

Psychologische Evaluation einer Gesundheitsförderung für ältere Langzeitarbeitslose

Inauguraldissertation

zur Erlangung des akademischen Grades

Doktor der Medizinwissenschaften (Dr. rer. hum.)

am Institut für Präventivmedizin
der Medizinischen Fakultät
der Universität Rostock

Rostock, 2011

vorgelegt von Markus Preuss
geb. am 04.02.1969 in Bad Neuenahr-Ahrweiler
aus Rostock

1. Gutachter:

Prof. Dr. med. Regina Stoll

Institut für Präventivmedizin, Universität Rostock

2. Gutachter:

Prof. Dr.-Ing. Günther Kundt

Institut für Biostatistik und Informatik in Medizin und Altersforschung, Universität Rostock

3. Gutachter:

Prof. Dr. phil. Erich Kasten

Institut für Medizinische Psychologie, Universität Göttingen

Datum der Einreichung: 14.07.2011

Datum der Verteidigung: 30.05.2012

Danksagung

Zuallererst möchte ich meiner lieben Familie danken, die mir viel Kraft und dem Leben neben der Dissertation Erfüllung gab sowie dem Ziel, die Arbeit durchzuführen und zum Abschluss zu bringen, einen unschätzbaren intrinsischen Wert verliehen hat. Ich danke meiner lieben Lebensgefährtin Dagmar und meinem lieben Sohn Aaron für ihre Zuneigung und ihre Unterstützung und auch für ihre Entbehrungen, die sie eingegangen sind, damit ich mich dieser Arbeit intensiv widmen konnte. Von Herzen möchte ich mich auch bei meiner Mutter Karin bedanken, die immer an mich glaubte und zu mir stand.

Frau Professor Regina Stoll danke ich dafür, dass sie mir dieses relevante Thema für meine Dissertation anvertraute, meine Arbeit anspornend und ermutigend betreute und dafür, dass sie mich als wissenschaftlichen Mitarbeiter über drei Jahre in ihr Forschungsteam aufnahm und den Rahmen der Studie "Fit50+" so verantwortungsvoll bereitstellte. Herrn Professor Erich Kasten danke ich für seine wissenschaftliche Beratung sowie für seine ausführliche, konstruktive Kritik an meiner Arbeit, die mich ausschlaggebend dazu motivierte, die Qualität der Dissertation auf das vorliegende Niveau anzuheben. Herrn Professor Kundt spreche ich meinen Dank für seine detaillierte statistische Beratung aus. Frau Dr. Steffi Kreuzfeld möchte ich neben ihrer Projektleitung vor allem für die motivierende und zugleich herzliche Zusammenarbeit vielmals danken sowie für ihre Einweihung in institutionelles und medizinisches Wissen zur Arbeitslosigkeitsforschung, das mir ohne sie verborgen geblieben wäre. Frau Dr. Dagmar Arndt danke ich für die produktive und inspirierende Zusammenarbeit, für ihre wertvolle Kritik an meiner Arbeit und dafür, dass sie mich in den Jahren immer wieder anregte, "über den Tellerrand zu blicken"; ich werde ihren Forschergeist vermissen. Frau Manuela Kukuk danke ich sehr für das Korrekturlesen dieser Arbeit und bei Herrn Dr. Thomas Klauer bedanke ich mich für die hilfreichen methodischen Tipps zur Bearbeitung der Studiendaten.

Ein Projekt wie die Studie "Fit50+" kann selbstverständlich nicht alleine durchgeführt werden, weswegen ich allen weiteren involvierten Kollegen meinen herzlichen Dank ausspreche. Namentlich nennen möchte ich Frau Wiedemann, Frau Sandra Engel, Herrn Dr. Matthias Weippert, Herrn Vowinkel und Frau Hempel. Auf diesem Wege möchte ich mich ebenso bei allen langzeiterwerbslosen Versuchsteilnehmern bedanken, die mir neben meinem wissenschaftlichen Interesse auch persönliche, menschliche Einblicke in die Thematik gewährten und ohne deren Teilnahme diese Studie überhaupt nicht möglich gewesen wäre.

Abkürzungsverzeichnis

-2LL	Likelihood-Funktion des negativen doppelten Wert des Logarithmus
ANOVA	Varianzanalyse
ARGE	Arbeitsgemeinschaft nach dem SGB II
BDI	Fragebogen "Beck-Depressions-Inventar"
χ^2	Chi-Quadrat Testkennwert
d	Cohen's d Effektstärke, standardisierter Mittelwertsunterschied
df	Freiheitsgrade
EG	Experimentalgruppe bzw. Interventionsgruppe
η^2	Eta-Quadrat, hier: Effektstärke der ANOVA für abhängige Stichproben
f	Effektstärke f , hier: Effektstärke der ANOVA unabhängige Stichproben
F	Testkennwert des F-Tests
FPI	Fragebogen "Freiburger Persönlichkeitsinventar"
GHQ	Fragebogen "General Health Questionnaire"
H_0	Nullhypothese
H_1	Alternativhypothese
IPM	Institut für Präventivmedizin, Medizinische Fakultät der Universität Rostock
K	Anzahl der Zwischensubjektfaktorgruppen der ANOVA mit Messwiederholung
KG	Kontrollgruppe
KSK	Körperliche Summenskala des SF12
L1, L2,...	Hypothesen zum Längsschnitt
LEB	Persönlichkeitsmerkmal "Lebenszufriedenheit" des FPI
M1, M2,...	Hypothesen zum interventionsmoderierenden Einfluss
MW	Mittelwert
N	Anzahl
p	Wahrscheinlichkeitswert (Signifikanzwert)
PSK	Psychische Summenskala des SF12
PSW	Fragebogen "Psychosoziales Wohlbefindenden"
Q	Anzahl der Innersubjektfaktorstufen der ANOVA mit Messwiederholung
Q1, Q2,...	Hypothesen zum Querschnitt
r	Korrelationskoeffizient nach Pearson
R^2	Quadriertes r (Varianzaufklärung)
Rho	Rangkorrelationskoeffizient nach Spearman
SD	Standardabweichung
SE	Standardfehler
SF12	Kurzversion des "Fragebogen zum Gesundheitszustand"
SOC	Fragebogen zum "Sense of coherence", Kohärenzsinn
SWE	Fragebogen zur "Allgemeinen Selbstwirksamkeitserwartung"
t	Testkennwert des t-Tests
T1	Erster Messzeitpunkt, vor der Intervention (prä)
T2	Zweiter Messzeitpunkt, unmittelbar nach Ende der Intervention (post)
T3	Dritter Messzeitpunkt, 6 Monate nach Ende der Intervention (follow-up)
T-Wert	T-transformierte Standardnorm (MW=50; SD=10)
VAS	"Visuelle Analogskala" zur Messung des subjektiven Gesundheitszustands
Vpn	Versuchsperson
z	z -transformierte Standardnorm (MW=0; SD=1)
Z	Testkennwert für nonparametrische Tests (z.B. Wilcoxon-Test, U-Test)

Inhaltsverzeichnis

1	Einleitung und empirischer Hintergrund	1
1.1	Existiert ein psychisches Erwerbslosigkeitssyndrom?	3
1.2	Psychologische Erklärungsansätze	4
1.3	Einflüsse auf die Folgen von Erwerbslosigkeit	7
1.4	Besonderheiten der Langzeitarbeitslosigkeit	12
1.5	Förderung von Erwerbslosen	15
1.6	Vorstellung von Interventionsstudien	16
1.6.1	Durchschnittlicher publizierter Interventionserfolg und Zeitstabilität	21
2	Methode	22
2.1	Förderung & Ethik	23
2.2	Design	23
2.3	Geplante Stichprobengröße	23
2.3.1	Verzicht auf Kontrollgruppe	23
2.3.2	Interventionsgruppen	24
2.3.3	Rekrutierung der Probanden	25
2.4	Interventionsprogramm	26
2.5	Follow-up	27
2.6	Untersuchungsablauf	28
2.7	Ebenen zur Erfassung der Gesundheit	28
2.7.1	Vorstellung der Messinstrumente	30
2.8	Ziele und Hypothesen zu den Querschnittsdaten	33
2.8.1	Querschnittshypothesen Q1, Q2, Q3	33
2.9	Ziele und Hypothesen zu den Längsschnittdaten	35
2.9.1	Risikoberechnung des Sportabbruchs	36
2.9.2	Längsschnitthypothesen L1, L2, L3, L4	37
2.9.3	Moderatorhypothesen M1, M2, M3, M4	39
2.9.4	Zusätzliche Ziele	41
2.10	Methodische Verfahren und Anmerkungen	41
2.10.1	Anmerkung zum unterschiedlichen N	41
2.10.2	Methode zum Test der Längsschnitthypothesen	42
2.10.2.1	Schema zur Prüfung der Längsschnitthypothesen L1-L4	43
2.10.2.2	Schema zur Prüfung der Moderatorhypothesen M1-M3	43
2.10.2.3	Methode zur Prüfung der Moderatorhypothese M4	44
2.10.2.4	Quantifizierung der Veränderungssensitivität der Instrumente	44
3	Querschnittsergebnisse	45
3.1	Mediennutzung	46
3.2	Erfassung der Gesundheit - Ableitung des Interventionsbedarfs	47
3.2.1	Q1: Test des Zusammenhangs von Antworttendenz und Gesundheit	50
3.2.2	Q2: Test des Geschlechtseffekts der psychischen Gesundheit	51
3.2.3	Q3: Test des geschlechtsspezifischen Effekts des Beziehungsstatus	51

4	Längsschnittergebnisse.....	53
4.1	Drop-Out-Analysen	53
4.1.1	Drop-Out-Analyse T2	54
4.1.2	Drop-Out-Analyse T3	55
4.2	Sportfortsetzeranalyse	56
4.2.1	T1 Modell: Sportfortsetzeranalyse mit Daten der Erstuntersuchung	56
4.2.2	Kombiniertes Modell: Sportfortsetzeranalyse mit Veränderungsdaten und Prädiktoren des T1-Modells	59
4.3	L1, L2, L3, L4: Tests der Längsschnitteffekte	60
4.3.1	Standardisierte Darstellung sämtlicher Längsschnitteffekte.....	62
4.3.2	Effektstärken der Interventionseffekte	63
4.4	M1, M2, M3, M4: Tests der Moderatoreffekte.....	66
4.4.1	Effektstärken der Moderatoreffekte	72
4.5	Veränderungssensitivität der Messinstrumente	72
4.5.1	Bestimmung der vergleichbaren Veränderungssensitivität	74
4.5.2	Optimale Stichprobenumfänge zum Erreichen von $1-\beta \geq 0,8$	75
4.5.3	Empfehlungen für Replikationsstudie mit N=100	76
5	Diskussion	77
5.1	Querschnittsergebnisse	77
5.2	Längsschnittergebnisse	81
5.2.1	Ergebnisse der Sportfortsetzeranalyse.....	81
5.2.2	Interventionsergebnisse der Längsschnittanalysen	83
5.2.3	Ergebnisse zur Moderation der Interventionseffekte.....	90
5.2.4	Fazit zum Interventionseffekt und Ausblick.....	95
6	Literatur.....	99
7	Anhang.....	109
7.1	Einwilligungserklärung	109
7.2	Soziodemographischer Fragebogen	110
7.3	GHQ (General Health Questionnaire "GHQ12")	112
7.4	BDI (Beck-Depressions-Inventar "BDI-V").....	113
7.5	SWE (Allgemeine Selbstwirksamkeitserwartung)	114
7.6	SOC (Sense of Coherence "SOC-L9")	115
7.7	PSW (Psychosoziales Wohlbefinden)	117
7.8	VAS (Visuelle Analogskala des Gesundheitszustandes des "EuroQOL").....	118
7.9	TV- und Medienkonsum	119
7.10	Zusätzliche Tabellen und Grafiken.....	120
8	Zusammenfassung der Dissertation.....	124

Abbildungsverzeichnis

Abb. 1.1: Effektstärken der psychischen Folgen von Arbeitslosigkeit.....	4
Abb. 1.2: Verbleib in Langzeitarbeitslosigkeit in Abhängigkeit vom Alter	9
Abb. 1.3: Psychische Beschwerden im Zusammenhang mit der Arbeitslosigkeitsdauer	11
Abb. 2.1: Darstellung der Studienphasen	23
Abb. 2.2: Darstellung der konzeptionellen Einordnung der Messgrößen.....	29
Abb. 3.1: Gesundheitsproblematische Anteile der Stichprobe	49
Abb. 4.1: z-standardisierte Veränderungen sämtlicher Längsschnittvariablen	62
Abb. 4.2: Darstellung der Interventions-Effektstärken.....	64
Abb. 4.3: Interaktionseffekt Kohorte * Innersubjektfaktor der SWE	66
Abb. 4.4: Interaktionseffekt Depressivität * Innersubjektfaktor des GHQ.....	68
Abb. 4.5: Interaktionseffekt Depressivität * Innersubjektfaktor des BDI.....	68
Abb. 4.6: Interaktionseffekt Depressivität * Innersubjektfaktor der PSK.....	69
Abb. 4.7: Interaktionseffekt Depressivität * Innersubjektfaktor des TV-Konsums	69
Abb. 4.8: Interaktionseffekt Geschlecht * Innersubjektfaktor der KSK.....	70
Abb. 4.9: Beispielhafte Darstellung zur Bestimmung des β -Fehlers.....	73
Abb. 4.10: Darstellung der Sensitivitäten der Innersubjektfaktor-Haupteffekte.....	74
Abb. 4.11: Visualisierung der Bestimmung optimaler Stichprobenumfänge für die Innersubjektfaktor-Haupteffekte	75
Abb. 5.1: Vergleich Effektstärken "Fit50+" mit Meta-Effektstärke von Paul und Moser	86
Abb. 7.1: T1: Durchschnittliche tägliche Nutzungsdauer für Unterhaltungsmedien	120
Abb. 7.2: Persönlichkeitsdimensionen des FPI.....	120
Abb. 7.3: T1-Ergebnisse des GHQ auf Itemebene.....	121
Abb. 7.4: T1-Ergebnisse des BDI auf Itemebene.....	121
Abb. 7.5: T1-Ergebnisse des SOC auf Itemebene	122
Abb. 7.6: T1-Ergebnisse der SWE auf Itemebene	122
Abb. 7.7: T1-Ergebnisse des PSW auf Itemebene.....	122
Abb. 7.8: Scree-Plot der Faktorenanalyse sämtlicher Variablen zu T1	123

Tabellenverzeichnis

Tab. 2.1: Inhalte der Gesundheitsedukation	26
Tab. 2.2: Inhalte des Trainingprogramms	27
Tab. 3.1: Soziodemographische Merkmale zwischen den Interventionskohorten	46
Tab. 3.2: Ableitung des Interventionsbedarfs.....	49
Tab. 3.3: Zusammenhänge der Messinstrumente mit FPI Merkmal "Offenheit"	50
Tab. 3.4: Gruppenunterschiede des Faktors "Geschlechts"	51
Tab. 3.5: Geschlechtsspezifische Gruppenunterschiede des Faktors "Beziehungsstatus"	52
Tab. 4.1: Gruppenunterschiede des Faktors "T2-Teilnahme"	55
Tab. 4.2: Gruppenunterschiede des Faktors "T3-Teilnahme"	55
Tab. 4.3: Gruppenunterschiede des Faktors "Sportfortsetzung" für T1-Variablen.....	56
Tab. 4.4: Ergebnisse der logistischen Regression des T1-Modells.....	57
Tab. 4.5: Beobachtete vs. vorhergesagte Anzahl von Sportfortsetzung und -abbruch	58
Tab. 4.6: Gruppenunterschiede des Faktors "Sportfortsetzung" für T2-T1-Variablen.....	59
Tab. 4.7: Ergebnisse der logistischen Regression des T2-T1-Modells.....	60
Tab. 4.8: Beobachtete vs. vorhergesagte Anzahl von Sportfortsetzung und -abbruch	60
Tab. 4.9: Tests der Längsschnitthypothesen L1, L2, L3, L4.....	61
Tab. 4.10: Effektstärken und Power der Interventionseffekte.....	64
Tab. 4.11: Effektstärke Cohen's d der Interventionseffekte.....	65
Tab. 4.12: Tests der Moderatorhypothese M1	66
Tab. 4.13: Tests der Moderatorhypothesen M2	67
Tab. 4.14: Tests der Moderatorhypothesen M3	70
Tab. 4.15: Tests der follow-up Moderatorhypothese M4	71
Tab. 4.16: Effektstärke und Power der Moderatoreffekte.....	72
Tab. 4.17: Power der Innersubjektfaktor-Haupteffekte.....	73
Tab. 4.18: Optimale Stichprobenumfänge zum Erreichen von $1-\beta \geq 0,8$	75
Tab. 7.1: Täglicher Konsum von Unterhaltungsmedien zu T1.....	120
Tab. 7.2: Prozentuale Antworthäufigkeiten auf Itemebene des SF12 zu T1	123
Tab. 7.3: Ladungen der Faktorenanalyse zu T1	123

1 Einleitung und empirischer Hintergrund

Es war ein Donnerstag, der 24. Oktober 1929, an dem sich an den US-amerikanischen Aktienmärkten das Platzen einer langjährigen Spekulationsblase in massiven Kurseinbrüchen entlud. Dies führte zur Panik unter den Anlegern und zum bisher folgenreichsten Börsencrash der Vergangenheit. Dieser Tag ging als "schwarzer Donnerstag" in die Geschichtsbücher ein. Durch die damals bereits international vernetzte Wirtschaft inklusive ihrer Kapitalströme blieben die Folgen des Crashes nicht auf die USA begrenzt. Stattdessen führten die anschließenden wirtschaftlichen Störungen in fast sämtlichen westlichen Ökonomien nahezu gleichzeitig zur ersten Weltwirtschaftskrise, welche die sog. "Goldenen Zwanziger" abrupt beendete. In der Weimarer Republik führte die ökonomische Krise zu kaskadenartigen Unternehmenszusammenbrüchen; es kam zum erstmaligen Auftreten der Massenarbeitslosigkeit. Diese erreichte im Jahr 1932 ihren Höhepunkt: In der deutschen Bevölkerung standen 6,1 Mio. Arbeitslosen nur noch 12 Mio. Beschäftigte gegenüber.

Die "Textilfabrik Marienthal" in Österreich war eines der Unternehmen, das von der damaligen Weltwirtschaftskrise getroffen wurde. Die Fabrik, ca. 100 Jahre zuvor gegründet, war gleichzeitig namensgebend für die Arbeitersiedlung, die sich im Laufe des Jahrhunderts um sie entwickelte. Als das Werk infolge der Weltwirtschaftskrise im Februar 1930 schließen musste, führte dies in der Gemeinde Marienthal fast zur kompletten Erwerbslosigkeit der dort lebenden Bürger. Das Schicksal der Ortschaft bot der Forschergruppe um Marie Jahoda, Paul Lazarsfeld und Hans Zeisel die Gelegenheit für eine Feldforschung, in welcher sie die sozio-psychologischen Auswirkungen der dort abrupt auftauchenden Erwerbslosigkeit erstmalig systematisch erfassten (Jahoda, Lazarsfeld, & Zeisel, 1933). Die Forscher begleiteten die Erwerbslosen von Marienthal im Längsschnitt mit Interviews, Fragebögen, strukturierten Verhaltensbeobachtungen und Haushaltserhebungen. Ihre wichtigste Erkenntnis war, dass die Erwerbslosigkeit nicht - wie man bis dahin annahm - zu einer aktiven Revolte der Betroffenen führte, sondern eine passive Resignation bis Apathie auslöste. Jahoda et al. (1933) dokumentierten erstmalig, dass das Auftreten von Erwerbslosigkeit auf der Ebene der Betroffenen nicht nur zu existentiellen finanziellen Einbußen führt, sondern sich in erheblichen Begleitschäden in den sozialen und psychischen Lebensbedingungen manifestiert.

Hollederer (2009) schreibt, dass es "den Arbeitslosen" als Typus genau so wenig gibt wie "den Beschäftigten". Erwerbslose bilden keine homogene Gruppe, sondern setzen sich aus Personen mit unterschiedlichen Hintergründen und Verweildauern in der Arbeitslosigkeit zusammen. Bereits die Anlässe der Arbeitslosigkeit lassen sich unterscheiden (Hradil & Schiener, 2001): Z.B. kommt es wetterbedingt in den Wintermonaten zum Auftreten der sog. *saisonalen Arbeitslosigkeit*. Je nach Zähltag ergeben sich also erhebliche Periodeneffekte. Demgegenüber führten die Einflüsse des gesellschaftlichen Strukturwandels von der Agrar-, zur Industrie-, Dienstleistungs- und Informationsgesellschaft zur sog. *strukturellen Arbeitslosigkeit*. Diesem Arbeitslosigkeitsanlass kann man im Hinblick auf das Qualifikationsniveau unterstellen, dass er selektiv bestimmte Personengruppen trifft. So fallen durch Automatisierungs- und Globalisierungsprozesse stetig Stellen für Arbeitnehmer weg.

Die Anlässe von Erwerbslosigkeit bedingen demnach einerseits eine Fluktuation und andererseits eine Segmentierung des Arbeitsmarktes. Im internationalen Vergleich hat Deutschland eine hohe sog. "Sockelarbeitslosigkeit". Diese ist geprägt durch einen großen Anteil an schwierig vermittelbaren, langjährig Erwerbslosen, welche meist folgende Arbeitslosigkeitsrisiken aufweisen: höheres Alter, geringe Qualifikationen und gesundheitliche Einschränkungen. Leider zeigte sich in den letzten Jahrzehnten, dass der "harte Kern" von Langzeiterwerbslosen selbst in Zeiten eines Konjunkturaufschwungs (z.B. Boom der IT-Branche bis 2001 oder das Wirtschaftswachstum von 2005-2007) nur sehr begrenzt abgebaut werden konnte, so dass die Sockelarbeitslosigkeit seit 1973, dem Beginn der ersten Ölkrise, kontinuierlich anstieg (Hollederer, 2009; Wiggert, 2009).

Die Arbeitsvermittler der Agenturen für Arbeit bezeichnen die Risiken der Arbeitslosigkeit aus ihrer Perspektive als "Vermittlungshemmnisse". Hollederer (2009) berechnete, dass nach dem Categorieschema der Arbeitsvermittler der Anteil der Erwerbslosen mit einem "gesundheitlichen Vermittlungshemmnis" im Jahr 2005 bei 24% im steuerfinanzierten Rechtskreis des SGB II¹ und bei 19% im SGB III²-Bereich lag. Zahlreiche Studien belegen, dass Erwerbslose im Vergleich zu Beschäftigten einen schlechteren Gesundheitszustand aufweisen (Überblicke z.B. Hollederer & Brand, 2006; Kieselbach, Winefield, & Boyd, 2006). Wie Reviews und diverse epidemiologische Surveys zur Thematik (z.B. Bundesgesundheits-survey, Mikrozensus etc.) sowie Auswertungen von Krankenkassen übereinstimmend berichten, liegt im Vergleich zu Beschäftigten das Krankheitsrisiko von Arbeitslosen deutlich höher und steigt zudem mit der Dauer der Erwerbslosigkeit an. So werden z.B. eine höhere Mortalität, mehr psychische Störungen, mehr Inanspruchnahme gesundheitlicher Leistungen (Arzneimittelverordnungen und Krankenhausbehandlungen), mehr Übergewicht und Adipositas, ein ungünstigeres Gesundheitsverhalten sowie ein erhöhter Suchtmittelkonsum (v.a. Alkohol- und Nikotinabusus) festgestellt (Hollederer, 2009). Für Erwerbslose mit gesundheitlichen Einschränkungen besteht definitiv eine verminderte Übergangswahrscheinlichkeit in den Arbeitsmarkt. Selbst in multivariaten Analysen mit Kontrolle von Alter, Geschlecht, Ausbildungsniveau und Arbeitslosigkeitsdauer spielt die Gesundheit eine entscheidende Rolle für die Wiedervermittlung (Cramer et al., 2002).

Während einerseits das höhere Morbiditätsrisiko von Erwerbslosen durch eine große Vielfalt von Erkrankungen gekennzeichnet sein mag, zeigen sich andererseits besonders in den psychischen Erkrankungen wesentlich höhere Quoten als bei den Berufstätigen. Bereits Jahoda et al. (1933) beschreiben für die *Arbeitslosen von Marienthal* sich längsschnittlich entwickelnde Symptome von Depressivität. Innerhalb der "sächsischen Längsschnittstudie zu Arbeitslosigkeitserfahrungen" berechneten Berth, Förster, Balck, Brähler und Stöbel-Richter (2008), dass die Variable "aktuelle Arbeitslosigkeit" mit einer Odds-Ratio von 7,1 den stärksten Prädiktor der psychischen Beanspruchung darstellte. Bereits die Variable "Bedrohung durch Arbeitsplatzverlust" ging mit einer Odds-Ratio von 5,3 mit einem erhöhten psychischen Beschwerdedruck einher. Laut Bundesverband der BKK stellen für Erwerbslose psychische Störungen die häufigste Ursache von stationären Krankenhausaufenthalten dar (BKK, 2009). Eine Auswertung der Techniker Krankenkasse ergab, dass bei erwerbslosen

¹ Zweites Sozialgesetzbuch zur Grundsicherung für Arbeitsuchende.

² Drittes Sozialgesetzbuch zur Arbeitsförderung.

Versicherten dreimal häufiger psychische Störungen diagnostiziert werden als bei Berufstätigen. Zudem werden Arbeitslose aufgrund von psychischen Beschwerden ca. 4-mal länger krank geschrieben als Erwerbstätige (TK, 2010). Für Arzneimittelverordnungen ist die Wahrscheinlichkeit der Verschreibung von Antidepressiva für Erwerbslose im Vergleich zu Berufstätigen fast 3-mal höher (Rothländer, Mühlpfordt, & Richter, 2012).

Aufgrund der Vielzahl der Quellen zur Thematik ließe sich die Liste von gesundheitlichen Unterschieden zwischen Arbeitslosen und Erwerbstätigen noch fortsetzen. Die bisher aufgeführten sollten jedoch als Argument dafür ausreichend sein, dass zumindest die Koinzidenz zwischen Arbeitslosigkeit und Krankheit mittlerweile sehr gut belegt ist. Demnach kann geschlussfolgert werden, dass die gesellschaftlichen Kosten der Erwerbslosigkeit weit aus höher liegen als die Summe der Transferleistungen im Rahmen des ALG I & II. Die Kosten der Arbeitslosigkeit werden nach Wiggert (2009) in volkswirtschaftliche, psychosoziale und Opportunitätskosten differenziert. Wie die Kausalkette zwischen den Dimensionen Erwerbslosigkeit und Krankheit aussieht, ist für die psychischen Störungen durch umfassende Meta-Analysen von Längsschnittdaten (Paul & Moser, 2001; Paul, Hassel, & Moser, 2006; Paul & Moser, 2009a) mittlerweile wesentlich besser geklärt als noch vor 10 Jahren. Trotzdem bleibt die Kausaldebatte um die Symptome der Erwerbslosigkeit weiterhin Gegenstand der Diskussion um die gesellschaftliche Dimension der Arbeitslosigkeit. Diese ist bei weitem nicht nur wissenschaftlich, sondern - vor allem in Zeiten des Wahlkampfes - politisch motiviert und wird entsprechend oft auch emotional geführt.

1.1 Existiert ein psychisches Erwerbslosigkeitssyndrom?

Paul et al. (2006) metaanalysierten 237 Querschnittsstudien, in denen Arbeitslose und Erwerbstätige im Hinblick auf ihr Befinden verglichen wurden. Die Studien stammten aus 26 westlichen Ländern und wurden zwischen 1963 und 2004 veröffentlicht. Die Metaanalytiker resümieren, dass ein breites Spektrum an Symptomen psychischer Beschwerden zur Erwerbslosigkeit in den Primärstudien angeführt wurde. In ihrer Analyse beschränkten sich die Autoren auf 6 in den Primärstudien häufig untersuchten Variablen der psychischen Gesundheit:

1. Unspezifische psych. Beschwerden (z.B. General Health Questionnaire: Goldberg, 1978)
2. Depressivität (z.B. Beck-Depressions-Inventar: Beck, Steer, & Hautzinger, 2005)
3. Angstsymptome (z.B. State Trait Anxiety Inventory: Spielberger, 1983)
4. Psychosomatische Symptome (z.B. Symptom Checkliste: Derogatis & Cleary, 1977)
5. Subjektives Wohlbefinden (z.B. Lebenszufriedenheit des FPI-R: Fahrenberg, et al., 2010)
6. Selbstwertgefühl (z.B. Self Esteem Scale: Rosenberg, 1965)

Paul et al. (2006) stellten fest, dass die 6 Variablen in deutlichem statistischen Zusammenhang standen und interpretieren, dass sie als Indikatoren für das abstrakte Konstrukt "Psychische Gesundheit" aufgefasst werden können. Die Metaanalyse legte einen klaren Zusammenhang offen: Erwerbslose zeigten im Vergleich deutlich mehr Symptome psychischer Beeinträchtigung und geschwälerten Wohlbefindens als Berufstätige. Die Unterschiede zwischen Arbeitslosen und Erwerbstätigen geben die Verfasser mit der Effektstärke *Cohen's d* an (Cohen, 1988) (vgl. folgende Abb.).

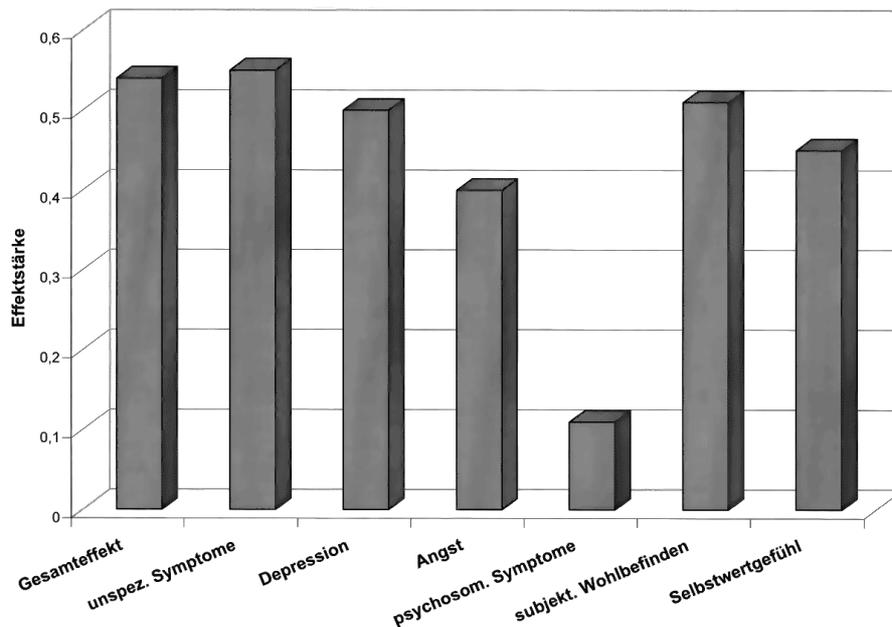


Abb. 1.1: Effektstärken der psychischen Folgen von Arbeitslosigkeit (Paul, et al., 2006, S. 42).

Mit $d=0,54$ war der Gesamteffekt von mittlerer Größe, wie auch die meisten Subskalen. Die psychosomatischen Symptome zeigten mit $d=0,11$ den kleinsten aller signifikanten Unterschiedseffekte. Da Effektstärken schwierig zu interpretierende Größen darstellen, metaanalysierten Paul et al. (2006) den Prozentsatz der Fälle von krankheitsrelevanten psychischen Störungen. In dieser Analyse benutzten sie ausschließlich etablierte klinische Screening-Tests für die Symptome einer depressiven Episode oder einer Angststörung. Für Erwerbslose ergab sich ein Anteil von 34%, während für Erwerbstätige ein Anteil von 16% berechnet wurde. Demnach war der Anteil derer, die psychisch so beeinträchtigt waren, dass sie möglicherweise psychotherapeutische oder psychiatrische Behandlungen benötigen, unter den Erwerbslosen mehr als doppelt so hoch wie unter den Berufstätigen.

1.2 Psychologische Erklärungsansätze

Um den Zusammenhang von seelischen Problemen und Arbeitslosigkeit zu erklären, wurden laut Paul et al. (2006) in der psychologischen Forschung die *Stress-Theorie* nach Lazarus und Alfert (1964) oder die *Theorie der erlernten Hilflosigkeit* nach Seligman (1978) oder die *Theorie der kognitiven Dissonanz* nach Festinger (2001) herangezogen. Fünf, mehr spezifische Erklärungsansätze haben sich darüber hinaus als einflussreich erwiesen:

Die **Deprivationstheorie** von Jahoda (1982, 1992) postuliert, dass die Erwerbsarbeit nicht nur die Funktion des Gelderwerbs erfüllt (*manifeste Funktion* von Arbeit), sondern zusätzlich fünf weitere sog. *latente Funktionen* bereitstellt: 1.) Die Auferlegung einer festen Zeitstruktur, 2.) Sozialkontakt, 3.) Status und Identität, 4.) die Teilhabe an kollektiven Zielen und 5.) regelmäßige Aktivität. Jahoda vertritt die Meinung, dass die latenten Funktionen mit basalen menschlichen Bedürfnissen korrespondieren. Die Möglichkeiten zur Inanspruchnahme der manifesten und der latenten Funktionen beeinflussen nach Jahodas Ansatz die psychische Gesundheit. Da Erwerbslose nur eingeschränkt Zugang zu den Funktionen haben, sind sie depriviert, welches zur Entwicklung von psychischen Beschwerden führt.

Vergleichbar mit Jahodas Ansatz geht auch das **Vitamin-Modell** von Warr und Jackson (1985) davon aus, dass externe Faktoren die psychische Instabilität des Erwerbslosen verantworten. Die Autoren messen der Erwerbsarbeit folgende Funktionen zu: 1.) Möglichkeit zur Ausübung von Kontrolle, 2.) Möglichkeiten zur Nutzung der eigenen Fähigkeiten, 3.) von außen gesetzte Ziele, 4.) Abwechslung, 5.) Klarheit und Eindeutigkeit der Umwelt, 6.) Verfügbarkeit von Geld, 7.) physische Sicherheit, 8.) Möglichkeit zu interpersonellem Kontakt und 9.) eine anerkannte soziale Position. Warr und Jackson (1985) hypothetisieren, dass die 9 Funktionen die psychische Gesundheit analog zur Wirkweise von Vitaminen beeinflussen. Niedrige Ausprägungen oder das Fehlen dieser 9 "Vitamine" führen zu psychischen Beschwerden. Höhere Ausprägungen der "Vitamine" führen zu positiven psychischen Resultaten, allerdings nur bis zu einem bestimmten Schwellenwert, danach könne eine weitere Erhöhung - wie im Fall von *Abwechslung* - sogar schädlich sein, behaupten die Autoren. Da Erwerbslose mit Umweltbedingungen konfrontiert sind, die nur wenige dieser "Vitamine" bereitstellen, führt dies zu psychischen Beschwerden. An dem Modell fällt auf, dass *Verfügbarkeit von Geld* erst an 6. Stelle genannt wird, mithin nur ein Punkt unter vielen ist und keinen Schwerpunkt, wie bei Jahoda, darstellt.

Das **Handlungsrestriktionsmodell** wurde als Reaktion auf Jahodas Theorie von Fryer (1986) entworfen und distanziert sich von dem situationszentrierten und somit deterministischen Menschenbild, welches den obigen Ansätzen zugrunde liegt. Fryer sieht vornehmlich den eingeschränkten Handlungsspielraum und die Unplanbarkeit der Lebenssituation von Erwerbslosen als ursächlich für deren psychisches Leid. Der Autor geht vom Konzept eines selbständig planenden und handelnden Individuums aus, das ein Bedürfnis nach Kontrolle hat und intrinsisch motiviert ist. Zum einen wird im Rahmen der Erwerbslosigkeit durch schwindende finanzielle Mittel (bis zur Verarmung) das Bedürfnis nach Selbstbestimmung frustriert. Bestimmte Rollen, wie z.B. die des Familienernährers, können nicht mehr eingenommen werden und Gefühle von Ohnmacht stellen sich ein. Da eigene Pläne als nicht mehr realisierbar empfunden werden, unterbleibt im Laufe der Arbeitslosigkeitsdauer das Setzen von Zielen. Zum anderen weiß der Arbeitslose nicht, wann er eine neue Stelle bekommen wird, wodurch die zeitliche Perspektive der Situation unklar bleibt. Weil die Zukunft nicht berechenbar ist, wird sie vom Betroffenen als bedrohlich wahrgenommen. Unplanbare und eingeschränkte Handlungsmöglichkeiten führen schließlich zu psychischen Problemen. Im Gegensatz zu Jahodas sowie Warr und Jacksons Überlegungen wird die Rolle einer eingeschränkten finanziellen Autonomie durch Fryer besonders betont.

Die **Inkongruenzhypothese** (Paul & Moser, 2006) fußt auf der empirischen Beobachtung, dass selbst unter Arbeitslosen eine starke innere Bindung und positive Einstellung gegenüber der Erwerbsarbeit ("employment commitment") festgestellt werden kann (vgl. auch Mohr, 2001). Die Arbeitsethik von Erwerbslosen und Berufstätigen unterscheidet sich kaum; beide Gruppen zeigen eine hohe Wertschätzung der Erwerbsarbeit, welche selbst im Verlauf der Langzeiterwerbslosigkeit erstaunlich stabil bleibt. Arbeitslose befinden sich daher dauerhaft in einer Situation mangelnder Passung ihrer Einstellungen zur Erwerbsarbeit mit ihrer realen Erwerbssituation. Dies erzeugt eine kognitive Dissonanz (Festinger, 2001). Situationen, die durch eine Dissonanz von Zielen und Einstellungen einerseits und den realen Gegebenheiten andererseits gekennzeichnet sind, werden in diversen psychologischen

Erklärungsansätzen als Ursache psychopathologischer Symptome angesehen (Grawe, 2004).

Die **Selektionshypothese** (z.B. Mastekaasa, 1996) postuliert, dass der Zusammenhang zwischen Arbeitslosigkeit und eingeschränkter psychischer Gesundheit schlicht daraus resultiert, dass Personen mit psychischen Problemen mit höherer Wahrscheinlichkeit ihren Arbeitsplatz verlieren und anschließend länger brauchen, eine neue Beschäftigung zu finden. Und selbst falls sie eine neue Stelle erhalten, haben sie durch ihre psychischen Beschwerden eine erneut höhere Wahrscheinlichkeit, wieder erwerbslos zu werden, so dass Menschen mit eingeschränkter psychischer Gesundheit im Laufe von mehreren Iterationen immer häufiger in die Erwerbslosigkeit samt längeren Verweildauern und schlussendlich in die Langzeitarbeitslosigkeit hineinselektiert werden.

Kausalhypothese vs. Selektionshypothese

Die Deprivationstheorie, das Vitaminmodell, das Handlungsrestriktionsmodell und die Inkongruenzhypothese sehen die Erwerbslosigkeit als kausale Ursache des negativen psychischen Befindens und sind somit als Kausalhypothesen zu bezeichnen, während die Selektionshypothese von einer umgekehrten Verursachung in dem Sinne ausgeht, dass psychische Beschwerden zur Arbeitslosigkeit führen. Die Debatte darüber, welcher Erklärungsklasse die größere Bedeutung zugeschrieben werden sollte, ist auch heute noch Gegenstand der Diskussion.

Wie bereits angedeutet, ist diese Diskussion mitnichten nur wissenschaftlich motiviert. Einerseits stellt die Selektionshypothese nach Mastekaasa (1996) zum Zusammenhang von Erwerbslosigkeit und psychischen Störungen eine plausible und zugleich simple Erklärung bereit, die das Sparsamkeitsprinzip einer Theorie nach "Okhams Rasiermesser" (z.B. Hamilton, 2010) am besten erfüllt. Andererseits könnte man dem Ansatz unterstellen, dass ihm eine politische Funktion innewohnt, weil er die Legislative von der Verantwortung für die Folgen der Erwerbslosigkeit zu entbinden vermag. Kausalhypothesen implizieren durch die Identifizierung der Erwerbslosigkeit als Ursache für psychische Störungen eine Haftung des wirtschaftlichen und politischen Systems, während die Selektionshypothese die Verantwortung auf die Betroffenen abzuwälzen vermag. Allerdings müssen sich Kausal- und Selektionseffekte, so konträr wie sie auf den ersten Blick wirken, nicht widersprechen. Es ist annehmbar, dass sie im Rahmen eines Zusammenspiels die Ebenen Arbeitslosigkeit und Gesundheit verknüpfen. Speziell für Langzeiterwerbslose wird von einigen Wissenschaftlern eine Zunahme von Selektionseffekten vermutet (z.B. Claussen, 1994; Zempel & Frese, 2000; Mohr, 2001). Jedoch beziehen die Autoren Selektionsprozesse bei langjährig Erwerbslosen nicht vornehmlich auf die psychische Gesundheit, sondern mehr auf das Qualifikationsniveau und das Alter der Betroffenen.

Beim metaanalytischen Test von Kausal- vs. Selektionshypothese fanden Paul et al. (2006) sowie Paul und Moser (2009a) mithilfe von 87 Längsschnittstudien, in denen Personen zwischen dem ersten und dem zweiten Erhebungszeitpunkt einen Wechsel im Erwerbsstatus durchmachten, empirische Belege für beide Erklärungsklassen. So kalkulierten sie beim Verlust eines Arbeitsplatzes $d=-0,19$ für die Verschlechterung des Befindens, beim Erhalt einer neuen Stelle $d=0,35$ für die Verbesserung der psychischen Gesundheit. Beide

Befunde stützen die Kausalhypothese. Selektionseffekte wurden mit $d=0,23$ für ein a priori schlechteres Befinden von Berufstätigen, die ihre Stelle im Längsschnitt einbüßten, festgestellt. Kontinuierlich Arbeitslose wiesen zu T1 ein um $d=0,15$ schlechteres Befinden auf, als die Erwerbslosen, welche im Verlauf eine neue Stelle fanden. Zu ihren berechneten Selektionseffekten bemerken Paul et al. allerdings selbstkritisch, dass es trotzdem unklar bleibe, ob die gefundenen Befindensunterschiede des 1. Messzeitpunkts tatsächlich Selektionseffekte widerspiegeln: Antizipationseffekte wären ihrer Meinung nach ebenfalls eine mögliche Erklärung. Deshalb empfehlen die Autoren, dass ihre Effektstärken nur als Obergrenzen möglicherweise existierender Selektionseffekte auf die psychische Gesundheit interpretiert werden sollen (Paul et al., 2006).

1.3 Einflüsse auf die Folgen von Erwerbslosigkeit

In der Erwerbslosigkeitsforschung hat die Suche nach soziodemographischen Randbedingungen, welche die negativen Folgen der Erwerbslosigkeit möglicherweise verstärken oder abschwächen, eine lange Tradition. In Anlehnung an Paul und Moser (2009b) werden im Folgenden potentielle Einflussgrößen inklusive empirischer Befunde vorgestellt.

Einfluss des Geschlechts

Vor allem in älteren Publikationen ging man davon aus, dass Männer stärker unter der Erwerbslosigkeit leiden müssten als Frauen. Die Begründung dazu stützt sich im Wesentlichen auf zwei Argumente: Zum einen wird behauptet, dass Männer Sinn und Selbstwert in der Gesellschaft fast ausschließlich durch Erwerbsarbeit erzielen können (Jahoda, 1983). Demnach leide der Mann besonders unter der Arbeitslosigkeit, weil er in seinen Augen bei der Erfüllung seiner zentralen Lebensaufgabe - der Rolle als Familienernährer - versage (Komarovsky, 1940). Zum anderen wurde für die Frau vermutet, dass sie eine schwächere innere Bindung an die Erwerbsarbeit habe, weil ihr die gesellschaftlich akzeptierte Alternativrolle als Hausfrau zur Verfügung stünde, die Ersatz für eine verlorene Erwerbsarbeit böte (Shamir, 1985). Ergänzend konstatiert Shamir, dass das mit der Erwerbslosigkeit verbundene soziale Stigma für den Mann ausgeprägter sei, da sein soziales Umfeld vornehmlich von ihm die Übernahme der Rolle als Familienernährer erwarte. Des Weiteren wurde angeführt, dass das Haushaltseinkommen von Paaren durchschnittlich stärker durch die Erwerbslosigkeit des Mannes reduziert werde, als durch die der Frau. Vor diesem Hintergrund wurde geschlussfolgert, dass Frauen in festen partnerschaftlichen Lebensgemeinschaften (v.a. Ehefrauen) im Falle ihrer Erwerbslosigkeit mehr finanzielle Unterstützung von ihren Partnern erwarten könnten als dies umgekehrt für die Männer der Fall sei (Paul & Moser, 2009b). Relativierend wurde argumentiert, dass im Zuge der Angleichung der Geschlechterrollen auch die berufliche Identifikation der Frau gewachsen sei und damit auch ihre psychische Abhängigkeit vom Vorhandensein eines Arbeitsplatzes zugenommen habe (Leana & Feldman, 1991). Das Alternativrollen-Argument wurde durch Mohr (1993) relativiert: Zum einen sei der soziale Status, den die Hausfrauenrolle gewähre, gering, zum anderen sei die Besetzung dieser Rolle für Frauen mittlerweile nur noch dann gesellschaftlich legitimiert, wenn kleine Kinder zu versorgen seien. Nach Mohr dürfte also nur eine Subgruppe der erwerbslosen Frauen vom Vorhandensein der Alternativrolle profitieren.

In der Metaanalyse von Murphy und Athanasou (1999) konnte innerhalb der Gruppe der Erwerbslosen kein Unterschiedseffekt des Geschlechts festgestellt werden. Im Querschnitt fanden auch Paul und Moser (2009a) nur geringe Unterschiede. Allerdings konnten sie längsschnittlich geschlechtsspezifische Unterschiede in der Verschlechterung des Befindens von der Berufstätigkeit zur Erwerbslosigkeit feststellen. Demnach waren die psychischen Verschlechterungen zwischen Männern und Frauen deutlich stärker ausgeprägt für die Männer, wenn die Arbeitslosigkeit auftrat. Eine weitergehende Analyse, die das Entstehungsjahr der Untersuchungen mit einbezog, zeigte, dass dieser Geschlechtseffekt nicht auf ältere Studien begrenzt war. Er trat auch auf, wenn nur jüngere Studien metaanalysiert wurden. Die zu erwartende Annäherung der negativen psychischen Effekte von Arbeitslosigkeit zwischen Männern und Frauen fanden sich demnach laut Paul und Moser (2009b) nicht.

Einfluss des Beziehungsstatus

Ob eine feste Lebensgemeinschaft eine moderierende Funktion bei der Bewältigung des Stressors "Erwerbslosigkeit" hat, wurde laut Paul und Moser (2009b) bisher nur selten untersucht. Laut der Autoren finden sich nur wenige Erörterungen dieses Themas in der Literatur (z.B. Lahelma, 1989), in welcher meist argumentiert wird, dass nicht der Beziehungsstatus als solches, sondern die damit verbundene soziale Unterstützung entscheidend sei (Leana & Feldman, 1991). Dabei wird auch darauf aufmerksam gemacht, dass in Partnerschaften von Erwerbslosen häufig das gegenteilige Phänomen auftreten kann, das mit dem Begriff "soziale Unterminierung" paraphrasiert wird (Vinokur & Van Ryn, 1993). Auf der konkreten Ebene drückt sich dies für den Erwerbslosen z.B. in häufiger Kritik, Nörgeleien und Kränkungen durch den Lebensgefährten aus und kann deshalb zu negativen Effekten der Partnerschaft führen.

Paul und Moser (2009a) untersuchten den Status der Beziehung als vermittelnde Variable auf die psychische Gesundheit und fanden keinen signifikanten Effekt. Demnach ist nicht davon auszugehen, dass sich das psychische Befinden signifikant unterschiedlich zwischen den Gruppen des Beziehungsstatus verschlechtert, wenn das Ereignis Arbeitslosigkeit auftritt. In der Metaanalyse von McKee-Ryan, Song, Wanberg und Kinicki (2005) zeigte sich innerhalb der Arbeitslosengruppe zumindest der erwartete positive Zusammenhang zwischen sozialer Unterstützung und psychischer Gesundheit.

Einfluss des Alters

Weil sich vornehmlich Menschen mittleren Alters ein eigenes Leben mit einer eigenen Familie aufbauen und somit in dieser Altersperiode die finanziellen familiären Verantwortlichkeiten am größten sind, vertritt Winefield (1995) die Ansicht, dass im Gegensatz zu jüngeren oder älteren Menschen, diejenigen mittleren Alters am stärksten unter der Erwerbslosigkeit leiden müssten. Im Gegensatz dazu hätten junge Leute bescheidenere Ansprüche, während bei älteren postuliert wird, dass sie die wichtigsten Anschaffungen bereits getätigt und familiäre Ziele erreicht hätten. Eine andere Ansicht ist, dass die Erwerbsarbeit für junge Menschen eine wertvolle Unterstützung bei der Bewältigung der zentralen Entwicklungsaufgaben 1.) Herausbildung einer eigenen Identität und 2.) Ablösung vom Elternhaus darstellt (Schober, 1978). Demnach müssten insbesondere junge Menschen Anzeichen des negativen psychischen Befindens beim Auftreten von Erwerbslosigkeit zeigen. Des Weiteren wurde behauptet

tet, dass ältere Menschen häufig ihre Karriereziele hinter sich hätten und daher eine auftretende Erwerbslosigkeit in eine Art vorgezogenen Ruhestand uminterpretieren könnten (Warr & Jackson, 1987). Zusammengefasst implizieren also sämtliche Argumente, dass ältere Erwerbslose am wenigsten unter ihrem Zustand leiden müssten.

McKee-Ryan et al. (2005) stellten einen Gruppenunterschied im Vergleich zwischen Schulabgängern und Erwachsenen fest; dieser zeigte einen stärkeren negativen Effekt der Erwerbslosigkeit auf das Befinden der Jugendlichen. Murphy und Athanasou (1999) fanden analog dazu keinen signifikanten Unterschied zwischen arbeitslosen Jugendlichen und Erwachsenen. In der Metaanalyse von Paul und Moser (2009a) wurde überprüft, ob es einen linearen oder einen kurvlinearen Zusammenhang zwischen dem Alter der Probanden und der Effektstärke für die negativen Auswirkungen der Arbeitslosigkeit gab. Kein Zusammenhang erwies sich als signifikant. Auch eine Begrenzung der Analyse auf männliche Stichproben zeigte keinen signifikanten Zusammenhang zwischen Alter und Beschwerdedruck. Paul und Moser (2009b) resümieren, dass ein Fazit zum vermittelnden Effekt des Alters nicht möglich sei und stellen es als bemerkenswert heraus, dass sich für die plausibel erwartbare kurvlineare Beziehung (mit milden negativen Auswirkungen der Arbeitslosigkeit bei jüngeren und älteren Menschen und starken negativen Folgen für Menschen mittleren Alters) in bisher keiner Analyse empirische Hinweise fanden.

Dies bedeutet im Umkehrschluss, dass bislang nicht davon auszugehen ist, dass eine Erwerbslosigkeit im fortgeschrittenen Alter mit geringerem psychischen Schaden einhergeht. Diese Auffassung ist vermutlich falsch. Es sollte auch Notiz davon genommen werden, dass über 50-jährige erwerbsfähige Menschen beim Verlust ihrer Stelle ein stark erhöhtes Risiko haben, über 12 Monate in der Arbeitslosigkeit zu verbleiben und damit den Status "Langzeitarbeitslos" zu erhalten. Mit dem Eintreffen dieses Status droht durch die reformierte Sozialgesetzgebung eine Verschlechterung der finanziellen Rahmenbedingungen. Die steigende Zuwachsrates des Langzeitarbeitslosigkeitsrisikos in Relation zum Alter ist seit Jahren bekannt. In folgender Grafik sind für die Jahre 1994 und 1996 die prozentualen Verbleiberaten in Abhängigkeit vom Alter der Betroffenen abgetragen.

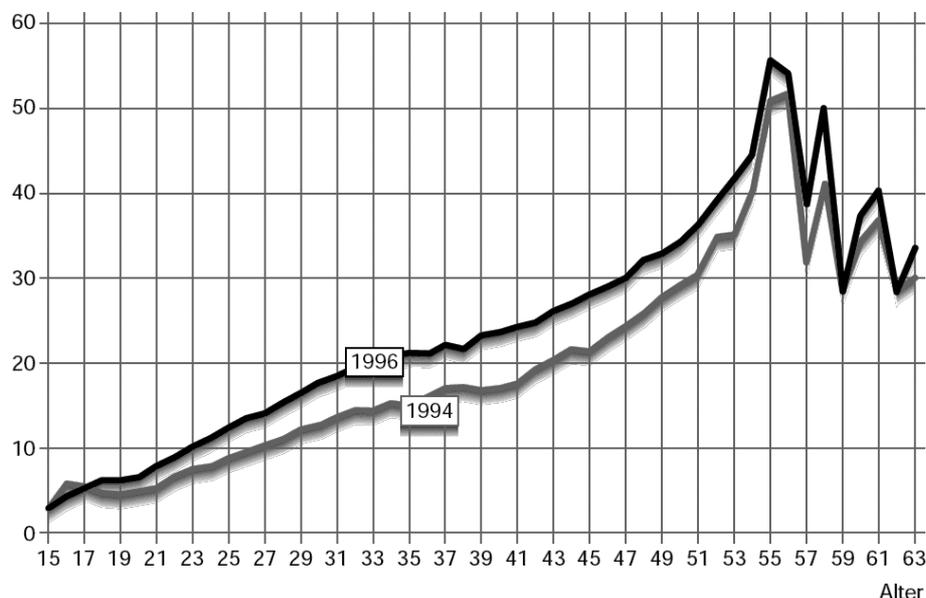


Abb. 1.2: Verbleib in Langzeitarbeitslosigkeit (in %) in Abhängigkeit vom Alter (IAB, 1999, S.8).

