampentests zeigen nahezu übereinstimmende Werte der durchschnittlichen VE_{max} zwischen der ersten und der zweiten Messung.

Alle Untersuchungen auf Korrelationen ergeben signifikante Ergebnisse und deuten auf hohe Zusammenhänge zwischen den Vor- und Nachtestergebnissen hin. Der größte Zusammenhang besteht zwischen dem 1000-m-Lauf mit $r = .79$ ($p = .000$) und dem 3000-m-FET ($r = .77$, $p = .000$). Der im Vergleich geringste Zusammenhang herrscht zwischen den Vor- und Nachuntersuchungen beider Rampentests. Die statistischen Zusammenhänge sind alle höchst signifikant (siehe Anhang A3).

Mittelwertvergleich zur Reliabilitätsuntersuchung des 3000-m-FET

Zum Abschluss der Reliabilitätsprüfungen sollen noch einmal die Mittelwerte verglichen und geprüft werden, ob und wie stark sich die Mittelwerte eines Parameters zwischen dem Vor- und Nachtest unterscheiden. Da der 3000-m-FET für diese Studie eine besondere Stellung einnimmt, wird auch ausschließlich dieser anhand eines t-Tests bei abhängigen Stichproben auf Mittelwertunterschiede hin geprüft. Hierbei werden die Mittelwerte aller Parameter verglichen, die zu den Messzeitpunkten t1 und t2 mit den gleichen Personen erhoben worden sind. Wie aus Tabelle 17 ersichtlich, unterscheiden sich die Mittelwerte des 3000-m-FET von Messzeitpunkt t1 und t2 bis auf das maximale Atemminutenvolumen ($p = .049$) nicht signifikant voneinander.

Tabelle 17

T-Test für abhängige Stichproben der erhobenen Parameter bei dem 3000-m-FET ($n = 29$).

Variable	3000-m-FET t1	3000-m-FET t2	t	p
Testdauer (Sek.)	255	254	.776	.444
VO _{2max} (l/Min.)	3818	3854	-.902	.375
Hf _{max} (Schläge/Min.)	175	176	-.771	.447
max. Laktatkonzentration (mmol/l)	9,9	9,5	1,34	.188
Werte der Borg-Skala	17	17	-.740	.466
VE _{max} (l/Min.)	147,7	155,3	2,06	.049

5.1.2.4.2 Ergebnisdarstellung der Validitätsüberprüfung des 3000-m-FET

Im Folgenden werden die Ergebnisse der Validitätsüberprüfung des 3000-m-FET dargestellt. Das Hauptaugenmerk liegt dabei auf der Inhaltsvalidität, also wie gut der 3000-m-FET das zu messende Merkmal (die Ausdauerleistungsfähigkeit) tatsächlich misst und auf der Konstruktvalidität, also inwiefern die Testergebnisse des 3000-m-FET mit Ergebnissen aus anderen Tests, die das gleiche Merkmal messen, übereinstimmen. Für die Validitätsüberprüfung wurden drei weitere Ausdauertests ausgewählt, die jede der Testpersonen mindestens einmal ausführte. Zum einen wurde der in der Bundeswehr verbreitete und als Bestandteil des Basis-Fitness-Tests (BFT) zur Personalauswahl durchgeführte 1000-m-Lauf herangezogen. Daneben kamen zwei Rampentests für den Vergleich zum Einsatz, die sowohl auf dem Laufband als auch auf dem Fahrradergometer im Labor unter standardisierten Bedingungen absolviert worden sind. Wie bei der Reliabilitätsüberprüfung sind die gleichen Kennwerte aus Spiroergometrie, Laktatmessung, Herzfrequenz und Borg-Skala für die Überprüfung der Gütekriterien herangezogen worden. Die Reihenfolge der Ergebnisdarstellung ergibt sich wieder aus der Relevanz der Parameter für diese Arbeit. Für die Validitätsstudie konnten insgesamt 145 Pbn gewonnen und getestet werden. Entsprechend der Qualität der Daten wird als Mittelwert, wie auch bei der Reliabilitätsüberprüfung, das arithmetische Mittel ($M = \bar{x}$) und die Standardabweichung (SD) angegeben, sowie der kleinste und größte Wert des jeweiligen Messwertbereiches aufgeführt.

Zeiten

Die Ergebnisse der benötigten Zeiten bzw. der Belastungsdauer werden in Sek. angegeben. Die beiden Streckentests, 3000-m-FET und der 1000-m-Lauf sollten in einer möglichst kurzen Zeit bewältigt werden. Bei den beiden Rampentests war dagegen entscheidend, die Belastung möglichst lange durchzuhalten und die nächste Wattstufe zu erreichen. Beim Vergleich der Minimum- und Maximum-Werte der Strecken- und Rampentests sollten hinsichtlich der Testdauer die unterschiedlichen Testkriterien berücksichtigt werden. Die benötigte Zeit stellt bei der Leistungsbeurteilung die entscheidende Größe dar, um den Testpersonen einen bestimmten Normwert zuzuordnen und einen bestimmten Fitnesszustand zuzuschreiben. Je kürzer eine Versuchsperson bspw. für die 3000-Meter-Strecke und den 1000-m-Lauf (Streckentests) gebraucht hat, desto höher wird das Fitnessniveau dieser Person eingeschätzt. Umge-

kehrt verhält es sich bei den beiden Rampentests, hier bedeutet eine höhere Zeit eine bessere Ausdauerleistungsfähigkeit. Die Tabelle 18 zeigt die deskriptive Statistik der Zeiten aller vier Ausdauerests.

Tabelle 18

Benötigte Zeiten für die verschiedenen Ausdauerests in Sek. (N = 145).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	226	165	208	307
<i>Max.</i>	347	390	847	883
<i>M</i>	263	241	552	582
<i>SD</i>	20	43	130	119

Im Durchschnitt benötigen die Versuchsteilnehmer 4:23 Min. ($SD = 20$ Sek.) für den 3000-m-FET. Die schnellste Zeit für die 3000-Meter-Strecke lag bei 3:46 Min., die langsamste Zeit bei 5:47 Min. Für den 1000-m-Lauf kann eine Durchschnittszeit von 4:01 Min. mit einer größeren SD von 43 Sek. im Vergleich zum 3000-m-FET gemessen werden. Der Rampentest auf dem Fahrradergometer wurde im Durchschnitt 9:12 Min. ($SD = 2:11$ Min.) durchgehalten, was den Beginn der Wattstufe von 370 Watt entspricht. Bei dem Rampentest auf dem Laufband brachen die Pbn den Test im Mittel bei 9:42 Min. ($SD = 1:59$ Min.) ab, was dem letzten Drittel der Geschwindigkeitsstufe von 17 km/h bzw. 18 Sek. vor dem Beginn der 18 km/h-Stufe entspricht. Nach den Testkriterien der beiden Streckentests mit einem Zeitrahmen für die Bewältigung der Strecken von 6:30 Min. haben alle Pbn die Anforderungen erfüllt. Eine Testperson absolvierte den 1000-m-Lauf genau in 6:30 Min. Der schnellste Läufer absolvierte den 1000-m-Lauf in 2:45 Min. und damit in einer wesentlich kürzeren Zeit (≈ 1 Min.) als die beste Person bei dem 3000-m-FET mit 3:46 Min.

Betrachtet man die Dauer der Tests differenziert nach Geschlecht, so zeigt sich, dass die 26 weiblichen Teilnehmer im Schnitt mit 4:52 Min. ($SD = 20$ Sek.) rund 36 Sek. mehr für die Bewältigung des 3000-m-FET benötigen als die männlichen Pbn ($M = 4:16$ Min.; $SD = 13$ Sek.). Die schnellste Probandin hat den 3000-m-FET in 4:15 Min. und die langsamste weibliche Teilnehmerin in 5:47 Min. bewältigt. Ähnlich verhält es sich bei der Betrachtung der Mittelwerte der Zeiten bei den anderen drei Ausdauerests. Die 119 Männer kamen im Durchschnitt auf 9:54 Min. ($SD = 1:37$ Min.) bei dem Rampentest auf dem Fahrradergometer und auf 10:12 Min. ($SD = 1:50$ Min.) bei dem Rampentest auf dem Laufband. Die Frauen dagegen

erreichten bei den Rampentests im Mittel Belastungszeiten von 6:04 Min. ($SD = 1:29$ Min.; Fahrradergometer) und 7:04 Min. ($SD = 1:22$ Min.; Laufband). In Bezug auf den 1000-m-Lauf konnte bei den Männer eine Durchschnittszeit von 3:45 Min. ($SD = 36$ Sek.) und bei den Frauen eine mittlere Zeit von 5:00 Min. ($SD = 42$ Sek.) vermerkt werden.

Tabelle 19

Korrelationen der benötigten Zeiten ($N = 145$).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.718	-.812	-.641
	Signifikanz (2-seitig)	.000	.000	.000

Im Hinblick auf die Zusammenhänge des 3000-m-FET mit den drei anderen Ausdauertests sei noch einmal auf die Unterschiedlichkeit der Testprotokolle verwiesen. Die Zielstellung bei den Streckentests war es, die jeweilige Strecke in möglichst kurzer Zeit zu absolvieren und bei den Rampentests dagegen entsprechend lange der kontinuierlich ansteigenden Belastung in Form einer Erhöhung der Wattzahl bzw. der Geschwindigkeit standzuhalten. Die Korrelationsanalyse in Tabelle 19 zeigt aufgrund dessen einen negativen Zusammenhang der benötigten Zeiten zwischen dem 3000-m-FET und den beiden Rampentests. Im Vergleich des 3000-m-FET mit den anderen Ausdauertests ergeben sich höchst signifikante Zusammenhänge. Die höchste Korrelation besteht zwischen dem 3000-m-FET und dem Rampentest auf dem Fahrradergometer ($r = -.81$; $p = .000$). Zwischen dem 3000-m-FET und dem 1000-m-Lauf besteht ein ebenfalls höchst signifikanter Zusammenhang mit einem Korrelationskoeffizienten in Höhe von $r = .72$ ($p = .000$). Ein negativer mittlerer Zusammenhang in Höhe von $r = -.64$ ($p = .000$) ist zwischen dem 3000-m-FET und dem Rampentest auf dem Laufband festzustellen. Der Vergleich zwischen den beiden Lauftests (in Tabelle 19 nicht abgebildet) zeigt ebenfalls einen hohen Zusammenhang von $r = -.80$ ($p = .000$), was auf die Ähnlichkeit der Belastungsart hindeutet.

Maximale Sauerstoffaufnahme ($VO_{2max/peak}$)

In Tabelle 20 sind die durchschnittlichen maximalen Sauerstoffaufnahme-Werte aller vier Tests dargestellt. Wie bereits detailliert erläutert ergibt sich wegen der Bedeutung der VO_{2max} als Indikator für die allgemeine Ausdauerleistungsfähigkeit eine besondere Relevanz dieses Parameters für das Anliegen im Rahmen der vorliegenden Arbeit.

Tabelle 20

Maximale absolute Sauerstoffaufnahme-Werte in l/Min. (N = 145).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	1779	1959	1670	2077
<i>Max.</i>	5591	5548	5137	5496
<i>M</i>	3520	3750	3543	3904
<i>SD</i>	736	725	751	771

Die im Mittel geringsten maximalen Sauerstoffaufnahme-Werte wurden bei dem 3000-m-FET festgestellt (3520 l/Min., $SD = 736$ l/Min.). Allerdings ist dieser Mittelwert lediglich geringfügig unterhalb desjenigen des 1000-m-Laufs (3750 l/Min., $SD = 725$ l/Min.). Auffällig ist jedoch, dass der maximale Wert der VO_{2max} bei dem 3000-m-FET gemessen wurde (5591 l/Min.). Den durchschnittlich höchsten VO_{2max} -Wert erreichten die Versuchsteilnehmer, wie es erwartet werden konnte, bei dem Rampentest auf dem Laufband (3904 l/Min., $SD = 771$ l/Min.). In Abbildung 26 sind die durchschnittlichen absoluten VO_{2max} -Werte der vier Ausdauertests gegenübergestellt.

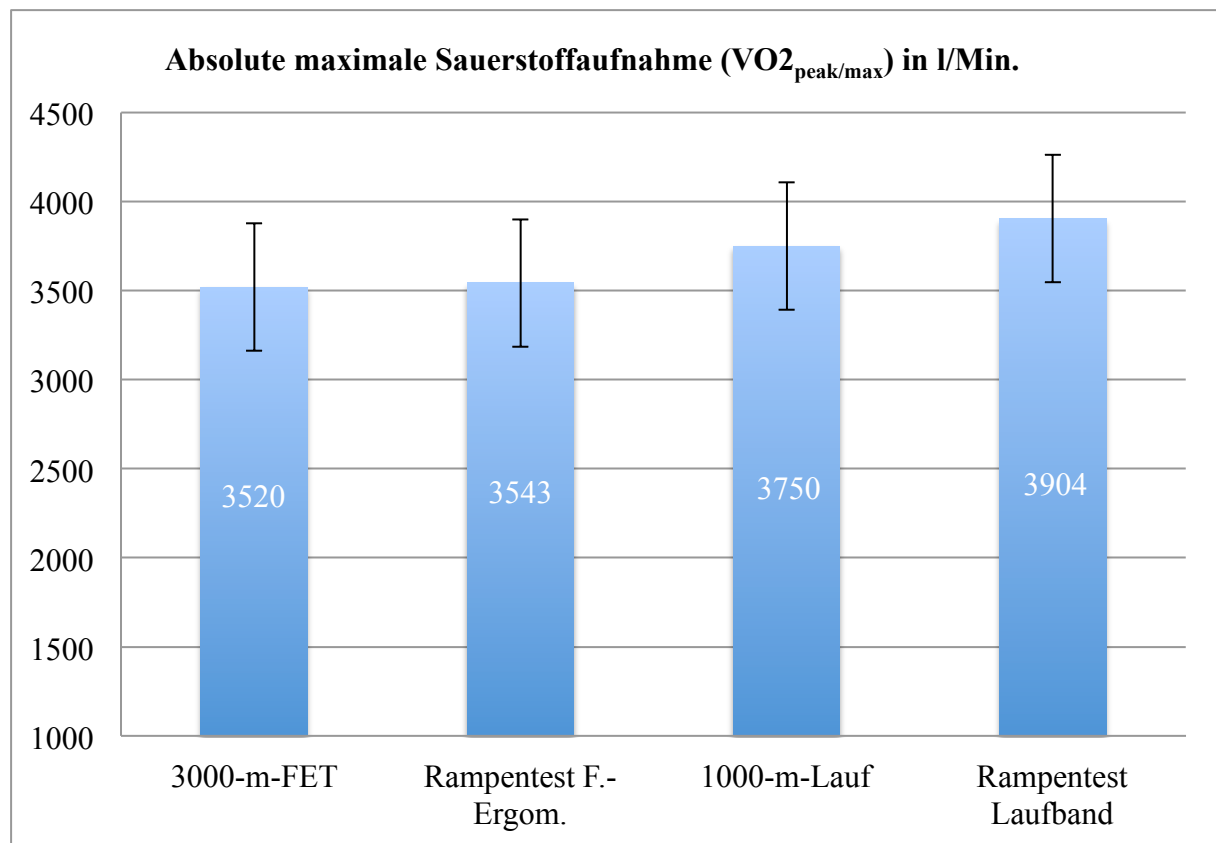


Abbildung 26. Absolute $VO_{2peak/max}$ -Mittelwerte in l/Min. der Pbn für alle vier Ausdauerarts (N = 145).

Es wurde bereits darauf hingewiesen, dass man angesichts der kleineren eingesetzten Muskelmasse bei Rad- bzw. Ergometerbelastungen aus trainingswissenschaftlicher Sicht bei der maximal erreichten Sauerstoffaufnahme nicht von der VO_{2max} , sondern von einer VO_{2peak} spricht. Der Einfachheit halber im Hinblick auf die Datenanalyse und -dokumentation der Vergleiche wird von einer Differenzierung der VO_{2max} und der VO_{2peak} im Folgenden weitestgehend abgesehen.

Tabelle 21

Korrelationen der absoluten maximalen Sauerstoffaufnahme-Werte (N = 145).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.920	.941	.911
	Signifikanz (2-seitig)	.000	.000	.000

Die Zusammenhangsanalyse zeigt, dass zwischen dem 3000-m-FET und den anderen drei Testformen ein höchst signifikanter Zusammenhang von $r > .90$ ($p = .000$) besteht. Der höchste Korrelationskoeffizient zeigt sich zwischen dem 3000-m-FET und dem Rampentest auf dem Fahrradergometer ($r = .94$; $p = .000$) und tendiert damit in Richtung eines so gut wie völligen Zusammenhangs. Der geringste, wenn man diesbezüglich (eben relativ gesehen) überhaupt von gering bzw. geringst sprechen kann, Zusammenhang zwischen den Tests zeigt sich bei dem 3000-m-FET und dem Rampentest auf dem Laufband ($r = .91$; $p = .000$). Der Vollständigkeit halber sei auch, wenn nicht in der Tabelle 21 enthalten, auf die Korrelation der VO_{2max} zwischen den beiden Lauftests hingewiesen, der ebenfalls mit $r = .95$ ($p = .000$) in Richtung eines völligen Zusammenhangs tendiert.

Normwerte für die VO_{2max} hinsichtlich einer Einordnung der Ausdauerleistungsfähigkeit existieren hauptsächlich für die relative VO_{2max} , die man im Verhältnis zum Körpergewicht berechnet (rel. $VO_{2max} = \text{ml/kg/Min.}$). Bei dem hier praktizierten Vergleich der Ausdauerests auf dem Fahrradergometer und auf dem Laufband ist der Fokus auf die absolute maximale Sauerstoffaufnahme gerichtet. Für den Vergleich der interindividuellen Ausdauerleistungsfähigkeit der Pbn ist v. a. bei Laufbelastungen die relative VO_{2max} , infolge der Miteinberechnung des Körpergewichts und der mit dem Gewicht zusammenhängenden zusätzlichen Arbeit gegen die Schwerkraft der bessere Parameter. Dagegen ist die Wirkung des Körpergewichts bei Ergometer- und Radbelastungen für den Energieaufwand von untergeordneter Rolle. Aufgrund dessen ist die absolute VO_{2max} für Vergleiche von Leistungen bei Rad- und Ergometerbelastungen zwischen Individuen ein geeigneteres Kriterium, da das Gewicht bei der absoluten VO_{2max} keine Berücksichtigung findet. Den Besonderheiten bei dem Vergleich der unterschiedlichen Ausdauerests soll Rechnung getragen werden, indem auch die relative VO_{2max} in die Datenanalyse miteinbezogen wird. Tabelle 22 zeigt im Folgenden die berechneten relativen VO_{2max} -Werte der Ausdauerests sowohl für die Fahrradergometer- als auch für die beiden Laufbandtests.

Tabelle 22

Die relativen VO_{2max} -Werte der einzelnen Ausdauertests in ml/kg/Min. ($N = 145$).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	25,9	29,9	23,9	29,9
<i>Max.</i>	65,0	64,5	62,3	69,5
<i>M</i>	45,2	48,1	45,4	50,1
<i>SD</i>	7,3	6,6	7,1	7,1

Die nachfolgende Tabelle 23 stellt die Normwerte der unterschiedlichen Leistungsniveaus dar. Im Weiteren sollen auch noch einmal die geschlechterspezifischen Unterschiede anhand der gewonnenen Daten veranschaulicht werden.

Tabelle 23

Durchschnitts- und Normwerte der VO_{2max} -Werte für unterschiedliche Leistungsniveaus (mod. nach Zintl, 1994, S. 56).

		Rel. VO_{2max}
Untrainierte	Frauen (20.-30. Lebensjahr)	32 - 38 ml/kg/Min.
	Männer (20.-30. Lebensjahr)	40 - 55 ml/kg/Min.
Hochtrainierte Ausdauer-sportler	Frauen	60 - 70 ml/kg/Min.
	Männer	80 - 90 ml/kg/Min.
Normwerte für Fitness-zustand	Frauen	35 - 38 ml/kg/Min.
	Männer	45 - 50 ml/kg/Min.
	Ausdauertrainierte	55 - 65 ml/kg/Min.
	Ausdaersportler (internationales Niveau)	65 - 80 ml/kg/Min.
	Ausdaersportler (internationales Spitzenniveau)	85 - 90 ml/kg/Min.

Betrachtet man die Korrelationskoeffizienten in der untenstehenden Tabelle 24, sind auch bei den Vergleichen zwischen dem 3000-m-FET und den übrigen Ausdauertests, insbesondere bei den Laufbandtests, höchst signifikante Zusammenhänge festzustellen. Die Korrelationsanalyse der relativen VO_{2max} zeigt nicht gleiche Werte, wie es bei der Analyse der absoluten VO_{2max} der Fall ist, jedoch befinden sich die Korrelationskoeffizienten alle deutlich im Bereich von $r > .80$.

Tabelle 24

Korrelationen der relativen VO_{2max} -Werte ($N = 145$).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.864	.899	.813
	Signifikanz (2-seitig)	.000	.000	.000

Wie der Tabelle 25 zu entnehmen ist, erscheint eine Differenzierung bei der Betrachtung der relativen VO_{2max} notwendig, um genaue Aussagen über die Ausdauerleistungsfähigkeiten treffen zu können. Insofern soll auch eine geschlechterspezifische Differenzierung der relativen VO_{2max} vorgenommen werden.

Tabelle 25

Deskriptive Statistik der relativen VO_{2max} -Werte (ml/kg/Min.) nach geschlechtsspezifischer Unterscheidung (weibl.: $n = 26$; männl.: $n = 119$).

		3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
Min.	weiblich	25,9	29,9	23,9	29,9
	männlich	30,8	36,2	30,8	37,9
Max.	weiblich	48,9	55,1	49,6	56,9
	männlich	65,0	64,5	62,3	69,5
M	weiblich	38,2	41,1	38,1	42,4
	männlich	46,8	49,7	47,1	51,8
SD	weiblich	6,5	5,6	7,0	6,7
	männlich	6,5	5,7	6,1	6,0

Nach der Normwerttabelle von Zintl (1994, S. 56; Tabelle 23) sind den weiblichen Pbn in Bezug auf die Mittelwerte der relativen VO_{2max} aller vier Ausdauertests ein überdurchschnittliches Fitnessniveau zuzuordnen ($M > 38$ ml/kg/Min.). Den weiblichen Testpersonen mit den höchsten Werten mit höher als 55 ml/kg/Min. wäre eine gute Ausdauerleistungsfähigkeit zuzuschreiben. Die Männer liegen auf Grundlage der Mittelweltergebnisse von $M \geq 45$ ml/kg/Min. bei dem 3000-m-FET, dem Rampentest (F.-Ergom.) und dem 1000-m-Lauf nach Zintl (1994) im Bereich der Fitnessnorm und bei dem Rampentest auf dem Laufband mit $M = 51,8$ ml/kg/Min. im Durchschnitt sogar etwas über der Norm. An dieser Stelle sei angemerkt, dass sich nach der Tabelle von Zintl (1994) für den Bereich „Normwerte für Fitnesszu-

stand“ keine präzise Aussage über den jeweiligen Fitnesszustand treffen lässt, d. h. es ist danach nicht zu erschließen, als wie fit die Person tatsächlich einzuordnen ist. Die Bereiche der relativen VO_{2max} der Männer und Frauen in der Kategorie „Normwerte für Fitnesszustand“ überschneidet sich mit den Bereichen für beide Geschlechter in der Kategorie „Untrainiert“.

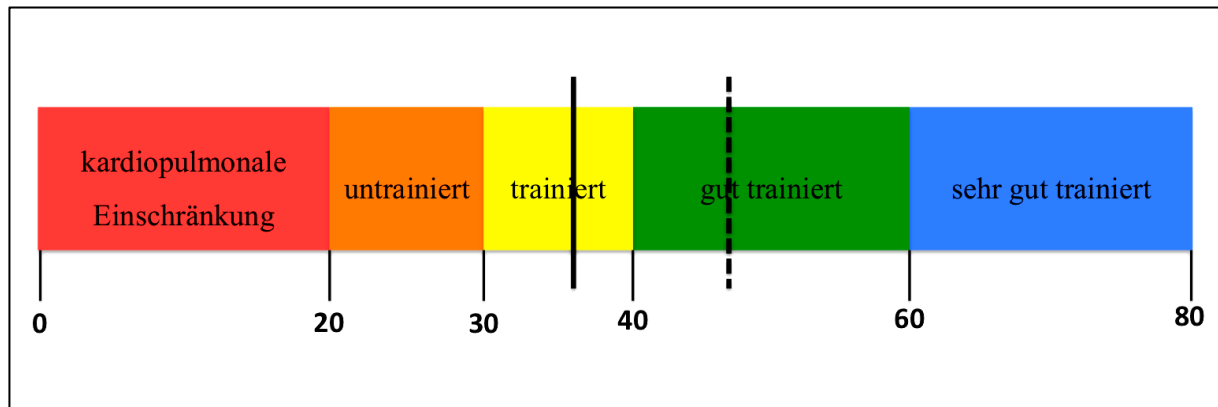


Abbildung 27. Einordnung der durchschnittlichen rel. VO_{2max} (ml/kg/Min.) des 3000-m-FET, differenziert nach Geschlecht (weibl.: $n = 26$; männl.: $n = 119$) in Anlehnung an Care Fusion (2013).

Abbildung 27 dokumentiert die Einordnung der relativen VO_{2max} des 3000-m-FET-Werte in einer Normierung nach der Spiroergometrie Software „Lab Manager“ von Care Fusion (2013), die nach der Testung eines Pbn eine Einordnung der erbrachten Ausdauerleistung bzw. des Fitnessniveaus ermöglicht. Die gestrichelte Linie zeigt den im Mittel erreichten Wert der männlichen Pbn im Bereich „gut trainiert“ mit einer durchschnittlichen relativen VO_{2max} von 46,8 ml/kg/Min. und die durchgezogene Linie im gelben Bereich („trainiert“) zeigt die Zuordnung der Durchschnittswerte der weiblichen Pbn mit einem Mittelwert von 38,2 ml/kg/Min. Im Vergleich lassen sich - je nachdem welche Normwerte für die Einordnung der relativen maximalen Sauerstoffaufnahme genommen werden - unterschiedliche Leistungsfähigkeiten auf Grundlage der rel. VO_{2max} festmachen. Liegen die Mittelwerte der weiblichen und männlichen Pbn zwischen dem 20. und 30. Lebensjahr nach Zintl (1994) im „untrainierten“ Bereich, sind die Mittelwerte der rel. VO_{2max} bei der Darstellung in Anlehnung an Fusion Care dagegen im „trainierten“ und „gut trainierten“ Bereich zu finden. Folgt man diesen Einordnungsvorschlägen, lässt sich eine klare Einordnung der rel. VO_{2max} -Werte in präzise abgestufte Leistungsbereiche nicht bewerkstelligen.

Maximale Herzfrequenzen

Die Tabelle 26 zeigt die durchschnittlichen maximalen Herzfrequenzwerte (Hf_{\max}) aller vier Tests. Daneben werden die am niedrigsten und höchsten gemessenen Hf_{\max} -Werte, sowie die Standardabweichungen (SD) der Mittelwerte (M) angegeben.

Tabelle 26

Maximale Herzfrequenzwerte in Schlägen/Min. ($N = 145$).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	135	158	147	170
<i>Max.</i>	193	209	200	212
<i>M</i>	176	182	181	191
<i>SD</i>	9,9	9,6	8,2	7,6

Die Belastungen bei den beiden Lauftests, speziell der Rampentest, haben zu den im Mittel höchsten Herzfrequenzwerten geführt. Auch die maximal erreichten Herzfrequenzwerte (209 und 212 Schlägen/Min.; $SD = 9,6$ bzw. $SD = 7,6$ Schlägen/Min.) wurden bei den beiden Tests auf dem Laufband gemessen. Der 1000-m-Lauf unterscheidet sich hinsichtlich der Hf_{\max} von dem 3000-m-FET im Mittel lediglich in sechs Schlägen/Min. mit einer Standardabweichung von rund acht bzw. zehn Schlägen/Min.

Aufgrund der Abnahme der Hf_{\max} pro Lebensdekade um etwa 6-8 Schlägen unabhängig vom Trainingszustand (Roecker, Niess, Horstmann, Striegel, Mayer & Dickhuth, 2002), erscheint an dieser Stelle eine Differenzierung der Hf_{\max} aller vier Tests im Hinblick auf das Alter angebracht. Aus Abbildung 28 wird deutlich, dass die durchschnittlichen Hf_{\max} -Werte bei allen vier Tests pro Lebensdekade abnehmen. Darüber hinaus kann durch die Faustformel $Hf_{\max} = 220$ minus Lebensalter bei den Laufbelastungen bzw. 200 minus Lebensalter für die Fahrradergometrie im Rahmen dieser Studie konstatiert werden, dass sich die Testpersonen im Durchschnitt bei allen Tests zur Überprüfung der Ausdauerleistung ausbelastet haben. Für diese Berechnung wurden die Mittelwerte der einzelnen Altersgruppen herangezogen. Bei der Altersgruppe der 30-39-jährigen ($n = 15$) ergab sich bspw. ein Mittelwert von 33 Jahren ($SD = 2,9$), der durch die Berechnung der Faustformel zu einer Hf_{\max} von 167 Schlägen/Min. für die Fahrradbelastungen und zu einer Hf_{\max} von 187 Schlägen/Min. bei den Belastungen auf dem Laufband führt. Aus dem nachstehendem Diagramm kann entnommen werden, dass

die Pbn der Gruppe der bis 29-jährigen ($n = 123$) im Mittel bei den Lauftests knapp unterhalb der Ausbelastungsgrenze geblieben sind. Die Altersgruppe der 40 bis 49-jährigen ($n = 5$) mit einem mittleren Alter von 43 Jahren ($SD = 3,5$ Jahren) hat sich im Durchschnitt bis auf den 1000-m-Lauf bei den anderen leistungsdiagnostischen Tests ausbelastet. Ansonsten konnten nach Berechnungen mit Hilfe der oben beschriebenen Faustformel bei allen Altersgruppen im Durchschnitt eine Ausbelastung im Hinblick auf die Herzfrequenz beobachtet werden. Je nach genaueren Berechnungen mit Hilfe von komplizierteren Formeln und bei einer zusätzlichen Differenzierung nach Geschlecht, lassen sich exaktere maximale Herzfrequenzwerte feststellen, wodurch sich andere, evtl. noch zutreffendere, Ausbelastungsgrenzen bestimmen ließen (siehe dazu z. B. Roecker et al., 2002).

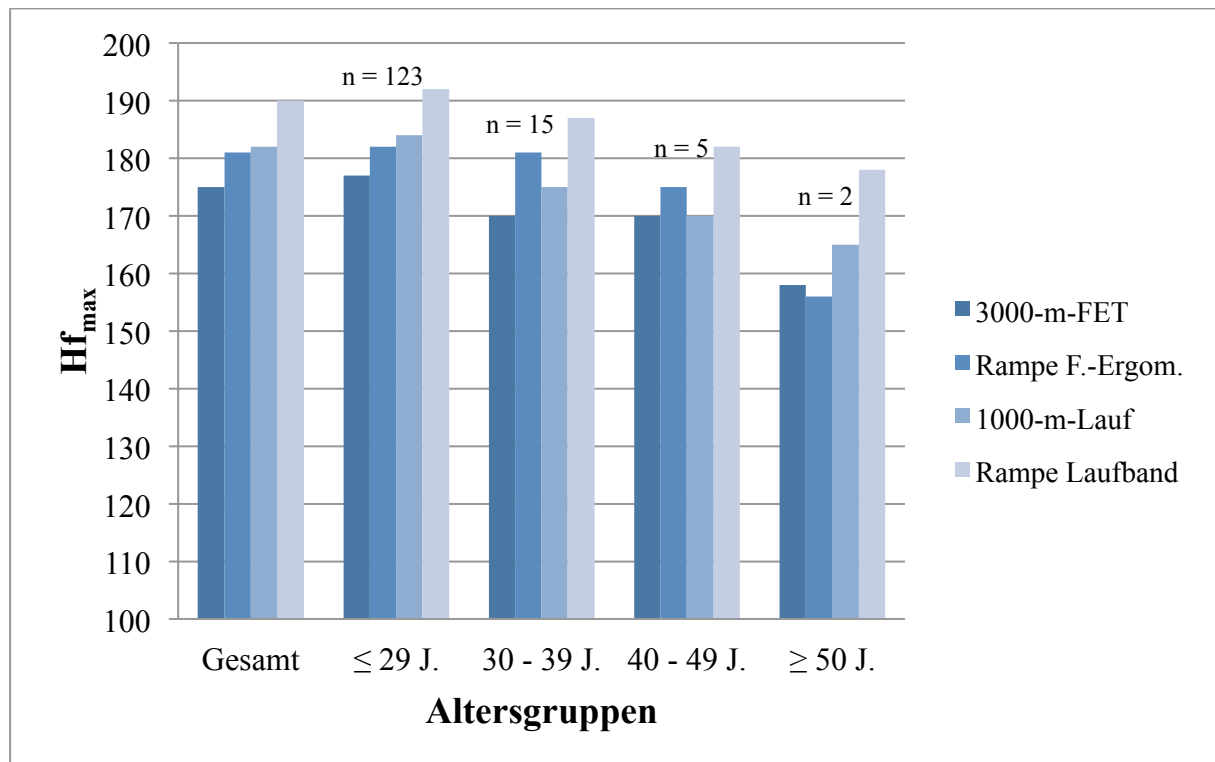


Abbildung 28. Hf_{\max} -Werte in Abhängigkeit vom Alter eingeteilt nach Lebensdekaden ($N = 145$).

Tabelle 27 zeigt die Zusammenhänge der maximalen Herzfrequenzen zwischen den Ausdauer-erests. Der höchste Zusammenhang besteht zwischen dem 3000-m-FET und dem 1000-m-Lauf ($r = .48$; $p = .000$). Zwischen dem 3000-m-FET und den beiden Rampentests bestehen ebenfalls mittlere und höchst signifikante Zusammenhänge in Höhe von $r = .48$ ($p = .000$) und $r = .42$ ($p = .000$).

Tabelle 27

Korrelationen zwischen den Hf_{max} -Werten bei verschiedenen Ausdauertests ($N = 145$).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.481	.464	.417
	Signifikanz (2-seitig)	.000	.000	.000

Laktatwerte

Die gemessenen Laktatkonzentration im Blut der Testpersonen werden in Tabelle 28 dargestellt. Bei Betrachtung der Laktatwerte darf der Einfluss des Untersuchungsdesigns, im speziellen die Pausenzeit zwischen den jeweiligen Tests, nicht unberücksichtigt bleiben. Da zu einem Testzeitpunkt jeweils zwei Tests hintereinander absolviert wurden (t1: 3000-m-FET und Rampentest F.-Ergom.; t2: 1000-m-Lauf und Rampentest Laufband) und die 30 bis 45-minütige Pausenzeit zwischen den Testungen keine vollständige Regeneration bzw. keinen vollständigen Laktatabbau gewährleistet, ist davon auszugehen, dass die Vorbelastungen die Laktatwerte nach den Rampentests beeinflussen. Insofern diente die Feststellung der Laktatkonzentrationen hier lediglich als objektiver Gradmesser für die Ausbelastung der Versuchsteilnehmer. Eine Laktatkonzentration von > 8 mmol/l als Maß wird i. d. R. für eine gute Ausbelastung akzeptiert (siehe z. B. aus der Fünten et al., 2013; Heck, 1990).

Tabelle 28

Maximale Laktatkonzentrationen in mmol/l ($N = 145$).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	4,2	2,9	4,1	4,2
<i>Max.</i>	21,4	19,0	15,2	17,6
<i>M</i>	11,7	9,5	10,2	9,3
<i>SD</i>	3,3	3,0	2,2	2,5

Die beiden Rad-Tests, speziell der 3000-m-FET, weisen die im Mittel höchsten Laktatwerte auf. Bei dem 3000-m-FET erreichen die Pbn im Durchschnitt einen Wert in Höhe von 11,7 mmol/l ($SD = 3,3$ mmol/l) und liegen somit deutlich im Ausbelastungsbereich. Die im Mittel

niedrigsten Laktatkonzentrationen sind bei dem Rampentest auf dem Laufband diagnostiziert worden (9,3 mmol/l, $SD = 2,5$ mmol/l). Durchschnittlich liegen die unmittelbar nach Testbeendigung analysierten Laktatwerte aller durchgeführten Tests deutlich im Bereich des Ausbelastungszustandes. Im Durchschnitt weisen die Pbn sogar Laktatwerte nach den beiden Fahrradergometertests über zehn mmol/l auf. Die höchsten Laktatwerte können nach dem 3000-m-FET festgestellt werden, was auch durch den am höchsten gemessenen Laktatwert von 21,4 mmol/l deutlich wird. Somit kommt es v. a. bei diesem Streckentest, der i. d. R. in 4:03 bis 4:43 Min. beendet wurde, zu einer relativ schnellen Übersäuerung der Muskulatur.

Tabelle 29

Korrelationen der maximalen Laktatkonzentrationen bei den verschiedenen Ausdauerests ($N = 145$).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.530	.610	.472
	Signifikanz (2-seitig)	.000	.000	.000

Die Korrelationstabelle 29 stellt die Zusammenhänge bezüglich der Höhe der Laktatkonzentrationen dar, die im Anschluss an die Tests gemessen werden konnten. Hierbei zeigt sich der höchste Zusammenhang zwischen dem 3000-m-FET und dem ebenfalls auf dem Fahrradergometer durchgeführten Rampentest mit einem Korrelationskoeffizienten von $r = .61$ ($p = .000$). Gleichmaßen weisen die beiden Streckentests einen mittleren und höchst signifikanten Zusammenhang in Höhe von $r = .53$ ($p = .000$) auf. Der niedrigste Zusammenhang hinsichtlich der gemessenen Laktatkonzentration besteht zwischen dem 3000-m-FET und dem Rampentest, der auf dem Laufband durchgeführt wurde ($r = .47$; $p = .000$). Wegen der gleichen Tätigkeit interessiert an dieser Stelle auch der Vergleich zwischen dem 1000-m-Lauf und dem Rampentest auf dem Laufband, der i. H. v. $r = .72$ ($p = .000$) auf eine starke Ähnlichkeit der Belastung hinsichtlich der Laktatkonzentration hindeutet.

Borg-Skala

Wie bereits beschrieben, wurde als subjektives Ausbelastungskriterium die Borg-Skala herangezogen, womit die Pbn direkt nach Testbeendigung und der Abnahme der Atemmaske ihre Einschätzung bezüglich des Beanspruchungsgrades abgeben sollten. Die im Mittel angegebenen Werte der Borg-Skala veranschaulichen, wie stark sich die Testpersonen von den jeweiligen Tests ausbelastet gefühlt haben. Aus Tabelle 30 wird ersichtlich, dass die im Durchschnitt höchsten Werte bei den beiden Rampentests ($M = 18$; $SD = 1,3$) dokumentiert werden konnten.

Tabelle 30

Angegebene Werte der Borg-Skala als Ausbelastungskriterium (N = 145).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	13	14	14	14
<i>Max.</i>	20	20	20	20
<i>M</i>	17	17	18	18
<i>SD</i>	1,3	1,6	1,3	1,3

Der 3000-m-FET wurde nach der Frage: „Wie anstrengend fanden Sie den Test?“ im Schnitt mit „sehr schwer“, was dem Wert 17 ($SD = 1,3$) auf der Borg-Skala entspricht, beantwortet. Die Korrelationsanalysen der Borg-Werte weisen Zusammenhänge im mittleren bis niedrigen Bereich $r < .50$ auf. Die Zusammenhangsanalyse zwischen dem 3000-m-FET und dem Rampentest auf dem Laufband und dem Fahrradergometer ist hinsichtlich der Werte der Borg-Skala weder signifikant noch besteht ein Zusammenhang. Der deutlichste Zusammenhang der Borg-Werte zeigt sich beim 3000-m-FET und dem 1000-m-Lauf ($r = .43$; $p = .000$; siehe Tabelle 31).

Tabelle 31

Korrelationen der Werte der Borg-Skala (N = 145).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.431	.113	.113
	Signifikanz (2-seitig)	.000	.181	.181

Respiratorischer Quotient (RQ)

Der RQ-Wert (häufig auch als RER = „respiratory exchange rate“ bezeichnet) stellt ein weiteres Ausbelastungskriterium dar, das als vierter Indikator Aufschluss darüber geben soll, ob sich die Untersuchungsteilnehmer tatsächlich ausbelastet haben. In erster Linie zeigt der RQ einen Stoffwechsellage bzw. eine -situation der Arbeitsmuskulatur an, d. h. er gibt Auskunft über eine bestimmte Inanspruchnahme von Kohlenhydraten und vom Fettstoffwechsel während Dauerbelastungen und ist deswegen ein metabolischer Parameter der Ausbelastung (Haber, 2007; Meyer, 2003). In der Literatur wird generell von einem Wert der Ausbelastung gesprochen, wenn ein RQ von $> 1,1$ vorliegt (siehe hierzu bspw. Dickhut, 2000; Haber, 2007).

„Ein RQ von > 1 zeigt an, dass ein Pbn oder Patient metabolisch ausbelastet war. Diese Feststellung ist unabhängig von der erbrachten Leistung, da der gleiche Wert von, sagen wir, 1,2 sowohl bei schwachen als auch bei hochtrainierten Pbn bei Ausbelastung auftritt. Ein RQ von > 1 ist daher ein Wert, der objektiv die Ausbelastung anzeigt. Wird eine Spiroergometrie bei einem RQ < 1 abgebrochen, so kann angenommen werden, dass zumindest metabolisch keine Ausbelastung vorgelegen hat“ (Haber, 2007, S. 151).

In der Tabelle 32 sind die durchschnittlichen RQ-Werte sowie die maximal erreichten und die minimal gemessenen RQ-Werte für alle vier Ausdauertests dokumentiert.

Tabelle 32

Deskriptive Statistik der RQ-Werte bei den verschiedenen Ausdauertests (N = 145).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	1,0	1,0	1,1	1,0
<i>Max.</i>	1,5	1,4	1,4	1,1
<i>M</i>	1,2	1,2	1,2	1,2
<i>SD</i>	0,1	0,1	0,1	0,1

Die Ergebnisse zeigen, dass die durchschnittlichen RQ-Werte deutlich die Grenze von über 1,1 bzw. über 1, was für eine Ausbelastung spricht, deutlich übersteigen und somit auch bei Betrachtung dieses Parameters von einer Ausbelastung der getesteten Pbn ausgegangen werden kann.

Atemminutenvolumen

Dem VE kommt, wie erläutert, was die Relevanz in dieser Arbeit im Vergleich zu den zuvor dargestellten Parametern betrifft, eine untergeordnete Rolle zu. Sie soll allerdings im Rahmen der Testvalidierung des 3000-m-FET nicht unbeachtet bleiben. Deswegen wird abschließend im Rahmen der Ergebnisdarstellung der Überprüfung der Testgütekriterien auch auf das VE_{\max} Bezug genommen. Für die genaue Darstellung der Ergebnisse sei auf Anhang A verwiesen.