Deutlich erkennbar ist die stetige Zuwachsrate und der sprunghafte Anstieg des Risikos ab ca. dem 50. auf das 55. Lebensjahr. In dem Alter weisen Menschen beim Verlust ihres Arbeitsplatzes die höchste Wahrscheinlichkeit auf, dauerhaft arbeitslos zu bleiben. Nach dem Gipfel nehmen die Verbleiberaten aufgrund artifizierter Gründe ab, welches durch in den Neunzigern noch forcierte Programme wie z.B. der Frühverrentung erklärbar ist. Es kann angenommen werden, dass die Bekanntheit des spezifischen Langzeitarbeitslosigkeitsrisikos für über 50-jährige Menschen Angst vor einem Arbeitsplatzverlust auslöst. Zusätzlich dürfte das Wissen um die erheblich geringeren Chancen auf Wiedereingliederung Sorge um einen langen Verbleib in der Erwerbslosigkeit auslösen und fatalistische Einstellungen im Sinne von "Ab 50 wird man nicht mehr gebraucht" begünstigen.

Einfluss der Arbeitslosigkeitsdauer

Laut Paul und Moser (2009b) wurden im Laufe der Theoriebildung zum Einfluss der Arbeitslosigkeitsdauer zahlreiche Stufenmodelle zur psychischen Verarbeitung der Erwerbslosigkeit formuliert. Bereits in den Dreißigern verfassten Zawadski und Lazarsfeld (1935) aufgrund der Ergebnisse der "*Arbeitslosen von Marienthal*" ein solches Modell, das 3 Jahre später in folgendes klassisches Modell umformuliert wurde:

- 1.) Zunächst tritt ein Schockzustand auf, dem dann eine aktive Suche nach Arbeit folgt, während der das Individuum zunächst noch optimistisch ist und nicht resigniert; es hält eine ungebrochene Haltung aufrecht.
- 2.) Dann, wenn alle Anstrengungen sich als nutzlos erweisen, wird das Individuum pessimistisch, ängstlich und erlebt akute seelische Not; dies ist die kritischste Phase.
- 3.) Dann schließlich wird das Individuum fatalistisch und passt sich an seinen neuen Zustand an, allerdings in einem eingeschränkten Rahmen. Es hat nun eine gebrochene Haltung" (Eisenberg & Lazarsfeld, 1938, S. 378; zit. nach Paul & Moser, 2009b).

Auch neuere Phasenmodelle (z.B. Hill, 1978; Amundson & Borgen, 1982; Winegardner, Simonetti, & Nykodym, 1984) formulieren meist eine Beschreibung, die dem generellen Verlauf des klassischen Modell ähnelt. Empirische Beweise für diese Verlaufsform finden sich z.B. in der Studie von Jones (1991), der 28 Längsschnittstudien zur Arbeitslosigkeitsdauer analysierte. Der spezielle Verlauf zur psychischen Arbeitslosigkeitsverarbeitung lässt sich funktional auch durch ein Zusammenspiel folgender Modelle erklären:

- 1.) Das **additive Stressmodell** (Jackson & Warr, 1984) prognostiziert, dass es für den Erwerbslosen im Rahmen seiner Bemühungen, eine neue Stelle zu finden, zu zunehmend schmerzhafter werdenden Zurückweisungen kommt. Gleichzeitig gehen die finanziellen Ersparnisse zuneige und die Unterstützungszahlungen werden im Verlauf geringer. Des Weiteren steigen die soziale Isolation sowie die Stigmatisierung. Nach dem Modell additiver Stressoren verschlechtert sich die psychische Gesundheit kontinuierlich mit der Arbeitslosigkeitsdauer (linearer Zusammenhang).
- 2.) Das **Adaptionsmodell** (Warr und Jackson, 1987) sagt voraus, dass Arbeitslose nach einiger Zeit eine bessere Vertrautheit mit ihrer Situation gewinnen, die es ihnen ermöglicht, Copingstrategien und andere Formen der Anpassung zu entwickeln. Einer anfänglichen Verschlechterung der Befindens folge demnach eine Stabilisierung auf niedrigem Niveau (kurvlinearer Zusammenhang).

Paul und Moser (2009a) fanden einen hochsignifikanten linearen Zusammenhang zwischen der Dauer der Arbeitslosigkeit und der psychischen Beanspruchung. Diese Relation stützt das additive Stressmodell. Der statistische Test auf einen kurvlinearen Zusammenhang war ebenfalls signifikant, welches darauf hinweist, dass auch der vermutete Adaptionsprozess existiert. Die Autoren zeigen in ihrer Publikation, dass die anfängliche zunehmende Verschlechterung des psychischen Befindens mit ca. 9 Monaten einen Höhepunkt zu erreichen scheint, dem eine partielle Verbesserung folgt und dann eine Stabilisierung auf einem Niveau, das schlechter ist als das Ausgangsbefinden zu Beginn der Arbeitslosigkeit, aber besser als das Befinden während der Krise um den 9. Monat (vgl. folgende Abb.).

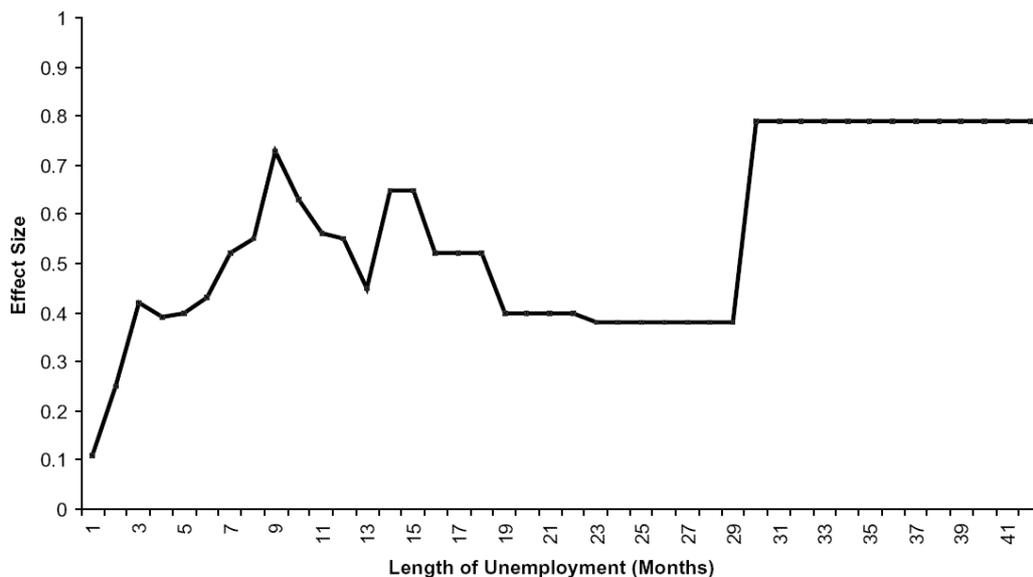


Abb. 1.3: Effektstärken psychischer Beschwerden im Zusammenhang mit der Arbeitslosigkeitsdauer (Paul & Moser, 2009a, S. 274). Die Werte ab dem 23. Monat repräsentieren keine detaillierten Veränderungen, da sie nur durch wenige Primärstudien repräsentiert werden.

Zusammenfassend kann demnach als erwiesen gelten, dass die Arbeitslosigkeitsdauer mit einer zunehmenden Verschlechterung der psychischen Gesundheit zusammenhängt, aber sich innerhalb der ersten zwei Jahre Hinweise auf eine Stabilisierung des Befindens (auf niedrigem Niveau) finden lassen. Die Entwicklung der psychischen Gesundheit bei extrem langfristiger Arbeitslosigkeit (>3 Jahre) konnten Paul und Moser (2009a) nicht mehr differenziert klären, da für diesen Zeitraum nur noch wenige Primärstudien vorhanden waren.

Lahelma (1989) nimmt an, dass in Bezug auf die psychische Beeinträchtigung eine weitere Steigerung nach 3-jähriger Erwerbslosigkeit nicht mehr stattfindet. In Bezug auf die Begrifflichkeiten stellt sich Mohr (2001) ausdrücklich dagegen, den möglichen Stillstand in der Zunahme von psychischen Beschwerden mit dem positiv konnotierten Begriff der "Adaption" zu bezeichnen. Sie beruft sich auf Zippay (1995), demzufolge entnommen werden kann, dass der Preis für den Stillstand der resignative Rückzug vom Arbeitsmarkt und somit von Misserfolgserfahrungen sei. Nach einer dreijährigen Arbeitslosigkeit steigt nach Shlay und Rossi (1992) das Risiko, dass der Betroffene sich immer weiter zurückzieht und zu den sog. "sozial Desintegrierten" (zu denen im Extremfall auch Obdachlose zählen) wechselt. Mohr (2001) schlägt vor, für die extrem langjährig Erwerbslosen die "Aufgabe der Wiedervermittlungshoffnung" zu erheben, wobei sie gleichzeitig zu bedenken gibt, dass sich eine valide

Ermittlung dieses Merkmals äußerst schwierig gestalten würde, da es stark von externen Faktoren, wie z.B. der regionalen Arbeitsmarktlage, abhängig ist.

Die Befunde, wonach bei über 3-jähriger Erwerbslosigkeit depressive Reaktionen im Durchschnitt nicht weiter zunehmen, schließen zeitlich begrenzte, kritische Erhöhungen der Symptomatik über die Zeit nicht aus: Strehmel und Mayring (1986) belegten in ihrer Untersuchung deutliche Verschlechterungen des Befindens bei extrem langfristig erwerbslosen Lehrern in Abhängigkeit von den jährlichen Einstellungsterminen und den damit verbundenen Absagen auf Bewerbungen. Zunahmen der Depressivität nach langjähriger Arbeitslosigkeit werden auch von Leana und Feldman (1992) in Verbindung mit Verschlechterungen der finanziellen Rahmenbedingungen dargestellt. Finanzielle Ressourcen spielen selbstverständlich eine zentrale Rolle bei der Bewältigung insbesondere der Langzeitarbeitslosigkeit. Frese und Mohr (1987) wiesen nach, dass die durch finanzielle Einschränkungen verursachte Unkontrollierbarkeit der eigenen Lebenssituation depressive Reaktionen bei älteren Langzeiterwerbslosen verursacht.

1.4 Besonderheiten der Langzeitarbeitslosigkeit

In Deutschland sind ziemlich genau 50% der amtlich registrierten Arbeitslosen ein Jahr oder länger erwerbslos und somit "langzeitarbeitslos" (STABU, 2006). Im EU-Vergleich zählt Deutschland zu den Ländern mit dem höchsten Anteil an Langzeiterwerbslosen. Höhere Anteile finden sich nur noch in süd- (z.B. Griechenland: 56%) oder in osteuropäischen (z.B. Polen: 55%) EU-Mitgliedsländern; Schweden (16%) hat den kleinsten Anteil von langfristig Erwerbslosen. Der tatsächliche Anteil der Betroffenen liegt in Deutschland allerdings höher. In den amtlichen Statistiken fehlen jene, die infolge des Arbeitsplatzverlustes frühzeitig in Rente gegangen sind. Auch diejenigen, welche durch zu hohe Einkünfte ihres Partners nicht zum Leistungsbezug berechtigt sind und sich von daher nicht arbeitslos melden, fehlen (Mohr & Otto, 2005). Nimmt der Betroffene an einer Maßnahme oder einer Umschulung teil, wird er für die Dauer nicht als erwerbslos gezählt. Langzeitarbeitslose, die aufgrund von Erkrankungen nicht mehr ausreichend arbeitsfähig sind, gehen ebenso nicht mehr in die Statistiken ein (Hollederer, 2002). Insofern gibt es Anhaltspunkte, die darauf hinweisen, dass die offiziellen Statistiken das wahre Ausmaß des Problems unterschätzt abbilden.

Zunahme von Selektionseffekten

Mohr (2001) vertritt die Ansicht, dass bei Langzeiterwerbslosen Kausal- und Selektionseffekte im Sinne eines Teufelskreises auftreten. Sie beruft sich auf Zempel und Frese, die auf Grundlage ihrer Längsschnittergebnisse konstatieren:

"...längere Erwerbslosigkeit führt zu psychischen Befindensbeeinträchtigungen und reduzierter Eigeninitiative, geringere Eigeninitiative verlängert die Dauer der Erwerbslosigkeit und die Folgen nehmen mit der Dauer weiter zu" (Zempel & Frese, 2000, S. 387).

Claussen (1994) spricht sich bei langfristig Erwerbslosen für das Wirken von psychischen Selektionseffekten aus, indem er die seelische Gesundheit als Kriterium für die Wiedereinstellungswahrscheinlichkeit beschreibt. In einer US-amerikanischen epidemiologischen Studie (N=89.267) belegten Brown et al. (2003), dass Langzeiterwerbslose gegenüber Beschäftigten eine annähernd doppelt so hohe Wahrscheinlichkeit aufweisen, innerhalb eines

definierten Zeitraumes psychisch zu erkranken. Bereits Klein (1990) dokumentiert für die alten Bundesländer, dass zwei Drittel aller wieder eingestellten Langzeitarbeitslosen innerhalb von 4 Jahren erneut erwerbslos wurden. Daher schätzt Mohr (2001), dass viele langfristig Erwerbslose mehrfache Selektionsprozesse auf dem Arbeitsmarkt durchmachen, welches zur stetigen Reduktion von Vermittlungschancen führt. Sie schlussfolgert, dass diese "mehrfach Selektierten" ein besonderes Risiko haben, Erwerbslose auf Dauer zu bleiben. Selektionsprozesse werden für die Rahmenbedingungen Alter, Qualifikationsniveau und Gesundheit angenommen.

Rückgang des Qualifikationsniveaus

Das berufsrelevante Lernen findet meist innerhalb der Arbeitstätigkeit selbst statt. Vor diesem Hintergrund bedeutet der Verlust einer Anstellung gleichzeitig den Verlust von Lern- und Qualifikationsmöglichkeiten (Mohr, 2001; Olbrich, 1989). Bedenkt man die Halbwertszeit von IT-gestützten Arbeitsmethoden, dann erzeugt eine Erwerbslosigkeit von wenigen Jahren eine Dequalifizierung, die ein bedeutsames Vermittlungshemmnis darstellen kann. Laut Mohr (2001) kommt es bei Langzeitarbeitslosen vermutlich auch zu einem Verfall der "soft-skills", wie z.B. Kommunikations- und Konfliktfähigkeiten oder die Befähigung zu lernen und zu organisieren. Gerade für ältere Langzeiterwerbslose ist die Dequalifizierung besonders fatal, da sie bereits aufgrund ihres Alters eine geringere Wiedereinstellungschance haben.

Einschränkung des sozialen Netzes

Klein und Strasser (1997) stellen fest, dass die (geschlechtsunabhängige) Scheidungswahrscheinlichkeit bei Langzeiterwerbslosen fast dreimal höher ist als in der Allgemeinbevölkerung. Bei langfristig erwerbslosen Männern ist der Prozentsatz der Ehen, die innerhalb der ersten 4 Jahre geschieden werden sogar mehr als 3-mal so hoch wie bei den männlichen Berufstätigen (20% vs. 6%). Neben dem Verlust regelmäßiger sozialer Kontakte am Arbeitsplatz verfügen die Betroffenen also auch seltener über feste lebenspartnerschaftliche Beziehungen als Erwerbstätige. Eine generelle Veränderung des sozialen Umfeldes im Laufe der langfristigen Erwerbslosigkeit ist anzunehmen: Jackson (1988) weist darauf hin, dass vor allem die sog. "weak ties" (lockere bzw. flüchtige Bekanntschaften und Beziehungen) nach längerer Erwerbslosigkeit drastisch abnehmen. Dies ist arbeitsmarkttechnisch eine kritische Veränderung, weil der Studie von Engelbrech (1987) zu entnehmen ist, dass vor allem das weitere soziale Netz – d.h. außerhalb des engen Familien- und Freundeskreises – Hilfen bei der Wiedervermittlung bietet. Viele Stellen werden über Mundpropaganda weitergegeben und v.a. ehemalige Kollegen oder Geschäftsbekanntschaften geben die entscheidenden Hinweise, wo eine Stelle frei wird, auf die es sich zu bewerben lohnt.

Psychische Wiedereinstellungseffekte

Bei der Aufnahme einer Erwerbsarbeit kommt es durchschnittlich zu einer Verbesserung des Befindens. Paul et al. (2006) berechneten, dass für Erwerbslose nach einer neuen Anstellung eine Verbesserung der psychischen Gesundheit mit $d=0,35$ resultiert. Auch für die Gruppe der Langzeitarbeitslosen stellt die Mehrheit der Studien laut Murphy und Athanasou (1999) eine Verbesserung der psychischen Gesundheit nach der Wiedervermittlung fest. Nach Verkleij (1989) sind es sogar vor allem die wiederbeschäftigten Langzeiterwerbslosen,

im Kontrast zu den kurzfristigen Arbeitslosen, die eine deutliche Steigerung ihres psychischen Befindens zeigen. Auch Claussen (1999) ermittelte bei langfristig Erwerbslosen nach dem beruflichen Wiedereinstieg allgemeine Verbesserungen des Befindens.

Claussen stellt jedoch auch fest, dass eine Erhöhung der psychischen Gesundheit nicht für alle Wiedereingestellten zu finden war. Der Autor vermutet, dass eine empfundene Arbeitsplatzunsicherheit, die für einen Teil der Wiedervermittelten vorlag, eine Reduzierung der negativen psychischen Symptomatik erschwerte (Claussen, 1999). In den letzten zwei Jahrzehnten weisen Studien – so z.B. Wanberg (1995) – vermehrt darauf hin, dass eine gesundheitliche Verbesserung bei Wiedereinstellung nicht für jene gilt, die in Tätigkeiten mit schlechter Bezahlung, hoher Repetitivität, fehlenden Entwicklungsmöglichkeiten und unsicherer Zukunftsperspektive vermittelt werden. Zu diesen Einschätzungen kommen auch die Vertreter des Konzepts der "Prekarisierung" (z.B. Plöger, 2010). Auch Leana und Feldman (1995) berichten, dass Wiederbeschäftigte sich nicht besser fühlten als Erwerbslose, wenn die neue Tätigkeit nicht ihren Vorstellungen entsprach und sie mit der Stelle unzufrieden waren. Nach Ansicht der Autoren ist es nicht entscheidend, irgendeine Arbeit zu haben, sondern einer Tätigkeit nachzugehen, die den eigenen Bedürfnissen entspricht. Übereinstimmend mit dieser Sichtweise belegen auch Broom et al. (2006) in ihrer Studie, dass sich keine Unterschiede des negativen psychischen Befindens zwischen Berufstätigen mit prekären Arbeitsbedingungen und Arbeitslosen finden ließen. Sobald die Stärke der von ihnen ermittelten beruflichen Stressoren einen bestimmten Schwellenwert überschritt, fühlten sich beide Gruppen gleichermaßen schlecht. Selbst Paul et al. (2006) sehen - trotz ihres Befundes für die positiven Wiedereinstellungseffekte - die Unterstellung einer "segensreichen" Funktion für jegliche Form der Erwerbsarbeit kritisch. Sie schreiben, dass ein solch verallgemeinernder Schluss eigentlich nicht zulässig sei, ohne dass Erwerbstätige nicht nur mit Arbeitslosen, sondern mit der gesamten Gruppe der sog. "out of labour force" (vgl. Rengers, 2005) verglichen werden. Angehörige der "out of labour force" stellen nämlich nicht nur Erwerbslose, sondern z.B. auch Rentner, Hausfrauen/-männer, Studierende oder finanziell wohlhabende Menschen (diejenigen, die nicht arbeiten und trotzdem finanzielle Autonomie genießen) dar. Erst beim Nachweis von positiven Unterschieden zwischen Erwerbstätigen und einer repräsentativen Stichprobe aus der gesamten "out of labour force" könne der Erwerbsarbeit ein "segensreicher" Effekt zugesprochen werden, schlussfolgern die Autoren (vgl. auch Paul, Geithner, & Moser, 2009).

Schlechte Prognose auf Rückkehr in den Arbeitsmarkt

Bereits vor zwei Jahrzehnten kam Klein (1990) zu dem pessimistisch stimmenden Befund, dass für die Zahl der langfristig Erwerbslosen kaum eine Reduzierung durch Wiedervermittlung feststellbar war. Mohr (2001) berichtet, dass knapp drei Viertel aller wiederbeschäftigten Langzeitarbeitslosen nach 4 Jahren erneut arbeitslos werden; mehr als 50% von ihnen sogar bereits in den ersten 9 Monaten nach der Wiedervermittlung. Dementsprechend schlussfolgert die Autorin, dass es sich bei stabil wiedervermittelten Langzeiterwerbslosen nur um eine Minderheit handelt. Mehr Bewerbungsaktivitäten scheinen einen besseren Vermittlungserfolg ebenso nicht zu garantieren: Zippay (1995) berichtet, dass ältere langfristig Erwerbslose sich in ihrem anfänglichen Arbeitssuchverhalten nicht von den jüngeren Wiedervermittelten unter-

schieden, sondern nur im Erfolg ihres Suchverhaltens. Auch ein hohes "employment commitment" hat laut Mohr (2001) keinen Einfluss auf die Wiedervermittlungsquote. Zudem stellen Warr und Jackson (1985) keinen Zusammenhang zwischen der Dauer der Arbeitslosigkeit und der Veränderung des commitments fest und auch Paul und Moser (2006) gehen von einer Stabilität der arbeitsethischen Einstellung aus, welche sie im Rahmen ihrer *Inkongruenzhypothese* (vgl. Kap. 1.2) als mögliche Ursache des Leidensdrucks von Arbeitslosen beschreiben. Pernice (1996) kommt auf der Grundlage seiner längsschnittlichen Erhebungen mit Erwerbslosen zu dem Ergebnis, dass die Reduktion eines anfangs hohen commitments zwar mit einer besseren psychosozialen Gesundheit einherging, aber nicht zu einer Verminderung der Wiedervermittlungsquote führte. Mohr (2001) konstatiert anhand amtlicher Statistiken, dass eine beeinträchtigte Gesundheit, das Alter, das Geschlecht und die regionale Arbeitsmarktlage herausragende Determinanten für die Wiedervermittlungschancen von Langzeitarbeitslosen darstellen. Klammert man die Gesundheit aus, sind dies sämtlich Merkmale, auf die die Betroffenen keinen Einfluss haben.

1.5 Förderung von Erwerbslosen

Langzeiterwerbslose haben per se eine ungünstige Wiedervermittlungsprognose, kaum Auswahl an neuen Arbeitsgelegenheiten und von daher ein erhöhtes Risiko, dass sie bei einer Wiedereinstellung eine Tätigkeit mit prekären Arbeitsbedingungen erhalten und/oder befristet im Rahmen von Zeitarbeit oder im Niedriglohnsektor eingestellt werden. Dass sich unter diesen Umständen nur wenige Argumente finden lassen, die eine psychische Stabilisierung der Betroffenen vorhersagen, wurde in den letzten Kapiteln diskutiert.

Es werden viele unterschiedliche Programme für Erwerbslose durchgeführt. Da solche Interventionen fast ausschließlich im Auftrag der Bundesagentur für Arbeit oder der ARGE³ (bzw. von den entsprechenden Behörden auf internationaler Ebene) vergeben werden, haben diese Maßnahmen den nachvollziehbaren Legitimationsdruck zu erfüllen, dass sie möglichst unmittelbar sog. "Vermittlungshemmnisse" abbauen sollen. Insofern werden vornehmlich Weiterbildungsmaßnahmen, subventionierte Arbeit oder Bewerbungstrainings angeboten. Meist erhofft man sich implizit einen positiven Nebeneffekt auf die psychische Verfassung der Zielpersonen, allerdings werden gezielte Interventionen zur Förderung der seelischen Gesundheit nur selten angewendet.

Will man Maßnahmen für Erwerbslose evaluieren, so stellt sich zuerst die Frage, welche Parameter als Erfolg definiert werden sollen. Vor allem der niedrige Anteil von stabil wiedervermittelten Langzeiterwerbslosen nebst der Tatsache, dass dieser erheblich vom regionalen Arbeitsmarkt abhängt, legt nahe, nicht allein in einer höheren Wiedervermittlungsquote das (vergleichbare) Erfolgskriterium zu suchen. Unter dem Druck dieses Kriteriums würde zudem die Gefahr steigen, dass lediglich den Erwerbslosen Förderungsmaßnahmen offeriert würden, für die ohnehin eine bessere Vermittlungschance besteht. Oder, dass Fördermaßnahmen nur noch in Regionen mit starkem Arbeitsmarkt als profitabel angesehen werden würden. Trube und Luschei (2000) fordern eine Erweiterung der Erfolgskriterien für Förderungsmaßnahmen auf qualitative Zielkriterien und eine Evaluation, die Informa-

³ Arbeitsgemeinschaft nach dem SGB II

tionen enthält über Bedingungen von Erfolg und Misserfolg. Ernst, Hartwig und Pokora (2008) verlangen für die Evaluation von Bildungs- und Beschäftigungsmaßnahmen für Erwerbslose "...zusätzlich die Erfolge sozialer Integration, verbesserter Gesundheit und sozialer Kompetenzentwicklung zu dokumentieren" (S. 268).

Mohr und Otto (2005) fordern, dass sich Maßnahmen für Erwerbslose vorwiegend an Risikogruppen richten sollten. Interventionsangebote müssten den Anspruch haben, langfristig Erwerbslose, ältere Personen, Geringqualifizierte, Frauen und Personen mit physischen oder psychischen Gesundheitsbeeinträchtigungen zu fördern. Amundson und Borgen (1987) sprechen sich bei Langzeitarbeitslosen dafür aus, primär Hilfe bei der Bewältigung der Erwerbslosigkeit anzubieten und erst danach Fähigkeiten für die Arbeitsuche aufzubauen. Der Aufbau einer *realistischen Erwartungshaltung* sei eine Gratwanderung, die es bei der Unterstützung von Langzeiterwerbslosen zu meistern gilt, weil übersteigerte Hoffnungen bei Nichterfüllung zur Enttäuschung führen und depressive Reaktionen auslösen können, resümieren Frese und Mohr (1987). Dass von älteren langfristig Erwerbslosen speziell supportives Verhalten bei der Stellensuche als stabilisierend erlebt wird, beschreibt Frese (1987). Rife und Belcher (1993) berichten, dass die Mehrheit ihrer befragten Langzeiterwerbslosen vor allem andere Arbeitslose als Quelle der sozialen Unterstützung nannte. Freunde und Familienmitglieder, die erwerbstätig oder in Rente waren, wurden von den Erwerbslosen kaum als Quelle der Unterstützung genannt. Insofern sollten Interventionen als Gruppenmaßnahme für langfristig Erwerbslose angeboten werden, damit sie die Gelegenheit erhalten, sich über ihr Bewältigungsverhalten austauschen zu können. Mielck (2005) beschreibt, dass sich Arbeitslose zwar auch subjektiv stärker krank fühlen, aber trotz dieses Bewusstseins weitaus weniger Angebote zur Prävention bzw. Gesundheitsförderung wahrnehmen als Berufstätige. Dies nicht zuletzt finanzielle Gründe: Selbst bei Erstattung der Kursgebühren durch die Krankenkasse müssen die Kosten von den Teilnehmern prinzipiell vorgestreckt werden, welches in Anbetracht der ALG-II-Regelsätze oft nicht leistbar ist. Daraus kann geschlossen werden, dass eine Gesundheitsförderung direkt von der Institution, welche die Langzeiterwerbslosen primär betreut (meist die ARGE) kostenlos angeboten werden sollte, um die potentiellen Teilnehmer erreichen zu können.

1.6 Vorstellung von Interventionsstudien

In der Recherche zu Forschungsberichten, in denen psychologische Maße als Evaluationskriterien für einen Interventionseffekt bei Erwerbslosen herangezogen wurden, findet man hauptsächlich sozialpädagogische, motivationsfördernde, psychologische bzw. "psychologienahe" Interventionen, aber keine primär körperliche Gesundheitsförderung.

Rothländer und Richter (2009) sowie **Rothländer et al. (2012)** beschreiben ihr Programm "AKTIVA" als psychosoziales Training für Erwerbslose. Bewährte kognitiv-behaviorale Techniken kamen zur Anwendung um die Zielgruppe in ihrer Handlungsfähigkeit zu stärken. Die Maßnahme bestand aus den Modulen "Aktivitätenplanung" (Zeitmanagement), "Konstruktives Denken" (Identifizierung und Realitätsüberprüfung destruktiver Denkschemata und Erlernen konstruktiver Gedanken), "soziale Kompetenz" (nach Hinsch und Pfingsten, 2002) und "systematisches Problemlösen" (Inhalte des Projektmanagements). Das Programm richtete

sich vornehmlich an Langzeitarbeitslose mit Qualifizierungsdefiziten und umfasste 24 Stunden innerhalb von 2-4 Wochen. Untersuchungsergebnisse wurden für T1 (vor dem Training), T2 (letzter Trainingstag) und T3 (3 Monate nach dem Training) für 418 Langzeiterwerbslose (EG: 236, KG: 182) analysiert. Über 2 verschiedene Zugangswege erfolgte die Rekrutierung, einerseits als verpflichtende Maßnahme und andererseits als freiwilliges Angebot. Fragebögen für körperliche, psychische und soziale Beschwerden, Selbstwirksamkeitserwartung und einer zur sozialen Unterstützung kamen zum Einsatz. Für T1-T2 berichten die Autoren eine signifikante Verringerung der psychischen Beschwerden für die EG, jedoch nicht für die KG. Das Interventionsergebnis blieb im follow-up nur für die Teilnehmer des freien Angebotes stabil. Auf die Trainingseffekte von Rothländer et al. (2012) wird noch im Diskussionsteil dieser Arbeit (Kap. 5.2.2) Bezug genommen.

Vastamaki, Moser und Paul (2009) untersuchten in Finnland 74 Langzeitarbeitslose, welche verschiedene Maßnahmen absolvierten. Diese bestanden aus beruflicher Fortbildung, Gruppenberatung, subventionierter Arbeit und Netzwerkarbeit. Als psychologisches Maß wurde der Kohärenzsinn (Antonovsky, 1987) untersucht. Als Kovariaten wurden workability (Arbeitsfähigkeit), Coping-Ressourcen, finanzielle Situation und Bildung erhoben. Die Daten wurden vor der Maßnahme (T1) und nach 6 Monaten (T2) (die Autoren sprechen von "follow-up") erhoben. Sie fanden Verbesserungen nur in einzelnen Komponenten des Kohärenzsinn. Lediglich die Probanden, die während oder nach der Intervention wiedereingestellt wurden, zeigten signifikante positive Effekte im Kohärenz-Gesamtscore.

Nitsche und Richter (2003) evaluierten das "TAURIS"-Projekt, in welchem Langzeiterwerbslose die Wahl hatten, in ehrenamtlicher gemeinnütziger Arbeit mit 14 Stunden Arbeit/Woche und einem finanziellen Ausgleich von € 78,-/Monat, eingesetzt zu werden. Die Autoren untersuchten u.a. den Einfluss auf unspezifische psychische Beschwerden mit dem GHQ (General Health Questionnaire, Goldberg & Hillier, 1979) und das psychosoziale Wohlbefinden. Ihre Analyse basiert auf 350 langzeitarbeitslose Probanden, die mindestens zweimalig im Abstand von 6 Monaten untersucht wurden. Nitsche und Richter (2003) errechneten, dass sich das Wohlbefinden der Teilnehmer so verbesserte, dass es sich nicht mehr von altersparallelisierten Erwerbstätigen unterschied und damit signifikant höher lag als das von denen ohne Maßnahme. Die positiven Interventionseffekte erwiesen sich allerdings nicht als zeitstabil, denn sie waren maximal für 6 Monate nachweisbar. Nach etwa einem Jahr hatten sich die positiven Effekte so abgemildert, dass sie auf das Anfangsniveau zurückfielen. Dass in unserer leistungsorientierten Gesellschaft nur der bezahlten Arbeit die entscheidende Bedeutung zur Stabilisierung von Selbstwert und Gesundheit zukäme, wurde von den Autoren als Ursache für die unsoliden Interventionseffekte interpretiert.

In den Studien von **Proudfoot, Guest, Carson, Dunn und Gray (1997)** bzw. **Proudfoot, Gray, Carson, Guest und Dunn (1999)** wurden 289 langzeitarbeitslose Probanden untersucht. Nach einem kognitiv-behavioralen Training waren unspezifische psychische Beschwerden (GHQ) gemindert und Selbstbewusstsein, Attributionsstil, Selbstwirksamkeitserwartung und Arbeitsmotivation signifikant gesteigert. Im follow-up war dies allerdings nur für die Probanden der Fall, die wieder eingestellt wurden. Bei den Probanden, die 3 bis 4 Monate nach dem Abschluss des Training noch erwerbslos waren, blieb nur die Verände-

rung des Attributionsstils stabil; alle anderen Variablen unterschieden sich nicht mehr von den Werten der Erstuntersuchung.

Vuori und Vesalainen (1999) untersuchten in Finnland 745 Probanden, die sich in der Übergangsphase von der Kurzzeit- zur Langzeitarbeitslosigkeit (knapp < 1 Jahr) befanden und von den Sachbearbeitern der Arbeitsämter so beurteilt wurden, dass sie ein erhöhtes Risiko hätten, in dauerhafter Arbeitslosigkeit zu verbleiben. Die Versuchspersonen nahmen an verschiedenen Maßnahmen teil, entweder an "guidance courses", in denen die Teilnehmer aktiviert und motiviert werden sollten, oder an beruflicher Fortbildung, oder an einer subventionierten Arbeitsstelle für 6 Monate. Das primäre Zielkriterium in der Studie stellte die veränderte Bewerbungsaktivität der Teilnehmer nebst Vermittlungsquote dar. Es wurden jedoch auch allgemeine psychische Beschwerden (GHQ) längsschnittlich erfasst. Fragebögen wurden zu T1 und ein Jahr später (T2) an die Teilnehmer verschickt. In der Auswertung zeigte sich, dass trotz inhaltlicher Verschiedenheit keine der Maßnahmen einen herausragenden Einfluss auf das Befinden der Probanden zu T2 hatte. Von den anderen Prädiktoren zeigte eine "gute finanzielle Situation" den einzigen signifikanten positiven Einfluss auf den GHQ. Nur falls sich Probanden noch in einer Maßnahme – egal, in welcher – zu T2 befanden, zeigte dies einen positiven Effekt auf den GHQ. Die Autoren interpretieren, dass es während einer Teilnahme prinzipiell zu positiven Effekten kommt, die sich danach aber rasch abbauen. Subgruppenanalysen demaskierten, dass negative Veränderungen des Befindens von T1 zu T2 davon abhingen, ob sich die finanzielle Situation zwischen den Messzeitpunkten (erneut) verschlechterte. Probanden, die an keiner der Maßnahmen teilgenommen hatten und kontinuierlich arbeitslos blieben, zeigten keinen Unterschied im GHQ zwischen T1 und T2, welches gegen einen Testeffekt des Instrumentes spricht. Personen, die zu T2 eine neue Anstellung hatten, verbesserten hochsignifikant ihr GHQ Ergebnis, welches von Vuori und Vesalainen (1999) als kausale Reduktion von unspezifischen psychischen Beschwerden durch den Wechsel in den Erwerbsstatus interpretiert wurde.

Vinokur, Price und Schul (1995) untersuchten mit 1801 Probanden die Auswirkungen ihrer "JOBS"-Workshops, die den Probanden verbesserte Bewerbungsstrategien sowie Möglichkeiten des konstruktiven Copings bei enttäuschenden Ergebnissen ihrer Bewerbungsaktivitäten näher brachten. Problemlösetraining, Strategien zur Entscheidungsfindung und motivierende Haltungen der Trainer den Erwerbslosen gegenüber waren weitere als wirksam vermutete Einflussgrößen. Die Workshops mit einem Umfang von 20 Stunden wurden innerhalb einer Woche abgehalten. Sozialarbeiter und arbeitslose Lehrer waren als Trainer engagiert, die vor ihrem Einsatz eine 240 Stunden umfassende Ausbildung absolvieren mussten. Probanden wurden in EG und KG randomisiert. Die Autoren untersuchten die Auswirkung ihrer Intervention auf Depressivität mit einer Risikogruppen-Stratifizierung, die sie aus den Daten der Depressionsscores von T1 bildeten. Hierbei zeigte sich, dass nur die Probanden aus der Hochrisikogruppe ihre Depressivität signifikant senkten. Die Höhe der erreichten Reduktionen verringerte sich in der EG von T2-T3 nicht signifikant. Die Daten der Personen aus der Niedrigrisikogruppe blieben unverändert zwischen allen Messzeitpunkten und zwar sowohl für EG und KG als auch zwischen EG und KG gab es (trotz hohem N) keine signifikanten Unterschiede. Die Autoren schließen, dass die Maßnahme am effektivsten in der Hochrisikogruppe war und dadurch die Möglichkeit gegeben wäre, sie bevorzugt

dieser zukünftig anzubieten. Allerdings geben sie zu bedenken, dass dann auf Vorbildeffekte, die von nichtdepressiven Teilnehmern ausgehen, welche die depressiven Teilnehmer via Modelllernen "mitziehen", möglicherweise verzichtet werden müsse, welches den Effekt des Trainings in unbekanntem Ausmaß schmälern könne. Der Faktor "finanzielles Auskommen" hatte eine signifikant moderierende Wirkung auf die Depressivität.

Vinokur, Schul, Vuori und Price (2000) untersuchten mit dem gleichen Probandenkollektiv wie in Vinokur et al. (1995) die Auswirkungen ihrer "JOBS"-Workshops, im 2-jährigen follow-up. Hierbei fanden sie, dass die Risikostratifizierung aus der Publikation von 1995 keinen Interaktionseffekt mehr mit der Versuchsgruppe (EG vs. KG) einging. Dementsprechend interpretieren sie, dass die Stratifizierung aufgrund des Depressionsscores von T1 nur kurzfristige Interaktionseffekte produzierte.

Vuori und Silvonen (2005) hatten in ihrer Studie 1153 arbeitslose Probanden zur Verfügung, die sie in EG und KG randomisierten. Die Teilnehmer durchliefen einen einwöchigen "Työhön" genannten Workshop mit Stärken-Schwächen-Analysen, Bewerbungstraining und ein Stressverarbeitungstraining für Hindernisse oder Absagen im Bewerbungsprozess. Mitglieder der Kontrollgruppe erhielten ein Booklet, das sie im Selbststudium durcharbeiten sollten. Die Probanden wurden zu Beginn (T1), nach dem Workshop (T2), nach 6-monatigen (T3) und 2-jährigen (T4) follow-up befragt. Als psychologische Maße benennen die Autoren Depressivität und Selbstbewusstsein. Sie geben an, dass die Probanden zu T2 ihre Depressionssymptome senken und ihr Selbstbewusstsein stärken konnten. Im 6-monatigen follow-up (T3) verschwanden diese Verbesserungen wieder, um dann nach T4 wieder signifikant aufzutauchen. Die Autoren erklären diesen Befund damit, dass die Teilnehmer der EG im 2-jährigen follow-up eine höhere Wahrscheinlichkeit hatten, eine Arbeit erhalten zu haben und durch diesen vermittelnden Umstand die positiven psychischen Effekte verspätet wieder zum Tragen kamen.

Creed, Machin und Hicks (1996) untersuchten 30 junge (19 Jahre) aber trotzdem langzeitarbeitslose Probanden mit einem Training, welches berufspraktisches Wissen bei Konstruktionsarbeiten, Soziale Fähigkeiten, Erste-Hilfe-Kurse, Umweltschutz und "Kurse zur Lebenskunde" für eine Zeit von max. 16 Wochen vermittelte. Die Kontrollgruppe bestand aus einer dem Alter entsprechenden Wartegruppe. Fragebogendaten zum Selbstbewusstsein und zu unspezifischen psychischen Beschwerden (GHQ) wurden vor (T1) und nach dem Training (T2) erhoben. Vor dem Training erfolgte eine Messung der Persönlichkeitsstruktur. Von besonderem Interesse war als erwartete Moderatorvariable des Interventionseffektes der Persönlichkeitsfaktor "Neurotizismus". Unterschiede in der Verbesserung der abhängigen Variablen konnten zwischen EG und KG nur für Selbstbewusstsein, nicht aber für den GHQ gefunden werden. Für die Stärkung des Selbstbewusstseins geben die Autoren zwischen T1-T2 eine mittlere Effektstärke an. Der Faktor Neurotizismus zeigte keinen moderierenden Einfluss. Ein follow-up wurde nicht durchgeführt.

Creed, Machin und Hicks (1999) untersuchten erneut ein kognitiv-behaviorales Training für 43 junge (19 Jahre) langzeitarbeitslose Probanden. Die KG (Wartegruppe) bestand aus 22 altersentsprechenden langzeitarbeitslosen Probanden. Das Training bestand aus einem 15-stündigen 3-Tage Programm, welches in Anlehnung an Seligmans Konzept des "gelernten Optimismus" aufgebaut war (Seligman, 1991). Als Maße kamen u.a. der GHQ

und eine Skala zum Selbstbewusstsein zum Einsatz, die vor (T1) und nach dem Training (T2) sowie in einem 4-monatigen follow-up (T3) erhoben wurden. Von T1-T2 ergaben sich für die EG signifikante Zugewinne im GHQ und im Selbstbewusstsein, nicht jedoch für die KG. Von T2-T3 verschlechterte sich der GHQ signifikant für die EG und kehrte ca. zum ursprünglichen Niveau zurück, während sich der Aufbau des Selbstbewusstseins trotz deskriptiven Rückgangs als relativ zeitstabil erwies.

Maysent und Spera (1995) beschreiben die Effekte eines Stressmanagement-Trainings für 76 kürzlich gekündigte Angestellte, die von einer Outplacement-Firma betreut wurden. Die Teilnehmer entstammten sämtlich dem bildungsnahen "white-collar"-Segment. Durch den äußerst kurzfristigen Beginn der Maßnahme nach der Kündigung bestand der Interventionseffekt hier darin, einer drohenden psychischen Verschlechterung Einhalt zu gebieten und war damit primärpräventiv entwickelt. Das Training fand an zwei Tagen für jeweils 3 Stunden statt und hatte die Trainingsinhalte: Stressverarbeitung, Selbstbewusstsein, Kontrollübernahme und Planung von konkreten Schritten. Die KG wurde als Wartegruppe realisiert, bestehend aus 37 Probanden, welche in Bezug auf Alter und Bildung mit der EG äquivalent war. Maysent und Spera testeten Interventionseffekte mit einem "Personal Strain Questionnaire", welche als eine von 8 Skalen die Subskala "Psychological Strain" enthielt. Die Autoren berichten, dass die EG nach dem Training keine Veränderung in der psychischen Beanspruchungsvariable aufwies, während die KG sich verschlechterte. Deskriptive Werte sind in der Publikation nicht angegeben. Ein follow-up wurde nicht durchgeführt.

Saam und Wodtke (1995) beschreiben ein kognitives Training für kürzlich gekündigte Manager. Sie untersuchten 42 Akademiker, die vormals im mittleren und höheren Management gearbeitet hatten und sich nach Kündigung in einem Outplacement-Programm befanden. Die Probanden wurden in EG und KG randomisiert. Interventionsteilnehmer bekamen Einzeltherapie-Sitzungen, in denen sie strukturiert lernten, irrationale und automatisch ablaufende Gedanken zu durchbrechen sowie Stress zu reduzieren. Probanden der KG erhielten Einzelsitzungen, in welchen ihnen unstrukturierte Empfehlungen zur Stressverarbeitung vermittelt wurden. Beide Gruppen erhielten 6 Beratungstermine. Saam und Wodtke setzten einen Angstfragebogen ein. Die Daten wurden vor (T1) und nach dem Training (T2) erhoben. Für die EG belegten sie eine signifikante Reduktion der situationsbezogene Angst und eine tendenzielle Minderung der Trait-Angst, während sich in der KG keine signifikanten Effekte zeigten. Ein follow-up wurde nicht durchgeführt, da nach 5 Monaten sämtliche Teilnehmer eine neue Anstellung gefunden hatten.

Frese et al. (2002) stellen eine Pilotstudie zur Erhöhung der Eigeninitiative bei Arbeitslosen vor, die sie mit 8 Arbeitslosen im prä-post-Vergleich evaluierten. Die Intervention wurde dreitägig angelegt und die Trainingsinhalte reichten von Erhöhung der Eigeninitiative über Selbstsicherheitstraining bis hin zu Übungen von selbstregulativem Verhalten. Neben einem selbst operationalisierten Eigeninitiative-Fragebogen kamen Messungen von Selbstwirksamkeitserwartung, Depressivität und psychosomatischen Beschwerden zum Einsatz. Frese et al. berichten eine tendenzielle Reduktion von Depressivität, eine signifikante Verringerung psychosomatischer Beschwerden und die Erhöhung der Selbstwirksamkeitserwartung. Ein follow-up wurde nicht durchgeführt.

1.6.1 Durchschnittlicher publizierter Interventionserfolg und Zeitstabilität

Aus den soeben narrativ vorgestellten Studien kann man das Fazit ableiten, dass die meisten publizierten Interventionen positive Effekte in verschiedenen psychologischen Konstrukten zu berichten wissen. Dies betrifft vornehmlich die Reduktion von unspezifischen psychischen Beschwerden und der Depressivität sowie die Erhöhung des Selbstbewusstseins. Falls mindestens drei Messzeitpunkte erhoben wurden, stellten sich im follow-up die Interventionseffekte entweder als relativ zeitstabil oder als instabil heraus. Numerische Rückgänge der Interventionserfolge waren in sämtlichen Studien zu finden, die ein follow-up durchführten und die deskriptiven Werte angaben. Es kann nicht abgeleitet werden, dass das Alter der Versuchsteilnehmer oder die Dauer der Arbeitslosigkeit für die Interventionseffektivität eine Rolle spielten. Im Gegensatz dazu war der Bildungsstand erheblich dafür, ob die Probanden im Anschluss wiedervermittelt wurden. Falls sie erhoben wurde, stellte die finanzielle Beanspruchung stets einen signifikanten Prädiktor für die psychische Gesundheit der Erwerbslosen dar.

Statt eines narrativen Reviews zur Thematik metaanalysierten Paul und Moser (2009a) systematisch, mit welchem Effekt durchschnittlich Interventionserfolge für die Zielgruppe erzielt werden. Sie verwendeten in ihrer Metaanalyse ausschließlich Interventionsstudien für Erwerbslose, in denen mindestens eine von 6 häufig untersuchten Indikatorvariablen zu psychischen Beschwerden im Längsschnitt getestet wurden (unspezifische psychische Belastungssymptome, Depressivitätssymptome, Angstsymptome, psychosomatische Symptome, Beeinträchtigung der Lebenszufriedenheit und vermindertes Selbstwertgefühl). Studien, in denen die Erwerbslosenstichprobe aus einer Population von Patienten aus medizinischen Einrichtungen gezogen worden war, schlossen die Autoren aus, um die Vergleichbarkeit des Interventionseffektes mit allgemeinen Arbeitslosen gewährleisten zu können. Die Metaanalyse des durchschnittlichen Erfolgs für 16 internationale Interventionsstudien geben Paul und Moser als Meta-Effektstärke Cohen's d für abhängige Stichproben an:

"As expected, intervention studies that tested the effectiveness of psychological or other programs intending to help the unemployed lead to reductions of distress symptoms. With $d=0.35$ this effect was of medium size and highly significant ($p<.001$). [...] Thus, we can conclude that the intervention programs meta-analysed here were effective because they were associated with an improvement in mental health that is not typical for permanently unemployed persons" (Paul & Moser, 2009a, S. 276).

Demnach konnte für die 16 Primärstudien von T1-T2 eine durchschnittliche Reduktion von psychischen Beschwerden mit $d=0,35$ festgestellt werden, welche hochsignifikant war. Zusätzlich betonen Paul und Moser (2009a) ausdrücklich, dass ihr metaanalyzierter Interventionserfolg den wahren mittleren Effekt von Interventionen, die mit Erwerbslosen getestet werden, durch einen annehmbaren **Publikationsbias** in unbekanntem Ausmaß überschätzen dürfte, weil erfolgreiche Interventionen eher publiziert werden als gescheiterte Interventionsversuche. Von daher wird ihr Meta-Effekt den wahren durchschnittlichen Interventionserfolg um einen Fehler zu hoch einschätzen, von dem niemand weiß, wie groß er ist, eben weil die Ergebnisse von missglückten Interventionen oft nicht verfügbar gemacht werden und deswegen auch nicht von den Autoren metaanalysiert werden konnten.

In einer separaten Publikation analysierten Paul und Moser (2009c), wie durchschnittlich **zeitstabil** ihr berechneter Meta-Interventionseffekt von $d=0,35$ war. Da viele Interventionsstudien kein follow-up durchführten, nutzten die Autoren folgende Vorgehensweise: Sie entdeckten, dass in den Primärstudien die Zeiträume vom ersten bis zum zweiten Messzeitpunkt sehr unterschiedlich waren; diese reichten von wenigen Wochen bis zu einem Jahr. Paul und Moser nutzten die Variabilität der T1-T2-Zeitintervalle, um einen mutmaßlichen Zusammenhang zwischen der vergangenen Zeit und kleiner werdenden Interventionseffekten in den Studien aufzudecken. Zu ihrem Ergebnis schreiben sie:

"Tatsächlich fand sich bei einer entsprechenden meta-regressionsanalytischen Auswertung ein hochsignifikanter Zusammenhang in dem Sinne, dass die Befindensverbesserung mit zunehmender Dauer des Zeitintervalls bis zum zweiten Messzeitpunkt immer weiter abnahm. Ein positiver Effekt auf die psychische Gesundheit tritt also kurz nach der Intervention auf, schwächt sich dann aber mehr und mehr ab" (Paul & Moser, 2009c, S. 130).

Demnach sind psychisch stabilisierende Interventionseffekte von bisherigen Studien mit Arbeitslosen durchschnittlich von kurzfristiger Dauer. Paul und Moser (2009c) schlussfolgern inhaltlich, dass für Erwerbslose, die trotz einer Intervention im Anschluss keine neue Stelle finden, ihr Hauptproblem (die Erwerbslosigkeit) ungelöst bleibt und dadurch eine negative psychische Symptomatik mit der Zeit wieder Überhand gewinnen kann.

2 Methode

Dass Arbeitslose eine schlechtere psychische und körperliche Gesundheit aufweisen als Berufstätige, ist aufgrund vieler stützender Befunde zur Thematik voraussetzbar. Für ältere, langjährig Erwerbslose kommt hinzu, dass diese vermutlich auch aufgrund einer eingeschränkten Gesundheit Selektionsprozesse durchlaufen haben, so dass gesundheitliche Beschwerden in dieser Gruppe wahrscheinlich ein gravierendes Vermittlungshemmnis darstellen. Bisherige, durch die Bundesagentur für Arbeit finanzierte Maßnahmen zur Förderung von älteren Langzeiterwerbslosen stützen sich fast ausschließlich auf den Abbau von Qualifikationsdefiziten oder auf die Motivation zur vermehrten Bewerbungsaktivität. Im gesundheitlichen Bereich steckt demnach viel ungenutztes Präventionspotential. Ein spezielles Förderungsprogramm könnte die körperliche Gesundheit stärken, das psychische Befinden verbessern und so zum Abbau von gesundheitlichen Vermittlungshemmnissen beitragen.

Das herausragende Merkmal der hier vorgestellten Interventionsmaßnahme stellt sicherlich der Schwerpunkt ihrer Förderung dar: Ältere Langzeiterwerbslose erhielten eine hauptsächlich körperliche Gesundheitsförderung. Nach einer wissensvermittelnden Edukation zur Erhöhung der Gesundheitskompetenz kam ein 8 bis 10-wöchiges Sportprogramm zur Stärkung der körperlichen Ausdauer zum Einsatz. Neben der Hoffnung, für die Zielgruppe gesundheitsbezogene Vermittlungshemmnisse abzubauen und damit die Chance auf Wiedereinstellung zu erhöhen, sollte die Gesundheitsförderung darüber hinaus wirksam sein, indem sie den Krankenkassen helfen könnte, Kosten für die Inanspruchnahme von medizinischen Diensten zu reduzieren.

In dieser Dissertation werden die psychischen Auswirkungen der Gesundheitsförderung evaluiert. Die körperlich messbaren Effekte der Intervention werden in einer separaten Arbeit am IPM in Rostock geprüft. Veränderte Wiedereinstellungschancen werden nicht examiniert, da für einen statistischen Nachweis dieser die Datenbasis zu klein wäre. Zudem bietet eine wirtschaftlich strukturschwache Region wie Mecklenburg-Vorpommern besonders erschwerte Rahmenbedingungen für eine Rückkehr in den ersten Arbeitsmarkt. Daher wird das Zielkriterium in dieser Arbeit auf die Erhöhung des psychischen und des subjektiven körperlichen Gesundheitszustandes beschränkt.

2.1 Förderung & Ethik

Die vorliegende Studie wurde unter dem Namen *"Fit50+"* durchgeführt und im Rahmen des BMBF-Programms *"Präventionsforschung zur Gesundheitsförderung und Primärprävention von älteren Menschen"* gefördert. Die Studie wurde von der Ethikkommission der Universität Rostock im Oktober 2006 geprüft und erhielt ein positives Votum (Aktenzeichen: A 26/2006).

2.2 Design

Es handelt sich um eine prospektive, hypothesenüberprüfende Zweigruppen-Längsschnittstudie mit 3 Messzeitpunkten, in welcher die Effekte einer Gesundheitsförderung mit einem körperlichen Ausdauertraining in zwei Intensitätsstufen in einem quasi-experimentellen Setting geprüft wurden. Die Messzeitpunkte lagen vor dem Training (T1), nach dem Training (T2) und nach einer follow-up-Periode (T3).



Abb. 2.1: Darstellung der Studienphasen.

2.3 Geplante Stichprobengröße

In jeder Längsschnittstudie ist zu erwarten, dass es im Laufe der Durchführung zu Studienabbruchern kommt. In dieser Studie konnten die Probanden z.B. wiedereingestellt werden, erkranken oder motivationale und persönliche Gründe konnten in der Studie zu einem Abbruch führen. Um trotz Abbrecher noch eine zufriedenstellende Datenbasis zur Verfügung zu haben, wurde anvisiert 120 Probanden zu rekrutieren.