Die Pbn erreichten bei den Rampentests die im Durchschnitt höchsten VE_{\max} -Werte mit 141 l/Min. ($SD = 29,2$ l/Min.) bei der Rampe auf dem Laufband und 140 l/Min. ($SD = 33,2$ l/Min.) bei der Fahrradergometrie. Bei dem 3000-m-FET wiesen die Testpersonen einen durchschnittlichen VE_{\max} -Wert von 137 l/Min. ($SD = 35,1$ l/Min.; siehe Anhang G2) auf. Zwischen dem 3000-m-FET und den anderen drei Ausdauertests zeigen sich hinsichtlich der maximal erreichten Atemminutenvolumina höchst signifikante Zusammenhänge i. H .v. $r > .70$ ($p = .000$). Die Befunde der Zusammenhangsanalysen der maximalen Atemminutenvolumina können Anhang A7 entnommen werden.

Leistung in Watt

Zusätzlich zu den spiroergometrischen Messgrößen $VO_{2\max}$ sowie der Belastungsdauer der vier Tests konnten durch die Tacx-Software noch weitere Leistungsparameter bezüglich der beiden Rad-Tests berücksichtigt werden, die die Ergebnisdarstellung der Befunde zur Testentwicklung hinsichtlich der physischen Belastbarkeitskomponente abschließen sollen. In der Tabelle 33 sind die durchschnittlichen maximalen Leistungen beider Tests in Watt, die absolute Durchschnittsleistung in Watt, sowie die relative Durchschnittsleistung in Watt/kg angegeben.

Tabelle 33

Durchschnittliche Leistungen (in Watt) bei den beiden Fahrrad-ergometertests (N = 145).

3000-m-FET	max. Leistung (Watt)	Durchschn. Leistung (Watt)	Rel. durchschnittl. Leistung (Watt/kg)
<i>Min.</i>	222	138	1,1
<i>Max.</i>	789	458	5,4
<i>M</i>	469	306	3,9
<i>SD</i>	114,1	61,7	0,7
Rampentest F.-Ergom.	max. Leistung (Watt)	Durchschn. Leistung (Watt)	Rel. durchschnittl. Leistung (Watt/kg)
<i>Min.</i>	230	89	1,5
<i>Max.</i>	566	325	4,1
<i>M</i>	411	227	2,9
<i>SD</i>	72,5	42,7	0,5

Die Pbn bewältigten den 3000-m-FET mit einer durchschnittlichen Wattleistung in Höhe von 306 Watt ($SD = 61,7$ Watt/kg) und einer durchschnittlichen Wattleistung im Verhältnis zum Körpergewicht von 3,9 Watt/kg ($SD = 0,7$ Watt/kg). Die höchste Leistung aller Pbn lag bei einer durchschnittlichen Wattleistung von 458 Watt und die geringste Leistung, mit der der 3000-m-FET im Durchschnitt bewältigt wurde, bei 138 Watt. Die im Mittel erreichte Leistung bei dem Rampentest lag bei rund 411 Watt ($SD = 72,5$ Watt), was etwa der zehnten Stufe des Rampentests (Abbruch nach 10 Min.) entspricht. Die maximal erreichte Stufe wurde bei ≈ 560 Watt gemessen und die geringste Stufe, bei der der Rampentest zum Abbruch geführt hat, lag bei ≈ 230 Watt.

5.2 Verfahrensentwicklung zur Erfassung der psychischen Belastbarkeitskomponente

Für die Erhebung der psychischen Parameter im Rahmen des psycho-physischen Belastbarkeitskonzepts kamen im wesentlichen zwei Instrumente zum Einsatz, mit denen die in dieser Arbeit relevanten Persönlichkeitsmerkmale Emotionale Stabilität und Belastbarkeit erfasst wurden. Zum einen kam bei allen Untersuchungen eine „Langform“ (ursprüngliche Skalen mit allen Items) der beiden Skalen zur emotionalen Stabilität und Belastbarkeit zum Einsatz. Diese wurden, wie bereits in Kapitel 4.3.1 beschrieben, aus einem gängigen, empirisch hinreichend evaluierten, in der Praxis weit verbreiteten und bereits bewährten Persönlichkeitsinventar herangezogen. Es handelt sich dabei um das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) von Hossiep und Paschen (1998). Darüber hinaus wurde ein weiteres und eigens für diese Studie entwickeltes Verfahren zur Erfassung der psychischen Komponente eingesetzt, das in diesem Abschnitt der Arbeit auf seine Nützlichkeit für zukünftige Untersuchungen in diesem Bereich geprüft werden soll. Dieses Verfahren beruht auf den beiden Skalen Emotionale Stabilität und Belastbarkeit des BIP. Auf deren Grundlage wurde eine „Kurzform“ dieser Skalen entwickelt. Die Herangehensweise der Entwicklung und der Sinn und Zweck dieser neu konzipierten Kurzform ist den Kapiteln 4.3.2 und 4.3.3 zu entnehmen. In diesem Abschnitt soll diese Kurzform auf die relevanten Gütekriterien hin untersucht werden und v. a. durch Zusammenhangsanalysen der Frage nachgegangen werden, inwiefern die Kurzform der Langform entspricht oder gar ebenbürtig ist und für die Überprüfung der psychischen Belastbarkeit (auch) ein geeignetes Instrument darstellt. Das besondere Anliegen der Entwicklung dieser Kurzform liegt in dem Interesse eines möglichst zeitökonomischen Verfahrens zum Screening im Rahmen der Datengenerierung zur Zusammenhangsüberprüfung der psycho-physischen Belastbarkeit.

5.2.1 Überprüfung der Kurz- und Langform-Skalen zur psychischen Belastbarkeit

An dieser Stelle sei nochmals die hohe Interkorrelation der beiden Skalen Emotionale Stabilität und Belastbarkeit, wie sie im BIP konzipiert sind, herausgestellt. Es wurde bereits erläutert (siehe Kapitel 4.3.1), dass die Autoren selbst (Hossiep & Paschen, 1998) darauf hinweisen, dass die beiden Skalen zu einer Skala für die Erfassung psychischer Belastbarkeit zusammengefasst werden könnten. Vor dem Hintergrund eines eingeschränkten Zeitfensters für solche

Testungen ergaben sich Fragen nach der Ökonomisierung des Fragebogenverfahrens und so die Überlegung, aus den beiden hoch interkorrelierenden Skalen eine Kurzform zu entwickeln, welche

- die beiden Merkmale als Komponenten psychischer Belastbarkeit erfasst, wie die Kombination der beiden Skalen der Langform.
- eine geeignete und zeiteffizientere Alternative für die Erfassung der psychischen Belastbarkeit darstellt, wenn sie sich als valide herausstellt.
- als zeitökonomisches Screeninginstrument nutzbar ist.

Für die Kurzform wurden aus beiden Skalen der Langform Items mit einer hohen Faktorladung sowohl positiv als auch negativ gepolter Items von $r_f > .65$ ausgewählt und aus diesen Items jeweils vier passende Adjektive für eine positiv und eine negativ formulierte „Wolke“ als zusammenfassende Darstellung der Dimension und Realisationsform der Kurzform abgeleitet (siehe Kapitel 4.3.2). Im Weiteren galt es dann zu prüfen, inwiefern die durch die beiden Wolken konzipierte Kurzform den beiden Skalen aus dem BIP entspricht und die Frage zu beantworten, ob die Kurzskala tatsächlich ein geeignetes und vergleichbares Instrument zur Erfassung der psychischen Belastbarkeit darstellt. Dazu konnte von folgender Feststellung ausgegangen werden: Die beiden Skalen aus dem BIP „Emotionale Stabilität“ und „Belastbarkeit“ korrelieren in hohem Maße ($r \geq .68; p \leq .001$) und können somit zu einer Skala (siehe Hossiep & Paschen, 1998) „psychische Belastbarkeit“ zusammengefasst werden. Darauf aufbauend wurde von folgenden Annahmen ausgegangen, die auch die nachfolgenden Ergebnisdarstellungen strukturieren:

(H6) Die Kurzform-Skala zur Erfassung der psychischen Belastbarkeit weist einen hohen Zusammenhang mit der Langform auf.

(H7) Sowohl die Lang- als auch die Kurzform-Skala zur Erfassung des Persönlichkeitsmerkmals „psychische Belastbarkeit“ weisen in Hinblick auf die divergente Validität einen negativen Zusammenhang mit der Skala „Neurotizismus“ des Persönlichkeitsinventars „NEO-FFI“ (Costa & McCrae, 1992; siehe Kapitel 4.3.3) auf.

(H8) Die Kurz- und Langform der Skalen zur psychischen Belastbarkeit weisen hinsichtlich der konvergenten Validität einen positiven Zusammenhang mit der Skala der allgemeinen

Selbstwirksamkeitserwartung (SWE; Schwarzer & Jerusalem, 1999; siehe Kapitel 4.3.3) auf.

In den folgenden Kapiteln soll diesen Überlegungen und Annahmen nachgegangen werden und anhand des erhobenen Datenpools analysiert werden, inwiefern sich insbesondere die aus Ökonomisierungszwecken eigens konzipierte Kurzskala für die praktische Anwendung bewährt, wie präzise das zu erfassende Merkmal psychische Belastbarkeit durch die Kurzskala gemessen wird und ob sich im Vergleich zur Langform-Skala Zusammenhänge zeigen, die auf eine Anwendbarkeit hinweisen. Für die detaillierte Darstellung der Entwicklung der beiden Skalen zur psychischen Belastbarkeit bzw. des methodischen Vorgehens sei auf Kapitel 4.3 verwiesen.

5.2.1.1 Korrelationsanalysen zur Kurz- und Langform der psychischen Belastbarkeitsskalen

Zum Vergleich der Kurz- und Langform der Skalen zur Messung des Persönlichkeitsmerkmals psychische Belastbarkeit wurde eine bivariate Korrelationsanalyse nach Pearson berechnet, um so zunächst zu analysieren, ob zwischen den beiden Instrumenten ein Zusammenhang besteht. Wie bereits in Kapitel 4.3.2 erläutert, wurde das positiv gepolte Item (Wolke) mit den Adjektiven „robust“, „standhaft“, „willensstark“ und „stabil“ vor der Analyse mit SPSS umgepolzt, sodass „trifft voll zu“ mit dem höchsten Wert (6) und die Antwortmöglichkeit „trifft überhaupt nicht zu“ mit dem geringsten Wert (1) versehen worden sind. Der höchste zu erreichende Wert bzw. Summenscore bei der Kurzskala ist demzufolge zwölf und der niedrigste Wert ist zwei. Auf Seiten der Langform-Skala mit den beiden Skalen des BIP wurden ebenfalls die entsprechenden Umpolungen vorgenommen. Bei der Skala Belastbarkeit betraf dies die Items 2, 4, 5, 6 und 8, bei der Skala Emotionale Stabilität die Items 1, 2, 3, 5, 10, 11, 12, 13, 15 und 16 (siehe Hossiep & Paschen, 1998). Nach Umpolung dieser Items ergibt sich aufgrund der insgesamt 29 Items der Langform-Skala ein maximaler Itemscore der psychischen Belastbarkeitsskala von 174 und ein minimaler Score von 29. In Tabelle 34 sind die durchschnittlichen Itemscores sowohl der langen als auch der kurzen Version der Skala des zu erfassenden Persönlichkeitsmerkmals psychische Belastbarkeit zu sehen. Bei allen Untersuchungen zur psycho-physischen Belastbarkeit, kam zur Erfassung der psychischen Belastbarkeitskomponente sowohl die eigens entwickelte Kurzskala als auch die Langform-Skala zum Einsatz, somit füllten alle Testpersonen beide Skalen der psychischen Belastbarkeit aus. Daraus ergibt sich eine Gesamtstichprobe von $N = 424$. Alle 424 Pbn bewältigten auch den 3000-m-FET, Details zur Stichprobe hinsichtlich der Alters- und Geschlechterverteilung sind im

Kapitel 5.3 zur Untersuchungen der psycho-physischen Belastbarkeit zu finden. Die Teilstichproben werden im Rahmen der jeweiligen Ergebnisdarstellung beschrieben. Tabelle 34 zeigt die Mittelwerte und deren Standardabweichung, sowie die am geringsten und höchsten erreichten Itemscores der psychischen Belastbarkeitsskalen.

Tabelle 34

Deskriptive Darstellung der Itemscores beider Skalen (Lang- und Kurzform) zur Erfassung der psychischen Belastbarkeit (N = 424).

	Kurzform-Skala: psychische Belastbarkeit	Langform-Skala: psychische Belastbarkeit
<i>Min.</i>	4	48
<i>Max</i>	12	171
<i>M</i>	9,7	119,2
<i>SD</i>	1,4	18,6

Grundsätzlich kann durch einen hohen Score bei der Skala psychische Belastbarkeit davon ausgegangen werden, dass sich solche Individuen robuster gegenüber Belastungen zeigen als Personen mit einem niedrigeren Score. Personen mit einem hohen Itemscore beschreiben sich selbst als eher gelassen, stress- und frustrationstolerant, resistent gegen Depressionen und auch bei hohen Belastungen grundsätzlich als leistungsfähig (siehe dazu auch Kapitel 2.5).

Tabelle 35 zeigt den Zusammenhang der Kurz- und Langform-Skalen zur Erfassung des Persönlichkeitsmerkmals psychische Belastbarkeit. Der Korrelationskoeffizient in Höhe von $r = .60$ ($p = .000$) weist auf einen mittleren (höchst signifikanten) Zusammenhang der beiden Skalen hin. Diesem Befund nach entspricht die Kurzskala psychische Belastbarkeit im mittleren Maße der auf dem BIP basierenden Langform-Skala.

Tabelle 35

Korrelationsanalyse der beiden Skalen (Kurz- und Langform) zur Erfassung der psychischen Belastbarkeit (N = 424).

		Langform-Skala: psychische Belast- barkeit	BIP-Skala: Emotionale Stabilität	BIP-Skala: Belastbarkeit
Kurzform-Skala: psychische Belastbarkeit	Korrelation nach Pearson	.597	.524	.545
	Signifikanz (2-seitig)	.000	.000	.000

Darüber hinaus wurden die beiden Skalen der Langform „Belastbarkeit“ und „Emotionale Stabilität“ noch einmal gesondert mit der Kurzform verglichen. Die Zusammenhangsüberprüfungen zeigen, dass beide Korrelationskoeffizienten jeweils i. H. v. $r > .50$ ($p = .000$) liegen. Interessant ist auch der Zusammenhang der beiden BIP-Skalen. Die Ergebnisse von Hossiep und Paschen (1998), die einen Zusammenhang i. H. v. $r = .68$ ($p = .000$) feststellten, und aufgrund dessen eine hohe Interkorrelation der beiden Skalen konstatierten, sodass sie auf die Möglichkeit der Integration der beiden Skalen zu einer Skala hinwiesen, konnte in dieser Untersuchung mit einem Korrelationskoeffizienten von $r = .60$ ($p = .000$) nicht ganz bestätigt werden. Das Ausmaß der Interkorrelation lässt es allerdings durchaus als gerechtfertigt erscheinen, weiterhin von einer Skala zur Erfassung der psychischen Belastbarkeit auszugehen.

5.2.1.2 Korrelationsanalysen zum State- und Trait-Aspekt mittels der Kurzform-Skalen

Bei einer Teilstichprobe ($n = 275$) kamen neben der Trait- auch zwei State-Versionen des Fragebogens der Kurzsкала zum Einsatz. In Leistungssituationen, wie bei Eignungsdiagnostiken, spielt auch das aktuelle Befinden eine wesentliche Rolle und soll deswegen in die Analyse miteinbezogen werden. Die State-Version der Kurzsкала wurde bei den Untersuchungen jeweils einmal vor und einmal nach dem 3000-m-FET abgefragt, um Veränderungen des Empfindens durch die körperliche Belastung zu erfassen. Beide Aspekte wurden mit der gleichen Skala (Kurzform zur Erfassung der psychischen Belastbarkeit) gemessen. Der Unterschied bestand lediglich in der Instruktion der Skala. Während für die Trait-Version der Kurzform die Anweisung lautete, anhand der in der Wolke zusammengefassten Adjektive anzugeben, inwiefern diese „im Allgemeinen“ zutreffen, lautete die Anweisung der beiden State-

Versionen (vorher und nachher), inwiefern die Adjektive in der Wolke „im Moment“ zutreffen. Da sich in Leistungssituationen Veränderungen hinsichtlich der abgefragten Adjektive „nervös“, „reizbar“, „instabil“, „entmutigt“ und „gelassen“, „willensstark“, „robust“, „standhaft“ in Abhängigkeit von der Situation ergeben können, erschien eine Abfrage bzw. Messung des aktuellen Befindens bzw. des aktuellen Zustands (State) neben der Messung der Eigenschaft bzw. des Persönlichkeitsmerkmals (Trait) als sinnvoll. In der nachstehenden Tabelle 36 sind die Ergebnisse der Korrelationsanalysen veranschaulicht.

Tabelle 36

Deskriptive Darstellung der Itemscores der beiden State-Kurzform Skalen zur Erfassung der psychischen Belastbarkeit (n = 275).

	Kurzform-Skala: psychische Belastbarkeit (State vorher)	Kurzform-Skala: psychische Belastbarkeit (State nachher)
<i>Min.</i>	2	4
<i>Max.</i>	12	12
<i>M</i>	9,8	9,8
<i>SD</i>	1,7	1,8

Die 275 Personen, die neben der Kurzform (Trait) auch die beiden State-Versionen (vorher und nachher) der psychischen Belastbarkeitsskala ausgefüllt hatten, erreichten durchschnittlich einen Itemscore von 9,8 ($SD = 1,7; 1,8$), sowohl bei der State-Version der Kurzskaala, die vor den physischen, als auch bei der State-Version, die nach den physischen Tests eingesetzt wurden, um eventuelle Befindensänderungen zu überprüfen.

Tabelle 37

Korrelationsanalyse der beiden eingesetzten Kurzform-Skalen (State und Trait; n = 275).

		Kurzform-Skala: psychische Belastbarkeit (State vorher)	Kurzform-Skala: psychische Belastbarkeit (State nachher)
Kurzform-Skala: psychische Belastbarkeit (Trait)	Korrelation nach Pearson	.559	.432
	Signifikanz (2-seitig)	.000	.000

Die Korrelationskoeffizienten weisen auf einen mittleren (höchst signifikanten) Zusammenhang zwischen der Kurzsкала für das Persönlichkeitsmerkmal (Trait) und den beiden Kurzscales, die den aktuellen Befindlichkeitszustand (State) ermitteln hin. Im Hinblick auf den Zusammenhang zwischen der Kurzsкала (State), die vor dem Test zum Einsatz kam und der Kurzsкала (State), die nach der Leistungsdiagnostik vorgelegt wurde, ergibt sich ebenfalls ein mittlerer Zusammenhang von $r = .45$ ($p = .000$). Dieser Befund verdeutlicht so geringfügige Befindensunterschiede (siehe auch die Mittelwerte und Streuungen; Tabelle 36), dass sie hier nicht weiter verfolgt werden sollen.

5.2.2 Analysen zur Überprüfung der konvergenten und divergenten Validität

Zur Überprüfung der Konstruktvalidität der psychischen Belastbarkeit, operationalisiert mittels der verschiedenen Skalen (Kurz- und Langform), wurde die Multitrait-Multimethod-Methode nach Campbell und Fiske (1959) herangezogen. Bei dieser Methode wird zum Nachweis der Konstruktvalidität zum einen ein systematischer Vergleich mit anderen Merkmalen (Traits), die das gleiche oder ein anderes Konstrukt messen, vorgenommen und zum anderen der Vergleich mit unterschiedlichen Messmethoden bzw. Instrumenten (Persönlichkeitsinventaren) durchgeführt. Bei der konvergenten Validitätsüberprüfung werden andere Testverfahren bzw. Instrumente (Fragebogen, Skalen, Inventare) ausgewählt, die das gleiche Merkmal erfassen. Hohe Korrelationen zwischen den verschiedenen Instrumenten, die das gleiche Merkmal messen, belegen nach dieser Methode die konvergente Konstruktvalidität des neu entwickelten Verfahrens. Die Skala zur Erfassung der psychischen Belastbarkeit, die sich aus den beiden Skalen des BIP „Emotionale Stabilität“ und „Belastbarkeit“ zusammensetzt, sowie die daraus entwickelte Kurzsкала, sollten demnach mit anderen Verfahren in hohem Maße korrelieren, die die gleichen Persönlichkeitsmerkmale messen. Nach Campbell und Fiske (1959) ist neben der konvergenten auch die divergente Validität eine bewährte methodische Herangehensweise für die Überprüfung der Konstruktvalidität. Diesem Konzept liegt die Überlegung zugrunde, dass Messungen von unterschiedlichen Persönlichkeitsmerkmalen miteinander geringer oder negativ korrelieren sollten. Demzufolge sollte ein neu entwickelter Fragebogen zur Erfassung des Merkmals Emotionale Stabilität bzw. Belastbarkeit gering oder negativ mit einem Verfahren zur Messung des Trait-fremden Konstrukts „Extraversion“ korrelieren (vgl. Schermelleh-Engel & Schweizer, 2008).

5.2.2.1 Psychische Belastbarkeit und Neurotizismus (NEO-FFI)

Für die Überprüfung der divergenten Validität erscheint das der emotionalen Stabilität komplementäre Merkmal emotionale Labilität am sinnvollsten, das durch die Neurotizismus-Skala des NEO-FFI (Costa & McCrae, 1992) erfasst wird. Die Neurotizismus-Skala des NEO-FFI besteht aus zwölf Items mit einer fünfstufigen Antwortkategorie, die die selbsteingeschätzte emotionale Labilität einer Person messen. Dementsprechend ergibt sich ein minimal erreichbarer Itemscore von 12 und ein maximal zu erreichender Itemscore von 60 für die Neurotizismus-Skala. Die nachfolgende Tabelle 38 zeigt u. a. die Mittelwerte der Neurotizismus-Skala.

Tabelle 38

Deskriptive Darstellung der Itemscores der Neurotizismus-Skala des NEO-FFI zur Erfassung der emotionalen Labilität (n = 155).

	Neurotizismus
<i>Min.</i>	14
<i>Max.</i>	50
<i>M</i>	25
<i>SD</i>	6,4

Aufgrund zweier Skalen, die gegensätzliche Merkmale erfassen, sollten dementsprechend die nachstehenden Ergebnisse den Erwartungen einer mittleren bis hohen negativen Korrelation zwischen den Skalen Neurotizismus und psychische Belastbarkeit gerecht werden. Wie aus Tabelle 39 zu entnehmen ist, konnte hinsichtlich der divergenten Validität ein Zusammenhang i. H. v. $r = -.62$ ($p = .000$) zwischen den Skalen Neurotizismus und psychische Belastbarkeit (Langform-Skala) als auch zwischen der Neurotizismus-Skala und der psychischen Belastbarkeit (Kurzskala) mit einem Korrelationskoeffizienten von $r = -.58$ ($p = .000$) nachgewiesen werden. Für diese Art der Überprüfung der Konstruktvalidität wurde einer Teilstichprobe von $n = 155$ neben den beiden Skalen (Kurz- und Langform) der psychischen Belastbarkeit auch die Neurotizismus-Skala des NEO-FFI vorgelegt.

Tabelle 39

Korrelationsanalyse zur divergenten Validitätsüberprüfung (n = 155).

		Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Neurotizismus-Skala des NEO-FFI (nach Costa & McCrae, 1992)	Korrelation nach Pearson	-,580	-,624
	Signifikanz (2-seitig)	,000	,000

Darüber hinaus kann bei der Berechnung der Zusammenhänge der Skalen Neurotizismus aus dem NEO-FFI und den einzelnen Skalen des BIP zur emotionalen Stabilität und zur Belastbarkeit ebenfalls höchst signifikante ($p = .000$) Zusammenhänge im mittleren Bereich mit $r = -.61$ bzw. $r = -.60$ festgestellt werden.

5.2.2.2 Psychische Belastbarkeit und Selbstwirksamkeit

Die Überprüfung der konvergenten Validität wurde mit Hilfe der Skala zur allgemeinen Selbstwirksamkeitserwartung (SWE) von Schwarzer und Jerusalem (1999) durchgeführt. Diese Skala besteht aus 10 Items mit vierstufigen Antwortkategorien. Danach ergeben sich Itemscores zwischen 10 und 40 (siehe Kapitel 4.3.3). In Tabelle 40 sind die itemscores der Skala zur allgemeinen Selbstwirksamkeit zusammengefasst.

Tabelle 40

Deskriptive Darstellung der Itemscores der Skala zur allgemeinen Selbstwirksamkeitserwartung (n = 225).

	Selbstwirksamkeit
<i>Min.</i>	20
<i>Max.</i>	40
<i>M</i>	32
<i>SD</i>	3,1

Der Überprüfung der konvergenten Validität der Skalen zur psychischen Belastbarkeit und der Skala zur allgemeinen Selbstwirksamkeitserwartung liegt die Überlegung zugrunde, dass Personen, die sich als psychisch belastbar einschätzen, auch eine höhere subjektive Selbstwirksamkeit aufweisen und vice versa. Demzufolge sollten die Skalen im mittelhohen Bereich miteinander korrelieren. Der Tabelle 41 kann man entnehmen, dass die Werte der Skala zur Erfassung der allgemeinen Selbstwirksamkeitserwartung und die psychische Belastbarkeitskala (Langform) mit $r = .55$ im mittleren Maße und höchst signifikant ($p = .000$) miteinander korrelieren. Dagegen konnte bei dem Vergleich zwischen der allgemeinen Selbstwirksamkeitserwartung und der hier entwickelten Kurzskala zur Erfassung der psychischen Belastbarkeit lediglich ein Korrelationskoeffizient von $r = .34$ ($p = .000$) berechnet werden. Im separaten Vergleich der beiden BIP-Skalen Belastbarkeit und Emotionale Stabilität mit der allgemeinen Selbstwirksamkeitserwartung zeigten sich ebenfalls jeweils höchst signifikante ($p = .000$) Zusammenhänge im mittleren Bereich (Belastbarkeit: $r = .46$; Emotionale Stabilität: $r = .48$).

Tabelle 41

Korrelationsanalyse zur konvergenten Validitätsüberprüfung ($n = 225$).

		Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Selbstwirksamkeits-Skala (nach Schwarzer & Jerusalem, 1999)	Korrelation nach Pearson	.341	.550
	Signifikanz (2-seitig)	.000	.000

Zusammenfassend verweisen die Ergebnisse zur Verfahrensentwicklung der psychischen Belastbarkeitskomponente auf eine hinreichende Konstruktvalidität der entwickelten Skalen, sowie auf die Brauchbarkeit der unterschiedlichen Skalenformen (Lang- und Kurzform-Skalen), worauf in der Diskussion (Kapitel 6) eingegangen wird.

5.3 Untersuchungen zur psycho-physischen Belastbarkeit

Die Analyse des Zusammenhangs ausgewählter psychischer und physischer Komponenten bildet ein wesentliches Anliegen dieser Arbeit. Zunächst werden die Rohwerte der verschiedenen Skalen zur Erfassung der psychischen Belastbarkeitskomponente sowie die Rohwerte der Skalen Neurotizismus und der allgemeinen Selbstwirksamkeitserwartung einer z-Transformation unterzogen, um die durch die unterschiedliche Itemanzahl pro Skala und deren unterschiedliche Abstufungen entstandenen Ungleichgewichtungen zu kompensieren bzw. sie so vergleichbar zu machen und die unterschiedlichen Antwortwahrscheinlichkeiten auszugleichen (siehe dazu z. B. Nitsch, 1970, S. 168 f.) Nach der z-Standardisierung werden die Ergebnisse der Erfassung der psychischen Komponente (Fragebogenverfahren) mit denen der physischen Komponente (Ausdauerstest) auf Zusammenhänge hin untersucht. Hierbei ergeben sich aufgrund der verschiedenen Studien und den dabei zusätzlich eingesetzten Verfahren (weitere Ausdauerstests mit der Erhebung spiroergometrischer Messgrößen, zusätzliche Fragebogenverfahren für die Überprüfung der Konstruktvalidität) teilweise unterschiedliche Stichprobenumfänge bei der Ergebnisdarstellung.

Grundlage jeder Untersuchung ist der 3000-m-FET sowie die Lang- und Kurzform-Skalen zur Erfassung der psychischen Belastbarkeit. Im Rahmen der Ergebnisdarstellung werden alle erhobenen Parameter berücksichtigt und ausgewertet. Es kann als methodische Restriktion angesehen werden, dass der 3000-m-FET nicht bei allen Untersuchungen mit den gleichen Fahrradergometern bzw. identischen Messinstrumenten durchgeführt werden konnte. Im Lauf der Entwicklung und Validierung des 3000-m-FET (siehe Kapitel 5.1) wurde die Entscheidung getroffen, den Test mit den Ergo-Trainern der Firma Tacx und der dazugehörigen Software weiterzuführen und von dem Einsatz anderer Ergometer abzusehen. Von einer Differenzierung der mit den unterschiedlichen Messgerätschaften erhobenen 3000-m-FET-Daten wird bei der folgenden Darstellung der Ergebnisse verzichtet, da sich nur marginale Unterschiede hinsichtlich der Zusammenhangsanalysen (psycho-physischer Belastbarkeit) nachweisen lassen. Bei der differenzierten Betrachtung psycho-physischer Zusammenhänge auf Grundlage von verschiedenen eingesetzten Messinstrumenten zur Erfassung der 3000-m-Zeit können keine bedeutenden Unterschiede hinsichtlich der Korrelationskoeffizienten (siehe Anhang A5 und A6) festgestellt werden. Im Anschluss an die Analyse der Korrelationen zwischen den psycho-physischen Merkmalen soll mit Hilfe von linearen Regressionsanalysen der Frage nachgegangen werden, ob ein Kausalmodell zwischen den Merkmalen angenommen werden

kann. Im ersten Zugang erfolgt die Berechnung der Regressionsanalysen nach dem Schrittweise-Verfahren, bei dem alle Prädiktorvariablen Berücksichtigung finden. Anschließend wird die Berechnung mit dem Einschlussverfahren durchgeführt, um den Einfluss nicht bedeutender Regressoren auszuschließen. Mit Hilfe der Regressionsanalyse kann neben der Aussage über den Zusammenhang auch die Richtung des Zusammenhangs, also inwieweit die unabhängigen Variablen (physische Merkmale) die abhängige Variablen (psychische Belastbarkeit) beeinflussen, überprüft werden (siehe Backhaus, Erichson, Plinke & Weiber, 2008, S. 46 ff.).

Bei dieser Analyse der erhobenen Daten wird davon ausgegangen, dass die physischen Variablen (Indikatoren für die physische Leistungsfähigkeit) die psychische Belastbarkeit, hier speziell die Merkmale Emotionale Stabilität und Belastbarkeit, beeinflussen. Aus diesem Grund wird ein multiples Regressionsproblem betrachtet, das sich aus mehreren unabhängigen Variablen, insbesondere der benötigten Zeit des 3000-m-FET und der aus der Spiroergometrie gewonnenen VO_{2max} , zusammensetzt. Des Weiteren werden auch Faktoren wie Alter, Geschlecht und der BMI neben den Indikatoren für die Ausdauerleistungsfähigkeit bzw. für das Fitnessniveau in die Analyse einbezogen, um auch den Einfluss dieser Prädiktoren auf die psychische Belastbarkeit zu analysieren. Vorrangig geht es darum, die Bestimmung der Stärke und die Richtung des jeweiligen Einflusses der unabhängigen Variable auf die abhängige Variable zu untersuchen. Dementsprechend wird hierbei auch von einer Ursachenanalyse gesprochen – mit bedacht formuliert: kann so zu einer Ursachenanalyse beigetragen werden (Backhaus et al. 2008, S. 49).

Darüber hinaus wird ein Extremgruppenvergleich vorgenommen. D. h. die besonders fitten Personen werden mit den weniger oder gar unfitten Personen der Gesamtstichprobe hinsichtlich der psychischen Belastbarkeit verglichen. Auf der Grundlage der Überprüfung, ob sich die Personen, die den 3000-m-FET in einer sehr schnellen Zeit bewältigen konnten, von der Gruppe der weniger ausdauertrainierten Personen grundsätzlich in der Tendenz der psychischen Belastbarkeit unterscheiden erfolgt die Berechnung möglicher Unterschiede zwischen den beiden Leistungsgruppen mit dem t-Test für unabhängige Stichproben. Ferner soll auch eine mittlere Leistungsgruppe bei der Berechnung der Unterschiede mit Hilfe einer einfaktoriellen Varianzanalyse (ANOVA) Berücksichtigung finden. Mit varianzanalytischen Verfahren können neben den Unterschieden auch die Wirkungen einer (oder mehrerer) unabhängiger Variablen auf eine (oder mehrere) abhängige Variable untersucht werden (Backhaus et al,

2008, S. 118). Die Leistungsgruppen (LG) werden hierbei neben der Zeit auch anhand der VO_{2max} als weiteren Indikator für die physische Leistungsfähigkeit bzw. Belastbarkeit in drei Gruppen eingeteilt. Des Weiteren wird im Rahmen der varianzanalytischen Berechnungen eine Differenzierung der physischen Belastbarkeit (Ausdauerleistungsfähigkeit) für beide Geschlechter separat vorgenommen und das Geschlecht sowie das Alter (eingeteilt in drei Altersgruppen) mit in die Berechnungen einbezogen, um eventuelle Interaktionseffekte im Hinblick auf die abhängige Variable (psychische Belastbarkeit) aufzuklären. Hierzu werden im Anschluss an die einfaktorielle auch zweifaktorielle Varianzanalysen berechnet. Für die Einteilung der Extremgruppen kommen unterschiedliche Verfahren zum Einsatz. So wird zum einen für die zweifaktorielle Varianzanalyse neben der „klassischen“ Differenzierung nach dem Prinzip „Mittelwert plus/minus der Standardabweichung“ auch die gesamte Stichprobe auf Grundlage der physischen Leistungsindikatoren (Zeit und VO_{2max}) gedrittelt, sodass eine in etwa gleich große Zellenverteilung erreicht werden kann. Dieses „Drittungsverfahren“ erscheint insbesondere vor dem Hintergrund sinnvoll, dass die weiblichen Versuchsteilnehmer (16 % der Gesamtstichprobe) deutlich schwächer repräsentiert sind als die Männer (84 %) und durch die Hinzunahme des Faktors Geschlecht die Zellenbesetzungen der weiblichen LG teilweise relativ klein ausfällt. Zusätzlich erscheint ein ergänzendes Verfahren bei der Einteilung der LG angebracht, da somit weitere interessante Effekte aufdeckbar sind, die für die Interpretation der Befundlage relevant sein könnten.

Für die Überprüfung der psycho-physischen Zusammenhänge und die Analysen hinsichtlich der Wechselwirkung der beiden Komponenten, wurden im Rahmen von mehreren Studien eine unterschiedliche Anzahl von Pbn mit Hilfe von verschiedenen Ausdauertests und anhand von diversen Skalen zur Erfassung der psychischen Belastbarkeit untersucht. Es ergeben sich somit für die jeweiligen Erhebungsinstrumente unterschiedliche Stichprobenumfänge. Die verschiedenen Studien und die dabei angewandten Verfahren und eingesetzten Instrumente bzw. Gerätschaften für die Erhebung der Daten hinsichtlich der psycho-physischen Belastbarkeit sind in Abbildung 29 differenziert verdeutlicht. Die abgerundeten, farbig unterlegten Kästchen dokumentieren die Stichprobengröße der jeweiligen Studien. Daneben sind jeweils die eingesetzten Verfahren zur Erhebung der physischen (farbig unterlegt gekennzeichnet) und psychischen Merkmale und die im Rahmen der entsprechenden Untersuchung herangezogenen Parameter bzw. die physischen und psychischen Belastbarkeitsindikatoren abgebildet.

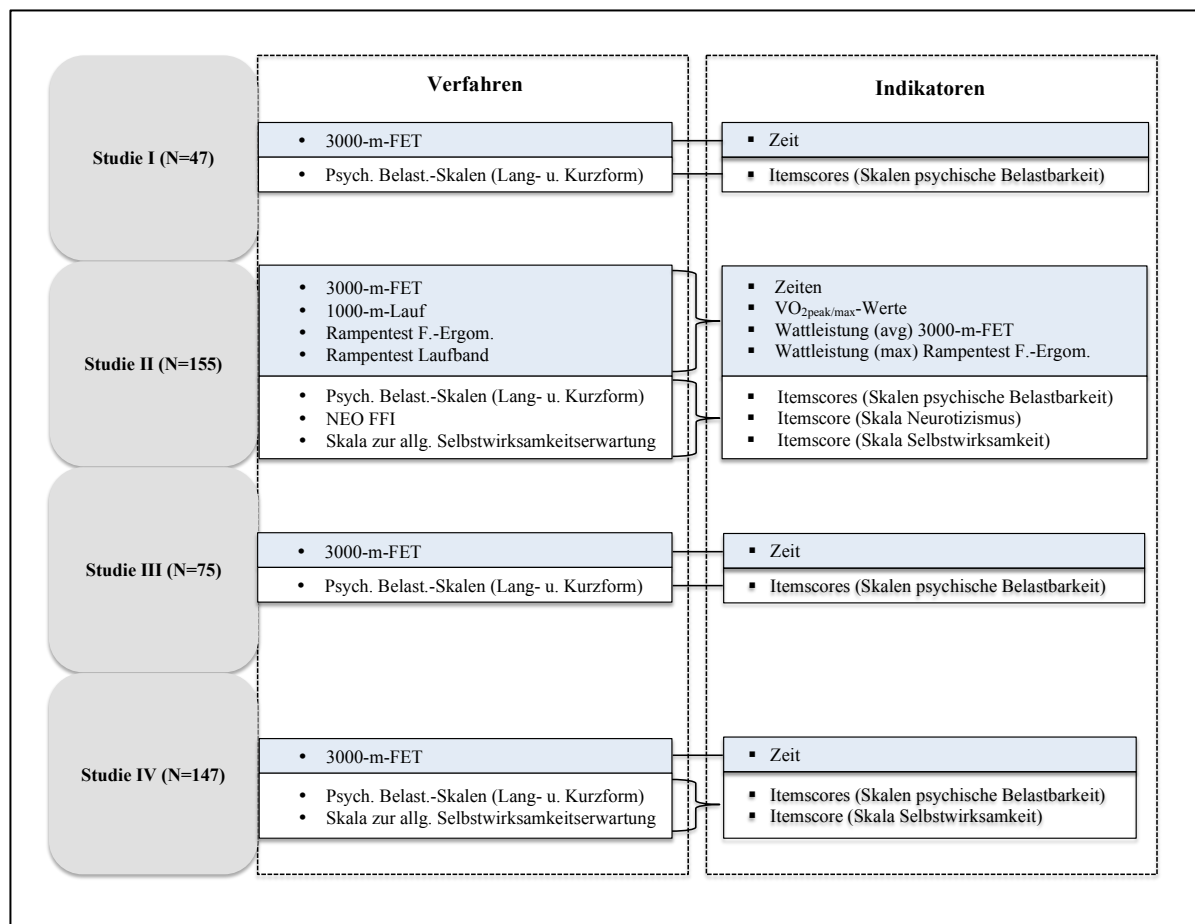


Abbildung 29. Darstellung der einzelnen Studien, deren Stichprobengröße, dabei eingesetzte Verfahren und korrespondierende Indikatoren.

Bei der Studie II beantworteten nicht alle Versuchsteilnehmer die Skala zur allgemeinen Selbstwirksamkeitserwartung, wodurch sich für dieses Instrument eine kleinere Stichprobe als die Summe der Teilnehmer aus den Studien II und IV ergibt. Die aufwendigste Studie bzw. die größte Anzahl an Parametern wurde mit der Studie II (Laborstudie) realisiert. Die Laborstudie ist in erster Linie zur Überprüfung der Gütekriterien des 3000-m-FET durchgeführt worden. Zugleich können durch die zur Validierung eingesetzten Ausdauer- und die Erhebung der spiroergometrischen Messgrößen weitere Parameter bzw. Indikatoren der physischen Belastbarkeit gewonnen werden, denen sowohl bei den einfachen Korrelationsanalysen als auch im Zuge der multivariaten Analyseverfahren und im Rahmen der varianzanalytischen Berechnungen bei den Ergebnisdarstellungen besondere Relevanz zukommt. Die Ergebnisdarstellung erfolgt in drei Schritten. Zunächst werden die Befunde der Korrelations- und dann der Regressionsanalysen berichtet. Anschließend erfolgt die Darstellung der Ergebnisse aus den Extremgruppenvergleichen und den Varianzanalysen.

5.3.1 Korrelationsanalytische Befunde

Im Folgenden sind die Ergebnisse der Zusammenhangsanalysen dargestellt, mit einem Fokus auf den Analysen zwischen der physischen Belastbarkeit in Sinne der Ausdauerleistungsfähigkeit und der psychischen Belastbarkeit, die durch die Itemscores der jeweiligen Persönlichkeitsskalen operationalisiert ist. Im ersten Schritt werden die Indikatoren der physischen Belastbarkeit mit der psychischen Belastbarkeit und anschließend mit dem Persönlichkeitsmerkmal Neurotizismus verglichen.

5.3.1.1 Ausdauerleistung und psychische Belastbarkeit

Die Ausdauerleistung stellt, wie bereits im theoretischen Rahmenkonzept erläutert, den Indikator zur Erfassung der physischen Belastbarkeitskomponente dar. Die durch das Testverfahren zur Erfassung der Ausdauerleistungsfähigkeit gewonnenen Indikatoren, insbesondere die gefahrene Zeit des 3000-m-FET, sowie die via Spiroergometrie gemessenen maximalen Sauerstoffaufnahme-Werte (VO_{2max} bzw. VO_{2peak}) und durchschnittlichen Wattleistungen, sollen im folgenden Schritt mit den psychischen Parametern, speziell mit den Itemscores der Persönlichkeitsskalen, verglichen werden. Dabei liegt der Fokus auf dem Vergleich zwischen den Indikatoren der physischen Belastbarkeit und den Itemscores der Lang- und Kurzform-Skalen der psychischen Belastbarkeit. Darüber hinaus soll auch ein Vergleich der physischen Parameter mit den zusätzlich erhobenen Persönlichkeitsskalen Neurotizismus und der allgemeinen Selbstwirksamkeitserwartung hinsichtlich der Überprüfung auf eventuelle Zusammenhänge durchgeführt werden, um so gegebenenfalls weitere Hinweise auf psycho-physische Interdependenzen zu erhalten.

Nach Abschluss der Validierungsstudien des 3000-m-FET, bei denen insgesamt 145 Personen vier Ausdauertests mittels einer Spiroergometrie absolvierten, kam der 3000-m-FET ohne eine Erhebung spiroergometrischer Messgrößen zum größten Teil als einziges Testverfahren hinsichtlich der physischen Belastbarkeit zum Einsatz. Aus diesem Grund stellt die 3000-m-Zeit den Hauptanteil der physischen Daten dar.

Vergleichbar verhält es sich bei der Erhebung der psychischen Belastbarkeitsparameter. Hierbei wurden bei allen Untersuchungen stets beide Formen der psychischen Belastbarkeitsskala eingesetzt, sodass sich eine Stichprobe von insgesamt $N = 424$ ergibt, die sowohl den

3000-m-FET zur Erfassung der physischen Belastbarkeit als auch die Lang- und Kurzformskalen zur Erfassung der psychischen Belastbarkeit ausfüllten. Das Durchschnittsalter der Gesamtstichprobe beläuft sich auf 29,1 Jahre ($SD = 9,8$). Der jüngste Teilnehmer war zum Zeitpunkt der Messung 16 und der älteste Teilnehmer 64 Jahre alt. Nach Einteilung der Teilnehmer in Altersgruppen nach Lebensdekaden stellt die größte Gruppe der Versuchsteilnehmer mit 61 Prozent die Gruppe der 20- bis 29-jährigen. Die zweitstärkste Altersgruppe sind die 30- bis 39-jährigen mit insgesamt 71 Pbn (16,8 %). Die Gruppe der Teilnehmer zwischen 40 und 49 Jahren machen mit 45 Pbn 10,6 Prozent der Stichprobe aus. Jünger als 20 Jahre waren 27 Personen (6,4 %) und die Gruppe der 50- bis 59-jährigen umfasst 4 Prozent der Gesamtstichprobe. Fünf Versuchsteilnehmer waren älter als 60 Jahre (1,2 %; vgl. Abbildung 30).

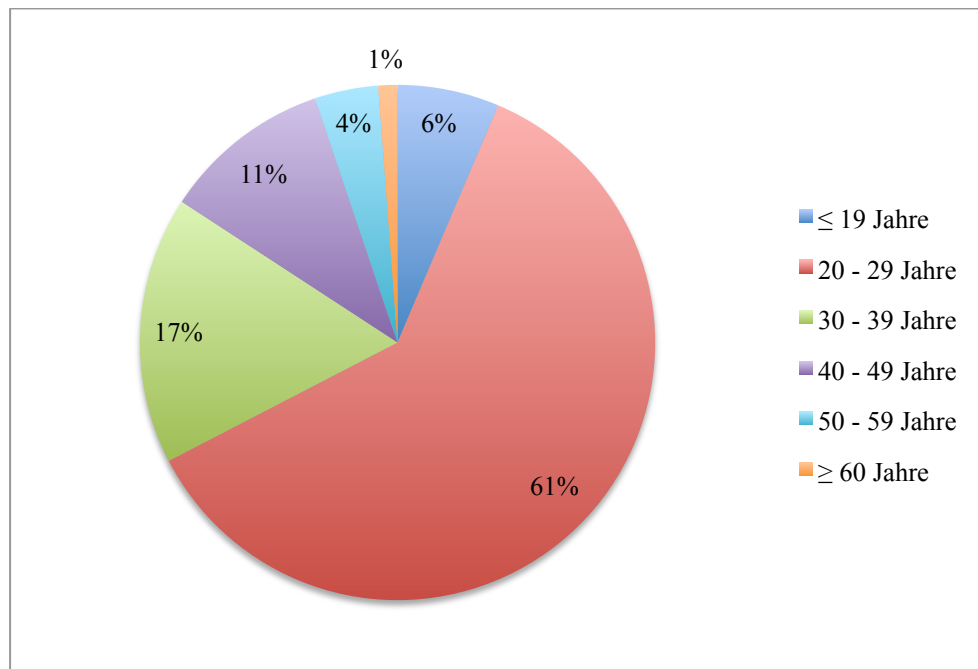


Abbildung 30. Altersverteilung der Gesamtstichprobe in Gruppen nach Lebensdekaden (N = 424).

Von den 424 auf die psycho-physische Belastbarkeit hin getesteten Personen sind 356 (84 %) männlich und 68 (16 %) weiblich. Die männlichen Probanden waren zum Zeitpunkt der Datenerhebung durchschnittlich 29 Jahre ($SD = 9,3$ Jahre) und die weiblichen Testpersonen im Mittel ca. 2 Jahre älter ($M = 30,7$; $SD = 12,1$ Jahre). Die Durchschnittszeit, die für die Bewältigung der 3000-m-Strecke auf dem Fahrradergometer benötigt wurde, lag bei rund 4:23 Min. (263 Sek.) mit einer SD von 30 Sek. Die 68 Teilnehmerinnen benötigten im Mittel 5:01 Min. (301 Sek.; $SD = 44$ Sek.) und die Männer durchschnittlich 04:16 Min. (256 Sek.; $SD = 20$ Sek.) für die 3000-m-Strecke. Die prozentuale Abweichung der durchschnittlichen

Zeit der Frauen zu den männlichen Teilnehmern beträgt rund 18 Prozent. Die schnellste Zeit betrug 3:20 Min., die langsamste Zeit 8:13 Min. Eine tabellarische Darstellung der Häufigkeitsverteilung der Zeiten des 3000-m-FET ist dem Anhang F1 zu entnehmen.

Tabelle 42

Zusammenhang zwischen dem physischen Indikator Zeit (3000-m-FET) und der psychischen Belastbarkeit (N = 424).

		Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Zeit 3000-m-FET	Korrelation nach Pearson	-.199	-.269
	Signifikanz (2-seitig)	.000	.000

Zwischen dem Parameter Zeit bei dem 3000-m-FET und der psychischen Belastbarkeit (Langform- und Kurzform-Skala) besteht ein niedriger negativer Zusammenhang, der in beiden Fällen höchst signifikant ($p = .000$; vgl. Tabelle 42) ist. Mit dem 3000-m-FET konnten insgesamt 424 Personen getestet werden, die vor dem Ausdauerstest auch die beiden Formen der Skala zur psychischen Belastbarkeit ausfüllten. Versuchsteilnehmer, die die 3000 Meter bei dem FET schnell zurücklegten, d. h. in einer kurzen Zeit bewältigten, erreichen in der Tendenz höhere Itemscores bei den Skalen der psychischen Belastbarkeit. D. h., Personen mit einer erhöhten Ausdauerleistung weisen eine ausgeprägtere psychische Belastbarkeit auf.

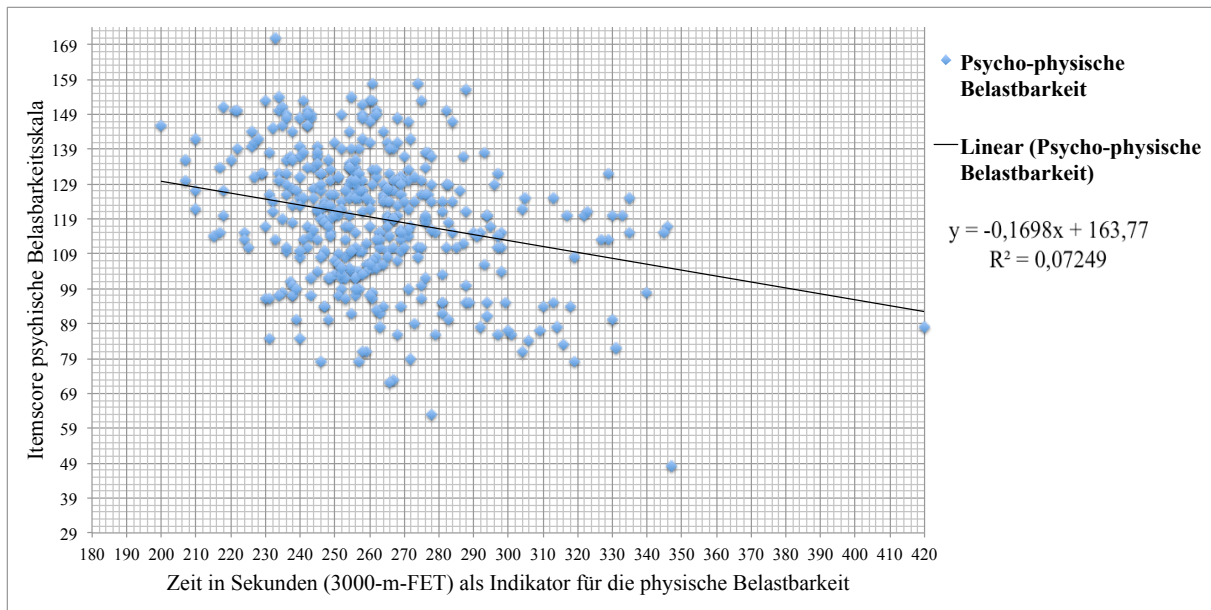


Abbildung 31. Zusammenhang zwischen objektiv erfasster (Zeit 3000-m-FET) physischer und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit (N = 424).

In Abbildung 31 ist der Zusammenhang ($r = -.27$; $p = .000$) zwischen der physischen und psychischen Belastbarkeit veranschaulicht. Die Regressionsgerade zeigt, dass Personen, die sich mittels der beiden Skalen aus dem BIP (Emotionale Stabilität und Belastbarkeit) als psychisch belastbar einschätzen, tendenziell eine bessere Ausdauerleistungsfähigkeit, operationalisiert über die Zeit beim 3000-m-FET, im Sinne eines höheren physischen Belastbarkeitsvermögens aufweisen. Je niedriger die Zeit (in Sek.), die für die 3000-m-Strecke benötigt wurde, desto höher der Itemscore der Skala zur psychischen Belastbarkeit. Mit Hilfe der Regressionsgeraden, die als Funktion im Diagramm dargestellt ist, wird die Stärke des Zusammenhangs der hier im Fokus stehenden Parameter quantifiziert. Das Bestimmtheitsmaß (Determinationskoeffizient: $R^2 = .07$) gibt den Anteil der Varianz der Kriteriumsvariable psychische Belastbarkeit an. Die im Rahmen der Laborstudie zur Überprüfung der Gütekriterien des 3000-m-FET ermittelten Daten des 1000-m-Laufs sowie der beiden Rampentests sollen ebenfalls zur Verdeutlichung psycho-physischer Zusammenhänge hinzugezogen werden. Ein deutlich höherer Zusammenhang zeigte sich zwischen der erzielten Zeit beim 1000-m-Lauf und den psychischen Parametern.

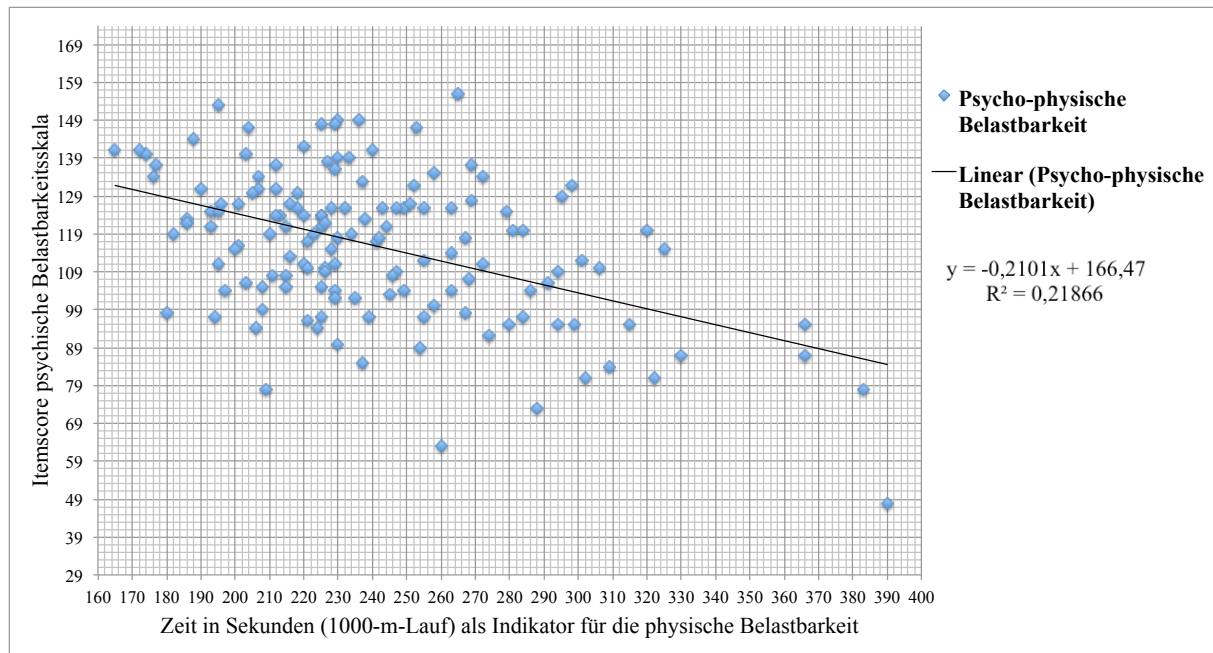


Abbildung 32. Zusammenhang zwischen objektiv erfasster physischer (Zeit 1000-m-Lauf) und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit (N = 145).

Betrachtet man die Zeit, die für die Bewältigung des 1000-m-Laufs gemessen wurde, zeigt sich bei der Analyse in Bezug auf die psychische Belastbarkeit (Langform; siehe Tabelle 43) mit $r = -.47$ ($p = .000$) ein mittlerer und höchst signifikanter Zusammenhang. Im Vergleich der 1000-m-Zeit mit der Kurzform-Skala der psychischen Belastbarkeit kann ebenfalls ein höchst signifikanter mittlerer Zusammenhang festgestellt werden ($r = -.45$; $p = .000$). Die Höhe der Korrelationskoeffizienten lassen darauf schließen, dass Personen, die den 1000-m-Lauf in einer kurzen Zeit (schnell bzw. zügig) zurückgelegt haben, mit deutlicher Tendenz höhere Werte bei der psychischen Belastbarkeit aufweisen. Dieser Zusammenhang war feststellbar mit beiden (Kurz- und Langform) Skalen. Die Kürze der benötigten Zeit für die Bewältigung der Strecke gilt als Indikator für die physische Belastbarkeit bzw. für die Ausdauerleistungsfähigkeit. Durch die Regressionsgrade ($R^2 = .22$) in Abbildung 32 wird der Zusammenhang des physischen Belastbarkeitsindikators Zeit 1000-m-Lauf und der psychischen Belastbarkeit (Langform-Skala) veranschaulicht.

Tabelle 43

Korrelationsanalyse des Parameters Zeit und der psychischen Belastbarkeit (N = 145).

	Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Zeit: 1000-m-Lauf	-.451***	-.468***
Zeit: Rampentest F.-Ergom.	.362***	.289***
Zeit: Rampentest Laufband	.372***	.348***

*** $p \leq .001$

Die Daten aus den beiden Rampentests weisen im Hinblick auf die psychischen Belastbarkeitsscores höchst signifikante positive Zusammenhänge im geringen Bereich auf, wobei der Korrelationskoeffizient von $r = .37$ ($p = .000$), also der Zusammenhang der Zeiten bei der Rampenbelastung auf dem Laufband und den Itemscores der Kurzskala, in Richtung eines mittleren Zusammenhangs gehen. Je länger die Pbn den Rampenbelastungen standhalten, desto höher sind tendenziell die Itemscores der subjektiv eingeschätzten psychischen Belastbarkeit sowohl im Hinblick auf die Lang- als auch auf die Kurzform der vorgelegten Skalen.

Tabelle 44

Korrelationsanalyse der Parameter durchschnittliche Wattleistung (3000-m-FET) bzw. maximale Wattleistung (Rampentest F.-Ergom.) und der psychischen Belastbarkeit (N = 145).

	Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Durchschn. Wattleistung 3000-m-FET	.227*	.335***
Maximale Wattleistung Rampentest F.-Ergom.	.352***	.236**

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Bei den Untersuchungen mit dem Gerät der Firma Tacx konnten bei dem 3000-m-FET bei 145 Pbn durch die vorhandene Software parallel die durchschnittlichen Leistungen der Probanden in Watt, sowie bei dem Rampentest auf dem Fahrradergometer die maximal erreichte Wattleistung erfasst werden, die auch als Merkmale der physischen Belastbarkeit mit in die

Analysen einbezogen werden sollen. Bei der Korrelationsanalyse zwischen der durchschnittlichen Leistung, gemessen in Watt, zeigt sich ein höchst signifikanter Zusammenhang ($p = .000$) in Bezug auf die Langform-Skala der psychischen Belastbarkeit i. H. v. $r = .34$. Bei der maximalen Wattleistung, die zum Zeitpunkt des Abbruchs des Rampentests festgehalten wurde, und der Langform-Skala konnte ein hoch signifikanter Zusammenhang mit einem Koeffizienten von $r = .24$ ($p = .004$) festgestellt werden. Hinsichtlich der Überprüfung des Zusammenhangs zwischen der maximal erreichten Wattleistung und der Kurzform-Skala der psychischen Belastbarkeit ist ein höchst signifikanter Korrelationskoeffizient von $r = .35$ ($p = .000$) festzustellen, der ebenfalls in Richtung eines mittleren Zusammenhangs tendiert (siehe Tabelle 44).

Tabelle 45

Korrelationsanalyse der Parameter $VO_{2peak/max}$ und der psychischen Belastbarkeit ($N = 145$).

	Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
VO_{2peak} : 3000-m-FET	.343***	.294***
VO_{2peak} : Rampentest F.-Ergom.	.390***	.289***
VO_{2max} : 1000-m-Lauf	.409***	.380***
VO_{2max} : Rampentest Laufband	.401***	.371***

*** $p \leq .001$

Zwischen der maximalen Sauerstoffaufnahme der vier Ausdauerests und der psychischen Belastbarkeit (Tabelle 45) können durchgehend höchst signifikante Zusammenhänge festgestellt werden. Die Korrelationskoeffizienten liegen, bis auf die Zusammenhänge der beiden Laufests und der Kurzskaala, mit $r < .39$ im niedrigen mit Tendenz zum mittleren Bereich. Während der VO_{2peak} -Wert, erfasst bei dem Rampentest auf dem Fahrradergometer, und die Werte der psychischen Belastbarkeit (Kurzskaala) einen mittleren Zusammenhang ($r = .39$; $p = .000$) aufweisen, zeigen sich zwischen der VO_{2peak} des 3000-m-FET mit der Kurzskaala ein geringer Zusammenhang von $r = .24$ ($p = .000$). Die Korrelationskoeffizienten der Zusammenhangsanalyse zwischen den VO_{2max} -Werten der beiden Laufests und der psychischen Belastbarkeit (Langform) weisen ebenfalls beide in Richtung eines höchst signifikanten

($p = .000$) und mittleren Zusammenhangs $r = .38$ (1000-m-Lauf) und $r = .37$ ($p = .000$). Die Zusammenhänge der maximalen VO_2 -Werte der Lauftests und der Kurzsкала liegen im mittleren Bereich und sind höchst signifikant ($r = .41$, $p = .000$; 1000-m-Lauf und $r = .40$; $p = .000$; Rampentest).

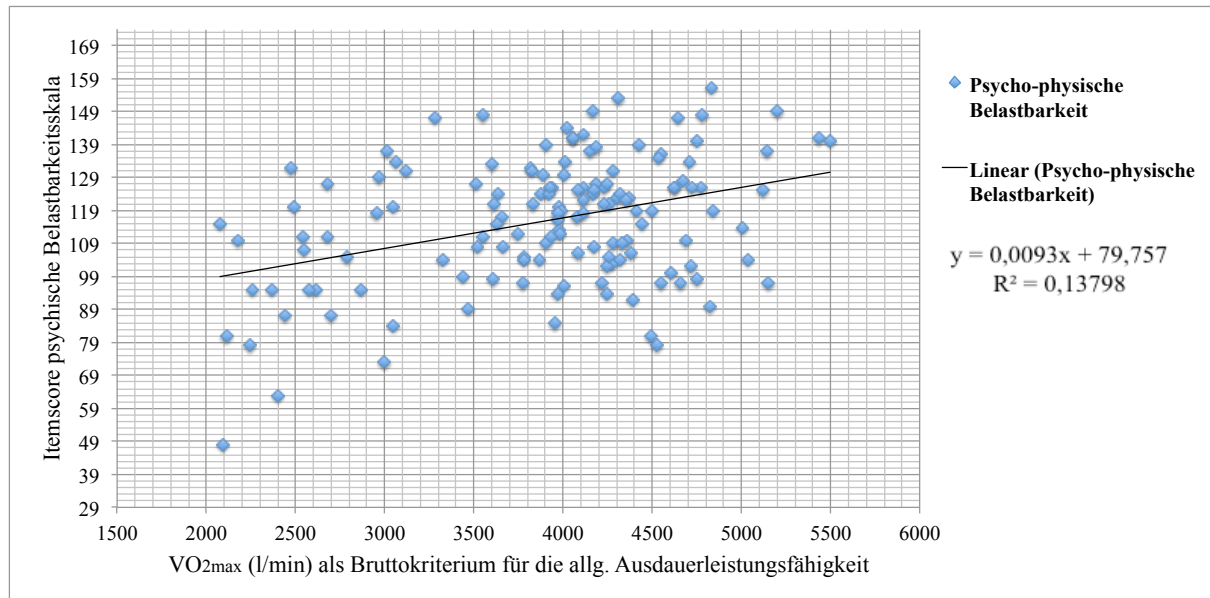


Abbildung 33. Zusammenhang zwischen objektiv erfasster physischer (VO_{2max}) und subjektiv eingeschätzter (Langform-Skala) psychischer Belastbarkeit ($N = 145$).

Das Streudiagramm und die darin enthaltene Regressionsgerade in Abbildung 33 verdeutlichen noch einmal den linearen Trend zwischen der maximalen Sauerstoffaufnahme (gemessen bei dem Rampentest auf dem Laufband) und dem Itemscore der psychischen Belastbarkeitsskala (Langform). Die Regressionsgerade verläuft durch den Schwerpunkt der abgebildeten Punkte des Scatter Plots (x-y Punktwolke) und stellt dabei die Gerade dar, bei der die Summe der quadratischen Abweichungen minimal ist. Hierbei werden für jeden Punkt im Scatter Plot die vertikale Entfernung von der Geraden gemessen, dieser Abstand quadriert und anschließend alle quadratischen Abweichungen (Residuen) aufsummiert. Zusammenfassend zeigt sich durch die Lage der Geraden im Koordinatensystem ein Trend, wobei die Steigung der VO_{2max} mit einer Erhöhung des Itemscores als Indikator für die psychische Belastbarkeit zu erkennen ist. Hierzu soll abschließend angemerkt werden, dass sich hinsichtlich der Berechnungen mit den z-transformierten Werten und den Rohwerten keine Unterschiede bei den Ergebnissen zeigen. Aus Gründen der Einfachheit wird im weiteren Verlauf der Ergebnisdarstellungen auf die Erwähnung von Spezifikationen der Werte verzichtet und die Darstellungen wie bereits ausgeführt, mit Hilfe der Skalenrohwerte veranschaulicht.

5.3.1.2 Ausdauerleistung und Neurotizismus

Zusätzlich sollen auch noch die Zusammenhänge der physischen Werte mit den Skalen Neurotizismus und Selbstwirksamkeit verglichen werden. Hierbei können in Bezug auf die Skala Neurotizismus vergleichbare Zusammenhänge in hohen Signifikanzbereichen wie auch bei dem für diese Arbeit wesentlichen Konstrukt psychische Belastbarkeit vermutet werden. Im Hinblick auf die Selbstwirksamkeit, gemessen mit der Skala zur allgemeinen Selbstwirksamkeitserwartung nach Schwarzer und Jerusalem (1999), können bei der Korrelationsanalyse keinerlei Zusammenhänge zu auch nur einem der physischen Parameter nachgewiesen werden (siehe Anhang A4). Aus diesem Grund wird bei der Darstellung der Ergebnisse der Zusammenhangsanalysen zwischen den Itemscores der Skala zur allgemeinen Selbstwirksamkeitserwartung und den physischen Belastbarkeitsindikatoren im weiteren Verlauf abgesehen.

Die Skala zur Erfassung des Persönlichkeitsmerkmals Neurotizismus, die vorwiegend für die Überprüfung der Konstruktvalidität (divergenter Aspekt) der psychischen Belastbarkeitskalen zum Einsatz kam, soll auch an dieser Stelle mit den physischen Belastbarkeitsindikatoren auf Zusammenhänge hin untersucht werden. Entsprechend der Befunde aus den Analysen mit den Skalen zur Erfassung der psychischen Belastbarkeit werden Korrelationen in ähnlicher Höhe jedoch mit umgedrehten Vorzeichen angenommen. Diese Vermutung basiert auf der Überlegung, dass die Neurotizismus-Skala (NEO-FFI) ein gegenteiliges Merkmal zur psychischen Belastbarkeit erfasst, nämlich die emotionale Labilität. Zeigt sich zwischen der erzielten Zeit bei dem 3000-m-FET und dem Itemscore der Langformskala „psychische Belastbarkeit“ ein negativer Zusammenhang, wird in Bezug auf den Itemscore der Skala Neurotizismus ein positiver Zusammenhang erwartet. D. h., je kürzer die Zeit, die zur Bewältigung der 3000-m-Strecke benötigt wurde, desto niedriger der selbsteingeschätzte Neurotizismus-Wert bzw. desto niedriger sind die Werte hinsichtlich der emotionalen Labilität, was darauf hindeutet, dass fittere Personen emotional stabiler sind.

Tabelle 46

Korrelationsanalyse der physischen Parameter und der emotionalen Labilität, erfasst mit der Skala Neurotizismus des NEO-FFI (N = 145)⁵.

	Neurotizismus
Zeit: 3000-m-FET	.290***
Zeit: 1000-m-Lauf	.421***
Zeit: Rampentest F.-Ergom.	-.333***
Zeit: Rampentest Laufband	-.316***
VO _{2peak} : 3000-m-FET	-.297***
VO _{2max} : 1000-m-Lauf	-.347***
VO _{2peak} : Rampentest F.-Ergom.	-.335***
VO _{2max} : Rampentest Laufband	-.314***
Durchschn. Wattleistung: 3000-m-FET	-.309**
Maximale Wattleistung: Rampentest F.-Ergom.	-.266***

** $p \leq .01$; *** $p \leq .001$

Die Zusammenhangsanalysen (vgl. Tabelle 46) zeigen - der Annahme entsprechend - zwischen den im Rahmen der Laboruntersuchungen erhobenen physischen Parametern und den Itemscores der Skala Neurotizismus zumeist höchst signifikante Zusammenhänge ($p \leq .001$) im niedrigen Bereich ($r \leq .39$), jedoch - wie erläutert - im Gegensatz zu den Skalen der psychischen Belastbarkeit mit negativem Vorzeichen. Eine Ausnahme zeigt sich auch hier wieder bei der 1000-m-Zeit und dem psychischen Faktor emotionale Labilität. Hier besteht ein höchst signifikanter mittlerer positiver Zusammenhang ($r = .42$; $p = .000$), dies deutet daraufhin, dass Personen mit einer guten Fitness bzw. Ausdauerleistungsfähigkeit, operationalisiert über die Zeit bzw. Testdauer zur Bewältigung der 1000-m-Strecke auf dem Laufband, tendenziell niedrigere subjektiv eingeschätzte Itemscores bei der Neurotizismus-Skala aufweisen.

⁵ N = 145; bis auf die Korrelationen der 3000-m-FET-Zeiten und der Itemscores der Neurotizismus-Skala, hier bestand die Stichprobe aus 155 Pbn.

5.3.2 Regressionsanalytische Befunde

Im Folgenden werden die Beziehungen zwischen den physischen und den psychischen Variablen analysiert, um im Nachgang zu den Korrelationsanalysen die Anteile an der Varianzaufklärung zu ermitteln. Die objektiv erfassten physischen Belastbarkeitsparameter bzw. die Indikatoren für die physische Belastbarkeit stellen dabei die Prädiktorvariablen dar, die mit der subjektiv eingeschätzten psychischen Belastbarkeit als Kriteriumsvariable in Beziehung gesetzt werden. Bei der Analyse der psycho-physischen Korrelate soll neben der Betrachtung des Zusammenhangs zweier isolierter Indikatoren auch der Einfluss bzw. der Zusammenhang mehrerer Indikatoren gleichzeitig auf die abhängige Variable (psychische Belastbarkeit) analysiert werden. Bei der multiplen Regressionsrechnung kann eine Gleichung zur Vorhersage einer Kriteriumsvariablen bei gleichzeitiger Berücksichtigung mehrerer Prädiktorvariablen bestimmt werden (Bortz, 1989 S. 547). Neben den Indikatoren für die physische Belastbarkeit werden soziodemographische Parameter wie Alter, Geschlecht und der BMI als weitere Prädiktoren in die Regressionsanalyse mit einbezogen. Die multiple Korrelationsstatistik ermöglicht es, „Beziehungen zwischen zwei oder mehreren Prädiktorvariablen und einer einzelnen Kriteriumsvariablen zu analysieren. Ein Ergebnis dieser Analyse besteht in einer Gleichung zur Vorhersage des Kriteriumswertes einer Person (multiple Regressionsgleichung) und dem multiplen Korrelationskoeffizienten R “ (Bortz, 1989, S. 556). Gegenstand der Regressionsanalysen sind auf der Prädiktoreseite somit die bei den Ausdauertests erfassten Zeiten, die im Rahmen der Laboruntersuchungen gemessenen VO_{2peak} - bzw. VO_{2max} -Werte, die mit der Tacx-Software gemessenen durchschnittlichen und maximalen Wattleistungen im Zuge der Fahrradergometertests und die aufgenommenen und z. T. berechneten soziodemographischen Parameter (Alter, Geschlecht und BMI) sowie auf der Kriterien-seite insbesondere die Daten der Lang- und Kurzform-Skalen zur Einschätzung der psychischen Belastbarkeit.

An dieser Stelle sei auf das Problem der Multikollinearität, speziell bei den Berechnungen der vorliegenden Regressionsanalysen, hingewiesen: „Das lineare Regressionsmodell basiert auf der Prämisse, dass die Regressoren nicht exakt linear abhängig sind. D. h. ein Regressor darf sich nicht als lineare Funktion der übrigen Regressoren darstellen lassen.“ (Backhaus, Erichson, Plinke & Weiber, 2003, S. 88) Regressoren sind in diesem Fall als Synonym für die unabhängige Variable bzw. Prädiktorvariablen zu sehen. Um den Grad an Multikollinearität aufzudecken und dem Problem zu begegnen, wurden vor den Regressionsberechnungen zu-

nächst in einer Korrelationsmatrix (siehe Anhang A1) die Prädiktorvariablen auf deren Zusammenhänge bzw. auf die Höhe der Korrelationskoeffizienten hin untersucht. Der Annahme entsprechend ergeben sich bei paarweiser Betrachtung zwischen den einzelnen unabhängigen Variablen zumeist Zusammenhänge im Bereich von $r \geq .80$. Angesichts der hohen Interkorrelationen der physischen Parameter Zeit, $VO_{2\text{peak}/\text{max}}$ -Werte und den Leistungen in Watt ist eine gewisse Gefahr der Multikollinearität bei Einbezug mehrerer physischer Parameter als Prädiktoren in die Regressionsberechnung gegeben. „Voneinander unabhängige Prädiktorvariablen, die jeweils hoch mit der Kriteriumsvariablen korrelieren, sind am besten zur Vorhersage einer Kriteriumsvariablen geeignet.“ (Bortz et al., 1989, S. 563) Auch wenn immer ein bestimmter Grad an Multikollinearität bei empirischen Daten besteht, sollte auf Grund der hohen Korrelationskoeffizienten zwischen den Regressoren von der Hinzunahme mehrerer unabhängiger Variablen in das Regressionsmodell abgesehen werden, da dies v. a. zu einer Redundanz in den Daten und damit zu einem Informationsverlust führt (Backhaus, et al. 2003).

Des Weiteren sei darauf hingewiesen, dass seitens der physischen Indikatoren für die Belastbarkeit die Variablen Zeit und $VO_{2\text{max}}$ bzw. $VO_{2\text{peak}}$, erhoben im Zuge der Laborstudie mit verschiedenen Testverfahren, als Einflussgröße jeweils viermal vorliegen. Vor dem Hintergrund dieser Besonderheiten werden bei den folgenden Regressionsanalysen die Berechnungen zunächst mit dem schrittweise (stepwise) Verfahren vorgenommen, bei dem alle physischen Parameter einbezogen werden. Bei dieser Methode werden die unabhängigen Variablen nacheinander und einzeln in die Regressionsgleichung aufgenommen. Bei der Analyse mit der entsprechenden Prozedur in SPSS[®] wird jeweils die Variable ausgewählt, die ein bestimmtes Gütekriterium maximiert. Als erstes führt das Programm bei dieser Methode eine einfache Regressionsanalyse mit der Variablen durch, die den höchsten Zusammenhang (positiv oder negativ) mit der Kriteriumsvariablen aufdeckt. Anschließend wird bei diesem Modell der Prädiktor mit dem größten partiellen Korrelationskoeffizienten ausgewählt, der den größten Beitrag zur Maximierung des Bestimmtheitsmaßes (R^2) leistet (Backhaus et al., 2003, S. 105)

Anschließend werden mit dem Einschlussverfahren jeweils ein Indikator für die physische Belastbarkeit und daneben die Parameter Alter, Geschlecht und BMI als Regressoren mit in die Regressionsberechnungen aufgenommen, um den tatsächlichen Einfluss auf die psychische Belastbarkeit zu bestimmen und festzustellen, in welcher Form die Prädiktoren zu kombinieren sind, um die psychische Belastbarkeit als abhängige Variable möglichst präzise zu

prognostizieren. Im Anschluss soll ein Ranking der Prädiktoren mit dem größten geschätzten Einfluss auf die Kriterienvariablen aufgezeigt werden.

Zunächst stehen wie bei den vorherigen Korrelationsberechnungen die beiden Indikatoren für die physische und psychische Belastbarkeit im Mittelpunkt der Regressionsanalyse. Wie bereits erwähnt sind der 3000-m-FET und die beiden Persönlichkeitsskalen (Lang- und Kurzform), die beiden Verfahren, womit alle 424 Untersuchungsteilnehmer untersucht wurden. Demzufolge stellt die Zeit des 3000-m-FET in dem ersten Regressionsmodell die Prädiktorvariable dar und die Itemscores der psychischen Belastbarkeitsskalen die Kriteriumsvariablen. Für die erste Regressionsberechnung (Methode: Einschluss) wird als Kriterium alleine der Itemscore der Langform-Skala in das Modell mitaufgenommen. Neben dem physischen Belastbarkeitsindikator Zeit (3000-m-FET) umfasst die erste Regressionsberechnung zusätzlich die Prädiktoren Alter, Geschlecht und BMI.

Tabelle 47

Regressionsanalyse hinsichtlich der Langform-Skala der psychischen Belastbarkeit (N = 145).

Prädiktor	Regressionskoeffizient B	Standardfehler	Standardisierter Regressionskoeffizient (β)
Alter	-.865	.249	-.281*
Geschlecht	-8.254	5.147	-.169
BMI	.473	.622	.067
Zeit 3000-m-FET	-.003	.095	-.144

Anmerkung: $R = .44$; $R^2 = .19$; R^2 korrigiert = .17; * $p \leq .05$.

Die Regressionsanalyse im Hinblick auf die psychische Belastbarkeit ergibt eine Varianzaufklärung für die vier in Tabelle 47 dargestellten Prädiktoren von $R^2 = 19,3$ Prozent. In der Regression ($R_{corr} = .17$; $F = 8,2$; $p \leq .001$) wird die psychische Belastbarkeit am besten durch das Geschlecht ($B = -8.25$) vorhergesagt. Daneben erweist sich das Alter als signifikante ($p \leq .05$) Prädiktorvariable ($\beta = -.28$). Ein im Anschluss an die oben dargestellten Korrelationsanalysen unerwartetes Ergebnis zeichnet sich bei der Betrachtung des Indikators für die physische Belastbarkeit ab. Entgegen dem höchst signifikanten Zusammenhang der bivariaten Korrelationsanalyse, stellt im Rahmen der hier berechneten multivariaten Analyse die Zeit 3000-m-FET weder einen signifikanten noch einen bedeutenden Prädiktor dar ($\beta = -.14$).

Ähnlich verhält es sich bei der Substitution der Kriteriumsvariable. Bei der Analyse des Modells, bei dem die Kurzform-Skala der psychischen Belastbarkeit als Kriteriumsvariable definiert wird, trägt lediglich das Geschlecht einen bedeutenden Beitrag zur Vorhersage des Einflusses auf die psychische Belastbarkeit mit einem Regressionskoeffizienten von $B = -.78$ ($p = .026$) bei (vgl. Tabelle 48).

Tabelle 48

Regressionsanalyse hinsichtlich der Kurzform-Skala der psychischen Belastbarkeit (N = 145).

Prädiktor	Regressionskoeffizient B	Standardfehler	Standardisierter Regressionskoeffizient (β)
Alter	-.039	.017	-.188*
Geschlecht	-.783	.348	-.237*
BMI	-.007	.042	-.014
Zeit 3000-m-FET	-.011	.006	-.173

Anmerkung: $R = .44$; $R^2 = .19$; R^2 korrigiert = .17; * $p \leq .05$.

5.3.3 Extremgruppenvergleiche und varianzanalytische Befunde

Die Untersuchungsteilnehmer werden anhand der Zeiten des 3000-m-FET in drei Gruppen aufgeteilt. Zur ersten Gruppe (Leistungsgruppe I) gehören die Untersuchungsteilnehmer, die den 3000-m-FET in einer überdurchschnittlichen, besonders schnellen Zeit zurückgelegt haben ($n_1=38$). Diese Personen benötigten für die 3000-m-Strecke maximal 233 Sek. (3:53 Min.). In der zweiten Gruppe (Leistungsgruppe II; $n_2=333$) befinden sich die Versuchsteilnehmer, die den Test im Zeitrahmen von 234 Sek. ($\geq 3:54$ Min.) bis 293 Sek. ($\leq 4:50$ Min.) absolviert haben. Diese Zeitspanne entspricht dem Mittelwert der Gesamtstichprobe (263 Sek.) plus/minus der *SD* (30 Sek.). Alle Versuchsteilnehmer der Leistungsgruppe III ($n_3=54$) benötigten 294 Sek. und mehr für die Bewältigung der 3000-m-Strecke auf dem Fahrradergometer, was für eine gering ausgeprägte Ausdauerleistungsfähigkeit spricht. Angesichts dieser Einteilung entstehen drei Leistungsgruppen, wobei Gruppe zwei (LG II) mit rund 79 Prozent der Pbn die meisten Untersuchungsteilnehmer umfasst, Leistungsgruppe eins (LG I) mit den schnellsten 3000-m-Zeiten ca. 9 Prozent und die leistungsschwächste Gruppe (LG III) ca. 13 Prozent aller Versuchsteilnehmer ausmacht. Der Vergleich der Extremgruppen soll zu einer Verdeutlichung der Unterschiede hinsichtlich der psycho-physischen Belastbarkeit beitragen. Dementsprechend ist die Erwartung, dass die Untersu-

chungsteilnehmer mit einer besonders hohen Ausprägung der Ausdauerleistungsfähigkeit (LG I) bei den psychischen Belastbarkeitsskalen (Lang- und Kurzform) signifikant höhere Werte aufweisen als die anderen beiden LG. Im Folgenden sollen die Unterschiede der psychischen Belastbarkeit in Abhängigkeit der verschiedenen physischen Leistungsniveaus bzw. LG im Detail analysiert werden.

Tabelle 49

Vergleich der LG I und III (Zeit: 3000-m-FET) im Hinblick auf die psychische Belastbarkeit (n = 92), berechnet mit dem t-Test für unabhängige Stichproben.

Variable	LG I (n = 38)		LG III (n = 54)		t-test	
	M	SD	M	SD	t	p
Psychische Belastbarkeit (Langform)	129,61	17,51	106,42	18,33	6,08	.000
Psychische Belastbarkeit (Kurzform)	10,26	1,50	9,18	1,57	3,32	.001

Durch Berechnung des t-Tests bei unabhängigen Stichproben können zwischen den beiden LG eins und drei höchst signifikante Unterschiede aufgezeigt werden (siehe Tabelle 49). Die Versuchsteilnehmer der Gruppe eins, die bei dem 3000-m-FET am besten abgeschnitten haben, weisen mit einem durchschnittlichen Rohwert von 130 ($SD = 18$) wesentlich ($p = .000$) höhere Werte bei der psychischen Belastbarkeitsskala (Langform) auf als die Gruppe der weniger körperlich Leistungsfähigen ($M = 106$; $SD = 18$). Bei differenzierter Betrachtung der Ausdauerleistung anhand des Indikators der absoluten VO_{2peak} , gemessen bei dem Rampentest auf dem Fahrradergometer, zeigen sich ebenfalls höchst signifikante Unterschiede ($p = .001$) hinsichtlich der psychischen Belastbarkeit. Für die Berechnung mit dem t-Test für unabhängige Stichproben wurden zunächst drei LG in Bezug auf die VO_{2peak} -Werte definiert. Die Gruppe der am stärksten Ausdauertrainierten (LG I, $n_1=23$) mit einer VO_{2peak} im Bereich von 4291 l/Min. und höher weisen im Mittel bei der psychischen Belastbarkeitsskala (Langform) einen Rohwert von 123 ($SD = 15$) auf. Hingegen kann bei den unfitten Untersuchungsteilnehmern (LG III, $n_3=25$) mit VO_{2peak} -Werten im Bereich von 2772 l/Min. und niedriger im Hinblick auf die psychische Belastbarkeitsskala (Langform) ein durchschnittlicher Rohwert

von 103 ($SD = 21$) festgestellt werden (vgl. Tabelle 50). Für den Extremgruppenvergleich mit Hilfe des t-Tests wurde LG II mit den mittelmäßig leistungsstarken Teilnehmern ausgeschlossen.

Tabelle 50

Vergleich der LG I und III (VO_{2peak} : Rampentest F.-Ergom.; in l/Min.) im Hinblick auf die psychische Belastbarkeit ($n = 48$), berechnet mit dem t-Test für unabhängige Stichproben.

Variable	LG I (n = 23)		LG III (n = 25)		t-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
Psychische Belastbarkeit (Langform)	122,83	15,49	102,92	21,38	3,7	.001
Psychische Belastbarkeit (Kurzform)	10,48	0,99	8,68	1,67	4,47	.000

Mit der Berechnung einer einfaktoriellen Varianzanalyse (ANOVA) soll im Folgenden die LG II ebenfalls Berücksichtigung bei den Gruppenvergleichen finden. Bei Einbezug der Gruppe mit den mittleren Leistungen hinsichtlich der 3000-m-Zeiten, zeigen sich ebenso höchst signifikante Unterschiede bei der psychischen Belastbarkeit. Die Mittelwerte der LG bezüglich der psychischen Belastbarkeit können aus Tabelle 51 entnommen werden.

Tabelle 51

Vergleich der drei LG (Zeit: 3000-m-FET) im Hinblick auf die Lang und Kurzform-Skala der psychischen Belastbarkeit ($N = 424$), berechnet mit Hilfe der einfaktoriellen Varianzanalyse (ANOVA).

Variable	LG I (n = 38)		LG II (n = 333)		LG III (n = 54)		VA	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>P</i>
Psychische Belastbarkeit (Langform)	129,61	17,51	120,05	17,80	106,42	18,33	20,67	.000
Psychische Belastbarkeit (Kurzform)	10,26	1,50	9,75	1,35	9,18	1,57	6,98	.001

Die Pbn der LG I weisen Rohwerte bei der Langform-Skala von 85 bis 171 auf. Die Spannweite bei der Probandengruppe mit einem durchschnittlichen Fitnessniveau liegt zwischen Rohwerten von Minimum 63 bis Maximum 158. Bei der Gruppe der weniger Ausdauertrainierten zeigen sich Werte bei der psychischen Belastbarkeitsskala von 48 bis 138. In Abbildung 34 werden die Unterschiede der psychischen Belastbarkeit unter den LG noch einmal verdeutlicht.