2.3.1 Verzicht auf Kontrollgruppe

Eine Kontrollgruppe wurde aus organisatorischen und finanziellen Gründen nicht in das Design aufgenommen. Der Verzicht auf eine Kontrollgruppe ist allerdings für psychologische Längsschnittuntersuchungen mit Erwerbslosen vertretbar, weil sich in bisherigen Studien

empirisch keine sog. "Testeffekte" zwischen den Messzeitpunkten für die Zielgruppe nachweisen lassen. Als "Testeffekt" bezeichnet man in der Sozialwissenschaft das Phänomen, dass Vpn sich bei mehrmaliger Befragung implizit aufgefordert fühlen können, eine Verbesserung ihres Befindens zu beschreiben (vgl. Winefield, 1995). Solche unerwünschten Effekte wiederholter Messungen sind keine Seltenheit und wurden bereits von Wohlwill als "ärgerliches Problem" (Wohlwill, 1977, S. 181) longitudinaler Forschungsdesigns bezeichnet. Paul et al. (2006) und Paul und Moser (2009a) stellten in ihrer Metaanalyse von Längsschnittdaten fest, dass es bei kontinuierlich Berufstätigen zu einer kleinen ($d=0,06$), aber signifikanten Steigerung der psychischen Gesundheit in der Mehrfachbefragung kam. Dies traf jedoch **nicht** auf die Erwerbslosen zu:

"Bei Personen, die zu beiden Erhebungszeitpunkten arbeitslos waren, zeigt sich keine signifikante Veränderung ihrer psychischen Gesundheit, wenn man Interventionsstudien aus der Datenbasis ausschließt. Interventionsstudien auszuschließen scheint gerechtfertigt, da bei ihnen eine Befindensveränderung ja explizit angestrebt wird. Nach Ausschluss dieser Studien betrug die Effektstärke für kontinuierlich arbeitslose Personen $d=-0,03$ und war nicht signifikant". (Paul et al., 2006, S. 44).

Die Autoren stellen für den diskrepanten Testeffekt zwischen Berufstätigen und Arbeitslosen die Erklärung in den Raum, dass es bei kontinuierlich Erwerbslosen eventuell auch zu einem schwachen Testeffekt käme, aber dieser dadurch neutralisiert würde, weil mit der Zeit ihre psychischen Beschwerden eigentlich zunähmen (vgl. Zusammenhang von Arbeitslosigkeitsdauer und psychischer Beanspruchung, Kap. 1.3). Paul et al. (2006) sowie Paul und Moser (2009a) erbrachten somit den Meta-Nachweis, dass kontinuierlich Erwerbslose im Längsschnitt nicht spontan ihre psychische Gesundheit verbessern, welches eine Kontrollgruppe in Form einer Wartegruppe für die Zielgruppe fast redundant macht.

2.3.2 Interventionsgruppen

Die Probanden waren über 50-jährig und langfristig Erwerbslose (>1 Jahr), die sämtlich aus dem Raum Schwerin stammten. Die Untersuchung erfolgte in zwei Erhebungswellen: Erst nachdem die erste Erhebungswelle inklusive follow-up abgeschlossen war, wurden neue Probanden für die zweite Erhebungsgruppe rekrutiert. Dies lag darin begründet, dass das körperliche Training nach der ersten Durchführung sportwissenschaftlich optimiert werden sollte. Insgesamt wurden 119 Probanden in zwei Erhebungswellen medizinisch-psychologisch untersucht und anschließend mit dem jeweiligen Ausdauertraining trainiert. Weil die Interventionsgruppen zeitlich disjunkt diagnostiziert und trainiert wurden, sind sie nachfolgend als "**Interventionskohorten**" bezeichnet.

Interventionskohorte 1 (N=63) wurde von September 2007 bis April 2008 rekrutiert. Für die letzten Versuchsteilnehmer erfolgte die follow-up-Untersuchung im Dezember 2008.

Interventionskohorte 2 (N=56) fand von März 2009 bis Juli 2009 Eingang in die Studie; die letzten Probanden wurden Ende März 2010 im follow-up untersucht.

Einschlusskriterien zur Aufnahme in die Stichprobe:

- Das Kriterium von "Langzeiterwerbslosigkeit" (mind. 1 Jahr arbeitslos) sollte erfüllt sein.
- Die Probanden sollten das 50. Lebensjahr vollendet haben.

Ausschlusskriterien zur Ablehnung in die Stichprobe:

- Körperliche Gesundheitszustände, bei denen ein sportliches Training kontraindiziert oder schlicht nicht möglich war oder eine akute, schwere Alkoholkrankheit oder Rückfall bei chronischer Alkoholkrankheit. Die Prüfung dieser Kriterien oblag einer Ärztin am IMP.
- Sehr mangelhafte Deutschkenntnisse, die befürchten ließen, dass die Teilnehmer die psychologischen Fragebögen sprachlich nicht verstehen könnten (Migrationshintergrund stellte kein Ausschlusskriterium dar). Die Einschätzung der Deutschkenntnisse erfolgte durch die "Jobmentoren" eines Bildungsträgers, welcher die Rekrutierung übernahm.
- Personen über 60 Jahre.

2.3.3 Rekrutierung der Probanden

Die ARGE Schwerin teilte einem privaten Bildungsträger ältere Langzeiterwerbslose zu. Dort wurden die "Kunden" (offizielle Bezeichnung) für einen Zeitraum von ca. 10 Monaten von einem "Jobmentor" betreut. Die Zuweisung durch die ARGE war für alle Kunden verpflichtend. Im Verlauf des Betreuungszeitraumes mussten zwei Bildungskurse gewählt werden. Neben der Teilnahme am Gesundheitsprogramm konnten die Kunden sich für weitere typische Qualifizierungsmaßnahmen (z.B. Computerkurse, kaufmännisches Wissen etc.) entscheiden. 70% der Kunden entschieden sich für den Gesundheitskurs.

Die Jobmentoren waren instruiert, die Kunden zu einer Teilnahme am Gesundheitsprogramm motivierend anzuregen. Falls der Jobmentor aufgrund seines Categorieschemas ein sog. "gesundheitliches Vermittlungshemmnis" feststellte, sollte verstärkt motiviert werden. Eine Teilnahme am Gesundheitsprogramm implizierte keinen Zwang an der Studie teilzunehmen. Zur Teilnahme an der Studie wurden die Kunden damit motiviert, dass sie am IMP die Gelegenheit hatten, umfassend medizinisch untersucht zu werden, inklusive ärztlicher Rückmeldung und individueller Gesundheitsberatung. Des Weiteren wurden die Probanden für die Untersuchungstage von den Verpflichtungen des Bildungsträgers freigestellt, erhielten das Fahrgeld von Schwerin nach Rostock und zurück sowie ein zusätzliches Probandengeld von € 15,- für jede durchgeführte Untersuchung.

Probanden sollten vor der Teilnahme am gesundheitsedukativen Kurs (vgl. Kap. 2.4) im IMP untersucht werden. Dies ließ sich aus organisatorischen Gründen leider nicht durchgängig gewährleisten, da die Gesundheitskurse des Bildungsträgers mit einer Teilnehmerzahl von 20 Kunden durchgeführt wurden und alle zeitgleich mit dem Unterricht beginnen mussten. Demgegenüber war die Rekrutierung der Studienteilnehmer mit höherem Aufwand und terminlichen Herausforderungen verbunden. Da die Teilnahme an den Untersuchungsterminen freiwillig und zudem für die Probanden zeitaufwändig war, musste bei den Terminabsprachen Rücksicht auf deren terminliche Wünsche genommen werden. Einige Kunden ließen sich auch erst zu einer Teilnahme an der Studie bewegen, nachdem sie einige Tage anhand des Gesundheitskurses mit den Inhalten vertraut waren. So wurden viele Probanden für die Eingangsdiagnostik erst während der wissensvermittelnden Kurse rekrutiert. Gewährleistet war jedoch stets, dass die Erstuntersuchung vor Beginn des 8- bzw. 10-wöchigen Sportprogramms stattfand.

2.4 Interventionsprogramm

Um die Integrationschancen älterer Langzeitarbeitsloser zu verbessern, wurde das Gesundheitsprogramm vom Bildungsträger in Schwerin entwickelt und seit Dezember 2006 durchgeführt. Mitarbeiter des IPM waren daran nicht beteiligt. Für den **gesundheitsedukativen Kurs** erreichte der Bildungsträger eine branchenübliche Zertifizierung und die ARGE Schwerin finanzierte den Unterricht mit Bildungsgutscheinen. Die Fachdozenten stammten aus Gesundheitsberufen. Die Gesundheitsedukation fand in den Seminarräumen des Bildungsträgers statt und umfasste folgende Kernthemen:

Tab. 2.1: Inhalte der Gesundheitsedukation.

Titel	Inhalte
Gesundheits-Check	- Analyse des Ist-Gesundheitszustandes - Motivation zur Bewegung und aktiv-sein - Überwindung der negativen Symptomatik von Arbeitslosigkeit
Gesundheit und Wohlbefinden	- Was ist Gesundheit? Zusammenhang Gesundheit und Wohlbefinden - Risiken von Bewegungsmangel, Herz-Kreislauf- und Abhängigkeitserkrankungen - Einschätzung der eigenen Leistungs- und Belastungsfähigkeit - Gesundheitspflege
Anatomisches Wissen, Rückenschule	- Gelenk- und Rückenprobleme, Erkrankungen am Bewegungsapparat - Analyse und Korrektur von Fehlhaltungen in unterschiedlichen Körperpositionen - Zusammenhänge zwischen Herzfrequenz, Atemvolumen und Belastung
Ernährungsberatung	- Stoffwechsel, Nahrungspyramide, Ernährungsplanung, Risiken falscher Ernährung - Ernährungsplanung, Vielseitigkeit und Genuss - Kochtipps zur schonenden Zubereitung des Essens - Übungen für eine finanziell sparsame und trotzdem gesunde Ernährungsweise
Umgang mit psychischen Belastungen	- Was bedeutet Stress? Identifikation von potentiellen Stressoren - Unterschied zw. Belastung und Beanspruchung, Möglichkeiten des Copings - Vorstellung von Atemübungen, Yoga, Taiji

Wie anhand des Curriculums ersichtlich, war das Hauptziel die Vermittlung von körperlicher Gesundheitskompetenz. Zwar wurde auch Wissen zu den Themen "Stress" und "Beanspruchung" vermittelt sowie Entspannungsverfahren vorgestellt, aber es fand keine psychologische Intervention z.B. in Form eines kognitiv-behavioralen Gruppentrainings statt. Der Kurs wurde in einem Umfang von 40 Stunden mit einer Gruppengröße von 20 Teilnehmern an 8 Terminen mit jeweils 5 Stunden durchgeführt. Der Unterricht fand an 2 Tagen/Woche statt, so dass er innerhalb von 4 Wochen absolvierbar war. Die Anwesenheit der Teilnehmer wurde stets kontrolliert.

Das **körperliche Training** wurde in einem Fitness-Studio in Schwerin durchgeführt. Es wurde zunächst über eigene Mittel des Bildungsträgers, später im Rahmen der "Perspektive 50plus" durch das Bundesministerium für Arbeit und Soziales finanziert. Ein Sozialpädagoge mit sportpädagogischer Qualifikation und/oder ein Fitnesstrainer (abh. von Interventionskohorte) betreuten die Probanden. Die Trainingsintensität musste adaptiv gestaltet werden und richtete sich nach dem Leistungsvermögen der Teilnehmer. Die meisten Probanden hatten 3 Jahrzehnte keinen Sport mehr getrieben und entsprechend niedrig war das Leistungsniveau zu Beginn. Viele Pausen und eine motivierende Haltung der Trainer ermöglichten eine behutsame Heranführung an die Trainingsinhalte. Im Laufe der Zeit besserte sich die Kondition der Probanden und die Intensität konnte gesteigert werden. Die Trainingsgruppen umfassten 12-15 Teilnehmer und waren stabil, so dass die Probanden die Gelegenheit hatten, sich

kennenzulernen. Die Trainer offerierten auch individuelle Beratungen zur Optimierung des Sportprogramms. Für jeden Trainingstermin erfolgte eine Anwesenheitsprüfung.

Tab. 2.2: Inhalte des Trainingprogramms.

Trainings-Schwerpunkte	Inhalte
Herz-Kreislauf Training	<ul style="list-style-type: none"> - Bestimmung und Verbesserung der Ausdauer und Belastbarkeit - Training mit Einsatz von Cardiogeräten (Crosstrainer, Rudermaschinen etc.) - Korrektur und Optimierung von Bewegungsabläufen - Herzschlagfrequenzgesteuertes Training
Circle-Training	<ul style="list-style-type: none"> - Training zur Stärkung der großen Muskelgruppen - Training der Kraft und Ausdauer mit Übungen und Trainingsgeräten - Verbesserte Wahrnehmung der subjektiven körperlichen Leistungsfähigkeit - Stärkung der Skelettmuskulatur
Rückentraining	<ul style="list-style-type: none"> - Kraft-Ausdauertraining zur Festigung des Halte- und Stützapparates - Übungen zum Verinnerlichen rückenschonenden Verhaltens - Trainingsübungen zum Stehen, Sitzen, Bücken, Heben
Haltungs- und Entspannungs-training	<ul style="list-style-type: none"> - Angebote zur Identifizierung individuell passender Entspannungs Übungen (Taiji, Qigong, Yoga, Atemübungen, "Traumreisen") - Stretching, Abbau von Verspannungen, Entlastung der Wirbelsäule und Gelenke

Die Kohorten unterschieden sich im körperlichen Training wie folgt:

1. Interventionskohorte: Das körperliche Training umfasste 40 Stunden, fand an 1-2 Terminen pro Woche statt und wurde von einem geschulten Fitnesstrainer durchgeführt. Neben dem körperlichen Training wurde die Zeit auch genutzt, um Trainingswissen zu vermitteln. Pro Trainingstermin wurden Übungen zu einem der in Tab. 2.2 genannten Schwerpunkte durchgeführt. Der Kurs konnte innerhalb von 8 Wochen absolviert werden.

2. Interventionskohorte: Nach der ersten Erhebungswelle wurde das Bewegungsprogramm auf das Ziel hin optimiert, ein verbessertes Training zur Prävention von Herz-Kreislaufkrankungen bereitzustellen. Hiermit sollten auch medizinische Hypothesen in Bezug auf die Verbesserung von körperlichen Variablen im Kontrast zur 1. Kohorte prüfbar gemacht werden. Zu diesem Zweck erfolgte eine Erhöhung der Trainingsfrequenz und eine stärkere Konzentration auf das Ausdauertraining. Pro Termin wurden sämtliche Schwerpunkte aus Tab. 2.2 angewendet. Die Durchführung des Programms unterstand einem Sozialpädagogen mit sporttherapeutischer Qualifikation in Assistenz mit dem Fitnesstrainer. Das Training umfasste 60 Stunden bei stets 2 Terminen pro Woche und war innerhalb von 10 Wochen absolvierbar.

2.5 Follow-up

Im 6-monatigen follow-up nach der Intervention und des zweiten Messzeitpunktes bekamen alle Studienteilnehmer die Offerte, kostenlos im selben Fitnesscenter weiter trainieren zu dürfen. Das follow-up-Training konnte frei genutzt werden, mit der einzigen Einschränkung, dass die Trainingszeiten auf den Vormittag zu begrenzen waren. Für Trainingsfrequenz, -dauer und -intensität gab es keine Vorgaben. An 2 Tagen/Woche konnte zudem am angeleiteten Circle-Training in der Gruppe teilgenommen werden.

2.6 Untersuchungsablauf

Die Untersuchungen wurden im IPM Rostock durchgeführt, in dem für die Studie zwei Labors eingerichtet waren: Ein medizinisches, welches eine promovierte Ärztin mit Assistenz einer MTA belegte und ein psychologisches Labor. Die psychologischen Untersuchungen wurden vom Verfasser dieser Arbeit oder einer weiteren im IPM angestellten promovierten Ärztin durchgeführt. Im psychologischen Labor fanden neben den psychometrischen Erhebungen auch medizinische Messungen zum Ruhe-Blutdruck und zur Herzfrequenzvariabilität während einer 40-minütigen Ruhephase statt. Die Erhebung der ausschließlich psychologischen Daten beanspruchte etwas über eine Stunde. Die Gesamtuntersuchungszeit im IPM umfasste inklusive Pause ca. 4 Stunden. An den Untersuchungstagen reisten morgens jeweils 2 Probanden an; um ca. 9 Uhr konnte die Diagnostik begonnen werden. Die Probanden wurden zeitgleich entweder zuerst medizinisch oder psychologisch untersucht. Nach ca. 2 Stunden wurde eine 15-minütige Pause mit Früchtetee und Gebäck angeboten, danach wechselten die Probanden die Labore und die Untersuchungen setzten sich vice versa fort. Die Gesamtuntersuchung konnte um ca. 13 Uhr abgeschlossen werden. Die Diagnosen der medizinischen Untersuchung wurden den Probanden inklusive einer medizinischen Gesundheitsberatung rückgemeldet, die Ergebnisse der psychologischen Erhebung nicht.

2.7 Ebenen zur Erfassung der Gesundheit

Das Konzept "Gesundheit" ist laut Verfassung der WHO als "Wohlergehen"⁴ aufzufassen, das nicht nur als Abwesenheit von Krankheit zu definieren ist. Einflussfaktoren auf die Gesundheit wurden in Anlehnung an die WHO-Definition mittels der folgenden drei Ebenen betrachtet. Die einzelnen Messverfahren hierzu werden detailliert im Anschluss vorgestellt.

1. Risikofaktoren der psychischen Gesundheit: Ermittlung von psychischen Beschwerden mit folgenden Instrumenten:

- *General Health Questionnaire (GHQ)*, zur Messung von unspezifischen psychischen Beschwerden.
- *Beck Depressions-Inventar (BDI)*, zur speziellen Messung von Depressivität.
- *SF12 Psychische Summenskala (PSK)*, zur Messung der erlebten Einschränkungen der Lebensqualität durch psychische Beschwerden.
- *Lebenszufriedenheit (LEB)*, zur Erfassung einer geminderten Lebenszufriedenheit.

2. Ressourcen der psychischen Gesundheit: Messung von relienzförderlichen bzw. salutogenetisch wirksamen psychischen Ressourcen. Hiermit sollte der Forderung der WHO nachgekommen werden, Gesundheit nicht ausschließlich als Abwesenheit von Beschwerden zu definieren. Folgende Instrumente dienten der Erhebung:

- *Sense of Coherence Skala (SOC)*, zur Messung des Kohärenzsinns.
- *Allgemeine Selbstwirksamkeitserwartung (SWE)*, zur Messung der erwarteten Kompetenzüberzeugung und mittelbar auch der Kontrollüberzeugung.

⁴ "Gesundheit ist ein Zustand des vollständigen körperlichen, geistigen und sozialen Wohlergehens und nicht nur das Fehlen von Krankheit oder Gebrechen." Verfassung der World Health Organisation (1947, S. 2).

→ *Psychosoziales Wohlbefinden (PSW)*, zur Messung ebendieses. Dieses Merkmal wird von seinen Autoren als salutogenetisch wirksam definiert.

3. Körperliches Wohlergehen: Erhebung der subjektiven körperlichen Gesundheit mit folgenden Messverfahren:

- *Visuelle Analogskala (VAS)*, zur Erfassung des körperlichen Gesundheitsempfindens.
- *SF12 Körperliche Summenskala (KSK)*, zur Messung der körperlichen gesundheitsbezogenen Lebensqualität.

4. Passives Freizeitverhalten: Als Messgröße eines prototypischen, passiven Freizeitverhaltens wurde der Fernsehkonsum erhoben. Obwohl das Freizeitverhalten selbst keine Gesundheitsvariable repräsentiert, wurde ihm unterstellt, einen vermittelnden Einfluss auf die Gesundheit zu haben.

Gesundheit	Psychische Gesundheit	Risikofaktoren	Unspezifische psychische Beschwerden (GHQ)
			Depressivität (BDI)
		Ressourcen	Psych. Einschränkung des Lebens (SF12-PSK)
			Verminderte Lebenszufriedenheit (LEB) ⁵
	Körperliche Gesundheit		Kohärenzsinn (SOC)
			Selbstwirksamkeitserwartung (SWE)
Passives Freizeitverhalten		Psychosoziales Wohlbefinden (PSW)	
		Subjektives körperliches Gesundheitsempfinden (VAS)	
		Körperliche Einschränkungen der Lebensqualität (SF12-KSK)	
		Fernsehkonsum (Zeit/Tag)	

Abb. 2.2: Darstellung der konzeptionellen Einordnung der Messgrößen.

Die Unterscheidung der psychischen Messgrößen in einerseits Risikofaktoren und andererseits Ressourcen fußt auf den konzeptionellen Entwurf der ressourcenorientierten Skalen: Sowohl der *Kohärenzsinn* (Antonovsky, 1987; Schumacher & Brähler, 2000) als auch die *Selbstwirksamkeitserwartung* (Schwarzer & Jerusalem, 1999) und das *psychosoziale Wohlbefinden* (Rimann & Udris, 1993) werden von den Testautoren ausdrücklich als salutogenetische bzw. resilienzförderliche Ressourcen des psychischen Wohlergehens definiert. Rein technisch unterscheiden sich ressourcenorientierte Skalen durch einen starken Überhang (oder ausschließlicher Verwendung) von semantisch positiv formulierten Items (z.B. "Ich spüre, wie schön das Leben ist") im Gegensatz zu denen von risikoorientierten Skalen, die häufiger (oder ausschließlich) inhaltlich negativ verfasste Items (z.B. "Ich fühle mich als Versager") beinhalten.

Anmerkung: Da ein Auftrag des BMBF-Programms lautete, potentiell geeignete Messinstrumente für die Evaluation des vorliegenden Gesundheitsprogramms zu identifizieren, wurden mehr Fragebögen als üblich für die Untersuchungen verwendet.

⁵ "LEB" wurde aus der einmaligen Anwendung des Persönlichkeitsfragebogens FPI-R extrahiert und stand damit nur für die Analysen im Querschnitt zur Verfügung.

2.7.1 Vorstellung der Messinstrumente

Im Folgenden werden die soeben kurz genannten Instrumente sowie die zu erhebenden Konstrukte detaillierter dargestellt. Die nicht käuflich zu erwerbenden Fragebögen sind im Anhang einsehbar.

Soziodemographischer Fragebogen

Während der psychologischen Erstuntersuchung wurden zunächst soziodemographische Merkmale abgefragt, welcher in starker Anlehnung an den Fragebogen des TAURIS Projektes (Nitsche & Richter, 2003) zusammengestellt war (Anhang 7.2).

Freiburger Persönlichkeitsinventar (FPI)

Einerseits werden Persönlichkeitsmerkmale als relativ zeitstabil aufgefasst und von daher könnte annehmbar sein, dass sie "prämorbid" bei der hiesigen Zielgruppe, also bereits vor der Arbeitslosigkeit, vorhanden waren. Ein auffälliges Persönlichkeitsprofil wäre demnach in Richtung der *Selektionshypothese* (Kap. 1.2) deutbar. Andererseits unterliegen auch Persönlichkeitsmerkmale Änderungen. Als persönlichkeitsverändernd vermutet man sog. *kritische Lebensereignisse*. Diese werden als so einschneidend empfunden, dass sie zu tief greifenden Erschütterungen des Selbst- und Weltbildes führen und die Persönlichkeit nachhaltig verändern. Der Arbeitsplatzverlust zählt laut Filipp und Aymanns (2010) als kritisches Lebensereignis. Zur Erfassung der Persönlichkeitsmerkmale wurde der FPI-R nach Fahrenberg et al. (2010) verwendet. Es werden 138 Fragen, die mit "stimmt" oder "stimmt nicht" zu beantworten sind, auf 12 Dimensionen der Persönlichkeit verdichtet. Die Persönlichkeitsfaktoren sind nicht unabhängig (orthogonal) voneinander, sondern wurden von den Testautoren nach inhaltlichen Kriterien faktorisiert und zusammengestellt. Für die hiesige Studie war das Persönlichkeitsmerkmal "Lebenszufriedenheit" interessant, da es inhaltlich mit dem Indikator "subjektives Wohlbefinden" der Metaanalyse von Paul et al. (2006) übereinstimmt. Eine *eingeschränkte Lebenszufriedenheit* wurde als Risikofaktor der psychischen Gesundheit definiert. Der FPI kam einmalig zur Anwendung und stand deshalb ausschließlich für Querschnittsanalysen zur Verfügung.

General Health Questionnaire (GHQ)

Der GHQ war in den von Paul et al. (2006) sowie Paul und Moser (2009a) metaanalysierten Primärstudien das am meisten verwendete Instrument, sowohl in den Querschnitts- als auch in den Längsschnittstudien mit Erwerbslosen. Er wurde als Indikator für *unspezifische psychische Beschwerden* als Risikofaktor der psychischen Gesundheit für sämtliche Erhebungszeitpunkte erhoben. Der GHQ kam in der deutschen 12-Item-Version (übersetzt von Linden et al., 1996) zum Einsatz. Dieser ursprünglich von Goldberg (1978) und Goldberg und Hillier (1979) entwickelte Fragebogen wird oft im klinischen Bereich für Veränderungsmessungen eingesetzt. Er gehört zur Gruppe der globalen klinischen Tests und dient der Identifikation von allgemeinen psychischen Beschwerden (Linden, et al., 1996). Die Items werden mit einer 4-stufigen Skala (0-3) beantwortet (vgl. Anhang 7.3).

Beck Depressions-Inventar (BDI)

Studien, welche die Interventionseffektivität für Arbeitslose prüfen, setzen häufig Depressivitätsskalen für die Veränderungsmessung ein (z.B. Vinokur, et al., 1995; Vuori & Silvonon, 2005). Der BDI wurde als Indikator eines Risikofaktors der psychischen Gesundheit erhoben und zu allen Messzeitpunkten eingesetzt. Das Depressions-Inventar von Beck ist eines der bekanntesten Instrumente in der klinisch-psychologischen Diagnostik zum Screening einer depressiven Episode. Zahlreiche Untersuchungen bescheinigen dem BDI eine gute Zuverlässigkeit und Validität (z.B. Richter et al., 1997; Kuhner, Burger, Keller, & Hautzinger, 2007). Schmitt et al. (2003) validierten eine hinsichtlich der Bearbeitungszeit ökonomisierte Form des BDI, die sie "BDI-V" nennen. In dieser teilt der Proband den Grad seiner Zustimmung zu jeweils 20 Aussagen auf einer 6-stufigen Skala (0-5) mit, anstatt wie im Original BDI, eine von vier abgestuften, semantisch verschiedenen Antwortmöglichkeiten auszuwählen (vgl. Anhang 7.4).

Sense of Coherence Skala (SOC)

Vastamaki et al. (2009) diskutieren Zugewinne im Kohärenzsinn von Arbeitslosen, die nach einer Wiedereinstellung im ersten Arbeitsmarkt feststellbar waren und weisen ferner darauf hin, dass die Veränderbarkeit des SOC in ihrer Stichprobe keinen Alterseffekt aufwies, welches die Skala als Instrument für eine Veränderungsmessung interessant machte. Das Modell der Salutogenese von Antonovsky (1987) rückte in Abkehr vom pathogenetischen Erklärungsansätzen die Frage in den Mittelpunkt, was Menschen trotz diverser, allgegenwärtiger Belastungen gesund erhält bzw. gesunden lässt (Margraf, 1998). Im salutogenetischen Modell erfolgt die Betrachtung von Gesundheit und Krankheit nicht dichotom, sondern als Endpunkte eines Kontinuums, in welchem ein Mensch nicht entweder gesund oder krank, sondern lediglich mehr oder weniger gesund oder krank ist (Antonovsky & Franke, 1997). Im Mittelpunkt des Salutogenesemodells steht das Konzept des **Kohärenzsинns** ("sense of coherence"), das definiert wird als

"...eine globale Orientierung, die zum Ausdruck bringt, in welchem Umfang man ein generalisiertes, überdauerndes und dynamisches Gefühl des Vertrauens besitzt, dass die eigene innere und äußere Umwelt vorhersagbar ist und dass mit großer Wahrscheinlichkeit die Dinge sich so entwickeln werden, wie man es vernünftigerweise erwarten kann" (Antonovsky, 1987; Übersetzung von Becker, 1997, S.10).

Das Kohärenzgefühl stellt eine Bewältigungsressource, die Menschen widerstandsfähiger gegenüber Stressoren macht, dar. Es umfasst folgende Komponenten:

1. **Verstehbarkeit** bezeichnet das Ausmaß, indem Reize, Ereignisse oder Entwicklungen als strukturiert, geordnet und vorhersehbar wahrgenommen werden.
2. **Handhabbarkeit** meint das Ausmaß, indem eine Person personale und soziale Ressourcen wahrnimmt, um in- und externe Anforderungen bewältigen zu können.
3. **Sinnhaftigkeit** reflektiert, inwieweit eine Person ihr Leben als sinnvoll empfindet und zumindest einige der vom Leben gestellten Anforderungen als Herausforderungen betrachtet, die Engagement und Investitionen verdienen.

In der hiesigen Untersuchung wurde die "Leipziger Kurzsкала" (Schumacher & Brähler, 2000) zur Erfassung des Kohärenzgefühls verwendet (Anhang 7.6). 2 Items reflektieren "Verstehbarkeit", 3 Items "Handhabbarkeit" und 4 Items messen "Sinnhaftigkeit". Die Zustimmung

oder Ablehnung zu den Items wird 7-stufig (1-7) erhoben. Der SOC wurde als Indikator für psychische Ressourcen erhoben und zu allen Terminen eingesetzt.

Allgemeine Selbstwirksamkeitserwartung (SWE)

Manche Interventionsstudien mit Arbeitslosen berichten Hinzugewinne der Selbstwirksamkeitserwartung (z.B. Frese, et al., 2002). Die SWE wurde als Indikator einer Ressource der psychischen Gesundheit zu allen Erhebungsterminen erhoben. Das Konzept der Selbstwirksamkeitserwartung ("perceived self-efficacy") wurde von Bandura (1977) als wesentlicher Aspekt seiner sozial-kognitiven Theorie eingeführt. Die Allgemeine Selbstwirksamkeitserwartung bezieht sich auf die persönliche Einschätzung der eigenen Kompetenzen, allgemein im täglichen Leben mit Schwierigkeiten und Barrieren zurecht zu kommen und kritische Anforderungssituationen aus eigener Kraft erfolgreich bewältigen zu können. Sie stellt eine personale Bewältigungsressource mit prädiktivem Wert für das Wohlbefinden und eine konstruktive Lebensbewältigung dar (Schwarzer, 1994). Die Skala wurde von Jerusalem und Schwarzer (1984) sowie Schwarzer und Jerusalem (1995) entwickelt. Die deutschsprachige Version ist in Schwarzer und Jerusalem (1999) publiziert. Die SWE umfasst 10 Items, die 4-stufig (1-4) zu beantworten sind (Anhang 7.5). Mittelbar stellt die Skala auch ein Indiz für die "Internale Kontrollüberzeugung"⁶ (Rotter, 1975) dar.

Psychosoziales Wohlbefinden (PSW)

Der Fragebogen kam bei der Evaluation des TAURIS Projektes (Nitsche & Richter, 2003), einer Intervention für Langzeiterwerbslose (vgl. Kap. 1.6), zum Einsatz. Die Skala misst salutogenetisch wirksame Ressourcen der psychischen Gesundheit und kam in sämtlichen Untersuchungsterminen zur Anwendung. Der Fragebogen zum psychosozialen Wohlbefinden ist ein von Rimann und Udris vorgestelltes Instrument zur kurzen Erfassung des "well being" mithilfe von 6 Items, die 4-fach gestuft sind (1-4) (Anhang 7.7). Die faktorielle Struktur dieses Screening Instrumentes ist unidimensional und soll laut der Autoren "...als Ressource, die einen salutogenetischen Beitrag für die Gesundheit leisten kann" (Rimann & Udris, 1993, S. 24) interpretiert werden. Für den Fragebogen existiert allerdings keine Vergleichsstichprobe anhand derer empirisch ermittelte Werte standardisiert werden können, weswegen sich aus dem Instrument kein Interventionsbedarf (vgl. Kap. 3.2) ableiten ließ.

Fragebogen zum Gesundheitszustand (SF12)

Da sich das Gesundheitsprogramm vornehmlich auf ein körperliches Training stützte, war die Änderung der gesundheitsbezogenen Lebensqualität interessant. Neben seiner weiten Verbreitung liegt der Vorteil des SF12 darin, dass die *psychische Summenskala (PSK)* und die *körperliche Summenskala (KSK)* als Indikatoren der Lebensqualität getrennt ausgewertet werden. Diese liefern sehr eigenständige Daten, d.h. sie korrelieren nicht miteinander. Die PSK wurde als Risikofaktor der psychischen Gesundheit und die KSK als Indikator für die körperliche Gesundheit erhoben. Beide kamen zu allen Messzeitpunkten zur Anwendung.

⁶ "Internale Kontrollüberzeugung": Inwieweit glaubt ein Mensch, dass er selbst für das Geschick in seinem Leben verantwortlich ist. Das Antonym ist die "externale Kontrollüberzeugung": Inwieweit glaubt ein Mensch, dass Umweltfaktoren (z.B. "die Gesellschaft") sein Leben schicksalhaft bestimmen.

Der SF12 entstammt dem SF36, welcher ein etabliertes Messinstrument für die gesundheitsbezogene Lebensqualitätsforschung darstellt (Ravens-Sieberer, 2002) und international psychometrisch normiert wurde. Die Testautoren beschreiben, dass der SF12 mit dem Ziel entwickelt wurde, 90% der Gesamtvarianz aus der körperlichen und psychischen Summenskala des SF36 einer US-amerikanischen Normstichprobe zu reproduzieren (Bullinger & Kirchberger, 1998, S. 65). Die 12 Items des Fragebogens werden mit unterschiedlicher Skalierung beantwortet.

Visuelle Analogskala des Gesundheitsempfindens (VAS)

Die VAS ist ein einfaches und zeitökonomisch einsetzbares Instrument zur Erfassung des subjektiven Gesundheitszustandes. Ihre hohe Auflösung (0-100) ließ hoffen, dass sich auch kleinere Veränderungen darin abbilden. Die Skala wurde zur Messung des körperlichen Gesundheitsempfindens zu allen Testzeitpunkten eingesetzt. Die VAS ist dem *EuroQOL*-Fragebogen (vgl. z.B. Greiner & Claes, 2007) entnommen. In diesem Instrument ist die Einschätzung der eigenen Gesundheit als Skala folgendermaßen operationalisiert: 0 wird vorgegeben als "schlechtester denkbare Gesundheitszustand" und 100 als "bester denkbare Gesundheitszustand" (Anhang 7.8).

Fernseh- und Medienkonsum

Da von Arbeitslosen gemeinhin angenommen wird, dass sie oftmals unstrukturiert und wenig aktiv den Tag verbringen (z.B. Jahoda, 1983; Warr & Jackson, 1985), lag es nahe, den Fernseh- und Medienkonsum der Probanden zu erheben. Neben dem Fernsehen wurden die Probanden auch nach ihren Gewohnheiten im Internet befragt. Es wurde "surfen" im Internet, also der auf Vergnügen oder Zeitvertreib abzielende Internetgebrauch erhoben. Des Weiteren erfolgte die Befragung nach der Nutzung von Computerspielen. Die tägliche Nutzungshäufigkeit ist pro Medienkategorie 8-stufig in 2-Stunden Schritten (vgl. Anhang 7.9) abgefragt. Der Fragebogen wurde selbst erstellt und zu allen Messzeitpunkten vorgelegt. Speziell der Fernsehkonsum wurde wegen seiner hohen Verbreitung als Indikator eines prototypischen passiven Freizeitverhaltens definiert.

2.8 Ziele und Hypothesen zu den Querschnittsdaten

Der Interventionsbedarf zur Gesundheit sollte anhand von standardisierten Messverfahren in Abgleich mit Normwerten von Vergleichsstudien festgestellt werden. Mit dem Vergleich von Ist- zu Sollwerten erfolgte die Schätzung des Interventionsbedarfs. Aufgrund der Literatur (z.B. Holleederer & Brand, 2006; Kieselbach, et al., 2006; Paul et al., 2006) war davon auszugehen, dass Langzeiterwerbslose höhere Quoten einer eingeschränkten Gesundheit aufweisen werden als in den jeweiligen Vergleichsstichproben zur Allgemeinbevölkerung.

2.8.1 Querschnittshypothesen Q1, Q2, Q3

Ziel der Formulierung der folgenden Querschnittshypothesen waren statistisch testbare Erwartungen darüber, durch welche Faktoren der erfasste Gesundheitszustand der Probanden zusätzlich beeinflusst sein könnte.

Q1: Hypothese zur Auswirkung von Antworttendenzen auf die erfasste Gesundheit

Die Schätzung des Interventionsbedarfs sollte gegen eventuelle Aggravationstendenzen abgesichert werden. Für jeden Probanden erfolgte vor der Untersuchung eine Aufklärung darüber, dass ihre Angaben dem Datenschutz unterliegen, sie in anonymisierter Form gespeichert und lediglich in Form von Gruppenmittelwerten veröffentlicht werden (vgl. Einverständniserklärung, Anhang 7.1). Es wurde ihnen verdeutlicht, dass sie durch ihre Beantwortung weder Vorteile zu erwarten hätten noch dass ihnen Nachteile bei wahrheitsgemäßer Beantwortung drohen könnten.

In Fragebogendaten ist es möglich, dass sich je nach Stichprobe verschiedene Biases ergeben. Problematisch für die Interpretation von psychologischen bzw. medizinischen Diagnosen sind Aggravations- bzw. Simulationstendenzen, d.h. wenn Probanden mit unwahrheitsgemäßen Antworten zur Übertreibung ihrer Beschwerdesymptomatik neigen. Mit solchen Tendenzen kann gerechnet werden, wenn postulierbar ist, dass sich die Subjekte durch die Aggravation eines eingeschränkten gesundheitlichen Zustandes Vorteile erhoffen (z.B. bescheinigte Arbeitsunfähigkeit zur vorzeitigen Beantragung der Rente oder im Rahmen des sog. "sekundären Krankheitsgewinns").

Demgegenüber bedeutet eine sozial erwünschte Antworttendenz, dass Menschen sich an gesellschaftliche Werte orientieren und ihre Angaben dahingehend modifizieren, die eigene Persönlichkeit in ein positives Licht zu rücken. Ein gewisser sozial erwünschter Bias ist völlig normal, aber diese Antworttendenz neigt zur verkleinerten Darstellung von (gesundheitlichen) Problemen und muss deshalb als Gegenphänomen zur Aggravationstendenz aufgefasst werden. Die Überlegung war, dass, falls ein Zusammenhang zwischen sozial erwünschtem Bias und sich verkleinernde gesundheitliche Probleme nachgewiesen werden könnte, eine Aggravationstendenz in der Stichprobe durchschnittlich falsifiziert wäre.

Das FPI Merkmal "Offenheit" stellt eine normierte 9-stufige Variable zur Einschätzung des Wahrheitsgehalts von Antworten bereit. Je höher ihr Wert, desto offener bzw. wahrhaftiger ist die Antworttendenz. Je verschlossener und unwahrhaftiger, desto sozial erwünschter ist die Antworttendenz laut Fahrenberg et al. (2010). Um die Wirkung eines sozial erwünschten Bias auf die Erfassung der Gesundheit abzuschätzen, sollten Zusammenhangsprüfungen des Offenheitsmerkmals mit den Gesundheitsindikatoren durchgeführt werden. Folgende Zusammenhänge wurden erwartet:

Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit LEB zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt negativ mit BDI zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit PSK zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit SOC zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit SWE zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit PSW zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit KSK zusammen
Verschlossene bzw. sozial erwünschte Antworttendenz hängt positiv mit VAS zusammen

Q2: Hypothese zur Auswirkung des Geschlechts auf die psychische Gesundheit

Sowohl Jahoda (1983) als auch Shamir (1985) gehen davon aus, dass Männer durch den Wegfall ihres Rollenverständnis als primärer Familienernährer psychisch stärker unter der Arbeitslosigkeit leiden müssten als Frauen. Murphy und Athanasou (1999) konnten in ihrer Metaanalyse keine Geschlechtseffekte in den Querschnittsdaten finden, Paul und Moser (2009a) geben an, dass sie innerhalb der Arbeitslosengruppe nur geringe Unterschiede zwischen Männern und Frauen fanden. Für die Querschnittsdaten dieser Studie sollte ein möglicher Geschlechtseffekt untersucht werden, wozu die Hypothese formuliert wurde, dass Männer psychisch stärker beansprucht sein sollten als Frauen. Folgende Prognosen wurden dementsprechend operationalisiert:

Männer zeigen niedrigere LEB als Frauen
Männer zeigen niedrigeren GHQ als Frauen
Männer zeigen höheren BDI als Frauen
Männer zeigen niedrigere PSK als Frauen
Männer zeigen niedrigeren SOC als Frauen
Männer zeigen niedrigere SWE als Frauen
Männer zeigen niedrigeres PSW als Frauen

Q3: Hypothese zum geschlechtsspezifischen Beziehungsstatus auf das Befinden

Lahelma (1989) ist der Auffassung, dass eine feste Beziehung sich abschwächend auf die negativen Folgen von Arbeitslosigkeit auswirken und damit das Befinden stärken sollte. Paul und Moser (2009b) fanden für den Beziehungsstatus keinen signifikanten Effekt. McKee-Ryan et al. (2005) konnten in ihrer Metaanalyse signifikante Moderatoreffekte nur für die aus einer Beziehung resultierende soziale Unterstützung nachweisen. In einem früheren Beitrag argumentieren Paul et al. (2006), dass speziell (Ehe-)Frauen im Falle der Arbeitslosigkeit mit einer stärkeren finanziellen Unterstützung ihres Partners rechnen könnten, als dies umgekehrt analog für die Männer der Fall sei. Demnach könnten sich Befindensvorteile in Bezug auf den Beziehungsstatus geschlechtsspezifisch äußern und sollten deshalb getrennt für erwerbslose Männer und Frauen untersucht werden. Folgende Prognosen wurden aufgestellt:

Männer oder Frauen in Beziehung haben höhere LEB als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben höheren GHQ als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben niedrigeren BDI als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben höhere PSK als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben höheren SOC als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben höhere SWE als diejenigen ohne Beziehung
Männer oder Frauen in Beziehung haben höheres PSW als diejenigen ohne Beziehung

2.9 Ziele und Hypothesen zu den Längsschnittdaten

Zuerst sollten Drop-Out-Analysen zwischen den Personen, die an der Erst- und Zweituntersuchung teilnahmen und denen, die nur zur Erstuntersuchung anwesend waren, durchge-

führt werden. Des Weiteren sollten sie zwischen den Probanden, die an allen drei Untersuchungsterminen teilnahmen und den restlichen Personen berechnet werden. Mit den Tests sollte geprüft werden, inwieweit aufgrund der Werte der Erstuntersuchung mit Selektionseffekten im Verlauf der Studie zu rechnen war.

2.9.1 Risikoberechnung des Sportabbruchs

Das Hauptziel des Ausdauertrainings war die Förderung der Gesundheit im Interventionszeitraum. Daneben wurde die Durchführung der Intervention auch von der Hoffnung getragen, in den Probanden ein körperlich aktiveres Bewegungsverhalten zu initialisieren und damit eine Weichenstellung hin zu einem gesünderen Lebensstil zu ermöglichen. Daher wurde als Interventionserfolg definiert, wenn die Teilnehmer ihre sportliche Aktivität im follow-up Zeitraum fortsetzen. Um die Hürde dazu möglichst klein zu halten, wurde jedem Teilnehmer offeriert, im selben Fitnesscenter, in dem zuvor die Intervention statt fand, kostenlos ein eigenständiges Training fortzusetzen. Das Angebot war von daher als "niedrigschwellig" zu bezeichnen. Zuerst sollte ermittelt werden, wie viele Versuchsteilnehmer die Aktivität aufrechterhielten. Des Weiteren interessierte, unter welchen Umständen der Sport besser aufrechterhalten bzw. eher abgebrochen wurde.

Empirischer Hintergrund

Es ließ sich keine Publikation finden, die sich mit der Aufrechterhaltungswahrscheinlichkeit sportlicher Aktivität bei Erwerbslosen beschäftigt. Zuerst wurde Literatur gesichtet, die sich mit Unterschieden zwischen älteren Menschen, die Sport treiben oder nicht, befasst (retrospektive Betrachtungsweise). Diese stammte aus dem Kreis der Präventivmedizin, Sportmedizin oder der medizinischen Altersforschung, in welcher vornehmlich körperliche Unterschiede zwischen den Gruppen berichtet werden. Zwei Publikationen, die auch soziopsychologische Differenzen berechneten, seien kurz beschrieben: Resnick (2001) analysierte die Daten von älteren Menschen, die entweder noch sportlich aktiv waren oder nicht. Sie identifizierten, dass diejenigen mit einem aktiven Bewegungsverhalten eine stärkere Selbstwirksamkeitserwartung (SWE) aufwiesen. Rhodes et al. (1999) zeigten in ihrem Review, dass eine hohe SWE, eine "positive Einstellung zum Sport" und eine stärkere soziale Unterstützung einen erhöhenden Einfluss darauf hatten, ob Menschen im Alter noch sportlich aktiv waren.

Studien, die sich explizit mit der Fortsetzungswahrscheinlichkeit eines zu T1 begonnenen Sportes im follow-up befassen (prospektive Betrachtungsweise), sind von der Anzahl her überschaubar; allerdings weisen deren Ergebnisse in verschiedene Richtungen: McAuley (1993) betont die Rolle einer zum ersten Messzeitpunkt hohen SWE als positiven Einflussfaktor für die Aufrechterhaltung eines begonnenen Sportprogramms für ältere Menschen. In einer separaten Publikation analysierten McAuley, Lox und Duncan (1993) die Veränderungen von körperlichen Variablen und der SWE nach einem 5-monatigen Training für ältere Probanden: Hier entdeckten sie, dass die durch das Training verbesserte SWE im Gegensatz zu den Veränderungen der körperlichen Variablen den einzigen Faktor darstellte, welcher die Fortsetzungswahrscheinlichkeit des Bewegungsverhaltens im follow-up vorhersagen

ließ. Zu einem anderen Ergebnis kommen Stigjelbout, Hopman-Rock, Crone, Lechner und Van Mechelen (2006): Sie testeten die Aufrechterhaltungswahrscheinlichkeit von verschiedenen Sportprogrammen für ältere Menschen im follow-up und konnten eine anfangs höhere SWE nur als signifikanten Prädiktor für die Absicht, den Sport fortzusetzen, belegen. In der tatsächlichen Aufrechterhaltungsquote spielte die SWE keine Rolle. Klonoff, Annechild und Landrine (1994) fanden einen umgekehrten Zusammenhang zwischen ursprünglich negativem Befinden und höherer Aufrechterhaltungswahrscheinlichkeit bei Frauen, denen die Gelegenheit geboten wurde, kostenlos an einem Aerobic Programm teilzunehmen: Es stellte sich heraus, dass zu T1 die Merkmale "hoher BMI", "körperliche Beschwerden" und "Ängstlichkeit" die Teilnahmewahrscheinlichkeit und Frequenz der Inanspruchnahme des Programms erhöhten. Williams und Lord (1995) untersuchten ältere weibliche Versuchsteilnehmer im 6-monatigen follow-up, nachdem diese an einem 12-monatigen Fitnessprogramm teilgenommen hatten und belegten, dass eine anfangs stärkere Depressivität die Fortsetzungswahrscheinlichkeit des Sportes erhöhte. Für Emery, Hauck und Blumenthal (1992) wiederum stellte die durch die sportliche Intervention herbeigeführte "reduzierte Ängstlichkeit" eine signifikante Vorhersagevariable zur höheren Fortsetzungswahrscheinlichkeit eines begonnenen Bewegungsverhaltens dar.

Bisherige Ergebnisse zur prospektiven Aufrechterhaltungswahrscheinlichkeit sind demnach nicht konsistent; so bleibt unklar, ob ein zu Beginn eher positives oder eher negatives Befinden die Fortsetzung eines begonnenen Sports erhöht. Ferner sieht es so aus, dass bisher nur wenige psychologische Konzepte zur Vorhersage des Phänomens verwendet wurden und zudem vom Einsatz der SWE dominiert waren. In keiner der genannten Studien kamen Kohärenzsinn, unspezifische psychische Beschwerden, Lebenszufriedenheit oder der Fernsehkonsum als potentielle Prädiktoren zur Anwendung.

Aufgrund der uneinheitlichen Befundlage wurden für die hiesigen Sportfortsetzeranalysen keine Hypothesen formuliert. Stattdessen sollte in den Daten explorativ nach Unterschieden zwischen Fortsetzern und Abbrechern gesucht werden. Nach der Bestimmung der Sportfortsetzungsquote sollten die Variablen, welche zwischen Aufrechterhaltung und Abbruch eine Trennschärfe aufweisen, als Prädiktoren für ein binär-logistisches Prädiktionsmodell zur Bestimmung des Sportabbruchrisikos verwendet werden.

2.9.2 Längsschnitthypothesen L1, L2, L3, L4

Ziel der Längsschnitthypothesen war die Formulierung von statistisch testbaren Aussagen zum Interventionseffekt. In den Hypothesen L1-L4 wurden für die Konstrukte Prognosen über die Veränderungsrichtung der Daten im Interventionszeitraum aufgestellt.

L1: Hypothesen zu Risikofaktoren der psychischen Gesundheit

Während im systematischen Review von Lawlor und Hopker (2001) noch Zweifel an der methodischen Qualität einiger Primärstudien zum Zusammenhang von Sport und ausgeglichenem seelischen Befinden laut wurde, ist in der umfassenden Längsschnittstudie von van Gool et al. (2007) gut belegt, dass eine sportliche Betätigung einen kausal positiven Einfluss

auf die psychische Gesundheit haben kann. Daher wurde die Hypothese aufgestellt, dass die Intervention sich senkend auf die Risikofaktoren der psychischen Gesundheit auswirken sollte, welches mit den Veränderungsprognosen folgender Konstrukte operationalisiert wurde:

- | |
|---|
| L1.1: Unspezifische psychische Beschwerden sinken im Interventionszeitraum (GHQ steigt) |
| L1.2: Depressivität sinkt im Interventionszeitraum (BDI sinkt) |
| L1.3: Psych. Einschränkung der Lebensqualität sinkt im Interventionszeitraum (PSK steigt) |

L2: Hypothesen zu Ressourcen der psychischen Gesundheit

Analog zu L1 wurde erwartet, dass sich positive Interventionseffekte auf die psychische Gesundheit auch erhöhend in den Ressourcen der psychischen Gesundheit manifestieren. So berichtet z.B. McAuley (1993) für die Selbstwirksamkeit eine Steigerung durch sportliche Aktivität bei älteren Probanden. Es wurde die Hypothese aufgestellt, dass das Training die psychischen Ressourcen steigert, welches mit der Veränderung folgender Konstrukte erfasst werden sollte:

- | |
|--|
| L2.1: Der Kohärenzsinn erhöht sich im Interventionszeitraum (SOC steigt) |
| L2.2: Die Selbstwirksamkeitserwartung steigt im Interventionszeitraum (SWE steigt) |
| L2.3: Das psychosoziale Wohlbefinden erhöht sich im Interventionszeitraum (PSW steigt) |

L3: Hypothesen zur subjektiven körperlichen Gesundheit

Ein angemessenes Ausdauertraining kann sich reduzierend auf Übergewicht, Bluthochdruck und Fettstoffwechselstörungen auswirken (z.B. Lee, Sui, & Blair, 2009). Diese Parameter stellen relevante Risiken von kardio- und zerebrovaskulären Erkrankungen dar. Des Weiteren gilt sportliche Aktivität als eine der effektivsten Methoden zur Behandlung von generalisierten Schmerzen, wie der Fibromyalgie oder somatoformen Schmerzstörungen (Busch, Barber, Overend, Peloso, & Schachter, 2007). Von daher war davon auszugehen, dass ein Bewegungsprogramm sowohl physiologisch als auch subjektiv wahrnehmbar die körperliche Gesundheit stärken sollte. Dies führte zur Formulierung der Hypothese, dass das subjektive körperbezogene Gesundheitsgefühl durch das Training steigen wird, welches folgendermaßen operationalisiert wurde:

- | |
|--|
| L3.1: Die körperliche Lebensqualität erhöht sich im Interventionszeitraum (KSK steigt) |
| L3.2: Der subjektive Gesundheitszustand steigt im Interventionszeitraum (VAS steigt) |

L4: Hypothese zum Freizeitverhalten

Es konnte angenommen werden, dass das sportbasierte Gesundheitsförderungsprogramm ein aktiveres Freizeitverhalten in den Langzeiterwerbslosen initiiert. Die Annahme eines aktiveren Lebensstils sollte sich logischerweise senkend auf ein passives Freizeitverhalten auswirken. Als prototypisch für das Letztere kann der TV-Konsum angesehen werden. Die Hypothese, dass sich das passive Freizeitverhalten reduzieren wird, resultierte in der Formulierung folgender Prognose:

- | |
|--|
| L4: Der TV-Konsum sinkt im Interventionszeitraum (TV-Zeit/Tag sinkt) |
|--|

2.9.3 Moderatorhypothesen M1, M2, M3, M4

Hypothesen zu Zwischensubjekt-Innersubjekt-Interaktionen wurden aufgestellt. In diesen wurde eine Annahme darüber getroffen, in welche Richtung ein Zwischensubjektfaktor den Interventionseffekt moderieren könnte.

M1: Hypothese zum Moderatoreffekt "Trainingsintensivierung"

Die Intensivierung des Ausdauertrainings in der 2. Interventionskohorte erfolgte mit dem Ziel, ein verbessertes Training für die Prävention von vaskulären Erkrankungen bereitzustellen. In Bezug auf die hier untersuchten Längsschnittvariablen war es von Interesse zu prüfen, ob das intensivere Bewegungsprogramms der 2. Kohorte im Vergleich zum moderaten Trainingsprogramm der 1. Kohorte zu einer verbesserten psychischen und subjektiv körperlichen Gesundheit führen würde oder ob das passive Freizeitverhalten stärker durch das intensive Training reduziert werden könnte. Folgende Unterschiedserwartungen wurden daher operationalisiert:

Interventionskohorte 2 zeigt stärkere Erhöhung im GHQ als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Reduktion im BDI als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im PSK als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im SOC als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im SWE als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im PSW als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im KSK als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Erhöhung im VAS als Interventionskohorte 1
Interventionskohorte 2 zeigt stärkere Reduktion des TV-Konsums als Interventionskohorte 1

M2 Hypothesen zum Moderatoreffekt "Depressivität"

Die Wirksamkeit von sportlicher Aktivität in der Depressionsbehandlung gilt im Allgemeinen als belegt (z.B. Blumenthal et al., 1999; Kessler et al., 2001) und wird auch als vorbeugende, antidepressive Maßnahme für ältere Menschen empfohlen (Strawbridge, Deleger, Roberts, & Kaplan, 2002). Daher wurde die Hypothese aufgestellt, dass Probanden mit Depressionssymptomen von klinisch relevanter Intensität vom Trainingsprogramm stärker in der Reduktion von Risikofaktoren der psychischen Gesundheit profitieren werden als nichtdepressive Probanden, welches mit folgenden Unterschieden messbar sein sollte:

M2.1: Depressive Probanden erreichen stärkere Erhöhung im GHQ als nichtdepressive
M2.2: Depressive Probanden erreichen stärkere Reduktion im BDI als nichtdepressive
M2.3: Depressive Probanden erreichen stärkere Erhöhung in der PSK als nichtdepressive

Depp, Schkade, Thompson und Jeste (2010) argumentieren in ihrem Übersichtsartikel zum TV-Konsum von älteren Menschen, dass diese das Fernsehen vermutlich auch zur Emotionsregulation gebrauchen. Es ist vorstellbar, dass die Reduktion des TV-Konsums vor allem für Probanden mit affektiven Störungen geschmälert sein könnte, da deren emotionale Balance möglicherweise stärker von dem Medium abhängt als bei Probanden ohne Depressivität. Daher wurde die Hypothese aufgestellt, dass Personen mit der Symptomatik eine schwächere Reduktion des TV-Konsums aufweisen könnten als Nichtdepressive:

M2.4: Depressive Probanden zeigen schwächere Reduktion des TV-Konsums als nichtdepressive

M3 Hypothesen zum Moderatoreffekt "Geschlecht"

Meist werden von Männern präventive Maßnahmen zur Verhinderung von Krankheiten (z.B. Impfungen, Vorsorgeuntersuchungen, Ernährungsberatung) schlechter angenommen als von Frauen (z.B. McQueen, Vernon, Meissner, Klabunde, & Rakowski, 2006). Speziell für den Sport aber postulieren Markland und Ingledew (1997), dass dieser von Männern besser akzeptiert sei, da Männer ein körperbetonteres und kompetitiveres Motiv beim treiben von Sport verfolgen würden als Frauen. Geschlechtsdifferenzen in der Wirksamkeit von Sport auf empfundene Gesundheitsverbesserungen finden sich in einigen Publikationen: Z.B. untersuchten Foy, Rejeski, Berry, Zaccaro und Woodard (2001) den moderierenden Einfluss des Geschlechtes auf die Wirksamkeit von Sport bei COPD-Patienten und berichten, dass die männlichen Patienten stärker in der gesundheitsbezogenen Lebensqualität profitierten als die weiblichen. Jette et al. (1996) zeigen auf Basis ihrer Längsschnittuntersuchung, dass sich für ältere Probanden nach einem Sportprogramm ein gesteigertes körperliches Wohlbefinden nur für die Männer aber nicht für die Frauen nachweisen ließ. Für einen möglichen Geschlechtseffekt wurde daher die Hypothese aufgestellt, dass Männer in den körperbezogenen Variablen stärker vom Training profitieren werden als die Frauen, welches folgendermaßen messbar sein sollte:

M3.1: Männer zeigen stärkere Erhöhung in der KSK als Frauen

M3.2: Männer zeigen stärkere Erhöhung in der VAS als Frauen

M4 Hypothese zum Moderatoreffekt "Sportfortsetzung" im follow-up

Nach der Interventionsphase bekamen die Teilnehmer das Angebot, im selben Fitnesscenter, in dem zuvor die Intervention statt gefunden hatte, kostenfrei weiter trainieren zu können. Es wurde die Hypothese aufgestellt, dass Probanden, die sich weiterhin sportlich betätigen, ein für die Gesundheit günstigeres Ergebnis in der follow-up-Veränderung aufweisen werden als die Sportabbrecher. Als *"für die Gesundheit günstigeres Ergebnis"* der Sportfortsetzer wurde definiert:

→ Eine stärkere Gesundheitsverbesserung der Fortsetzer als die Abbrecher und/oder

→ eine stärkere Gesundheitsverschlechterung der Abbrecher als die Fortsetzer im follow-up.

Folgende Unterschiede wurden in diesem Sinne für das follow-up erwartet:

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im GHQ als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im BDI als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im PSK als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im SOC als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im SWE als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im PSW als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im KSK als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im VAS als Abbrecher

Fortsetzer zeigen im follow-up ein günstigeres Ergebnis im TV-Konsum als Abbrecher

2.9.4 Zusätzliche Ziele

Effektstärkenberechnungen der Hypothesentests: Um die Stärke der Interventions- und Moderator-effekte vergleichend gegenüberstellen zu können und sie zudem mit Effekten aus anderen publizierten Interventionsstudien vergleichbar zu machen, sollten neben den statistischen Tests auch Effektstärkenberechnungen durchgeführt werden.

Sensitivitätsprüfungen der Messinstrumente: Da ein Auftrag des BMBF lautete, geeignete Instrumente zur Erfassung von Veränderungen im hiesigen Design zu identifizieren, sollte untersucht werden, mit welcher Veränderungssensitivität (Power) sich Unterschiede zwischen sämtlichen Messzeitpunkten (T1, T2, T3) nachweisen lassen.

2.10 Methodische Verfahren und Anmerkungen

Für Unterschiedstests wurden je nach Voraussetzung t-Tests für unabhängige oder abhängige Stichproben durchgeführt, als nonparametrische Alternative kamen Mann-Whitney-U-Tests bzw. Wilcoxon-Tests oder Chi-Quadrat-Tests zum Einsatz. Bivariate Zusammenhänge wurden mit Pearson-Korrelationen oder nonparametrisch als Spearman-Rho bestimmt. Die Prädiktion einer zweistufigen Variable war als binär-logistische Regressionsanalyse realisiert. Risiken wurden als Odds und Odds-Ratios berechnet.

Für sämtliche Prüfungen kamen die statistischen Tests zweiseitig mit einem Signifikanzniveau von $p=0,05$ zur Anwendung. Per default werden deskriptive Werte in dieser Arbeit als Mittelwert \pm Standardabweichung (z.B. $50,00 \pm 10,00$) berichtet; beim Bericht von anderen Maßen der zentralen Tendenz, wie z.B. dem Median, wird explizit darauf aufmerksam gemacht. Statistische Analysen und Variablenberechnungen wurden mit dem Programmpaket SPSS 15.0 durchgeführt. Effektstärkenberechnungen und Poweranalysen erfolgten zusätzlich mit dem Programm G*Power 3.12.

2.10.1 Anmerkung zum unterschiedlichen N

Der GHQ, SOC, SWE, PSW und SF12 kamen von Beginn der Studie zum Einsatz. Der BDI, FPI, VAS und TV-Konsum wurden später in das Design der Studie aufgenommen, weshalb deren N verkleinert ist. 2 Versuchspersonen mit Migrationshintergrund hatten in den Fragebögen FPI, BDI, SOC und SWE zu starke sprachliche Verständnisschwierigkeiten, sodass diese Tests für diese 2 Probanden nicht ausgewertet werden konnten.

Die Längsschnitthypothesen der Zeiteffekte (L1-L4) wurden pro Instrument zusammen mit den Moderatorhypothesen (M1-M3) jeweils innerhalb einer ANOVA mit Messwiederholung getestet. Da der Zwischensubjektfaktor "Depressivität" in den Längsschnitthypothesentests der Zeiteffekte zum GHQ und zur PSK (L1) sowie zum TV-Konsum (L4) mit getestet wurde, reduzierte sich das N dieser Messgrößen um die nicht vorhandenen T1-Werte des BDI, der zur Stellung der Depressivitätsdiagnose vonnöten war.

Sonstige fehlende Daten in den Ergebnissen sind als "Missing Data" (fehlende Beantwortung eines Items oder Tests zum Messzeitpunkt) zustande gekommen. D.h. abgesehen von den oben beschriebenen Einschränkungen des N wurden sämtliche vorhandenen Daten

ausgewertet. Keine Versuchsperson wurde im Nachhinein aufgrund irgendwelcher weiterer Kriterien (z.B. Ausreißer o.ä.) ausgeschlossen.

2.10.2 Methode zum Test der Längsschnitthypothesen

Um gleichzeitig die Stabilität von möglichen Interventionseffekten im follow-up beurteilen zu können, erfolgten die Tests der Längsschnitt- und Moderatorhypothesen nur für die Probanden, die an sämtlichen Messzeitpunkten teilgenommen hatten. Zur Ermittlung der Zeiteffekte wurde für jedes Fragebogeninstrument eine multivariate Varianzanalyse mit Messwiederholung für die Untersuchungszeitpunkte T1, T2, T3 (Faktorstufen des Innersubjektfaktors) berechnet. Innerhalb desselben Modells erfolgten die Tests der Moderatorhypothesen, weswegen neben dem Innersubjektfaktor auch Zwischensubjektfaktoren mit getestet wurden.

Sämtliche Messverfahren zur Erfassung der Gesundheit erfüllen laut der Test-Autoren die Normalverteilungsannahme in der Population. Hingegen war die Variable zum TV-Konsum zwar unimodal aber nicht normalverteilt, sondern zeigte eine Rechtsschiefe. Da es im multivariaten Fall mit der Aufnahme von Zwischensubjektfaktoren kein entsprechendes nichtparametrisches Verfahren gibt, wurde für die Prüfung der Hypothesen zum TV-Konsum dennoch die ANOVA mit Messwiederholung herangezogen, da Varianzanalysen gegen Abweichungen von der Normalverteilung als relativ robust eingeschätzt werden (Zöfel, 2003). Bei schiefen Populationsverteilungen sind Abweichungen von der Normalverteilung nach Bortz (2005) auch zu vernachlässigen. Für jede ANOVA wurde mit dem Test nach Mauchly (1940) geprüft, ob Sphärizität zwischen den Innersubjektfaktorstufen angenommen werden konnte. War dies nicht der Fall, erfolgten in den Analysen Korrekturen der Freiheitsgrade nach Greenhouse und Geisser (1959).

In Bezug auf eine mutmaßliche Zeitstabilität der Interventionseffekte wurde keine absolute Zeitstabilität (z.B. signifikante Erhöhung des Kohärenzsinn mit keinerlei Rückgang im follow-up) erwartet. Die im Theorieteil (Kap. 1.6) beschriebenen Interventionsstudien zur Verbesserung der psychischen Gesundheit von Erwerbslosen zeigten deskriptiv alle einen Rückgang des Interventionserfolges, wenn ein follow-up durchgeführt wurde. Auch der meta-analytisierte durchschnittliche Interventionseffekt von Paul und Moser wurde von den Autoren als eher kurzfristig wirksam eingeschätzt (Paul & Moser, 2009c; vgl. Kap. 1.6.1). Von daher sollte ein statistischer Test angewandt werden, der Interventionseffekte, die im follow-up wieder vollständig in sich zusammenbrechen, sensitiv von denen zu unterscheiden weiß, in denen deskriptiv zwar ein Rückgang des Interventionseffektes im follow-up beobachtbar ist, dieser aber nicht bis zum Ausgangsniveau zurückkehrt. Im Gegensatz zur absoluten Zeitstabilität sollte also eine *relative Zeitstabilität* prüfbar sein. Zur Beurteilung dieser wurde der Helmert-Kontrast verwendet. Hierbei wird "...der Effekt jeder Faktorstufe, außer der letzten, mit dem mittleren Effekt der folgenden Faktorstufen verglichen" (Rudolf & Müller, 2004, S. 100). Mit dem Helmert-Kontrast ist bei 3 Messzeitpunkten prüfbar, ob ein Interventionseffekt selbst unter Einbezug seines (erwartbaren) Rückgangs im follow-up noch signifikant nachweisbar ist.

2.10.2.1 Schema zur Prüfung der Längsschnitthypothesen L1-L4

Die Prüfungen der Zeiteffekte in Kap. 4.3 (Tab. 4.9) erfolgten nach folgendem Schema:

- (A) Für den Haupteffekt des Innersubjektfaktors (T1, T2, T3) erfolgte der Test der Nullhypothesenannahme, dass sich die Daten zwischen den Zeitpunkten nicht unterscheiden. Konnte die Annahme abgelehnt werden, war ein signifikanter Haupteffekt des Innersubjektfaktors und somit eine Veränderungssensitivität (vgl. Kap. 2.10.2.4) des Instrumentes als belegt.
- (B) Zur Prüfung der Interventionshypothese wurde der hypothesenrelevante Kontrast der Messwiederholung im Untersuchungszeitpaar T1-T2 examiniert. Es erfolgte der Test der Nullhypothese, dass die Daten zwischen den Zeitpunkten unverändert waren. Konnte dies signifikant abgelehnt werden, galt der Unterschied im Zeitpaar als belegt.

Falls die Nullhypothese in (B) signifikant falsifiziert war, wurden zur Beurteilung der Veränderungsrichtung die deskriptiven Randmittel der Innersubjektfaktorstufen inspiziert. Es erfolgte die Kontrolle, ob der Unterschied zwischen T1-T2 deskriptiv die vorausgesagte Änderung der Interventionshypothese signalisierte. War dies der Fall, galt die Interventionshypothese als bestätigt. Falls weder in (A) noch in (B) die Nullhypothesen zurückgewiesen werden konnten, entfiel die Examinierung der deskriptiven Werte, da von einer statistischen Gleichheit der Daten zwischen den Zeitpunkten auszugehen war.

- (C) Falls die Interventionshypothese bestätigt war, wurde zur Beurteilung der Zeitstabilität des Interventionseffektes der Helmert-Kontrast getestet. Es wurde die Nullhypothese geprüft, dass sich T1 nicht von T2 unter Hinzunahme von T3 unterscheidet. Falls die Nullhypothese falsifiziert werden konnte, galt der mittlere Unterschied als belegt.

Falls die Hypothese zum Interventionseffekt bestätigt war und gleichzeitig eine Signifikanz des Helmert-Kontrastes belegt werden konnte (C), galt eine relative Zeitstabilität des Interventionseffektes als belegt. Falls in diesem Setting der Helmert-Kontrast eine Signifikanz vermissen ließ, wurde der Effekt der bestätigten Interventionshypothese als zeitlich instabil interpretiert.

2.10.2.2 Schema zur Prüfung der Moderatorhypothesen M1-M3

Interventionsmoderierende Einflüsse wurden als Interaktionseffekt des Zwischensubjektfaktors mit dem Innersubjektfaktor getestet. Mit der Prüfung der Moderatoreffekte innerhalb desselben varianzanalytischen Modells wie die Hypothesentests zu L1-L4 war gewährleistet, dass eine Alphafehlerkumulierung verhindert wird, weil das Signifikanzniveau von $p=0,05$ jeweils auf das gesamte Modell seine Anwendung fand. Als maßgeblich moderierend wurde ein differenzierter Interventionseffekt unter Einschluss des follow-up definiert. Von daher erfolgte die Prüfung der Moderatorhypothesen über den Helmert-Kontrast des Interaktionseffektes von Zwischensubjekt- und Innersubjektfaktor. Im Kap. 4.4 (Tab. 4.12 bis 4.14) wurden die Hypothesen folgendermaßen getestet:

- (a) Für den Test der Moderatorhypothese wurde die Annahme der Nullhypothese der Helmert-Interaktion geprüft, die besagt, dass der Zwischensubjektfaktor keinen Einfluss auf die Veränderung des Innersubjektfaktors von T1 vs. T2 & T3 hat. Konnte die Nullhypothese abgelehnt werden, galt ein differenzierter Einfluss des Zwischensubjektfaktors als belegt.
- (b) Innerhalb der gleichen Prozedur wurde die Nullhypotheseannahme geprüft, dass der Zwischensubjektfaktor keinen Unterschied zwischen den Faktorstufen T2 und T3 des Innersubjektfaktors bedingt.

Falls die Nullhypothese in (a) falsifiziert war, erfolgte eine deskriptive Erfassung der Randmittel der Stufen des Innersubjektfaktors für die Zwischensubjektfaktorgruppen. Es wurde kontrolliert, ob die durch den Zwischensubjektfaktor bedingten unterschiedlichen Veränderungsrichtungen in die hypothetisierten Richtungen wiesen. War dies der Fall, galt die Moderatorhypothese als bestätigt. War der Interaktionseffekt ausschließlich signifikant für die Interaktion T2-T3 (b), wurde dies als isolierte Wirkung des Zwischensubjektfaktors für die Veränderung im follow-up-Zeitraum interpretiert.

2.10.2.3 Prüfung der Moderatorhypothese M4

Die zu prüfenden Konstrukte der Moderatorhypotesentests zu M4 in Kap. 4.4 (Tab. 4.15) wurden in separaten⁷ ANOVAS getestet. Trotz neuer ANOVAS erfolgten die Tests für dieselben Probandenkollektive, wie in den Hypotesentests von M1-M3 (bzw. L1-L4).

"Sportfortsetzung" war der Zwischensubjektfaktor, der mit dem follow-up interagieren sollte, weswegen T2 und T3 des jeweiligen Instrumentes als Faktorstufen des Innersubjektfaktors definiert waren. Um möglichen, divergierenden Ausgangsbedingungen zwischen den Gruppen der "Sportfortsetzung" Rechnung zu tragen, wurden die follow-up-ANOVAS für die Daten von T1 kontrolliert. Der Interaktionseffekt von "Sportfortsetzung" mit dem Innersubjektfaktor wurde getestet und bei Signifikanz die Randmittel deskriptiv auf Hypothesenkonformität inspiziert. Zusätzlich erfolgte der Signifikanztest der Kontrollvariable T1, um ihren Einfluss auf die follow-up Veränderung einschätzen zu können.

2.10.2.4 Quantifizierung der Veränderungssensitivität der Instrumente

Als Maß der Veränderungssensitivität wurde definiert, mit welcher Stärke der Haupteffekt des Innersubjektfaktors, also der Effekt der Variabilität sämtlicher Messzeitpunkte, eines Fragebogeninstrumentes Signifikanz erreichen konnte. Dies ist die Teststärke welche auch als "Power" bezeichnet wird und die Sensitivität reflektiert. In den Hypotesentests L1-L4 wurde die Signifikanz des Innersubjektfaktor-Haupteffekts unter dem Punkt (A) geprüft (vgl. Kap. 2.10.2.1). Da die Signifikanztests auf unterschiedliche Stichprobengrößen basierten, wurden die Poweranalysen auf ein identisches N hochgerechnet, um eine Vergleichbarkeit

⁷ Die Eröffnung der separaten ANOVA für die follow-up Effekte der Hypothese M4 lag primär darin begründet, dass sich nicht noch ein weiterer Zwischensubjektfaktor in den obigen ANOVAS zu M1-M3 (die zusammen mit L1-L4 getestet wurden) hätten aufnehmen lassen, ohne zu verhindern, dass die Zellengrößen der Zwischensubjektgruppen dann $N < 5$ gewesen wären.

zwischen den Instrumenten herstellen zu können. Die Poweranalyse nach Cohen (1988) ist eine etablierte Möglichkeit zur Berechnung der Sensitivität eines Tests. Power ist definiert als $1-\beta$ und stellt somit die Gegenwahrscheinlichkeit zum β -Fehler dar. Beim Hypothesentesten reflektiert der β -Fehler das Risiko, die H_1 bei Gültigkeit dieser zu Gunsten der H_0 zu verwerfen und damit einen falsch-negativen Fehler (auch "Fehlertyp II" genannt) zu begehen (Bortz, 2005). $1-\beta$ stellt somit die Chance dar, die H_1 bei Gültigkeit dieser zu akzeptieren und damit eine richtig-positive Entscheidung zu treffen, welches im Fall des Innersubjektfaktor-Haupteffekts der ANOVA mit Messwiederholung die Stärke der Veränderungssensitivität erfasst.

3 Querschnittsergebnisse

Die Stichprobe setzte sich aus $N=119$, 46 männlichen (39%) und 73 weiblichen (61%) Probanden zusammen. Das durchschnittliche Alter zum Zeitpunkt der Erstuntersuchung betrug $53,7\pm 3,0$ Jahre (Spanne: ca. 48 bis 59 Jahre), welches zwischen Männern ($54,6\pm 3,2$) und Frauen ($53,2\pm 2,7$) statistisch gleich war.

Die Variable "Arbeitslosigkeitsdauer" (Spanne: ca. 1 bis 20 Jahre) erwies sich weder als normalverteilt noch als unipolar. Im Median waren die Probanden 5,1 Jahre erwerbslos. Frauen (Median: 6,1 Jahre) waren signifikant für längere Zeit arbeitslos als Männer (Median: 4,2 Jahre) ($Z=2,130$; $p=0,033$).

In einer Ehe oder festen Lebensgemeinschaft lebten 44 Probanden (38%). 48 waren geschieden oder getrennt (40%), 23 waren ledig (19%), 4 verwitwet (3%). Somit ließen sich 75 Probanden (62% der Stichprobe; 67% der Männer; 60% der Frauen) zur Kategorie "ohne Lebensgemeinschaft" zusammenfassen.

103 Probanden hatten Kinder (87% der Stichprobe, 74% der Männer, 95% der Frauen), die meist erwachsen waren. Lediglich 24 Subjekte (20%) gaben an, derzeit mit Kindern zusammen zu leben. Dass sich Kinder im Haushalt befanden, war nicht deckungsgleich mit der Aussage, ob der Proband in Partnerschaft lebte. Von den 24 Personen mit Kindern im Haushalt lebten 11 allein mit ihrem Kind (15% der Singles), während 13 mit Kindern inkl. einer Partnerschaft lebten (32% der Lebensgemeinschaften).

Schulische Abschlüsse lagen in der Mehrzahl der Fälle mit 57% beim 10-klässigen POS-Abschluss, welcher mit dem Realschulabschluss vergleichbar ist, gefolgt vom 8. Klasse POS-Abschluss mit 30%, der etwa dem Volksschul- bzw. Hauptschulabschluss gleicht. 10% hatten das Abitur und 2,5% gaben einen nicht kategorisierbaren Abschluss an. Nach der Schulzeit erreichten 63% eine Berufsausbildung, 9% verfügten über keinerlei Ausbildung, 9% hatten ein Universitätsstudium abgeschlossen, 8% erreichten einen Fachschulabschluss, 5% den Teilfacharbeiter und 4% schlossen ein Fachhochschulstudium ab. Ein Proband machte keine Angaben zur beruflichen Ausbildung.

Des Weiteren bezeichneten sich 62% der Gesamtstichprobe als "Hauptverdiener" bevor sie arbeitslos wurden; Männer (80%) signifikant häufiger als Frauen (51%) ($\chi^2_{(df=1)}=10,621$; $p=0,001$). Da der Beziehungsstand vor der Arbeitslosigkeit nicht abgefragt wurde, lassen sich leider keine Rückschlüsse ziehen, ob die Angabe "Hauptverdiener" aufgrund der Dominanz von Single-Haushalten bereits vor der Arbeitslosigkeit bestand.

Überprüfung der Zufallsverteilung zwischen den Interventionsgruppen

Weil die Versuchspersonen den Interventionsgruppen nicht randomisiert zugeteilt wurden, sondern sie per Gelegenheit Eingang in diese fanden, sollte geprüft werden, ob sich die Teilnehmer zwischen den Gruppen in soziodemographischen Merkmalen unterscheiden. Hierbei ergaben sich Verteilungen wie in folgender Tab. 3.1 dargestellt.

Tab. 3.1: Soziodemographische Merkmale zwischen den Interventionskohorten.

	1. Kohorte (N=63)	2. Kohorte (N=56)	Unterschiedstest
Alter	53,7±2,5 Jahre	53,7±3,4 Jahre	$t(df=117)=0,110; p=0,912$
Geschlecht	♂N=25; ♀N=38	♂N=21; ♀N=35	$\chi^2(df=1)=0,060; p=0,807$
Dauer der Arbeitslosigkeit	Median=4,4 Jahre	Median=5,5 Jahre	$Z=1,669; p=0,095$
Feste Lebensgemeinschaft?	nein: N=35; ja: N=28	nein: N=40; ja: N=16	$\chi^2(df=1)=3,205; p=0,073$
Kinder vorhanden?	nein: N=10; ja: N=53	nein: N=6; ja: N=50	$\chi^2(df=1)=0,678; p=0,410$
Ausbildung ⁸	Berufsausbildung: N=49 Akademiker: N=10 Ungelernte: N=4	Berufsausbildung: N=42 Akademiker: N=6 Ungelernte: N=7	$\chi^2(df=2)=1,823; p=0,402$

Signifikante Unterschiede ließen sich zwischen den Interventionskohorten nicht finden. Lediglich die Arbeitslosigkeitsdauer und die Verfügbarkeit einer festen Lebensgemeinschaft zeigten tendenzielle Verschiedenheiten zwischen den Gruppen.

Finanzielles Auskommen

Die Probanden sollten mit einem Schulnotensystem von 1 ≈ "sehr gut" bis 6 ≈ "katastrophal" ihre finanzielle Situation bewerten (vgl. Anhang 7.2). 37% bezeichneten ihre monetäre Situation als "mangelhaft", 22% als "ausreichend", 17% als "befriedigend", 15% als "katastrophal", 8% "gut" und nur 1% als "sehr gut". Als finanzielle Problemkategorien (Mehrfachantworten waren möglich) erhielt die höchste Zustimmung, dass "Ausgaben stark eingeschränkt" sind (87%) sowie "Anschaffungen nicht getätigt werden können" (78%), gefolgt von "Ersparnisse aufgebraucht" (61%), "Schulden" (24%) und "Probleme mit Vertragszahlungen" (23%). "Keine finanzielle Probleme" wurde am wenigsten (9%) gewählt.

3.1 Mediennutzung

Für den Fragebogen zur Mediennutzung, aus dem der **TV-Konsum** abgeleitet wurde, ergaben sich Häufigkeiten, wie im Anhang (Tab. 7.1) dargestellt. Niemand gab an, kein Fernsehen zu schauen. Dieses Medium wurde also zu 100% von der Stichprobe genutzt, welches es als Indikator für ein passives Freizeitverhalten brauchbar machte. 30% der Probanden "surften" im Internet und immerhin noch 14% verbrachten Teile ihrer Freizeit mit Computerspielen. Nach der Überführung der Zeitkategorien in numerische Werte (z.B. "1-2h." = 1,5h. etc.) betrug die durchschnittliche Mediennutzungsdauer der Stichprobe 4,7±2,1 Stunden

⁸ Teilfacharbeiter, Facharbeiter und Fachschule wurden zur Kategorie "Berufsausbildung", Fachhochschule und Universität zur Kategorie "Akademiker" zusammengefasst.

täglich. Diese setzte sich aus den gemittelten Zeiten für Fernsehen ($3,8 \pm 1,7$ h.), Internet ($0,7 \pm 1,2$ h.) und Computerspiele ($0,3 \pm 0,7$ h.) zusammen (vgl. Abb. 7.1 im Anhang).

Die *Arbeitsgemeinschaft Fernsehforschung* (AGF, 2010) berichtet für die relevanten Jahre der Erstuntersuchung (2007, 2008, 2009) einen täglichen Fernsehkonsum von durchschnittlich 3,48h. für die deutsche Bevölkerung. Standardabweichungen werden von der AGF leider nicht bereitgestellt. Ein TV-Konsum von 3,5h./Tag ist eine Schätzung, die des Öfteren in Deutschland erhoben wird (vgl. auch Winterhoff-Spurk, 2005). Gemessen an 3,5h. lag der selbstberichtete TV-Konsum der Stichprobe (3,8h.) also nur geringfügig höher.

3.2 Erfassung der Gesundheit - Ableitung des Interventionsbedarfs

Mit den psychometrischen Messverfahren wurden die Indikatoren zum Gesundheitszustand anhand der empfohlenen Vorgehensweise des jeweiligen Manuals bzw. der Publikation zur Vergleichsstichprobe bestimmt. Dies erforderte meist eine Summierung der Item-Rohwerte mit anschließender Transformation in eine Standardnorm, welche einen Vergleich mit der jeweiligen Normstichprobe erlaubte. In sämtlichen Publikationen wurde von den Autoren angegeben, dass das zu messende Merkmal in der Population normalverteilt sei, weshalb auf die Prüfung der empirischen Normalverteilung mit dem Kolmogorow-Smirnow-Test nach den Empfehlungen von Altman und Bland (1995) verzichtet wurde. Als Schwellenwert zur Bestimmung von "gesundheitlich problematischen Fällen" wurde entweder der in den Publikationen empfohlene Grenzwert oder eine Abweichung von mehr als 1 Standardabweichung vom normierten Mittelwert verwendet. Falls der empirisch ermittelte prozentuale Anteil gesundheitlich problematischer Fälle höher als der aufgrund der Vergleichsstichproben erwartete war, wurde dies als Interventionsbedarf interpretiert. Grafiken bzw. Tabellen zu den Rohwerten der Messinstrumente auf Itemebene finden sich im Anhang 7.10.

Die Rohwerte der Persönlichkeitsfaktoren des **FPI** wurden in alters- und geschlechtsspezifische Stanine nach Fahrenberg et al. (2010) transformiert. Stanine hat eine Spanne von 1 bis 9, wobei die Normstichprobe mit $MW=5$ ($SD=1,96$) standardisiert ist. Für Stanine gilt demnach 3,04 bis 6,96 als Normbereich. Die empirisch ermittelten Stanine der Persönlichkeitsfaktoren lagen im unauffälligen Bereich. (vgl. Anhang Abb. 7.2). Das Merkmal *Lebenszufriedenheit* (**LEB**) lag am weitesten unterhalb des Normstichprobenmittelwertes. Da eine "eingeschränkte Lebenszufriedenheit" als Risikofaktor der psychischen Gesundheit definiert war, erfolgte die Bestimmung des Anteils von Probanden mit dieser Auffälligkeit. Stanine ≤ 3 wurde als auffälliger Bereich gezählt.

Für den **GHQ** wurden negativ formulierte Items umgepolt⁹, damit numerisch positive Werte konsistent einen positiven psychischen Zustand darstellen. Um die Normen zum GHQ anwendbar zu machen, musste die Likert-Skala (0-1-2-3) in das sog. "GHQ-Scoring" (0-0-1-1) nach Piccinelli, Bisoffi, Bon, Cunico und Tansella (1993) umkodiert werden¹⁰. Pro Vpn wurde ein Durchschnittswert berechnet, der zwischen 0 und 1 variieren konnte. Für den deutschen Sprachraum repräsentative Normen stammen von Goldberg et al. (1997) und

⁹ Z.B. Item 2: "Ich konnte nur wenig schlafen, weil ich mir Sorgen machte": Die Likert-Skala 0-1-2-3 wurde für dieses negativ formulierte Item nach 3-2-1-0 umgepolt.

¹⁰ Die Auswertung der Längsschnittdaten (Kap. 4) erfolgte jedoch weiterhin über die Likert Skala des GHQ.

greifen auf eine internationale Stichprobe aus 15 Zentren zurück, anhand derer spezifische "Thresholds" für Fallidentifikationen berichtet werden. Maßgeblich für Deutschland sind die Analysen der Studienzentren Mainz und Berlin, für die ein Schwellenwert von $< \frac{2}{3}$ berichtet wird (Goldberg, et al., 1997, S. 194). Dieser korrespondiert laut der Autoren mit einer Sensitivität von 78,1% und Spezifität von 73,1% für die Identifikation von "Fälle mit unspezifischen psychischen Beschwerden".

Die Itemsommen des **BDI** wurden in geschlechtsspezifische T-Standardnormen (MW=50; SD=10) nach der Allgemeinbevölkerungsstichprobe von Schmitt et al. (2006) transformiert. Die Autoren empfehlen $T > 60$ als Schwellenwert für die Identifikation von klinisch relevanten Depressivitätswerten. Dieser ist laut Schmitt et al. mit einer Sensitivität von 92% und einer Spezifität von 91% für die Diagnose einer depressiven Störung validiert.

Für den **SOC** beschreiben Schumacher und Brähler (2000), dass sich bisherige Publikationen stets mit der Schwierigkeit konfrontiert sahen, die von Antonovsky (1987) unterstellten Komponenten des Kohärenzgefühls (Handhabbarkeit, Verstehbarkeit, Sinnhaftigkeit) innerhalb einer Faktorenanalyse nachzuweisen. Aufgrund der mangelnden faktoriellen Eindeutigkeit der Komponenten empfehlen Schumacher und Brähler, dass ihre Skala als Generalfaktor analysiert werden sollte. Für den hier verwendeten SOC existiert eine Normierung von Schumacher, Wilz, Gunzelmann und Brähler (2000). Aus den Itemsommen wurden anhand der Normen geschlechtsspezifische Prozentränge gebildet. Um die Vergleichbarkeit zu anderen in dieser Arbeit vorgestellten Skalen zu erleichtern, erfolgte eine zusätzliche Transformation der Prozentränge in T-Werte. Das Kriterium $T < 40$ wurde benutzt, um "Fälle mit unterdurchschnittlichem Kohärenzgefühl" zu identifizieren.

In der **SWE** wurden die summierten Item-Rohwerte in alters- und geschlechtsspezifische T-Werte anhand der bevölkerungsrepräsentativen Stichprobe von Hinz, Schumacher, Albani, Schmid und Brähler (2006) transformiert. Der Schwellenwert von $T < 40$ kam für die Identifikation von "Fälle mit unterdurchschnittlicher SWE" zur Anwendung.

Der **SF12** wurde nach den Vorgaben von Bullinger und Kirchberger (1998) spezifisch als *körperliche Summenskala (KSK)* und *psychische Summenskala (PSK)* ausgewertet. Die standardisierten Summenskalen (MW=50; SD=10) entsprechen der T-Skala. Höhere Skalenwerte stehen für eine höhere Lebensqualität. Sowohl für die KSK, als auch für die PSK wurde ein Schwellenwert von $T < 40$ festgelegt, um "Fälle mit erhöhter körperlicher (KSK) bzw. psychischer (PSK) Einschränkung der Lebensqualität" zu identifizieren.

Je höher die **VAS** (Spanne: 0-100), desto besser die selbst eingeschätzte Gesundheit. Für Vergleichswerte wurde auf die Studie von Hinz, Klaiberg, Brähler und König (2006) zurückgegriffen. Die Publikation berichtet u.a. Durchschnittswerte von 41-60-jährigen Personen und referiert für Männer einen VAS von $78,8 \pm 15,2$, für Frauen einen von $76,2 \pm 16,9$. Probanden, die mehr als eine Standardabweichung unter denen von Hinz, Klaiberg et al. (2006) berichteten entfernt lagen, wurden als "Fälle mit unterdurchschnittlichem körperlichen Gesundheitsgefühl" identifiziert.

In folgender Abb. 3.1 sind die gesundheitsproblematischen Fälle, die mithilfe der Normen der oben genannten Vergleichsstudien identifiziert wurden, als prozentuale Häufigkeiten dargestellt (vgl. auch darauffolgende Tab. 3.2).

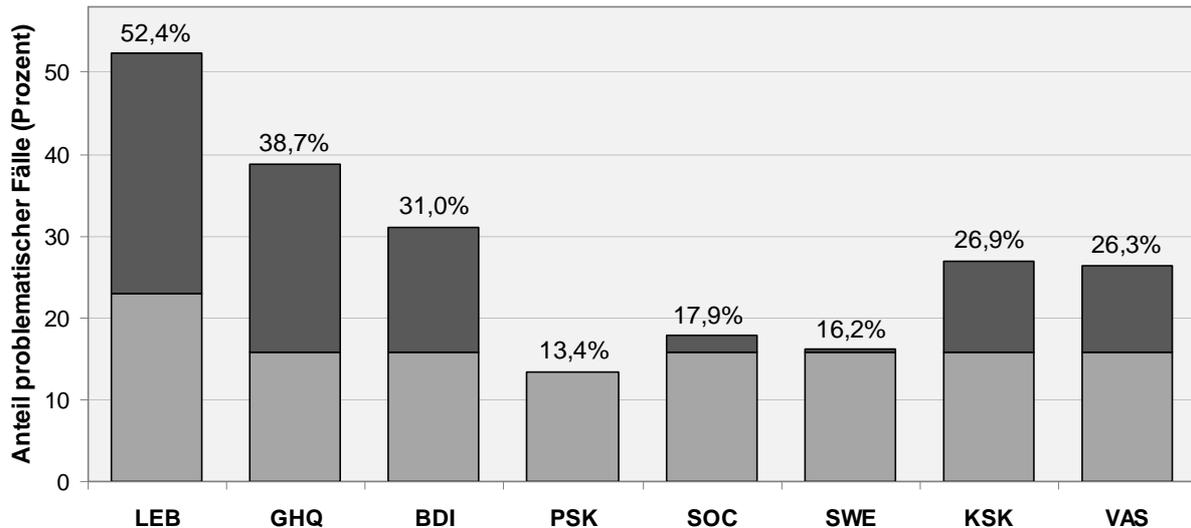


Abb. 3.1: Gesundheitsproblematische Anteile der Stichprobe als prozentuale Häufigkeiten.

In einer Normalverteilung befinden sich 68,27% der Stichprobe innerhalb 1 Standardabweichung um den Mittelwert und demnach 31,73% außerhalb. Die Hälfte davon entspricht dem Anteil, der einseitig im entweder über- oder unterdurchschnittlichen Bereich maximal erwartet werden darf; dies sind 15,87%. Dieser ist in den Säulen für die Messverfahren GHQ bis VAS hellgrau markiert. Anteile ab 15,9% weichen von den Normstichproben ab, wurden in den Säulen dunkelgrau gefärbt und stellen damit den Interventionsbedarf dar. Der Anteil des Merkmals LEB wurde anhand von Stanine-Normen berechnet, in welcher 23% dem Anteil entspricht, der laut Standardnorm für das Merkmal im einseitigen Bereich (hier: unterdurchschnittliche Lebenszufriedenheit) maximal erwartet werden darf.

Tab. 3.2: Ableitung des Interventionsbedarfs.

	Vergleichsstudie	N	Wert	Stichprobe MW±SD	Fall Kriterium	Beobachtet	Erwartet	Abweichung	Ø Abweichung
LEB	Fahrenberg et al. (2010)	84	Stanine	3,734±1,773	≤ 3	52,4%	23,0%	+128,3%	+87,8%
GHQ	Goldberg et al. (1997)	119	GHQ Score	0,735±0,217	< 0,67	38,7%	15,9%	+143,4%	
BDI	Schmitt et al. (2006)	87	T-Wert	54,99±9,470	> 60	31,0%	15,9%	+95,0%	
PSK	Bullinger & Kirchnerberger (1998)	119	T-Wert	50,53±9,324	< 40	13,4%	15,9%	-15,7%	
SOC	Schumacher et al. (2000)	117	T-Wert	48,51±9,351	< 40	17,9%	15,9%	+13,0%	+7,6%
SWE	Hinz, Schumacher et al. (2006)	117	T-Wert	50,49±10,14	< 40	16,2%	15,9%	+2,0%	
KSK	Bullinger & Kirchnerberger (1998)	119	T-Wert	45,68±9,331	< 40	26,9%	15,9%	+69,2%	+67,3%
VAS	Hinz, Klaiberg et al. (2006)	76	VAS Score	74,18±15,85 71,69±18,20	♂ < 63,6 ♀ < 59,3	26,3%	15,9%	+65,4%	

In Tab. 3.2 zeigt "Abweichung" den Unterschied vom beobachteten Anteil gesundheitsproblematischer Fälle in Relation zum erwarteten der Vergleichsstudie. Die "durchschnittliche Abweichung" bezieht sich auf die konzeptionelle Einordnung der Messgrößen in *psychische*

Risikofaktoren, psychische Ressourcen und subjektive körperliche Gesundheit (vgl. Abb. 2.2, S. 29). Die Risikofaktoren GHQ, LEB und BDI indizierten den höchsten Interventionsbedarf: Unspezifische Beschwerden erforderten damit am stärksten eine Linderung, die Lebenszufriedenheit war gravierend eingeschränkt und Symptome von Depressivität kamen ca. doppelt so häufig vor wie in der Allgemeinbevölkerung. In der Höhe gefolgt werden diese von den Abweichungen der KSK und der VAS, die zu zwei Drittel häufiger eine eingeschränkte körperliche Gesundheit sowie eine geminderte körperliche Lebensqualität signalisierten und damit einen Bedarf nach medizinischer Intervention deutlich machten. Bemerkenswert ist, dass die Anteile zu auffällig eingeschränkten psychischen Ressourcen, indiziert durch SOC und SWE, kaum Abweichungen zu den erwarteten Anteilen zeigten. In Relation zu den anderen Konzepten können sowohl die Selbstwirksamkeitserwartung als auch der Kohärenzsinn als einigermaßen unauffällig bezeichnet werden, so dass hier von einem relativ geringen Interventionsbedarf auszugehen war. Der Anteil der PSK, der eine eingeschränkte Lebensqualität aufgrund psychischer Probleme reflektiert und inhaltlich wie LEB, GHQ und BDI als psychischer Risikofaktor definiert war, wich als einziger nicht vom erwarteten problematischen Anteil seiner Normstichprobe ab sondern lag etwas darunter.

Es ist darauf hinzuweisen, dass der "wahre" Interventionsbedarf eventuell unterschätzt sein könnte. In den bevölkerungsrepräsentativen Stichproben sind neben Berufstätigen natürlich auch Erwerbslose und zudem weitere Angehörige der sog. "out of labour force" (Rengers, 2005) erfasst (relevant für die Stichprobe: z.B. ältere Hausfrauen/-männer, Frührentner, über 50-jährige krankheitsbedingt berufsunfähige Menschen, finanziell wohlhabende ältere Menschen, die nicht arbeiten). Für eine genauere Bestimmung des Interventionsbedarfs hätte die Stichprobe ausschließlich mit erwerbstätigen 50-60-Jährigen verglichen werden müssen. Da aber solch differenzierte Vergleichsdaten für die Messverfahren nicht sämtlich verfügbar waren, wurde für alle Instrumente entschieden, die Daten von parallelisierten Stichproben aus der Allgemeinbevölkerung als Norm zu verwenden.

3.2.1 Q1: Test des Zusammenhangs von Antworttendenz und Gesundheit

Zur Aufdeckung von Zusammenhängen mit Antworttendenzen wurden zwischen dem Merkmal "Offenheit" des FPI und den Gesundheitsindikatoren Korrelationen berechnet.

Tab. 3.3: Korrelative Zusammenhänge der Gesundheitsindikatoren mit dem FPI Merkmal "Offenheit".

	LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS
N	84	84	82	84	82	82	84	84	76
r	-0,035	-0,257	0,265	-0,255	-0,282	-0,265	-0,084	-0,062	-0,220
p	0,754	0,018	0,027	0,019	0,010	0,016	0,449	0,578	0,085

Es zeigte sich, dass die Antworttendenz in 5 Messverfahren signifikante Zusammenhänge im folgenden Sinn offenbarte: Je offener die Antworttendenz, desto eher stimmte der Proband Aussagen im Zusammenhang mit Depressivität (BDI) zu, desto eher gab er zu, durch psychische Probleme eingeschränkt zu sein (PSK), desto niedriger war sein Kohärenzgefühl (SOC), desto niedriger schätzte er seine allgemeine Selbstwirksamkeitserwartung (SWE) ein und desto geringer war das GHQ-Ergebnis, welches auf mehr allgemeine psychische Beschwerden hinwies. Eine Tendenz für diesen Zusammenhang ergab sich auch für ein niedri-

geres körperliches Gesundheitsempfinden (VAS). Umgekehrt zusammengefasst lautet die inhaltliche Relation, dass eine verschlossene Antworttendenz (sozial erwünschter Bias) zur Verkleinerung von potentiellen Problemen im Zusammenhang mit der Gesundheit führte. Weil die Variable "Offenheit" (vgl. Anhang Abb. 7.2) durchschnittlich einen Wert von $4,3 \pm 1,80$ aufwies (Normstichprobe $5 \pm 1,96$) und demnach die Probanden eher sozial erwünscht statt offen antworteten, ist zu interpretieren, dass die Antworttendenz bei der Ermittlung des Interventionsbedarfs eher zu einer Unter- statt Überschätzung geführt haben wird.

Dass der gesundheitliche Interventionsbedarf anhand "unwahrheitsgemäßer" Antworttendenzen der Probanden aggraviert sein könnte, gilt durchschnittlich als falsifiziert. In 5 von 9 Indikatoren zur Gesundheit zeigte sich der umgekehrte Zusammenhang: Je weniger "wahrheitsgemäß" und damit sozial erwünschter die Antworttendenz, desto eher neigten die Probanden dazu, vor allem psychische Probleme abgemildert darzustellen.

3.2.2 Q2: Test des Geschlechtseffekts der psychischen Gesundheit

Geschlechtsunterschiede wurden für die Messinstrumente zur Erfassung der psychischen Gesundheit mit t-Tests für unabhängige Stichproben berechnet.

Tab. 3.4: Gruppenunterschiede des Zwischensubjektfaktors "Geschlecht".

		LEB	GHQ	BDI	PSK	SOC	SWE	PSW
Geschlecht	♂	41	46	38	46	46	46	46
	♀	43	73	49	73	71	71	73
df		82	117	85	117	115	115	117
t		1,994	1,242	0,300	2,306	0,704	0,174	2,266
p		0,049	0,221	0,765	0,023	0,491	0,863	0,025

Männer hatten eine signifikant niedrigere Lebenszufriedenheit (LEB) als Frauen ($3,34 \pm 1,53$ vs. $4,09 \pm 1,91$), während Frauen eine signifikant psychisch eingeschränktere Lebensqualität (PSK) als Männer aufwiesen ($48,99 \pm 10,17$ vs. $52,97 \pm 7,25$). Diese Unterschiede beziehen sich auf bereits geschlechtsadjustierte Standardnormen, weshalb sie als bedeutsam aufgefasst werden sollten. Für den Unterschied im PSW, in dem Frauen ein niedrigeres psychosoziales Wohlbefinden aufwiesen als Männer ($2,64 \pm 0,57$ vs. $2,88 \pm 0,58$) könnte es sich evt. auch um einen normalen Geschlechtsunterschied handeln, da Rimann und Udris (1993) für dieses Instrument keine Normdaten bereitstellen.

Dass männliche Langzeiterwerbslose eine höhere Beanspruchung ihrer psychischen Gesundheit als weibliche aufweisen, konnte nur für eine niedrigere Lebenszufriedenheit bestätigt werden. Hingegen zeigten Frauen eine psychisch eingeschränktere Lebensqualität und ein niedrigeres psychosoziales Wohlbefinden als die Männer.

3.2.3 Q3: Test des geschlechtsspezifischen Effekts des Beziehungsstatus

Psychische Unterschiede in Abhängigkeit vom Beziehungsstatus wurden mit t-Tests für unabhängige Stichproben geprüft. Für die Gesamtstichprobe wiesen diejenigen mit einer

festen Partnerschaft im Vergleich mit den Alleinstehenden eine höhere Lebenszufriedenheit (LEB: $4,32 \pm 1,99$ vs. $3,38 \pm 1,54$) ($t_{(df=82)}=2,437$; $p=0,017$) und einen höheren Kohärenzsinn (SOC: $50,80 \pm 9,50$ vs. $47,13 \pm 9,04$) ($t_{(df=117)}=2,089$; $p=0,039$) auf. Geschlechtsspezifisch ergaben die Tests zwischen den Gruppen des Beziehungsstatus folgende Ergebnisse:

Tab. 3.5: Geschlechtsspezifische Gruppenunterschiede des Zwischensubjektfaktors "Beziehungsstatus".

			LEB	GHQ	BDI	PSK	SOC	SWE	PSW
Männer	Beziehung	Nein	27	31	25	31	31	31	31
		Ja	14	15	13	15	15	15	15
		df	39	44	36	44	44	44	44
		<i>t</i>	-0,908	-0,090	-0,099	1,108	-0,220	-0,381	-0,068
		<i>p</i>	0,369	0,928	0,922	0,274	0,827	0,705	0,946
		<i>d</i>	0,282	0,030	0,033	0,349	0,067	0,120	0,030
Frauen	Beziehung	Nein	26	44	29	44	42	42	44
		Ja	17	29	20	29	29	29	29
		df	41	71	47	71	69	69	71
		<i>t</i>	-2,297	-2,244	1,001	-2,215	-2,437	-0,779	-2,107
		<i>p</i>	0,027	0,028	0,322	0,030	0,017	0,439	0,039
		<i>d</i>	0,698	0,537	0,291	0,530	0,588	0,188	0,504

Innerhalb der Gruppe der männlichen Untersuchungsteilnehmer zeigte sich in keinem Test ein signifikanter Unterschied zwischen den Männern, die in einer festen Lebensgemeinschaft lebten vs. denen, die über keine partnerschaftliche Beziehung verfügten.

Innerhalb der Gruppe der weiblichen Versuchspersonen erzielten die mit einer festen Beziehung signifikante Befindensvorteile im Vergleich zu denen ohne Partnerschaft. Für die weiblichen Probanden löste das Vorhandensein einer festen Beziehung jeweils signifikant mehr Lebenszufriedenheit (LEB: $4,88 \pm 2,03$ vs. $3,58 \pm 1,68$), weniger unspezifische psychische Beschwerden (GHQ: $2,16 \pm 0,45$ vs. $1,90 \pm 0,50$), weniger psychische Einschränkungen der Lebensqualität (PSK: $52,15 \pm 8,16$ vs. $46,91 \pm 10,89$), ein höheres Kohärenzgefühl (SOC: $52,15 \pm 8,16$ vs. $46,82 \pm 9,66$) und mehr psychosoziales Wohlbefinden (PSW: $2,81 \pm 0,49$ vs. $2,53 \pm 0,59$) aus.

Weil das N der weiblichen Probanden in der Stichprobe überwog und von daher deren Gruppenunterschiede eine höhere Wahrscheinlichkeit hatten, das Signifikanzkriterium zu unterschreiten als die der männlichen Versuchsteilnehmer, wurden zusätzlich Effektstärkenberechnungen nach den Vorgaben von Faul, Erdfelder, Lang und Buchner (2007) durchgeführt. In obiger Tab. 3.5 indiziert "*d*" unabhängig von der Stichprobengröße den Schätzer der Effektstärke des Zwischensubjektfaktors "Beziehungsstatus". Die Inspektion von *d* offenbart, dass sich innerhalb der weiblichen Probanden für die signifikanten Tests durchweg "mittlere" Effektstärken mit $d \geq 0,5$ finden ließen, während die Unterschiedseffekte der Männer sämtlich unter 0,5 lagen und damit allerhöchstens einen "kleinen" Effekt nach den Konventionen von Cohen (1988, 1992) erreichten¹¹. Im Unterschiedstest zur PSK, welcher für die männlichen Probanden den stärksten Effekt enthielt, deutete sich inhaltlich sogar der umgekehrte Unterschied in dem Sinne an, dass Single-Männer einen Befindensvorteil gegenüber denen mit Beziehung hätten ($53,79 \pm 6,76$ vs. $51,27 \pm 8,14$).

¹¹ Nach Cohen (1988, 1992) indiziert $d \geq 0,20$ einen "kleinen" Effekt, $d \geq 0,50$ einen "mittleren" Effekt und $d \geq 0,80$ einen "großen" Effekt.

Da die Effekte des Beziehungsstatus exklusiv für die weiblichen Versuchsteilnehmer so deutlich in die hypothetisierte Richtung verliefen, wurde für diese explorativ getestet, ob sich das Vorhandensein einer Beziehung evt. auch in einer besseren finanziellen Situation widerspiegeln könnte. Paul et al. (2006) formulierten ja die Überlegung, dass v.a. weibliche Erwerbslose in fester Beziehung eine monetäre Unterstützung von ihren Partnern erwarten könnten. Hierzu wurde für die Versuchsteilnehmerinnen ein Unterschiedstest der Variable "finanzielle Selbsteinschätzung" (vgl. Kap. 5.1 bzw. Anhang 7.2) in Abhängigkeit vom Beziehungsstatus geprüft. Lediglich deskriptiv deutete sich an, dass Frauen mit Beziehung ihre finanzielle Situation etwas besser ($4,17 \pm 1,17 \approx$ "ausreichend") als die Alleinstehenden ($4,59 \pm 1,00 \approx$ "mangelhaft") einschätzten; der t-Test hierzu offenbarte jedoch keinen signifikanten Unterschied ($t_{(df=71)}=1,640$; $p=0,105$).

Innerhalb der Gruppe der weiblichen Probanden löste das Vorhandensein einer festen Beziehung Befindensvorteile in 5 (von 7) Indikatoren der psychischen Gesundheit aus. Innerhalb des Kollektivs der männlichen Versuchspersonen konnten keinerlei Befindensunterschiede in Abhängigkeit vom Beziehungsstatus objektiviert werden.

4 Längsschnittergebnisse

Im Folgenden finden sich die Ergebnisse, die aufgrund der Auswertung der Längsschnittdaten zustande gekommen sind. Zuerst werden Drop-Out-Analysen vorgestellt, dann die Sportfortsetzeranalyse und die Tests zu den Längsschnitt- und Moderatorhypothesen. Mit den Sensitivitätsanalysen wird Kap. 4 abgeschlossen.

4.1 Drop-Out-Analysen

Dass nicht sämtliche Probanden an allen Wiederholungsmessungen teilnehmen, ist in longitudinalen Forschungsdesigns immanent. Folgende Quoten wurden hierzu ermittelt:

Teilnehmer- und Drop-Out-Quote T2

- An T2 nahmen 94 von ursprünglich 119 Vpn teil (79% der Stichprobe, 85% der Männer, 75% der Frauen).
- Von den 25 Nichtteilnehmern (21% der Stichprobe, 15% der Männer, 25% der Frauen) wurden im Nachhinein folgende Abbruchgründe ermittelt:
 - 6 Probanden hatten eine Erwerbsarbeit aufgenommen (5% der Stichprobe),
 - 8 gaben an, erkrankt zu sein (6,7% der Stichprobe),
 - 11 hatten das Interesse an der Teilnahme vermutlich verloren (9,3% der Stichprobe).

Teilnehmer- und Drop-Out-Quote T3

- An T3 nahmen 60 Probanden teil (50% der Stichprobe, 67% der Männer, 40% der Frauen). Weitere 14 willigten aber zumindest ein, trotz physischer Nichtteilnahme an der Untersuchung im IPM in Rostock, postalisch zugeschickte Fragebögen zu beantworten; dies entspricht 74 Probanden (62% der Stichprobe, 70% der Männer, 58% der Frauen).