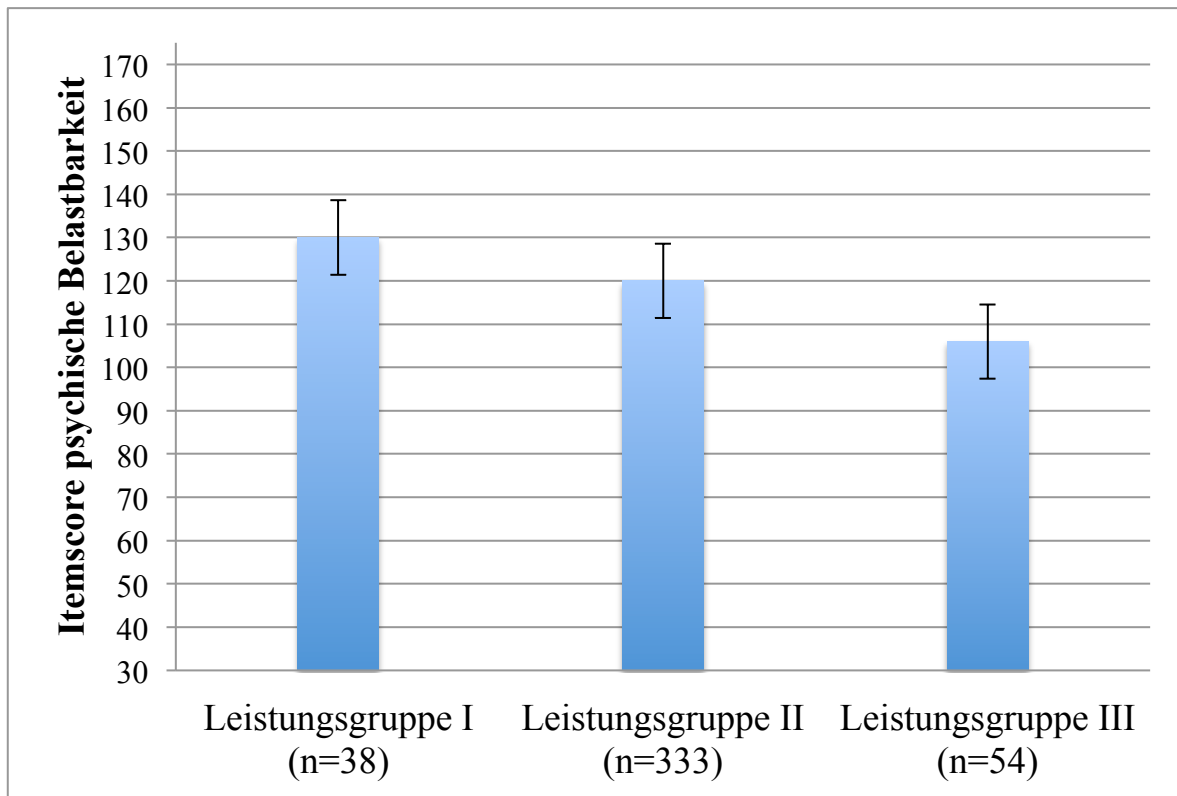


Abbildung 34. Vergleich der LG (aufgeteilt nach den Zeiten des 3000-m-FET) hinsichtlich der Langform-Skala psychische Belastbarkeit (N = 424).

Zuletzt wird im Rahmen der Gruppenvergleiche eine zweifaktorielle Varianzanalyse berechnet. Hierbei kann zur Erklärung der Zielvariablen neben den LG das Geschlecht als weiterer Faktor berücksichtigt werden. Faktor L sind hierbei die LG (dreistufig) und Faktor G steht für das Geschlecht (zweistufig). Durch den Levene-Test kann, wie auch bei den obigen Berechnungen, eine Gleichheit der Varianzen festgestellt werden, dadurch sind die Voraussetzungen für weitere Berechnungen erfüllt. In Tabelle 52 sind die beiden unabhängigen Variablen (Faktoren L und G) zur Erklärung der psychischen Belastbarkeit (Lang und Kurzform-Skala) abgebildet. Die Regressionsanalysen haben gezeigt, dass lediglich das Geschlecht und das Alter einen Einfluss auf die psychische Belastbarkeit hat (siehe Kapitel 5.3.2). Dadurch hat sich die Vermutung einer Mitbestimmung des Geschlechts als varianzaufklärende Variable hinsichtlich der psychischen Belastbarkeit bestärkt. Der Vergleich der Extremgruppen mit Hilfe der zweifaktoriellen Varianzanalyse neben der Verdeutlichung der Unterschiede hinsichtlich der psycho-physischen Belastbarkeit auch mögliche Wechselwirkungen zwischen den beiden Faktoren L und G aufdecken.

Tabelle 52

Zweifaktorielle Varianzanalyse der Faktoren L (LG) und G (Geschlecht) hinsichtlich beider Skalen der psychischen Belastbarkeit ($N = 424$).

Skalen		Männer			Frauen			Zweifaktorielle ANOVA [F-Werte]		
		Leist.- Grp. I (n = 36)	Leist.- Grp. II (n = 309)	Leist.- Grp. III (n = 13)	Leist.- Grp. I (n = 2)	Leist.- Grp. II (n = 23)	Leist.- Grp. III (n = 41)	Faktor L (Leist.-Grp.)	Faktor G (Geschlecht)	L*G
Psychische Belastbarkeit (Langform)	<i>M</i>	129,08	120,84	115,23	139,00	109,52	103,63	6,36**	0,81	1,31
	<i>SD</i>	17,83	17,36	14,06	4,24	20,54	18,78			
Psychische Belastbarkeit (Kurzform)	<i>M</i>	10,08	9,69	9,60	11,00	8,75	9,07	4,41*	0,33	1,95
	<i>SD</i>	1,31	1,25	1,43	1,73	2,12	1,61			

* $p \leq .05$; ** $p \leq .01$

Die Ergebnisse der zweifaktoriellen Varianzanalyse zeigen einen hohen signifikanten Unterschied bei den LG im Hinblick auf die psychische Belastbarkeit (Langform-Skala: $p = .002$; Kurzform-Skala: $p = .013$). Bei dem Faktor Geschlecht kann bezüglich beider Skalen der psychischen Belastbarkeit keine Varianz aufgeklärt werden. Es besteht darüber hinaus auch kein Interaktionseffekt zwischen den Faktoren L und G (Langform-Skala: $p = .271$; Kurzform-Skala: $p = .143$). Dementsprechend erscheint es naheliegend, dass die hier durch die Varianzanalyse festgestellten signifikanten Unterschiede in der Ausprägung der psychischen Belastbarkeit zwischen den LG mit einem Einfluss der physischen Belastbarkeit in Verbindung gebracht werden. Zur Aufklärung darüber, inwiefern diese Befundlage auf eine wesentliche Verbindung hinweist, sind auch die weiteren Analysen mit heranzuziehen, worauf dann in der Diskussion einzugehen sein wird.

Bei den obigen Berechnungen ist insbesondere bei der zweifaktoriellen Varianzanalyse der Mittelwert plus/minus der Standardabweichung der Gesamtstichprobe herangezogen worden,

so dass zunächst keine separate geschlechterspezifische Bildung der LG vorgenommen wurde. Folglich fällt bei genauerer Betrachtung der Zellenverteilung der weiblichen Versuchsteilnehmer auf, dass LG I lediglich zwei Versuchsteilnehmer umfasst. Dementsprechend ist eine spezifizierte Einteilung der Frauen und der Männer im Hinblick auf das Leistungskriterium erforderlich, d. h. sowohl bei der Bildung der weiblichen und männlichen LG müssen die Leistungen geschlechterspezifisch vorgenommen und somit der Mittelwert und die Standardabweichung gesondert berechnet werden. Danach ergibt sich folgendes Bild: Der Mittelwert der weiblichen Versuchsteilnehmer bei dem 3000-m-FET liegt zeitlich bei 301 Sek. (5:01 Min.) mit einer *SD* von 42,26 Sek. Anhand dieser Werte ergibt sich für die Frauen eine neue Einteilung der Leistung in drei Gruppen, wobei die Versuchsteilnehmerinnen der LG I Zeiten von 259 Sek. (4:19 Min.) oder besser, die Teilnehmerinnen, die der LG II zugeordnet werden, Zeiten zwischen 260 bis 343 Sek. (4:20 bis 5:43 Min.) aufweisen und Teilnehmerinnen der LG III den 3000-m-FET mit Zeiten von 344 Sek. (5:44 Min.) und schlechter bewältigten. Die Durchschnittszeit des 3000-m-FET der Männer liegt bei 256 Sek. (4:16 Min.) mit einer *SD* von 20 Sek. Somit ergibt sich eine Einteilung der LG, bei der die LG I Pbn mit Zeiten bis (\leq) 236 Sek. (\leq 3:56 Min.) beinhaltet. LG II umfasst die Probanden, die für die 3000 Meter Strecke Zeiten von 237 bis 275 Sek. (3:57 bis 4:35 Min.) benötigten und LG III die unfitten Versuchsteilnehmer, bei denen Zeiten von (\geq) 276 Sek. (\geq 4:36 Min.) oder mehr gemessen wurden. Tabelle 53 zeigt die Ergebnisse der varianzanalytischen Berechnungen auf Grundlage der differenzierter eingeteilten LG beider Geschlechter.

Tabelle 53

Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem $M \pm SD$ der 3000-m-FET-Zeiten ($N = 424$).

Skalen		Männer			Frauen			Zweifaktorielle ANOVA [F-Werte]		
		Leist.- Grp. I (n = 52)	Leist.- Grp. II (n = 262)	Leist.- Grp. III (n = 44)	Leist.- Grp. I (n = 6)	Leist.- Grp. II (n = 54)	Leist.- Grp. III (n = 6)	Faktor L (Leist.-Grp.)	Faktor G (Geschlecht)	L*G
Psychische Belastbarkeit (Langform)	<i>M</i>	129,90	120,31	118,36	126,33	104,87	104,16	7,83***	9,04**	1,10
	<i>SD</i>	17,81	17,14	16,46	15,90	18,01	31,82			
Psychische Belastbarkeit (Kurzform)	<i>M</i>	10,25	9,83	9,34	10,33	9,03	8,66	5,10**	2,54	0,98
	<i>SD</i>	1,39	1,24	1,56	1,36	1,69	1,75			

** $p \leq .01$; *** $p \leq .001$

Basierend auf der geschlechterspezifischeren Einteilung der LG ergibt sich nun ein anderes Bild bei der Zellenverteilung, insbesondere bei den weiblichen Probandinnen. Die Zellenbesetzung bei den physisch leistungsstärksten Frauen hat sich aufgrund der gesonderten Klassifizierung verdreifacht. Hingegen hat sich der Anteil der leistungsschwächsten Gruppe stark reduziert, sodass die LG II der Frauen den deutlich größten Anteil an Versuchsteilnehmerinnen umfasst. Die Ergebnisse zeigen ein ähnliches Bild wie bei der allgemeinen Einteilung der LG ohne Berücksichtigung der geschlechterspezifischen Leistungen ($M \pm SD$). Der Summenscore der Langform-Skala zeigt einen höchst signifikanten Unterschied zwischen den LG hinsichtlich der psychischen Belastbarkeit. Im Gegensatz zu der ersten zweifaktoriellen Varianzanalyse ist bei dieser Berechnung ein hoch signifikanter Unterschied der psychischen Belastbarkeit (Langform) auch bei dem Faktor Geschlecht (G) nachzuweisen. Ein Interaktionseffekt zwischen den LG und den Geschlechtern (L*G) bleibt bei beiden Skalen aus (Langform-Skala: $F = 1,10$; Kurzform-Skala: $F = 0,98$; Abbildung 35). Die Unterschiede der durch-

schnittlichen Itemscores, erfasst durch die Langform-Skala der psychischen Belastbarkeit, werden speziell zwischen der LG I und den anderen beiden LG deutlich. Die Männer weisen hier einen prozentualen Unterschied hinsichtlich des Itemscores mit rund 13 Prozent zwischen den LG I und LG III auf. Bei den Frauen ist der Unterschied bei der selbsteingeschätzten psychischen Belastbarkeit zwischen den LG I und III mit rund 23 Prozent noch stärker ausgeprägt (LG I: $M = 126$, $SD = 16$; LG III: $M = 104$, $SD = 32$). Die Prozentunterschiede beziehen sich auf die empirisch ermittelten Rohwerte der Langform-Skala der psychischen Belastbarkeit (Männer: $Min.: 78$; $Max.: 171$; Spannweite: $R = 93$ und Frauen: $Min.: 48$; $Max.: 142$; Spannweite: $R = 94$).

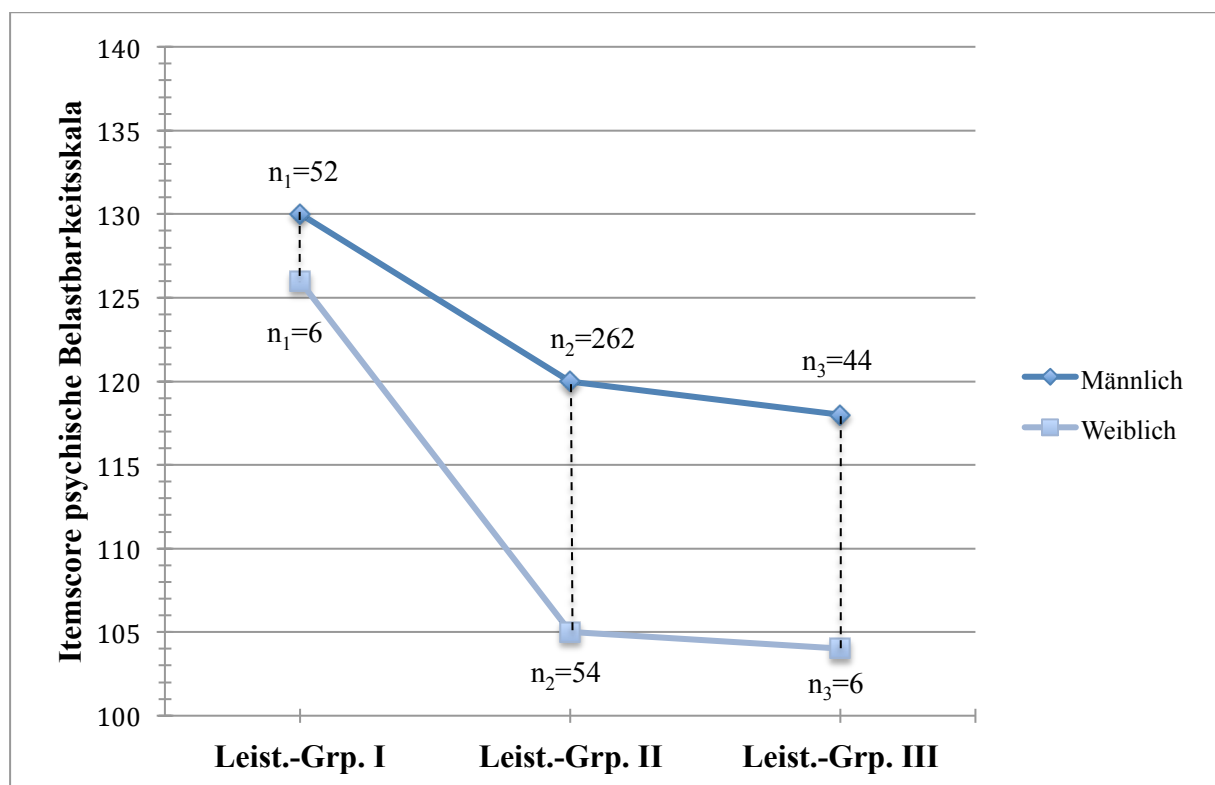


Abbildung 35. Profildiagramm der LG (eingeteilt nach der Berechnung: $M \pm SD$ der 3000-Zeiten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; $N = 424$).

In Abbildung 35 ist der geschlechterspezifische Unterschied der drei LG im Hinblick auf die psychische Belastbarkeit veranschaulicht. Durch das Profildiagramm fällt auf, dass v. a. die LG I sich bei beiden Geschlechtern deutlich psychisch belastbarer einschätzen als die Probanden der anderen beiden LG. Dieser Unterschied ist bei den weiblichen Versuchsteilnehmerinnen deutlich stärker ausgeprägt als bei den männlichen Pbn. Insgesamt verdeutlicht sich aber auch, dass die männlichen Untersuchungsteilnehmer in allen drei LG eine höhere psychische

Belastbarkeit ($M = 121,47$; $SD = 17,47$) aufweisen als die Frauen ($M = 107,05$; $SD = 20,02$; $t = 6,04$, $p = .000$). Dieser Unterschied kommt am stärksten bei den LG II und III zum tragen und zeigt sich in einem rund 15 Prozent höheren Itemscore der psychischen Belastbarkeitskala (Langform) bei den Männer im Vergleich zu den Frauen. In der Tendenz lassen sich also signifikante psychische Belastbarkeitsunterschiede zwischen den verschiedenen LG, als auch zwischen den Geschlechtern feststellen.

Eine dritte Möglichkeit der Differenzierung der LG besteht in der geschlechterspezifischen Aufteilung der Leistungen durch Drittelung der 3000-m-Zeiten. Dieses Vorgehen ist dann sinnvoll, wenn eine ungefähre Gleichverteilung der Zellenbestzungen angestrebt werden soll. Abschließend zu den zweifaktoriellen Varianzanalysen wird dieses Verfahren angewendet, um eventuelle weitere Unterschiede aufgrund einer anderen Gruppenzusammensetzung herauszufinden. Dafür wurden die 3000-m-Zeiten der Frauen und der Männer wieder in drei LG aufgeteilt, wobei die Einteilung der LG nicht anhand der Mittelwerte und der zugehörigen Standardabweichung vorgenommen wurde, sondern durch eine Drittelung der 3000-m-Zeiten sowohl für die männlichen als auch für die weiblichen Pbn. Die leistungsstärksten Frauen wiesen nach dieser Einteilung Zeiten von 281 Sek. (4:41 Min.) oder weniger auf, LG II erreichten Zeiten zwischen 282 und 310 Sek. (4:42 bis 5:10 Min.) und die leistungsschwächsten Frauen (LG III) Zeiten von 311 Sek. (5:11 Min.) oder mehr. Bei den Männer ergab sich nach diesem Verfahren folgende Zusammensetzung: LG I ≤ 247 Sek. ($\leq 4:07$ Min.), LG II 248 bis 262 Sek. (4:08 bis 4:22 Min.) und LG III ≥ 263 Sek. ($\geq 4:23$ Min.). Die Mittelwertunterschiede der geschlechterspezifischen LG im Hinblick auf beide Skalen der psychischen Belastbarkeit können Tabelle 54 entnommen werden.

Tabelle 54

Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach der Drittelung der 3000-m-FET-Zeiten ($N = 424$).

Skalen		Männer			Frauen			Zweifaktorielle ANOVA [F-Werte]		
		Leist.- Grp. I (n = 118)	Leist.- Grp. II (n = 125)	Leist.- Grp. III (n = 115)	Leist.- Grp. I (n = 21)	Leist.- Grp. II (n = 23)	Leist.- Grp. III (n = 22)	Faktor L (Leist.-Grp.)	Faktor G (Geschlecht)	L*G
Psychische Belastbarkeit (Langform)	<i>M</i>	125,44	119,30	119,70	111,95	104,78	103,86	3,54*	37,75***	0,08
	<i>SD</i>	18,12	17,52	16,14	22,72	15,86	21,12			
Psychische Belastbarkeit (Kurzform)	<i>M</i>	10,07	9,76	9,65	9,28	8,91	9,18	1,22	14,35***	0,39
	<i>SD</i>	1,36	1,22	1,34	1,51	1,61	1,56			

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Auf Grundlage dieser Einteilung liefert die zweifaktorielle Varianzanalyse weitere Aufschlüsse. Wie aus Tabelle 54 hervorgeht, ergeben sich zum einen durch die oben dargestellte Einteilung entsprechend der Logik wesentlich größere Zellenbesetzungen der LG I und III für beide Geschlechter. Zum anderen zeigen sich auch weniger starke Unterschiede der LG im Hinblick auf die psychische Belastbarkeit. Vergleicht man die Mittelwerte der einzelnen LG hinsichtlich der Langform-Skala kann festgestellt werden, dass auch bei dieser Berechnung signifikante Unterschiede zwischen den LG vorhanden sind, jedoch im Vergleich zu den obigen Varianzanalysen in abgeschwächter Form (Faktor L: $F = 3,54$; $p = .030$). Durch die Erfassung der psychischen Belastbarkeit mit Hilfe der Kurzskala, kann anhand der Mittelwertvergleiche kein signifikanter Unterschied zwischen den LG auf Basis dieser Einteilung festgestellt werden (Faktor L: $F = 1,22$; $p = .296$). Andererseits erweist sich der Unterschied des Faktors Geschlecht (G) bei dieser Berechnung als höchst signifikant (Langform-Skala: $F = 37,75$; $p = .000$; Kurzform-Skala: $F = 14,35$; $p = .000$) in Bezug auf beide Skalen der psychischen

Belastbarkeit.

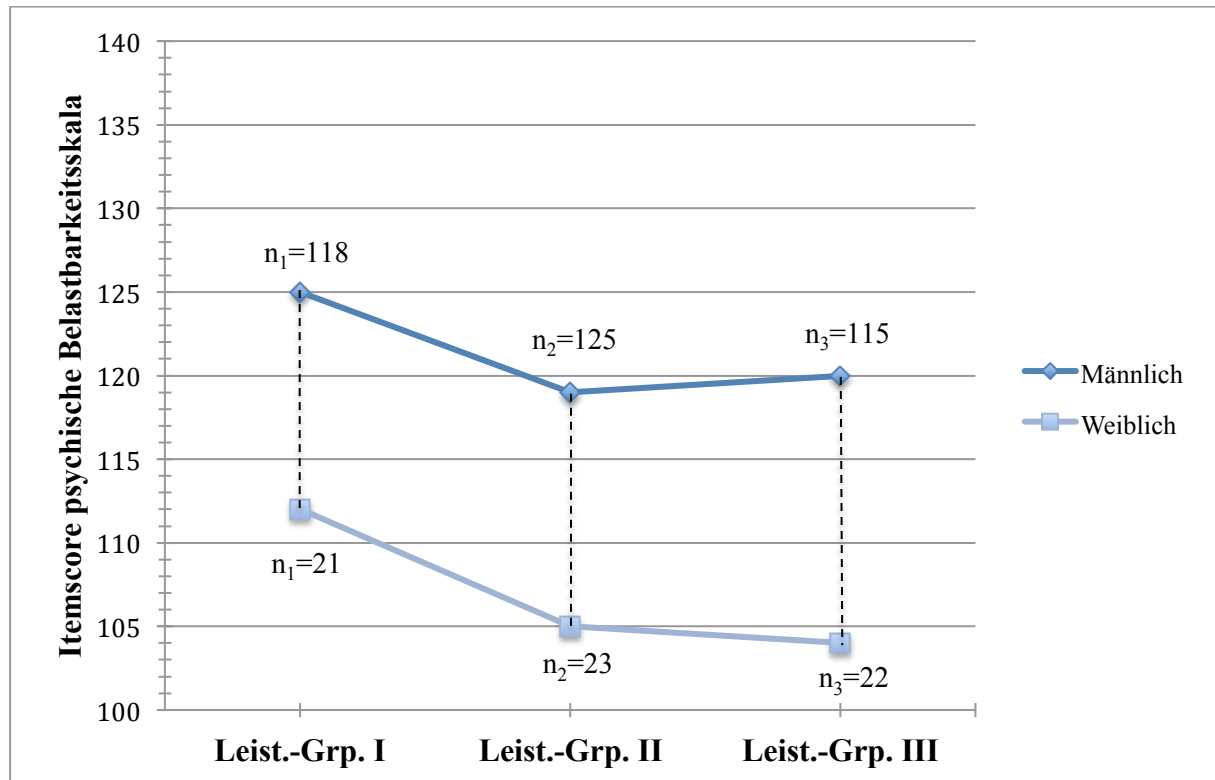


Abbildung 36. Profildiagramm der LG (eingeteilt nach Drittelung der 3000-Zeiten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; N = 424).

In Abbildung 36 ist zu erkennen, dass auf Grundlage dieser Gruppenkomposition die Unterschiede der Itemscore-Mittelwerte der psychischen Belastbarkeit (Langform) zwischen den LG nicht so deutlich ausgeprägt sind wie bei der Einteilung mit Hilfe der Berechnung $M \pm SD$. Wesentlich signifikantere Unterschiede bei der psychischen Belastbarkeit sind hierbei, wie auch aus Tabelle 54 hervorgeht, zwischen den Geschlechtern erkennbar. Anhand der folgenden Berechnung soll der Unterschied der psychischen Belastbarkeit zwischen den Geschlechtern aufgezeigt werden: Die männlichen Versuchsteilnehmer erreichten einen durchschnittlichen Summenscore bei der Langform-Skala von rund 121 und die weiblichen Probanden von 107, dies entspricht einer gemittelten Differenz des Summenscores von 14. Der am höchsten erreichte Summenscore der Frauen liegt bei 142 und bei den Männer bei 171, was im Mittel einem maximalen Itemscore beider Geschlechter von insgesamt 156,5 entspricht. Der Wert von 156,5 repräsentiert nach dieser empirischen Berechnung die 100 Prozent des erreichten Itemscores. Somit ergibt sich ein durchschnittlicher prozentualer Unterschied hinsichtlich der psychischen Belastbarkeit zwischen den Geschlechtern von 8,9 Pro-

zent.

Als zusätzlicher Vergleich der psychischen Belastbarkeit bezüglich der verschiedenen Leistungsklassen und des Geschlechts soll nun mittels der gemessenen VO_{2max} des Rampentests auf dem Laufband eine weitere Variable für die Bildung der LG herangezogen werden. Hierbei werden die Versuchsteilnehmerinnen und -teilnehmer wie bei der vorherigen Analyse anhand der ermittelten VO_{2max} -Werte in drei LG eingeteilt. Für die Einteilung werden die Mittelwerte und die Standardabweichungen der VO_{2max} -Werte beider Geschlechter separat berechnet. Der durchschnittliche VO_{2max} -Wert der Frauen beträgt bei dem Rampentest, der auf dem Laufband durchgeführt wurde, 2768 l/Min. ($SD = 664,64$ l/Min.) und für die Männer 4170 l/Min. ($SD = 506,48$ l/Min.). Nach diesem Verfahren ergeben sich wie in Tabelle 55 dargestellt, folgende LG:

Tabelle 55

LG der Frauen und Männer bestimmt durch die VO_{2max} -Werte (l/Min.) des Rampentests auf dem Laufband (N = 145).

Leistungsgruppen	Frauen (n = 27)	Männer (n = 118)
Leist.-Grp. I	≥ 3434	≥ 4678
Leist.-Grp. II	2103 - 3433	3665 - 4677
Leist.-Grp. III	≤ 2102	≤ 3664

Auf Grundlage dieser Einteilung der LG ergeben sich, wie aus Tabelle 56 zu entnehmen, korrespondierende Ergebnisse zu den bisherigen Analysen, die nach dem gleichen Verfahren jedoch mit dem Parameter Zeit (3000-m-FET) vorgenommen wurden. Es zeigen sich sowohl bei den Geschlechtern als auch bei den verschiedenen eingesetzten Skalen signifikante, teilweise hoch signifikante Unterschiede zwischen den LG im Hinblick auf die psychische Belastbarkeit ($F = 4,04$; $p = .020$ bzw. $F = 4,49$; $p = .013$). Ferner ist festzustellen, dass - wie bei den obigen Analysen - Interaktionseffekte zwischen den LG und dem Faktor Geschlecht ausbleiben und der Faktor Geschlecht erneut hoch bzw. höchst signifikante Unterschiede bei der psychischen Belastbarkeit aufweist ($F = 8,67$; $p = .004$ bzw. $F = 14,53$; $p = .000$). Durch die Berechnung mit dem Parameter VO_{2max} als Indikator für die allgemeine Ausdauerleistungsfähigkeit sind dementsprechend signifikante und hoch signifikante Unterschiede zwischen den LG bei der psychischen Belastbarkeit festzuhalten.

Tabelle 56

Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Geschlecht (G) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem $M \pm SD$ der VO_{2max} -Werte des Rampentests auf dem Laufband ($N = 145$).

Skalen		Männer			Frauen			Zweifaktorielle ANOVA [F-Werte]		
		Leist.- Grp. I (n = 18)	Leist.- Grp. II (n = 81)	Leist.- Grp. III (n = 19)	Leist.- Grp. I (n = 3)	Leist.- Grp. II (n = 22)	Leist.- Grp. III (n = 2)	Faktor L (Leist.-Grp.)	Faktor G (Geschlecht)	L*G
Psychische Belastbarkeit (Langform)	<i>M</i>	125,11	118,05	115,63	121,00	103,27	81,50	4,04*	8,67**	1,49
	<i>SD</i>	19,77	16,09	18,60	17,52	20,03	47,37			
Psychische Belastbarkeit (Kurzform)	<i>M</i>	10,44	9,78	9,68	9,66	8,72	7,00	4,49**	14,53***	1,70
	<i>SD</i>	0,98	1,06	1,20	0,57	1,66	1,42			

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Die Abbildung 37 soll die Unterschiede für die Berechnung auf Grundlage der Einteilung der LG durch die VO_{2max} -Werte noch einmal verdeutlichen. Bei der Betrachtung des Profildia-gramms fällt auf, dass die Unterschiede der psychischen Belastbarkeit zwischen den weiblichen LG wieder besonders deutlich auftreten. An dieser Stelle sei nochmals auf die Zellenbe-setzung der weiblichen Extremgruppen verwiesen, die aufgrund des Verfahrens zur Einteilung der LG relativ klein ausfällt. Die Männer haben bei der Kurzsкала einen durchschnittlichen Itemscore von 9,86 und eine Spannweite von $R = 5$ (*Minimum* = 7; *Maximum* = 12), die Frauen haben hingegen im Mittel einen Summenscore bei der Kurzsкала von 8,70 mit einer Spannweite von $R = 6$ (*Minimum* = 5; *Maximum* = 11). In Bezug auf die empirisch ermittel-ten Rohwerte der Kurzform-Skala weisen die Versuchsteilnehmerinnen eine rund 10 Prozent geringere psychische Belastbarkeit im Vergleich zu den männlichen Versuchspersonen auf.

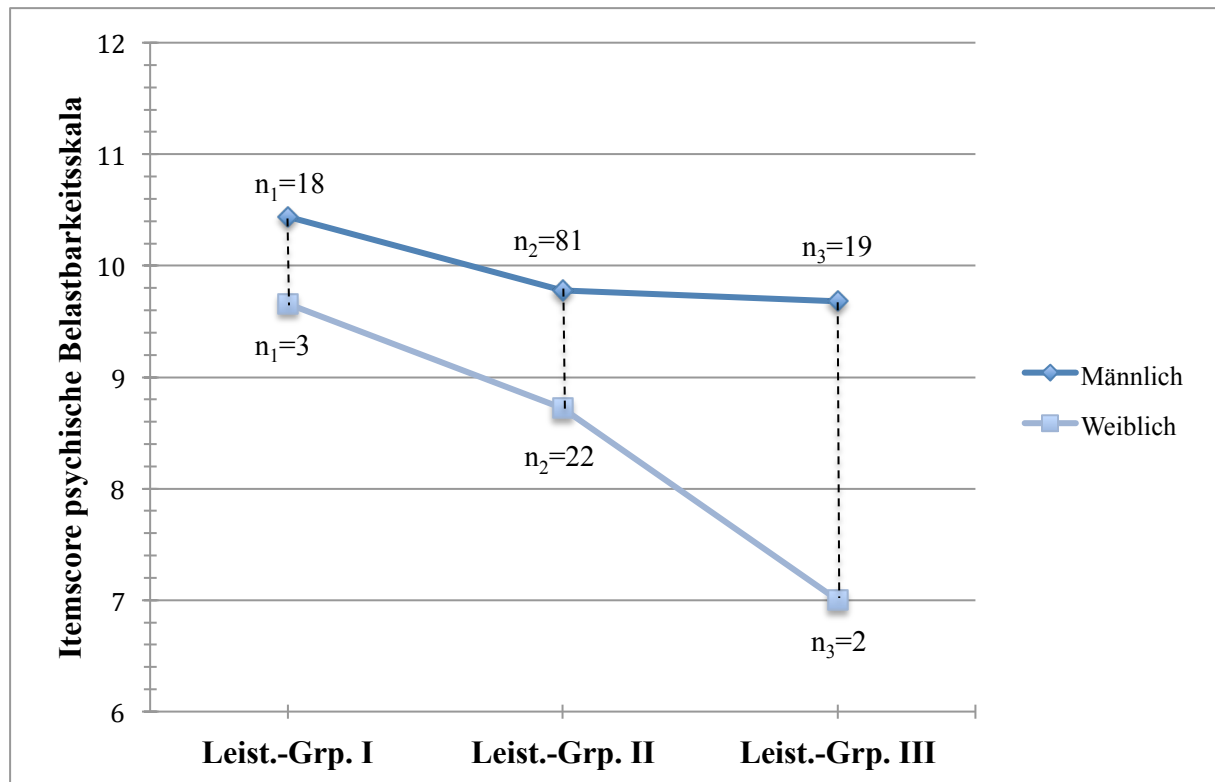


Abbildung 37. Profildiagramm der LG (eingeteilt nach den VO_{2max} -Werten) differenziert nach Geschlecht in Abhängigkeit der psychischen Belastbarkeit (Kurzform-Skala; N = 145).

Abschließend zur (zweifaktoriellen) varianzanalytischen Betrachtung der Ergebnisse soll aufgrund der Befunde der Regressionsberechnungen ergänzend zum Geschlecht auch der Faktor Alter Berücksichtigung finden. Das Alter hat wie auch das Geschlecht als einzige Variable bei den Regressionsanalysen zur Varianzaufklärung beigetragen und soll deswegen ebenfalls im Rahmen der varianzanalytischen Berechnungen mit einbezogen werden. Hierfür werden die Versuchsteilnehmer wieder in drei Altersgruppen sowie in drei LG eingeteilt. Die Einteilung der Altersgruppen erfolgt nach dem Drittelungsverfahren, sodass sich die Versuchsteilnehmer gleichmäßiger auf die Zellen verteilen. Nach dieser Komposition der Altersgruppen, die mit Hilfe des Statistik-Programms SPSS[®] vorgenommen wurde, stellt die Gruppe der ≤ 23 -Jährigen die jüngste bzw. die Altersgruppe I, Personen zwischen 24 und 29 Jahren Altersgruppe II und Personen mit einem Alter von 30 Jahren oder älter die Altersgruppe III dar. Für die Einteilung der LG wurde wieder eine Drittelung der Leistung aller Pbn (differenziert nach Geschlecht) des 3000-m-FET durchgeführt (s. o.). In Anbetracht der drei Alters- und Leistungsgruppen ergeben sich folgende Unterschiede im Hinblick auf die psychische Belastbarkeit (Lang- und Kurzform-Skala; Tabelle 57):

Tabelle 57

Zweifaktorielle Varianzanalyse mit dem Einbezug der Faktoren Alter (A) und LG (L) für beide Skalenformen der psychischen Belastbarkeit, eingeteilt nach dem Drittelungsverfahren für die Faktoren L und A ($N = 424$).

Skalen		Alters.-Grp. I (n = 151)			Alters.-Grp. II (n = 137)			Alters.-Grp. III (n = 136)			Zweifaktorielle ANOVA [F-Werte]		
		Leist.- Grp. I (n = 70)	Leist.- Grp. II (n = 42)	Leist.- Grp. III (n = 39)	Leist.- Grp. I (n = 49)	Leist.- Grp. II (n = 58)	Leist.- Grp. III (n = 30)	Leist.- Grp. I (n = 21)	Leist.- Grp. II (n = 47)	Leist.- Grp. III (n = 68)	Faktor L	Faktor A	L*A
Psychische Belastbarkeit (Langform)	M	125,47	123,00	122,10	119,57	114,16	122,79	125,55	115,02	111,97	3,62*	3,85*	2,23
	SD	19,04	20,01	15,01	20,67	18,17	17,19	16,80	14,77	18,60			
Psychische Belastbarkeit (Kurzform)	M	10,04	9,85	9,84	9,61	9,25	9,44	10,50	9,85	9,46	3,65*	5,05**	1,24
	SD	1,37	1,33	1,06	1,78	1,28	1,42	1,05	1,30	1,52			

* $p \leq .05$; ** $p \leq .01$

Es zeigen sich für beide Skalenformen signifikante Unterschiede zwischen den LG (Langform-Skala: $F = 3,62$; $p = .028$ und Kurzform-Skala: $F = 3,65$; $p = 0.27$). Unter den Altersgruppen besteht in Bezug auf die psychische Belastbarkeit, erfasst per Langform-Skala, ebenfalls ein signifikanter Unterschied ($F = 3,85$; $p = .022$). Ein deutlicherer Unterschied kann hier durch den Einsatz der Kurzsкала der psychischen Belastbarkeit ($F = 5,05$; $p = .007$) festgestellt werden. Es ergeben sich keine Interaktionseffekte zwischen den Altersgruppen und den LG. In Abbildung 38 werden die Unterschiede der psychischen Belastbarkeit, erfasst mit Hilfe der Kurzsкала, im Hinblick auf die Faktoren A und L deutlich veranschaulicht. Hierbei fällt auf, dass die Altersgruppe II (24 bis 29-Jährigen) die geringsten psychischen Belastbarkeitswerte aufweisen. Gleichwohl bei dieser Altersgruppe die psychische Belastbarkeit nicht, wie bei den bisherigen Analysen mit einem höheren physischen Leistungsniveau, zunimmt sondern bei der LG II die niedrigsten Itemscores vorzufinden sind. Des Weiteren fällt auf, dass bei der Gruppe der ältesten Probanden (Altersgruppe III: ≥ 30 Jahre) die psychische Belast-

barkeit mit einem höheren physischen Leistungsniveau einhergeht. Die fittesten Personen dieser Altersgruppe weisen im Mittel höhere Itemscores auf ($\approx 6,5\%$ bei der Langform bzw. $\approx 12\%$ bei der Kurzskala) als die physisch leistungsstärksten Personen der anderen beiden Altersgruppen (I und II). Die Gruppe der jüngsten Untersuchungsteilnehmer weisen ebenfalls zwischen den physisch Leistungsstarken und den weniger fitten Personen, wenn auch nicht so stark wie bei Altersgruppe III, höhere Itemscores bei der Kurzskala auf.

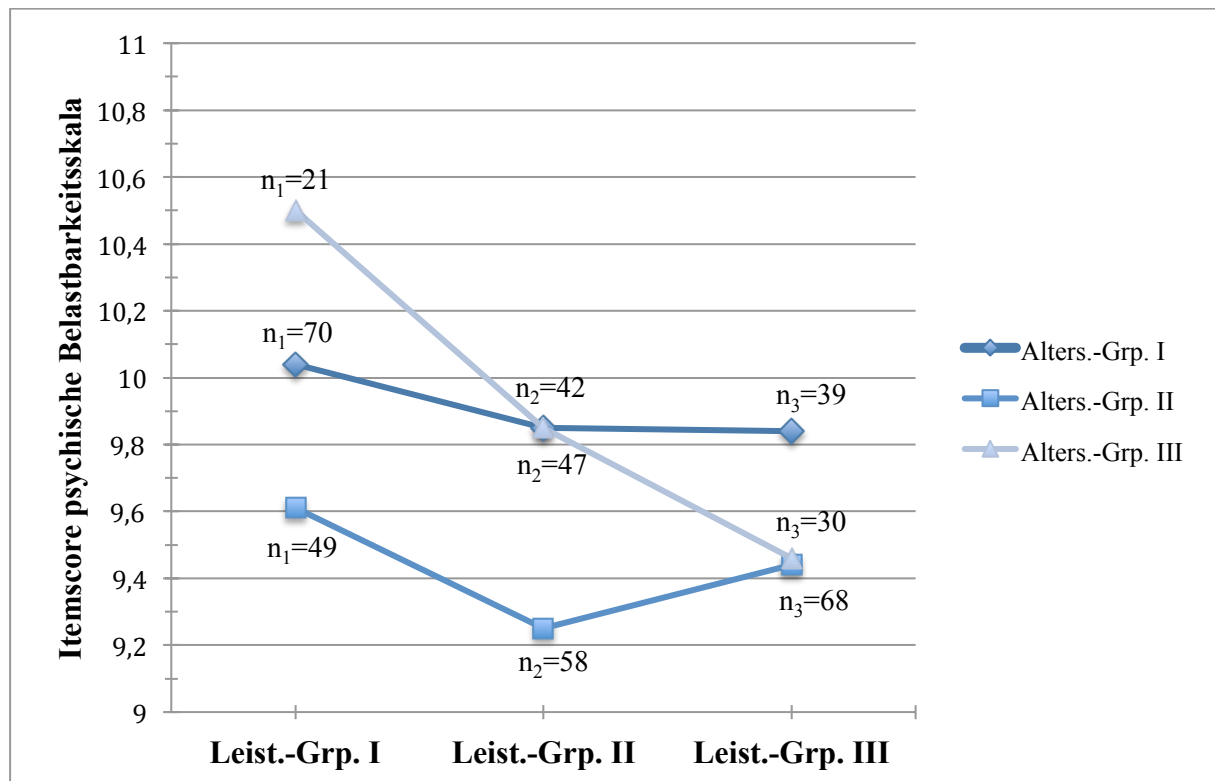


Abbildung 38. Profildiagramm der LG (Faktor L) und Altersgruppen (Faktor A) (jeweils eingeteilt nach dem Drittelungsverfahren) in Abhängigkeit der psychischen Belastbarkeit (Kurzform-Skala; N = 424).

Die jüngste Gruppe der bis einschließlich 23-Jährigen weisen bei den Leistungsschwächsten die höchsten psychischen Belastbarkeitswerte hinsichtlich der Kurzform auf. Die weniger fitten Personen (LG III) zeigen bei den Altersgruppen II und III nahezu identische Werte im Hinblick auf die psychische Belastbarkeit. Daneben sind ebenfalls im Rahmen der LG II nahezu identische Itemscores bei den Altersgruppen I und III vorzufinden. Eine sichtbar prägnantere Differenzierung der psychischen Belastbarkeits-Werte zwischen den Altersgruppen zeigt sich bei den Personen mit dem physisch höchsten Leistungsniveau (LG I). Hierbei ist festzustellen, dass die 21 getesteten Personen der Altersgruppe III (≥ 30 -Jährigen) sich im Vergleich zu den anderen beiden Altersgruppen (I und II) - wie bereits erwähnt - im Hinblick

auf die psychische Belastbarkeit deutlich unterscheiden.

An dieser Stelle scheint eine Darstellung der Ausprägungen der psychischen Belastbarkeit zwischen den drei Leistungs- und Altersgruppen seitens der Langform-Skala angezeigt (Abbildung 39), um die Ergebnisse beider Skalenformen besser gegenüberzustellen. Die Befunde lassen sich zum Teil genau replizieren und können demnach auf ähnliche Weise interpretiert werden. Gegenüber der Kurzsкала kommen bei der Darstellung der Langform-Skala die Ausprägungen aufgrund der höheren Itemanzahl deutlicher zum Vorschein. Die Unterschiede innerhalb der einzelnen Leistungs- und Altersgruppen zeigen - wie auch bei der Kurzsкала - ein nahezu identisches Bild bei den Ausprägungen der psychischen Belastbarkeit (vgl. Abbildung 38).

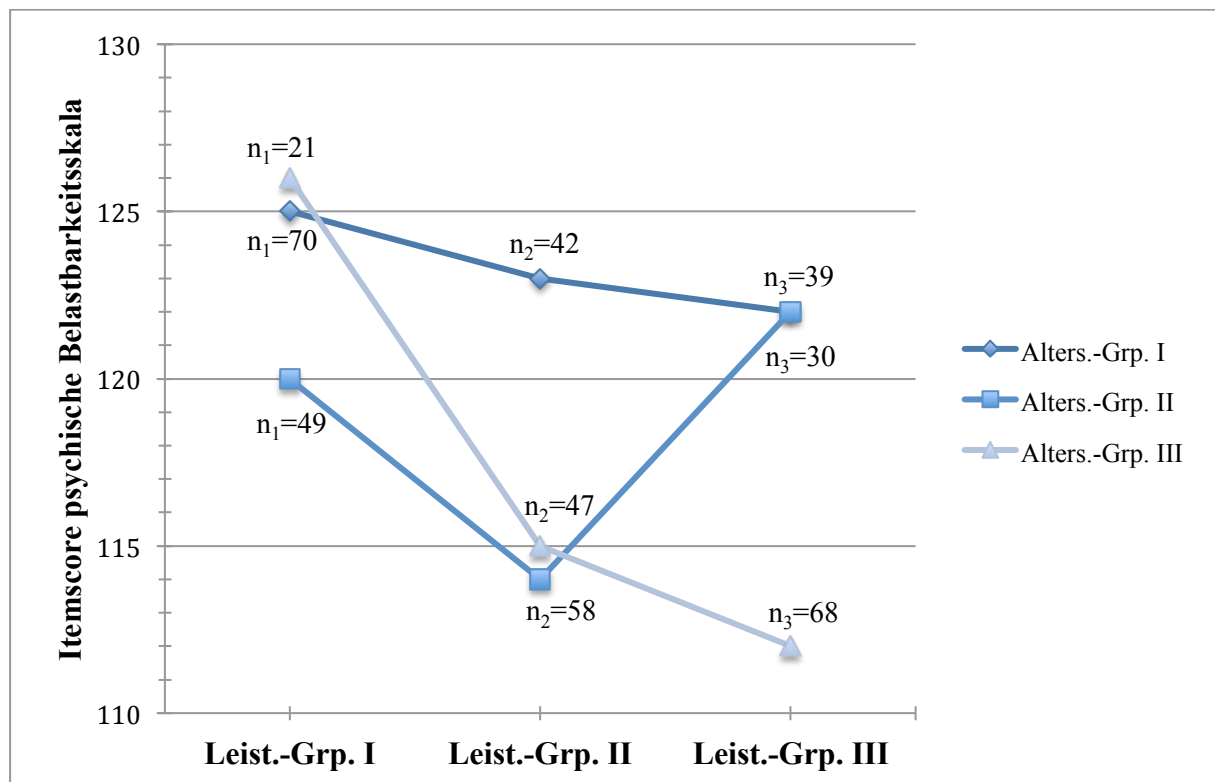


Abbildung 39. Profildiagramm der LG (Faktor L) und Altersgruppen (Faktor A) (jeweils eingeteilt nach dem Drittelungsverfahren) in Abhängigkeit der psychischen Belastbarkeit (Langform-Skala; N = 424).

Die Befunde deuten auch hier wieder darauf hin, dass im Rahmen der ältesten Gruppe die psychische Belastbarkeit mit einer besseren physischen Belastbarkeit einhergeht. In geringerer Ausprägung ist dies auch bei den jüngsten Untersuchungsteilnehmer zu beobachten (Altersgruppe I). Ebenso zeigt sich wie bei der Kurzform-Skala zur psychischen Belastbarkeit

innerhalb der Altersgruppe II, dass die Personen mit einer durchschnittlichen Fitness die niedrigsten Itemscores aufweisen. Ein Unterschied zwischen den beiden Skalen besteht lediglich darin, dass bei den 24 bis 29-Jährigen die LG III etwas höhere psychische Belastbarkeitswerte aufweist als LG I. Genau umgekehrt ist dies bei den Ergebnissen durch Erfassung mit Hilfe der Kurzskala der Fall. Zusammenfassend führt der Einsatz der Kurzform- und Langform-Skalen zu vergleichbaren Ergebnissen.

6 DISKUSSION

Nach der detaillierten Darstellung der Befunde in den vorherigen Kapiteln, gilt es nun, die Ergebnisse im Rückbezug auf das Anliegen der Arbeit und die theoretischen Erörterungen zu interpretieren und zu diskutieren. Hierzu erscheint eine Unterteilung in drei Bereiche sinnvoll: Als erstes werden die Befunde der Verfahrensentwicklung zur Erfassung der physischen, dann der psychischen Belastbarkeitskomponente nach methodischen Gesichtspunkten betrachtet und unter Berücksichtigung methodologischer Aspekte diskutiert. Im Anschluss daran sollen die Ergebnisse aus den Untersuchungen zur psycho-physischen Belastbarkeit einer eingehenden Betrachtung unterzogen werden.

Bei der Umsetzung des Anliegens, ein Verfahren für die Personalauswahl - hier speziell bei der Bundeswehr - zu entwickeln, welches dazu beitragen kann, die psycho-physische Belastbarkeit angemessen zu berücksichtigen und abschätzen zu können, ging es im ersten Schritt konkret um eine Verfahrensentwicklung zur Erfassung der physischen Belastbarkeitskomponente, in diesem Fall speziell der Ausdauerleistungsfähigkeit. Dazu wird die Konzeption, Entwicklung, Erprobung und Überprüfung des 3000-m-FET in den Mittelpunkt der Betrachtung gestellt. Wesentlich für das Anliegen (das in Verbindung mit einem eigens dazu angelegten Projekt verfolgt wurde) ist die Frage, ob der neu entwickelte 3000-m-FET nicht nur nach testtheoretischen Kriterien, sondern auch unter Beachtung einer justiziabel sicheren Einsetzbarkeit ein geeignetes Testverfahren zur Messung einer Komponente im Sinne einer Basis-Ausdauerleistung darstellt. Zur Beantwortung dieser Frage sind die Ergebnisse der Laborstudie zur Überprüfung der Gütekriterien (siehe Kapitel 5.1.2) grundlegend wichtig. In dem daran anschließenden zweiten Schritt, der sich auf die psychische Komponente der psycho-physischen Belastbarkeit bezieht, stehen bei der Verfahrensentwicklung zur Erfassung der psychischen Belastbarkeit zum einen die Befunde der Konstruktvalidierung und zum anderen der Vergleich der auf den BIP-Skalen (Emotionale Stabilität und Belastbarkeit) basierenden Langform-Skala und der daraus entwickelten Kurzform-Skala im Fokus. Im Anschluss daran werden die Ergebnisse aus den Korrelations-, Regressions- und Varianzanalysen zur psycho-physischen Belastbarkeit zunächst im Einzelnen und dann im Verbund erörtert. Die drei wesentlichen Bereiche dieser Arbeit, nämlich die Verfahrensentwicklung zur Erfassung der physischen und psychischen Belastbarkeitskomponenten sowie die Kopplung dieser beiden Verfahren im Sinne der Operationalisierung des Konstrukts der psycho-physischen Belastbarkeit, stellen die Grundlage eines Assessments psycho-physischer Belastbarkeit im Kontext der Per-

sonalauswahl dar. Im Zuge dieser Arbeitsschritte ließ sich dann auch die Frage nach psychophysischen Zusammenhängen konkretisieren und spezifisch analysieren.

Der 3000-m-FET wurde entwickelt und im Rahmen der Studien zu dieser Arbeit erprobt und validiert, um ein geeignetes Verfahren zur Überprüfung der Ausdauerleistungsfähigkeit für die Personalgewinnung der Bundeswehr zu implementieren. Die Testung der Basisfitness der (zukünftigen) Mitarbeiter der Bundeswehr dient neben der Personalauswahl auch zur jährlichen Feststellung und Dokumentation der körperlichen Leistungsfähigkeit und Belastbarkeit (siehe hierzu Leyk et al., 2013). Weiterhin stehen im Rahmen der Verfahrensentwicklung und speziell der Verfahrensevaluierung, die Reliabilitätsprüfung und die Validierung zur Erfassung der psychischen Belastbarkeit mit den beiden herangezogenen Skalen des BIP berufsrelevante Merkmale im Fokus. Dementsprechend stellt die Verbindung dieser beiden Verfahren eine Möglichkeit eines Assessments der psycho-physischen Belastbarkeit dar, welches im Personalauswahlverfahren in verschiedenen Bereichen auch über den bundeswehrspezifischen Einsatz hinaus als Screeningtool Anwendung finden könnte. Die Ergebnisse der drei Hauptanliegen sollen nachfolgend interpretiert und in der Diskussion daraufhin eingeschätzt werden, inwiefern sie einzeln betrachtet und in Verbindung gesehen eine nach testtheoretischen und den genannten pragmatischen Kriterien geeignete Möglichkeit zur Erfassung psychophysischer Belastbarkeit bei der Personalauswahl darstellen. Mit diesen Verfahren werden auf Seiten der psychischen Komponente subjektive Daten (Selbsteinschätzung mit Persönlichkeitsskalen) und Daten „objektiver Art“ durch die Testung der Ausdauer auf Seiten der physischen Komponente, so wie es bspw. Egloff und Gruhn (1998, S. 53; siehe Kapitel 3.1) ansprechen, miteinander in Verbindung gebracht.

Die Überprüfung der Zuverlässigkeit des 3000-m-FET belegt, dass von einer sehr zufriedenstellenden Retest-Reliabilität auszugehen ist. Bei Betrachtung der Befunde des Vor- und Nachtests ($n = 29$) zeigen sich bei der benötigten Zeit deutliche ($r > .90$) Zusammenhänge, die zugleich höchst signifikant sind ($p = .000$). Vergleicht man den Parameter VO_{2max} bzw. VO_{2peak} von Testung t1 mit t2, zeigen sich ebenfalls Korrelationen mit $r > .90$ ($p = .000$). Anhand der Korrelationsanalysen der Indikatoren für die physische Belastbarkeit lassen sich starke Zusammenhänge zwischen den Vor- und Nachtests nachweisen. Dies ist sowohl bei dem 3000-m-FET als auch bei den anderen drei Ausdauertests festzustellen, die dementsprechend eine sehr gute Reliabilität dieses Verfahrens belegen. Die benötigten Zeiten sowie die erreichten maximalen Sauerstoffaufnahme-Werte sind zu beiden Messzeitpunkten nahezu

identisch. Ähnlich hohe Zusammenhänge sind bei den anderen Parametern festzustellen, die im Zuge der Laborstudie gemessen wurden. Die 29 Untersuchungsteilnehmer der Reliabilitätsstudie weisen bei beiden Zeitpunkten annähernd die gleichen maximalen Herzfrequenzwerte auf (bspw. 3000-m-FET: $r = .81$; $p = .000$), sowie ähnlich hohe Laktatkonzentrationen mit Zusammenhängen zwischen Testung t1 und t2 von $r = .88$ (3000-m-FET) und beim Atemminutenvolumina mit einem höchst signifikanten Korrelationskoeffizienten von $r > .74$. Auch bei dem subjektiven Ausbelastungskriterium mit dem Kriterium des Wertes auf der der Borg-Skala zeigen sich mittlere bis hohe Zusammenhänge ($r > .55$) mit $p \leq .01$ zwischen den Testungen, die zu unterschiedlichen Zeitpunkten durchgeführt wurden. Es ist wichtig festzustellen, dass damit auch deutlich wird, dass sich die Untersuchungsteilnehmer bei allen vier Testungen zu beiden Zeitpunkten etwa gleich angestrengt gefühlt und subjektiv ausbelastet haben. Darüber hinaus lassen sich durch die Berechnung des t-Tests für abhängige Stichproben keinerlei signifikante Unterschiede bei dem 3000-m-FET zwischen den ermittelten Werten des Vor- und Nachtests feststellen.

Trotz der annähernd identischen Testergebnisse kann aus Tabelle 7 (Testdauer) entnommen werden, dass die Probanden bei den Streckentests (3000-m-FET und 1000-m-Lauf), sowie bei dem Rampentest auf dem Fahrradergometer zum Zeitpunkt t2 etwas besser waren, d. h. bei den Streckentests die Distanz schneller bewältigt und bei dem Rampentest wenige Sekunden der Belastung länger standgehalten haben, obwohl sich die leistungslimitierenden Faktoren (Tabellen 9, 11 und 13) und das subjektive Beanspruchungsempfinden (Tabelle 15) zwischen den zwei Testzeitpunkten unwesentlich unterscheiden. Für eine Erklärung dieser Leistungsverbesserung, insbesondere bei dem 3000-m-FET kann die Central Governor Theory von Noakes (2001, 2012) aufgegriffen werden. Hiernach wird angenommen, dass der Abruf der vollen Leistungsfähigkeit bzw. die Fähigkeit an seine Grenzen zu gehen und sich auszubelasten neben der Tagesform und motivationalen Aspekten, wie die Gegenwart von Konkurrenten, die Selbstüberzeugung, etc., wesentlich auch von der Vorerfahrung abhängt. Die Ausschöpfung der leistungslimitierenden physiologischen Faktoren wie bspw. der VO_{2max} , oder die der maximalen Herzfrequenz wird in erster Linie vom Gehirn reguliert. Die Leistungsverbesserung bei dem Nachtest kann somit plausibel primär auf die Vorerfahrung (t1) und weniger auf die physiologischen Faktoren zurückgeführt werden. Der Faktor VO_{2max} ist nach Noakes (2001, 2012) ein eher minderwertiger Indikator zur Erklärung der vollen Ausbelastung einer Person. Weiterhin muss mit ins Kalkül genommen werden, dass aufgrund der Vorerfahrung durch den Vortest andere Trittfrequenzen gewählt wurden, die zu verbesserten Pacing-

Strategien und somit zu schnelleren Zeiten geführt haben. Nach Noakes (2001, 2012) ist das Gehirn, das leistungslimitierende Organ, welches die Muskelzellen bzw. -arbeit reguliert und vorwiegend eine protektive Funktion einnimmt, sodass der Körper in Form von Ermüdungszuständen vor Schäden geschützt wird. In diesem Zusammenhang ist insbesondere die Vorerfahrung ein wichtiger zu berücksichtigender Aspekt, der zur Ausschöpfung der Leistungsreserven, die ein Mensch in einer Belastungssituation partizipiert, eine wesentliche Rolle spielt.

Die Funktion des Gehirns und des Zentralnervensystems (ZNS) als primär leistungslimitierendes Organ verweist auf einen möglichen Erklärungsansatz für das Ausbelastungsempfinden bzw. für das Gefühl, die Grenze der Belastbarkeit zu erreichen, den auch andere Autoren aufgreifen. Hollmann und Strüder (2009, S. 328) treffen hierzu die Aussage, dass

„vom Gehirn aus ein Schutzmechanismus eingeleitet wird, da bei Arbeit unter Sauerstoffmangelbedingungen und bei ansteigender Belastungsintensität und dem damit weiter abfallenden arteriellen O_2 -Partialdruck das Nervensystem gefährdet ist. (. . .) Hier scheint dem Gehirn, insbesondere dem präfrontalen Kortex, eine steuernde und leistungsbegrenzende Rolle zuzukommen, indem eine Limitierung der Mobilisation von motorischen Einheiten stattfindet.“

Da alle vier Tests ähnlich hohe Zusammenhänge zwischen den Testzeitpunkten aufweisen, kann diese Befundlage als Beleg für eine hohe Reliabilität bzw. Messzuverlässigkeit des entwickelten 3000-m-FET wie der anderen drei Testverfahren gelten. Die Hypothese 5 kann somit als bestätigt gelten und die Annahme angenommen werden, dass bei wiederholter Durchführung zu einem späteren Zeitpunkt die Befunde mit den Werten der ersten Testung in hohem Maße übereinstimmen (H5, siehe Kapitel 5.1.2).

Die Ergebnisse der Validitätsüberprüfung des 3000-m-FET lassen insgesamt auf eine gute Inhalts- und Kriteriumsvalidität schließen. Bei Betrachtung der Befunde der Laborstudie und der dabei ermittelten Parameter bei dem Vergleich des 3000-m-FET mit den drei weiteren eingesetzten Testverfahren zur Messung der Ausdauerleistungsfähigkeit zeigt sich, dass insbesondere die absoluten VO_{2max} bzw. VO_{2peak} -Werte in hohem Maße miteinander korrelieren ($r > .90$, $p = .000$) und nahezu gleichermaßen ($r > .80$) die relativen VO_{2max} bzw. VO_{2peak} -Werte der vier Tests in Zusammenhang stehen. Zugleich zeigen sich bei den Zeiten bzw. der Testdauer zwischen den Strecken- und Rampentests wie erwartet hohe negative Korrelationen, die sich als höchst signifikant erweisen. Insofern haben die Personen, die bei dem Rampentest den Belastungen in Form von steigender Geschwindigkeit oder zunehmender Watt-

zahl länger standgehalten haben, wie im Vorhinein angenommen, auch die Vorgaben der Streckentests i. d. R. in kurzer Zeit bewältigt. Diese Befundlage spricht dafür, dass man davon ausgehen kann, dass der 3000-m-FET das zu messende Kriterium, nämlich die Ausdauerleistung, tatsächlich misst und somit auch als ein hinreichend valider Test für die Feststellung des Konstrukts Ausdauerleistungsfähigkeit zu beurteilen ist. Die Testverfahren zur Erfassung der Ausdauerleistung, speziell mit Bezug auf den 3000-m-FET und insbesondere hinsichtlich der Parameter Zeit und VO_{2max} , haben sich bewährt und Hypothese 4 kann demnach als bestätigt gelten (H4, siehe Kapitel 5.1.2).

Die erreichten maximale Sauerstoffaufnahme-Werte des 3000-m-FET sind im Vergleich zu den VO_{2max} -Werten der anderen drei Ausdauertests, insbesondere im Vergleich zu den Tests mit Laufbelastungen, aufgrund weniger beteiligter Muskelgruppen und Mitochondrien, entsprechend der funktionalen Zusammenhänge, nicht ganz so hoch und weichen mit einer rund 9,6 Prozent geringeren VO_{2max} von dem im Durchschnitt erreichten Maximalwerten des Laufband-Rampentests ab (VO_{2peak} 3000-m-FET: $M = 3520$ l/Min.; VO_{2max} Rampentest-Laufband: $M = 3904$ l/Min.). Trotz der geringeren maximalen Sauerstoffaufnahme-Werte bei dem 3000-m-FET ist der höchste VO_{2max} -Wert (5591 l/Min.) der gesamten spiroergometrischen Studie bei dem 3000-m-FET gemessen worden. Dies könnte daran liegen, dass dieser Pbn eine gut trainierte radspezifische Ausdauerleistungsfähigkeit besitzt und die entsprechende Muskulatur besonders gut ausgebildet ist und ausbelastet werden kann, sodass die tatsächliche VO_{2max} auch bei dieser Art der Belastung zu mobilisieren ist (vgl. Meyer & Kindermann, 1999). An diesem Einzelbeispiel lässt sich gleichzeitig verdeutlichen, dass auch bei solch einem Belastungsschema (Kurzzeit- oder Mittelzeitausdauerbelastungen mit dem Fahrradergometer) mit verhältnismäßig kurzer Belastungsdauer auf dem Ergometer die tatsächliche VO_{2max} durchaus erreicht werden kann. Auch der Vergleich der physischen Belastbarkeitsindikatoren, wie die durchschnittliche Wattleistung des 3000-m-FET, sowie die maximal erreichte Wattzahl des Rampentests auf dem Fahrradergometer deuten klar darauf hin, dass ein deutlicher Zusammenhang ($r = .79$; $p = .000$) zwischen den erbrachten Leistungen beider Fahrradergometer-tests besteht. Aufgrund der gleichzeitig erfassten Ausbelastungsparameter maximale Laktatkonzentration, maximale Herzfrequenz, maximaler RQ und Wert der Borg-Skala kann konstatiert werden, dass die Pbn im Durchschnitt an ihre körperliche Belastungsgrenze gegangen sind und die Indikatoren für die physische Belastbarkeit bzw. der Ausdauerleistungsfähigkeit, insbesondere die VO_{2max} , somit als aussagekräftige Parameter zur Bestimmung der körperlichen Leistungsfähigkeit herangezogen werden können.

Die Korrelationsanalysen der vier Tests hinsichtlich der Ausbelastungsparameter maximale Herzfrequenzen und Laktatwerte zeigen mittlere und höchst signifikante Zusammenhänge. Die höchsten Laktatkonzentrationen ($M = 11,7$ mmol/l) wurden bei dem 3000-m-FET gemessen und liegen im Mittel mit rund 20 Prozent höher als der durchschnittliche Laktatwert bei den Pbn, der bei den Rampentests auf dem Laufband festgestellt wurde ($M = 9,3$ mmol/l). Hiermit bestätigt sich die Annahme, dass die kurze intensive Belastung des 3000-m-FET unter Berücksichtigung der weniger eingesetzten Muskelgruppen bei der radspezifischen Belastung zu einer stärkeren Laktatbildung bzw. Übersäuerung der entsprechenden Muskulatur führt, als dies bei den Laufbelastungen der Fall ist. Hierzu verweisen Hollmann und Strüder (2009, S. 344) auf Untersuchungen bei denen sowohl bei ausdauertrainierten als auch -untrainierten Personen „signifikant höhere Lactatspiegel im arteriellen Blut bei Radfahrbelastungen im Vergleich zur Laufarbeit, was als Zeichen eines bereits lokal überforderten aeroben Stoffwechsels gedeutet werden muß.“

Die Ergebnisse der Ausbelastungsindikatoren, speziell der Laktatanalysen, deuten klar daraufhin, dass bei der Belastung des 3000-m-FET die anaerob-laktazide Energiebereitstellung eine wesentlich größere Rolle für die Leistungserbringung spielt als bei den anderen Testverfahren, für die eine längere sowie stetig steigende Belastung in Form der Geschwindigkeit oder der Wattstufe (Widerstand) charakteristisch sind. Dies ist insbesondere auf die extreme Beschleunigung der Pbn zu Beginn der Testphase des 3000-m-FET zurückzuführen. Exemplarisch wird durch das Layout der Tacx-Trainer-Software (siehe Abbildung 14) graphisch die Geschwindigkeit, Trittfrequenz und die Leistung während der Testung aufgezeichnet und deutlich gemacht, wie die Pbn meist äußerst intensiv die erste Min. der Testphase bewältigen und sich die Leistung nach den ersten 40 bis 60 Sek. i. d. R. auf einem konstanten Niveau einpendelt. Entsprechend der intensiven Mittelzeitausdauerbelastung bei dem 3000-m-FET wird die Energiegewinnung während der ca. ersten 60 Sek. überwiegend aus der Verstoffwechslung von Kohlenhydraten (Glukose) gewonnen. Diese werden nicht vollständig abgebaut (anaerobe Glykolyse), sodass sich infolge dessen in der beanspruchten Beinmuskulatur verstärkt Laktat anhäuft. Je nach Höhe des Widerstandes durch den eingelegten Gang kann die Belastung des 3000-m-FET im Allgemeinen als Kraftausdauerbelastung bezeichnet werden, die bei einem besonders erfolgreichen Abschneiden (kurze Testzeit) auch eine große anaerobe Kapazität bzw. Säuretoleranz der Testperson voraussetzt.

Angesichts der Befundlage ist festzustellen, dass der neu entwickelte 3000-m-FET im Rahmen eines Screenings des Fitnesszustands bei der Personalauswahl (in diesem Fall der Bundeswehr) ein durchaus geeignetes Verfahren ist, das zum einen das zu erfassende Kriterium der allgemeinen Ausdauerleistungsfähigkeit äußerst zufriedenstellend ermittelt und zum anderen ein standardisiert durchzuführendes, also von infrastrukturellen Gegebenheiten und Witterungsbedingungen unabhängiges, Testverfahren darstellt, welches auch zeitökonomisch durchführbar ist. Darüber hinaus bietet das zunächst für die Anliegen der Bundeswehr im Rahmen ihrer Personalgewinnung entwickelte Testverfahren zur Überprüfung der allgemeinen Ausdauerleistungsfähigkeit eine eignungsdiagnostische Möglichkeit, auf diese Art und Weise das Anliegen eines Screenings des Fitnesszustands von Bewerbern auch im Personalauswahlverfahren in anderen Bereichen zu berücksichtigen. Durch die sitzende Position während der Testung ist es des Weiteren möglich, parallel spiroergometrische Parameter zur weitergehenden Beurteilung der physischen Leistungsfähigkeit bzw. Belastbarkeit zu erfassen, oder andere Untersuchungen zur Feststellung des Gesundheitszustandes, wie bspw. Aufzeichnungen eines Elektrogramms (EKG; siehe z. B. auch aus der Fünten et al., 2013), bei Bedarf relativ einfach zu realisieren. Ein weiterer, insbesondere gesundheitlich zu beachtender Gesichtspunkt ist der Faktor des eigenen Körpergewichts, der bei dem 3000-m-FET im Vergleich zu Laufbelastungen eine geringe(re) Rolle spielt, sodass eine Fahrradergometerbelastung für Personen mit höherem Gewicht oder auch fortgeschrittenen Alters unproblematisch(er) ist (siehe z. B. Kamon & Pandolf, 1972).

Insgesamt sprechen die Ergebnisse der Laborstudie zur Überprüfung der Gütekriterien und zur Evaluation des 3000-m-FET unter Berücksichtigung der angeführten und erläuterten Kriterien für den Einsatz des so konzipierten Verfahrens im Kontext der Personalauswahl in den Karrierecentern der Bundeswehr. Die Befunde deuten darauf hin, dass das entwickelte Testverfahren die physische Belastbarkeit im Sinne der Ausdauerleistungsfähigkeit hinreichend bestimmbar macht und auf Grundlage dessen im Rahmen weiterreichender Untersuchungen zur psycho-physischen Belastbarkeit herangezogen werden kann.

Das zweite herausgestellte Anliegen dieser Arbeit bestand in der Konzeption eines Verfahrens zur Erfassung des Persönlichkeitsmerkmals psychische Belastbarkeit sowie der Entwicklung eines möglichst zeiteffizient durchführbaren Screeningstools dieser Persönlichkeitseigenschaft. Die Kombination der Skalen des BIP Emotionale Stabilität und Belastbarkeit zu einer psychischen Belastbarkeitsskala (vgl. Hossiep & Paschen, 1998) erscheint insbesondere

im Kontext der Personalauswahl, der in dieser Arbeit im Fokus steht, aufgrund der Berufsbezogenheit der Items als praktisch und passend. Darüber hinaus stellt die aus der so konstruierten psychischen Belastbarkeitsskala abgeleitete Kurzskala ein Verfahren dar, das auf Grundlage der Befunde zumindest als ein geeignetes Screeningtool herangezogen werden kann, um die psychische Belastbarkeit sozusagen im Schnellverfahren einzuschätzen. Der Zusammenhang zwischen der Lang- und Kurzform ist mit einem Korrelationskoeffizienten i. H. v. $r = .60$ ($p = .000$) angezeigt und deutet darauf hin, dass die Langform für eine genauere Erfassung der psychischen Belastbarkeit aufgrund der Anzahl der Items der Kurzform vorzuziehen ist. Dies kann zum Zwecke eines groben und zeiteffizienten Screenings jedoch als ein durchaus brauchbares Instrument für die Praxis eingeschätzt und empfohlen werden. Ferner weisen die Ergebnisse der Varianzanalysen auf eine hohe Ähnlichkeit der beiden Verfahren hinsichtlich der Messung der Dimensionsausprägung hin. Beim Vergleich der Leistungs- (Ausdauerleistung) und Altersgruppen im Hinblick auf die psychische Belastbarkeit zeigen sich sowohl mit dem Einsatz der Lang- als auch der Kurzform (LF und KF) bei den Berechnungen mit Hilfe der zweifaktoriellen Varianzanalyse annähernd kongruente Ergebnisse. Dies spricht dafür, dass auch durch die Kurzskala im Vergleich mit der Langform-Skala Unterschiede in fast identischer Ausprägung zwischen den Gruppen hinsichtlich der psychischen Belastbarkeit aufgedeckt werden können. Die Hypothese H6 (siehe Kapitel 5.2) kann somit nicht uneingeschränkt bestätigt werden. Der Zusammenhang liegt mit $r = .60$ nicht in einem hohen Bereich, kann jedoch als durchaus beachtlich eingestuft werden. Dies ist so zu deuten, dass damit die Hypothese zwar zu relativieren ist, aber auch nicht als völlig widerlegt erscheint. Mit der Länge einer Skala gehen sicherlich auch weitere Nuancierungen, Differenzierungen oder Spezifikationen im Hinblick auf die angezielte Dimension ein. Bei einer Reduktion der Skala kann dies nicht vollständig abgebildet werden. Weiterhin könnte sich durch die Art der Itemdarstellung eine Akzentuierung ergeben, die in folgenden Analysen nachzugehen wäre.

Die Befunde aus den Vergleichen der beiden Skalen zur psychischen Belastbarkeit einerseits mit der dazu divergenten Skala Neurotizismus und andererseits mit der dazu konvergenten Skala zur allgemeinen Selbstwirksamkeit deuten darauf hin, dass bei beiden Skalenformen zur Erfassung der psychischen Belastbarkeit von einer hinreichenden Konstruktvalidität gesprochen werden kann. Die Hypothesen H7 und H8 (siehe Kapitel 5.2) können somit angenommen werden. Die Ergebnisse zur Überprüfung der konvergenten Validität zeigen hochsignifikante positive Zusammenhänge, die in der angenommenen Richtung liegen (KF:

$r = .34$; $p = .000$ und LF: $r = .55$; $p = .000$). Dementsprechend kann die Annahme, dass psychisch belastbare(re) Personen i. d. R. sich selbst auch als allgemein selbstwirksam(er) einschätzen, als bestätigt gelten. Weiterhin konnten bei den divergenten Validitätsüberprüfungen der beiden Skalenformen mit der Neurotizismus-Skala aus dem NEO-FFI noch deutlichere Zusammenhänge, diesmal annahmeentsprechend mit negativen Vorzeichen, festgestellt (KF: $r = -.58$; $p = .000$ und LF: $r = -.62$; $p = .000$) werden. Dementsprechend weisen die Untersuchungsteilnehmer, die sich als psychisch belastbar(er) einschätzen zumeist auch niedrige(re) Itemscores bei der Neurotizismus-Skala auf, womit das divergente Merkmal, nämlich die emotionale Labilität, gemessen wird. Zudem zeigt sich auch durch den etwa gleichermaßen ausgeprägten Zusammenhang beim Vergleich der Werte der Langform- und der Neurotizismus-Skala sowie bei dem Vergleich zwischen der Kurzform- und der Neurotizismus-Skala (Werte), dass auch hier eine starke Ähnlichkeit zwischen der Lang- und Kurzform zur Erfassung der psychischen Belastbarkeit besteht. Dies belegt für beide Formen ihre Brauchbarkeit.

Die beiden evaluierten Verfahren zur Erfassung der physischen und psychischen Belastbarkeit wurden im Teil drei der Untersuchungen miteinander kombiniert und so die physische mit der psychischen Belastbarkeitskomponente operational verknüpft. Dabei kam hauptsächlich der 3000-m-FET zur Operationalisierung der physischen Leistungsfähigkeit bzw. Belastbarkeit zur Anwendung, wie er für den Einsatz im Rahmen des Personalauswahlprozesses in der Bundeswehr überprüft wurde. Zur Feststellung des Persönlichkeitsmerkmals psychische Belastbarkeit kam vorwiegend eine Langform-Skala zum Einsatz, die auf einem Persönlichkeitsinventar fußt, welches sich bereits in verschiedensten berufsbezogenen Bereichen bewährt hat, sowie eine daraus entwickelte Kurzsкала, die in Form einer positiven und negativen „Wolke“ einschlägige, die psychische Belastbarkeit umschreibende, Adjektive beinhaltet. Diese beiden Verfahren lassen sich im Anschluss an die vorgenommene Interpretation und Diskussion der Befunde als nützliche Verfahren einordnen. In Kombination mit dem Verfahren zur Testung der physischen Belastbarkeitskomponente bieten sie im Kontext der Personalauswahl ein geeignetes und passendes Instrumentarium zum Assessment der psychophysischen Belastbarkeit. Die folgenden Überlegungen und Erläuterungen betreffen die Verfahrenskombination und die Möglichkeit, damit ein geeignetes Instrumentarium für ein Assessment psycho-physischer Belastbarkeit zur Verfügung zu stellen.

Der Vergleich des Summenscores der Kurzsкала mit der Zeit des 3000-m-FET weist einen geringen, höchst signifikanten Zusammenhang auf ($r = .20$; $p = .000$). Beim Vergleich der

Langform-Skala und der 3000-m-FET-Zeit zeigt sich ein etwas höherer und ebenfalls höchst signifikanter Zusammenhang ($r = .27$; $p = .000$). Die Unterschiede der beiden Korrelationen weisen darauf hin, dass die beiden Skalenformen das Merkmal nicht völlig gleich abbilden und der Unterschied wahrscheinlich darauf zurückzuführen ist, dass die Langform-Skala mit den 29 Items das Merkmal psychische Belastbarkeit eventuell etwas differenzierter (s. o.) und ggf. stabiler erfasst. Evtl. ist es auch so, dass ggf. Antworttendenzen weniger Einfluss auf das Ergebnis nehmen. Auch wenn die Korrelationen in einem geringen Bereich liegen, verweisen die gleichen Vorzeichen und die hohen Signifikanzen auf die Brauchbarkeit des Screeningverfahrens, was auch durch die weiteren Indikatoren der Ausdauerleistungsfähigkeit belegt wird und nachgewiesen werden kann.

Zu dieser Argumentationslinie passend zeigen die Ergebnisse der Korrelationsanalysen zwischen der absoluten VO_{2peak} des 3000-m-FET und den beiden psychischen Belastbarkeitsskalen etwas deutlichere Zusammenhänge, insbesondere zwischen der Kurzsкала und der erreichten VO_{2peak} des 3000-m-FET ($r = .343$; $p = .000$). Die VO_{2max} bzw. VO_{2peak} ist ein Indikator, der funktional näher am zu messenden Merkmal liegt. Das unter Berücksichtigung funktionaler Zusammenhänge die VO_{2max} als proximaler Indikator (Parameter der sehr eng mit der kardiopulmonalen Leistungsfähigkeit assoziiert wird) für die Ausdauerleistungsfähigkeit gesehen wird, während die benötigte Zeit für die Bewältigung der 3000 Meter Strecke eher als distaler Indikator einzuordnen ist. Demgegenüber ist interessant, dass durch die Erfassung der Ausdauerleistung mit dem Laufband bzw. bei den beiden Laufbelastungen (speziell bei dem 1000-m-Lauf) das zu messende Merkmal tatsächlich differenzierter erfasst wird (siehe Scatterplot, Abbildung 32). Wobei anzumerken bleibt, dass gerade zur Feststellung der maximalen Sauerstoffaufnahme bzw. der allgemeinen Ausdauerleistungsfähigkeit ein höherer Zeitaufwand sowie eine Laufbelastung erforderlich ist. Die Kehrseite der Testökonomie (3000-m-FET; Kurzsкала) spiegelt sich in einer Reduzierung der Genauigkeit bei der Feststellung des Kriteriums wieder, was bei Testentwicklungen und bei der Durchführung von zeiteffizienten Testmaßnahmen durchaus nicht ungewöhnlich ist.

Die Ergebnisse der Zusammenhangsanalysen zwischen den Indikatoren der physischen Belastbarkeit und den Indikatoren der psychischen Belastbarkeit, die mit den beschriebenen Testverfahren (3000-m-FET und psychische Belastbarkeitsskalen: Langform- und Kurzform) erfasst wurden, weisen je nach Parameter seitens der physischen Belastbarkeit (Testdauer, Wattleistung oder VO_{2peak}) und je nach Verfahren auf Seiten der psychischen Belastbarkeit

(Kurzform- oder Langform-Skala) Unterschiede hinsichtlich der Höhe des Korrelationskoeffizienten auf. Dies bedeutet, dass sich die Ausprägung der hier untersuchten psychophysischen Zusammenhänge in Abhängigkeit von den eingesetzten Verfahren und den daraus gewonnenen Parametern unterschiedlich darstellt. Die psycho-physischen Zusammenhänge fallen je nach dem, welche Verfahren man miteinander kombiniert bzw. welche Indikatoren man miteinander vergleicht, mehr oder weniger deutlich aus. Dies ist bspw. bei Betrachtung der Ergebnisse der Korrelationsanalysen zwischen den Itemscores der psychischen Belastbarkeits-Skalen oder auch der Neurotizismus-Skala und den Zeiten oder VO_{2max} -Werten des 1000-m-Laufs zu erkennen. Im Vergleich dieser psychischen und physischen Belastbarkeitsindikatoren zeigen sich sichtbar höhere Zusammenhänge. Ein möglicher erklärender Hinweis auf die höheren Zusammenhänge zwischen den Leistungsindikatoren des 1000-m-Laufs (Testdauer) und der psychischen Belastbarkeit liegt in der Streuung (die ausgeprägtere Streuung schlägt sich in der Berechnung des Koeffizienten nieder) des Leistungsindikators Zeit.

Die Zusammenhangsanalysen (siehe Scatterplot; Abbildung 32) der 1000-m-Zeit und der Langform-Skala zeigen eine größere Streuung als beim 3000-m-FET hinsichtlich des physischen Belastbarkeitsindikators, bei dem die Zeiten enger zusammenliegen, was auch durch die Punktwolke in Abbildung 31 erkennbar ist. Der höhere Zusammenhang zwischen der Zeit sowie der VO_{2max} des 1000-m-Lauftests und der psychischen Belastbarkeit kann dadurch begründet werden, dass der 1000-m-Lauf das Konstrukt Ausdauerleistungsfähigkeit differenzierter erfasst als der 3000-m-FET. Bei Betrachtung der Zusammenhangsanalysen zwischen den Itemscores der Neurotizismus-Skala und den Indikatoren für die Leistungsfähigkeit bzw. physischen Belastbarkeit (Testdauer, VO_{2max} bzw. VO_{2peak} und durchschnittliche bzw. maximale Wattleistung), die anhand der verschiedenen Ausdauerests erfasst wurden, ergeben sich ebenfalls nahezu übereinstimmende Befunde, die in die erwartete Richtung deuten und im Gegensatz zu den psychischen Belastbarkeitsskalen ein umgekehrtes Vorzeichen aufweisen. Auch hier zeigen sich zwischen den physischen Belastbarkeitsindikatoren des 1000-m-Laufs und den Neurotizismus-Werten höhere Zusammenhänge im Vergleich mit den anderen Testverfahren, die die Ausdauer messen (siehe Tabelle 46). Im Falle dieser Zusammenhangsanalysen sei erneut auf die Stichprobengröße verwiesen, die bei dem 3000-m-FET für den Belastbarkeitsindikator Zeit deutlich größer ist ($N = 424$; siehe Kapitel 5.3) als bei den im Rahmen der Laborstudie erhobenen Leistungsindikatoren der anderen Ausdauerestverfahren (1000-m-Lauf sowie die beiden Rampentests), wobei insgesamt eine deutlich kleinere Stich-

probe (N = 145; siehe Kapitel 5.1.2) getestet worden ist. Es gibt eine Vielzahl von Faktoren, die den Zusammenhang zwischen physischer und psychischer Belastbarkeit beeinflussen können. Die Höhe des Korrelationskoeffizienten ist dabei auch im Wesentlichen abhängig von verschiedenen Faktoren, wie bspw. der Größe der Stichprobe und der Heterogenität der Stichprobe. Diese nimmt im Allgemeinen mit der Stichprobengröße zu und führt i. d. R. zu geringeren Ausprägungen der Zusammenhänge. Durch die Befunde der Korrelationsanalysen kann festgehalten werden, dass der psycho-physische Zusammenhang speziell bei dem Vergleich der Zeit des 1000-m-Laufs und der Langformskala zur psychischen Belastbarkeit am deutlichsten nachweisbar ist. Die Hypothese H1 ist mit den hier herangezogenen Verfahren und den daraus gewonnenen Indikatoren für die jeweilige Belastbarkeitskomponente bestätigt worden.

Der psycho-physische Zusammenhang, der über die beiden in dieser Arbeit entwickelten Testverfahren (3000-m-FET und Kurzskala psychische Belastbarkeit) gemessen wurde, weist in die erwartete Richtung und hat sich als höchst signifikant herausgestellt. Der Korrelationskoeffizient zwischen der Zeit des 3000-m-FET und dem Itemscore der Kurzskala ist im Vergleich zu den anderen Zusammenhangsanalysen relativ gering, was sicherlich dem Anliegen geschuldet ist, den psycho-physischen Zusammenhang möglichst auch ökonomisch zu erfassen. Durch die Ökonomisierung erkaufte man sich Einschränkungen insbesondere im Hinblick auf die Reliabilität. Dies scheint in dem Zusammenhang akzeptabel, da es ja darum geht, lediglich eine Information zur Orientierung zu gewinnen bzw. ein Screening im Rahmen der Personalauswahl durchzuführen. Dazu geht es lediglich um die Erfüllung eines Minimal Kriteriums, um Hinweise für den Selektionsprozess zu generieren. Mit dem Ziel eine spezifische und differentielle Individualdiagnostik zu betreiben muss man an die Erhöhung der Präzision denken. Dies haben insbesondere die unterschiedlichen Ausprägungen der Korrelationsanalysen deutlich gemacht, die je nach Verfahrensverbund und je nach dem, wie differenziert und präzise das ausgewählte Verfahren das zu messende Merkmal erfasst, auch zu deutlich stärkeren Ausprägungen der Zusammenhänge bzw. genauer der Zusammenhangsberechnungen führen. Zusammenfassend weisen die Befunde der Korrelationsanalysen darauf hin, dass ein Zusammenhang zwischen psychischen und physischen Merkmalen konstatiert werden kann und sich dies in den Befunden aller hier angewendeten Verfahren und deren Kombinationen zeigt. Dementsprechend lassen sich auch die eingesetzten Verfahren je nach Anliegen als durchaus geeignet beurteilen, um damit die psycho-physische Belastbarkeit zumindest im Groben festzustellen.

Aus den anschließend durchgeführten Regressionsanalysen ließen sich zunächst keine weiterführenden Hinweise mehr gewinnen, die evtl. zusätzliche Informationen zu den hier analysierten Zusammenhängen und Vergleichen ergeben hätten. Es konnten dann aber Hinweise gewonnen werden, die auf einen Einfluss der Prädiktoren Alter und Geschlecht hinsichtlich der Kriteriumsvariablen psychische Belastbarkeit schließen lassen. Aufgrund dessen wurden diese Faktoren mit in die Berechnungen der Varianzanalysen einbezogen, deren Befunde im Folgenden diskutiert werden sollen.

Wie bereits in Kapitel 5.3.2 erwähnt, ergeben sich durch die hohe Interkorrelation der Indikatoren für die physische Belastbarkeit bzw. der erklärenden Variablen Probleme der Multikollinearität (siehe Anhang A1). Aus diesem Grund wurde neben den o. g. Prädiktoren und dem BMI für die Analysen mit dem Einschluss- und dem Schrittweise-Verfahren jeweils nur ein Indikator der physischen Belastbarkeit als Prädiktor hinzugezogen. Durch die Berechnungen konnten in Bezug auf die physischen Belastbarkeitsindikatoren keine Effekte auf die psychische Belastbarkeit nachgewiesen werden. Insofern kann die Hypothese H2 nicht bestätigt werden. Es stellt sich hierzu die Frage, ob ggf. die Berücksichtigung weiterer Indikatoren physischer Belastbarkeit bzw. Leistungsfähigkeit, die weniger miteinander korrelieren, beim regressionsanalytischen Vorgehen zu aufklärenderen Ergebnissen hätte führen können. D. h., in Folgeuntersuchungen ist daran zu denken, dass weitere Indikatoren mit Hilfe auch anderer Testverfahren oder Testungen anderer konditioneller Fähigkeiten zu berücksichtigen, die dann in regressionsanalytische Berechnungen einbezogen werden können.