- Für die 45 Nichtteilnehmer (38% der Stichprobe, 30% der Männer, 43% der Frauen) wurden als Abbruchgründe für den dritten Untersuchungszeitpunkt folgende Gründe im Nachhinein ermittelt:
 - 1 Proband war verstorben,
 - 8 hatten eine Arbeit aufgenommen (6,7% der Stichprobe),
 - 13 gaben an, erkrankt zu sein (10,9% der Stichprobe),
 - 23 hatten das Interesse an der Teilnahme vermutlich verloren (19,3% der Stichprobe).

Die Ursachen zu den Drop-out-Quoten T2 und T3 sind nicht disjunkt (z.B. für Teilnehmer, die zwischen T1 und T2 das Interesse an der Studie verloren hatten, traf dieser Grund zwischen T2 und T3 immer noch zu). Insofern ergeben sich die Abbruchgründe von T3 auch aus denen von T2. Folgendes ist zu den Abbruch-Ursachen anzumerken:

Arbeitsaufnahme: Insgesamt nahmen 17 Teilnehmer während des Studienzeitraums eine Erwerbsarbeit auf (14% der Stichprobe). Von diesen 17 nahmen jedoch 7 an sämtlichen Untersuchungen teil und 5 weitere an der Untersuchung T2. Die Aufnahme von Erwerbsarbeit machte einen Studienabbruch demnach zwar wahrscheinlicher, stellte jedoch keine hinreichende Bedingung dar.

Erkrankung: Der Abbruchgrund "Erkrankung" fußt auf die mündliche Selbstauskunft des Probanden. Medizinische Belege dafür waren nicht einforderbar, da die Studienteilnahme auf Freiwilligkeit (Ethik) beruhte. Ob eine Versuchsperson erkrankt war, wurde für die Statistik nur dann erhoben, wenn der Proband den Wiederholungsmessungen fernblieb, d.h. eine mögliche Erkrankung wurde in der Analyse nicht ausgewertet, wenn der Proband trotzdem weiter an der Studie bzw. dem Sport teilnahm. Die Erkenntnis der Salutogeneseforschung, dass die psychische Resilienz einen erheblichen Einfluss darauf hat, ob sich Menschen bei gleicher Symptomatik noch gesund oder bereits krank fühlen (Antonovsky & Franke, 1997), wird als weiteres Argument dafür herangezogen, dass die mündliche Selbstauskunft einer Erkrankung nicht als hinreichender Abbruchgrund gewertet wurde.

Verlorenes Interesse: Dieser Abbruchgrund setzt sich aus den Probanden zusammen, die ein mangelndes Interesse frank zugaben und denen, die sich trotz mehrfacher Kontaktaufnahme nicht auf Termine für die Wiederholungsmessung festlegen bzw. Untersuchungstermine unangemessen oft verschieben wollten. Auch Versuchsteilnehmer, welche mehrfach die vereinbarten Untersuchungstermine sehr kurzfristig absagten bzw. trotz Termins einfach nicht erschienen, wurden in diese Gruppe zusammengefasst. Hier wird ein "verlorenes Interesse" der Einfachheit halber vorausgesetzt, obwohl evt. relevante private Gründe vorlagen, die sich jedoch nicht reliabel erheben ließen. Da eine Arbeitsaufnahme sich heterogen auf die weitere Studienteilnahme auswirkte und dies vermutlich auch für den Abbruchgrund "Erkrankung" zutraf, wurden diese Abbruchursachen nicht differenziert analysiert.

4.1.1 Drop-Out-Analyse T2

Zur Generalisierbarkeit der Studienergebnisse ist eine Aufklärung wichtig, ob in Bezug auf die Variablen der Erstuntersuchung selektiv bestimmte Probanden den Wiederholungsmessungen fernblieben. So ließe sich z.B. spekulieren, dass Depressivität eine motivationale

Antriebsschwäche verursacht, welche die weitere Teilnahmewahrscheinlichkeit senkt. Es wurde getestet, ob T2-Teilnehmer sich von den Nichtteilnehmern aufgrund der Ergebnisse der Eingangsdiagnostik unterschieden:

Tab. 4.1: Tests der Gruppenunterschiede des Zwischensubjektfaktors "T2-Teilnahme".

		LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum ¹²
T2	Nein	13	25	16	25	24	24	25	25	16	17
	Ja	71	94	71	94	93	93	94	94	60	67
df		82	117	85	117	115	115	117	117	74	N=84
t		-1,629	-0,433	0,107	-1,053	-0,137	-0,151	-1,208	-1,918	0,361	Z=-0,949
p		0,107	0,666	0,915	0,295	0,891	0,881	0,230	0,058	0,719	0,343

In der Tabelle finden sich die Tests der Unterschiede zwischen den Gruppen der T2-Teilnahme. Die Annahme der Nullhypothese konnte für alle Prüfungen beibehalten werden.

A priori unterschiedliche Ausprägungen in den Variablen zur Erfassung der Gesundheit und zum TV-Konsum konnten zwischen den Teilnehmern der Erst- und Zweituntersuchung vs. denen, die nur an der Erstuntersuchung teilnahmen, nicht festgestellt werden.

4.1.2 Drop-Out-Analyse T3

Im Folgenden finden sich die Tests der abhängigen Variablen zwischen den 74 Probanden, die an allen Untersuchungen teilnahmen (inklusive derer, die zu T3 die Fragebögen postalisch zur Verfügung stellten) im Vergleich mit den 44 restlichen¹³ Probanden (Teilnahme ausschließlich an T1 oder an T1&T2).

Tab. 4.2: Tests der Gruppenunterschiede des Zwischensubjektfaktors "T3-Teilnahme".

		LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum ¹⁴
T3	Nein	17	44	27	44	43	43	44	44	25	27
	Ja	67	74	60	74	73	73	74	74	51	57
df		82	116	85	116	114	114	116	116	74	N=84
t		-1,606	-0,817	-0,621	-1,367	-0,122	-0,975	-2,528	-0,703	1,477	Z=-2,649
p		0,112	0,415	0,536	0,174	0,903	0,332	0,013	0,483	0,144	0,008

Es wurde ein signifikanter Unterschied für ein schwächeres psychosoziales Wohlbefinden (PSW) der Nichtteilnehmer ($2,55 \pm 0,54$) gegenüber den T3-Teilnehmern ($2,82 \pm 0,58$) sichtbar. Zudem war der Fernsehkonsum mit +53 Minuten hochsignifikant für die Nichtteilnehmer ($263 \text{min} \pm 90 \text{min}$) gegenüber den Studienkomplettierern ($210 \text{min} \pm 104 \text{min}$) erhöht.

A priori unterschiedliche Werte fanden sich zwischen den Studienkomplettierern im Vergleich zu den Studienabbrechern für ein signifikant höheres psychosoziales Wohlbefinden und einem hochsignifikant niedrigeren Fernsehkonsum. Für die restlichen Variablen konnte die Annahme der Nullhypothese beibehalten werden.

¹² Mann-Whitney-U-Test

¹³ Ein zwischen T2 und T3 verstorbener Versuchsteilnehmer wurde nicht in die Analysen einbezogen.

¹⁴ Mann-Whitney-U-Test

4.2 Sportfortsetzeranalyse

43 Probanden setzten den Sport im follow-up-Zeitraum freiwillig fort (36% der Stichprobe, 41% der Männer, 33% der Frauen). Für die 76 Versuchspersonen, die den Sport nach der Intervention nicht aufrecht erhielten (64% der Stichprobe, 59% der Männer, 67% der Frauen) kommen dieselben Ursachen infrage, die vorhin für den Studienabbruch (Kap. 4.1) diskutiert wurden: Ein Proband war verstorben. Von den Sportabbrechern nahmen 12 eine Arbeit auf, 14 gaben an, erkrankt gewesen zu sein und 49 Probanden hatten vermutlich kein Interesse, den Sport fortzusetzen. Umgekehrt betrachtet hatten von den 43 Sportfortsetzern aber 5 eine Arbeit gefunden und somit trotz ihrer neuen Anstellung weiter trainiert. Ebenso ist es möglich, dass Probanden trotz einer Erkrankung weiter trainierten. Beide Ursachen stellten also durchaus Wahrscheinlichkeitserhöhungen, aber keine hinreichenden Ursachen dafür dar, die sportliche Betätigung im 6-monatigen follow-up gänzlich abubrechen. Insofern sollte auch in der folgenden Analyse eine differenzierte Betrachtung der Abbruchgründe "Arbeit" und "Erkrankung" unberücksichtigt bleiben. Es wurden Analysen zur Beantwortung der Fragestellung durchgeführt, ob sich retrospektiv das Risiko "vorhersagen" lässt, dass der begonnene Sport von den Teilnehmern abgebrochen wurde¹⁵. Hierzu erfolgten Berechnungen von Odds und Odds-Ratios auf der Basis von logistischen Regressionen.

4.2.1 T1 Modell: Sportfortsetzeranalyse mit Daten der Erstuntersuchung

Ob Versuchspersonen, die nach der Intervention die sportliche Betätigung abbrachen, sich möglicherweise bereits in ihren Daten der Eingangsuntersuchung von denen unterschieden, die den Sport fortsetzten, wurde analog zur Vorgehensweise für die Drop-Out-Quoten analysiert: Zur Entdeckung von potentiell geeigneten Prädiktoren wurden für die abhängigen Variablen von T1 zwischen den Gruppen "Sportfortsetzer" und "Sportabbrecher" Unterschiedstests berechnet¹⁶.

Tab. 4.3: Tests der Gruppenunterschiede des Zwischensubjektfaktors "Sportfortsetzung" für T1-Variablen.

		LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum ¹⁷	
Forts. Sport	Nein	45	75	55	75	73	73	75	75	49	53	
	Ja	39	43	32	43	43	43	43	43	27	31	
		df	82	116	85	116	114	114	116	116	74	N=84
		t	-1,329	-0,996	-0,214	-1,023	-0,463	-1,227	-1,983	-1,598	-0,326	z=-2,911
		p	0,187	0,321	0,831	0,308	0,644	0,222	0,050	0,113	0,746	0,004

In den Tests bildete der tägliche Fernsehkonsum mit einer Differenz von +64 Minuten das distinkteste Unterscheidungsmerkmal der Sportabbrecher (251±102min) gegenüber den Fortsetzern (187±90min). Eine starke Tendenz ließ sich für ein a priori niedrigeres psychosoziales Wohlbefinden der Sportabbrecher (2,64±0,59) im Vergleich mit den Fortsetzern (2,86±0,53) ausmachen.

¹⁵ Dass das "Risiko des Sportabbruchs" statt der "Chance der Sportfortsetzung" bestimmt wurde, liegt alleine darin begründet, dass Odds-Ratios in der medizinischen Forschung traditionell zur Prädiktion des unerwünschten Ereignisses (z.B. Erkrankung) statt des gewünschten Resultats (z.B. Gesundung) herangezogen werden.

¹⁶ Die Prüfungen erfolgten unter Ausschluss des nach T2 verstorbenen Versuchsteilnehmers.

¹⁷ Mann-Whitney-U-Test.

Auf Itemebene stimmten die Fortsetzer den folgenden signifikant stärker zu (U-Test):

PSW Item 1: "Ich fühle mich unbeschwert und gut aufgelegt" ($Z=1,979$; $p=0,048$)

PSW Item 6: "Ich fühle mich voller Energie und Tatkraft" ($Z=2,749$; $p=0,006$)

SOC Item 7: "Ich fühle mich selten als Verlierer" ($Z=2,577$; $p=0,010$)

Eine Tendenz zur erhöhten Zustimmung ergab sich für die Abbrecher für:

SF12 Item 12: "Meine sozialen Kontakte sind eingeschränkt" ($Z=1,810$; $p=0,070$)

Um zu entscheiden, welche Variablen den größten Einfluss darauf hatten, ob ein Proband, der zu T1 befragt wurde, sich später gegen die Fortsetzung des Trainings entschied, wurde eine schrittweise logistische Regressionsanalyse berechnet. Mindestens tendenziell ($p < 0,1$) trennscharfe Messgrößen wurden als Prädiktoren für den Regressand "Sportabbruch" eingesetzt. Zwischen allen Prädiktoren ergaben sich vollständige Daten für 82 Probanden¹⁸.

Tab. 4.4: Ergebnisse der binären logistischen Regressionsanalyse für die Prädiktoren aus T1.

Modellkomponenten	Regress. Koeffizient	SE	Wert	p
Fernsehkonsument	0,466	0,167	7,734	0,005
SOC-7 ("Fühle mich selten als Verlierer")	-0,314	0,151	4,337	0,037
Konstante	0,185	0,826	0,050	0,823
PSW-1 ("Fühle mich unbeschwert und gut aufgelegt")	-	-	1,707	0,191
PSW-6 ("Fühle mich voller Energie")	-	-	1,007	0,316
SF12-12 ("Soziale Kontakte sind eingeschränkt")	-	-	0,501	0,479
PSW	-	-	0,037	0,847

In den Tests der Tab. 4.4 wurde geprüft, ob sich die Vorhersage des Modells signifikant ändert, wenn der betreffenden Komponente eine Null zugewiesen wird. Die Variable Fernsehkonsum ($p=0,005$) und das SOC-Item 7 ($p=0,037$) erwiesen sich als signifikante Koeffizienten. Die restlichen Prädiktoren fanden keinen Einzug in das Regressionsmodell, da sie es nicht mehr signifikant verbesserten. Die Anpassungsgüte des Modells wurde mit der Likelihood-Funktion beurteilt; als Maß wurde der negative doppelte Wert des Logarithmus benutzt. Dieser betrug $-2LL=96,11$ und war mit Hinzunahme der signifikanten Prädiktoren kleiner als der Ursprungswert, der nur die Konstante enthielt. Dieser Unterschied war hochsignifikant ($\chi^2=12,642$; $p=0,002$); die Varianzaufklärung ergab $R^2=0,194$ (nach Nagelkerke).

Prädiktionsmodell, Odds und Odds-Ratios

Mit den Regressionskoeffizienten wurde für jede Versuchsperson die Wahrscheinlichkeit berechnet, mit der das Ereignis "Sportabbruch" eintritt. Die Risikowahrscheinlichkeit eines Falles wurde nach dem Ansatz

$$p = 1 / (1 + e^{-z})$$

berechnet, wobei

$$z = b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n + a$$

x_i sind die Werte der unabhängigen Variablen, b_i sind Koeffizienten, deren Berechnung Aufgabe der logistischen Regression war, a ist die Konstante; e repräsentiert die Eulersche Zahl¹⁹. Ergibt sich für p ein Wert $< 0,5$, nimmt man an, dass das Ereignis nicht eintritt. Im

¹⁸ Hauptsächliche N-limitierende Variable zwischen den Prädiktoren stellte die Variable "TV-Konsum" dar.

¹⁹ $e=2,718281828459\dots$

anderen Fall nimmt man das Eintreffen des Ereignisses an. Im Prädiktionsmodell resultierte für den Exponenten z nach Einsetzen der Koeffizienten

$$z = 0,446 * [\text{Fernsehkonsument}] + (-0,314 * [\text{SOC-7}]) + 0,185$$

Inhaltlich bedeutet die Beziehung: Je höher der Fernsehkonsum war und je weniger der Aussage "Ich fühle mich selten als Verlierer" zugestimmt wurde, desto größer der Exponent z , welches nach Einsatz in die Formel zu einer steigenden Wahrscheinlichkeit des Ereignisses führt. Beispiel: Für einen Probanden, der täglich **3,5** Stunden fernsah und das Item "Ich fühle mich selten als Verlierer" mit **2** (\approx "stimme kaum zu") beantwortete, resultiert

$$z = 0,466 * 3,5 + (-0,314 * 2) + 0,185 = 1,188$$

und hiermit

$$p = 1 / (1 + e^{-1,188}) = 0,767$$

Da im Beispiel die Wahrscheinlichkeit des Sportabbruchs 77% beträgt, wäre das Eintreffen des Ereignisses angenommen. Im Folgenden findet sich eine Klassifizierungstabelle, in der für die 82 Fälle die beobachtete Gruppenzugehörigkeit der aufgrund des berechneten Modells vorhergesagten gegenübergestellt ist:

Tab. 4.5: Beobachtete und vorhergesagte Häufigkeiten der Sportfortsetzung und des -abbruchs.

Beobachtet		Vorhergesagt		
		Sport		Prozentsatz der Richtigen
		Abbruch	Fortsetzung	
Sport	Abbruch	41	10	80,4%
	Fortsetzung	15	16	51,6%
Gesamtprozentsatz				69,5%

Demnach sind von den 51 Sportabbrechern 41 richtig klassifiziert (80%) und von den 31 Sportfortsetzern sind mit 16 ca. die Hälfte (52%) korrekt beurteilt worden. Insgesamt waren 57 von 82 Fällen richtig zugeordnet; das ergibt eine Trefferquote von 69,5%.

Die Berechnungen von Odds und Odds-Ratios zum Vergleich der Prädiktion mit den tatsächlich eintreffenden Ereignissen erfolgten nach Bland und Altman (2000). Für die Prädiktion "Sportabbruch" ergab sich eine Odds von 1,661 (CI=1,127 – 2,449), dass der Sport tatsächlich abgebrochen wurde. Die Odds, dass sich trotz der Prädiktion der Sport fortsetzte, betrug 0,380 (CI=0,198 – 0,729).

Um eine Aussage darüber fällen zu können, mit welcher erhöhten Wahrscheinlichkeit beim Vorliegen der Riskoprädiktion das Ereignis eintrat als die Gegenwahrscheinlichkeit, wurde das Odds-Ratio berechnet. Dieses betrug **OR=4,373 (CI=1,630 – 11,734)**. Das 95%-Konfidenzintervall des Odds-Ratios umschließt nicht 1 und ist demzufolge statistisch signifikant. Erhält ein Fall die retrospektive "Vorhersage", dass das Training nicht fortgesetzt wird, hatte dieser das 4,4-fach erhöhte Risiko, den Sport tatsächlich abubrechen, als den Sport – entgegen der Vorhersage – fortgesetzt zu haben.

4.2.2 Kombiniertes Modell: Sportfortsetzeranalyse mit Veränderungsdaten und Prädiktoren des T1-Modells

Es wurde geprüft, inwieweit mögliche Veränderungsunterschiede nach der Intervention zur Wahrscheinlichkeitserhöhung geführt haben, dass Probanden im follow-up den Sport abbrechen. Hierzu wurden aus den Fragebögen und den Items Differenzwerte (T2 minus T1) gebildet. Hatte z.B. ein Proband vor dem Training ein BDI=55 und nach dem Training ein BDI=50, resultierte ein Differenzwert von $\text{Dif}_{(T2-T1)}\text{BDI}=-5$, also eine Reduktion der Depressivität um 5 Standardwerte.

Tab. 4.6: Tests der Unterschiede des Zwischensubjektfaktors "Sportfortsetzung" für Differenzwerte T2-T1.

		Dif _(T2-T1) GHQ	Dif _(T2-T1) BDI	Dif _(T2-T1) PSK	Dif _(T2-T1) SOC	Dif _(T2-T1) SWE	Dif _(T2-T1) PSW	Dif _(T2-T1) KSK	Dif _(T2-T1) VAS	Dif _(T2-T1) TV ²⁰
Forts. Sport	Nein	51	40	46	50	50	51	46	34	37
	Ja	42	30	42	42	42	42	42	26	30
		df	68	86	90	90	91	86	58	N=67
		t	0,398	-0,398	-0,074	1,249	1,599	0,663	-0,426	Z=-0,929
		p	0,692	0,641	0,692	0,941	0,215	0,113	0,509	0,353

Wie man sieht, unterschieden sich Sportfortsetzer nicht signifikant von den Abbrechern in den Veränderungswerten der Messverfahren. Signifikante Unterschiede zwischen den Gruppen konnten jedoch in den Differenzen der folgenden Items gefunden werden (U-Test):

- Dif_(T2-T1)BDI Item 7 "Ich bin von mir enttäuscht" ($Z=2,022$; $p=0,043$)
- Dif_(T2-T1)BDI Item 8 "Ich werfe mir Fehler und Schwächen vor" ($Z=2,033$; $p=0,042$)
- Dif_(T2-T1)BDI Item 9 "Ich denke daran, mir etwas anzutun" ($Z=2,193$; $p=0,028$)
- Dif_(T2-T1)SF12 Item 9 "Ich fühle mich ruhig und gelassen" ($Z=2,995$; $p=0,003$)
- Dif_(T2-T1)PSW Item 2 "Ich fühle mich anderen unterlegen" ($Z=2,326$; $p=0,020$)

In den ersten 4 Items hatten die Fortsetzer signifikant gesundheitsförderlichere Veränderungen nach der Intervention als die Abbrecher, also weniger eigene Enttäuschung (BDI-7), weniger Selbstvorwürfe (BDI-8), weniger selbstbestrafende Gedanken (BDI-9) und mehr Ruhe und Gelassenheit (SF12-9). Im PSW Item 2 hatten die Abbrecher eine höhere positive Veränderung, also weniger Unterlegenheitsgefühle. Die trennscharfen Items wurden wie im vorherigen Kap. 4.2.1 als Prädiktoren benutzt, um das Risiko des Sportabbruchs mittels einer logistischen Regression vorherzusagen. Sie wurden zusätzlich mit den signifikanten Koeffizienten des T1-Modells (Fernsehkonsument und SOC-7: "Fühle mich selten als Verlierer") innerhalb der Analyse kombiniert. Vollständige Daten zwischen sämtlichen Prädiktoren lagen für 60 Versuchspersonen vor.

Der folgenden Tab. 4.7 ist entnehmbar, dass die Veränderungen des SF12-9 ("Ich fühle mich ruhig und gelassen") ($p=0,001$) sowie die des BDI-07 ("Ich bin von mir enttäuscht") ($p=0,036$) und der T1-Fernsehkonsument ($p=0,011$) sich als signifikante Regressionskoeffizienten erwiesen, während die restlichen Komponenten entbehrlich blieben. Die Anpassungsgüte betrug $-2LL=58,16$ und war mit $\chi^2=24,953$ ($p=0,000$) hochsignifikant besser angepasst als von der Nullhypothesenannahme. Die Varianzaufklärung erreichte $R^2=0,454$ (nach Nagelkerke).

²⁰ Mann-Whitney-U-Test

Tab. 4.7: Ergebnisse der binären logistischen Regressionsanalyse für die Prädiktoren der Differenzwerte und der signifikanten Koeffizienten des T1-Modells.

Modellkomponenten	Regress. Koeffizient	SE	Wert	p
Dif _(T2-T1) SF12-9 ("Fühle mich ruhig und gelassen")	1,299	0,403	10,373	0,001
Fernsehkonsum	0,557	0,219	6,455	0,011
Dif _(T2-T1) BDI-07 ("Ich bin von mir enttäuscht")	0,584	0,278	4,404	0,036
Konstante	-1,846	0,857	4,641	0,031
SOC-7 ("Fühle mich selten als Verlierer")	-	-	3,070	0,080
Dif _(T2-T1) BDI-09 ("Ich denke daran, mir was anzutun")	-	-	2,348	0,125
Dif _(T2-T1) PSW-2 ("Ich fühle mich anderen unterlegen")	-	-	1,287	0,257
Dif _(T2-T1) BDI-08 ("Werfe mir Fehler und Schwächen vor")	-	-	0,905	0,341

Erneut wurde das Risiko des Sportabbruchs berechnet. Die Gleichung für den Exponenten z der Formel $p = 1 / (1 + e^{-z})$ lautete:

$$z = 1,299 * [\text{Dif}_{(T2-T1)}\text{SF12-9}] + 0,557 * [\text{Fernsehkonsum}] + 0,584 * [\text{Dif}_{(T2-T1)}\text{BDI-07}] - 1,846$$

Inhaltlich Bedeutung: Je weniger Ruhe und Gelassenheit der Proband nach dem Training hinzugewinnen konnte und je höher sein ursprünglicher Fernsehkonsum war und je weniger er Gefühle eigener Enttäuschung im Interventionszeitraum abbauen konnte, desto größer war das Sportabbruchrisiko. Im Folgenden findet sich die Klassifizierungstabelle, in der die beobachtete Gruppenzugehörigkeit der Modellvorhersage gegenübergestellt ist:

Tab. 4.8: Beobachtete und vorhergesagte Häufigkeiten der Sportfortsetzung und des -abbruchs.

Beobachtet		Vorhergesagt		
		Sport		Prozentsatz der Richtigen
		Abbruch	Fortsetzung	
Sport	Abbruch	25	6	80,6%
	Fortsetzung	9	20	69,0%
Gesamtprozentsatz				75,0%

Demnach wurden von den 31 Abbrechern 25 richtig klassifiziert (81%) und von den 29 Sportfortsetzern 20 (69%). Insgesamt waren mit dem Modell 45 von 60 Fällen korrekt beurteilt; dies ergibt eine Trefferquote von 75%.

Für die Risikoprädiktion ergab sich eine Odds von 2,599 (CI=1,471 – 4,592) zum tatsächlichen Sportabbruch. Die Odds, dass sich trotzdem der Sport fortsetzte betrug 0,281 (CI=0,131 – 0,599). Das Odds-Ratio des Risikos erreichte **OR=9,259 (CI=2,821 – 30,394)**. Weil das Konfidenzintervall 1 nicht umschließt ist das Odds-Ratio signifikant. Erhielt ein Proband die Klassifizierung "Sportabbruch", hatte er ein um den Faktor 9,3 höheres Risiko, den Sport tatsächlich abgebrochen statt fortgesetzt zu haben.

4.3 L1, L2, L3, L4: Tests der Längsschnitteffekte

In folgender Tab. 4.9 sind die Längsschnitthypothesentests dargestellt. Die Hypothesen L1-L4 wurden nach dem in Kap. 2.10.2.1 vorgestellten Schema geprüft. Pro Messinstrument ist die Prüfung des Innersubjektfaktors aus der ANOVA mit Messwiederholung dargestellt. Der dort dargestellte Innersubjektfaktor-Haupteffekt (**A**) ist erst in Kap. 4.5 von Wichtigkeit und wird deswegen hier noch nicht besprochen.

Tab. 4.9: Test der Längsschnitthypothesen L1-L4.

	L1 Risikofaktoren			L2 Ressourcen			L3 Körper		L4	
	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum	
	N	56	57	56	72	72	69	70	50	52
	K	4	4	4	2	2	2	4	4	4
(A) <i>Innersubj.-Haupteffekt T1, T2, T3</i>	df1	2	1,783	2	1,82	2	2	2	2	2
	df2	104	94,34	104	127,4	140	134	132	92	96
	F	9,058	2,029	3,762	1,220	0,986	1,531	0,457	1,659	4,334
	p	0,000	0,142	0,026	0,296	0,376	0,220	0,634	0,196	0,016
(B) <i>Innersubj.-Kontrast T1 vs. T2</i>	df1	1	1	1	1	1	1	1	1	1
	df2	52	53	52	70	70	67	66	46	48
	F	17,750	5,136	5,697	0,051	0,284	2,055	0,615	2,443	7,532
	p	0,000	0,028	0,021	0,823	0,596	0,156	0,436	0,125	0,008
(C) <i>Helmert-Kontrast T1 vs. T2&T3</i>	df1	1	1	1	-	-	-	-	-	1
	df2	52	53	52	-	-	-	-	-	48
	F	7,647	2,257	0,991	-	-	-	-	-	5,420
	p	0,008	0,139	0,324	-	-	-	-	-	0,024

K: Zahl der Gruppen der Zwischensubjektffaktoren, die im Rahmen der Moderatorprüfungen (Kap. 4.4) mitgetestet wurden. BDI und SOC erfüllten nicht die Sphäritätsannahme, weshalb ihre Freiheitsgrade unter (A) nach Greenhouse und Geisser (1959) korrigiert wurden.

Unter (B) erfolgten die Tests der im Kap. 2.9.2 aufgestellten Interventionshypothesen. Anhand der *p*-Werte sieht man, dass im Gesundheitsförderungszeitraum

im GHQ (T1: 1,890±0,488; T2: 2,131±0,434; T3: 1,962±0,546),

im BDI (T1: 58,50±6,69; T2: 56,45±9,22; T3: 58,16±9,20),

im PSK (T1: 49,11±7,7; T2: 51,92±7,27; T3: 48,57±9,64) und

im TV-Konsum (T1: 214,6±106,7min.; T2: 178,7±98,3min.; T3: 197,9±129,3 min.)

signifikante Veränderungen zwischen T1 und T2 ausgelöst wurden. Die Inspektion der deskriptiven Werte (Randmittel der ANOVA) ergab, dass diese sich von T1 nach T2 stets in die hypothetisierte Richtung veränderten, weswegen sämtliche Hypothesen unter L1 und die von L4 bestätigt sind. Die Tests unter L2 und L3 zeigten hingegen keine Signifikanz für eine Veränderung von T1 nach T2.

Die Hypothesen unter L1 und die von L4 sind bestätigt: Durch die Intervention sanken unspezifische psychische Beschwerden (GHQ), Depressivität (BDI), psychische Einschränkungen der Lebensqualität (PSK) und das passive Freizeitverhalten (TV-Konsum). Die Hypothesen unter L2 und die von L3 sind nicht bestätigt: Psychische Ressourcen (SOC, SWE, PSW) konnten nicht signifikant gesteigert werden und auch die erwartete Erhöhung der subjektiven körperlichen Gesundheit (KSK, VAS) war nicht signifikant belegbar.

Unter (C) wurde nach dem im Kap. 2.10.2.1 dargestellten Schema der Helmert-Kontrast getestet, falls der Interventionshypothesentest unter (B) bestätigt war. Der Helmert-Kontrast war im GHQ und im TV-Konsum signifikant, im BDI und in der PSK hingegen nicht.

Unter Einschluss der 6-monatigen follow-up-Veränderung waren relativ zeitstabile Interventionseffekte für die Reduktion von unspezifischen psychischen Beschwerden (GHQ) und der Senkung des passiven Freizeitverhaltens (TV-Konsum) belegbar. Demgegenüber ist für die Reduktion von Depressivität (BDI) und der Minderung der psychischen Einschränkung der Lebensqualität (PSK) von lediglich kurzfristig wirksamen Interventionseffekten auszugehen.

4.3.1 Standardisierte Darstellung sämtlicher Längsschnitteffekte

Um die Interventions- nebenst follow-up-Veränderungen vergleichend innerhalb einer Übersichtsgrafik darstellen zu können, wurden die Daten der Messverfahren jeweils pro Erhebungszeitpunkt anhand der zu T1 Randmittel und Standardabweichungen z-transformiert. Nach der Transformation beginnen alle Messverfahren mit dem Ursprung Null im Zeithorizont von T1, womit das Ausmaß ihrer Veränderungen nach T2 und nach T3 vergleichend erfasst werden kann. Die z-Werte des BDI und des TV-Konsums wurden zudem umgepolt, damit für sämtliche Maße nach der Standardisierung eine numerisch positive Veränderungsrichtung gleichzeitig eine gesundheitsförderliche Verbesserung darstellt.

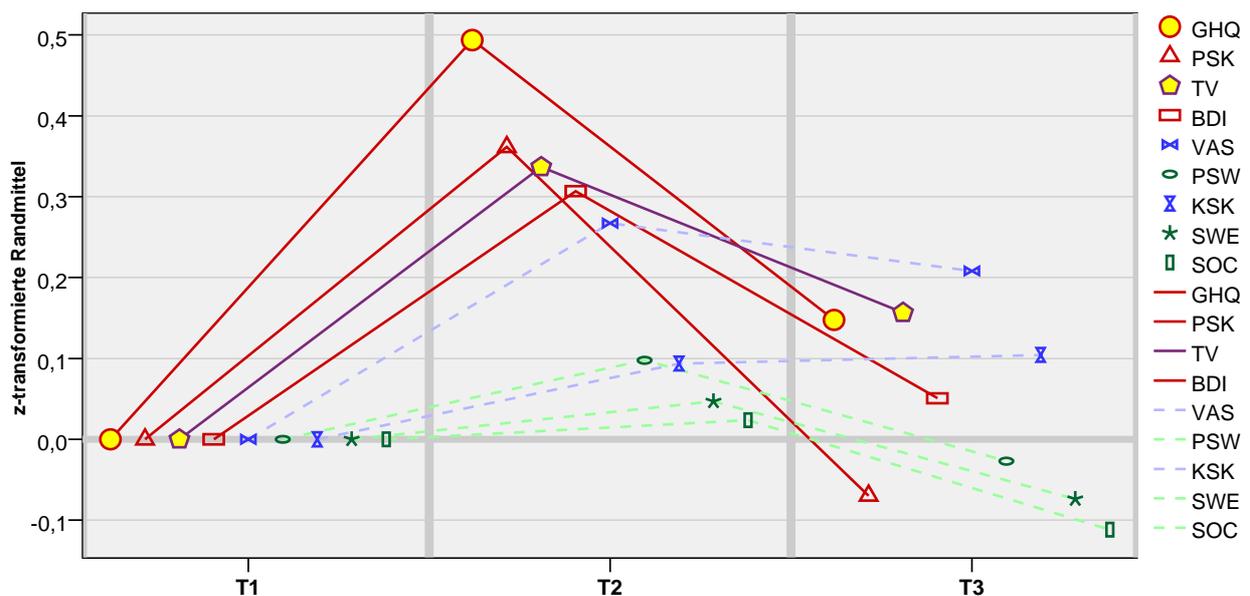


Abb. 4.1: z-standardisierte Veränderungen sämtlicher Längsschnittvariablen. Die Veränderungen von BDI und TV-Konsum sind umgepolt dargestellt (weitere Erläuterung siehe Text).

Die Variablen der Interventionshypothesen zu **L1 (psychische Risikofaktoren)** sind rot, die von **L2 (psychische Ressourcen)** sind grün und die von **L3 (Körper)** sind blau gefärbt. Die Variable zur Hypothese **L4 (TV-Konsum)** ist lila dargestellt. Die Längsschnittsdaten der Messverfahren, deren Interventionshypothesen bestätigt werden konnten, sind mit einer durchgezogenen Linie verbunden. Die nichtsignifikanten "Veränderungen" der restlichen Messverfahren sind mit abgeschwächter Farbe und gestrichelten Verbindungen angedeutet. Die Icons der Messverfahren, die einen signifikanten **Helmert-Kontrast** erreichten, haben eine gelbe Füllung, die restlichen sind ungefüllt dargestellt.

Klar ist erkennbar, dass im GHQ die höchsten standardisierten Mittelwertsunterschiede zw. T1-T2 gefunden wurden, gefolgt von der PSK, TV-Konsum und BDI. Es visualisiert sich bei der standardisierten Betrachtung, dass die Intervention in den psychischen Risikofaktoren relativ starke Veränderungen bewirkte, wenn man sie mit den Effekten zu den psychischen Ressourcen vergleicht, die über die Zeitpunkte nahezu unverändert blieben. Die Veränderungshöhe der Variablen zur körperlichen Gesundheit lagen deskriptiv im Mittel über die der psychischen Ressourcen.

Auch kristallisiert sich auf einen Blick heraus, wie sich die signifikanten Interventionseffekte, die zusätzlich einen signifikanten Helmert-Kontrast erreichten (GHQ, TV-Konsum) von den restlichen signifikanten Interventionseffekten (PSK, BDI) unterscheiden. Nach dem signifikanten Interventionseffekt kommt es für GHQ und TV zwar auch zu einer Abwärtskorrektur, allerdings bewegt sich diese weniger stark auf das Ursprungsniveau zurück als im BDI und in der PSK. Der Helmert-Kontrast konnte somit die signifikanten Trainingseffekte in gleich große Gruppen trennen: Zwei Interventionseffekte erreichten eine Zeitstabilität, zwei nicht.

Schließlich ist dem Diagramm zu entnehmen, dass die Veränderungsrichtung zw. T1-T2 in allen Messgrößen zumindest die hypothetisierte Richtung andeutete; d.h. Ausschläge unterhalb des Startwertes (Null) fanden im Zeitraum von T1 nach T2 nicht statt.

4.3.2 Effektstärken der Interventionseffekte

Anhand der Innersubjektkontraste der Messwiederholung T1-T2 wurden die Interventionshypothesen L1-L4 jeweils unter dem Punkt **(B)** der Tab. 4.9 getestet. Eine Möglichkeit, die Interventionsresultate zu vergleichen, könnte darin bestehen, die p -Werte der Kontraste T1-T2 gegenüber zu stellen. Diese Methode ist bei unterschiedlichem N natürlich problematisch und wurde in der psychologischen Forschung immer wieder kritisch reflektiert, wenn es um die vergleichende Bewertung von Experimenten anhand der p -Werte ging (z.B. Bredenkamp, 1969; Hager, 2006).

Der Einsatz des BDI, VAS und TV-Konsum erfolgte zu einem späteren Zeitpunkt in der Studie, welches deren N verkleinerte. Des Weiteren kam die Depressivitätsdiagnose als Zwischensubjektfaktor der ANOVAS für GHQ, PSK, TV-Konsum (und BDI) zur Anwendung, welches deren N um die nicht vorhandenen T1-Werte des BDI (der zur Stellung der Depressivitätsdiagnose vonnöten war) verkleinerte. Weil die Hypothesentests auf unterschiedliche Stichprobengrößen fußten und p -Werte vom N beeinflusst werden, wäre über den Signifikanzvergleich keine faire Betrachtung zwischen den Interventionsergebnissen gewährleistet.

Demgegenüber stellt die **Effektstärke η^2** (Eta-Quadrat) einen von der Stichprobengröße unabhängigen Effekt dar²¹. η^2 spiegelt den Anteil der aufgeklärten Variabilität zwischen den Messzeitpunkten wider und erreichte z.B. im Interventionseffekt des GHQ eine Aufklärung von 25,4% (vgl. folgende Abb. 4.2). Da es mit η^2 möglich ist, die Trainingserfolge unabhängig vom N zu vergleichen, wurden in nachfolgender Abb. 4.2 die Effektstärken absteigend geordnet dargestellt.

²¹ In der ANOVA mit Messwiederholung wird η^2 sowohl von den Mittelwertsunterschieden und Standardabweichungen als auch von der Korrelation zwischen den Messzeitpunkten beeinflusst (Rasch, Friese, Hofmann, & Naumann, 2010). Insofern stellt η^2 hier den Unterschiedseffekt bei abhängigen Stichproben dar. Zur Generalisierung ist zu bemerken, dass η^2 die Effektgröße der Population stets etwas überschätzt (Rasch, et al., 2010, S 114); die wahren Populationseffekte werden daher etwas niedriger als die hier empirisch berechneten sein.

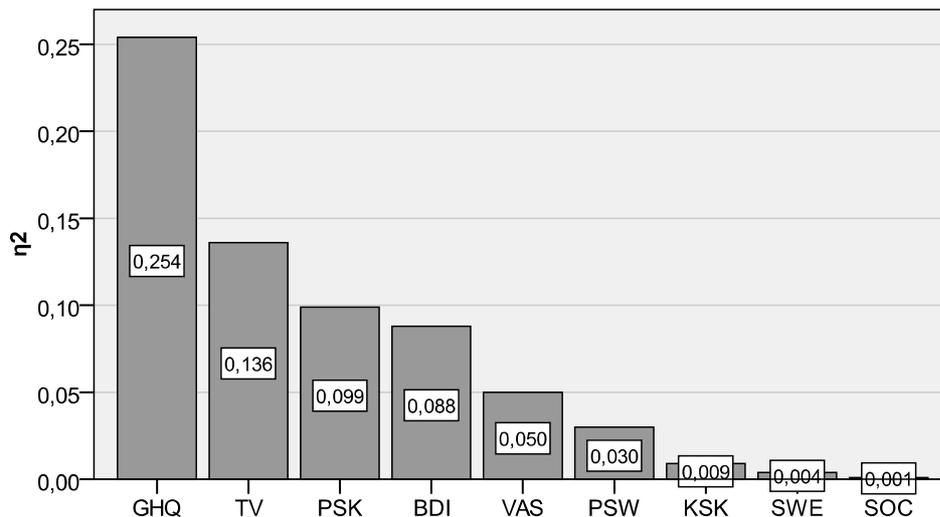


Abb. 4.2: Darstellung der Interventions-Effektstärken.

Es ist erkennbar, dass im GHQ der mit Abstand höchste Interventionseffekt zu finden war. Er überragt den zweithöchsten Effekt (Reduktion des TV-Konsums) annähernd um das Doppelte. Für die signifikanten Interventionseffekte ergibt sich folgende Reihenfolge ihrer Effektivität: In Relation zu den restlichen Effekten gelang es in der Intervention vor allem, unspezifische psychische Beschwerden zu lindern (GHQ), gefolgt von der Senkung des passiven Lebensstils (TV), gefolgt von der Minderung einer psychisch eingeschränkten Lebensqualität (PSK) und gefolgt von der Senkung der Depressivität (BDI). In Tab. 4.10 wurde η^2 zu Klassifikationszwecken mit folgenden Konventionen von Cohen (1988, 1992) beurteilt: $\eta^2 \geq 0,010$: "kleiner Effekt"; $\eta^2 \geq 0,059$: "mittlerer Effekt"; $\eta^2 \geq 0,138$: "großer Effekt".

Tab. 4.10: Effektstärke η^2 des Innersubjektcontrasts (T1-T2) und statistische Power des Interventionshypothesentests nebst durchschnittlicher Effektstärke pro inhaltlicher Interventionsebene.

	p	$r_{(T1,T2)}$	η^2	η^2 Cohen Class.	N	1- β	Inhaltliche Ebene	$\emptyset \eta^2$	$\emptyset \eta^2$ Cohen Class
GHQ	0,000	0,706	0,254	Groß	56	0,985	Psychische Risikofaktoren	0,147	Groß
BDI	0,028	0,766	0,088	Mittel	57	0,605			
PSK	0,021	0,474	0,099	Mittel	56	0,649			
SOC	0,823	0,575	0,001	-	72	0,056	Psychische Ressourcen	0,012	Klein
SWE	0,596	0,699	0,004	-	72	0,082			
PSW	0,156	0,830	0,030	Klein	69	0,293	Subj. Körperliche Gesundheit	0,030	Klein
KSK	0,436	0,482	0,009	-	70	0,121			
VAS	0,125	0,077	0,050	Klein	50	0,334			
TV	0,008	0,591	0,136	Mittel	52	0,767	Passiver Le- bensstil	0,136	Mittel

p : Signifikanz des Innersubjektcontrasts T1-T2 (vgl. (B) in Tab. 4.9); $r_{(T1,T2)}$: Korrelation zwischen den Messzeitpunkten T1 und T2; η^2 : Effektstärke des Innersubjektcontrasts T1-T2; 1- β : Power des Hypothesentests; Cohen Class.: Klassifikation der Effektstärke nach den Konventionen von Cohen (1988, 1992).

Nach Cohen (1988, 1992) erreichte der Interventionseffekt des GHQ einen "großen" Effekt, die Interventionsergebnisse zum BDI, PSK und TV-Konsum ergaben jeweils "mittlere" Effekte und die von PSW und VAS jeweils "kleine" Effekte. Die Interventionsresultate des SOC, SWE und KSK lagen noch unterhalb der Schwelle von "kleinen" Effekten.

Zusammengefasst für die inhaltlichen Ebenen der Gesundheit zeigte sich für die Reduktion von *psychischen Risikofaktoren* durchschnittlichen ein "großer" Effekt, welches hauptsächlich dem relativ überragenden Effekt im GHQ geschuldet ist. Demgegenüber gelang der Aufbau von *psychischen Ressourcen* im Mittel nur knapp oberhalb der Schwelle eines "kleinen" Effektes. Die Erhöhung der (*subjektiven*) *körperlichen Gesundheit* erreichte durchschnittlich einen "kleinen" Effekt. Die Reduktion des *passiven Lebensstils* (TV-Konsum) erreichte einen "mittleren" Effekt und lag knapp unterhalb der Schwelle eines "großen".

Es kristallisiert sich heraus, dass in dieser Studie nur "große" bis "mittlere" Interventionseffekte ihre Wirkung signifikant (p) unter Beweis stellen konnten; "kleine" Effektstärken (und darunter) waren sämtlich nicht mehr statistisch abgesichert. Des Weiteren ist in Tabelle 4.10 die Power ($1-\beta$) berichtet, die den Schätzer der Teststärke der jeweiligen Interventionshypothese (H_1) darstellt. Innerhalb der signifikanten Hypothesentest (p) betrug die kleinste Power $1-\beta=0,61$ (BDI). Powergrößen darunter führten nicht mehr zur Signifikanz.

Mit dem Innersubjektkontrast der Messwiederholung T1-T2 wurden die hiesigen Interventionshypothesen getestet, weshalb η^2 die adäquate Effektgröße darstellt. Trotzdem sind in folgender Tab. 4.11 die Effektstärken des Unterschiedes zwischen T1-T2 zusätzlich als **Cohen's d** angegeben, welche die klassische Effektgröße für einfache paarweise Vergleiche darstellt. Z.B. Metaanalysen, wie die von Paul und Moser (2009a) verwenden d wenn Unterschiedseffekte berichtet werden.

Tab. 4.11: Effektstärke Cohen's d für paarweisen Vergleich (T1-T2) bei abhängigen Stichproben nebst durchschnittlicher Effektstärke pro inhaltlicher Interventionsebene.

	T1	T2	$r_{(T1,T2)}$	d	Inhaltliche Ebene	\bar{d}
GHQ	1,890±0,49	2,131±0,43	0,706	0,675	Psychische Risikofaktoren	0,462
BDI	58,50±6,69	56,45±9,22	0,766	0,345		
PSK	49,11±7,70	51,92±7,27	0,474	0,366		
SOC	48,61±8,96	48,83±8,33	0,575	0,028	Psychische Ressourcen	0,087
SWE	51,38±10,1	51,86±9,40	0,699	0,063		
PSW	2,822±0,58	2,878±0,54	0,830	0,171		
KSK	46,72±8,98	47,56±8,61	0,482	0,094	Subj. körperliche Gesundheit	0,158
VAS	70,94±22,1	76,84±16,6	0,077	0,222		
TV	214,6±106,0	178,7±98,3	0,591	0,386	Passiver Lebensstil	0,386

T1, T2: Randmittel der Untersuchungszeitpunkte; $r_{(T1,T2)}$: Korrelation zwischen den Messzeitpunkten T1 und T2; d : Effektstärke Cohen's d für paarweise Vergleiche bei abhängigen Stichproben.

Cohen's d wurde nach der Methode von Faul et al. (2007) für paarweise Vergleiche bei abhängigen Stichproben bestimmt. Für die Berechnung benötigt man die deskriptiven Werte und die Korrelation zwischen den Messzeitpunkten T1 und T2 (vgl. Tab. 4.11). Die Interventionseffektstärken in Form von d werden benötigt, um im Diskussionsteil die Effektivität des Programms mit anderen Studien vergleichen zu können (siehe Kap. 5.2.2).

4.4 M1, M2, M3, M4: Tests der Moderatoreffekte

Im Folgenden sind die Tests zu den Zwischensubjekt-Innersubjekt-Interaktionen nach dem im Kap. 2.10.2.2 vorgestellten Schema dargestellt, beginnend mit den Tests der **M1** Moderatorhypothese zur Intensitätssteigerung des Trainings in der 2. Interventionskohorte.

Tab. 4.12: Tests der Moderatorhypothese **M1**.

		GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum
Interventions-Kohorte	1.	17	19	20	33	33	30	34	11	13
	2.	39	38	36	39	39	39	36	39	39
(a) Helmert-Interaktion T1 vs. T2&T3	df1	1	1	1	1	1	1	1	1	1
	df2	52	53	52	70	70	67	66	46	48
	F	0,991	0,375	2,075	0,298	9,758	0,606	0,566	0,001	0,060
	p	0,324	0,543	0,156	0,593	0,003	0,439	0,455	0,972	0,807
(b) follow-up-Interaktion T2 vs. T3	df1	1	1	1	1	1	1	1	1	1
	df2	52	53	52	70	70	67	66	46	48
	F	0,156	0,649	2,075	2,236	1,750	2,767	0,197	2,260	2,314
	p	0,695	0,424	0,156	0,133	0,190	0,102	0,659	0,140	0,135

Außer in der SWE zeigte der Zwischensubjektfaktor "Interventionskohorte" keinen signifikant moderierenden Einfluss auf die Veränderung der Zeiteffekte (a). Die Inspektion der Daten für die SWE ergab, dass die 2. Kohorte, ausgehend von T1 ($49,31 \pm 10,35$) nach T2 ($51,26 \pm 9,41$) eine Verbesserung erfuhr, um in Richtung T3 ($51,31 \pm 12,12$) stabil zu bleiben, während unerwarteter Weise in der 1. Kohorte der gegenteilige Effekt zu beobachten war: Hier wurde im Zeitverlauf von T1 ($53,46 \pm 9,68$) über T2 ($52,45 \pm 9,31$) und nach T3 ($49,97 \pm 10,74$) die SWE stetig kleiner (vgl. folgende Abb. 4.3).

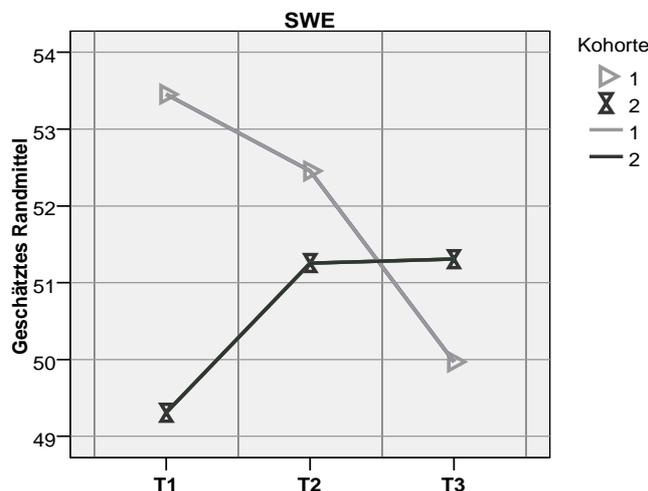


Abb. 4.3: Interaktionseffekt Kohorte * Innersubjektfaktor der SWE.

Während der Interaktionskontrast zwischen den Gruppen in die hypothetisierte Richtung verlief (2. Kohorte profitierte deskriptiv tatsächlich mehr als die 1. Kohorte), war jedoch unerwartet, dass sich die 1. Kohorte über die Zeit scheinbar verschlechterte. Um zu prüfen, ob innerhalb der Kohorten Unterschiede zwischen T1-T2 belegbar waren, wurden zusätzlich kohortenspezifische paarweise Vergleiche getestet. In den t-Tests für abhängige Stichproben erfolgte eine Anpassung der Signifikanzschwelle für die zweifache Testung als Šidák-Kor-

rektur²² nach Abdi (2007). Für die 2. Kohorte zeigte sich keine signifikante Erhöhung der SWE zwischen T1-T2 ($t_{(df=38)}=1,469$; $p=0,150$). Auch für die 1. Kohorte war die Verringerung der SWE zwischen T1-T2 nicht signifikant ($t_{(df=32)}=-0,884$; $p=0,383$). Dass die 2. Kohorte für sich genommen im Messzeitpaar T1-T2 keine signifikante Erhöhung der SWE zeigte und sich in der 1. Kohorte keine signifikante Verschlechterung offenbarte, liefert einen Hinweis darauf, dass der signifikante Interaktionseffekt sich nicht gezeigt hätte, wenn die (erwartete) numerische Erhöhung der SWE in der 2. Kohorte nicht gleichzeitig mit der (unerwarteten) numerischen Verringerung der 1. Kohorte hätte interagieren können. Um aufzuklären, ob der Interaktionseffekt als Phänomen einer "Regression zur Mitte" aufgefasst werden könnte, wurden die verschiedenen anmutenden Ausgangsbedingungen zu T1 zwischen den Gruppen (1. Kohorte: $53,46 \pm 9,68$; 2. Kohorte: $49,31 \pm 10,35$) mit dem t-Test für unabhängige Stichproben geprüft: Hier offenbarte sich ein tendenzieller Unterschied ($t_{(df=70)}=1,745$ $p=0,085$).

Die Intensivierung des Ausdauertrainings bewirkte in 8 von 9 Moderatorprüfungen der Hypothese M1 keine differenzierten Interventionserfolge. Lediglich in der SWE ist die Annahme der Hypothese prinzipiell zu bestätigen. Allerdings bleibt es uneindeutig, ob der Interaktionseffekt mit den Veränderungen der SWE tatsächlich auf ein überlegenes Training der 2. Kohorte zurückzuführen ist.

In folgender Tabelle sind die Tests der Moderatorhypothesen **M2** zur Prüfung des intervenionsmoderierenden Einfluss des Zwischensubjektfaktors "Depressivität" dargestellt:

Tab. 4.13: Tests der Moderatorhypothesen **M2**.

		GHQ	BDI	PSK	TV-Konsum
Depressivität	Nein	39	40	40	35
	Ja	17	17	16	17
(a) <i>Helmert-Interaktion</i> T1 vs. T2&T3	df1	1	1	1	1
	df2	52	53	52	48
	F	5,036	5,610	4,457	0,473
	p	0,029	0,022	0,040	0,495
(b) <i>follow-up-Interaktion</i> T2 vs. T3	df1	1	1	1	1
	df2	52	53	52	48
	F	5,070	2,457	1,593	6,677
	p	0,029	0,123	0,213	0,013

Im GHQ war der Helmert-Kontrast der Interaktion **(a)** signifikant, welches als Beleg dafür zu werten ist, dass die Depressivitätsdiagnose zu einem unterschiedlichen Trainingserfolg führte. Der Kontrast zwischen T2-T3 des GHQ war ebenso signifikant **(b)**, welches zeigt, dass die Depressivität auch isoliert im follow-up für differenzierte Effekte sorgte. Deskriptiv zeigte sich im GHQ, dass die Probanden mit Depressivität die stärkste Verbesserung aufwiesen (T1: $1,455 \pm 0,816$; T2: $1,879 \pm 0,733$), die sich im follow-up abschwächte (T3: $1,598 \pm 0,913$), während die restlichen Probanden nur eine sehr schwache Erhöhung zwischen T1-T2 offenbarten (T1: $2,325 \pm 0,531$; T2: $2,383 \pm 0,479$), die wieder auf das Ausgangsniveau zurückkehrte (T3: $2,326 \pm 0,591$) (vgl. folgende Abb. 4.4).

²² Adjustierung nach Šidák für Hochsignifikanz: $p=0,00501$; (einfache) Signifikanz: $p=0,02532$; Trend: $p=0,05132$.