Die Ergebnisse der Korrelationsanalysen zu den psycho-physischen Zusammenhängen finden in den varianzanalytischen Berechnungen Bestätigung. Die Antwort auf die übergeordnete Frage, inwiefern fittere Personen psychisch belastbarer sind als andere, kann v. a. anhand der Extremgruppenvergleiche und mit Hilfe von Varianzanalysen aufgezeigt werden. Die dabei ermittelten Werte lassen signifikante Unterschiede hinsichtlich der psychischen Belastbarkeit zwischen verschiedenen Leistungsgruppen (LG) erkennen. Eingangs des empirisch-methodischen Ansatzes (Kapitel 4) wurde die Hypothese (H3) aufgestellt, nach der von einem Unterschied der psychischen Belastbarkeit zwischen den LG in Bezug auf die körperliche Leistungsfähigkeit in Form der hier erfassten Ausdauerleistungsfähigkeit ausgegangen wird. Diese Annahme konnte durch verschiedene Berechnungen mit unterschiedlichen Einteilungen der Leistungsniveaus bestätigt werden. Die Untersuchungsteilnehmer wurden zunächst an-

hand der gemessenen Indikatoren für die physische Belastbarkeit in drei Gruppen eingeteilt. Für die Einteilung der LG bzw. der Differenzierung der Leistungsniveaus wurden verschiedene Verfahren angewendet wie bspw. die Einteilung auf Grundlage des Mittelwertes plus/minus der Standardabweichung, sowie durch das „Drittungsverfahren“, wobei u. a. ausgehend von der Spannweite der Zeiten des 3000-m-FET die Pbn nach drei Zeitabschnitten aufgeteilt wurden, sodass eine etwa gleich große bzw. ausgeglichene Zellenbesetzung erreicht werden konnte. Darüber hinaus wurden aufgrund der Ergebnisse der Regressionsberechnungen die Faktoren Alter und Geschlecht als zusätzliche Faktoren mit in die Varianzanalysen einbezogen. In Anbetracht dieser Faktoren und durch die Differenzierung von geschlechter-, sowie altersspezifischen LG haben sich ebenfalls signifikante Unterschiede herausgestellt.

Im Rahmen der einfaktoriellen Varianzanalyse zeigen sich deutliche Unterschiede (LF: $F = 20,67$ bzw. KF: $F = 6,98$; $p \leq .001$; siehe Tabelle 51) in Bezug auf die Itemscores bei beiden Skalenformen der psychischen Belastbarkeit zwischen den LG I, II und III (eingeteilt auf Grundlage der Zeit des 3000-m-FET). Dies ist im Anschluss an die festgestellte grobe Differenzierung der Versuchsgruppen bzw. Stichproben außer in vorliegenden Untersuchungen (siehe Kapitel 3) nicht überraschend und müsste in weiterführenden Untersuchungen stärker berücksichtigt werden. Danach kann darauf geschlossen werden, dass sich die physisch fitten Personen als signifikant psychisch belastbarer einschätzen als die weniger körperlich leistungsfähigen Teilnehmer der Untersuchungen. Diese Annahme wird durch die durchschnittlichen Itemscores der Langform-Skala zur psychischen Belastbarkeit (LG I: $M: 130$, $SD: 18$; LG II: $M: 120$, $SD: 18$ und LG III: $M: 106$, $SD: 18$; siehe Tabelle 51) bestätigt. Herauszustellen ist, dass sich damit deutliche Hinweise zur weiteren Aufklärung eines psycho-physischen Zusammenhangs mit spezifischem Bezug zur psycho-physischen Belastbarkeit insbesondere im Verbund mit den o. a. Korrelationsanalysen ergeben.

Diese Analysen, zum einen die des Zusammenhangs (Korrelationen zwischen physischen und psychischen Parametern) und zum anderen, die der Gruppenunterschiede (Einteilungskriterium: Ausdauerleistungsfähigkeit), verweisen zusammengenommen auf die Berechtigung der Annahme eines Einflusses der physischen Belastbarkeit auf die psychische Belastbarkeit, die allerdings letztlich nur mit Hilfe experimentell angelegter Studien im Sinne einer Wenn-dann-Aussage überprüfbar, belegbar oder gar nachweisbar ist. Besonders interessant und relevant dürfte dies für entsprechend angelegte Interventionsstudien mit einem Experimental-Kontrollgruppen-Design sein. Zugleich ist zu erkennen, dass sowohl der Einsatz der Lang- als

auch die Kurzform-Skala zu ähnlichen Ergebnissen führt, was für die Berechtigung des Einsatzes im Rahmen eines zeitökonomischen Screenings auch für die Kurzform-Skala spricht.

Darüber hinaus können auch bei den zweifaktoriellen Varianzanalysen signifikante Unterschiede zwischen den LG (unabhängig vom Einteilungsverfahren) festgestellt werden, die sich auch bei einer Differenzierung der geschlechterspezifischen Leistungen abbilden (siehe Tabelle 52 und 53). Besonders extreme Unterschiede im Hinblick auf die psychische Belastbarkeit (erfasst mit der Langform-Skala) ergeben sich im Vergleich der LG der Frauen. Hierbei fällt auf, dass insbesondere die fittesten Frauen eine deutlich höhere psychische Belastbarkeit aufweisen (23 %) als dies bei den Frauen der LG II und III, also den unteren Leistungsniveaus zuzuordnen sind, der Fall ist. Dieser Befund zeigt sich in ähnlicher, jedoch weniger starken Ausprägung auch bei den Männern, bei denen sich auch die ausdauerleistungsfähigen Personen als tendenziell psychisch belastbarer einschätzen als die Pbn der LG II und III (13 %). Die Ergebnisse, die in den Tabellen 51 und 52 abgebildet sind, zeigen, dass sowohl mit als auch ohne geschlechterspezifische Differenzierung der Leistungsfähigkeit signifikante Unterschiede im Hinblick auf die psychische Belastbarkeit bestehen. Unter Berücksichtigung einer separaten Einteilung der Leistungen (bspw. Zeit: 3000-m-FET) für die Männer als auch für die Frauen kommen die geschlechterspezifischen Unterschiede im Hinblick auf die psychische Belastbarkeit in Abhängigkeit des Fitnesszustands noch deutlicher zum Ausdruck und zeigen die extremsten Ausprägungen zwischen den fittesten und unfittesten weiblichen Pbn.

Des Weiteren ergeben auch die Varianzanalysen zwischen den LG auf Grundlage des physischen Belastbarkeitsindikators VO_{2max} (Rampentest Laufband; siehe Tabelle 56) signifikante Differenzen hinsichtlich der psychischen Belastbarkeit bei beiden Geschlechtern. Bei den Ergebnissen dieser Berechnungen fällt auf, dass der Unterschied durch den Einsatz der Kurzskala ($F = 4,49, p = .013$) größer ausfällt als bei der Erfassung mit Hilfe der Langform-Skala ($F = 4,04, p = .020$; Tabelle 56). Bezüglich des F -Wertes und der Signifikanzen zeigen sich nur geringfügige Unterschiede zwischen den beiden Skalenformen, die je nach herangezogenen Leistungsindikator und je nach Einteilungsverfahren der LG, die Ausprägung der psychischen Belastbarkeit in etwa gleich erfassen, was ebenfalls für die Ähnlichkeit bzw. Ebenbürtigkeit dieser beiden Skalenformen spricht.

Ferner deuten auch die Ergebnisse der Berechnungen mit Hinzunahme des Faktors Alter in eine entsprechende Richtung, wobei sich auch ein ähnliches Bild unabhängig von der eingesetzten Skalenform ergibt (siehe Abbildung 38 und 39). Des Weiteren zeigen sich vergleichbare Unterschiede zwischen den Altersgruppen in Abhängigkeit der physischen Belastbarkeit, wie dies auch bei den geschlechterspezifischen Analysen der Fall ist. Die Ergebnisse der zweifaktoriellen Varianzanalyse mit den Faktoren Leistungs- und Altersgruppe deuten darauf hin, dass sich insbesondere körperlich leistungsfähige Personen im Alter von 30 Jahren oder älter hinsichtlich der psychischen Belastbarkeit signifikant von den weniger physisch belastbaren Personen der gleichen Altersgruppe unterscheiden. Durch die Profildiagramme wird veranschaulicht, wie die Ausprägung der psychischen Belastbarkeit, gemessen durch die Itemscores der Skalen, speziell bei der ältesten Gruppe mit einem höherem körperlichen Leistungsniveau zunimmt. Die Altersgruppen der 24 bis 29-Jährigen weisen auch höhere psychische Belastbarkeits-Werte (Kurzskala) bei den besonders fitten im Gegensatz zu den weniger fitten Personen der gleichen Altersgruppe auf. Jedoch stellen bei dieser Altersgruppe die am wenigsten psychisch belastbaren Personen die LG II dar. Gemessen mit der Langform-Skala ergibt sich ein kontroverses Bild, wobei die am psychisch belastbarsten 24 bis 29-Jährigen als die am wenigsten fitten erscheinen (Abbildung 39). Dagegen ergeben sich bei der jüngsten Gruppe (≤ 23 Jahre) auch Ergebnisse, die darauf hindeuten, dass im Hinblick auf die psychische Belastbarkeit diejenigen, die bei dem 3000-m-FET am besten abgeschnitten haben, sich selbst auch als psychisch belastbarer einschätzen als die weniger Fitten.

Dieser Zusammenhang könnte vor dem Hintergrund der spezifischen Lebensphase zu erklären sein. Aufgrund der Befunde ist zu vermuten, dass in der Altersgruppe der 24 bis 29-jährigen verhältnismäßig wenig Sport getrieben wird bzw. regelmäßig betriebener Sport eher weniger anzunehmen ist. In dieser Lebensphase sind bekanntlich verschiedenste Entwicklungsaufgaben zu bewältigen. Dazu sind insbesondere die Gestaltung der beruflichen Karriere, der Aufbau der eigenen Familie und die Entwicklung der Statuspersönlichkeit zu nennen (siehe hierzu die umfangreiche Literatur zur Entwicklungspsychologie der Lebensspanne, u. a. Baltes, 1990; Baltes & Eckensberger, 1979; Krampen & Reichle, 2002). Innerhalb dieser Lebensphase gibt es vielfältige Umstände bzw. Engagements, die auch zu einer psychischen Belastbarkeit beitragen können - ohne dass sich dies dann auf Sport bzw. sportlich begründete Fitness zurückführen lässt. In anderen Lebensphasen, speziell im fortgeschrittenen Alter, nimmt dann die körperliche Fitness vor allem auch für die Gesundheitssicherung und Prävention, z. B. Vorbeugung von Bewegungsmangel-/Zivilisationskrankheiten, eine bedeu-

tendere Stellung ein. Bei jungen Menschen (hier: Altersgruppe I) zeigt sich u. a. durch die zeitlichen Arrangements mit hinreichend Freizeiten, dass auch ausreichend Raum für Freizeitbeschäftigungen wie bspw. sportliche Aktivitäten vorhanden sind. Sport kann auch aus sozialen Gründen in dieser Phase eine größere Rolle spielen als es bei größerer Bindung und Verpflichtung auf berufs- und statusbezogene Anforderungen der Fall ist und wirkt sich gegebenenfalls bei einer regelmäßigen Ausübung positiv auf psychische Merkmale im Sinne einer höheren psychischen Belastbarkeit aus. Sportliches Engagement und dessen Auswirkungen (u. a. Fitness) ist durchaus in Abhängigkeit von der Lebensphase und den spezifischen Entwicklungsaufgaben zu sehen. Diese Erklärungsfolie lässt sich weiterhin heranziehen, um die Befunde der Kurzform-Skala, die insgesamt ein relativ heterogenes Gesamtbild verdeutlichen, in einen Deutungszusammenhang zu bringen. Der mit dem Konzept der Entwicklungspsychologie der Lebensspanne angesprochene Entwicklungsverlauf kann auch zur Begründung der formulierten Hypothesen herangezogen werden und weist darauf hin, dass sich in Abhängigkeit vom Alter bzw. der jeweiligen Lebensphase und des Geschlechts ein strukturierter Lebensstil mit einem regelmäßigen Engagement für sportliche Betätigungen bzw. für die körperliche Fitness, in der psychischen Belastbarkeit und der emotionalen Stabilität niederschlägt.

Je älter man wird, desto größer ist die Wahrscheinlichkeit, dass man über einen längeren Zeitraum einen aktiven Lebensstil verfolgen kann und wenn darin auch sportliche oder zumindest körperliche Aktivität integriert ist, wird sich die Betätigung sowohl im Hinblick auf den körperlichen Fitnesszustand als auch auf die psychische Belastbarkeit bzw. die emotionale Stabilität auswirken. Solche Untersuchungen sollten demnach über ganze Lebensabschnitte hinweg angelegt sein, um diverse Effekte empirisch angemessen überprüfen zu können - wobei als langfristig definierte Studien, die sich dieser Thematik annehmen, der Zeitraum durchaus länger als nur Wochen oder Monate und eher übereinige Jahre dauern sollten, evtl. sogar über Lebensphasen hinweg.

Im Rückbezug auf das theoretische Rahmenkonzept (Kapitel 2) und die entsprechenden Ausführungen zum Stand der Forschung (Kapitel 3.2) ist festzustellen, dass aus arbeitswissenschaftlicher Perspektive psychische Belastungen im Sinne der Überbeanspruchung sich auf die Gesundheit und das Wohlbefinden auswirken bzw. sich auf körperlicher Ebene bemerkbar machen können. Darüber hinaus haben verschiedene Untersuchungen gezeigt (siehe Kapitel 3; z. B. Martinsen, 1987, 2002; Mutrie, 1988, 2000, Fuchs, 2003), dass sich sportliche Aktivität positiv auf Depressions- und Burn-Out-Erkrankungen sowie Stress auswirkt. Wie in

Kapitel 2.2 angesprochen lassen sich andererseits Überlegungen dazu generieren, inwiefern sich die körperliche Leistungsfähigkeit auf die Resilienz bzw. psychische Belastbarkeit auswirkt und Zusammenhänge zwischen der physischen Leistungsfähigkeit und psychischen Belastbarkeit bestehen und/oder Unterschiede zwischen verschiedenen Leistungsgruppen hinsichtlich der psychischen Belastbarkeit existieren. Die Befunde aus den durchgeführten Untersuchungen zu den hier analysierten Zusammenhängen zwischen den physischen und psychischen Merkmalen können zwar kein eindeutiges Bild darüber liefern, dass physisch fitte(re) Personen generell auch psychisch - insbesondere im Hinblick auf berufliche Anforderungen - belastbarer und resistenter gegenüber psychischen Belastungen sind. Jedoch weisen die dazu vorgenommenen statistischen Analysen (Kapitel 5.3) in Richtung der formulierten und erläuterten Hypothesen (Kapitel 4) und werden insbesondere durch die Varianzanalysen bekräftigt. Eine allgemeine Aussage derart, dass physisch fittere Individuen - wie es hier auf den beruflichen Kontext bezogen interessiert - auch psychisch belastbarer sind, kann aufgrund der Ergebnisse aus den Gruppenvergleichen nicht ganz eindeutig behauptet werden. Im Anschluss an die varianzanalytischen Berechnungen kann jedoch angenommen werden, dass sich physisch fitte bzw. leistungsfähigere Personen tendenziell als psychisch belastbarer einschätzen. Auch in Bezug auf Beanspruchungen, wie es Weber (1985; siehe Kapitel 2.2) formuliert, lassen sich Beanspruchungen in Arbeitssituationen nicht in rein psychische und physische Beanspruchungen differenzieren, sondern sie stehen vielmehr immer in einer Wechselbeziehung zueinander. Ausdauertrainierten Personen wird oft auch eine stärkere psychische Widerstandskraft zugeschrieben (siehe z. B. Schnabel et al., 2008). Die Befunde bestätigen diese Annahme insofern, als dass hier bei den besser ausdauertrainierten Pbn ein durchschnittlich höherer Itemscore mit den psychischen Belastbarkeitsskalen feststellbar ist.

In Abwägung der berichteten Ergebnisse und der darauf beruhenden Überlegungen, wie sie gerade angestellt wurden, lässt sich festhalten, dass die Konzeption zur Beziehung zwischen Belastung, psycho-physischer Belastbarkeit und Beanspruchung (siehe Kapitel 2.2, besonders Abbildung 4) durchaus Bestätigung findet. Für weitergehende bzw. weitreichende Konsequenzen im Rahmen der Personalgewinnung ist allerdings noch keine hinreichende Basis gegeben. Dazu ist auch auf die Problematik einer Schlussfolgerung von Gruppenuntersuchungen und entsprechenden Statistiken auf individuelle Fälle hinzuweisen.

Weiterführende Hinweise lassen sich aus der Untersuchung größerer Stichproben und daraus zu berechnender Normwerte gewinnen. Für die Entscheidungsfindung in der Personalge-

winnung ist es nicht nur erforderlich, das empirische sondern auch das theoretische Bezugssystem hinsichtlich der konkreten Anforderungen zu spezifizieren. Zudem sollte auch die Veränderbarkeit bzw. Trainierbarkeit im Rahmen des Personalmanagements Berücksichtigung finden. Dies veranschaulicht die folgende Abbildung 40.

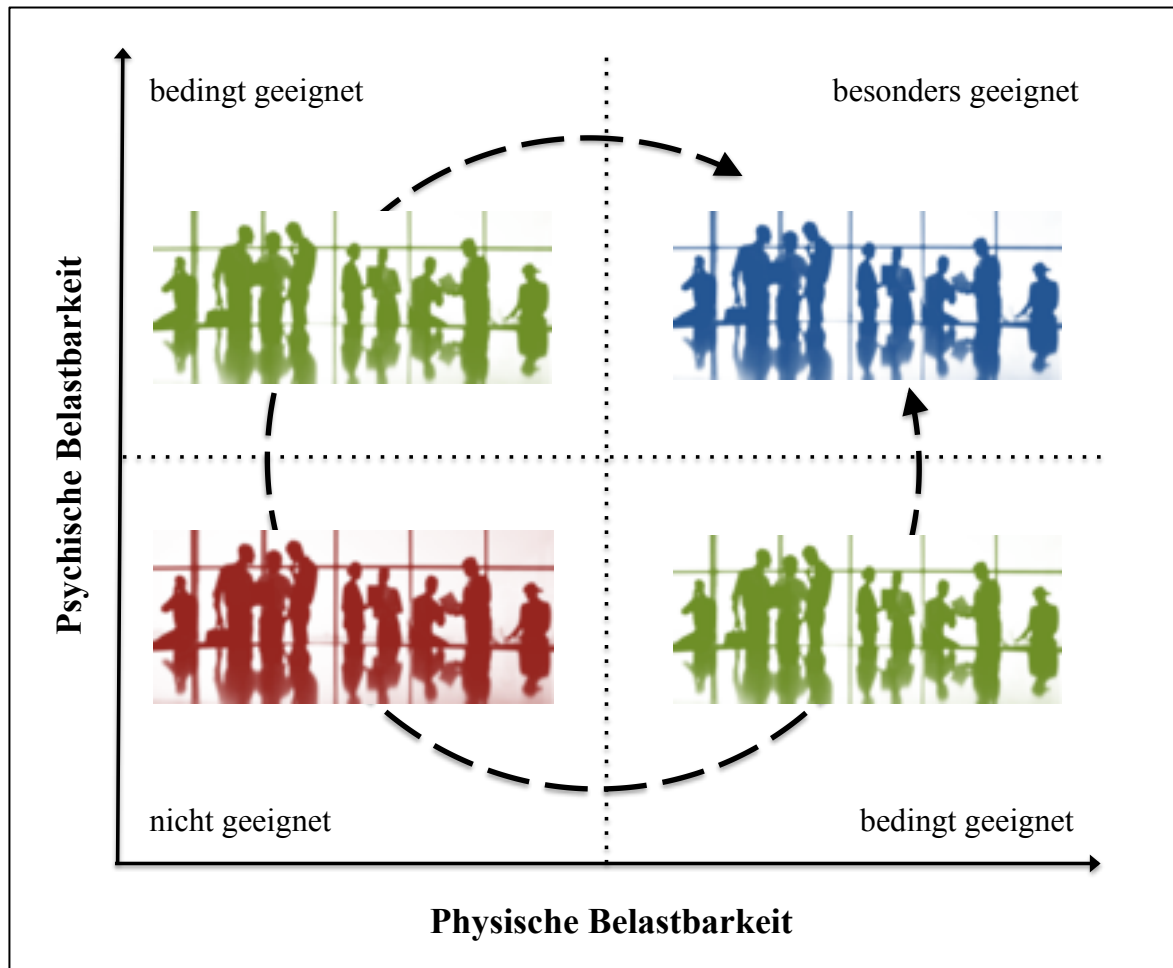


Abbildung 40. Berufsbedingte Eignung in Abhängigkeit des Ausprägungsgrades der psychophysischen Belastbarkeit.

Mit der Abbildung 40 soll darauf hingewiesen werden, dass ein wesentlicher Aspekt des Personalmanagements sich auf die Personalentwicklung bezieht und damit auch auf die Trainierbarkeit abhebt. Im Hinblick auf die psycho-physische Belastbarkeit bezieht sich dies hier speziell auf die Trainierbarkeit der physischen Komponente und der damit angenommenen und angestrebten psychischen Veränderung. Damit ist zugleich ein Ansatz und Anknüpfungspunkt für weiterführende Untersuchungen gegeben, wie es im folgenden Kapitel aufgegriffen ist.

7 SCHLUSSKOMMENTAR UND AUSBLICK

In diesem Kapitel soll es im Anschluss an die Diskussion und die damit verbundene Interpretation der Befunde darum gehen, die Ergebnisse der Untersuchungen im Abgleich mit den zu Beginn formulierten Anliegen, den theoretischen sowie methodischen Ansätzen sowohl hinsichtlich der pragmatischen als auch der wissenschaftlichen Relevanz zu reflektieren. Darüber hinaus wird in diesem Abschnitt ein Ausblick mit Überlegungen für nachfolgende Untersuchungen zu dieser Thematik gegeben, sowie Möglichkeiten aufgezeigt, dass hier entworfene Konzept weiter auszuarbeiten bzw. zu spezifizieren.

Eine aktuelle Beschäftigung mit dem psycho-physischen Zusammenhang ist insbesondere vor dem Hintergrund neuester gesellschaftlicher Entwicklungen der modernen Arbeitswelt und speziell einer Verschiebung in der Krankheitsstatistik von Herz-Kreislauf-Leiden hin zu Phänomenen wie bspw. Burn-Out, Depressionen und anderen psychischen Erkrankungen angezeigt. Inwieweit physische Aspekte der Belastbarkeit im Sinne der physischen Leistungsfähigkeit und Fitness auch psychische Faktoren bedingen und beeinflussen - und umgekehrt - wurde bereits im Rahmen sportwissenschaftlicher Fragestellungen versucht nachzugehen. Jedoch sind bisher keine wissenschaftlichen Arbeiten bekannt geworden, die sich dem psycho-physischen Zusammenhang hinsichtlich der Belastbarkeit im Kontext der Personalauswahl widmen. Mit dieser Arbeit soll neben der Analyse psycho-physischer Zusammenhänge und gruppenbezogene Unterschiede eine Möglichkeit eines Assessments in Form eines Screenings der psycho-physischen Belastbarkeit aufgezeigt werden und ferner auch nicht-militärischen Bereichen, wie bspw. der Wirtschaft und Industrie, interessante Hinweise geliefert werden, die in Zukunft dazu führen können, dass auch physische Aspekte der Belastbarkeit im Kontext der Personalauswahl Berücksichtigung finden.

Bei Personalentscheidungen ist den Verantwortlichen meistens daran gelegen viele Informationen über den Bewerber herauszufiltern und ein umfangreiches Bild über die Person zu erhalten, sodass eine möglichst genaue Vorhersage über die zukünftige Performance getroffen werden kann und um den bestmöglichen Bewerber auszuwählen. Diese ist auch je nach Beruf und Branche abhängig von der Belastbarkeit des Angestellten. Hierzu gehört u. a. die Fähigkeit, mit Belastungen und Stresssituationen umgehen zu können und emotional stabil zu bleiben, um auch gegen die Anfälligkeit der oben genannten Krankheiten besser gewappnet zu sein (siehe z. B. Judge et al., 2001; Judge & Bono, 2009; Bruk-Lee et al., 2009; Bamberg &

Keller, 2013). Hierbei spielen gesundheitliche Aspekte, insbesondere körperliche Aspekte, wie die physische Fitness eine bedeutende Rolle. „Der Arbeitgeber hat ein berechtigtes Interesse daran, dass der Bewerber die für den zu besetzenden Arbeitsplatz erforderliche körperliche Eignung mitbringt“ (Franken & Melzer, 2007, S. 86). Auch aus rechtlicher Perspektive besteht die Berechtigung, im Rahmen der Eignungsbeurteilung die körperlichen Voraussetzungen festzustellen, insofern es um die Ermittlung arbeitsbezogener Anforderungen geht. Allerdings besteht, wie dies auch bei psychologischen Eignungstests der Fall ist, keine Pflicht des Bewerbers, sich untersuchen zu lassen. Dies kann jedoch durchaus unter der Einwilligung der sich bewerbenden Person erfolgen (siehe Franken & Melzer, 2007). Mit dieser Arbeit wurden dazu Möglichkeiten aufgezeigt und Grundlagen hinsichtlich eines Screenings der psycho-physischen Belastbarkeit entwickelt.

Dazu wurde der psycho-physische Zusammenhang mit Bezug auf die Gesundheits- und Leistungsthematik diskutiert und im Rahmen der Belastungs-Beanspruchungsforschung spezifiziert. Das Konstrukt der psycho-physischen Belastbarkeit wird in diesem Rahmen konzeptualisiert und eingeordnet. Bei der Operationalisierung wird ein weiteres, in diesem Fall spezifisch: praktisch-diagnostisches Anliegen aufgegriffen, nämlich zeitökonomische Verfahren bzw. die zeitliche Ökonomisierung von Verfahren zu prüfen.

In dieser Arbeit konnten zum einen signifikante Zusammenhänge anhand der hier angewendeten Testverfahren zwischen physischen und psychischen Aspekten der (berufsbezogenen) Belastbarkeit herausgefunden und zum anderen teilweise deutliche Differenzen zwischen Gruppen mit unterschiedlichem physischen Leistungsniveau im Hinblick auf die psychische Belastbarkeit aufgedeckt werden. Demzufolge leistet diese Arbeit einen Beitrag, womit eine Möglichkeit aufgezeigt wird, psycho-physische Zusammenhänge sowie Unterschiede hinsichtlich der psychischen Belastbarkeit zwischen verschiedenen Leistungsgruppen im Kontext der Personalauswahl zu überprüfen. Daneben sollte die Relevanz dieser Thematik vor dem Hintergrund der erläuterten Problemstellung und aktueller Entwicklungen herausgestellt werden, um nebenbei auch Anregungen für Veränderungen der Strukturen von Personalauswahlprozessen zu geben. Pauschale Aussagen darüber, dass fitte(re) Menschen auch allgemein psychisch stabil(er) und belastbar(er) sind oder generell wenig(er) anfällig für Burn-Out-Syndrome o. ä. psychische Erkrankungen sind, können durch die hier gewonnen Erkenntnisse nicht getroffen werden. Es bleibt weiterhin offen und zu klären, mit welchen Verfahren die psycho-physischen Zusammenhänge hinsichtlich der (berufsbezogenen) Belastbarkeit nach-

weisbar gemacht und am besten erfasst werden können. Außerdem sollten sich zukünftige Studien mit den psycho-physischen Belastbarkeitszusammenhängen bestimmter Berufe und deren Anforderungen beschäftigen, sodass in Zukunft evtl. auch Kennwerte oder sogar Normwerte für die psycho-physische Belastbarkeit definiert werden können, die Orientierungshilfen für die verantwortlichen Personen und Entscheidungsträger im Personalauswahlprozess liefern.

Aus methodischen Gesichtspunkten stellen die Kennwerte für die physische Leistungsfähigkeit VO_{2max} bzw. VO_{2peak} , die Wattleistung sowie die Testdauer und auf der psychischen Seite die Itemscores aus den Persönlichkeitsskalen geeignete Parameter dar, die psycho-physischen Zusammenhänge messbar zu machen. Jedoch könnten in weiteren Untersuchungen auch zusätzliche bzw. andere Indikatoren sowohl für die physische, als auch für die psychische Belastbarkeit herangezogen werden. Für die Überprüfung der Ausdauerleistungsfähigkeit könnten bspw. die Bestimmung der ventilatorischen Schwellen (VT 1; VT 2: respiratorischer Kompensationspunkt, RCP) zusätzlich herangezogen werden. Für die Erfassung der Persönlichkeitsmerkmale wäre denkbar, wie in Kapitel 2.3 erwähnt, weitere Skalen und/oder Inventare zur Selbsteinschätzung der psychischen Belastbarkeit einzusetzen. Außerdem könnte es sinnvoll sein, die Selbsteinschätzung mit Fremdbeurteilungen und -einschätzungen von geschultem Personal zu ergänzen. Somit wäre es u. U. möglich ein differenzierteres Profil über die psycho-physische Belastbarkeit einer Person zu erhalten. Dies würde allerdings aufgrund des Umfangs der Verfahren bzw. der Verfahrensverbände und -kombinationen einer, wie in dieser Arbeit verfolgten, zeiteffizienten Screeningmethode widersprechen und evtl. einen zeitlich und materiell akzeptablen Aufwand übersteigen. Wenn das Anliegen in einem Screening besteht, wäre zu prüfen ob andere Parameter den Zweck besser erfüllen als die gewählten. Dies zu klären bleibt weiteren Studien vorbehalten. Besteht das Anliegen aber darin ein differenzierteres Bild zu erhalten ist daran zu denken, die genannten Parameter zu berücksichtigen.

Aus wissenschaftlicher Perspektive stehen sowohl die beiden Verfahrensentwicklungen, als auch die ausgewählten Konstrukte und ihre Verbindung (psycho-physisch) und die methodische Verbindung (Leistungsdiagnostik und Selbsteinschätzung) im Fokus. Die Verknüpfung objektiver Leistungsdaten mit subjektiven Einschätzungen stellt zwar ein durchaus bekanntes aber vielleicht hier besonders relevantes Problem dar, welches u. a. auch unter Berücksichtigung der jeweils (hier notwendigerweise verschiedenen) gewählten Gerätekonfigurationen zu weniger stringenten Befunden führen könnte. Darauf ist in nachfolgenden Untersuchungen für

die weitere Aufklärung des thematischen Komplexes zu achten. Die Verbindung des hier konzipierten Testverfahrens zur Überprüfung der Ausdauerleistung mit dem Projekt PersG BFT und den seitens der Auftraggeber definierten Vorgaben führte zunächst zu folgender Frage: Ist es möglich durch einen maximal 10 Minuten dauernden Fahrradergometertest und den dafür bereitgestellten Ergometern die Ausdauerleistung adequat zu überprüfen?

Die größte Anforderung bestand nach Meinung des Autors darin, im Rahmen der Vorgaben bzw. den einzuhaltenden Kriterien der Bundeswehr ein adäquates Testverfahren zur Überprüfung der Ausdauerleistungsfähigkeit zu entwickeln. Neben den vorgegebenen Kriterien zur Verfahrenskonzeption (bspw. zeitlicher Rahmen der Überprüfung, Durchführung mit einem Fahrradergometer, möglichst gleiche Bedingungen unabhängig von soziodemographischen Unterschieden) bestand eine weitere Herausforderung darin, dass die anfangs zur Verfügung gestellten Ergometer die Umsetzung eines für die Bundeswehr geeigneten Testdesigns stark erschwerten. Für die Umsetzung des Testdesigns wurden wie in Kapitel 5.1 anstelle die von der Bundeswehr angeschafften Ergometer die Hometrainer der Firma Tacx in Kombination mit Rennrädern des Herstellers Cannondale für die weiteren Untersuchungen eingesetzt. Diese ermöglichten weitestgehend die Realisierung des Testdesigns, das im Rahmen der vorgegebenen Kriterien seitens der Auftraggeber konzipiert und anschließend evaluiert werden sollte.

Durch die angesprochenen Schwierigkeiten bei der Umsetzung des konzipierten Testdesigns mit den verfügbaren Gerätekonfigurationen ergab sich die Notwendigkeit, Modifikationen dieser Konfigurationen vorzunehmen und im Hinblick auf die spezifischen Anforderungen (Stichwort: „Strukturähnlichkeit“ des Fahrradergometertests mit dem Lauftest) auf Machbarkeit und Brauchbarkeit zu überprüfen. Dies führte im Verlaufe der Verfahrensentwicklungen dazu, dass Datenerhebungen durchzuführen waren, für die verschiedene „Gelegenheiten“ (Studien im weiteren Sinne) wahrgenommen werden mussten. Die Datensammlungen sind dann in der Arbeit als Studien mit verschiedenen eingesetzten Verfahren dargestellt (Abbildung 29), die anschließend in die Gesamtstichprobe eingegangen sind. Dementsprechend wurde nicht einem vorab entworfenen Untersuchungsplan und einer Untersuchungsstrategie für die vollständige Datenerhebung gefolgt. Bei solchen Untersuchungsanliegen (geräteunterstützte Verfahrensentwicklungen) ist daran zu denken, die mit den theoretischen und methodischen Anforderungen gekoppelten Ansprüche an die Erhebungsinstrumente (Ergometer) im Vorhinein auf ihre Realisierbarkeit hin zu prüfen. Dies wurde erst im Verlaufe der Untersuchungsdurchführung deutlich.

Für die Überprüfung der Minimalanforderungen hinsichtlich der Ausdauerleistungsfähigkeit konnte anhand der Laborstudie nachgewiesen werden, dass der 3000-m-FET durchaus ein valides und zufriedenstellendes Testverfahren darstellt die Ausdauerleistungsfähigkeit im Zuge des Personalauswahlprozesses zu screenen. Allerdings wird bei diesem Ergometertest überwiegend die Kraftausdauer gefordert und aufgrund der relativ kurzen (Mittelzeitausdauer: 3:30 - 6:30 Min.) und intensiven Belastungsdauer darüber hinaus die Laktattoleranz der Testperson zum leistungslimitierenden Faktor. Aus gesundheitlicher Sicht empfiehlt es sich daher im Rahmen der Personalgewinnung und vor dem Hintergrund der Absicht einer Feststellung der Basisfitness von einem zu kurzem Zeitrahmen für die Testung abzusehen. Man könnte so zu einem Belastungsprotokoll tendieren, das die Belastung langsamer steigert und nicht zu abrupt hohen und durchgängig intensiven Belastungen führt, die in derart hohen Laktatkonzentrationen im Blut bzw. zu starken Übersäuerungen resultieren.

Im Hinblick auf die Studie von Midgley et al. (2008), bei der überprüft wurde, inwiefern auch kürzere Belastungsprotokolle zur Erfassung der maximalen Sauerstoffaufnahme als geeignet erscheinen, konnten mit der hier durchgeführten Laborstudie weitere interessante Ergebnisse geliefert werden. Die Befunde zeigen, dass Ausbelastungstests zur Erfassung der VO_{2max} mehr als fünf Minuten dauern sollten, um die tatsächliche VO_{2max} festzustellen. Bei den Rampentests, die mit einer durchschnittlichen Belastungsdauer von 9:12 Min. (Fahrradergometer) und 9:42 (Laufband) deutlich länger waren als die Streckentests (3000-m-FET: $M = 4:23$ Min.; 1000-m-Lauf: $M = 4:01$ Min.) erreichten die Pbn i. d. R. auch höhere VO_{2max} -Werte. Sollten im Rahmen von zukünftigen Personalauswahlprozessen die körperliche Leistungsfähigkeit in Form der Ausdauer durch den Parameter VO_{2max} bestimmt werden, empfiehlt es sich zur Erfassung dieses Parameters Belastungsprotokolle zu wählen, die aufgrund der oben genannten Aspekte eine stetig ansteigende Belastung beinhalten und hinsichtlich der Testzeit in Abhängigkeit von Belastungsgerät mindestens fünf bzw. sieben Minuten dauern.

Auf Grundlage der durchgeführten Untersuchungen und den daraus gewonnenen Befunden erscheinen die entwickelten Verfahren als geeignet und es steht nun an, in Bezug auf bestimmte Personen bzw. Berufsgruppen Normwerte zu erstellen, um die getesteten Personen hinsichtlich der physischen und psychischen Belastbarkeit angemessen einordnen zu können. Hierbei wäre es u. U. indiziert, zusätzliche Verfahren zur Erfassung weiterer Facetten der psycho-physischen Belastbarkeit hinzuzuziehen und dies einerseits im Sinne eines differen-

zierten nomologischen Netzwerkes und andererseits eines konsistent angelegten Assessment Centers zu nutzen.

In einem nächsten Schritt könnten zu den Querschnittsuntersuchungen und den hier konzipierten Verfahren Längsschnittuntersuchungen und/oder Interventionsstudien zur psychophysischen Belastbarkeit durchgeführt werden, die darauf abzielen, eventuelle Veränderungen bzw. Entwicklungen der physischen Fitness (unabhängige Variable) auf die psychische Belastbarkeit (abhängige Variable) zu evaluieren und daran anschließend Konsequenzen begründet ableiten zu können. Auch bei der Konzeption spezieller Programme und Maßnahmen, bspw. im Sinne der betrieblichen Gesundheitsförderung, könnten solche Verfahren eingesetzt werden und in Form eines Monitorings die Fitness bzw. körperliche Leistungsfähigkeit der Arbeitnehmer systematisch zu verfolgen und Veränderungen der psycho-physischen Belastbarkeit der Belegschaft zu dokumentieren. In diesem Kontext wäre es interessant, inwieweit sich durch eine Sicherung des erreichten Fitnessniveaus oder sogar durch eine gesteigerte physische Leistungsfähigkeit auch psychische Aspekte der Belastbarkeit stabilisieren oder (positiv) verändern. Darüber hinaus könnte das hier ausgearbeitete Konzept und die aufgezeigte Möglichkeit, psycho-physische Zusammenhänge zu operationalisieren, auch im Bereich der Talentauswahl Anwendung finden. Speziell in Sportarten, bei denen psychische Aspekte eine starke Rolle spielen, wie dies bspw. beim Golfsport oder Motorsport der Fall ist und wobei physische Fähigkeiten im Vergleich zu körperbetonten Sportarten eine geringer anzusetzende Erfolgskomponente darstellen, wäre ebenfalls interessant, inwieweit sich physisch fittere (ausdauerleistungsfähigere) Nachwuchssportler hinsichtlich psychischer (Belastbarkeits-) Aspekte von weniger fitten Talenten in diesen Sportarten unterscheiden. Auch vor dem Hintergrund einer angestrebten leistungssportlichen Karriere, die mit Niederlagen, Verletzungen und vergleichbaren Rückschlägen verbunden ist, erscheinen Aspekte der psychischen (insbesondere emotionalen) Belastbarkeit von essentieller Bedeutung für den Werdegang im (Hoch-) Leistungssport. Insofern wäre auch eine Übertragung des hier vorgestellten Konzepts mit entsprechender Spezifikation der Verfahren im Rahmen der Talentauswahl im Leistungssport ein denkbarer und überprüfenswerter Ansatz.

Ergänzend zu den Verfahrensentwicklungen und der Überprüfung von psycho-physischen Zusammenhängen könnten Experten befragt werden, inwieweit solche Aspekte angemessen berücksichtigt und Überprüfungen der psycho-physischen Belastbarkeit im Rahmen der Personalauswahl zweckmäßig platziert werden können und ob sich solche Maßnahmen im Rah-

men von Personalauswahlverfahren in Wirtschaft und Industrie eignen und langfristig bewähren. Des Weiteren lässt sich an bedarfsorientierte Spezifikationen des hier ausgearbeiteten Konzepts etwa im Hinblick auf die Bezugskonzepte und Konstrukte bzw. die Differenzierung des nomologischen Netzwerks in Verbindung mit den entsprechenden Operationalisierungen v. a. im Kontext der Personalauswahl denken, die sich in ähnlichen aber anders akzentuierten Kontexten als nützlich erweisen könnten, um anforderungsbezogene Empfehlungen für die Entscheidungsträger bei der Personalauswahl und -entwicklung zu geben bzw. diese daraufhin zu beraten.

Mit Blick auf die Gütekriterien ließen sich zur Erweiterung der Validitätstudien, insbesondere der Untersuchung zur konvergenten und divergenten Konstruktvalidität, auch andere Konstrukte heranziehen und darüber das Konzept weiter differenzieren. Der Beitrag in dieser Arbeit stellt lediglich einen ersten Schritt dar. Auch wenn die Überprüfung der Retest Reliabilität gute Ergebnisse zeigte, wären weitere Studien mit verschiedenen Zeiträumen für eine weitere Absicherung zweckmäßig, um auch Aufschluss über die Stabilität des Verfahrens über große Zeiträume zu gewinnen.

Hinsichtlich der erkenntnistheoretischen Relevanz wäre es in Bezug auf die hier hauptsächlich verwendeten Persönlichkeitsskalen, die sich auf spezifische Subskalen eines (berufsbezogenen) Persönlichkeitsinventars beziehen, in weiteren Untersuchungen aufschlussreich, alle (z. B. die hier zur Prüfung der Konstruktvalidität eingesetzten) Verfahren einer gemeinsamen Faktorenanalyse zu unterziehen, um über die faktorielle Struktur die Ordnung der Konstrukte (gleiche Ebene, Facetten etc.) zu prüfen und so auch empirisch gewonnene Hinweise zur Konzeptentwicklung zu erhalten. Dies wäre im Hinblick auf die auch hier aufgegriffenen Konstrukte der mental toughness, hardiness und resilience oder Resilienz, auf die z. B. in der Sportpsychologie häufig rekurriert wird, überaus sinnvoll und sicher nützlich.

8 ZUSAMMENFASSUNG

Mit der vorliegenden Arbeit wurden drei zentrale Anliegen angegangen. Diese bestehen zum Ersten darin, die physische Belastbarkeit für die Personalgewinnung für spezifische Handlungsfelder zu konzipieren und zu operationalisieren und dabei besonders die Kriterien der (Zeit-) Ökonomie und Standardisierung zu berücksichtigen. Zum Zweiten wurde das Verfahren zur Messung der physischen Belastbarkeit mit der psychischen Belastbarkeit in Verbindung gebracht. Die Operationalisierung der psychischen Belastbarkeit ist dann mit dem Verfahren zur Testung der physischen Belastbarkeit in einen Verfahrensverbund gestellt worden, sodass ein Assessment psycho-physischer Belastbarkeit in der Praxis der Personalgewinnung - hier speziell der Bundeswehr - ermöglicht werden konnte. Drittens ging es darum, mit diesen Instrumenten die spezifischen psycho-physischen Zusammenhänge zu prüfen und der Frage nachzugehen, ob physisch (ausdauer-) leistungsfähige(re) Personen auch eine ausgeprägte(-re) psychische Belastbarkeit aufweisen.