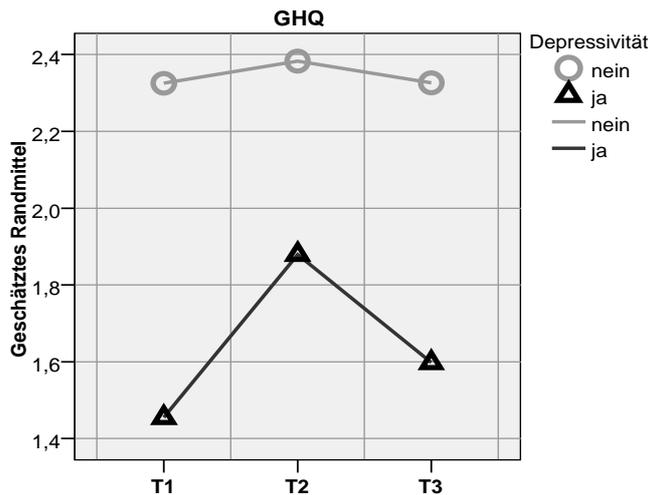


Abb. 4.4: Interaktionseffekt Depressivität * Innersubjektfaktor im GHQ.

Auch im BDI belegte die signifikante Helmert-Interaktion (a) (Tab. 4.13) einen unterschiedlichen Interventionserfolg. Auch hier profitierten hauptsächlich die Probanden mit Depressivität vom Training. Von diesen senkte sich die Depressivität (T1: $66,41 \pm 11,362$; T2: $61,50 \pm 15,66$), um sich im follow-up wieder zu erhöhen (T3: $65,16 \pm 15,61$), während die Werte der Nichtdepressiven relativ unverändert blieben (T1: $50,59 \pm 7,07$; T2: $51,40 \pm 9,74$; T3: $51,17 \pm 9,72$) (vgl. folgende Abb. 4.5).

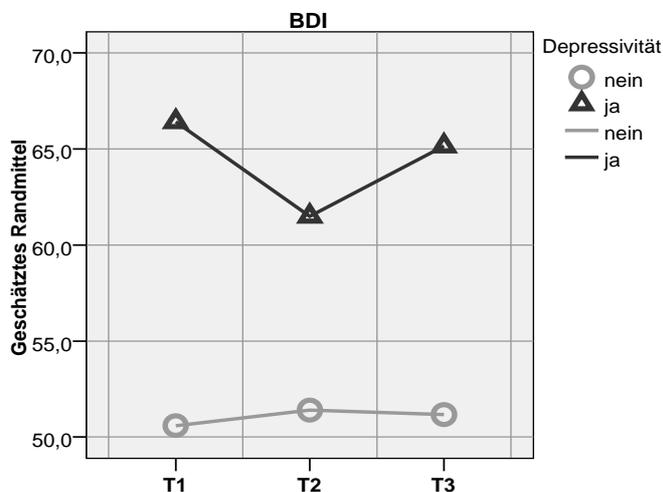


Abb. 4.5: Interaktionseffekt Depressivität * Innersubjektfaktor im BDI.

Ebenso in der PSK zeigte die signifikante Helmert-Interaktion (a) (Tab. 4.13) eine signifikante Differenzierung im mittleren Therapieerfolg zwischen den Gruppen an. Auch hier bestätigte sich erneut die Hypothese, dass hauptsächlich Probanden mit Depressivitätsdiagnose eine gesundheitsförderliche Veränderung im Interventionszeitraum aufweisen (T1: $42,81 \pm 13,29$; T2: $48,86 \pm 12,44$; T3: $43,83 \pm 16,51$), während die restlichen Probanden wenig Änderung erfuhren (T1: $55,42 \pm 8,05$; T2: $54,98 \pm 7,54$; T3: $53,32 \pm 10,00$) (vgl. folgende Abb. 4.6).

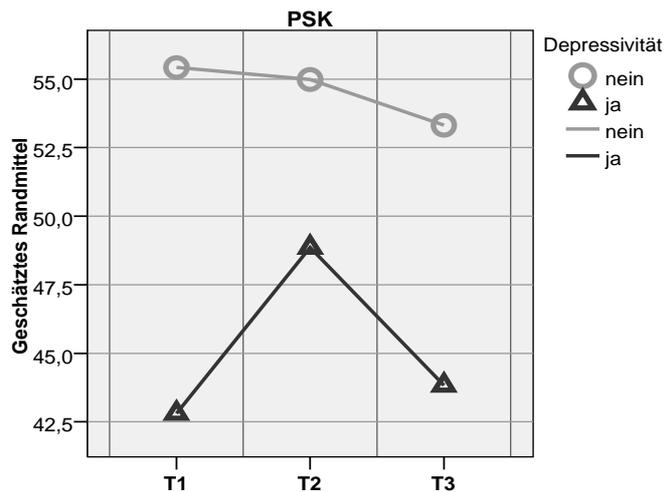


Abb. 4.6: Interaktionseffekt Depressivität * Innersubjektfaktor der PSK.

Im TV-Konsum deutete die Helmert-Interaktion (a) (Tab. 4.13) zwischen T1 und den folgenden Messzeitpunkten nicht auf einen differentiellen Interventionseffekt hin. Im Gegensatz dazu wies der Innersubjektkontrast zwischen T2-T3 ein signifikantes Ergebnis auf (b). Diese Diskrepanz zeigt, dass "Depressivität" keinen differentiellen Interventionseffekt, sondern einen selektiven Effekt im follow-up-Zeitraum bewirkte. Deskriptiv wiesen die depressiven Probanden einen Rückgang ihrer zuvor gewonnenen Reduktion des TV-Konsums im follow-up auf (T1: 236±176min; T2: 177±162min; T3: 226±204min), während sich die Reduktion des Fernsehkonsums bei den restlichen Teilnehmern im follow-up fortsetzte (T1: 194±121min; T2: 180±111min; T3: 170±140min) (vgl. folgende Abb. 4.7).

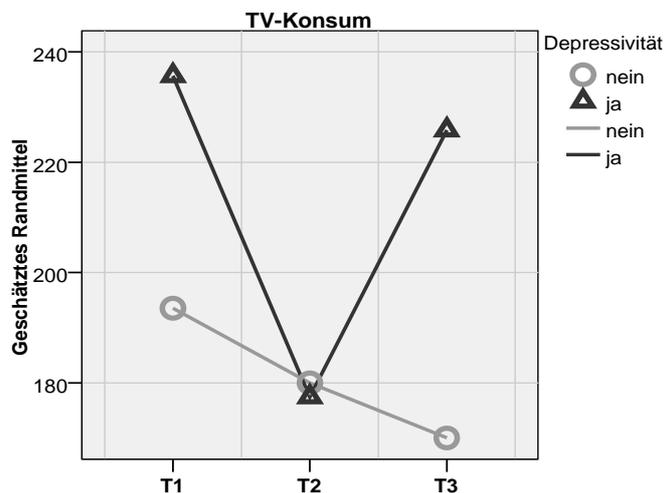


Abb. 4.7: Interaktionseffekt Depressivität * Innersubjektfaktor des TV-Konsums.

Die Hypothesen unter M2 konnten größtenteils bestätigt werden: Personen mit Depressivität profitierten signifikant stärker in der Reduktion von unspezifischen psychischen Beschwerden (GHQ), in der Senkung von Depressivität (BDI) und in der psychischen Einschränkungsminderung der Lebensqualität (PSK) als nichtdepressive Probanden. In der Reduktion des passiven Freizeitverhaltens (TV-Konsum) konnte durch die Depressivitätsdiagnose ein selektiver Rückgang des Interventionserfolges im follow-up belegt werden.

In folgender Tabelle 4.14 werden die Tests der Moderatorhypothesen **M3** zur Prüfung des moderierenden Einfluss des Geschlechts auf den Therapieerfolg gezeigt:

Tab. 4.14: Tests der Moderatorhypothesen **M3**.

		KSK	VAS
Geschlecht	♂	32	17
	♀	38	33
(a) <i>Helmert-Interaktion</i> <i>T1 vs. T2&T3</i>	df1	1	1
	df2	66	46
	F	2,951	0,408
	p	0,091	0,526
(b) <i>follow-up-Interaktion</i> <i>T2 vs. T3</i>	df1	1	1
	df2	66	46
	F	0,244	0,375
	p	0,623	0,543

Die Tests der Moderatorhypothesen zum Geschlechtseffekt waren nicht signifikant (a). In der KSK zeigte sich in der Helmert-Interaktion aber ein Trend, dass Männer und Frauen unterschiedlich von der Intervention profitierten. Deskriptiv offenbarte sich in der KSK ein hypothesenkonformer stärkerer Zuwachs von körperlicher Lebensqualität für die Männer (T1: 46,32±13,20; T2: 48,70±12,68; T3: 49,26±13,34) im Vergleich zu den Frauen (T1: 47,11±12,17; T2: 46,42±11,67; T3: 46,05±12,28) (vgl. folgende Abb. 4.8).

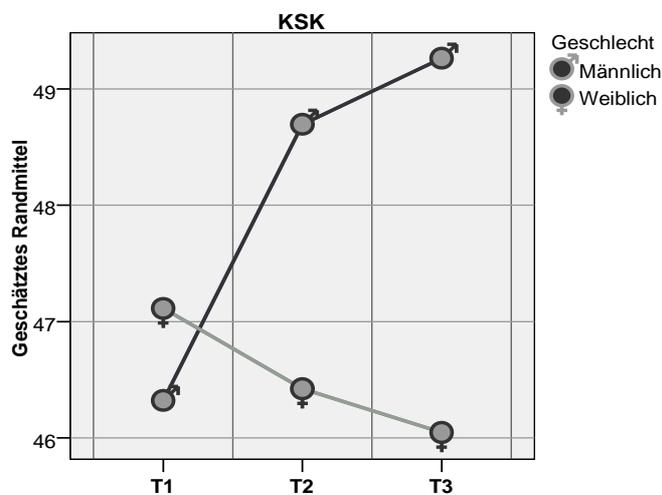


Abb. 4.8: Interaktionseffekt Geschlecht * Innersubjektfaktor der KSK.

Die Geschlechter wiesen ähnliche Ausgangsbedingungen auf und entwickelten sich stetig in verschiedene Richtungen. Der tendenzielle Interaktionseffekt war hauptsächlich dem selektiven Lebensqualitätszuwachs der männlichen Probanden geschuldet, denn die Werte der weiblichen Teilnehmer änderten sich vergleichsweise wenig zwischen den Zeitpunkten.

Zur Aufklärung der sich aufdrängenden Fragestellung, ob zumindest die Männer eine signifikante Veränderung der körperlichen Lebensqualität zwischen den Messzeitpunkten T1-T2 offenbaren würden, wurde ein t-Test für abhängige Stichproben berechnet. Hierbei zeigte sich eine tendenzielle Erhöhung der körperlichen Lebensqualität ($t_{(df=31)}=1,891$; $p=0,068$) für die Männer im Interventionszeitraum.

Die Hypothesen unter M3 konnten nicht bestätigt werden: Männer profitierten nicht signifikant stärker in der Erhöhung der körperlichen Lebensqualität (KSK) oder in der Besserung des körperlichen Gesundheitsempfindens (VAS) als Frauen. In der KSK weisen die Ergebnisse aber als statistischer Trend in die hypothetisch angenommene Richtung.

Folgende Tab. 4.15 zeigt die Tests der Moderatorhypothese **M4** nach dem in Kap. 2.10.2.3 vorgestellten Schema zur Prüfung des Einfluss der Sportfortsetzung vs. -abbruch im follow-up:

Tab. 4.15: Tests der follow-up Moderatorhypothese **M4**.

		GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum
Sportfortsetzung	Nein	27	27	25	30	30	30	28	24	24
	Ja	29	30	31	42	42	39	42	26	28
Interaktion T2-T3 * Sportfortsetzung	df2	53	54	53	69	69	66	67	47	49
	$F_{(df1=1)}$	1,689	0,179	0,359	0,199	2,983	0,718	3,075	0,428	0,220
	p	0,199	0,674	0,552	0,657	0,089	0,400	0,084	0,516	0,641
Einfluss Kontrollvariable T1	$F_{(df1=1)}$	4,885	0,053	0,922	2,966	7,323	0,007	0,979	3,164	11,40
	p	0,031	0,819	0,341	0,090	0,009	0,936	0,326	0,082	0,001

"Gesundheitlich positive Effekte" einer Fortsetzung des Sports nach T2 sollten sich dadurch auszeichnen, dass die abhängigen Variablen eine selektive Verbesserung der Sportfortsetzer und/oder eine selektive Verschlechterung der Sportabbrecher im follow-up-Zeitraum aufzeigen. Wie in der Tabelle ersichtlich, zeigte sich in keinem Messinstrument ein signifikanter Interaktionseffekt des Faktors "Sportfortsetzung" mit dem Messzeitpaar T2-T3. Für die Selbstwirksamkeitserwartung (SWE) und die körperliche Lebensqualität (KSK) deutete sich eine tendenzielle Beeinflussung des follow-up aber in der hypothetischen Richtung an. In diesen erzielte eine Aufrechterhaltung des Sports ein tendenziell gesundheitsförderlicheres Ergebnis, weil sich die Werte der Sportfortsetzer in SWE und KSK deskriptiv verbesserten ([SWE: T2: 50,95±6,70; T3: 51,16±6,97] [KSK: T2: 46,46±7,32; T3: 47,85±8,16]) während gleichzeitig die Sportabbrecher im follow-up deskriptiv eine Verschlechterung in diesen Indikatoren ([SWE: T2: 53,01±6,70; T3: 50,04±6,98] [KSK: T2: 48,44±7,33; T3: 46,60±8,21]) offenbarten.

Die jeweilige Kontrollvariable "T1" zeigte im GHQ, in der SWE und im TV-Konsum signifikante Einflüsse. Da für die Kontrollvariablen keine Hypothesen formuliert waren, verbleibt lediglich anzumerken, dass die Veränderung im follow-up auch signifikant von den ursprünglichen Werten der jeweiligen T1 Variable abhing. Dies ist nicht ungewöhnlich, wenn man sich bei longitudinalen Daten das Phänomen in Erinnerung ruft, dass einem starken Auf- oder Abwärtstrend von T1 nach T2 anschließend eher eine Korrektur von T2 nach T3 in die entgegengesetzte Richtung folgt, als das dies nach relativ unveränderten Werten der Fall ist (populäres Beispiel: der sog. "JoJo-Effekt" in der Gewichtsreduktion).

Hypothese M4 konnte für sämtliche Variablen nicht bestätigt werden. Probanden, die nach T2 den Sport fortsetzten, zeigten in den Gesundheitsindikatoren keine signifikant positiveren Ergebnisse im follow-up als die Sportabbrecher. In der KSK und in der SWE deutete sich der hypothetisierte Gruppenunterschied der follow-up-Entwicklung lediglich als Trend an.

4.4.1 Effektstärken der Moderatoreffekte

In den Hypothesen M1-M3 wurde die moderierende Wirkung von Zwischensubjektfaktoren auf die Interventionseffekte im Helmert-Kontrast der Interaktion jeweils unter dem Punkt **(a)** geprüft (Tab. 4.12 - 4.14). In der Hypothese M4 erfolgte der Test von "Sportfortsetzung" als Interaktionseffekt mit dem follow-up (Tab. 4.15). In folgender Tabelle sind die Effektstärken nebst Power der Tests aufgeführt, die anzeigen, dass der jeweilige Zwischensubjektfaktor mindestens tendenziell ($p < 0,1$) die Änderung der Längsschnittdaten moderierte.

Tab. 4.16: Effektstärke und Power der Zwischensubjektfaktor-Innersubjektfaktor-Interaktionen.

	p	Q	η^2	η^2 Cohen Class.	N	$1-\beta$
GHQ * Depressivität	0,029	3	0,088	Mittel	56	0,596
BDI * Depressivität	0,022	3	0,096	Mittel	57	0,643
PSK * Depressivität	0,040	3	0,079	Mittel	56	0,545
SWE * Interventions-Kohorte	0,003	3	0,122	Mittel	72	0,869
KSK * Geschlecht	0,091	3	0,043	Klein	70	0,395
KSK * Sportfortsetzung	0,084	2	0,044	Klein	70	0,409
SWE * Sportfortsetzung	0,089	2	0,041	Klein	72	0,399

p : Signifikanz des Interaktionseffekts (vgl. **(a)** in Tab. 4.12-4.14); Q: Anzahl der Messungen; η^2 : Interaktions-Effektstärke; $1-\beta$: Power des Interaktionstests; Cohen Class.: Klassifikation des Effekts nach Cohen (1988, 1992).

Die Depressivitätsdiagnose führte im GHQ, BDI und PSK zu differenzierten Interventionseffekten mit jeweils einem "mittleren" Effekt nach Cohen (1988, 1992). Der Effekt der Interventionskohorte in der SWE war ebenfalls von "mittlerer" Größe. Die tendenziell wirksamen Moderatoren Geschlecht und Sportfortsetzung erreichten jeweils einen "kleinen" Effekt auf die Veränderungsrichtung der Messzeitpunkte. Moderatoreffekte mit mindestens mittlerer Effektstärke konnten ihre Wirkung signifikant (p) unter Beweis stellen; kleine Effektstärken verfehlten das Signifikanzkriterium. Die kleinste Power einer signifikanten Interaktion betrug $1-\beta=0,55$ (PSK * Depressivität); eine Power darunter führte nicht mehr zur Signifikanz.

4.5 Veränderungssensitivität der Messinstrumente

Ziel der abschließenden Analysen war eine herzustellende Vergleichbarkeit der Veränderungssensitivität der Messinstrumente für die hiesige Forschungsfrage mit der Vorgehensweise wie im Kap. 2.10.2.4 beschrieben. In den Hypothesentests L1-L4 wurde der Innersubjekt-Haupteffekt jeweils unter dem Punkt **(A)** geprüft (vgl. Tab. 4.9). Mit der Poweranalyse nach Cohen (1988) wurde die Power ($1-\beta$) der jeweiligen H_1 zum Haupteffekt quantifiziert und damit die Veränderungssensitivität bestimmt (vgl. folgende Beispielgrafik in Abb. 4.9).

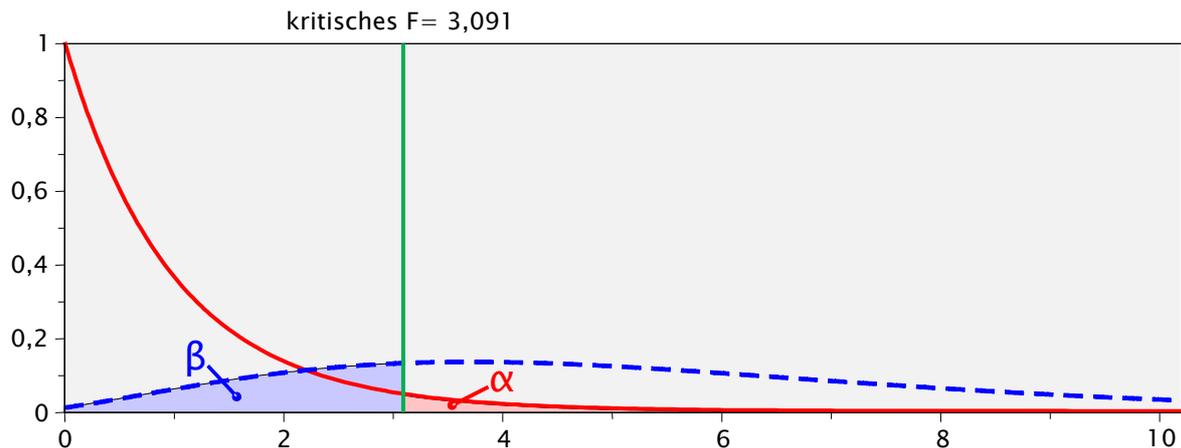


Abb. 4.9: Beispielhafte Darstellung zur Bestimmung des β -Fehlers des Innersubjektfaktor-Haupteffekts der Variable *TV-Konsum*: Die H_0 ist rot, die H_1 blau gestrichelt. Der β -Fehler entspricht der Fläche von H_1 , die beim Unterschreiten des kritischen F-Wertes entsteht. Der α -Fehler ist auf 0,05 festgelegt. Für den Innersubjekt-Haupteffekt der Variable *TV-Konsum* ergibt sich ein β -Fehler=0,26. Damit hat die H_1 eine Power von $1-\beta=0,74$.

Innerhalb der multivariaten Varianzanalyse mit Messwiederholung hängt die β -Fehlergröße des Innersubjektfaktor-Haupteffekts hauptsächlich²³ von folgenden Parametern ab: Gewähltes α -Fehlerniveau, Innersubjektfaktor-Effektstärke f , mittlere Korrelation²⁴ zwischen den Messzeitpunkten $\varnothing r_{(T1,T2,T3)}$ und Stichprobengröße N . Hierbei gilt: Je größer α und je größer f und je höher $\varnothing r_{(T1,T2,T3)}$ und je umfangreicher N , desto kleiner wird β und umso größer wird die Power $1-\beta$. Da $\alpha=0,05$ für sämtliche Tests festgelegt war und die Anzahl der Messungen für den Haupteffekt stets 3 betrug, konnten nur die restlichen Parameter variieren. Die Effektstärke f und die Power der Innersubjektfaktor-Haupteffekte wurden mit der Methode von Faul et al. (2007) nach Rasch et al. (2010) berechnet. Es ergaben sich folgende Werte:

Tab. 4.17: Übersicht der Innersubjektfaktor-Haupteffekte und deren erzielte Power.

	p	Q	K	f	$\varnothing r_{(T1,T2,T3)}$	η^2	η Cohen Class.	N	ε	$1-\beta$
GHQ	0,000	3	4	0,167	0,745	0,148	Groß	56	1	0,972
BDI	0,142	3	4	0,079	0,715	0,037	Klein	57	0,89	0,402
PSK	0,026	3	4	0,153	0,475	0,068	Mittel	56	1	0,676
SOC	0,296	3	2	0,062	0,670	0,017	Klein	72	0,91	0,251
SWE	0,376	3	2	0,049	0,740	0,014	Klein	72	1	0,219
PSW	0,220	3	2	0,056	0,795	0,022	Klein	69	1	0,321
KSK	0,634	3	4	0,047	0,495	0,007	-	70	1	0,123
VAS	0,196	3	4	0,125	0,277	0,035	Klein	50	1	0,342
TV	0,016	3	4	0,128	0,705	0,083	Mittel	52	1	0,740

p : Signifikanz des Innersubjektfaktor-Haupteffektes (vgl. (A) in Tab. 4.9); Q: Anzahl der Messungen; K: Anzahl der Zwischensubjektfaktorgruppen; f : Effektstärke f für unabhängige Stichproben; $\varnothing r_{(T1,T2,T3)}$: mittlere Korrelation zwischen den Messzeitpunkten; η^2 : Effektstärke für abhängige Stichproben; ε : Nicht-Sphäritätskorrektur; $1-\beta$: Power; Cohen Class.: Klassifikation der Effektstärke nach Cohen (1988, 1992).

²³ Des Weiteren wird die Power auch durch die Anzahl der Innersubjektfaktorstufen sowie der Zwischensubjektfaktorgruppen und der Nicht-Sphäritätskorrektur beeinflusst.

²⁴ Da Skalenwerte von Korrelationen nicht äquidistant (und damit nicht intervallskaliert) sind, wurden die Korrelationskoeffizienten jedes Messverfahren für $T1*T2$, $T1*T3$ und $T2*T3$ in Fischers Z-Werte nach den Vorgaben von Glass und Stanley (1970, S. 534) transformiert. Anhand der Fischers Z-Werte erfolgte die Durchschnittsbildung mit anschließender Rücktransformation als durchschnittlicher Korrelationskoeffizient.

Die Effektstärke f in Tab. 4.17 repräsentiert die Effektstärke der Varianzanalyse für Effekte bei unabhängigen Prüfgrößen. Sie wurde berechnet, weil sie zur Hochrechnung der Power (vgl. folgendes Kapitel) bei Variationen von N nach der Methode von Faul et al. (2007) vonnöten ist. Das Wirken der Effektstärke f auf die Power des Tests zum Haupteffekt wird in der ANOVA mit Messwiederholung erheblich von der durchschnittlichen Korrelation der Messzeitwerte beeinflusst (Rasch, et al., 2010, S. 130). Im Gegensatz zur Effektgröße f beinhaltet die Effektstärke η^2 der Tab. 4.17 bereits die Korrelation der Messzeitwerte und stellt deshalb die Effektstärke für abhängige Stichproben dar.

4.5.1 Bestimmung der vergleichbaren Veränderungssensitivität

In der Hochrechnung der Power mit der Methode von Faul et al. (2007) nach den Empfehlungen von Rasch et al. (2010) wurde die Power jedes Innersubjektfaktor-Haupteffektes für ein $N=72$ berechnet. 72 entspricht der maximalen Stichprobengröße, die in dieser Studie für einen Vergleich aller 3 Messzeitpunkte vorhanden war. Diese traf auf den SWE und den SOC zu (vgl. Tab. 4.17). Nach Hochrechnung der Power für die übrigen Messverfahren resultierten folgende Sensitivitäten für den Innersubjektfaktor-Haupteffekt:

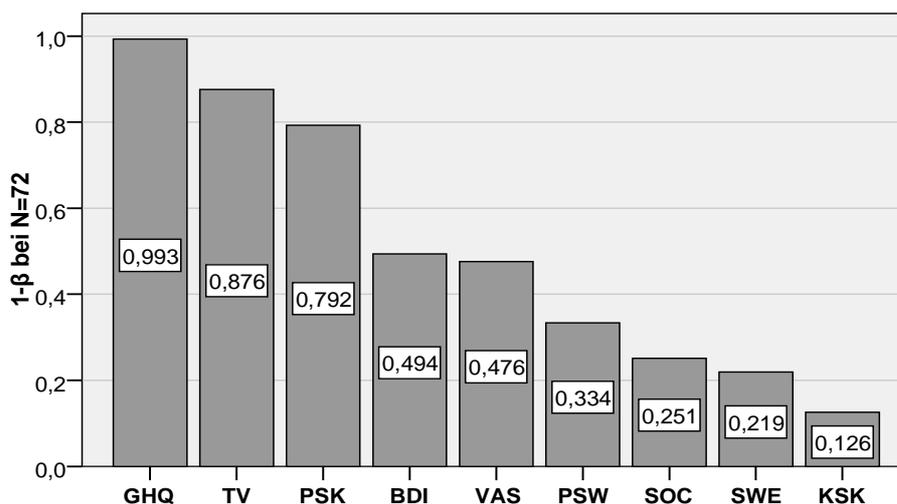


Abb. 4.10: Darstellung der Sensitivitäten (Power) der Innersubjektfaktor-Haupteffekte bei $N=72$.

Mit der Hochrechnung auf das identische N sind die Sensitivitäten der Messverfahren schließlich miteinander vergleichbar und deswegen ist deren Reihenfolge in der Grafik absteigend dargestellt, beginnend mit dem veränderungssensitivsten und endend mit dem unsensitivsten Messinstrument. Bei $N=72$ hätte der Haupteffekt des Innersubjektfaktors des GHQ mit 99% die größte Wahrscheinlichkeit erzielt, Veränderungen zwischen den Messzeitpunkten signifikant werden zu lassen. Die Chance auf Signifikanz des Haupteffekts des TV-Konsums hätte 88% und die des PSK hätte 79% betragen. Im Mittelfeld lagen BDI und VAS etwa gleichauf und die PSW etwas niedriger. Am unsensitivsten erwiesen sich die Messinstrumente KSK, SWE und SOC.

4.5.2 Optimale Stichprobenumfänge zum Erreichen von $1-\beta \geq 0,8$

Nach den Empfehlungen von Cohen (1988) sowie Westermann (2000) sollte das N für den statistischen Nachweis eines Effektes so gewählt werden, dass $1-\beta$ mindestens den Wert 0,8 erreicht. Dieses Kriterium könnte in der hiesigen Studie zu der Schlussfolgerung führen, dass für die Mehrzahl der eingesetzten Messverfahren die Stichprobengröße nicht ausreichte, um statistisch absicherbare Unterschiede zwischen den Erhebungszeitpunkten nachzuweisen. Von daher wurde mit der Methode von Faul et al. (2007) berechnet, wie umfangreich die Stichprobengröße jeweils hätte sein müssen, um eine Power von $1-\beta \geq 0,8$ zu erzielen.

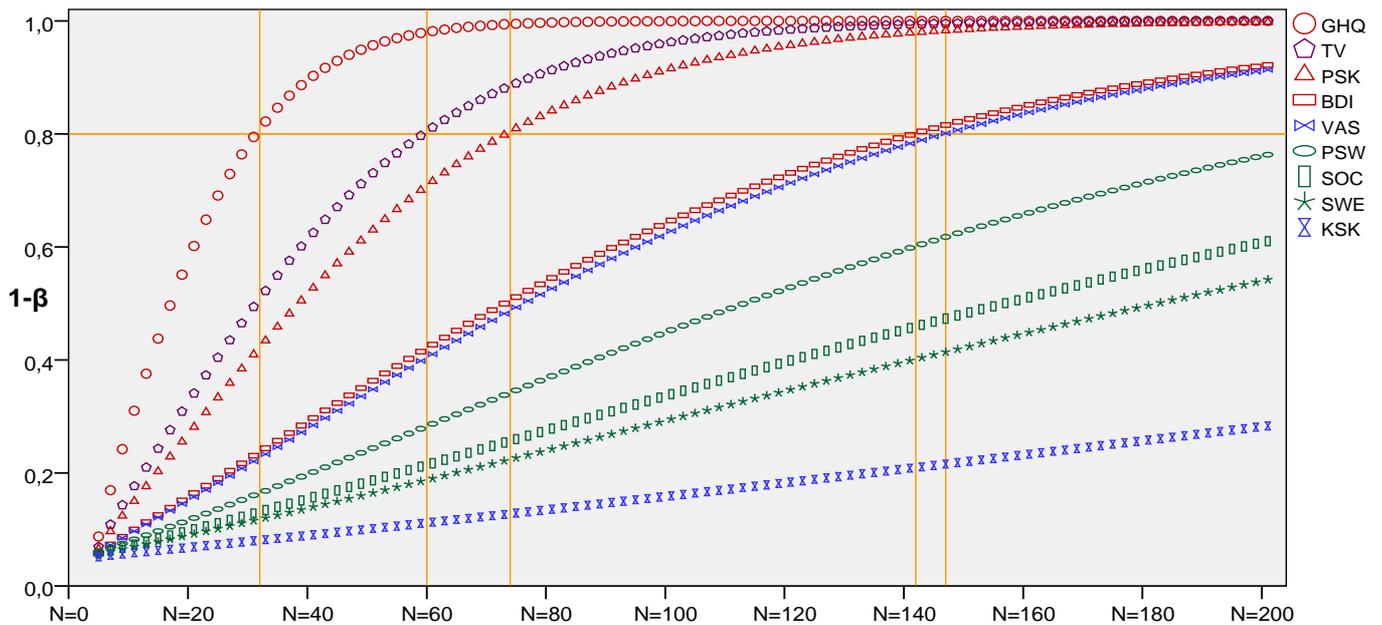


Abb. 4.11: Visualisierung der Bestimmung optimaler Stichprobenumfänge für die Innersubjektfaktor-Haupteffekte der jeweiligen Messinstrumente (siehe Legende), dargestellt von $N=5$ bis $N=201$ in 2er-Schritten. Der Schnittpunkt eines Graphen mit $1-\beta=0,8$ (y-Achse) identifiziert auf der x-Achse den optimalen Stichprobenumfang zum Erreichen des Power-Kriteriums (siehe Tab. 4.18 für konkrete Ergebnisse).

Im Diagramm sind Messgrößen zu psychischen Risikofaktoren rot, die zu psychischen Ressourcen grün und die zum Körper blau dargestellt; TV-Konsum ist lila. Während in der Grafik aus Darstellungsgründen nur die Hochrechnung von $N=5$ bis $N=201$ in 2er Schritten visualisiert ist, wurden die zugrunde liegenden Kalkulationen von $N=5$ in 1er Schritten bis zu einer Stichprobengröße von $N=1000$ hochgerechnet. Hierbei ergaben sich für die Messverfahren folgende optimale Stichprobenumfänge für das Aufdecken eines signifikanten Haupteffekts des Innersubjektfaktors:

Tab. 4.18: Optimale Stichprobenumfänge zum Erreichen von $1-\beta \geq 0,8$ für den Innersubjektfaktor-Haupteffekt.

	GHQ	TV-Konsum	PSK	BDI	VAS	PSW	SOC	SWE	KSK
$1-\beta \geq 0,8$	N=32	N=60	N=74	N=142	N=147	N=219	N=306	N=354	N=740

Der GHQ hätte bereits mit $N=32$ das Powerkriterium von 0,8 erzielt. TV-Konsum würde ab $N=60$ das Niveau erreichen und die PSK wäre ab $N=74$ dem Kriterium begegnet. Wie man sieht, liegen die optimalen Stichprobenumfänge für den TV-Konsum und die PSK höher als das tatsächliche N der Studie, in welcher - trotz suboptimaler Stichprobenumfänge - sowohl

die PSK (N=56) als auch der TV-Konsum (N=52) einen signifikanten Innersubjektfaktor-Haupteffekt erreichten. Durch die in Tab. 4.18 absteigende Ordnung der Messverfahren nach ihrer Power resultieren als optimale Stichprobengrößen zum Erreichen des Powerkriteriums für den BDI, VAS, PSW, SOC, SWE und KSK immer größere Stichprobenumfänge. Die KSK würde laut Hochrechnung sogar N=740 Probanden benötigen, um den Haupteffekt des Innersubjektfaktors mit einer Wahrscheinlichkeit von 0,8 signifikant werden zu lassen. Spätestens für die KSK sollte dann aber die Debatte um die "Signifikanz vs. Relevanz" eines Effektes einsetzen.

4.5.3 Empfehlungen für Replikationsstudie mit N=100

Falls eine Replikationsstudie z.B. mit einem N=100 geplant werden sollte, könnten laut obiger Hochrechnung die Messinstrumente GHQ, TV-Konsum und PSK vorbehaltlos empfohlen werden, da anzunehmen ist, dass diese eine Veränderungssensitivität von $1-\beta \geq 0,8$ bereitstellen würden. Obwohl der BDI das Powerkriterium laut Berechnung nicht erreicht, eignet er sich aber zur Diagnose von Depressivität, welche sich als bedeutender Zwischensubjektfaktor in den Interaktionen mit den Zeiteffekten der Risikofaktoren demaskierte. Von daher empfiehlt es sich, trotzdem den BDI oder ein anderes Maß zum Screening einer depressiven Störung anzuwenden. Zur Veränderungsmessung der subjektiven körperlichen Gesundheit sollte aber dringend empfohlen werden, die KSK durch ein veränderungsempfänglicheres Maß zu ersetzen (detaillierte Beschwerdelisten könnten hier vielleicht sensitiver sein). Für die ressourcenorientierten Instrumente SOC, SWE und PSW wäre es sehr von der Fragestellung abhängig, ob diese künftig eingesetzt werden sollten, da sie sich als so unsensitiv erwiesen, dass sie bei einem N=100 nicht empfohlen werden können.

5 Diskussion

Die meisten der hier untersuchten Langzeitarbeitslosen waren mehr als 5 Jahre erwerbslos. Der Bildungs- und Ausbildungsstand in Form des Realschulabschluss (POS) bzw. in Form der abgeschlossenen Berufsausbildung dominierte die Stichprobe. Völlig Ungelernte kamen in der Stichprobe etwas seltener vor als Personen mit einem Hochschulstudium. Der hohe Anteil von Probanden ohne feste Lebensgemeinschaft reflektiert möglicherweise die Ergebnisse von Klein und Strasser (1997), nachdem die Scheidungswahrscheinlichkeit von langfristig Erwerbslosen um ein Vielfaches höher als in der Allgemeinbevölkerung liegt. Die Probanden gaben mehrheitlich an, mangelhafte finanzielle Mittel zur Verfügung zu haben; fast alle bezogen die Transferleistung ALG-II. Finanzielle Spielräume waren stark eingeschränkt, die Ersparnisse größtenteils aufgebraucht und jeder Vierte berichtete über Probleme mit Schulden.

5.1 Querschnittsergebnisse

Der FPI wies durchschnittlich kein auffälliges Persönlichkeitsprofil für die Stichprobe auf. Die stärkste Abweichung vom Normstichprobenmittelwert zeigte das für die Studie als relevant definierte Merkmal "Lebenszufriedenheit" (vgl. Abb. 7.2). Dessen Divergenz vom normierten Mittel war vor allem der signifikant niedrigeren Lebenszufriedenheit der Männer geschuldet (vgl. Kap. 3.2.2). Implizit gegen die Selektionshypothese spricht das Ergebnis zum Faktor "Emotionalität" (Neurotizismus). Eine hohe Emotionalität wird als Neigung zur Depressivität und anderen psychopathologischen Symptomen aufgefasst (z.B. Hammarstrom & Janlert, 1997). Das Emotionalitätsmerkmal lag mit Stanine=5,2 aber fast zentral im Normbereich (vgl. Abb. 7.2), so dass sich aus diesem z.B. keine erhöhte Depressivitäts-Prädisposition für die Stichprobe ableiten lässt, obwohl mit dem klinischen Screening-Instrument BDI ein ca. doppelt so hoher Anteil für unter Depressivität leidende Probanden im Vergleich zur Allgemeinbevölkerung objektiviert wurde (vgl. Tab. 3.2). Der Anteil von Probanden mit depressiver Symptomatik stimmt in etwa mit Paul et al. (2006) überein, die beim Vergleich von Erwerbslosen mit Berufstätigen ein etwas mehr als doppelt so hohes Auftreten von Depressivität bei den Arbeitslosen feststellten.

Der anhand repräsentativer Stichproben objektivierter Anteil psychisch problematischer Fälle lag für die Fragebögen, welche psychische Risikofaktoren abbildeten, durchschnittlich etwa doppelt so hoch wie zu erwarten gewesen wäre (vgl. Tab. 3.2). Im Vergleich zur Allgemeinbevölkerung beinhaltete die Stichprobe deutlich mehr Fälle von eingeschränkter Lebenszufriedenheit (LEB) (+128%), einen gravierend höheren Anteil von Personen mit unspezifischen psychischen Beschwerden (GHQ) (+143%) und etwa doppelt so viele Fälle von Depressivität (BDI) (+95%). Der Interventionsbedarf für eine Reduktion von psychischen Beschwerden musste von daher als hoch eingeschätzt werden.

Im Vergleich zur Allgemeinbevölkerung waren für eine eingeschränkte Selbstwirksamkeitserwartung (SWE) (+2%) und für einen mangelnden Kohärenzsinn (SOC) (+13%) relativ wenige Abweichungen feststellbar, welches einen besonderen Interventionsbedarf für den Aufbau von Ressourcen der psychischen Gesundheit durchschnittlich negierte.

Zusammengefasst zeigte die Ermittlung des Interventionsbedarfs einen wesentlich höheren Anteil für Personen mit psychischen Risikofaktoren im Vergleich zum Anteil von Probanden mit unterdurchschnittlichen psychischen Ressourcen. Gleichzeitig muss festgehalten werden, dass weder eine erhöhte Depressivitäts-Prädisposition (in Form einer erhöhten Emotionalität) noch ein eingeschränktes Kohärenzgefühl noch eine mangelnde Selbstwirksamkeitserwartung die hohen psychischen Beschwerden der Probanden zufriedenstellend erklären können. In Analogie zur Gesundheitsdefinition der WHO muss gelten, dass das psychische Wohlergehen mehr als die Abwesenheit von psychischen Beschwerden (Risikofaktoren) bedeutet. Hieraus sollte im Umkehrschluss resultieren: *Psychische Beschwerden sind mehr als die Abwesenheit von Ressourcen der psychischen Gesundheit*. Es kann vorausgesetzt werden, dass mangelnde psychische Ressourcen sowie eine hohe Emotionalität als mögliche *interne Ursachen* von psychischen Beschwerden fungieren. Können jedoch weder in der Emotionalität noch in den Ressourcen substantielle Abweichungen von der Allgemeinbevölkerung festgestellt werden, sind folglich die auslösenden Bedingungen der psychischen Beschwerden als *externe Ursachen* zu suchen. Von den möglichen externen Auslösebedingungen sollte die Erwerbslosigkeit samt ihrer finanziellen sowie soziopsychologischen Begleitschäden (z.B. Frustration der Selbstbestimmung, vermehrtes Scheitern von Beziehungen, Unplanbarkeit der Zukunft, gesellschaftliches Stigma etc.) als wahrscheinlichste Kandidatin infrage kommen.

Innerhalb einer explorativen Faktorenanalyse (vgl. Abb. 7.8; Tab. 7.3) konnte gezeigt werden, dass die VAS und die KSK gemeinsam das körperbezogene Empfinden der Gesundheit auf einem Faktor reflektierten und sich damit prinzipiell von dem anderen Faktor, der die Variablen zur Erfassung der psychischen Gesundheit repräsentierte, unterschieden. Dass sich die Probanden in Relation zu den Vergleichsstichproben häufiger in ihrer körperlichen Lebensqualität beeinträchtigt fühlten (KSK) (+69%) und öfter ein vermindertes Gesundheitsgefühl aufwiesen (VAS) (+65%), muss als erhöhter Bedarf an medizinischen Interventionen für die Zielgruppe interpretiert werden. Die Stichprobe war demnach nicht nur durch psychische Beschwerden beansprucht, sondern fühlte sich auch durch körperliche Gesundheitsprobleme deutlich stärker beeinträchtigt als die Werte der Vergleichsstudien erwarten ließen.

Auch wenn durch fast alle psychischen Risikofaktor-Fragebögen evident wurde, dass die Stichprobe im Mittel beträchtliche seelische Probleme hatte, fühlten sich die Probanden aber nicht häufiger durch ihr psychisches Funktionsempfinden in ihrer Lebensqualität (PSK) eingeschränkt als die Normen des SF-12 nahe legten. Dies traf signifikant vor allem auf die Männer zu (vgl. Kap. 3.2.2) und ist erklärungsbedürftig.

Ein Argument könnte sein, dass die Hürde zur Bejahung der PSK-Items zu hoch war (Itemschwierigkeit): In diesen wird z.B. explizit gefragt, ob man aufgrund *seelischer Probleme* weniger geschafft hätte oder ob man aufgrund *seelischer Probleme* nicht so sorgfältig arbeiten könne oder ob man durch seine *seelischen Probleme* nicht dazu in der Lage wäre, seine sozialen Kontakte zu pflegen. Eine zumindest tendenzielle Bejahung dieser Fragen setzt natürlich voraus, dass der Proband bereits für sich akzeptiert hat, psychische Probleme zu haben. Diese voraussetzende Eigen-Pathologisierung für die Bejahung der Items könnte evtl. verstärkt für die Männer eine Hürde dargestellt haben.

Eine andere Erklärungsmöglichkeit wäre, dass sich im Ergebnis der PSK die verminderte Anforderung des Alltags von Arbeitslosen widerspiegelt, die zudem von einem Geschlechtseffekt moderiert sein könnte: Weil Langzeitarbeitslose nicht dem Druck einer zu erfüllenden Erwerbstätigkeit ausgesetzt sind, fühlen sie sich in ihrem psychischen Funktions- und Rollenverständnis evtl. weniger deutlich eingeschränkt als man es aufgrund ihrer hohen Beschwerden erwartet hätte. Da zudem für Männer im Durchschnitt eine stärkere Identifikation mit der Erwerbsarbeit vorausgesetzt wird (z.B. Jahoda, 1983), könnte die Abwesenheit von täglich zu erfüllenden Arbeitspflichten von den männlichen Probanden als besonders niedrige Anforderung ihres Alltags empfunden worden sein. Dies würde erklären, warum sich erwerbslose Männer durch ihre Psyche weniger in ihrem Funktions- und Rollenverständnis eingeschränkt fühlen als erwerbslose Frauen.

Einfluss von Antworttendenz auf die Erfassung der Gesundheit

In der Stichprobe konnte der ermittelte Interventionsbedarf implizit gegen mögliche Simulations- bzw. Aggravationstendenzen abgesichert werden, denn das Merkmal "Offenheit" des FPI legte klare Zusammenhänge einer sozial erwünschten Antworttendenz mit den meisten Variablen zur psychischen Gesundheit offen. Je verschlossener und damit sozial erwünschter die Probanden antworteten, desto eher neigten sie zu einer verkleinerten Darstellung von psychischen Problemen. Da "Offenheit" mit Stanine=4,3 etwas unterhalb der erwarteten (Stanine=5,0) lag (vgl. Abb. 7.2), ist zu schließen, dass der Interventionsbedarf aufgrund von sozial erwünschten Antworttendenzen eher unter- statt überschätzt wurde. Auf jeden Fall aber lässt sich eine übertriebene Ermittlung des Interventionsbedarfs aufgrund von "unwahrscheinlich gemäßen" Beantwortungen der Fragebögen durchschnittlich ausschließen.

Geschlechtseffekt auf die psychische Gesundheit

Die Tests, ob Männer im Querschnitt stärker unter den Folgen der Arbeitslosigkeit leiden als Frauen, konnten nur für eine signifikant niedrigere Lebenszufriedenheit (LEB) der Männer bestätigt werden. Auf der anderen Seite fühlten die weiblichen Versuchsteilnehmer sich häufiger psychisch in ihrem Funktionsverständnis (PSK) eingeschränkt als die Männer (mögliche Gründe dazu wurden im vorletzten Abschnitt diskutiert). Auch der Fragebogen zum psychosozialen Wohlbefinden (PSW) zeigte niedrigere Werte für die weiblichen Probanden. Mangels geschlechtsadjustierter Normen könnte es sich beim letztgenannten Instrument aber auch um einen gewöhnlichen Geschlechtseffekt handeln. Eine mögliche Interpretation wäre, dass die beiden letztgenannten Effekte zu Ungunsten der Frauen die signifikant längere Arbeitslosigkeitsdauer der weiblichen Versuchsteilnehmer (vgl. Anfang Kap. 3) widerspiegeln. Insgesamt konnten zwischen Männern und Frauen nur wenige Unterschiede entdeckt werden, die zudem in verschiedene Richtungen wiesen, so dass keine Aussage darüber zulässig ist, welches Geschlecht insgesamt psychisch stärker litt. Deshalb sollte der hiesige Befund als Bestätigung der Ergebnisse von Paul und Moser (2009a) eingeordnet werden, die innerhalb der Arbeitslosengruppe auch nur geringe Unterschiede zwischen Männern und Frauen fanden. Murphy und Athanasou (1999) konnten in ihrer Meta-Querschnittsanalyse gar keine Geschlechtseffekte objektivieren.

Geschlechtsspezifischer Beziehungsstatus

Dass der Beziehungsstatus einen signifikanten Einfluss auf die psychische Gesundheit von älteren Langzeiterwerbslosen haben könnte, manifestierte sich hingegen äußerst geschlechtsspezifisch. Der Beziehungsstatus führte bei den Männern zu keinerlei signifikanten Befindensunterschieden und auch die Stärke der männerspezifischen Beziehungseffekte erreichte nach den Konventionen von Cohen (1988, 1992) im Höchstfall eine "kleine" Effektstärke. Demgegenüber zeigten Frauen in Beziehung gegenüber denen ohne Partnerschaft signifikante Befindensvorteile in 5 (von 7) Indikatoren zur psychischen Gesundheit mit jeweils mindestens einem "mittleren" Effekt. Die Lebenszufriedenheit, der Kohärenzsinn und unspezifische psychische Beschwerden offenbarten innerhalb der Frauen die höchsten Effekte zwischen den Kategorien des Beziehungsstatus.

Inmitten der Erwerbslosigkeit als (externe) Belastung und der psychischen Gesundheit als Beanspruchungsebene fungiert für die weiblichen Erwerbslosen der Beziehungsstatus demnach als relevanter Moderator zwischen der Belastungs- und Beanspruchungsdimension. Von daher sind - je nach Präferenz - zwei Aussagen zulässig: 1.) Eine Langzeiterwerbslosigkeit trifft Frauen ohne eine partnerschaftliche Beziehung psychisch besonders schwerwiegend oder 2.) Für Frauen mit Beziehung führt eine Langzeitarbeitslosigkeit deutlich weniger stark zu psychischen Problemen.

Warum sich die frauenspezifischen Beziehungseffekte so deutlich demaskierten, konnte mit möglichen Unterschieden in der finanziellen Autonomie nicht erklärt werden, da sich keine signifikanten Unterschiede zwischen Frauen mit und ohne Beziehung in der Bewertung des finanziellen Auskommens finden ließen.

Zur Erklärung von möglichen Befindens-Unterschieden zwischen den Geschlechtern wurde in der psychologischen Erwerbslosigkeitsforschung immer wieder das sog. "Alternativrollen-Argument" (z.B. Shamir, 1985; Paul et al., 2006; Paul & Moser, 2009b) aufgegriffen. Dieses Argument hat den Vorteil, dass es die weiblichen Befindensdifferenzen, die durch den Status der Beziehung differenzierbar werden, recht plausibel erklären kann, denn die alternative Rolle als Hausfrau bietet sich nur für Frauen in Beziehungen an, während sie für alleinstehende Frauen nicht verfügbar ist. Zudem offenbarte die soziodemographische Analyse (Anfang Kap. 3), dass vor allem in partnerschaftlichen Beziehungen noch Kinder im Haushalt anzutreffen waren. Es ist annehmbar, dass die Übernahme eines sich anbietenden alternativen Rollenverständnisses für erwerbslose Frauen erleichtert wird, wenn dieses auch auf die Betreuungsfunktion von Kindern erweitert werden kann. Sich im Laufe der Zeit nicht mehr über die Erwerbslosigkeit definieren zu müssen, sondern die Identifikation als Hausfrau zur Verfügung zu haben, könnte die innere Bindung an die (meist vor Jahren) verlorene Erwerbsarbeit deutlich schmälern und damit zu weniger psychischem Leid beitragen. Zudem lässt sich mutmaßen, dass für über 50-jährige Menschen ein eher traditionelles Rollenverständnis zutrifft, welches für ältere weibliche Erwerbslose zu einer besseren sozialen Akzeptanz der Alternativrolle in ihrer Peer-Gruppe sorgen könnte.

Vor dem Hintergrund des Alternativrollen-Arguments lassen sich zudem die fehlenden männerspezifischen Unterschiede des Beziehungsstatus auf die psychische Gesundheit erklären, da für Männer, insbesondere für die mit traditionellem Rollenverständnis, eine solch

gesellschaftlich legitimierte Alternativrolle nicht verfügbar ist. Zudem wurde für erwerbslose Männer argumentiert, dass diese einem höheren gesellschaftlichen Stigma (z.B. Jahoda, 1983) sowie einer vermehrten "sozialen Unterminierung" durch den Lebenspartner ausgesetzt sind (Vinokur & Van Ryn, 1993), wenn sie innerhalb ihrer Beziehung nicht die primär ihnen zugedachte Rolle als Versorger erfüllen. Diese sehr männerspezifischen Effekte könnten die Wirkung einer normalerweise als positiv erlebten Lebenspartnerschaft so deutlich schmälern, dass erwerbslose Männer in Beziehungen sich nicht signifikant besser fühlen als die ohne Partnerschaft.

Für das ins Feld geführte Argument des "traditionellen Rollenverständnis" wäre es interessant, ob sich die geschlechtsspezifischen Befindensunterschiede in Abhängigkeit vom Beziehungsstatus auch in einer Stichprobe mit entweder wesentlich jüngeren Langzeiterwerbslosen (z.B. 30-Jährige) oder in 20 Jahren mit den künftigen 50-jährigen langfristig Arbeitslosen replizieren ließen. Wäre dies nicht der Fall, könnten die hier gefunden Effekte einem traditionellen Rollenverständnis zugeschrieben werden, welches als Geburts-Kohorteneffekt interpretierbar wäre.

5.2 Längsschnittergebnisse

Die Drop-out Analysen ergaben, dass von T1 ausgehend 79% an T2 und noch 62% an T3 teilnahmen. Die Probandenzahl von T3 setzt sich allerdings zusammen aus Probanden, die tatsächlich zur Untersuchung in Rostock anwesend waren (50%) und denen, die nicht mehr anreisen wollten, aber sich dazu bereit erklärten, postalisch die Fragebögen zu beantworten (12%). Die Drop-out-Quoten waren sicherlich davon beeinflusst, dass die Studienteilnahme für die Probanden eine neunmonatige Dauer umfasste und ein erheblicher zeitlicher Aufwand in Kauf genommen werden musste, um an den Untersuchungen teilzunehmen. So betrug die An- und Abreise von Schwerin nach Rostock mindestens 2 Stunden und die Untersuchungsdauer im IPM war 4-stündig.

Zwischen T2-Teilnehmern und Nichtteilnehmern konnten keine signifikanten Abweichungen in den Daten der Erstuntersuchung ermittelt werden. Teilnehmer an allen 3 Untersuchungsterminen zeigten jedoch hochsignifikant weniger Fernsehkonsum und ein höheres psychosoziales Wohlbefinden (PSW) zu T1, wenn sie mit denen verglichen wurden, die nicht an allen Untersuchungsterminen teilnahmen.

Dass es zwischen Studienkomplettierern und -abbrechern außer in der PSW zu sonst keinen signifikanten Unterschieden in den gesundheitsbezogenen körperlichen und psychischen Variablen kam, spricht für eine Generalisierbarkeit der Ergebnisse der Längsschnittauswertung. So konnte nicht festgestellt werden, dass selektiv Probanden mit ursprünglich stark erhöhtem gesundheitlichen Beschwerdedruck den Folgeuntersuchungen fernblieben.

5.2.1 Ergebnisse der Sportfortsetzeranalyse

In der hiesigen Studie konnte im Gegensatz zu denen im Kap. 2.9.1 vorgestellten Publikationen von McAuley et al. (1993), Rhodes et al. (1999) sowie Resnick (2001) kein prädiktiver Nutzen der SWE für die Bestimmung der Aufrechterhaltungswahrscheinlichkeit des Sportes im follow-up festgestellt werden, weder in der ursprünglichen Messung noch in den Verände-

rungsdaten dieses Messinstrumentes. Selbst auf Itemebene des SWE fanden sich keine Unterschiede zwischen Sportfortsetzern und -abbrechern. Insofern ist für ältere Langzeiterwerbslose die Selbstwirksamkeitserwartung als bedeutender Faktor für die Aufrechterhaltung eines begonnenen Sports mit den hiesigen Ergebnissen nicht belegbar.

Im Gegensatz zur Studie von Williams und Lord (1995) war der BDI als Gesamtest nicht trennscharf, aber immerhin konnten mit 3 seiner Veränderungs-Items signifikant stärkere Reduktionen für die Sportfortsetzer im Vergleich mit den -abbrechern nachgewiesen werden. Von daher deutet sich an, dass sich reduzierende depressivitätsassoziierte Symptome im Interventionszeitraum eine positive Rolle für die Aufrechterhaltung des begonnenen Sportes spielten. Nicht vorhandene Trennschärfen für die Gesamtestergebnisse trafen auf sämtliche psychometrische Instrumente zu. Neben dem Fernsehkonsum ließen sich ausschließlich mit einzelnen Items und Veränderungs-Items der Testverfahren BDI, SF12 und SOC konvergierende Logit-Modelle zur Vorhersage des Sportabbruchrisikos konstruieren.

Vor allem die Rolle des Fernsehkonsums zu T1 zeigte in der Sportfortsetzeranalyse einen bedeutenden Einfluss auf die Wahrscheinlichkeit, dass Probanden nach der Interventionsmaßnahme den Sport abbrachen. Ein ursprünglich hoher TV-Konsum erhöhte sowohl im *T1-Modell* (Kap. 4.2.1) als auch im *kombinierten Modell* (Kap. 4.2.2) die korrekte Vorhersage des Abbruchrisikos.

Das *kombinierte Modell*, welches aus trennscharfen Veränderungs-Items (T2 minus T1) und den signifikanten Items der Erstuntersuchung zusammengesetzt war, ist dem *T1-Modell*, welches nur die Variablen der ersten Messung enthält, vorzuziehen, da es eine höhere Trefferquote (75% vs. 69,5%) sowie eine größere Varianzaufklärung ($R^2=0,45$ vs. $R^2=0,19$) bereitstellte. Des Weiteren erzeugte das *kombinierte Modell* eine doppelt so hohe Odds-Ratio für das Risiko des Sportabbruchs wie das *T1-Modell* (OR=9,3 vs. OR=4,4). Die Kombination der T1-Variablen mit den Veränderungsvariablen stellte für die Risikovorhersage demnach einen deutlichen Mehrwert dar.

Da im *kombinierten Modell* neben dem ursprünglichen Fernsehkonsum die signifikanten psychologischen Koeffizienten sämtlich auf Veränderungs-Items fußten, kann konstatiert werden, dass die Entscheidung der älteren Langzeitarbeitslosen, den Sport fortzusetzen oder nicht, wesentlich davon abhing, in welche Richtung sich der Interventionserfolg für sie entwickelte. Die Zusammenhangsreihenfolge der Regressionsgleichung lautete inhaltlich:

→ Je weniger sich der Teilnehmende nach dem Training entspannter fühlte und je höher sein ursprünglicher Fernsehkonsum war und je weniger er Gefühle eigener Enttäuschung im Laufe der Intervention abbauen konnte, desto größer war das Sportabbruchrisiko im follow-up.

Dieses Modell gibt einen Hinweis darauf, welche Effekte durch eine ähnliche Intervention künftig intensiviert werden sollten. Die Veränderungs-Items legen nahe, dass zur Erhöhung der Aufrechterhaltungswahrscheinlichkeit sowohl eine *hinzugewonnene Entspannung* als auch eine *Reduktion von Gefühlen eigener Enttäuschung* verstärkt werden sollten. Ein zusätzliches Üben von Entspannungsverfahren und kognitiv-behaviorale Techniken, welche ältere Langzeitarbeitslose darin anregen, Erfolgserlebnisse durch das Training besser wahr-

zunehmen bzw. stärker für sich selbst anzuerkennen, könnten für eine nachhaltigere Annahme eines aktiven Lebensstils hilfreich sein.

Aus den Koeffizienten des *T1-Modells* hätte man sicherlich die Schlussfolgerung gezogen, primär den Fernsehkonsum der Teilnehmer zu senken. Wie anhand der Suche nach Prädiktoren für das Veränderungsmodell ermittelt wurde, unterschieden sich die Sportfortsetzer jedoch nicht von den -abbrechern in der erzielten Reduktion des TV-Konsums. Nichtsdestoweniger ist die Variable "Fernsehkonsum zu T1" offensichtlich so wichtig, dass sie selbst im kombinierten Modell als Determinante für die Aufrechterhaltungswahrscheinlichkeit des Sportes hochsignifikant an zweiter Stelle steht. Das *T1-Modell* bietet trotz kleinerer Odds-Ratio und Varianzaufklärung den Vorteil, dass die dazu notwendigen Variablen einfacher zu erheben sind und es für die Risikovorhersage auf eine größere Fallzahl zurückgreifen kann. Zudem gibt das *T1-Modell* Hinweise darauf, womit das Risiko des Sportabbruchs von vorneherein determiniert erscheint und dies war vor allem für den hohen Fernsehkonsum belegbar.

Für den TV-Konsum kann somit geschlossen werden, dass eine hohe Ausprägung ältere, langfristig Erwerbslose gewissermaßen daran "hindert", ein aktiveres Freizeitverhalten nachhaltig anzunehmen. Gleichzeitig kann geschlussfolgert werden, dass in dieser Studie die Langzeiterwerbslosen mit einem besonders stark ausgeprägten passiven Lebensstil (also die, die man hauptsächlich erreichen wollte) nicht so gut für die Fortsetzung des Sportes aktiviert werden konnten wie die restlichen Probanden. Wie die Drop-out-Analyse ergab, wirkte der Fernsehkonsum aber nicht nur der Fortsetzung des Sportes entgegen, sondern "hinderte" die Probanden auch an der Komplettierung der Studie. Insofern unterstreicht die Drop-out-Analyse zusätzlich die Wichtigkeit der Erhebung des Fernsehnutzungsverhaltens für künftige Interventionsstudien mit älteren Langzeiterwerbslosen.

5.2.2 Interventionsergebnisse der Längsschnittanalysen

Die Hypothesen unter **L1**, welche die Erwartung aufstellten, dass die Intervention zu einer Reduktion von Risikofaktoren der psychischen Gesundheit führen wird, konnte für sämtliche Konstrukte bestätigt werden. Es zeigte sich eine hochsignifikante Abnahme von allgemeinen psychischen Beschwerden (GHQ) mit einem "großen" Effekt. Daneben führte die Intervention zu einer signifikanten psychischen Einschränkungsminderung der Lebensqualität (PSK) mit einem "mittleren" Effekt. Die signifikante Reduktion von Depressivität (BDI) indizierte ebenso einen "mittleren" Effekt (Klassifikationen nach Cohen, 1988, 1992; vgl. Tab. 4.10 für konkrete Ergebnisse). Im Gegensatz dazu waren die Hypothesen von **L2**, in denen erwartet wurde, dass die Intervention sich erhöhend auf die Ressourcen der psychischen Gesundheit auswirkt, für sämtliche Konstrukte nicht belegbar. Weder die Selbstwirksamkeitserwartung (SWE) noch der Kohärenzsinn (SOC) oder das psychosoziale Wohlbefinden (PSW) konnten signifikant aufgebaut werden. Die Hypothesen von **L3**, welche die Erhöhung der körperlich wahrgenommenen Gesundheit prognostizierten, konnten ebenso nicht bestätigt werden. Weder das subjektive Gesundheitsempfinden (VAS) noch die körperliche gesundheitsbezogene Lebensqualität (KSK) erfuhren eine signifikante Steigerung im Trainingszeitraum. Die Hypothese **L4**, die eine Abnahme des passiven Lebensstils anhand der Reduzierung des

Fernsehkonzums vorhersagte, konnte hochsignifikant bestätigt werden. Sie erreichte einen "mittleren" Interventionseffekt und lag äußerst knapp unter der Schwelle eines "großen" Effektes.

Vergleichende Einordnung der Interventionseffekte

Für die Einordnung der Interventionseffekte des hiesigen Programms "Fit50+" existieren keine Vergleichsstudien, die ältere Langzeiterwerbslose körperlich trainierten und deren Interventionseffekte mit psychometrischen Maßen erhoben wurden. Im Publikationsbereich der Präventivmedizin, Sportmedizin und der medizinischen Altersforschung existieren zwar Studien zum Ausdauertraining für ältere Probanden, allerdings messen diese hauptsächlich körperliche Variablen; psychologische Maße kommen wenn, dann nur am Rande des wissenschaftlichen Interesses vor. Allerdings sollte die hiesige Studie nicht die Frage beantworten, ob sich ein Ausdauertraining auch positiv auf das gesundheitliche Wohlergehen von älteren Langzeiterwerbslosen auswirkt, sondern ob das Wohlergehen älterer, langfristiger Erwerbsloser auch mit einer primär körperlichen Gesundheitsförderung - statt z.B. einer primär sozialpädagogischen Intervention - gesteigert werden kann. Insofern muss sich die hiesige Studie mit Interventionen für Erwerbslose messen lassen, welche die Aufgabe hatten, die psychosoziale Gesundheit zu fördern.

Vergleich mit der Interventionsstudie "AktivA"

Für einen direkten Vergleich mit "Fit50+" erscheint die Studie "AktivA" von Rothländer et al. (2012) bzw. Rothländer und Richter (2009) am besten geeignet. Diese Studie wurde ebenfalls in Ostdeutschland und in einer Region mit vergleichbar schwachem Arbeitsmarkt (Sachsen) durchgeführt wie in Mecklenburg-Vorpommern. Des Weiteren untersuchten Rothländer et al. hauptsächlich ältere Erwerbslose, die in der Mehrzahl langzeiterwerbslos waren, welches die Zielgruppen vergleichbar macht. Schließlich wurde "AktivA" einigermaßen zeitgleich zur vorliegenden Studie "Fit50+" durchgeführt, so dass nicht mit Periodeneffekten zu rechnen ist, die einen Vergleich erschweren könnten. "AktivA" war eine kognitiv-behaviorale Intervention für Erwerbslose (vgl. Kap. 1.6 für eine genauere Beschreibung).

Rothländer et al. rekrutierten ihre Versuchsteilnehmer über zwei Zugangswege: Einerseits über regionale Beschäftigungsträger (verpflichtendes Angebot²⁵), andererseits über Zeitungsanzeigen und Aushänge (freies Angebot). Die Interventionsgruppen (N=236) wurden über Kontrollgruppen (N=182), deren Teilnehmer meist eine herkömmliche Qualifizierungsmaßnahme erhielten, abgesichert. Als Instrument wurde der Fragebogen für körperliche, psychische und soziale Symptome (KÖPS, Manz, 1998) eingesetzt. Rothländer et al. berichten für beide Interventionsgruppen Mittelwerte und Standardabweichungen zu den Messzeitpunkten, unter anderem für die Skala "KÖPS-psychisch", welche das Ausmaß von psychischen Beschwerden reflektiert (Rothländer et al., 2012, S. 123, Tab. 6: Daten von KÖPS-psychisch zu T1 u. T2).

Aus den deskriptiven Werten der Studie lassen sich Interventionseffektstärken als Cohen's *d* schätzen (Quelle: eigene Berechnung). Da zur Berechnung von *d* für abhängige Stichproben die Korrelation zwischen den Messzeitpunkten T1 und T2 vonnöten ist, diese

²⁵ Qualifizierungsmaßnahme nach dem SGB II

aber nicht publiziert ist (was üblich ist), kam die in dieser Dissertation berechnete mittlere Korrelationen für psychische Beschwerden²⁶ zw. T1 und T2 als Korrelationsschätzer für die Berechnung der "AktivA"-Effektstärke zum Einsatz. Es ergab sich folgende Schätzung²⁷ des Interventionseffekts für die *Reduktion von psychischen Beschwerden* von "AktivA":

Teilnehmer des verpflichtenden Angebotes (N=139): (T1: 8,5±8,6; T2: 5,9±7,1): **d=0,40**

Teilnehmer des freien Angebotes (N=110): (T1: 12,4±9,3; T2: 8,4±7,5): **d=0,57**

In der hiesigen Gesundheitsförderung gelang die Reduktion von psychischen Beschwerden durchschnittlich mit einem **d=0,46** (vgl. Tab. 4.11 oder nachfolgende Abb. 5.1). Demnach hätte "Fit50+" eine Effektstärke, die sich mindestens mit der von "AktivA" des verpflichtenden Angebotes vergleichen ließe. Der "AktivA"-Interventionseffekt für die Teilnehmer des freien Angebotes liegt jedoch höher. Rothländer et al. (2012) erklären den größeren Erfolg des freien Zugangs mit zu T1 stärkeren psychischen Beschwerden der freien Teilnehmer und einer anzunehmenden höheren Veränderungsmotivation dieser, da sie sich aktiv um eine Teilnahme bemühen mussten.

Ob sich "Fit50+" eher mit dem Effekt des verpflichtenden oder eher mit dem des freien Zugangs von "AktivA" messen sollte, ist nicht ganz klar. Einerseits war für unsrige Probanden die Teilnahme an einer Weiterbildungsmaßnahme rechtlich verpflichtend, andererseits konnten sie aber auswählen, an welchem Qualifizierungsangebot sie teilnehmen wollten, so dass es trotz Verpflichtung noch Auswahlmöglichkeiten gab. Immerhin entschieden sich 30% der Kunden des Bildungsträgers in Schwerin gegen eine Teilnahme am Gesundheitsprogramm. Nähme man der Einfachheit halber an, dass im Vergleich der Rekrutierungswege die Einordnung von "Fit50+" etwa dazwischen liegen müsste, ließen sich die vom Zugang abhängigen Interventionseffekte von "AktivA" mitteln, welches zu einem $\emptyset d=0,49$ für die durchschnittliche geschätzte Senkung von psychischen Beschwerden resultieren würde. Demnach wäre die "Fit50+"-Interventionseffektstärke von $d=0,46$ in etwa vergleichbar mit der mittleren von "AktivA".

Vergleich mit durchschnittlichem publizierten Interventionserfolg

Der für die Vergleichbarkeit repräsentativste Befund zur psychischen Interventionseffektivität von verschiedenen Programmen, die mit erwerbslosen Probanden durchgeführt und publiziert wurden, stammt von Paul und Moser (2009a). In ihrem systematischen Review meta-analysierten sie den Trainingserfolg von 16 internationalen Interventionsstudien und geben ihn als Cohen's *d* für abhängige Stichproben mit $d=0,35$ für die *Reduktion von psychischen Beschwerden* an (vgl. Kap. 1.6.1 für eine genaue Darstellung). Ein Vergleich der Ergebnisse von "Fit50+" mit dem Meta-Interventionserfolg ist mit den Cohen's *d* Effektstärken aus Tab. 4.11 möglich. Grafisch dargestellt verhalten sich die hiesigen Interventionseffekte folgendermaßen mit dem metaanalysierten Interventionserfolg:

²⁶ Für psychische Beschwerden kam die hiesige durchschnittliche Korrelation zwischen den Messzeitpunkten T1 und T2 der Instrumente GHQ, BDI und PSK zur Anwendung (vgl. Tab. 4.11): Die T1*T2-Korrelation des GHQ ($r_{(T1,T2)}=0,706$), des BDI ($r_{(T1,T2)}=0,766$) und der PSK ($r_{(T1,T2)}=0,474$) wurden in Fischer-Z-Werte nach Glass und Stanley (1970, S. 534) transformiert; anschließend erfolgte die Durchschnittsbildung und Rücktransformation als durchschnittlicher Korrelationskoeffizient. Dieser ergab ein von $\emptyset r_{(T1,T2)}=0,665$ für "psychische Beschwerden".

²⁷ Cohen's *d* für abhängige Stichproben nach der Methode von Faul et al. (2007).

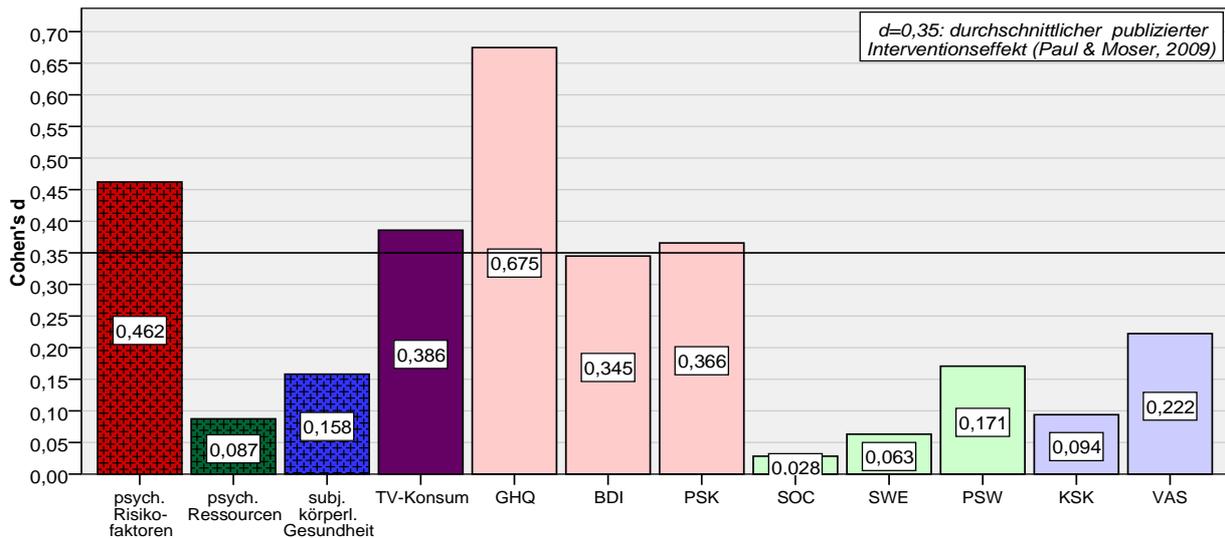


Abb. 5.1: Vergleich der Effektstärken von "Fit50+" mit Meta-Effektstärke von Paul und Moser (2009a, S. 276).

Grafik von links nach rechts: Der Meta-Interventionseffekt von Paul und Moser (2009a) ist als Linie bei $d=0,35$ eingezeichnet. Das durchschnittliche d der hiesigen Interventionseffekte zur Reduktion von psychischen Beschwerden (GHQ, BDI, PSK) ist als **psych. Risikofaktoren** dargestellt. Das durchschnittliche d der Trainingseffekte für den Aufbau von Ressourcen (SOC, SWE, PSW) ist in **psych. Ressourcen** zu finden. Der durchschnittliche Interventionserfolg für den Aufbau der körperlichen Gesundheit (KSK und VAS) ist in **subj. körperl. Gesundheit** repräsentiert. **TV-Konsum** steht als einzelne Variable für die Effektstärke zur Minderung des passiven Freizeitverhaltens. Im Anschluss an diese befinden sich die Effektstärken der einzelnen Messinstrumente, aus denen die zusammengefassten gemittelt wurden (entsprechende Gruppen-Farbe abgeschwächt dargestellt).

Inhaltlich vergleichbar mit dem Meta-Effekt von Paul und Moser ist der hiesige zusammengefasste Interventionseffekt zu "psychischen Risikofaktoren" und die zugrunde liegenden Effektstärken des GHQ, BDI und PSK, weil die Metaanalytiker explizit Interventionsresultate zur *Reduktion von psychischen Beschwerden* ("reductions of distress symptoms", Paul & Moser, 2009, S. 276) zusammenfassten (vgl. Kap. 1.6.1).

Der mittlere Effekt zur Reduktion von psychischen Beschwerden von "Fit50+" liegt mit $d=0,46$ über dem Meta-Effekt von Paul und Moser. Der Interventionseffekt zur Senkung von unspezifischen psychischen Beschwerden (GHQ) liegt mit $d=0,68$ fast doppelt so hoch wie der Meta-Effekt. Der Interventionseffekt zur Reduktion der Depressivität (BDI) entspricht mit $d=0,35$ der Meta-Effektstärke, ebenso wie der Effekt zur psychischen Einschränkungsminderung der Lebensqualität (PSK) von $d=0,37$. Bis auf den Effekt zur Senkung des Fernsehkonsums liegen die restlichen Effektstärken zwar unter dem Meta-Effekt, aber sie sind inhaltlich nicht mit diesem vergleichbar und daher nur zum Überblick dargestellt.

Insofern bleibt festzuhalten, dass im durchschnittlichen Vergleich mit 16 verschiedenen Programmen für Erwerbslose der Trainingserfolg von "Fit50+" im BDI und in der PSK gut mithalten kann und das Resultat des GHQ den Meta-Effekt um fast das Doppelte übersteigt. Da der hiesige gemittelte Effekt zur Reduktion von psychischen Beschwerden ($d=0,46$) den inhaltlich entsprechenden Meta-Effekt ($d=0,35$) übertrifft, ist im durchschnittlichen Vergleich

zu den 16 metaanalytisierten Interventionsstudien die hier durchgeführte Gesundheitsförderung als etwas erfolgreicher (+30%) zu bewerten.

Vergleich mit Kontrollgruppeneffekten

Das hier evaluierte Programm hatte keine Kontrollgruppe, wodurch sich seine Interventionseffekte statistisch nur gegen die Nullhypothesenannahme der Zeiteffekte absichern ließen. Wie in Kap. 2.3.1 dargestellt, metaanalytisierten Paul et al. (2006) sowie Paul und Moser (2009a) unter anderem, dass Erwerbslose ohne eine Intervention im Längsschnitt keine sog. "Testeffekte" aufweisen, d.h. nicht spontan ihre psychische Gesundheit verbessern, wenn sie zu mehreren Messzeitpunkten befragt werden. Paul und Moser errechneten für erwerbslose Studienteilnehmer eine nichtsignifikante Effektstärke von $d=0,03$ für "Veränderungen" von psychischen Beschwerden zwischen T1 und T2. Auch Vuori und Vesalainen (1999) berichteten für ihre Erwerbslosen-Kontrollgruppe, die keine Maßnahme erhielt (Wartegruppe), dass sich in deren Werten zum "psychological distress" (gemessen mit dem GHQ) keine Unterschiede zwischen den Messzeitpunkten finden ließen. Vor diesem Hintergrund kann geschlossen werden, dass die hiesigen signifikanten Interventionseffekte anhand der Tests gegen die Nullhypothesenannahme der Zeiteffekte zumindest implizit gegen eine Wartegruppe abgesichert sein sollten, eben weil es für Erwerbslose untypisch ist, dass sich ihre psychische Gesundheit durch eine pure Beobachtung (=Teilnahme an einer Längsschnittstudie ohne Intervention) verbessert. Insofern sollte ein "Hawthorne-Effekt" (z.B. Hsueh, 2002) als Erklärung der signifikanten Interventionseffekte nicht in Frage kommen.

Gegen nicht näher bestimmbare sog. "Placeboeffekte", die daraus resultieren, dass Erwerbslose an irgendeiner Maßnahme teilnehmen, ist die hiesige Studie mangels Kontrollgruppe ebenfalls nicht abgesichert. Die Studie "AktivA" von Rothländer et al. (2012) genügt jedoch recht gut dieser Design-Anforderung, denn die Autoren benutzten zur Absicherung ihres kognitiv-behavioralen Interventionseffekts Kontrollgruppen, die "AktivA" zwar nicht erhielten, aber zumeist aus Teilnehmern von herkömmlichen Bildungsmaßnahmen für Erwerbslose aus kooperierenden Bildungsträgern rekrutiert wurden (Rothländer & Richter, 2009; Rothländer, et al., 2012). In den Auswertungen erzielten neben den Interventionsteilnehmern von "AktivA" auch die Kontrollgruppensubjekte teilweise signifikante Verbesserungen des Befindens. Zu den Effekten der Kontrollgruppenteilnehmer schreiben sie:

"Dies kann möglicherweise ein Hinweis auf die Selbstwert stabilisierende und sozial integrative Qualität der Bildungs- und Beschäftigungsmaßnahmen sein, in denen sich die meisten Untersuchungsteilnehmer/innen befanden" (Rothländer & Richter, 2009, S. 15).

In der Publikation von 2012 berichtet die Forschergruppe für die Kontrollgruppen Mittelwerte und Standardabweichungen zu den Messzeitpunkten T1 und T2 u.a. zum Indikator "psychische Beschwerden". Für die Kontrollgruppe zur Absicherung des verpflichtenden Angebots (N=97) lässt sich für den Effekt *Reduktion von psychischen Beschwerden* (T1: $8,5\pm 8,4$; T2: $7,6\pm 8,2$) eine Effektstärke von $d=0,13$ schätzen. Für die Kontrollgruppe zum freien Angebot (N=85) (T1: $11,6\pm 10,8$; T2: $9,9\pm 10,3$) ist ein Effekt von $d=0,20$ schätzbar (Quelle: eigene

Berechnung²⁸). Unterstellt man bei der Teilnahme an einer herkömmlichen Bildungsmaßnahme für Erwerbslose "Placeboeffekte" für die *Reduktion von psychischen Beschwerden*, wäre diese "Placebo-Effektstärke" anhand der Studie von Rothländer et al. (2012) mit einem durchschnittlichen $d=0,17$ schätzbar. Rothländer et al. (2012) konnten die Interventionseffekte von "AktivA" signifikant gegen die der Kontrollgruppen absichern.

Da mit "Fit50+" für die Reduktion von psychischen Beschwerden ein durchschnittlicher Interventionseffekt von $d=0,46$ erreicht wurde, wäre das hiesige Interventionsresultat 2,7-mal stärker als der inhaltlich entsprechende Kontrollgruppeneffekt von $d=0,17$, der anhand Rothländer et al. (2012) geschätzt wurde. Von daher soll trotz des Mangels einer eigenen Kontrollgruppe postuliert werden, dass sich der hiesige Interventionserfolg zur Minderung von psychischen Beschwerden nicht zufriedenstellend mit möglicherweise existierenden "Placeboeffekten" von herkömmlichen Maßnahmen erklären lässt.

Zeitstabilität der signifikanten Interventionsergebnisse

Deskriptiv zeigte sich in allen signifikanten Interventionseffekten von "Fit50+" im follow-up ein Rückgang der zuvor gewonnenen Verbesserungen (vgl. Abb. 4.1, S. 62). Rein numerisch betrachtet manifestierte sich also in keinem Messverfahren eine absolute Zeitstabilität.

Im GHQ und im TV-Konsum zeigte sich aber trotz der deskriptiven Rückgänge der Trainingseffekte jeweils ein signifikanter Helmert-Kontrast. Dies signalisiert, dass die Effekte der Intervention für diese Konstrukte trotz Hinzunahme ihrer Rückgänge im follow-up noch belegbar waren. Von daher werden sowohl der Abbau von unspezifischen psychischen Beschwerden (GHQ) als auch die Reduktion des passiven Lebensstils (TV-Konsum) als relativ zeitstabile Interventionseffekte interpretiert.

Im Gegensatz dazu erreichte im BDI und in der PSK der Helmert-Kontrast keine Signifikanz; in diesen Messverfahren war unter Einbezug des follow-up der Interventionseffekt nicht mehr nachweisbar. Deswegen ist sowohl für die Reduktion von Depressivität (BDI) als auch für die psychische Einschränkungsminderung der Lebensqualität (PSK) von einer zeitlich begrenzten Wirkung auszugehen, die sich innerhalb von 6 Monaten nach der Intervention wieder auflöst.

Eine vergleichende Einordnung der Trainings-Zeitstabilität mit anderen Interventionsstudien ist schwierig (bis unmöglich), weil in vielen Untersuchungen entweder gar kein follow-up durchgeführt wurde (z.B. Saam & Wodtke, 1995; Maysent & Spera, 1995; Creed, et al., 1996; Frese, et al., 2002; Vastamaki, et al., 2009) oder die follow-up-Zeiträume unterschiedlich ausfallen. Sie reichen von 3 Monaten nach der Intervention (z.B. Rothländer, et al., 2012; Proudfoot, et al., 1997) bis zu 2 Jahren (z.B. Vuori & Silvonen, 2005). Dazu kommen noch verschiedene statistische Methoden für den Test des Kriteriums "Zeitstabilität". Es wäre nur mit erheblichem Aufwand möglich, valide Vergleiche durchzuführen und zudem müsste auf die Interventionsstudien, die kein follow-up durchführten, verzichtet werden.

Wie in Kap. 1.6.1 ausführlicher dargestellt, kamen Paul und Moser (2009c) für den Meta-Interventionseffekt ($d=0,35$) der 16 Primärstudien zu folgender Meta-Einschätzung der

²⁸ Wie auf Seite 85 wurde auch für die Kontrollgruppen Cohen's d für den Unterschied zw. T1-T2 für abhängige Stichproben nach Faul et al. (2007) berechnet. Als Korrelationsschätzer kam erneut die hiesige mittlere Korrelation für *psychische Beschwerden* zwischen T1-T2 mit $\emptyset r_{(T1,T2)}=0,665$ zum Einsatz.

zeitlichen Stabilität: "Ein positiver Effekt auf die psychische Gesundheit tritt kurz nach der Intervention auf, schwächt sich dann aber mehr und mehr ab" (Paul & Moser, 2009c, S. 130). Die Autoren interpretieren, dass für Erwerbslose, die trotz einer Intervention im Anschluss keine neue Stelle finden, ihr Hauptproblem (die Erwerbslosigkeit) ungelöst bleibt und dadurch eine negative psychische Symptomatik mit der Zeit wieder Überhand gewinnen kann (Paul & Moser, 2009c). In diesem Sinne sollen auch die hiesigen Interventionseffekte, die keine zeitliche Stabilität aufwiesen, interpretiert werden.

Nichtsignifikante Interventionserfolge

In Bezug auf die Ressourcen der psychischen Gesundheit ist das Resultat als eindeutig zu bewerten. Die wohl bedeutendsten Konstrukte der Gesundheitspsychologie "Selbstwirksamkeit" (und damit mittelbar auch "Kontrollüberzeugung") sowie "Kohärenzsinn" zeigten in der longitudinalen Auswertung keine nennenswerten Effekte. Auch das zusätzliche salutogenetische Instrument zum psychosozialen Wohlbefinden konnte keine signifikante Änderung zwischen den Zeitpunkten T1-T2 für sich verbuchen. Es ist zu schließen, dass das Förderprogramm keine Steigerung der Ressourcen der psychischen Gesundheit herbeiführen konnte. Warum diese keine Verbesserung erfuhren, könnte aber zumindest teilweise damit erklärt werden, dass sich für das Kollektiv zu T1 kaum Einschränkungen in Relation zur Allgemeinbevölkerung objektivieren ließen. Der Anteil von Probanden mit eingeschränkter Selbstwirksamkeitserwartung und unterdurchschnittlichem Kohärenzsinn lag nur etwas über dem Erwartungswert der Normen, so dass nicht von einem deutlichen Interventionsbedarf auszugehen war. Statistisch gesehen lässt sich in den Messinstrumenten somit die Wirkung eines Deckeneffektes für positive Veränderungen nicht ausschließen. Zusätzlich ist v.a. der SOC dafür bekannt, dass er für erwachsene Menschen ein Trait, also einen relativ zeitstabilen Wert erfasst (Antonovsky & Franke, 1997). Die Publikation von Vastamaki et al. (2009) weist zwar darauf hin, dass eine Erhöhung des SOC, selbst für ältere Arbeitslose, resultieren kann, wenn diese auf dem Arbeitsmarkt wiederingestellt werden; im Vergleich dazu war der Erfolg der hiesigen Intervention aber offensichtlich viel zu schwach.

Schwieriger ist die Interpretation der nicht bestätigten Hypothesen zum körperbezogenen Gesundheitsempfinden (VAS) und zur dazugehörigen Lebensqualität (KSK). Bereits der gesunde Menschenverstand vermutet für ein primär auf die körperliche Aktivierung aufbauendes Gesundheitsförderungsprogramm auf der subjektiv-körperlichen Ebene gut messbare Erfolge. Auch der Anteil von Probanden mit einer unterdurchschnittlichen körperlichen Lebensqualität, gemessen mit der KSK, und der Anteil von Probanden mit unterdurchschnittlichem Gesundheitsempfinden, gemessen mit der VAS, lagen relativ zum erwarteten Anteil zu zwei Drittel höher, so dass in Bezug auf die körperliche Gesundheit von einem relevanten Interventionsbedarf auszugehen war.

Die Mittelwerte des VAS zwischen T1-T2 unterschieden sich aber immerhin so stark, dass die Effektstärke des VAS ($\eta^2=0,050$) in der Nähe eines "mittleren" Effektes²⁹ lag. Ein mittlerer Effekt wird in der Interventionsforschung meist als relevanter Effekt interpretiert. Wenn ein nahezu relevanter Effekt in einer Analyse nicht signifikant wird, könnte auch ange-

²⁹ "Mittlerer Effekt" nach Cohen (1988, 1992): $\eta^2 \geq 0,059$

nommen werden, dass er "unter-powered" erhoben wurde, also die Stichprobengröße im Fall der VAS (N=50) nicht ausreichte, um den Effekt statistisch abzusichern.

Für die KSK ließe sich argumentieren, dass dadurch, dass sie ausschließlich anhand der körperlichen Einschränkungen die dazugehörige Lebensqualität misst, nur ein kleines Veränderungsergebnis resultiert und dafür kommen folgende Gründe infrage: An anderer Stelle³⁰ konnte für ein Sub-Sample unserer Stichprobe gezeigt werden, dass allein die Summe der medizinischen Befunde negativ mit der KSK zusammenhing ($r=-0,49$; $p<0,01$). Krankheiten bzw. Gebrechen oder Schmerzen senkten demnach substantiell die körperbezogene Lebensqualität. Während manche der erhobenen medizinischen Befunde prinzipiell einer Linderung durch Sport zugänglich sind, traf dies jedoch nicht auf alle bzw. nicht auf alle gleichermaßen zu (z.B. Diabetes Mellitus). Des Weiteren war der Interventionszeitraum für eine relevante Verbesserung von bestimmten medizinischen Diagnosen (z.B. Adipositas) wahrscheinlich zu kurz. Für andere Befunde, wie z.B. der Hypertonie, war eine objektiv feststellbare Senkung des Blutdrucks für die Probanden möglicherweise nicht ausreichend subjektiv erlebbar mit der Reduktion von physischen Einschränkungen verknüpft und konnte von daher nicht zu einer Erhöhung der körperbezogenen Lebensqualität beitragen. Eventuell war also die KSK für die hiesige Intervention einfach nicht veränderungssensitiv genug oder der Trainingszeitraum war für eine wahrnehmbare Erhöhung der körperlichen Gesundheit noch zu kurz bemessen.

5.2.3 Ergebnisse zur Moderation der Interventionseffekte

Intensivierung des Trainings in der 2. Interventionskohorte

Für den Einfluss des Zwischensubjektfaktors "Kohorte" auf die Veränderung der Messzeitwerte sollten die Ergebnisse insgesamt als eindeutig interpretiert werden. Die Intensivierung des Trainings in der 2. Interventionskohorte führte in 8 (von 9) statistischen Prüfungen nicht zu differenzierten Trainingseffekten. Somit ist in diesen die Steigerung des Ausdauertrainings als wirkungslos zu interpretieren.

Der Innersubjektfaktor der Selbstwirksamkeitserwartung (SWE) zeigte zwar einen signifikanten Interaktionseffekt mit dem Zwischensubjektfaktor "Interventionskohorte", allerdings ist zu interpretieren, dass dieser nicht zustande gekommen wäre, wenn die numerische SWE-Steigerung von Kohorte 2 nicht gleichzeitig mit der numerischen Senkung von Kohorte 1 hätte interagieren können (vgl. Abb. 4.3, S. 66). Die zusätzlichen, kohortenspezifischen t-Tests des Effekts zwischen T1-T2 ergaben für beide Gruppen nichtsignifikante Ergebnisse, welches signalisiert, dass die Unterschiede der Messzeitpunkte T1-T2 für beide Trainingsgruppen noch im Rahmen von Zufallsschwankungen lagen. Demgegenüber erreichten die zu T1 verschiedenen anmutenden Ausgangsbedingungen zwischen den Kohorten immerhin einen statistischen Trend. Von daher sollte der Interaktionseffekt von "Interventionskohorte" mit dem Zeiteffekt der SWE eher als Phänomen einer "Regression zur Mitte"

³⁰ Preuss M., Kreuzfeld S., Arndt D., Kumar M., Neubert S., Weippert M., Stoll R. *Körperbezogene Lebensqualität und Psychologische Faktoren bei älteren Langzeitarbeitslosen*. Abstract. 49. Wissenschaftliche Jahrestagung der DGAUM e.V. Aachen, Germany, pp. 96, March 11th-14th, 2009.

interpretiert werden, als das man ihm unterstellen sollte, dass er auf tatsächlich unterschiedlich wirksamen Trainingsbedingungen fußt.

Unter gemeinsamer Würdigung dieses Umstandes mit den nichtsignifikanten Interaktionseffekten der restlichen 8 Messverfahren soll daher geschlossen werden, dass die Intensivierung des Ausdauertrainings keinen Einfluss auf die Interventionsresultate hatte, weder zur psychischen und subjektiv-körperlichen Gesundheit noch auf die Änderung des passiven Freizeitverhaltens. Für die psychischen Variablen deckt sich dies mit den Ergebnissen von Aoyagi, Park und Shephard (2010), die auf Basis ihrer Längsschnittergebnisse für N=183 ältere Menschen schlussfolgern, dass ein regelmäßiges, angemessenes körperliches Training für eine Verbesserung der psychischen Gesundheit ausreicht und eine darüber hinausgehende Intensitätssteigerung keine stärkeren Befindensvorteile erbrachte.

Effekte der Sportfortsetzung im follow-up

Ein auf den ersten Blick ähnliches Ergebnis wie zur Intensivierung des Trainings zeigte sich für den Zwischensubjektfaktor "Sportfortsetzung" auf die Veränderungsmessungen zwischen T2-T3. Mit keiner der 9 follow-up-Veränderungen ging die Sportfortsetzung eine signifikante Interaktion ein. Es konnten lediglich tendenzielle Befindensvorteile durch die Fortsetzung des Sportes für die körperliche Lebensqualität (KSK) und die Selbstwirksamkeitserwartung (SWE) gefunden werden. Da die Interaktionen mit der KSK und der SWE jeweils einen "kleinen" Effekt nach Cohen (1988, 1992) mit statistischem Trend produzierten, ist zu schließen, dass die Aufrechterhaltung des Sportes im follow-up allenfalls geringe gesundheitsförderliche Resultate andeutete.

Diskussion der Sporteffekte und mutmaßliche weitere Einflussgrößen

Das Ergebnis der Sportfortsetzung im follow-up wirkt neben seinem primären Resultat darüber hinaus enttäuschend, weil man auf die Idee kommen könnte, damit die Rolle des Sportes für die Wirksamkeit des Trainings anzuzweifeln. Wenn man den hiesigen Effekt der Gesundheitsförderung hauptsächlich auf den Sport zurückführt, erscheint es doch auf den ersten Blick verwunderlich, dass seine Fortsetzung im Kontrast zum Abbruch höchstens kleine gesundheitsförderliche Effekte im follow-up andeutete. Ein derart globaler Zweifel wäre jedoch einzuschränken, da die Sporteffekte im follow-up aus folgenden Gründen nicht mit denen der Intervention vergleichbar sein sollten:

Erstens war die Intensität, mit der die Sportfortsetzer im follow-up freiwillig und eigenständig weitertrainierten, nicht vorgegeben. Während in der Interventionsphase die Stärke des Ausdauertrainings stets von den Sportpädagogen sicher gestellt war und regelmäßig, meist zweimal wöchentlich, trainiert wurde, bestimmte der Proband im follow-up selbst, wie oft und mit welchem Schwierigkeitsgrad er sich mit dem körperlichen Training konfrontieren wollte. Es wäre demnach möglich, dass die nichtsignifikanten follow-up-Unterschiede zwischen Fortsetzern und Abbrechern daraus resultierten, dass das eigenständige Training von den Sportfortsetzern nicht mehr ausreichend intensiv genutzt wurde.

Zweitens lässt sich voraussetzen, dass die psychischen Effekte des körperlichen Trainings in den Anfangswochen größer ausgefallen sein dürften als im weiteren zeitlichen Verlauf und vor allem im follow-up. Just vor dem Bewegungsprogramm hatten die meisten Teilnehmer seit Jahrzehnten keinen Sport mehr getrieben. Entsprechend groß waren die

ersten Erfolgserlebnisse durch das Wiederentdecken von verloren geglaubten physischen Funktionen. Es ist bekannt, dass sich Erfolgserlebnisse mit fortschreitender Dauer eines begonnenen Trainingsprogramms abschwächen, weswegen das Ausbleiben von abgesicherten Unterschieden zwischen Abbrechern und Fortsetzern im follow-up auch psychischen Gewöhnungseffekten geschuldet sein könnte.

Drittens führten die Probanden das freiwillige Training im follow-up meist solitär durch, während das Ausdauerprogramm der Intervention in Gruppen mit jeweils 15 Langzeitarbeitslosen stattfand. Das Gruppentraining beinhaltete die unterstützende Anleitung durch die Trainer und wahrscheinlich auch motivierende Synergien der Gruppe, u.a. sind Effekte des Modelllernens annehmbar. Diese motivationalen Faktoren fehlten beim solitären Training im follow-up.

Aller Einschränkungen der Kritik zum Trotz darf natürlich nicht ausgeschlossen werden, dass neben den sportassoziierten Effekten auch weitere Einflussgrößen zum hiesigen Interventionsergebnis beitragen. So ist sicherlich die Gesundheitsedukation zu nennen, welche stets vor dem Sportprogramm durchgeführt wurde und vorbereitend die körperbezogene Gesundheitskompetenz der Teilnehmer fördern sollte. Obwohl eine psychologische Intervention nicht durchgeführt wurde, kann aber davon ausgegangen werden, dass der wissensvermittelnde Kurs die Probanden kognitiv auf eine Verbesserung ihres Wohlergehens fokussierte und damit evt. Priming-Effekte auslöste. Einschränkend ist allerdings ins Gedächtnis zu rufen, dass viele Teilnehmer erst während des theoretischen Kurses rekrutiert werden konnten und somit die Messbarkeit des gesundheitsedukativen Effekts auf die Psyche in vielen Fällen leider "verschenkt" wurde. Eine differenzierte Beurteilung dieses Effektes wäre allerdings nur dann möglich gewesen, wenn im Studiendesign zwischen dem wissensvermittelnden Kurs und dem Bewegungsprogramm ein weiterer Messzeitpunkt eingeführt worden wäre.

Des Weiteren ist anzunehmen, dass die Langzeitarbeitslosen im Gruppentraining auch davon profitierten, dass sie soziale Beziehungen mit Ihresgleichen eingehen konnten (sozialer Effekt). Wie psychisch stabilisierend für Langzeitarbeitslose der Kontakt zu anderen Erwerbslosen sein kann, wurde ja in der Publikation von Rife und Belcher (1993) (vgl. Kap. 1.5) empirisch belegt. Daneben könnte die pure Teilnahme an der Maßnahme auch schwierig bestimmbare Effekte produziert haben. Diese lassen sich mit Aussagen der Teilnehmer z.B. in Richtung von "Hauptsache, mal etwas anderes" oder "Ich habe das Gefühl, dass mich das irgendwie weiter bringt", also als *Abwechslung vom Alltag* bzw. *Wahrnehmung von Entwicklungsmöglichkeiten* (exemplarisch) paraphrasieren.

Von diesen Effekten ist natürlich annehmbar, dass sie auf sämtliche Maßnahmen für Erwerbslose zutreffen sollten. Für den *sozialen Effekt* sollte es prinzipiell egal sein, ob die Teilnehmer gemeinsam ein Gesundheitsförderungsprogramm oder z.B. einen Computerkurs oder ein Bewerbungstraining absolvieren; soziale Kontaktmöglichkeiten wären ja jeweils gegeben. Analog dazu ließen sich leicht Argumente für mutmaßliche Effekte wie *Abwechslung vom Alltag* oder *Wahrnehmung von Entwicklungsmöglichkeiten* finden. Für beide kann angenommen werden, dass sie auch in herkömmlichen Qualifizierungsmaßnahmen für Langzeitarbeitslose ihre Wirkung entfalten. Zu einem in diese Richtung interpretierbaren

Ergebnis kommen Vuori und Vesalainen (1999) (vgl. Kap. 1.6), die über ihre Evaluation von 3 herkömmlichen, aber inhaltlich unterschiedlichen Maßnahmen für Arbeitslose berichten, dass diese zwar sämtlich psychisch stabilisierende Interventionseffekte produzierten (und gegen eine Wartegruppe abgesichert waren), aber diese zwischen den Maßnahmen nicht unterschiedlich stark ausfielen. Hier könnte man die - trotz inhaltlich unterschiedlicher Programme - gleich hohen Interventionseffekte auf die ihnen innewohnenden Gemeinsamkeiten zurückführen. Die Annahme, dass die Gemeinsamkeiten der von Vuori und Vesalainen untersuchten Maßnahmen u.a. darin bestanden, dass die Erwerbslosen soziale Beziehungen knüpfen konnten, ihnen Abwechslung vom Alltag sowie neue Entwicklungsmöglichkeiten geboten wurden, wäre zumindest nicht unplausibel.

Eventuell repräsentieren solche beispielhaft angesprochenen Effekte tatsächlich psychisch wirksame Komponenten von sämtlichen Förderungen für Erwerbslose, die bei herkömmlichen Maßnahmen leichtfertig und zudem diskreditierend unter dem Begriff "Placeboeffekt" subsumiert werden. Dass solche Effekte auch im hiesigen Gesundheitsförderungsprogramm eine Rolle spielten, darf logischerweise nicht abgestritten werden. Wie in der Diskussion um die Kontrollgruppeneffekte festgestellt wurde, liegen die signifikanten Interventionseffekte zur Reduktion von psychischen Beschwerden ($d=0,46$) von "Fit50+" jedoch deutlich höher als inhaltlich entsprechende Kontrollgruppeneffekte ($d=0,17$), die auf Basis der Studie von Rothländer et al. (2012) geschätzt wurden. Diese würden nach obiger Argumentationslinie unspezifisch wirksame "Placeboeffekte" repräsentieren, da die Kontrollgruppensubjekte der "AktivA"-Studie meist an üblichen Qualifizierungsmaßnahmen teilnahmen. Spezifisch wirksame Gesundheitsförderungseffekte könnten daher als Differenz zwischen den hiesigen Interventions- und den geschätzten "Placeboeffekten" ausgedrückt werden, welches zu einem $d=0,46-0,17=0,29$ resultieren würde. An dieser Schätzung ließe sich kritisieren, dass man nicht der Einfachheit halber von einem additiven Zusammenhang zwischen Gesundheitsförderungs- und "Placeboeffekten" ausgehen dürfe, da argumentiert werden könnte, dass diese Interaktionen miteinander eingehen; z.B. ließe sich spekulieren, dass Sport möglicherweise den "sozialen Effekt" verstärkt. Eine genaue, abschließende Differenzierung ist aufgrund der Datenlage zwar nicht möglich; Tatsache ist jedoch, dass selbst die möglicherweise einen "Placeboeffekt" verstärkenden Wirkungen durch den Sport nicht hätten auftreten können, wenn das Bewegungsprogramm nicht Hauptbestandteil der hiesigen Gesundheitsförderung gewesen wäre, weshalb es unangemessen wäre, solche Effekte mit dem Begriff "unspezifisch wirksam" auszulegen.

Depressivität

Als bedeutend erwies sich der Zwischensubjektfaktor "Depressivität" in stets signifikanten Interaktionen mit jeweils "mittlerer" Effektstärke mit sämtlichen Risikofaktoren der psychischen Gesundheit: Versuchsteilnehmer, die eine Depressivität mit relevanter Intensität aufwiesen, profitierten von allen Teilnehmern am stärksten in der Reduktion von unspezifischen psychischen Beschwerden (GHQ), in der Senkung von Depressivität (BDI) und in der Minderung ihrer durch psychische Probleme verursachten Einschränkungen der Lebensqualität (PSK). Insofern bestätigt sich für mutmaßlich depressive Langzeiterwerbslose eine psychisch stabilisierende und antidepressive Wirkung von Sport; eine Erkenntnis, die für ältere,

nicht zwingend erwerbslose, depressive Menschen ja bereits belegt ist (z.B. Blumenthal, et al., 1999; Kessler, et al., 2001; van Gool, et al., 2007). Des Weiteren stimmen die hiesigen Ergebnisse auch mit Vinokur et al. (1995) überein, die ebenfalls zu dem Ergebnis kamen, dass vornehmlich die von Depressivität betroffenen Probanden von ihrer Fördermaßnahme für Erwerbslose profitierten.

Im TV-Konsum führte die Depressivitätsdiagnose nicht zu einem differenzierten Interventionseffekt, aber dafür signifikant zu einem selektiv ungünstigeren follow-up-Ergebnis mit "mittlerer" Effektstärke. Inhaltlich ist zu interpretieren, dass die Diagnose nicht, wie prognostiziert, zu einer verminderten Reduktion des TV-Konsums führte. Stattdessen zeigte sich selektiv für die Probanden mit Depressivität, dass diese nach der Intervention zu ihren alten TV-Gewohnheiten zurückkehrten, während dies für die restlichen Probanden nicht der Fall war. Im Terminus für Abhängigkeitserkrankungen gesprochen erlebten die Probanden mit depressiver Symptomatik im follow-up also einen "Rückfall" im TV-Konsum. Vor dem Hintergrund, dass sich der Interventionseffekt zur Verminderung der Depressivität als zeitlich instabil erwies, wäre spekulierbar, dass der selektive Rückfall zum passiven Freizeitverhalten als Ausgleich für die verloren gegangene (weil nur kurzzeitig wirksame) Linderung der depressiven Symptomatik diene. Eventuell spiegelt der follow-up-Effekt auch eine stärkere Abhängigkeit und damit eine höhere Rückfallwahrscheinlichkeit von Depressiven zum Medium TV wider. So interpretieren auch Depp et al. (2010), dass der TV-Konsum vornehmlich von älteren Menschen mit affektiven Problemen als Copingverhalten zur Regulierung ihrer emotionalen Balance eingesetzt wird.

Geschlecht

Während der Zwischensubjektfaktor "Geschlecht" im VAS nicht zu differenzierten mittleren Interventionseffekten führte, deutete er eine geschlechtsspezifische Wirkung des Ausdauertrainings in der KSK immerhin als Trend an, die sich als "kleiner" Effekt manifestierte. Die deskriptive Inspektion der Daten ergab, dass exklusiv die Männer von der Erhöhung ihrer körperlichen Lebensqualität profitierten. Ein gruppenspezifischer Test für die Männer zwischen T1-T2 wies auf eine tendenzielle Stärkung ihrer körperlichen Lebensqualität hin. Mangels Signifikanz des Interaktionseffektes dürfen die Unterschiede zwischen Männer und Frauen allerdings nur vorsichtig interpretiert und allenfalls als Hinweis auf eine mögliche geschlechtsspezifische Wirkung des Ausdauertrainings verstanden werden. Dies stünde in Übereinstimmung mit den Ergebnissen von Jette et al. (1996), die für ein Ausdauertraining männerspezifische Steigerungen der körperlichen Lebensqualität bei älteren Menschen nachweisen konnten.

Falls dieser Befund in künftigen Studien mit Erwerbslosen signifikant bestätigt werden könnte, wäre dies für die präventive Praxis eine relevante Erkenntnis. Während Männer Präventionsangebote generell deutlich weniger gut als Frauen annehmen, (oft wird von einem Verhältnis von 20% zu 80% gesprochen, z.B. McQueen, et al., 2006), trifft dieses Missverhältnis ebenso auf Erwerbslose zu (Mielck, 2005). Obwohl empirisch erwiesen ist, dass unter den Erwerbslosen die Männer einen ebenso hohen gesundheitsbezogenen Interventionsbedarf wie die Frauen aufweisen, nehmen sie herkömmliche Präventionsangebote

zur Stärkung ihrer psychischen (z.B. psychosoziale Beratung) und körperlichen (z.B. Ernährungsberatung) Gesundheit weitaus weniger häufig an (Berth, et al., 2008).

Wie sich in der hiesigen Untersuchung zeigte, setzten die männlichen Studienteilnehmer aber zum einen häufiger als die weiblichen ihre Teilnahme bis zum Ende fort (67% der Männer, 40% der Frauen)³¹ und zum anderen zeigten sie auch im follow-up eine höhere Aufrechterhaltung des Sportes (41% der Männer, 33% der Frauen). Aufgrund dieser Quoten ist zu schließen, dass mit dem sportbasierten Gesundheitsprogramm dieses Mal ein Präventionsangebot bereitgestellt wurde, das - im Kontrast zu anderen Angeboten - sehr gut von den Männern angenommen wurde. Die oben diskutierte *mögliche* männerspezifische Effektivität des Sportes wäre als Erklärung für die gute Akzeptanz des Programms in Betracht zu ziehen. Für erwerbslose Männer mit offensichtlichen gesundheitlichen und psychischen Problemen, bei denen sich jedoch durch ein einführendes Gespräch herausstellt, dass sie etablierten Präventionsangeboten eher skeptisch bis ablehnend gegenüber stehen, könnte ein Sportprogramm zur Förderung ihrer körperlichen und psychischen Gesundheit somit künftig eine interessante Alternative darstellen.

5.2.4 Fazit zum Interventionseffekt und Ausblick

Die Frage, ob das hier evaluierte Gesundheitsprogramm psychologisch wirksam war, ist zu bejahen. Die Indikatoren zu psychischen Risikofaktoren zeigten zusammenfassend den höchsten Interventionsbedarf und auch mit deren Senkung war das Training am erfolgreichsten. Weil es für Erwerbslose untypisch ist, dass sich ihre psychischen Beschwerden spontan abbauen (Paul et al., 2006; Paul & Moser, 2009a), kann geschlossen werden, dass die hiesigen Veränderungsergebnisse die einer Warte-Kontrollgruppe übertroffen hätten. Psychische Ressourcen konnten über ihr Anfangsniveau zwar nicht weiter gesteigert werden, allerdings signalisierte die Eingangsdiagnose auch keinen besonderen Bedarf an dieser Erhöhung. Für die subjektiven körperbezogenen Messgrößen bleibt es allerdings offen, ob die Trainingszeit sowohl zu kurz als auch die Messinstrumente nicht ausreichend veränderungssensitiv waren; Argumente ließen sich für beide Sichtweisen finden.