Nachdem auf die historischen Anknüpfungspunkte und die Relevanz dieser Arbeit für die Personalauswahl, modellhaft aufgezeigt an der Bundeswehr, hingewiesen wurde, ist die Entwicklung eines theoretischen Rahmenkonzepts dargestellt worden. Hierzu wurde das Belastungs-Beanspruchungskonzept herangezogen und das Konstrukt der psycho-physischen Belastbarkeit integriert. Dazu ergänzend ist im Stand der Forschung auf Untersuchungen zur Persönlichkeit, Stress, sowie psychische Gesundheit in Verbindung mit Sport bzw. physische Fitness und körperliche Aktivität Bezug genommen worden. Vor dem Hintergrund der Zunahme psychischer Erkrankungen, insbesondere Depression und Burn-Out, wurden Beiträge zum Thema Stress und Emotionale Stabilität bzw. Neurotizismus gesichtet und auf die hier wesentlichen Konzepte eingegangen.

Von grundlegender Bedeutung für die Arbeit war es, die psychische und physische Belastbarkeit konzeptuell in einen systematischen Zusammenhang einzuordnen und zu operationalisieren, sodass eventuelle Zusammenhänge dieser beiden Komponenten mit Hilfe der hier entwickelten Verfahren festgestellt werden können. Diese Verfahren wurden nach der klassischen Testtheorie und den entsprechenden Testgütekriterien konzipiert. Der Anspruch besteht darin, die psycho-physischen Zusammenhänge in Form eines Screenings erfassen zu können und dementsprechend die Konzeptualisierung der Verfahren nach zeitökonomischen Gesichtspunkten umzusetzen. Dafür wurde zur Überprüfung der physischen Belastbarkeit ein

Verfahren entwickelt, das v. a. mit Bezug auf die zeiteffiziente Durchführbarkeit, die Standardisierung und die Justiziabilität speziell für die Erfassung einer Basisfitness angelegt ist.

In dem dann erläuterten methodisch-empirischen Ansatz für die Verfahrensentwicklungen zur Erfassung der physischen und psychischen Belastbarkeitskomponente stand zum einen die Ausdauerleistungsfähigkeit als physische Komponente der Belastbarkeit und zum anderen die psychische Komponente in Form der beiden Persönlichkeitskonstrukte Emotionale Stabilität und Belastbarkeit im Fokus. In Verbindung mit der Konzeptualisierung wurde die Verfahrensentwicklung zur psycho-physischen Belastbarkeit seitens der psychischen Belastbarkeitskomponente mit den Skalen Emotionale Stabilität und Belastbarkeit des Bochumer Inventars zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) realisiert, die dafür zu einer Skala „psychische Belastbarkeit“ kombiniert wurden. Für die Umsetzung der Konzeptualisierung der physischen Belastbarkeitskomponente wurde auf die Ausdauerleistungsfähigkeit als Teilkomponente der konditionellen Fähigkeiten zurückgegriffen und diese durch einen hier eigens entwickelten Fahrradergometertest operationalisiert. Mit dem Ziel der zeitökonomischen Erfassung der psychischen Belastbarkeit wurde auf Grundlage der beiden zusammengeführten Skalen des BIP eine Kurzskala entworfen, die im Rahmen der Datenerhebungen zur psycho-physischen Belastbarkeit parallel zur Langform-Skala zum Einsatz kam. Die Ausdauerleistungsfähigkeit sowie die Kombination der Faktoren Emotionale Stabilität und Belastbarkeit repräsentieren die beiden Komponenten des Konstrukts der psycho-physischen Belastbarkeit. Mit dem Fahrradergometertest einerseits und der Skalenkombination (Emotionale Stabilität mit Belastbarkeit) andererseits wurde das Konstrukt der psycho-physischen Belastbarkeit operationalisiert. Dieser Verfahrensverbund stellte die Grundlage zur Untersuchung psychophysischer Zusammenhänge bzw. zur Analyse von Vergleichen unterschiedlicher Leistungsgruppen bzw. Gruppen mit unterschiedlichem Fitnessniveau dar.

Aus den Studien zu den Verfahrensentwicklungen, an denen insgesamt 424 Personen teilnahmen, ist folgendes herauszustellen: Der konzipierte Fahrradergometertest (3000-m-FET) erfüllt hinsichtlich der Reliabilität und Validität die Gütekriterien. Hierzu wurde speziell eine Laborstudie mit dem Einsatz spiroergometrischer Verfahren durchgeführt und der 3000-m-FET mit anderen Ausdauertests verglichen. Dazu wurden mit anderen Testverfahren zur Feststellung der Ausdauerleistungsfähigkeit zusätzliche Leistungsindikatoren erfasst und diese dann auch in den Untersuchungen zu Zusammenhängen und Gruppenunterschieden berücksichtigt. Die Überprüfung der Skalen zur psychischen Belastbarkeitserfassung ergab, dass (a)

die Skalenkombination in einer Langform mit einer neu-konzipierten Kurzform mit einem Koeffizienten von $r = .60$ höchst signifikant ($p \leq .000$) korreliert. (b) Im Hinblick auf die konvergente Validitätsüberprüfung mit Hilfe der Skala zur allgemeinen Selbstwirksamkeit ist ein höchst signifikanter ($p \leq .000$) positiver Zusammenhang ($r = .34$ für die Kurzform und $r = .55$ für die Langform) zu verzeichnen. (c) Bei der Überprüfung der divergenten Validität zeigte sich ein ebenfalls höchst signifikanter ($p \leq .000$) nun allerdings theorieentsprechend negativer Zusammenhang von $r = -.58$ für die Kurzform und $r = -.62$ für die Langform. Die Itemscores dieser Skalen wurden ebenfalls für die psycho-physischen Zusammenhängeüberprüfungen herangezogen.

Die Analysen zu psycho-physischen Zusammenhängen zeigten Korrelationen zwischen der benötigten Zeit beim 3000-m-FET und den Scores der Kurz- sowie Langform-Skalen der psychischen Belastbarkeit von $r = -.20$ für die Kurzform und $r = -.27$ für die Langform (beide höchst signifikant mit $p \leq .000$). Das negative Vorzeichen verdeutlicht, dass fitte Personen hohe Werte in der psychischen Belastbarkeit aufweisen. Zur weiteren Analyse wurden (Extrem-) Gruppenvergleiche mit Hilfe von Varianzanalysen vorgenommen. Dabei zeigten sich fittere Personen gegenüber Personen mit geringerer ausgeprägten Fitness in Abhängigkeit von der Gruppenkomposition (Kriterium bzw. Parameter der Aufteilung und Größe) als mehr oder weniger deutlich psychisch belastbarer.

Mit dem entwickelten Instrumentarium zur Überprüfung der psycho-physischen Belastbarkeit kann ein geeignetes Hilfsmittel für die Personalauswahl, hier speziell und modellhaft aufgezeigt anhand des Anliegens der Personalgewinnung in der Bundeswehr, zur Verfügung gestellt werden. Im Hinblick auf die Annahme eines psycho-physischen Zusammenhangs kann aufgrund der empirischen Analysen speziell bezogen auf die hier betrachteten Aspekte und mit den eingesetzten Verfahren zwar eine signifikante aber nur gering ausgeprägte Korrelation festgestellt werden. Allerdings lässt sich aufgrund der vorgenommenen Gruppenvergleiche von mehr und weniger fitten Personen schlussfolgern, dass fittere Personen sich als psychisch belastbarer einschätzen. Um die psycho-physische Belastbarkeit detailliert(er) und präzise(r) zu erfassen, ist der Rückgriff auf andere Ausdauer-testverfahren oder die Messung von weiteren physischen Grundlagen, sowie die Messung zusätzlicher Persönlichkeitsmerkmale, die ähnliche Konstrukte im Sinne der Belastbarkeit erfassen, ratsam und angezeigt. Je umfassender, detaillierter oder spezifischer ein Instrument ein zu messendes Merkmal misst, desto genauer können anschließend Aussagen über die (in diesem Fall: psycho-physischen)

Zusammenhänge getroffen werden. Auf der Seite der Persönlichkeitsmerkmale ist dafür ggf. auch neben der Selbsteinschätzung die Fremdbeurteilung mit einzubeziehen, um u. a. auch Antworttendenzen oder der Tendenz sozial erwünscht zu antworten (social desirability), wie sie in Fragebogenverfahren und Instrumenten der Selbsteinschätzung immer wieder auftreten können, zu begegnen.

Zur Orientierung für die Praxis der Personalauswahl lassen sich anhand der Befunde signifikante Zusammenhänge zwischen der so bestimmten physischen und psychischen Belastbarkeit konstatieren. Inwiefern man in Zukunft tatsächlich auch physische Aspekte im Rahmen der Personalauswahl (stärker) berücksichtigen sollte, ist sicher abhängig vom jeweiligen Anliegen, den spezifischen Anforderungen und vom gegebenen Kontext. Die hier erarbeiteten Hintergründe und dargestellten Befunde lassen sich dann aufgreifen, wenn sich solche Aspekte für spezifische Bereiche auch über den militärischen Tätigkeitsbereich hinaus als wichtig oder relevant erweisen. Dass Herangehensweisen in der Personalauswahl aus dem militärischen in den wirtschaftlichen Bereich übernommen wurden, hat sich in der Historie des Assessment Centers bereits als nützlich erwiesen. Für die Zukunft ergeben sich dahingehende Hinweise, dass den physischen Faktoren, wie der körperlichen Fitness und Leistungsfähigkeit, die bisher - wenn überhaupt - hinsichtlich der betrieblichen Gesundheitsförderung in Unternehmen eine Rolle spielten, bereits im Rahmen von Personalauswahlprozessen stärkere Beachtung finden und ihnen eine größere Relevanz beigemessen werden kann.

9 LITERATURVERZEICHNIS

- American College of Sports Medicine (2006). *ACSM's guidelines for exercise testing and prescription* (6th ed.). Baltimore, MD: Lippincott, Williams & Wilkins.
- American College of Sports Medicine Position Stand (1998). The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine and Science in Sports and Exercise*, 30, 975-991.
- Allen, M. S., Greenless, I. & Jones, M. (2013). Personality in sport: a comprehensive review. In *International Review of Sport and Exercise Psychology*, 6 (1), 184-208.
- Allmer, H. (1996). *Erholung und Gesundheit*. Göttingen: Hogrefe.
- Alpers, G. W., Mühlberger, A. & Pauli, P. (2009). Psychophysiologie der Emotion. In V. Brandstätter & J. H. Otto (Hrsg.), *Handbuch der Allgemeinen Psychologie. Motivation und Emotion* (S. 412-421). Göttingen: Hogrefe.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2003). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2008). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze. *Psychologische Rundschau*, 41, 1-24.
- Baltes, P. B. & Eckensberger, L. H. (Hrsg.). (1979). *Entwicklungspsychologie der Lebensspanne*. Stuttgart: Kohlhammer.
- Bamber, E. & Keller, M. (2013). Stressbewältigung. In W. Sarges (Hrsg.), *Management-Diagnostik* (S. 278-285). Göttingen: Hogrefe.
- Bamberg, E. (2007). Belastung, Beanspruchung, Stress. In H. Schuler & K. Sonntag, (Hrsg.), *Handbuch der Arbeits- und Organisationspsychologie* (S. 141-148). Göttingen: Hogrefe.
- Barthel, E. (1989). *Nutzen eignungsdiagnostischer Verfahren bei der Bewerberauswahl*. Frankfurt am Main: Peter Lang.
- Barrick, M. R. & Mount, M. K. (1991). The Big-Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Berger, G., Pargman, D. & Weinberg, R. S. (2007). *Foundations of Exercise Psychology*. Morgantown, WV: Fitness Information Technology.
- Berger, B. G. (1994). Coping with stress: The effectiveness of exercise and other techniques. *Quest*, 46 (1), 100-119.

- Bloss, H. A. (1992). *Fitness-Lexikon* (2. Aufl.). Düsseldorf: ECON-Taschenbuchverlag.
- Borg, G. (1970). Perceived exertion as an indicator of somatic stress. *Scandinavian Journal of Rehabilitation Medicine*, 2, 92-98.
- Borg, G. (1978). Psychological assessment of physical effort. In U. Simri (Hrsg.), *Proceedings of the 1978 International Symposium on Psychological Assessment in Sport* (S. 49-57). Netanya, Israel: Wingate Institute.
- Borg, G. (1961). Perceived exertion in relation to physical work load and pulse-rate. *Kungliga Fysiografiska Sällskapet I Lund Forhandlingar*, 31, 105-115.
- Borg, G. & Dahlstrom, H. (1962). A pilot study of perceived exertion and physical working capacity. *Acta Soc. Med. Upsal.*, 67, 21-27.
- Borg, G. & Lindblad, I. (1976). The determination of subjective intensities in verbal descriptions of symptoms. *Internal Report from the Institute of Applied Psychology*, 75, University of Stockholm.
- Bortz, J. (1989). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bös, K. (1998). Fitnessdiagnose. In K. Bös & W. Brehm (Hrsg.), *Gesundheitssport. Ein Handbuch* (S. 188-198). Schorndorf: Hofmann.
- Brähler, E., Hintz, A. & Scheer, J. W. (2008). *GBB – 24-Der Gießener Beschwerdebogen*. Bern: Huber.
- Breucker, G. (2013). Europa investiert in die psychische Gesundheit am Arbeitsplatz. Gesund im Job. *Das Magazin für Betriebliches Gesundheitsmanagement*, 6, 4-5.
- Brown, J. D. (1990). Exercise, fitness, and mental health. In C. Bouchard, R. J., Shepard, T., Stephens, J. R., Sutton & B. D. McPherson (Hrsg.), *Exercise, fitness and health* (S. 607-626). Champaign, IL: Human Kinetics.
- Bruk-Lee, V., Khoury, H. A., Nixon, A. E., Goh, A. & Spector, P. E. (2009). Replicating and extending past personality/job satisfaction meta-analyses. *Human Performance*, 22, 156-189.
- Buchfuhrer, M. J., Hansen, J. E., Robinson, T. E., Sue, D. Y., Wasserman, K. & Whipp, B. J. (1983). Optimizing the exercise protocol for cardiopulmonary assessment. In *Journal of Applied Physiology*, 55 (5), 1558-1564.
- Bundeszentrale für politische Bildung (2011). Herausforderungen für die Personalgewinnung der Bundeswehr. Online: <http://www.bpb.de/apuz/59657/herausforderungen-fuer-die-personalgewinnung-der-bundeswehr?p=1> (Zugriff am 12.02.2013).
- Bundeswehr (2011). Online: <http://www.deutschesheer.de/portal/a/heer/> (Zugriff am 16.11.2013)
- Bunge, M. (1980). *The mind-body problem: A psychobiological approach*. New York: Pergamon Press.

- Busch, C. (2004). *Stressmanagement für Teams. Entwicklung und Evaluation eines Trainings im Call Center*. Hamburg: Dr. Kovac.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cattell, R. B. (1946). *The description and measurement of personality*. New York: World Book.
- Cattell, R. B., Cattell A. K. & Cattell, H. E. P. (1993). *16PF Fifth Edition Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Clarke, H. H. (1976). *Application of Measurement to Health and Physical Education*. Englewood Cliffs: Prentice Hall.
- Clayton, R. P. (1991). Stress reactivity. Hemodynamic adjustments in trained and untrained humans. *Medicine and Science in Sports and Exercise*, 23, 873-881.
- Cohen, S., Karmack, T. & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour*, 24 (4), 385-396.
- Conrad, W. (ohne Jahresangabe). *Untersuchungen über Persönlichkeitsunterschiede in Abhängigkeit von Niveau und Art des sportlichen Leistungsvermögens*. Manuskript.
- Costa, P. T., Jr., McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, Florida: Psychological Assessment Resources.
- Costa, P. T. & McCrae, R. R. (1989). *The NEO PI/FFI manual supplement*. Odessa, Florida: Psychological Assessment Resources.
- Costa, P.T., McCrae, R. R. & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness - a revision of the NEO personality-inventory. *Personality and Individual Differences*, 12, 887-898.
- Costa, P. T. & McCrae, R. R. (1992). *The NEO Personality Inventory (revised) manual*. Odessa, FL: Psychological Assessment Resources.
- Crust, L. & Clough, P. J. (2011). Developing mental toughness: From research to practice. *Journal of Sport Psychology in Action*, 2, 21-32.
- DAK Gesundheitsreport (2012). Online:
[http://www.presse.dak.de/ps.nsf/Show/7BDD14663C6ABDEAC125799D00476AB3/\\$File/120214_Krankmeldungen%20erreichen%202011%20hoechsten%20Stand%20seit%2015%20Jahren_I.pdf](http://www.presse.dak.de/ps.nsf/Show/7BDD14663C6ABDEAC125799D00476AB3/$File/120214_Krankmeldungen%20erreichen%202011%20hoechsten%20Stand%20seit%2015%20Jahren_I.pdf) (Zugriff am 11.03.2013).
- DAK Gesundheitsreport (2013). *Analyse der Arbeitsunfähigkeitsdaten. Update psychische Erkrankungen - Sind wir heute anders Krank?* Online: http://www.bptk.de/fileadmin/user_upload/News/BPtk/2013/20130228/20130228_DAK-Gesundheitsreport_2013.pdf (Zugriff am 08.05.2014).

- De Maizière, T. (2012). Rede des Verteidigungsministers anlässlich des Beförderungs- und Abschlussappells an der Offiziersschule des Heeres. Online: <http://www.bmvg.de/portal/a/bmvg/> (Zugriff am 10. Oktober 2013)
- De Marées, H. (2002). *Sportphysiologie*. Köln: Sport & Buch Strauß.
- Dickhut, H.-H. (2000). *Einführung in die Sport- und Leistungsmedizin*. Schorndorf: Hofmann.
- Dienstbier, R. A. (1991). Behavioral correlates of sympathoadrenal reactivity: The toughness model. *Medicine and Science in Sports and Exercise*, 23, 846-852.
- Dieter, U. L. (2004). *Sportmedizinisches Leistungsprofil von Handballspielerinnen in der nationalen und internationalen Spitzenklasse*. Unveröff. Diss., Justus-Liebig-Universität Gießen.
- Dimsdale, A. (1968). Two personality dimensions of a small sample of British athletes. *Bulletin of the British Psychology*, 21, 171-172.
- Doyne, E. J., Chambless, D. L. & Beutler, L. E. (1983). Aerobic exercise as treatment for depression in women. *Behavioural Therapy*, 14, 434-440.
- Drenth, P. J. D., Therry, H. & Wolff de, C. J. (1998). *Handbook of work and organizational psychology* (Band 3). Hove: Psychology Press.
- DuBois, P. (1970). *A History of psychological Testing*. Boston: Allyn and Bacon.
- Dustman, R. E., Ruhling, R. O., Russell, E. M., Shearer, D. E., Bonekat, W., Shigeoka, J. W., Wood, J. S., & Bradford, D. C. (1984). Aerobic exercise training and improved neuropsychological function of older individuals. *Neurobiology of Aging*, 5, 35-42.
- Egloff, B. & Gruhn, A. J. (1998). Gründe für und Veränderung an Ausdauersport: Zur Rolle von Persönlichkeitsvariablen. *Zeitschrift für Sportpsychologie*, 2, 46-55.
- Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.
- Eysenck, H. J. (1952). *The scientific study of personality*. London: Routledge (& Kegan Paul).
- Eysenck, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.
- Eysenck, H. J. & Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.
- Eysenck, H. J. & Eysenck S. B. (1964). *The Eysenck Personality Inventory*. London: University of London Press.
- Eysenck, H. J., Nias, D. K. B. & Cox, D. N. (1982). Sport and personality. In *Advances in Behaviour Research and Therapy*, 4, 1-56.

- Fahrenberg, J. & Selg, H. (1970). *Das Freiburger Persönlichkeitsinventar*. Göttingen: Hogrefe.
- Faller, M. (1993). *Innere Kündigung. Ursachen und Folgen*. München: Hampp.
- Fay, E. (2002). *Das Assessment-Center in der Praxis. Konzepte-Erfahrungen-Innovationen*. Göttingen: Vandenhoeck & Ruprecht.
- Fisseni, H.-J. & Preusser, I. (2007). *Assessment-Center. Eine Einführung in Theorie und Praxis*. Göttingen: Hogrefe.
- Folkins, C. H. & Sime, W. E. (1981). Physical fitness training and mental health. *American Psychologist*, 36, 373-389.
- Fuchs, R. (1997). *Psychologie und körperliche Bewegung*. Göttingen: Hogrefe.
- Fuchs, R. (2003). *Sport, Gesundheit und Public Health*. Göttingen: Hogrefe.
- Fünten, K. aus der, Faude, O., Skorski, S. & Meyer, T. (2013). Sportmedizin. In A. Güllich & M. Krüger (Hrsg.), *Sport. Das Lehrbuch für das Sportstudium* (S. 171-180). Springer: Berlin.
- Franken, T. & Marc, M. (2007). Rechtliche Grundlagen der Eignungsbeurteilung. In M. John & G. W. Maier (Hrsg.), *Eignungsdiagnostik in der Personalarbeit. Grundlagen, Methoden, Erfahrungen* (S. 73-102). Düsseldorf: Symposium.
- French, J. R. P., Rodgers, W. & Cobb, S. (1974). Adjustment as person-environment fit. In G. V. Coelho, D. A. Hamburg & J. E. Adams (Hrsg.), *Coping and adaptation* (S. 316-333). New York: Basic Books.
- Freudenberger, H. J. (1980). *Burnout*. New York: Doubleday.
- Frey, D. Kaminski, S., Streicher, B. & Niesta, D. (2007). Leistungsförderung. In H. Schuler & K. Sonntag (Hrsg.), *Handbuch der Arbeits- und Organisationspsychologie* (S. 633-640). Göttingen: Hogrefe.
- Fröhner, G. (1993). *Die Belastbarkeit als zentrale Größe im Nachwuchstraining*. (Trainerbibliothek 30). Münster: Philippka.
- Guthke, J. (1991). Das Lerntestkonzept in der Eignungsdiagnostik. In H. Schuler & U. Funke (Hrsg.), *Eignungsdiagnostik in Forschung und Praxis* (S. 33-35). Stuttgart: Verlag für Angewandte Psychologie.
- Graf, C., Hartmann, U., Platen, P., Rost, R. & Schänzer, W. (2001). Biologische Grundlagen. In R. Rost (Hrsg.), *Lehrbuch der Sportmedizin* (S. 25-154). Köln: Deutscher Ärzte Verlag.
- Greif, S. (1991). Streß in der Arbeit – Einführung und Grundbegriffe. In S. Greif, N. Semmer & E. Bamberg (Hrsg.), *Psychischer Streß am Arbeitsplatz* (S. 1-28). Göttingen: Hogrefe.

- Grosser, M., Starischka, S., & Zimmermann, E. (2008). *Das neue Konditionstraining*. München: BLV.
- Haber, P. (2007). *Lungenfunktion und Spiroergometrie*. Springer: Wien.
- Hacker, W. (1978). *Allgemeine Arbeits- und Ingenieurpsychologie*. Bern: Huber.
- Hagberg, J. M., Mullin, J. P., Bahrke, M. & Limburg, J. (1979). Physiological profiles and selected psychological characteristics of national class American cyclists. *Journal of Sports Medicine*, 19, 342-346.
- Hathaway, S. R. & McKinley, J. C. (1943). *Manual for the Minnesota Multiphasic Personality Inventory*. New York: Psychological Corporation.
- Heck, H. (1990). *Energiestoffwechsel und medizinische Leistungsdiagnostik*. Schorndorf: Hofmann
- Heck, H. & Schulz, H. (1998). Gütekriterien in der sportmedizinischen Leistungsdiagnostik. In D. Jeschke & R. Lorenz (Hrsg.), *Sportartspezifische Leistungsdiagnostik. Energetische Aspekte* (S. 13-26). Köln: Sport und Buch Strauß.
- Hollmann, W. (1990). *Training, Grundlagen und Anpassungsprozesse* (Studienbrief der Trainerakademie Köln des Deutschen Sportbundes). Schorndorf. Hofmann.
- Hollmann, W. & Strüder, H. K. (2009). *Sportmedizin. Grundlagen für körperliche Aktivität, Training und Präventivmedizin*. Stuttgart: Schattauer.
- Holmes, D. S. & Roth, D. L. (1985). Association of aerobic fitness with pulse rate and subjective responses to psychological stress. *Psychophysiology*, 22, 525-529.
- Hong, S., Farag, N. H., Nelesen, R. A., Ziegler, M. G. & Mills, P. J. (2004). Effects of regular exercise on lymphocyte subsets and CD62L after psychological vs. Physical stress. *Journal of Psychosomatic Research*, 56, 363-370.
- Hossiep, R. & Paschen, M. (1998). *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung*. Göttingen: Hogrefe.
- Hossiep, R. & Paschen, M. (2003). *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung*. Göttingen: Hogrefe.
- Hossiep, R. & Bräutigam, S. (2006). Personalauswahl und -entwicklung mit dem Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP). In W. Simon (Hrsg.), *Persönlichkeitsmodelle und Persönlichkeitstests* (S. 136-158). Offenbach: Gabal.
- Hossiep, R. & Krüger, C. (2012). *BIP-6F, Bochumer Inventar zur Persönlichkeitsbeschreibung – 6 Faktoren*. Göttingen: Hogrefe.
- Höft, S. (2007). Assessment Center. In H. Schuler & K. Sonntag (Hrsg.), *Handbuch der Arbeits- und Organisationspsychologie* (S. 475-482). Göttingen: Hogrefe.

- Hull, C. L. (1928). *Aptitude testing*. Yonkers: World Book.
- Izard, C. E. (1981). *Die Emotionen des Menschen*. Weinheim: Beltz.
- Jones, G., Hanton, S. & Connaughton, D. (2002). What is this thing called mental toughness? An investigation of elite sport performers. *Journal of Applied Sport Psychology*, 14, 205-218.
- Judge, T. A., Thoresen, C., Bono, J. E. & Patton, G. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, 127, 376-407.
- Judge, T. A., Klinger, R., Simon, L. S. & Yang, I. W. F. (2008). The contributions of personality to organizational behavior and psychology: Findings, criticisms, and future research directions. *Social and Personality Psychology Compass*, 2, 1982-2000.
- Judge, T. A. & Bono, J. E. (2009). Relationship of core self-evaluations traits – self-esteem, generalized self efficacy, locus of control and emotional stability – with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86, 80-92.
- Kaluza, G. (1996). *Gelassen und sicher im Streß. Psychologisches Programm zur Gesundheitsförderung*. Berlin: Springer.
- Kamon, E. & Pandolf, K. B. (1972). Maximal aerobic power during uphill running and cycling. *Journal of Applied Physiology*, 4, 467.
- Kaufmann, I., Pornschlegel, H. & Udriș, I. (1982). Arbeitsbelastung und Beanspruchung. In L. Zimmermann (Hrsg.), *Humane Arbeit – Leitfaden für Arbeitnehmer (Band 5). Belastungen und Streß bei der Arbeit* (S. 13-48). Reinbek: Rowohlt.
- Keller, S. & Seraganian, P. (1984). Physical fitness level and autonomic reactivity to psychosocial stress. *Journal of Psychosomatic Research*, 28, 279-287.
- Kenny, D. A. & Kashy, D. A. (1992). Analysis of the Multitrait-Multimethod Matrix by confirmatory factor-analysis. *Psychological Bulletin*, 112, 165-172.
- Kernen, H. (1997). *Burnout-Prophylaxe im Management. Erfolgreiches individuelles und institutionelles Ressourcenmanagement*. Bern: Haupt.
- Kerr, J. H. & Vlaswinkel, E. H. (1990). Effects of exercise on anxiety: A review. *Anxiety Research*, 2, 309-325.
- Kieschke, U. (2003). *Arbeit, Persönlichkeit und Gesundheit. Beiträge zu einer differentiellen Psychologie beruflichen Belastungsgeschehens*. Berlin: Logos.
- Klaperski, S., von Dawans, B., Heinrichs, M. & Fuchs, R. (2013). Does the level of physical exercise affect physiological and psychological responses to psychosocial stress in women? *Psychology of Sport & Exercise*, 14, 266-274.
- Kleinbeck, U. & Rutenfranz, J. (1987). *Arbeitspsychologie*. Göttingen: Hogrefe.

- Kleinginna, P. R. jr. & Kleinginna, A. M. (1981). A categorized list of emotion definitions with suggestions for a consensual definition. *Motivation and Emotion*, 5, 345-355.
- Klimmer, F. (1978). *Analyse mentaler und emotionaler Beanspruchung des arbeitenden Menschen*. Unveröff. Dissertation, TH Darmstadt.
- Knapp, B. (1965). The personality of lawn tennis players. *Bulletin of the British Psychological Society*, 18, 21-23.
- Krampen, G. & Reichle, B. (2002). Frühes Erwachsenenalter. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 319-349). Weinheim: Beltz/PVU.
- Krystek, U., Becherer, D. & Deichmann, K. (1995). *Innere Kündigung: Ursachen, Wirkungen und Lösungsansätze*. München: Hampp.
- Kuhn, K. (2007). Arbeitsbedingte Einflüsse bei der Entstehung chronischer Krankheiten. In B. Badura, H. Schellschmidt & C. Vetter (Hrsg.), *Fehlzeiten-Report 2006. Chronische Krankheiten. Zahlen, Daten, Analysen aus allen Branchen der Wirtschaft* (S. 25-43). Berlin: Springer.
- Kuhl, J. (2000). Handlungs- und Lageorientierung. In W. Sarges (Hrsg.), *Management Diagnostik* (S. 303-316). Göttingen: Hogrefe.
- Kunath, P. (1975). Das Problem der Trainings- und Wettkampfbelastung in der Sicht des Sportpsychologen. In P. Kunath (Gesamtleitung Autorenkollektiv), *Psychologie im Sport* (S. 118-129). Berlin: Sportverlag.
- Lang, P. J. (1993). The three-system approach to emotion. In N. Birbaumer & A. Öhman (Hrsg.), *The Structure of Emotion* (S. 18-30). Seattle: Hogrefe & Huber.
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Lazarus, R. S. & Launier, R. (1981). Streßbezogene Transaktion zwischen Person und Umwelt. In J. R. Nitsch (Hrsg.), *Streß – Theorien, Untersuchungen, Maßnahmen* (S. 213-259). Bern: Huber.
- Leitl, M. (2014). Was ist ... Resilienz? *Harvard Business Manager*, 2, 46-49.
- Leyk, D., Rhode, U., Gorges, W., Ridder, D., Wunderlich, M., Dinklage, C., Sievert, A., Rüter, T., Essfeld, D., (2006). Physical Performance, Body Weight and BMI of Young Adults in Germany 2000-2004: Results of the Physical-Fitness-Test Study. *International Journal of Sports Medicine*, 27, 642-647.
- Leyk, D. Witzki, A., Gorges, W., Rhode, U., Lison, A., Rondé, M., Wömpener, H., Schlattmann, A., Dobmeier, H., Rüter, T. & Wunderlich, M. (2010). Körperliche Leistungsfähigkeit, Körpermasse und Risikofaktoren von 18 bis 35-jährigen Soldaten: Ergebnisse der Evaluierungsstudie zum Basis-Fitness-Test (BFT). *Wehrmedizinische Monatszeitschrift*, 54 (11-12), 278-283.

- Leyk, D. (2012). Übergewicht, sinkende Fitness und Belastbarkeit – auch in der Bundeswehr? Online: <http://www.sanitaetsdienst-bundeswehr.de/portal/a/sanitaetsdienst/> (Zugriff am 16.12.2013)
- Leyk, D., Franke, E., Hofmann, M., Klein, G., Weller, N., Hackfort, D., Löllgen, H. & Piekarski, C. (2013). Gesundheits- und Fitnessförderung in der Bundeswehr: Von ressourcenorientierter Präventionsforschung zur Umsetzung in die Fläche. *Wehrmedizinische Monatszeitschrift*, 57 (7), 162-166.
- Liebert, C. (2006). *Emotionale Belastbarkeit und Umgang mit emotionalen Belastungen. Eine Auswertungshilfe für qualitative entscheidungsorientierte Gespräche*. TU Dresden: Diplomarbeit.
- Lienert, G. A. & Ratz, U. (1994). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lienert, G. A. & Ratz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- Lohmann-Haislah, A. (2012). *Stressreport Deutschland 2012. Psychische Anforderungen, Ressourcen und Befinden*. Dortmund, Berlin, Dresden: Bundesanstalt für Arbeitsmedizin und Arbeitsschutz.
- Long, B. C. (1991). Physiological and psychological stress recovery of physically fit and unfit women. *Canadian Journal of Behavioral Science*, 23, 53-65.
- Maddi, S. R. (2006). Hardiness: The courage to grow from stresses. *Journal of Positive Psychology*, 1, 160-168.
- Maddi, S. R. & Khoshaba, D. M. (2005). *Resilience at work: How to succeed no matter what life throws at you*. New York: Amacom.
- Maddi, S. R., Matthews, M. D., Kelly, R. D., Villarreal, B. & White, M. (2012). The Role of Hardiness and Grit in Predicting Performance and Retention of USMA Cadets. *Military Psychology*, 24 (1), 19-28.
- Martens, R. (1975). The paradigmatic crisis in American sport personology. *Sportwissenschaft*, 5 (1), 9-24.
- Martin, D. & Nicolaus, J. (1998). Leistungsvoraussetzungen und sportliche Leistungsfähigkeit von Kindern. Kinder und Jugendliche im Leistungssport. Beiträge des internationalen, interdisziplinären Symposiums „Kinderleistungen“, 07.-10.11.1966 in Saarbrücken. *Schriftenreihe des Bundesinstituts für Sportwissenschaft*, 95, 347-3639. Schorndorf: Hofmann.
- Martinsen, E. W. (1987). The role of aerobic exercise in the treatment of depression. *Stress Medicine*, 3, 93-100.
- Martinsen, E. W., Medhus, A. & Sandvik, L. (1985). Effects of aerobic exercise on depression: A controlled study. *British Medical Journal*, 291, 109.

- Martinsen, E. W. (2002). The Role of Exercise in the Management of Depression. In L. D. Zaichkowsky & D. L. Mostofsky (Hrsg.), *Medical and Psychological Aspects of Sport and Exercise* (S. 205-214). Morgantown: Fitness Information Technology.
- McCann, I. L. & Holmes, D. S. (1984). Influence of aerobic exercise on depression. *Journal of Personality and Social Psychology*, 46, 1142-1147.
- McClelland, D. C. (1987). *Human Motivation*. Cambridge: Cambridge University Press.
- McCrae, R. R. & Costa Jr., P. T. (1985). Updating Norman's Adequate Taxonomy': *intelligence* and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49, 710-721.
- McCrae, R. R. & Costa Jr., P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McKelvie, J. S., Lemieux, P. & Stout, D. (2002). Extraversion and neuroticism in contact athletes, no contact athletes and non-athletes: A research note. *Athletic Insight*, 5 (3). Online: <http://www.athleticinsight.com/Vol5Iss3/ExtraversionNeuroticism.htm>. (Zugriff am 05.02.2014)
- McKinnon, D. W. (1958). An Assessment Study of Air Force Officers. In *Wright Air Development Center Technical Report*, 91 (5), Personnel Laboratory, Lackland Air Force Base, Texas.
- Meili, R. & Steingrüber, H.-S. (1978). *Lehrbuch der psychologischen Diagnostik*. Stuttgart: Huber.
- Meyer, T. & Kindermann, W. (1999). Die maximale Sauerstoffaufnahme (VO_{2max}). *Deutsche Zeitschrift für Sportmedizin*, 50 (9), 285-286.
- Meyer, T. (2003). Der Respiratorische Quotient (RQ). *Deutsche Zeitschrift für Sportmedizin*, 54 (1), 29-30.
- Midgley, A. W., Bentley, D. J., Luttikholt, H., McNaughton, L. R. & Millet, G. P. (2008). Challenging a Dogma of Exercise Physiology. *Sports Medicine*, 38 (6), 441-447.
- Moser, K. & Schuler, H. (2013). Geschichte der Management-Diagnostik. In W. Sarges (Hrsg.), *Management-Diagnostik* (S. 33-43). Göttingen: Hogrefe.
- Morgan, W. P. (1968). Personality characteristics of wrestlers participating in the world championships. *Journal of Sports Medicine and Physical Fitness*, 8, 212-216.
- Morgan, W. P. & Costill, D. L. (1972). Psychological characteristics of the marathon runner. *Journal of Sports Medicine and Physical Fitness*, 12, 42-46.
- Morgan, W. P. (1980). The trait psychology controversy. *Research Quarterly for Exercise and Sport*, 51 (1), 50-76.

- Morgan, W. P. & Goldston, S. E. (1987). *Exercise and mental health*. Washington, DC: Hemisphere.
- Morris, T. (1995). Self-efficacy in Sport and Exercise. In T. Morris & J. Summer (Hrsg.), *Sport Psychology. Theorie, Applications and Issues* (S. 143-172). Brisbane: John Wiley & Sons
- Mummendey, H. D. (1983). Sport und Persönlichkeit – Versuch einer Tertiäranalyse. *Bielefelder Arbeiten zur Sozialpsychologie*, 97, 1-22.
- Mussel, P. (2012). Persönlichkeitsaspekte intellektueller Leistungen. *Report Psychologie*, 37, 440-448.
- Mutrie, N. (1988). Exercise as a treatment for moderate depression in the UK health service. In S. Biddle (Hrsg.), *Sport, health, psychology and exercise symposium* [Proceedings]. Bisham Abbey National Sports Centre: Buckinghamshire.
- Mutrie, N. (2000). The relationship between physical activity and clinically defined depression. In S. Biddle, K. Fox & S. Boutcher (Hrsg.), *Physical activity and psychological well-being* (S. 46-62). London: Routledge.
- Münzberger, E. (2008). Modularer Lehrbrief „Einführung in die Arbeitsmedizin“ Online: <http://arbmed.med.uni-rostock.de/lehrbrief/arbphys.htm#PhysBel> (Zugriff am 05.01.2014).
- National Audit Office (2006). Recruitment and Retention in the Armed Forces, Ministry of Defence, *Report by the Comptroller and Auditor General*, HC 1633-I Session 2005-2006, 32.
- Nerdinger, F. W., Blickle, G. & Schaper, N. (2008). *Arbeits- und Organisationspsychologie*. Heidelberg: Springer.
- Nitsch, J. R. & Udris, I. (1976). *Beanspruchung im Sport*. Bad Homburg: Limpert.
- Nitsch, J. R. (1981). Zur Gegenstandsbestimmung der Streßforschung. In J.R. Nitsch (Hrsg.), *Stress – Theorien, Untersuchungen, Maßnahmen* (S. 29-51). Bern: Huber.
- Nitsch, J. R. (1982). Zum Problem der Personenspezifität von Stressreaktionen. In B. Kirkcaldy (Hrsg.), *Individual differences in sport behaviour* (S. 183-199). Köln: bps.
- Nitsch, J. R. (1985). Emotionen und Handlungsregulationen. In G. Schilling & K. Herren (1985). *Angst, Freude und Leistung im Sport* (S. 37-60). Magglingen: Eigendruck.
- Noakes, T. D. (2001). *Lore of Running*. Champaign, IL: Human Kinetics.
- Noakes, T. D. (2012). Fatigue is a Brain-Derived Emotion that Regulates the Exercise Behavior to Ensure the Protection of Whole Body Homeostasis. In *Front Physiology*, (82) 3. Online: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3323922/> (Zugriff am 24.04.2014).
- Noble, B. J. & Robertson, R. J. (1996). *Perceived exertion*. Champaign, IL: Human Kinetics.

- Ones, D. S., Viswesvaran, C. & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679-703.
- Oppolzer, A. (1989): *Handbuch Arbeitsgestaltung. Leitfaden für eine menschengerechte Arbeitsorganisation*. Hamburg: VSA.
- Ostendorf, F. & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae: NEO-PI-R*. Göttingen: Hogrefe.
- Pahmeier, I. (1994). *Sportliche Aktivität als Bewältigungshilfe bei gesundheitlichen Beeinträchtigungen*. Frankfurt: Verlag Harri Deutsch.
- Perlman, B. & Hartman, E. A. (1982). Burnout: Summary and future research. *Human Relations*, 35, 283-305.
- Piedmont, L. R., Hill, C. D. & Blanco, S. (1997). Predicting athletic performance using the five-factor model of personality. *Personality and Individual Differences*, 27, 769-777.
- Piedmont, R. L. (1995). Big five adjective marker scales for use with college students. *Psychological Reports*, 77, 160-162.
- Plante, G. T. & Rodin, J. (1990). Physical Fitness and Enhanced Psychological Health. *Current Psychology: Research & Reviews*, Spring 1990, 9 (1), 3-24.
- Plante, T., Coscarelli, L. & Ford, M. (2001). Does exercising with another enhance the stress-reducing benefits of exercise? *International Journal of Stress Management*, 8, 201-213.
- Plate, T. (2006). *Personalauswahl in Unternehmensberatungen: Validität und Nutzen der Eignungsdiagnostik*. Wiesbaden: Deutscher Universitäts Verlag.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellermann (Hrsg.), *Emotion. Theory, research and experience (Band 1): Theories of emotion* (S. 3-33). New York: Academic Press.
- Pokan, R., Förster, H., Hofmann, P., Hörtnagel, H., Ledl-Kurkowski, E. & Wonisch, M. (2004). *Kompendium der Sportmedizin. Physiologie, innere Medizin und Pädiatrie*. Wien: Springer.
- Pronk, N. P., Martinson, B., Kessler, R. C., Beck, A. L., Simon, G. E. & Wang, P. (2004). The association between work performance and physical activity, cardiorespiratory fitness, and obesity. *Journal of Occupational and Environmental Medicine*, 46, 19-25.
- Rammsayer, T. & Weber, H. (2010). *Differentielle Psychologie – Persönlichkeitsforschung*. Göttingen: Hogrefe.
- Reeve, C. L. & Bonaccio, S. (2008). Does test anxiety induce measurement bias in cognitive ability tests? *Intelligence*, 36, 526-538.

- Reiner, M., Niermann, C., Krapf, F. & Woll, A. (2013). Stress Sport und Beschwerdewahrnehmung. Puffereffekte von Sport und körperlicher Aktivität? *Sportwissenschaft*, 43, 264-275.
- Richter, P. & Hacker, W. (1998). *Belastung und Beanspruchung. Streß, Ermüdung und Burn-out im Arbeitsleben*. Heidelberg: Asanger.
- Rohde U., Erley O., Rütter T., Wunderlich M. & Leyk D. (2007). Leistungsanforderungen bei typischen soldatischen Einsatzbelastungen. *Wehrmedizin und Wehrpharmazie*, 51, 138-142.
- Roberts, B. W. & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126 (1), 3-25.
- Robertson, I. T. & Kinder, A. (1993). Personality and Job competences: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology*, 66, 225-244.
- Roecker, K., Niess, A., Horstmann, T., Striegel, H., Mayer, F. & Dickhuth, H. H. (2002). Heart rate prescriptions from performance and anthropometrical characters. *Medicine and Science in Sports and Exercise*, 34, 881-887.
- Rohmert, W. & Rutenfranz, J. (1975). *Arbeitswissenschaftliche Beurteilung der Belastung und Beanspruchung an unterschiedlichen industriellen Arbeitsplätzen*. Bonn: Bundesminister für Arbeit und Sozialordnung.
- Rueter, M., Mutrie, N. & Harris, D. (1982). *Running as an adjunct to counseling in the treatment of depression*. Unpublished manuscript, The Pennsylvania State University.
- Rushall, B. S. (1972). Three studies relating personality variables to football performance. *International Journal of Sport Psychology*, 3, 12-24.
- Rutenfranz, J. (1972). Arbeitsmedizinische Aspekte des Streßproblems. In J. Nitsch (Hrsg.) *Streß, Theorien, Untersuchungen, Maßnahmen* (S. 379-390). Bern: Huber.
- Sack, H.-G. (1975). *Sportliche Betätigung und Persönlichkeit*. Ahrensburg: Czwalina.
- Sack, H.-G. (1982). Interindividuelle Persönlichkeitsunterschiede und Sportengagement. In B. D. Kirkcaldy (Hrsg.), *Individual differences in sport behaviour* (S. 99-158). Köln: bps.
- Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, 46, 929-938.
- Sarges, W. (1995). *Management-Diagnostik*. Göttingen: Hogrefe.
- Sarges, W. (2001). *Weiterentwicklung der Assessment Center-Methode*. Göttingen: Hogrefe.
- Schnabel, G., Harre, H. D. & Krug, J. (2008). *Trainingslehre – Trainingswissenschaft*. Aachen: Meyer & Meyer.

- Schermelleh-Engel, K. & Schweizer, K. (2008). Multitrait-Multimethod-Analysen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 325-341). Berlin: Springer.
- Schermelleh-Engel, K. & Werner, C. (2007). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 325-342). Heidelberg: Springer.
- Schlattmann, A. & Hackfort, D. (1988). *Die funktionale Bedeutung der Kontrolle positiver Emotionen im Sport*. Unveröffentl. Zwischenbericht an das Bundesinstitut für Sportwissenschaft.
- Schlicht, W. (1993). Does physical exercise reduce anxious emotions? A meta-analysis. *Anxiety, Stress and Coping*, 6, 275-288.
- Schlicht, W. (1995). *Wohlbefinden und Gesundheit durch Sport*. Schorndorf: Hofmann.
- Schlick, C., Bruder, R. & Luczak, H. (2010). *Arbeitswissenschaft*. Heidelberg: Springer.
- Schmidt-Atzert, L. & Ströhm, W. (1983). Ein Beitrag zur Taxonomie der Emotionswörter. *Psychologische Beiträge*, 25, 126-141.
- Schuler, H. (1996). *Psychologische Personalauswahl*. Göttingen: Verlag für Angewandte Psychologie.
- Schuler, H. (2000). *Psychologische Personalauswahl. Einführung in die Berufseignungsdiagnostik*. Göttingen: Verlag für Angewandte Psychologie.
- Schuler, H. & Funke, U. (1991). *Eignungsdiagnostik in Forschung und Praxis*. Stuttgart: Verlag für Angewandte Psychologie.
- Schwinger, M., Olbricht, S. & Stiensmeier-Pelster, J. (2013). Der Weg von der Persönlichkeit zu sportlichen Leistungen. Ein hierarchisches Modell. *Zeitschrift für Sportpsychologie*, 20 (3), 81-93.
- Schöpfflug, W. (1983). Coping efficiency and situational demands. In G. R. J. Hockey (Hrsg.), *Stress and fatigue in human performance* (S. 299-330). London: Wiley.
- Semmer, N. (1988). Streß. In R. Asanger & G. Wenninger (Hrsg.), *Handwörterbuch Psychologie* (S. 744-752). München: Psychologie Verlags Union.
- Sexton, H., Mære, Å. & Dahl, N. H. (1989). Exercise intensity and reduction in neurotic symptoms. *Acta Psychiatrica Scandinavica*, 80, 231-235.
- Sharkey, B. J. (1990). *Physiology of Fitness*. Champaign: Human Kinetics.
- Sime, W. E. (1987). Exercise in the treatment and prevention of depression. In W. P. Morgan & S. E. Goldston (Hrsg.), *Exercise and mental health* (S. 145-152). Washington, DC: Hemisphere.

- Simonov, P. V. (1975). *Higher nervous activity of men. Motivational-emotional aspects*. Moscow: Izdatelstvo Nauka.
- Singer, R. (2000). Sport und Persönlichkeit. In H. Gabler, J. R. Nitsch & R. Singer (Hrsg.), *Einführung in die Sportpsychologie* (S. 289-332). Schorndorf: Hofmann.
- Sinyor, D., Golden, M., Steinert, Y. & Seraganian, P. (1986). Experimental manipulation of aerobic fitness and the response to psychosocial stress. Heart rate and self-report measures. *Psychosomatic Medicine*, 48, 324-337.
- Sinyor, D., Schwartz, S. G., Peronnet, F., Brisson, G. & Seraganian, P. (1983). Aerobic fitness level and reactivity to psychosocial stress: Physiological, biochemical, and subjective measures. *Psychosomatic Medicine*, 65, 205-217.
- Spalding, T. W., Lyon, L. A., Steel, D. H. & Hatfield, B. D. (2004). Aerobic exercise training and cardiovascular reactivity to psychological stress in sedentary young normotensive men and women. *Psychophysiology*, 41, 552-562.
- Spector, P. E., Zapf, D., Chen, P. Y. & Frese, M. (1998). Why Negative Affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior*, 21, 79-95.
- Statistisches Bundesamt (2010). Hohe Kosten durch Demenz und Depressionen. Pressemitteilung Nr. 280. Online: https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2010/08/PD10_280_23631.html (Zugriff am 07.05.2014).
- Stephoe, A. (1994). Aerobic exercise, stress and health. In J. R. Nitsch & R. Seile (Hrsg.), *Bewegung und Sport* (S. 78-91). St. Augustin: Academia.
- Stern, W. (1935). *Allgemeine Psychologie auf personalistischer Grundlage*. Haag: Nijhoff.
- Stone, P. T. (1961). *Is the athlete different in personality from the non-athlete*. Paper read at Eastern Conference of British Association of Sports and Medicine.
- Sonntag, K. (2007). Theorien der Arbeitstätigkeit. In H. Schuler & K. Sonntag (Hrsg.), *Handbuch der Arbeits- und Organisationspsychologie* (S. 35-42). Göttingen: Hogrefe.
- Sonnentag, S. (2012). Psychological Detachment from work during leisure time: The benefits of mentally disengaging from work. *Psychological Science*, 21 (2), 114-118.
- Hickmann, C. & Pollmer, C. (2013). Einsatz mit offener Wunde. *Süddeutsche Zeitung*, Nr. 273, 2.
- Thorndike, R. L. (1949). *Personnel Selection*. New York: Wiley.
- Udris, I. & Frese, M. (1999). Belastung und Beanspruchung. In: C. G. Hoyos, & D. Frey, (Hrsg.), *Arbeits- und Organisationspsychologie* (S. 429-445). Weinheim: Psychologie Verlags Union.

- Uhle, T. & Treier, M. (2011). *Betriebliches Gesundheitsmanagement. Gesundheitsförderung in der Arbeitswelt Mitarbeiter einbinden, Prozesse gestalten, Erfolge messen*. Berlin: Springer.
- Ulmer, H.-V. *Belastung und Beanspruchung, Beanspruchungsregulation und Zielantizipation*.
Online: <http://www.med.uni-jena.de/motorik/pdf/ulmer2.pdf> (Zugriff am 22.12.2013).
- Van Andel, G. E. & Austin, D. R. (1984). Physical fitness and mental health: A review of the literature. *Adapted Physical Activity Quarterly*, 1, 207-220.
- Vanek, M. & Hosek, V. (1977). *Zur Persönlichkeit des Sportlers*. Schorndorf: Hofmann.
- Vealey, R. S. (1989). Sport personology: A paradigmatic and methodological analysis. *Journal of Sport and Exercise Psychology*, 11 (2), 216-235.
- Wankel, L. M. (1985). Personal and Situational Factors Affecting Exercise Involvement: The Importance of Enjoyment. *Research Quarterly for Exercise and Sport*, 56 (3), 275-282.
- Weber, P. (1985). *Probleme der Messung subjektiver Empfindungen*. Eschborn: Fachbuchhandlung für Psychologie.
- Weisung IGF (2009). *Weisung zur Ausbildung und zum Erhalt der Individuellen Grundfertigkeiten*. Berlin: Bundesministerium der Verteidigung.
- Wenzel, M. (2008). *Psychologische Grundlagen für die Allokation von Studieninteressenten an wissenschaftlichen Hochschulen*. Unveröff. Diss., Ruhr-Universität Bochum.
- Westhoff, K., Steinborn, M. B. & Schurz, A. (2013). Was versteht man unter dem Konstrukt „Gewissenhaftigkeit“? *Report psychologie*, 38 (5), 200 ff.
- Westhoff, K. (2005). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430* (2. überarb. Aufl.). Lengerich: Pabst Science Publishers.
- Westhoff, K. & Kluck, M. L. (2003). *Psychologische Gutachten schreiben und beurteilen*. (4. vollst. überarb. und erw. Aufl.). Berlin: Springer.
- WHO (2013). *BMI classification* Oline:
http://apps.who.int/bmi/index.jsp?introPage=intro_3.html (Zugriff am 03.12.2013).
- Wiggins, J. S. (1973). *Personality: And Prediction Principles of Personality Assessment*. Reading: Addison-Wesley Publishing Company.
- Wittchen, H.-U., Schönfeld, S., Kirschbaum, C., Trautmann, S., Thurau, C., Siegert, J., Höfler, M., Hauffa, R. & Zimmermann, P. (2013). Rates of mental disorders among german soldiers deployed to Afghanistan: Increased risk of PTSD or of mental disorders in general? *Depression and Anxiety*, 2 (1), 1-7.
- Witzki, A., Sievert, A., Gorges, W., Glitz, J. K., Günter, K. & Leyk, D. (2010) Psychophysiologische Leistungsparameter im Soldatenberuf: Multi-methodales Vorgehen eröffnet neue Perspektiven. *Wehrmedizinische Monatszeitschrift*, 54, 11-12, 283.

-
- Young, R. & Ismail, A. H. (1977). Comparison of Selected Psychological and Personality Variables in Regular and Nonregular Adult Male Exercisers. *Research Quarterly*, 48, 617-622.
- Zapf, D. & Semmer, N. (2004). Stress und Gesundheit in Organisationen. In H. Schuler (Hrsg.), *Enzyklopädie der Psychologie, Organisationspsychologie – Grundlagen und Personalpsychologie*, 3, (S. 1007-1112). Göttingen: Hogrefe.
- Zintl, F. (1994). *Ausdauertraining*. München: BLV

Anhang

Die Skalen „Emotionale Stabilität“ und „Belastbarkeit“ des BIP, sowie die Neurotizismus-Skala des NEO-FFI wurden dem Original entsprechend eingesetzt und sind aus urheberrechtlichen Gründen im Anhang nicht wiedergegeben.

Anhang A: Korrelationstabellen

Tabelle A1

Interkorrelationen der physischen Belastbarkeitsindikatoren ($N = 145$).

	Zeit: 3000-m-FET	Zeit: Rampe F.- Ergometer	Zeit: 1000-m-Lauf	Zeit: Rampe Laufband	VO_{2max} : 3000-m-FET	VO_{2max} : Rampe F.- Ergometer	VO_{2max} : 1000-m-Lauf	VO_{2max} : Rampe Laufband	VO_{2max} : 1000-m-Lauf	VO_{2max} : Rampe Laufband	Durchsch. Wattleistung: 3000-m-FET	Max. Wattleistung: Rampe F.- Ergometer
Zeit: 3000-m-FET												
Zeit: Rampe F.- Ergometer	-.812**											
Zeit: 1000-m-Lauf	.718**	-.591**										
Zeit: Rampe Laufband	-.641**	.680**	-.796**									
VO_{2max} : 3000-m-FET	-.854**	.886**	-.657**	.714**								
VO_{2max} : Rampe F.- Ergometer	-.800**	.928**	-.600**	.670**	.941**							
VO_{2max} : 1000-m-Lauf	-.780**	.818**	-.674**	.681**	.920**	.901**						
VO_{2max} : Rampe Laufband	-.733**	.835**	-.621**	.694**	.911**	.915**	.954**					
Durchsch. Wattleistung: 3000-m-FET	-.930**	.872**	-.640**	.651**	.907**	.857**	.820**	.781**				
	-.724**	.904**	-.542**	.587**	.810**	.874**	.764**	.795**				
												.789**

** $p \leq .001$

Tabelle A2

Interkorrelationen der Skalen des BIP (nach Hossiep & Paschen, 1998, S. 31).

Skala	LM	GM	FM	Ge	Fl	HO	Sen	Ko	Soz	TO	Du	ESt	Bel	SB
Leistungs- motivation (LM)	1.00													
Gestaltungs- motivation (GM)	.58	1.00												
Führungs- motivation (FM)	.50	.61	1.00											
Gewissen- haftigkeit (Ge)	.22	-.05	-.03	1.00										
Flexibilität (Fl)	.47	.55	.56	-.27	1.00									
Handlungs- orientierung (HO)	.36	.29	.42	.32	.39	1.00								
Sensitivität (Sen)	.20	.15	.33	.06	.32	.37	1.00							
Kontakt- fähigkeit (Ko)	.23	.33	.50	-.08	.43	.32	.55	1.00						
Soziabilität (Soz)	-.17	-.35	-.22	.06	-.06	.06	.37	.13	1.00					
Teamorien- tierung (TO)	.05	.09	.26	-.11	.30	.19	.23	.33	.27	1.00				
Durchsetzungs- stärke (Du)	.44	.63	.75	.03	.48	.41	.28	-.44	-.36	.15	1.00			
Emotionale Stabilität (ESt)	.19	.24	.41	-.08	.51	.45	.30	.36	.12	.22	.39	1.00		
Belastbarkeit (Bel)	.52	.40	.51	.09	.61	.59	.35	.33	.09	.21	.43	.68	1.00	
Selbstbewusst- sein (SB)	.36	.51	.67	-.04	.56	.47	.36	.55	.17	.18	.67	.66	.59	1.00

N = 5245; bis auf die Korrelationen der Skala Gewissenhaftigkeit mit den Skalen Führungsmotivation und Durchsetzungsstärke sind alle Korrelationen auf dem 1%-Niveau signifikant.

Tabelle A3

Korrelationen der VE_{max} -Werte zwischen Vor- und Nachtest ($n = 29$).

t1	t2	Korrelation nach Pearson	Signifikanz
3000-m-FET	3000-m-FET	.768	.000
1000-m-Lauf	1000-m-Lauf	.788	.000
Rampentest F.-Ergom.	Rampentest F.-Ergom.	.746	.000
Rampentest Laufband	Rampentest Laufband	.754	.000

Tabelle A4

Korrelationen zwischen den physischen Belastbarkeitsindikatoren und den Itemscores der Skala zur allgemeinen Selbstwirksamkeit ($N = 145$).

	Skala zur allg. Selbstwirksamkeit
Zeit: 3000-m-FET	-.038
Zeit: 1000-m-Lauf	-.208
Zeit: Rampentest F.-Ergom.	.134
Zeit: Rampentest Laufband	.160
VO_{2peak} : 3000-m-FET	.173
VO_{2max} : 1000-m-Lauf	.135
VO_{2peak} : Rampentest F.-Ergom.	.102
VO_{2max} : Rampentest Laufband	.137
Durchschn. Wattleistung: 3000-m-FET	.168
Maximale Wattleistung: Rampentest F.-Ergom.	-.013

Anmerkung: $N = 145$; bis auf die Korrelationen der 3000-m-FET-Zeiten mit den Itemscores der Skala zur allg. Selbstwirksamkeit, hier bestand die Stichprobe aus 225 Pbn. Keine der Korrelationen hat sich als signifikant erwiesen.

Tabelle A5

Korrelation zwischen der 3000-m-FET-Zeit, gemessen mit dem Ergometer der Firma Ergo Fit und den Itemscores der Skala zur psychischen Belastbarkeit (Lang- und Kurzform; n = 82).

		Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Zeit 3000-m-FET	Korrelation nach Pearson	-.180	-.243
	Signifikanz (2-seitig)	.004	.027

Tabelle A6

Korrelation zwischen der 3000-m-FET-Zeit, gemessen mit den Tacx-Trainern und den Itemscores der Skala zur psychischen Belastbarkeit (Lang- und Kurzform; n = 342).

		Kurzform-Skala: psychische Belastbarkeit (Trait)	Langform-Skala: psychische Belastbarkeit (Trait)
Zeit 3000-m-FET	Korrelation nach Pearson	-.189	-.250
	Signifikanz (2-seitig)	.000	.000

Tabelle A7

Korrelationen der VE_{max} -Werte zwischen den verschiedenen Ausdauerstestverfahren (N = 145).

		1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
3000-m-FET	Korrelation nach Pearson	.782	.735	.705
	Signifikanz (2-seitig)	.000	.000	.000

Anhang B: Fragebögen psychische Belastbarkeit (Kurzform-Skalen)

Bitte stufen Sie jede der folgenden 29 Aussagen so ein, wie es für Sie persönlich **im Allgemeinen** zutrifft.

Wählen Sie für jede Aussage eines der sechs Kästchen von „**trifft voll zu**“ bis „**trifft überhaupt nicht zu**“ aus, und lassen Sie bitte keine Aussage unbeantwortet.

Die Daten werden anonym behandelt und können nicht auf Ihre Person zurückgeführt werden.

Vielen Dank für Ihre Mitarbeit!

VPN-Nr.

Bitte geben Sie anhand der folgenden Merkmale an, inwiefern diese **im Allgemeinen** auf Sie zutreffen.

1.

robust, standhaft, willensstark, gelassen

—————

trifft voll zu

trifft überhaupt nicht zu

2.

nervös, reizbar, entmutigt, instabil

—————

trifft voll zu

trifft überhaupt nicht zu

Abbildung B2: Kurzform der psychischen Belastbarkeitsskala (Trait).

VPN-Nr.

vorher

Bitte geben Sie anhand der folgenden Merkmale an, inwiefern diese **im Moment** auf Sie zutreffen.

1.

robust, standhaft, willensstark, gelassen

—————

trifft voll zu

trifft überhaupt nicht zu

2.

nervös, reizbar, entmutigt, instabil

—————

trifft voll zu

trifft überhaupt nicht zu

Abbildung B3: Kurzform der psychischen Belastbarkeitsskala (State, vorher).

nachher

Bitte geben Sie anhand der folgenden Merkmale an, inwiefern diese **im Moment** auf Sie zutreffen.

1.

robust, standhaft, willensstark, gelassen

trifft voll zu

trifft überhaupt nicht zu

2.

nervös, reizbar, entmutigt, instabil

trifft voll zu

trifft überhaupt nicht zu

Abbildung B4: Kurzform der psychischen Belastbarkeitsskala (State, nachher).

Anhang C: Fragebogen zur Konstruktvalidierung

<p>1.) Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>2.) Die Lösung schwieriger Probleme gelingt mir immer, wenn ich mich darum bemühe.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>3.) Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>4.) In unerwarteten Situationen weiß ich immer, wie ich mich verhalten soll.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>5.) Auch bei überraschenden Ereignissen glaube ich, dass ich gut mit ihnen zurechtkommen kann.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>6.) Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>7.) Was auch immer passiert, ich werde schon klarkommen.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>8.) Für jedes Problem kann ich eine Lösung finden.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>9.) Wenn eine neue Sache auf mich zukommt, weiß ich, wie ich damit umgehen kann.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>
<p>10.) Wenn ein Problem auftaucht, kann ich es aus eigener Kraft meistern.</p> <p><input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/> — <input type="checkbox"/></p> <p>stimmt nicht stimmt kaum stimmt eher stimmt genau</p>

Abbildung C1: Skala zur allgemeinen Selbstwirksamkeitserwartung (nach Jerusalem & Schwarzer, 1999).

Anhang D: E-Mail zur Probandengewinnung

Sehr geehrte Damen und Herren,

im Rahmen einer wissenschaftlichen Studie des Departments für Sportwissenschaft suchen wir Probandinnen und Probanden für eine leistungsdiagnostische Untersuchung (Spiroergometrie und Laktat) mit dem Fahrradergometer und mit dem Laufband. Wir suchen Teilnehmer aller Altersklassen unabhängig von einem bestimmten Fitnesslevel oder Leistungsniveau.

Jeder Teilnehmer erhält eine Vergütung in Form eines Frottier Sporthandtuchs (50x100 cm) und eines 16 GB USB-Sticks.

Bei Interesse melden Sie sich bitte bei:

gregor.hackfort@unibw.de

Mit freundlichen Grüßen,

Dipl.-Sportwiss. Gregor Hackfort
Department für Sportwissenschaft
Universität der Bundeswehr München
Werner-Heisenberg-Weg 39
85577 Neubiberg
[REDACTED]

Abbildung D1: Wortlaut der E-Mail an die Mitarbeiter und Studenten der Universität der Bundeswehr München.

Anhang E: Deskriptive Statistik der einzelnen Items der psychischen Belastbarkeits- skala

Tabelle E1

Itemstatistik: Skala zur psychischen Belastbarkeit (N = 424).

Item	<i>M</i>	<i>SD</i>
Belastbarkeit Item 01	4,36	1,15
Belastbarkeit Item 02	4,33	1,09
Belastbarkeit Item 03	4,70	1,23
Belastbarkeit Item 04	4,27	1,02
Belastbarkeit Item 05	4,47	1,00
Belastbarkeit Item 06	4,53	1,25
Belastbarkeit Item 07	4,57	1,07
Belastbarkeit Item 08	4,00	1,01
Belastbarkeit Item 09	4,07	1,28
Belastbarkeit Item 10	4,20	1,14
Belastbarkeit Item 11	4,02	1,27
Belastbarkeit Item 12	4,02	1,31
Belastbarkeit Item 13	3,79	1,42
Emotionale Stabilität 01	3,68	1,26
Emotionale Stabilität 02	3,19	1,35
Emotionale Stabilität 03	3,45	1,30
Emotionale Stabilität 04	3,73	1,45
Emotionale Stabilität 05	3,81	1,33
Emotionale Stabilität 06	3,69	1,52
Emotionale Stabilität 07	4,24	1,29
Emotionale Stabilität 08	4,14	1,31
Emotionale Stabilität 09	4,74	1,14
Emotionale Stabilität 10	4,34	1,31
Emotionale Stabilität 11	4,11	1,21
Emotionale Stabilität 12	4,60	1,03
Emotionale Stabilität 13	4,44	1,12
Emotionale Stabilität 14	3,70	1,19
Emotionale Stabilität 15	4,27	1,25
Emotionale Stabilität 16	3,72	1,49

Anhang F: Häufigkeitsverteilung der 3000-m-FET-Zeiten

Tabelle F1

Tabellarische Darstellung der Häufigkeitsverteilung der Zeiten des 3000-m-FET ($N = 424$).

Zeiten (in Sek.)	absolute Häufigkeit	Prozente
≤ 220	13	3,1
221 - 235	34	8,0
236 - 250	93	22,0
251 - 265	129	30,5
266 - 280	75	17,7
281 - 295	35	8,3
296 - 310	18	4,3
311 - 325	10	2,4
≥ 326	16	3,8

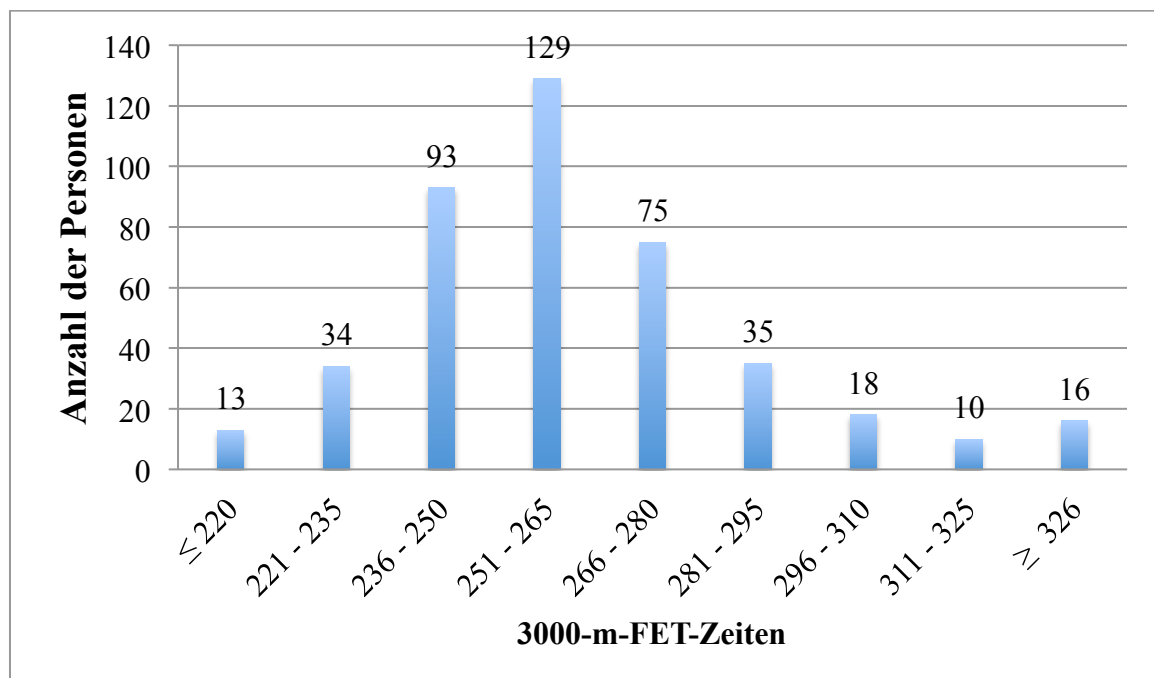


Abbildung F2: Häufigkeitsverteilung der 3000-m-FET-Zeiten ($N = 424$).

Anhang G: Deskriptive Statistik des maximalen Atemminutenvolumens (VE_{max})

Tabelle G1

VE_{max} -Werte in l/Min. zwischen Vor- und Nachtest ($n = 29$).

t1	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	74	97	116	117
<i>Max.</i>	206	196	210	208
<i>M</i>	147,66	140,93	158,93	157,10
<i>SD</i>	30,66	24,33	23,95	21,37
t2	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	101	97	113	114
<i>Max.</i>	226	193	209	201
<i>M</i>	155,34	145,93	158,59	155,17
<i>SD</i>	27,66	21,66	26,38	22,39

Tabelle G2

Deskriptive Statistik des VE_{max} in l/Min. ($N = 145$).

	3000-m-FET	1000-m-Lauf	Rampentest F.-Ergom.	Rampentest Laufband
<i>Min.</i>	56	58	52	63
<i>Max.</i>	206	196	210	212
<i>M</i>	137	133	140	141
<i>SD</i>	35,1	29,2	33,2	29,2

Anhang H: Häufigkeitsverteilungen der Itemscores der Persönlichkeitsskalen

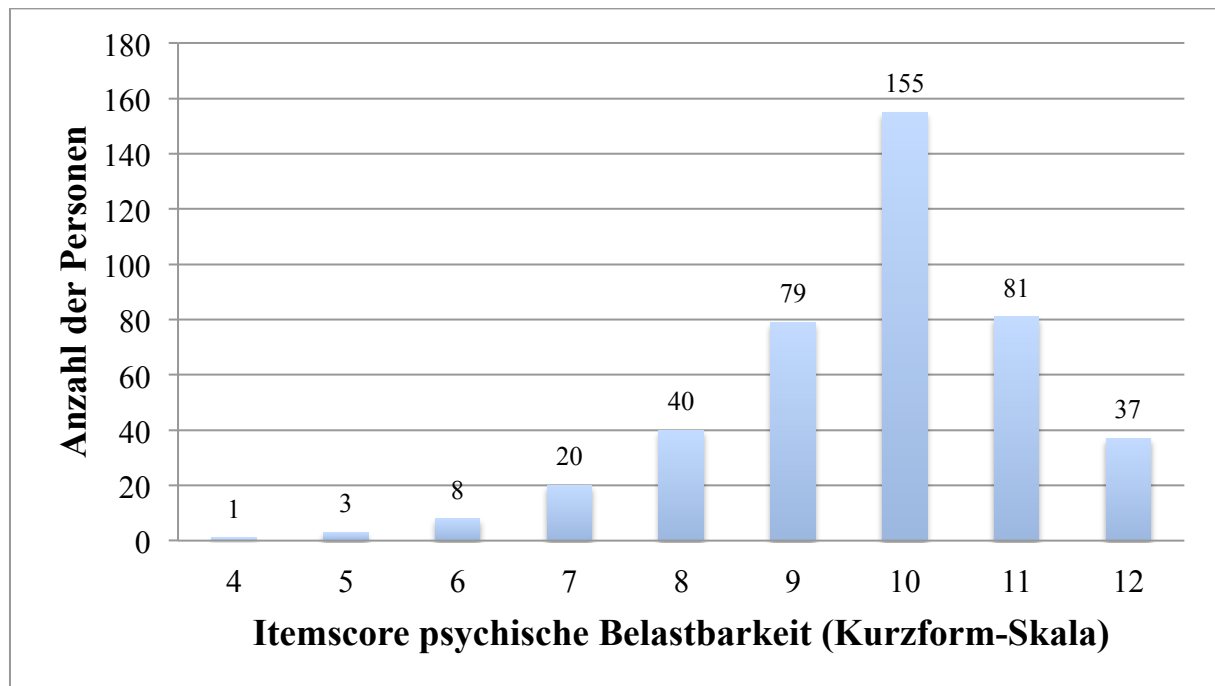


Abbildung H1: Häufigkeitsverteilung der Itemscores der psychischen Belastbarkeitsskala (Kurzform; N = 424).

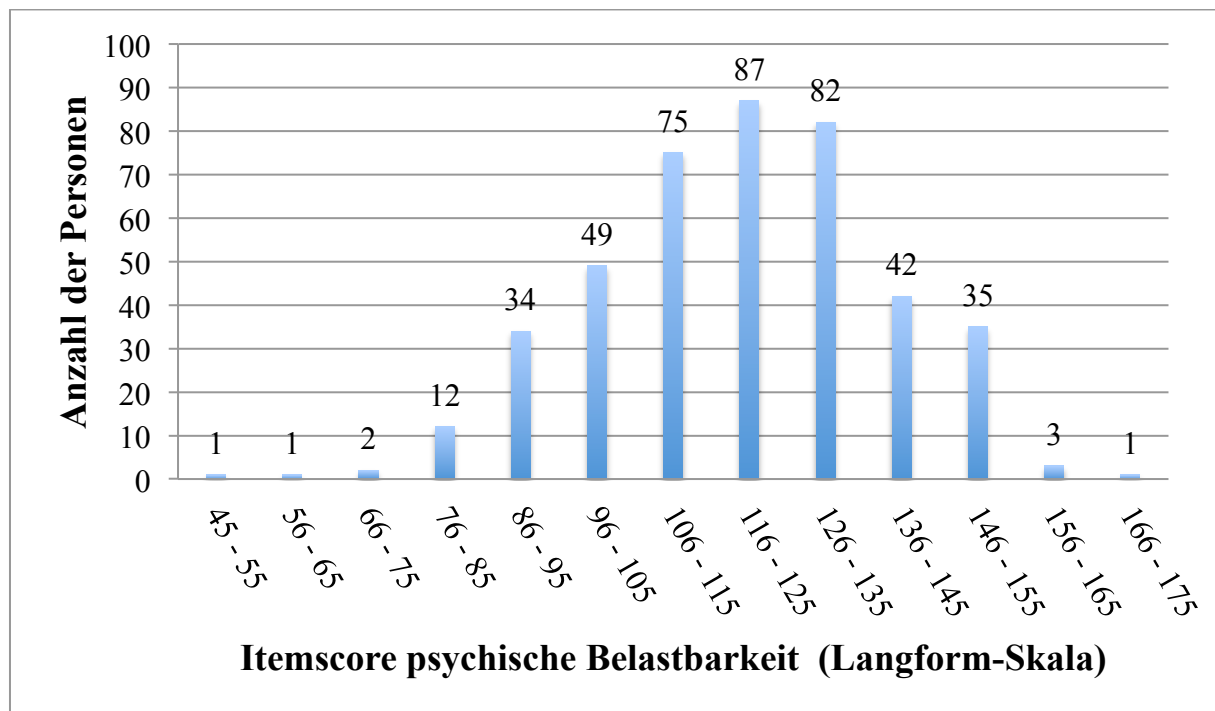


Abbildung H2: Häufigkeitsverteilung der Itemscores der psychischen Belastbarkeitsskala (Langform; N = 424).

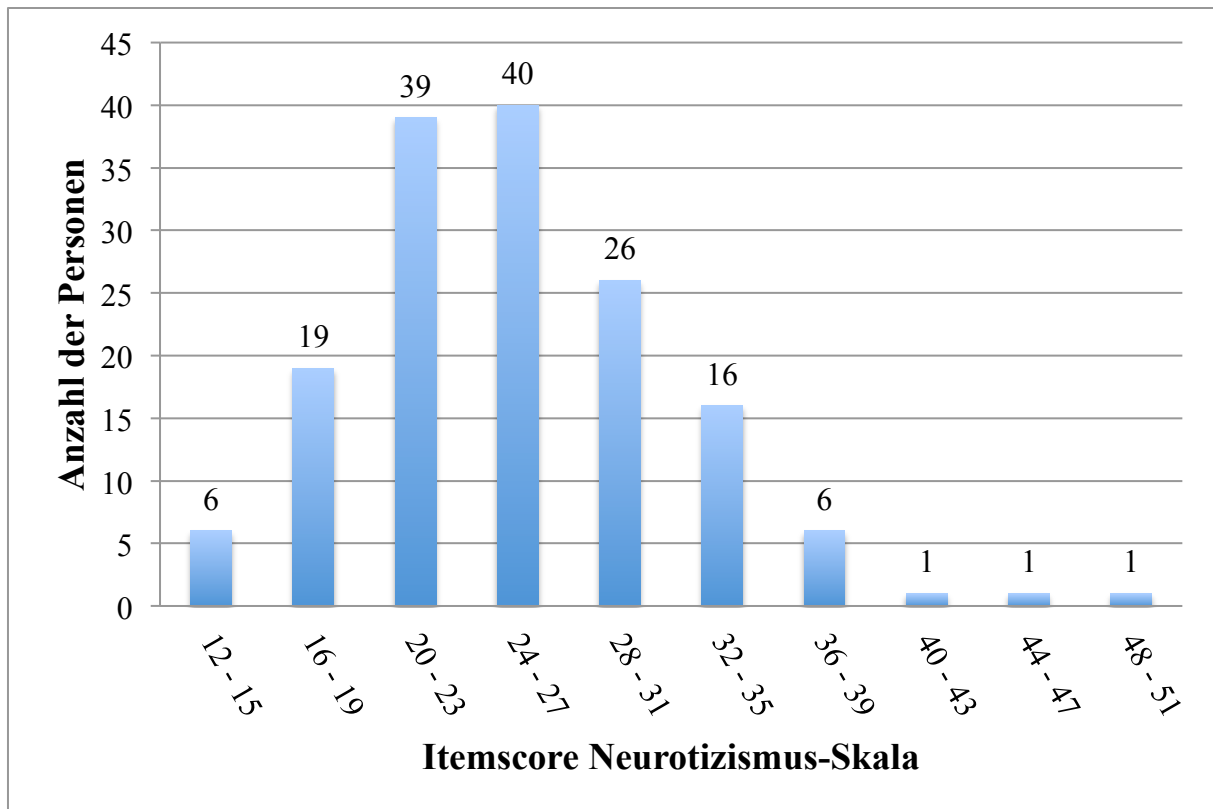


Abbildung H3: Häufigkeitsverteilung der Itemscores der Neurotizismus-Skala (N = 155).

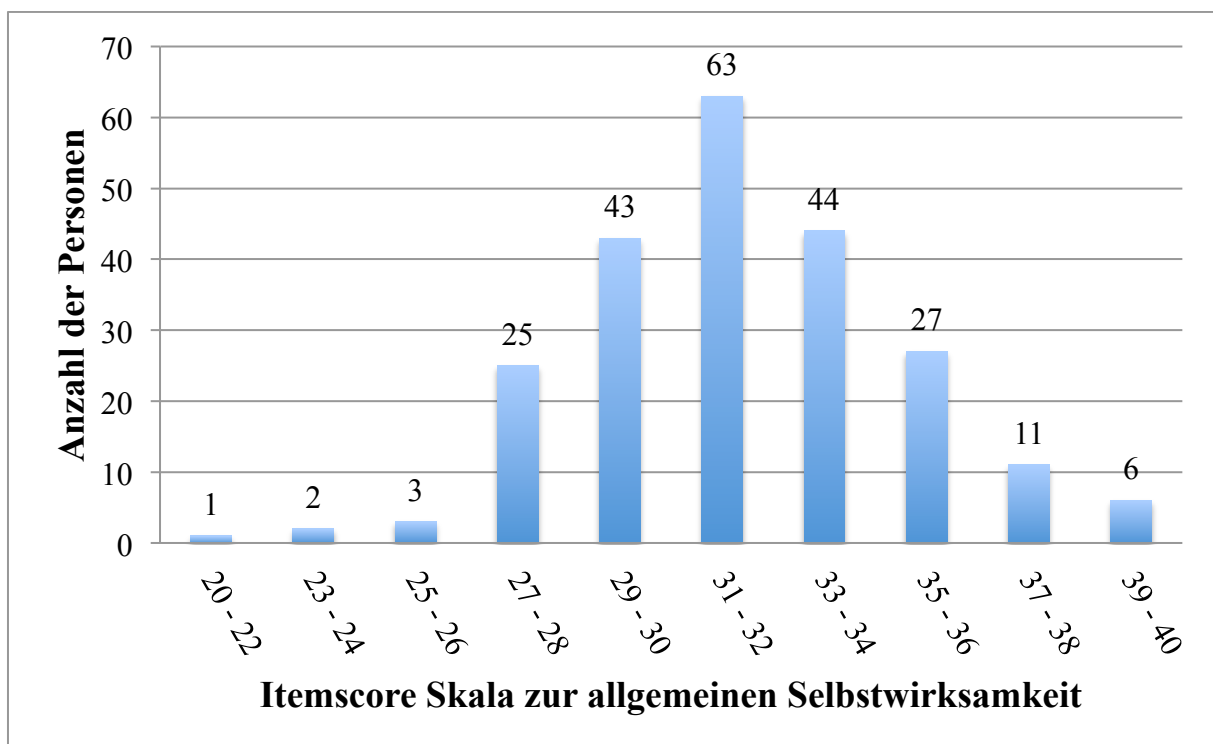


Abbildung H4: Häufigkeitsverteilung der Itemscores der Skala zur allgemeinen Selbstwirksamkeit (N = 225).

Abstract

Psychische Erkrankungen und Syndrome wie Burnout und Depressionen haben in der Krankheitsstatistik zugenommen, schränken die Produktivität am Arbeitsplatz ein, führen zugleich zu betriebs- und volkswirtschaftlichen Schäden und verursachen dem deutschen Gesundheitssystem immense Versorgungskosten. Sowohl physische und psychische Faktoren, als auch deren Zusammenhänge sind für die Gesundheit und das Wohlbefinden der Arbeitnehmer von besonderer Bedeutung. Das Hauptanliegen dieser Arbeit besteht darin, Wechselwirkungen bzw. Zusammenhänge zwischen physischen und psychischen Aspekten der Belastbarkeit zu analysieren und der Frage nachzugehen, inwieweit sich Personen mit differenziertem körperlichen Leistungsniveau hinsichtlich des Persönlichkeitsmerkmals Psychische Belastbarkeit voneinander unterscheiden. Für die Überprüfung der psychophysischen Belastbarkeit wurde zum einen ein Testverfahren zur Erfassung der Ausdauerleistungsfähigkeit (Fahrradergometer-Test) entwickelt und dieses im Rahmen einer Laborstudie mit einer Teilstichprobe anhand dreier weiterer Testverfahren (mit Einsatz einer Spiroergometrie) zur Leistungsdiagnostik validiert. Zum anderen wurden zur Überprüfung der psychischen Komponente zwei Skalen eines berufsbezogenen Persönlichkeitsinventars (emotionale Stabilität und Belastbarkeit) für die Operationalisierung der psychischen Belastbarkeit herangezogen, sowie eine daraus abgeleitete Kurzfassung für ein zeiteffizientes Screening konzipiert und eingesetzt. Im Rahmen einer Querschnittsuntersuchung wurden insgesamt 424 (356 männliche und 68 weibliche) Versuchsteilnehmer mit einem Durchschnittsalter von 29,1 ($SD=9,8$) Jahren auf die Ausdauerleistungsfähigkeit und anhand des Selbstbeschreibungsverfahrens zur psychischen Belastbarkeit hin untersucht. Die Ergebnisse zeigen, dass ein signifikanter Zusammenhang zwischen dem physischen Fitnessniveau, insbesondere hinsichtlich der Parameter Testzeit, VO_{2max}/VO_{2peak} , sowie der durchschnittlichen Wattleistung und dem Itemscore der psychischen Belastbarkeitsskalen besteht. Varianzanalytisch konnten signifikante Unterschiede zwischen Gruppen mit differenziertem Leistungsniveau in Hinblick auf die Itemscores der kombinierten Persönlichkeitsskalen emotionale Stabilität und Belastbarkeit nachgewiesen werden. Die Analysen zeigen insbesondere bei dem Vergleich der physisch Leistungsstärksten und weniger körperlich Leistungsfähigen hoch signifikante Unterschiede in Bezug auf die psychische Belastbarkeit. Die Befunde deuten darauf hin, dass physische Fitness sich auch auf psychische Faktoren auswirkt. Mit den hier erprobten Verfahren ist die Möglichkeit eines Assessments aufgezeigt worden, die psycho-physische Belastbarkeit in Form eines Screenings

zu erfassen. Dies ist im Kontext der Personalauswahl diskutiert worden. Der Ansatz ist als strategische Herangehensweise für ein Assessment psycho-physischer Aspekte der Belastbarkeit zu sehen und kann neben gesundheitlichen Belangen auch für die Personalauswahl und Personalentwicklung in Unternehmen, der Industrie und/oder im militärischen Bereich relevant sein.

Mental disorders and emotional syndromes such as burnout and depression has increased in recent years in statistics of absenteeism as well as costs and expenditures for healthcare in Germany. Factors of physical and mental resilience as well as their interrelation are of vital importance for the health and well-being of employees. The main objective of this thesis is to analyze the interaction and to measure correlations between indicators of physical and mental resilience and to find indices to what extent people with different individual physical performance levels differ with regard to the personality factor of mental resilience. For an assessment of these two factors, a special designed endurance performance test (cycling ergometer-test) as well as additional methods for endurance performance testing (partly with a spiroergometry) and two subscales of a standardized job related personality inventory (emotional stability and mental toughness) have been developed to operationalize the concept of psycho-physical resilience. Furthermore, a short version of the questionnaire was designed for a time-efficient screening of the personality factor mental resilience. In a cross-sectional study 424 (356 male and 68 female) participants with an average age of 29,1 ($SD=9,8$) years were tested in terms of the endurance performance level and the self-reported mental resilience. The results provided evidence that there is a significant relationship between the physical fitness-level, especially with respect to the parameters test-time, VO_{2max}/VO_{2peak} and the average power (Watt) and the item score of the mental resilience scale. The analysis of variance provide prove for significant differences between the performance groups depending on the item score of the combined scales emotional stability and mental toughness. A comparison of several groups within participants with various physical fitness levels unrevealed high significant differences especially in terms of individuals very high or very low in endurance performance. In addition, the findings demonstrate the influence of an active life-style on physical fitness and psychological factors. By such an assessment-strategy the possibility is given for a screening of psycho-physical resilience as it is discussed in the context of human resources and personnel selection. This approach can be regarded as a new strategy for an assessment of psycho-physical aspects of resilience and could be relevant for

various purposes of health related issues and employee acquisition or personnel development in fields, e.g., of business, industry and/or military.