Die Frage, ob die hiesige Gesundheitsförderung psychologisch empfohlen werden kann, sollte ebenso bejaht werden. Der von Paul und Moser (2009a) auf Basis von 16 Interventionsstudien für Erwerbslose berechnete Meta-Interventionseffekt zur Linderung von psychischen Beschwerden wurde vom hiesigen Effekt durchschnittlich um 30% übertroffen. Auf Basis der Daten von Rothländer et al. (2012) ließ sich schätzen, dass die hiesige Gesundheitsförderung vergleichbare psychische Verbesserungen vorzuweisen hatte wie eine kognitiv-behaviorale Gruppenintervention. Des Weiteren konnte geschätzt werden, dass die hier erzielten Linderungen psychischer Beschwerden weitaus stärker wirksam waren, als man es von "Placeboeffekten" von herkömmlichen Maßnahmen für die Zielgruppe hätte erwarten können. Die Auswertung des 6-monatigen follow-up signalisierte für die Reduktion von unspezifischen psychischen Beschwerden und für die Senkung des passiven Lebensstils noch eine relative Zeitstabilität, während die Effekte zur Linderung der Depressivität und

³¹ Teilnahmequote an sämtlichen Untersuchungen im IPM in Rostock.

zur Verkleinerung der psychischen Einschränkung der Lebensqualität allerdings auf den Interventionszeitraum begrenzt blieben.

Weil sich in den Analysen zeigte, dass die Gesundheitsförderung im Mittel hauptsächlich für Studienteilnehmer mit depressiver Symptomatik ihre Wirkung entfalten konnte, lässt sich ableiten, dass das Programm bevorzugt oder gar exklusiv Langzeiterwerbslosen mit der Symptomatik angeboten werden sollte. Während einer solchen Empfehlung auf den ersten Blick nichts entgegenzusetzen wäre, sollte auf den zweiten Blick das Argument von Vinokur et al. (1995) in Betracht gezogen werden, dass bei einem exklusiven Angebot für eine depressive Zielgruppe auf Vorbildeffekte, die von nichtdepressiven Teilnehmern ausgehen könnten, verzichtet werden müsste, welches den Trainingseffekt in unbekanntem Ausmaß schmälern könnte.

Sowohl die geschlechtsspezifischen Quoten der Studienkomplettierung als auch die der Sportfortsetzer im follow-up lassen darauf schließen, dass die männlichen Teilnehmer das Gesundheitsförderungsprogramm etwas besser als die weiblichen akzeptierten, während von sonstigen Präventionsangeboten gemeinhin berichtet wird, dass diese besser von Frauen angenommen werden. Von daher kann das hiesige Training vor allem für erwerbslose Männer als eine Alternative mit bevorzugter Akzeptanz empfohlen werden.

Auch wenn sich für die Sportfortsetzung im follow-up allerhöchstens kleine gesundheitsförderliche Effekte andeuteten, bleibt die Aufrechterhaltung eines begonnenen aktiven Bewegungsverhaltens allein aus medizinischen Gründen in jedem Fall erstrebenswert, z.B. zur Prävention von vaskulären Erkrankungen (vgl. Sui, LaMonte, Laditka, Hardin, Chase, et al., 2007) und sollte daher auch in zukünftigen Programmen als Ziel definiert werden, um einen nachhaltigen Effekt für die Gesundheit bereitstellen zu können. Dass zur Erhöhung der Aufrechterhaltungswahrscheinlichkeit die Wahrnehmung der psychischen Interventionseffekte verstärkt werden sollten, wurde im Kap. 5.2.1 anhand der Daten aus Kap. 4.2.2 hergeleitet.

Für die weitere Forschung wäre es für die psychische Stabilisierung von Langzeiterwerbslosen wichtig zu wissen, ob eine kognitiv-behaviorale Gruppenintervention (wie z.B. "AktivA" von Rothländer, et al., 2012), durch die Hinzunahme des hiesigen Sportprogramms sinnvoll so kombiniert werden könnte, dass sich beide Prozesse gegenseitig befruchten und zu einem höheren und/oder zeitstabileren Interventionseffekt führen. Wie bereits erwähnt, hatten die meisten der hier untersuchten älteren Langzeitarbeitslosen seit Jahrzehnten keinen Sport mehr betrieben. Für die Teilnehmer bedeutete das gewissermaßen einen Neustart dieser Form der Selbsterfahrung. Es kam zum Wiederentdecken von physischen Funktionen und körperlicher Leistungsfähigkeit, von denen viele Probanden angenommen hatten, sie längst verloren zu haben. Durch die Verstärkung, sich körperlich aktiv neu auszuprobieren kam es zu Erfolgserlebnissen, von denen angenommen werden kann, dass diese psychische Transfereffekte produziert haben. Diese kann man mit folgender Aussage eines Teilnehmers exemplarisch umschreiben: "Jetzt, wo ich Sport treibe, obwohl ich mir das schon gar nicht mehr zugtraut hatte, beginne ich mich zu fragen, welche anderen Dinge ich im Leben noch verändern könnte."

Auch eine Psychotherapie erstrebt nicht generell die Aufnahme von neuen Verhaltensweisen sondern zielt, wenn möglich, auch auf das Wiederentdecken und -beleben bzw. dem Ausbau eines vorhandenen, psychisch förderlichen Verhaltensrepertoires ab (Ressourcenorientierung). Von daher sind die Prozesse zur (Wieder-)Aufnahme eines körperlichen Bewegungsverhaltens und zur psychischen Verhaltensmodifikation durchaus vergleichbar. Psychotherapeutische Behandlungen laufen oft Gefahr, dass notwendige Verhaltensänderungen von den Individuen zwar eingesehen, intendiert und sogar geplant werden, aber es an der konkreten Umsetzung mangeln kann. Kuhl (2001) spricht hier vom Verharren in einer "Lageorientierung", obwohl eine "Handlungsorientierung" für eine konkrete Umsetzung erforderlich sei. Im Rahmen des *Transtheoretischen Modells der Verhaltensänderung* (Prochaska & Velicer, 1997) wird angenommen, dass die konkrete Handlung wie eine Stufe mit der Hilfe von bestimmten Vorbereitungen erklommen werden muss. Zum Erklimmen der Handlungsstufe einer Verhaltensänderung könnte der Sport eine wertvolle Hilfe bieten, denn die Aufnahme einer sportlichen Betätigung fordert vom Individuum von Anfang an ein aktives, konkretes Handeln, welches relativ rasch verfügbare Erfolgserlebnisse bereitstellen kann. Daraus entstehende psychische Transfereffekte sollten dazu nutzbar sein, ein kognitiv-behaviorales Training flankierend darin zu unterstützen, die Teilnehmer zur handlungsorientierten Umsetzung ihrer psychischen Ziele zu aktivieren. Eine kognitiv-behaviorale Gruppenintervention könnte ihrerseits helfen, sportbezogene Erfolgserlebnisse samt ihrer positiven seelischen Effekte besser wahrzunehmen bzw. mehr für sich anzuerkennen und so die Aufrechterhaltung des Sports nach der Interventionsphase wahrscheinlicher zu machen.

Der Arbeitsmarkt verändert sich zusehends durch sich beschleunigende wirtschaftliche Transformationen. Nach dem Wandel von der Agrar-, zur Industrie- und zur Dienstleistungsgesellschaft spricht man mittlerweile von der Informationsgesellschaft. Der technische Fortschritt, dadurch entstehende Automatisierungsprozesse und die Globalisierung - mitsamt ihrer wirtschaftlichen Krisen - beschleunigen die Veränderung der Arbeitswelt. Die weitere Verlegung von teilweise kompletten Berufssparten ins Ausland wird wohl auch künftig zu irreversiblen Verlusten der damit verbundenen Arbeitsplätze führen. Die Erfahrungen der vergangenen Jahrzehnte zeigte, dass vor allem der Bedarf nach eher bildungsfernen Arbeitsmöglichkeiten kontinuierlich sinkt. Da Bildung negativ über einen Kohorteneffekt mit dem Alter verknüpft ist, haben ältere Menschen beim Verlust ihrer Arbeit besonders schlechte Chancen auf einen Wiedereinstieg. Neben denen im Theorieteil dargestellten Risikobedingungen älterer, erwerbsloser Menschen steigt nach einer Wiedereinstellung die Gefahr eines erneuten Arbeitsplatzverlustes heutzutage auch dadurch, dass neue Arbeitsmöglichkeiten mit steigender Tendenz entweder befristet oder im Rahmen von Arbeitnehmerüberlassungen angeboten werden. Dazu gesellt sich der gesetzliche Beschluss, das Rentenalter auf 67 anzuheben. Von diesen Rahmenbedingungen kann sämtlich angenommen werden, dass zukünftig mit einer weiteren Zunahme von älteren Langzeiterwerbslosen zu rechnen ist, womit ein erhöhter Handlungsbedarf entsteht.

Auf der gesundheitlichen Ebene wurde im Theorieteil dieser Arbeit der empirische Beweis von Paul und Moser (2009a) nachgezeichnet, dass eine Erwerbslosigkeit nicht nur mit psychischen Problemen zusammenhängt, sondern sie auch kausal auslöst. Hierdurch ist ein

Bedarf an effektiven Interventionen gegeben und der Auftrag an die Präventionsforschung muss lauten, mithilfe wissenschaftlicher Evaluationen möglichst effektive gesundheitsförderliche Programme zu identifizieren und sie weiter zu entwickeln. Hierbei sollten auch bislang ungewöhnliche, neue Interventionsmöglichkeiten exploriert werden. Ein solch neuer Weg stellt sicherlich die Linderung von psychischen Beschwerden älterer Langzeiterwerbsloser mithilfe eines körperlichen Bewegungsprogramms dar, denn die Anwendung von Sport wäre auf den ersten Blick sicherlich nicht die Methode der ersten Wahl gewesen, wenn es hauptsächlich um die psychische Stabilisierung der Zielgruppe gegangen wäre. Obwohl die körperliche Aktivierung primär das Ziel verfolgte, die physische Gesundheit zu fördern, konnte sie sekundär auch ansehnliche psychische Interventionseffekte produzieren und sich in der Effektstärke auch mit Programmen, die auf die seelische Stabilisierung von Erwerbslosen spezialisiert sind, messen lassen. Eine Kombination von psychischen und physischen Interventionswegen sollte für die Präventionsforschung als der nächste logische Schritt für die Zielgruppe angesehen werden. Die oben angerissenen Herausforderungen einer sich stetig ändernden Arbeitswelt zu Ungunsten von erwerbsfähigen Menschen ab 50 Jahren sollten als Argument dafür ausreichend sein, dass weitere Forschungsprojekte in diesem Bereich nicht nur wissenschaftlich interessant, sondern auch gesellschaftlich relevant wären.

6 Literatur

- Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. In N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 111-119). CA, Thousand Oaks.
- AGF (2010). Entwicklung der durchschnittlichen Sehdauer pro Tag/Person in Minuten. Retrieved 25.10.2010, from <http://agf.de/daten/zuschauermarkt/sehdauer/>
- Altman, D. G., & Bland, J. M. (1995). The normal distribution. *BMJ*, 310, 298.
- Amundson, N. E., & Borgen, W. A. (1982). The dynamics of unemployment: Job loss and job search. *Personnel & Guidance Journal*, 60(9), 562-564.
- Amundson, N. E., & Borgen, W. A. (1987). Coping with unemployment: What helps and what hinders. *Journal of Employment Counseling*, 24(3), 97-106.
- Antonovsky, A. (1987). *Unraveling the mystery of health. How people manage stress and stay well*. San Francisco: Jossey-Bass.
- Antonovsky, A., & Franke, A. (1997). *Salutogenese. Zur Entmystifizierung der Gesundheit*. Tübingen: DGVT-Verl.
- Aoyagi, Y., Park, H., Park, S., & Shephard, R. J. (2010). Habitual physical activity and health-related quality of life in older adults: interactions between the amount and intensity of activity (the Nakanojo Study). *Quality of Life Research*, 19(3), 333-338.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Beck, A. T., Steer, R. A., & Hautzinger, M. (2005). *Beck-Depressions-Inventar (BDI). Testhandbuch* (2. Aufl.). Bern: Huber.
- Becker, P. (1997). *Psychologie der seelischen Gesundheit. Theorien, Modelle, Diagnostik* (2. Aufl.). Göttingen: Hogrefe.
- Berth, H., Förster, P., Balck, F., Brähler, E., & Stöbel-Richter, Y. (2008). Arbeitslosigkeitserfahrungen, Arbeitsplatzunsicherheit und der Bedarf an psychosozialer Versorgung. *Gesundheitswesen*, 70(5), 289-294.
- BKK (2009). *BKK Gesundheitsreport 2009*. Essen: BKK Bundesverband.
- Bland, J. M., & Altman, D. G. (2000). Statistics notes: The odds ratio. *BMJ*, 320, 1468.
- Blumenthal, J. A., Babyak, M. A., Moore, K. A., Craighead, W. E., Herman, S., Khatri, P., et al. (1999). Effects of exercise training on older patients with major depression. *Archives of Internal Medicine*, 159(19), 2349-2356.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. (6. Aufl.). Berlin; Heidelberg; New York: Springer.
- Bredenkamp, J. (1969). The application of significance tests in theory-testing experiments. *Psychologische Beiträge*, 11(3), 275-285.
- Broom, D. H., D'Souza, R. M., Strazdins, L., Butterworth, P., Parslow, R., & Rodgers, B. (2006). The lesser evil: bad jobs or unemployment? A survey of mid-aged Australians. *Social Science & Medicine*, 63(3), 575-586.
- Brown, D. W., Balluz, L. S., Ford, E. S., Giles, W. H., Strine, T. W., Moriarty, D. G., et al. (2003). Associations between Short- and Long-Term Unemployment and Frequent Mental Distress among a National Sample of Men and Women. *Journal of Occupational and Environmental Medicine*, 45(11), 1159-1166.

- Bullinger, M., & Kirchberger, I. (1998). *SF-36. Fragebogen zum Gesundheitszustand - Handanweisung*. Göttingen: Hogrefe.
- Busch, A. J., Barber, K. A., Overend, T. J., Peloso, P. M., & Schachter, C. L. (2007). Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews*(4).
- Claussen, B. (1994). Psychologically and biochemically assessed stress in a follow-up study of long-term unemployed. *Work & Stress*, 8(1), 4-18.
- Claussen, B. (1999). Alcohol disorders and re-employment in a 5-year follow-up of long-term unemployed. *Addiction*, 94(1), 133-138.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cramer, R., Gilberg, R., Hess, D., Marwinski, K., Schröder, H., & Smid, M. (2002). *Suchintensität und Einstellungen Arbeitsloser*. Nürnberg: IAB.
- Creed, P. A., Machin, M. A., & Hicks, R. E. (1996). Neuroticism and mental health outcomes for long-term unemployed youth attending occupational skills training programs. *Personality and Individual Differences*, 21(4), 537-544.
- Creed, P. A., Machin, M. A., & Hicks, R. E. (1999). Improving Mental Health Status and Coping Abilities for Long-Term Unemployed Youth Using Cognitive-Behaviour Therapy Based Training Interventions. *Journal of organizational behavior*, 20(6), 963-978.
- Depp, C. A., Schkade, D. A., Thompson, W. K., & Jeste, D. V. (2010). Age, Affective Experience, and Television Use. *American Journal of Preventive Medicine*, 39(2), 173-179.
- Derogatis, L. R., & Cleary, P. A. (1977). Factorial invariance across gender for the primary symptom dimensions of the SCL-90. *British Journal of Social & Clinical Psychology*, 16(4), 347-356.
- Emery, C. F., Hauck, E. R., & Blumenthal, J. A. (1992). Exercise adherence or maintenance among older adults: 1-year follow-up study. *Psychology and Aging*, 7(3), 466-470.
- Engelbrech, G. (1987). Erwerbsverhalten und Berufsverlauf von Frauen. Ergebnisse neuerer Untersuchungen im Überblick. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 20(2), 181-196.
- Ernst, S., Hartwig, U., & Pokora, F. (2008). Social Scoring. Evaluation qualifizierender Beschäftigung. *WSI Mitteilungen*. Heft 5 (pp. 267-273).
- Fahrenberg, J., Hampel, R., & Selg, H. (2010). *Freiburger Persönlichkeitsinventar (FPI-R); Manual* (8. Aufl.). Göttingen: Hogrefe.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Festinger, L. (2001). *A theory of cognitive dissonance* (Repr. ed.). Stanford: Stanford University Press.
- Filipp, S.-H., & Aymanns, P. (2010). *Kritische Lebensereignisse und Lebenskrisen. Vom Umgang mit den Schattenseiten des Lebens*. Stuttgart: Kohlhammer.

- Foy, C. G., Rejeski, W. J., Berry, M. J., Zaccaro, D., & Woodard, C. M. (2001). Gender moderates the effects of exercise therapy on health-related quality of life among COPD patients. *Chest*, 119(1), 70-76.
- Frese, M. (1987). Alleviating depression in the unemployed: Adequate financial support, hope and early retirement. *Social Science & Medicine*, 25(2), 213-215.
- Frese, M., Garman, G., Garmeister, K., Halemba, K., Hortig, A., Pulwitt, T., et al. (2002). Training to increase personal initiative of the unemployed: Report on a pilot study. *Zeitschrift für Arbeits und Organisationspsychologie*, 46(2), 89-97.
- Frese, M., & Mohr, G. (1978). Die psychopathologischen Folgen des Entzugs von Arbeit: Der Fall Arbeitslosigkeit. In M. Frese (Ed.), *Industrielle Psychopathologie* (pp. 343). Bern: Huber.
- Frese, M., & Mohr, G. (1987). Prolonged unemployment and depression in older workers: A longitudinal study of intervening variables. *Social Science & Medicine*, 25(2), 173-178.
- Fryer, D. (1986). Employment deprivation and personal agency during unemployment: A critical discussion of Jahoda's explanation of the psychological effects of unemployment. *Social Behaviour*, 1(1), 3-23.
- Fryer, D., & McKenna, S. (1987). The laying off hands - Unemployment and the experience of time. In S. Fineman & A. Sinfield (Eds.), *Unemployment. Personal and social consequences. Social science paperbacks*. (pp. 260). London: Tavistock.
- Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Goldberg, D. P. (1978). *Manual for the General Health Questionnaire*. Windsor: National Foundation for Educational Research.
- Goldberg, D. P., Gater, R., Sartorius, N., Üstün, T. B., Picinelli, M., Gureje, O., et al. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27(1), 191-198.
- Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9(1), 139-145.
- Grawe, K. (2004). *Psychological therapy*. Cambridge: Hogrefe & Huber.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.
- Greiner, W., & Claes, C. (2007). Der EQ-5D der EuroQol-Gruppe. *Gesundheitsökonomische Evaluationen* (pp. 403-414). Berlin: Springer.
- Hager, W. (2006). Die Fallibilität empirischer Daten und die Notwendigkeit der Kontrolle der Wahrscheinlichkeiten falscher Entscheidungen *Zeitschrift für Psychologie*, 214(1), 10-23.
- Hamilton, R. (2010). Ockham's Razor or Procrustes' Axe? Why we should reject philosophical speculation that ignores fact. *Journal of evaluation in clinical practice*, 16(2), 282-284.
- Hammarstrom, A., & Janlert, U. (1997). Nervous and depressive symptoms in a longitudinal study of youth unemployment--selection or exposure? *Journal of Adolescence*, 20(3), 293-305.
- Hill, J. (1978). The psychological impact of unemployment. *New Society*, 19, 118-120.
- Hinz, A., Klaiberg, A., Brähler, E., & König, H.-H. (2006). The Quality of Life Questionnaire EQ-5D: Modelling and Norm Values for the General Population. *Psychotherapie, Psychosomatik, medizinische Psychologie*, 56(2), 42-48.

- Hinz, A., Schumacher, J., Albani, C., Schmid, G., & Brähler, E. (2006). Bevölkerungsrepräsentative Normierung der Skala zur Allgemeinen Selbstwirksamkeitserwartung. *Diagnostica*, 52(1), 26-32.
- Hollederer, A. (2002). Aufsätze - Arbeitslosigkeit und Gesundheit. Ein Überblick über empirische Befunde und die Arbeitslosen- und Krankenkassenstatistik. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 35(3), 411-428.
- Hollederer, A. (2009). Gesundheit und Krankheit von Arbeitslosen sowie Chancen und Grenzen arbeitsmarktintegrativer Gesundheitsförderung. In A. Holleder (Ed.), *Gesundheit von Arbeitslosen fördern* (pp. 12-38). Frankfurt/Main: Fachhochschulverl.
- Hollederer, A., & Brand, H. (Eds.). (2006). *Arbeitslosigkeit, Gesundheit und Krankheit*. Bern: Huber.
- Hradil, S., & Schiener, J. (2001). *Soziale Ungleichheit in Deutschland*. Opladen: Leske & Budrich.
- Hsueh, Y. (2002). The Hawthorne Experiments and the Introduction of Jean Piaget in American Industrial Psychology, 1929-1932. *History of psychology*, 5(2), 163-189.
- IAB (1999). Langzeitarbeitslosigkeit früh erkennen gezielt bekämpfen. *IAB-Materialien* 1, 8-9.
- Jackson, P. R. (1988). Personal networks, support mobilization and unemployment. *Psychological Medicine*, 18(2), 397-404.
- Jackson, P. R., & Warr, P. B. (1984). Unemployment and psychological ill-health: The moderating role of duration and age. *Psychological Medicine*, 14(3), 605-614.
- Jahoda, M. (1982). *Employment and unemployment. A socialpsychological analysis*. Cambridge M: Cambridge University Press.
- Jahoda, M. (1983). Die sozialpsychologische Bedeutung von Arbeit und Arbeitslosigkeit. In M. Jahoda, T. Kieselbach, T. Leithäuser (Eds.), *Arbeit, Arbeitslosigkeit und Persönlichkeitsentwicklung*. Bremen: Universität, Fachbereich 9, 1-8.
- Jahoda, M. (1992). Reflections on Marienthal and after. *Journal of Occupational and Organizational Psychology*, 65(4), 355-358.
- Jahoda, M., Lazarsfeld, P. F., & Zeisel, H. (1933). *Die Arbeitslosen von Marienthal. Ein soziographischer Versuch über die Wirkungen langandauernder Arbeitslosigkeit*. Frankfurt/Main: Suhrkamp. [zitiert nach Paul und Moser, 2009b].
- Janssen, J., & Laatz, W. (2010). *Statistische Datenanalyse mit SPSS. Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests* (7. Aufl.). Berlin: Springer.
- Jerusalem, M., & Schwarzer, R. (1984). Selbstwirksamkeit. In R. Schwarzer (Ed.), *Skalen zur Befindlichkeit und Persönlichkeit* (pp. 15-28). Berlin: Institut für Psychologie, Freie Universität Berlin.
- Jette, A. M., Harris, B. A., Sleeper, L., Lachman, M. E., Heislein, D., Giorgetti, M., et al. (1996). A home-based exercise program for nondisabled older adults. *Journal of the American Geriatrics Society*, 44(6), 644-649.
- Jones, L. (1991). Specifying the temporal relationship between job loss and consequences: Implication for service delivery. *Journal of Applied Social Sciences*, 16(1), 1991-1992.
- Kessler, R. C., Soukup, J., Davis, R. B., Foster, D. F., Wilkey, S. A., Van Rompay, M. I., et al. (2001). The use of complementary and alternative therapies to treat anxiety and depression in the United States. *American Journal of Psychiatry*, 158(2), 289-294.
- Kieselbach, T., Winefield, A. H., & Boyd, C. (2006). *Unemployment and Health. International and interdisciplinary perspectives*. Bowen Hills, Qld.: Australian Academic Press.

- Klein, G., & Strasser, H. (Eds.). (1997). *Schwer vermittelbar. Zur Theorie und Empirie der Langzeitarbeitslosigkeit*. Opladen: Westdeutscher Verlag.
- Klein, T. (1990). Arbeitslosigkeit und Wiederbeschäftigung im Erwerbsverlauf. Theorieansätze und empirische Befunde. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 42(4), 688-705.
- Klonoff, E. A., Annechild, A., & Landrine, H. (1994). Predicting exercise adherence in women: The role of psychological and physiological factors. *Preventive Medicine*, 23(2), 257-262.
- Komarovsky, M. (1940). *The unemployed man and his family; the effect of unemployment upon the status of the man in fifty-nine families*. New York: Dryden Press.
- Kuhl, J. (2001). *Motivation und Persönlichkeit. Interaktionen psychischer Systeme*. Göttingen: Hogrefe.
- Kuhner, C., Burger, C., Keller, F., & Hautzinger, M. (2007). Reliability and validity of the Revised Beck Depression Inventory (BDI-II). Results from German samples. *Nervenarzt*, 78(6), 651-656.
- Lahelma, E. (1989). Unemployment, re-employment and mental well-being. A panel survey of industrial jobseekers in Finland. *Scandinavian Journal of Social Medicine*, 43 Suppl, 1-170.
- Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *BMJ*, 322(7289), 763-767.
- Lazarus, R. S., & Alfert, E. (1964). Short-Circuiting of Threat by Experimentally Altering Cognitive Appraisal. *Journal of Abnormal Psychology*, 69, 195-205.
- Leana, C. R., & Feldman, D. C. (1991). Gender differences in responses to unemployment. *Journal of Vocational Behavior*, 38(1), 65-77.
- Leana, C. R., & Feldman, D. C. (1992). *Coping with job loss. How individuals, organizations, and communities respond to layoffs*. New York: Lexington Books.
- Leana, C. R., & Feldman, D. C. (1995). Finding new jobs after a plant closing: Antecedents and outcomes of the occurrence and quality of reemployment. *Human Relations*, 48(12), 1381-1401.
- Lee, D. C., Sui, X., & Blair, S. N. (2009). Does physical activity ameliorate the health hazards of obesity? *British Journal of Sports Medicine*, 43(1), 49-51.
- Linden, M., Maier, W., Achberger, M., Herr, R., Helmchen, H., & Benkert, O. (1996). Psychische Erkrankungen und ihre Behandlung in Allgemeinarztpraxen in Deutschland. Ergebnisse aus einer Studie der Weltgesundheitsorganisation (WHO). *Der Nervenarzt*, 67(3), 205-215.
- Manz, R. (1998). *KÖPS. Ein Fragebogen für körperliche, psychische und soziale Symptome: Manual*. Frankfurt/Main: Swets Test Services.
- Markland, D., & Ingledew, D. K. (1997). The measurement of exercise motives: Factorial validity and invariance across gender of a revised Exercise Motivations Inventory. *British Journal of Health Psychology*, 2(4), 361-376.
- Mastekaasa, A. (1996). Unemployment and health: Selection effects. *Journal of Community & Applied Social Psychology*, 6(3), 189-205.
- Mauchly, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution. *The Annals of Mathematical Statistics*, 11(2), 204-209.

- Maysent, M., & Spera, S. (1995). Coping with job loss and career stress: Effectiveness of stress management training with outplaced employees. In *Job stress interventions* (pp. 159-170). Washington DC: American Psychological Association.
- McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of Behavioral Medicine*, 16(1), 103-113.
- McAuley, E., Lox, C., & Duncan, T. E. (1993). Long-term maintenance of exercise, self-efficacy, and physiological change in older adults. *Journals of Gerontology*, 48(4).
- McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of Applied Psychology*, 90(1), 53-76.
- McQueen, A., Vernon, S. W., Meissner, H. I., Klabunde, C. N., & Rakowski, W. (2006). Are there gender differences in colorectal cancer test use prevalence and correlates? *Cancer Epidemiology Biomarkers and Prevention*, 15(4), 782-791.
- Mielck, A. (2005). *Soziale Ungleichheit und Gesundheit. Einführung in die aktuelle Diskussion*. Bern: Huber.
- Mohr, G. (1993). Industriebearbeiterinnen sieben Jahre später: Arbeitssuchende Frauen im Vergleich mit wieder erwerbstätigen Frauen und Hausfrauen. In G. Mohr (Ed.), *Ausgezählt. Theoretische und empirische Beiträge zur Psychologie der Frauenerwerbslosigkeit* (pp. 187-204). Weinheim: Deutscher Studien Verlag.
- Mohr, G. (2001). Langzeiterwerbslosigkeit. In J. Zempel, J. Bacher & K. Moser (Eds.), *Erwerbslosigkeit. Ursachen, Auswirkungen und Interventionen*. (pp. 111-131). Opladen: Leske & Budrich.
- Mohr, G., & Otto, K. (2005). Langzeiterwerbslosigkeit: Welche Interventionen machen aus psychologischer Sicht Sinn? *Zeitschrift fuer Psychotraumatologie und Psychologische Medizin*, 3(4), 45-63.
- Murphy, G. C., & Athanasou, J. A. (1999). The effect of unemployment on mental health. *Journal of Occupational and Organizational Psychology*, 72(1), 83-99.
- Nitsche, I., & Richter, P. (2003). *Tätigkeiten ausserhalb der Erwerbsarbeit. Evaluation des TAURIS-Projektes*. Münster: Lit.
- Olbrich, E. (1989). Erfolgreiches Altern aus funktionalistischer und interpretativer Perspektive. Successful aging from a functional and interpretative perspective. *Zeitschrift für Gerontopsychologie und -psychiatrie*, 2, 314-318.
- Paul, K. I., Geithner, E., & Moser, K. (2009). Latent deprivation among people who are employed, unemployed, or out of the labor force. *The Journal of Psychology*, 143(5), 477-491.
- Paul, K. I., Hassel, A., & Moser, K. (2006). Die Auswirkungen von Arbeitslosigkeit auf die psychische Gesundheit. Befunde einer quantitativen Forschungsintegration In A. Holleederer & H. Brand (Eds.), *Arbeitslosigkeit, Gesundheit und Krankheit* (pp. 35-51). Bern: Huber.
- Paul, K. I., & Moser, K. (2001). Negatives psychisches Befinden als Wirkung und als Ursache von Arbeitslosigkeit: Ergebnisse einer Metaanalyse. In J. Zempel, J. Bacher & K. Moser (Eds.), *Erwerbslosigkeit. Ursachen, Auswirkungen und Interventionen*. (pp. 83-110). Opladen: Leske & Budrich.
- Paul, K. I., & Moser, K. (2006). Incongruence as an explanation for the negative mental health effects of unemployment: Meta-analytic evidence. *Journal of Occupational and Organizational Psychology*, 79(4), 595-622.
- Paul, K. I., & Moser, K. (2009a). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, 74(3), 264-282.

- Paul, K. I., & Moser, K. (2009b). Metaanalytische Moderatorenanalysen zu den psychischen Auswirkungen der Arbeitslosigkeit. Ein Überblick. In A. Holleder (Ed.), *Gesundheit von Arbeitslosen fördern* (pp. 39-61). Frankfurt/Main: Fachhochschulverlag.
- Paul, K. I., & Moser, K. (2009c). Wie wirken sich Interventionsmaßnahmen auf die psychische Gesundheit Arbeitsloser aus, wenn die Stellensuche erfolglos bleibt?. Erste Befunde einer Metaanalyse. In A. Holleder (Ed.), *Gesundheit von Arbeitslosen fördern* (pp. 124-134). Frankfurt/Main: Fachhochschulverlag.
- Pernice, R. (1996). Methodological issues in unemployment research: Quantitative and/or qualitative approaches? *Journal of Occupational and Organizational Psychology*, 69(4), 339-349.
- Piccinelli, M., Bisoffi, G., Bon, M. G., Cunico, L., & Tansella, M. (1993). Validity and test-retest reliability of the Italian version of the 12-item General Health Questionnaire in general practice: a comparison between three scoring methods. *Comprehensive Psychiatry*, 34(3), 198-205.
- Plöger, P. (2010). *Arbeits-sammler, Jobnomaden und Berufsartisten. Viel gelernt und nichts gewonnen? Das Paradox der neuen Arbeitswelt*. München: Hanser.
- Prochaska, J. O., & Velicer, W. F. (1997). The Transtheoretical Model of Health Behavior Change. *American journal of health promotion*, 12(1), 38-48.
- Proudfoot, J., Gray, J., Carson, J., Guest, D., & Dunn, G. (1999). Psychological training improves mental health and job-finding among unemployed people. *International Archives of Occupational and Environmental Health*, 72 Suppl, 40-42.
- Proudfoot, J., Guest, D., Carson, J., Dunn, G., & Gray, J. (1997). Effect of cognitive-behavioural training on job-finding among long-term unemployed people. *Lancet*, 350(9071), 96-100.
- Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2010). *Quantitative Methoden Band 2. Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Berlin: Springer.
- Ravens-Sieberer, U. (2002). Measuring and monitoring quality-of-life in population surveys: still a challenge for public health research. *Sozial- und Präventivmedizin*, 47(4), 203-204.
- Rengers, M. (2005). Das international vereinbarte Labour-Force-Konzept *destatis - Auszug aus Wirtschaft und Statistik* (pp. 1369-1383). Wiesbaden: Statistisches Bundesamt.
- Resnick, B. (2001). A prediction model of aerobic exercise in older adults living in a continuing-care retirement community. *Journal of Aging and Health*, 13(2), 287-310.
- Rhodes, R. E., Martin, A. D., Taunton, J. E., Rhodes, E. C., Donnelly, M., & Elliot, J. (1999). Factors associated with exercise adherence among older adults. An individual perspective. *Sports Medicine*, 28(6), 397-411.
- Richter, P., Werner, J., Bastine, R., Heerlein, A., Kick, H., & Sauer, H. (1997). Measuring treatment outcome by the Beck Depression Inventory. *Psychopathology*, 30(4), 234-240.
- Rife, J. C., & Belcher, J. R. (1993). Social support and job search intensity among older unemployed workers: Implications for employment counselors. *Journal of Employment Counseling*, 30(3), 98-107.
- Rimann, M., & Udris, I. (1993). *Belastungen und Gesundheitsressourcen im Berufs- und Privatbereich. Eine quantitative Studie. Forschungsprojekt SALUTE*. Zürich: Institut für Arbeitspsychologie Eidgenössische Technische Hochschule.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

- Rothländer, K., Mühlfordt, S., & Richter, P. (2012). Evaluation des Gesundheitsförderungsprogramms "Aktive Bewältigung von Arbeitslosigkeit (AktivA)". *Zeitschrift für Gesundheitspsychologie*, 20(3), 115-127.
- Rothländer, K., & Richter, P. (2009). *Gesund und mittendrin trotz Erwerbslosigkeit?! Ansätze zur Förderung der psycho-sozialen Gesundheit*. Berlin: Friedrich-Ebert-Stiftung.
- Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology*, 43(1), 56-67.
- Rudolf, M., & Müller, J. (2004). *Multivariate Verfahren. Eine praxisorientierte Einführung mit Anwendungsbeispielen in SPSS*. Göttingen: Hogrefe.
- Saam, R. H., & Wodtke, K. H. (1995). A cognitive stress reduction program for recently unemployed managers. *Career Development Quarterly*, 44(1), 43-51.
- Schmitt, A., Beckmann, M., Dusi, D., Maes, J., Schiller, A., & Schonauer, K. (2003). Messgüte des vereinfachten Beck-Depressions-Inventars (BDI-V). *Diagnostica*, 49(4), 147-156.
- Schmitt, M., Altstotter-Gleich, C., Hinz, A., Maes, J., & Brahler, E. (2006). Normwerte für das Vereinfachte Beck-Depressions-Inventar (BDI-V) in der Allgemeinbevölkerung. *Diagnostica*, 52(2), 51-59.
- Schober, K. (1978). Arbeitslose Jugendliche. Belastungen und Reaktionen der Betroffenen ; aus der Untersuchung des IAB über Ursachen und Auswirkungen der Arbeitslosigkeit. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 11(2), 198-215.
- Schumacher, J., & Brähler, E. (2000). Sense of Coherence Scale (SOC-Skala). Retrieved 01.03.2010 from http://www.praxis-schumacher.net/material/soc_scales.pdf.
- Schumacher, J., Wilz, G., Gunzelmann, T., & Brähler, E. (2000). Antonovsky's Sense of Coherence Scale - Its Validation in a Population-Based Sample and the Development of a New Short Scale. *Psychotherapie, Psychosomatik, medizinische Psychologie*, 50(12), 472.
- Schwarzer, R. (1994). Generalized self-efficacy: Assessment of a personal coping resource. *Diagnostica*, 40(2), 105-123.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, M. Wright & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-Nelson.
- Schwarzer, R., & Jerusalem, M. (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.
- Seligman, M. E. (1978). Learned helplessness as a model of depression. Comment and integration. *Journal of Abnormal Psychology*, 87(1), 165-179.
- Seligman, M. P. (1991). *Learned optimism*. New York: A.A. Knopf.
- Shamir, B. (1985). Sex differences in psychological adjustment to unemployment and re-employment: A question of commitment, alternatives or finance? *Social Problems*, 33(1), 67-79.
- Shlay, A. B., & Rossi, P. H. (1992). Social science research and contemporary studies of homelessness. *Annual Review of Sociology*, 18, 129-160.
- Spielberger, C. D. (1983). *STAI State-trait anxiety inventory for adults Form Y. Review set; manual, test, scoring key*. Redwood City, CA: Mind Garden.

- STABU (2006). *Im Blickpunkt: Deutschland in der EU 2006*. Wiesbaden: Statistisches Bundesamt.
- Stiggelbout, M., Hopman-Rock, M., Crone, M., Lechner, L., & Van Mechelen, W. (2006). Predicting older adults' maintenance in exercise participation using an integrated social psychological model. *Health Education Research*, 21(1), 1-14.
- Strawbridge, W. J., Deleger, S., Roberts, R. E., & Kaplan, G. A. (2002). Physical activity reduces the risk of subsequent depression for older adults. *American Journal of Epidemiology*, 156(4), 328-334.
- Strehmel, P., & Mayring, P. (1986). Arbeitslosigkeit: Belastungen und kognitive Verarbeitung: Ausgewählte Ergebnisse aus einer Längsschnittuntersuchung. *Augsburger Berichte zur Entwicklungspsychologie und Pädagogischen Psychologie* (Bd. 2). Augsburg: Forschungsstelle.
- Sui, X., La Monte, M. J., Laditka, J. N., Hardin, J. W., Chase, N., Hooker, S. P., & Blair, S. N. (2007). Cardiorespiratory Fitness and Adiposity as Mortality Predictors in Older Adults. *The Journal of the American Medical Association*, 289(21), 2507-2516.
- TK (2010). *Gesundheitsreport 2010. Gesundheitliche Veränderungen bei Berufstätigen und Arbeitslosen von 2000 bis 2009* (Bd. 24). Hamburg: TK.
- Trube, A., & Luschei, F. (2000). *Entwicklungs- und Vermittlungs-Assistenz (EVA) - ein Instrument zur Wiedereingliederung Langzeitarbeitsloser*. Siegen: Zentrum für Planung und Evaluation Sozialer Dienste.
- van Gool, C. H., Kempen, G. I., Bosma, H., van Boxtel, M. P., Jolles, J., & van Eijk, J. T. (2007). Associations between lifestyle and depressed mood: longitudinal results from the Maastricht Aging Study. *American Journal of Public Health*, 97(5), 887-894.
- Vastamaki, J., Moser, K., & Paul, K. I. (2009). How stable is sense of coherence? Changes following an intervention for unemployed individuals. *Scandinavian Journal of Psychology*, 50(2), 161-171.
- Verkleij, H. (1989). Vulnerabilities of very long term unemployed in the Netherlands: Results of a longitudinal study. In B. Starrin (Ed.), *Unemployment, poverty, and quality of working life. Some European experiences* (pp. 366). Berlin: Ed. Sigma.
- Vinokur, A. D., Price, R. H., & Schul, Y. (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. *American Journal of Community Psychology*, 23(1), 39-74.
- Vinokur, A. D., Schul, Y., Vuori, J., & Price, R. H. (2000). Two years after a job loss: long-term impact of the JOBS program on reemployment and mental health. *Journal of Occupational Health Psychology*, 5(1), 32-47.
- Vinokur, A. D., & Van Ryn, M. (1993). Social support and undermining in close relationships: Their independent effects on the mental health of unemployed persons. *Journal of Personality and Social Psychology*, 65(2), 350-359.
- Vuori, J., & Silvonon, J. (2005). The benefits of a preventive job search program on reemployment and mental health at 2-year follow-up. *Journal of Occupational and Organizational Psychology*, 78(1), 43-52.
- Vuori, J., & Vesalainen, J. (1999). Labour market interventions as predictors of reemployment, job seeking activity and psychological distress among the unemployed. *Journal of Occupational and Organizational Psychology*, 72(4), 523-538.
- Wanberg, C. R. (1995). A longitudinal study of the effects of unemployment and quality of reemployment. *Journal of Vocational Behavior*, 46(1), 40-54.
- Warr, P., & Jackson, P. (1985). Factors influencing the psychological impact of prolonged unemployment and of re-employment. *Psychological Medicine*, 15(4), 795-807.

- Warr, P., & Jackson, P. (1987). Adapting to the unemployed role: a longitudinal investigation. *Social Science & Medicine*, 25(11), 1219-1224.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur psychologischen Methodenlehre*. Göttingen: Hogrefe.
- Wiggert, A. (2009). *Kosten von Arbeitslosigkeit*. Hamburg: Igel Verlag.
- Williams, P., & Lord, S. R. (1995). Predictors of adherence to a structured exercise program for older women. *Psychology and Aging*, 10(4), 617-624.
- Winefield, A. H. (1995). Unemployment: Its psychological costs. In C. Cooper (Ed.), *International Review of Industrial and Organizational Psychology*. Chichester: Wiley & Sons.
- Winegardner, D., Simonetti, J. L., & Nykodym, N. (1984). Unemployment: The living death? *Journal of Employment Counseling*, 21(4), 149-155.
- Winterhoff-Spurk, P. (2005). *Kalte Herzen. wie das Fernsehen unseren Charakter formt* (2. Aufl.). Stuttgart: Klett-Cotta.
- Wohlwill, J. F. (1977). *Strategien entwicklungspsychologischer Forschung*. Stuttgart: Klett-Cotta.
- Zawadski, B., & Lazarsfeld, P. (1935). The psychological consequences of unemployment. *Journal of Social Psychology*, 6, 224-251. [zitiert nach Paul und Moser, 2009b].
- Zempel, J., & Frese, M. (2000). Prädiktoren der Erwerbslosigkeit und Wiederbeschäftigung. *Verhaltenstherapie und psychosoziale Praxis*, 32(3), 379-390.
- Zippay, A. (1995). Tracing behavioral changes among discouraged workers: What happens to the work ethic? *Psychological Reports*, 76(2), 531-543.
- Zöfel, P. (2003). *Statistik für Psychologen*. München: Pearson Studium.

7 Anhang

7.1 Einwilligungserklärung

Institut für Präventivmedizin der Universität Rostock,
St.-Georg-Str. 108, 18055 Rostock, Tel.: 0381/494-9957

Probandeneinwilligung

„Evaluation von Maßnahmen zur Verbesserung der physischen und psychischen Arbeitsfähigkeit älterer Langzeitarbeitsloser“

Ich, _____, wurde von meinem Arzt vollständig über Wesen, Bedeutung und Tragweite der geplanten Untersuchungen mit dem o. g. Titel aufgeklärt. Ich habe den Aufklärungstext gelesen und verstanden. Ich hatte die Möglichkeit, Fragen zu stellen, und habe die Antworten verstanden und akzeptiere sie. Mein Arzt hat mich über die mit der Teilnahme an der Studie verbundenen Risiken und den möglichen Nutzen informiert.

Ich hatte ausreichend Zeit, mich zur Teilnahme an dieser Studie zu entscheiden und weiß, dass die Teilnahme an diesen Untersuchungen freiwillig ist. Ich weiß, dass ich jederzeit und ohne Angabe von Gründen diese Zustimmung widerrufen kann, ohne dass dadurch Nachteile für mich entstehen.

Ich bin darüber informiert worden, dass bei Untersuchungen zur körperlichen Leistungsfähigkeit (Fahrradergometrien) an jungen, gesunden Probanden und Sportlern bisher keine unerwünschten Effekte beobachtet worden sind. Mir ist bekannt, dass mit zunehmendem Alter das Risiko für unerwünschte Effekte bei Leistungsuntersuchungen ansteigt und dass alle Angaben zu meinem Gesundheitszustand von entscheidender Bedeutung für die risikoarme Durchführung der Untersuchungen sind.

Die Fragen zum eigenen Gesundheitszustand und zu etwaigen Vorerkrankungen habe ich wahrheitsgemäß beantwortet.

Risiken, die im Zusammenhang mit den Untersuchungen zu diesem Forschungsvorhaben stehen, werden durch den Versicherungsschutz des Klinikums gedeckt.

Einwilligungserklärung zum Datenschutz

Bei wissenschaftlichen Untersuchungen werden persönliche Daten und medizinische Befunde über Sie erhoben. Die Erhebung, Weitergabe, Speicherung und Auswertung dieser Angaben über Ihre Gesundheit erfolgt nach gesetzlichen Bestimmungen, d.h. ohne die nachfolgende Einwilligung können Sie nicht an der Untersuchung teilnehmen.

Mir ist bekannt, dass meine persönlichen Daten in verschlüsselter (anonymisierter) Form gespeichert werden. Mir wurde versichert, dass dabei die Grundsätze des Datenschutzes beachtet werden, d.h. meine personenbezogenen Daten vertraulich behandelt und nicht an andere weitergegeben werden.

Mit meinem Einverständnis zur Teilnahme an den Untersuchungen erkläre ich gleichzeitig mein Einverständnis für die wissenschaftliche Veröffentlichung der Ergebnisse unter Beachtung datenschutzrechtlicher Bestimmungen.

Ich habe eine Kopie der Patienteninformation und dieser Einwilligungserklärung erhalten. Ich erkläre hiermit meine freiwillige Teilnahme an dieser Studie.

Ort, Datum

Unterschrift des Patienten/Probanden

7.2 Soziodemographischer Fragebogen

- 1. Welchen Familienstand haben Sie?**
- Verheiratet bzw. dauerhaft mit einem Partner zusammenlebend
 - geschieden bzw. dauerhaft vom Partner getrennt lebend
 - verwitwet
 - ledig

2. Haben Sie Kinder?

- Nein
- Ja → Wenn ja, tragen Sie bitte in folgende Tabelle die Anzahl der Kinder ein und wie viele davon noch in Ihrem Haushalt leben.

	Anzahl der Kinder	davon noch im Haushalt
Kinder bis 6 Jahre		
Kinder von 7-14 Jahren		
Kinder von 14-18 Jahren		
Kinder älter als 18 Jahre		

3. Welchen allgemeinbildenden Schulabschluss haben Sie erworben?

- POS bzw. Realschule
- Abitur
- 8 Klassen Volks- /Hauptschule bzw. 8 Klassen Abschluss
- Sonderschule
- anderer Schulabschluss, und zwar _____
- Keinen Schulabschluss

4. Welche abgeschlossene Berufs- bzw. Hochschulausbildung haben Sie?

- Teilfacharbeiterabschluss, und zwar als _____
- Facharbeiterabschluss, und zwar als _____
- Fachschulabschluss, und zwar als _____
- Fachhochschulabschluss, und zwar als _____
- Hochschul-/ Universitätsabschluss, und zwar als _____
- Anderer Abschluss, und zwar _____
- keine

5. Wie oft und seit wann sind Sie arbeitslos? Bitte ABM's, Umschulungen usw. als Unterbrechung der Arbeitslosigkeit angeben.

seit: _____

davon ABMs, Umschulungen, Unterbrechungen in der Arbeitslosigkeit _____ (Anzahl)

6. Wie viele Personen leben in Ihrem Haushalt und wie viele davon sind arbeitslos? Bitte rechnen Sie sich selbst immer dazu.

Personen im Haushalt _____ (bitte Anzahl eintragen)
davon zurzeit arbeitslos _____ (bitte Anzahl eintragen)

7. Waren Sie vor Ihrer Arbeitslosigkeit der Hauptverdiener in Ihrem Haushalt?

- Ja
 Nein

8. Wie hoch ist derzeit das gesamte monatliche Nettoeinkommen in Ihrem Haushalt?

- | | |
|---|--|
| <input type="checkbox"/> bis 400 EUR | <input type="checkbox"/> 1000 - 1200 EUR |
| <input type="checkbox"/> 400 - 600 EUR | <input type="checkbox"/> 1200 - 1500 EUR |
| <input type="checkbox"/> 600 - 800 EUR | <input type="checkbox"/> 1500 - 1800 EUR |
| <input type="checkbox"/> 800 - 1000 EUR | <input type="checkbox"/> über 2000 EUR |

9. Wie beurteilen Sie selbst Ihre wirtschaftliche Lage?

- Sehr gut
 Gut
 Befriedigend
 Ausreichend
 Mangelhaft
 Katastrophal

10. Hat sich Ihre Arbeitslosigkeit auf folgende wirtschaftliche Bereiche ausgewirkt? Mehrfachnennungen sind hier möglich. Bitte kreuzen Sie aber nur Zutreffendes an.

- Persönliche Ausgaben werden eingeschränkt.
 Vorgesehene Anschaffungen bzw. Urlaubsreisen werden zurückgestellt.
 Ersparnisse werden teilweise bzw. ganz aufgebraucht.
 Probleme mit Verträgen (Ratenzahlungen, Sparverträge, Versicherungsbeiträge)
 Probleme mit der Mietzahlung
 Es sind Schulden entstanden.
 Es gibt keine finanziellen Probleme.

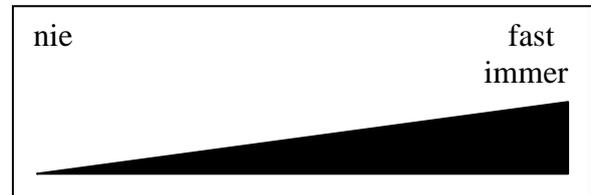
7.3 GHQ (General Health Questionnaire "GHQ12")

Die folgenden Fragen beziehen sich auf Ihr allgemeines Befinden in der letzten Zeit, d.h. den vergangenen vier Wochen. Bitte schätzen Sie ein, welche der vier Antwortmöglichkeiten jeweils für Sie am ehesten zutrifft.

		sehr selten	manch- mal	ziemlich oft	fast immer
1	Konnten Sie sich auf Dinge konzentrieren, die Sie zu tun hatten?				
2	Haben Sie wenig geschlafen, weil Sie Sorgen hatten?				
3	Hatten Sie das Gefühl, im Alltagsgeschehen nützlich zu sein bzw. gebraucht zu werden?				
4	Fühlten Sie sich imstande, Entscheidungen zu treffen?				
5	Hatten Sie das Gefühl ständig unter Belastung und Anspannung zu stehen?				
6	Hatten Sie das Gefühl, die anstehenden Schwierigkeiten meistern zu können?				
7	Konnten Sie ihr normales Alltagsleben genießen?				
8	Fühlten Sie sich imstande, sich den zu lösenden Problemen zu stellen?				
9	Hatten Sie den Eindruck, unglücklich bzw. niedergeschlagen zu sein?				
10	Hatten Sie das Gefühl, Ihr Selbstvertrauen verloren zu haben?				
11	Fühlten Sie sich als wertlose Person?				
12	Meinten Sie – alles in allem- recht glücklich zu sein?				

7.4 BDI-V (Beck-Depressions-Inventar "BDI-V")

In diesem Fragebogen geht es um Ihr gegenwärtiges Lebensgefühl. Bitte geben Sie zu jeder Frage an, wie häufig Sie die genannte Stimmung oder Sichtweise erleben.



1	Ich bin traurig	0	1	2	3	4	5
2	Ich sehe mutlos in die Zukunft	0	1	2	3	4	5
3	Ich fühle mich als Versager(in).	0	1	2	3	4	5
4	Es fällt mir schwer, etwas zu genießen.	0	1	2	3	4	5
5	Ich habe Schuldgefühle.	0	1	2	3	4	5
6	Ich fühle mich bestraft.	0	1	2	3	4	5
7	Ich bin von mir enttäuscht.	0	1	2	3	4	5
8	Ich werfe mir Fehler und Schwächen vor.	0	1	2	3	4	5
9	Ich denke daran, mir etwas anzutun.	0	1	2	3	4	5
10	Ich weine.	0	1	2	3	4	5
11	Ich fühle mich gereizt und verärgert.	0	1	2	3	4	5
12	Mir fehlt das Interesse an Menschen.	0	1	2	3	4	5
13	Ich schiebe Entscheidungen vor mir her.	0	1	2	3	4	5
14	Ich bin besorgt um mein Aussehen.	0	1	2	3	4	5
15	Ich muss mich zu jeder Tätigkeit zwingen.	0	1	2	3	4	5
16	Ich habe Schlafstörungen.	0	1	2	3	4	5
17	Ich bin müde und lustlos.	0	1	2	3	4	5
18	Ich habe keinen Appetit.	0	1	2	3	4	5
19	Ich mache mir Sorgen um meine Gesundheit	0	1	2	3	4	5
20	Sex ist mir gleichgültig.	0	1	2	3	4	5

7.5 SWE (Allgemeine Selbstwirksamkeitserwartung)

Frage	stimmt nicht	stimmt kaum	stimmt eher	stimmt genau
1. Die Lösung schwieriger Probleme gelingt mir immer, wenn ich mich darum bemühe.				
2. Wenn mir jemand Widerstand leistet, finde ich Mittel und Wege mich durchzusetzen.				
3. Es bereitet mir keine Schwierigkeiten, meine Absichten und Ziele zu verwirklichen.				
4. Auch bei überraschenden Ereignissen glaube ich, daß ich gut damit zurechtkommen werde.				
5. In unerwarteten Situationen weiß ich immer, wie ich mich verhalten soll.				
6. Für jedes Problem habe ich eine Lösung.				
7. Schwierigkeiten sehe ich gelassen entgegen, weil ich mich immer auf meine Fähigkeiten verlassen kann.				
8. Wenn ich mit einem Problem konfrontiert werde, habe ich meist mehrere Ideen, wie ich damit fertig werde.				
9. Wenn ich mit einer neuen Sache konfrontiert werde, weiß ich, wie ich damit umgehen kann.				
10. Was auch immer passiert, ich werde schon klarkommen.				

7.7 PSW (Psychosoziales Wohlbefinden)

Hier sind sechs Aussagen über Verhaltensweisen, Gedanken und Gefühle notiert, die man mehr oder weniger oft hat. Bitte schätzen Sie ein, wie diese Aussagen derzeit in Ihrem Leben auf Sie zutreffen. Kreuzen Sie das am besten zutreffende Feld pro Zeile an.

		Antwortmöglichkeiten			
		fast nie	manchmal	oft	immer
1	Ich bin unbeschwert und gut aufgelegt.				
2	Ich fühle mich anderen Menschen unterlegen.				
3	Ich bin ein ruhiger ausgeglichener Mensch.				
4	Ich blicke voller Zuversicht in die Zukunft.				
5	Ich bin stolz auf meinen Körper.				
6	Ich fühle mich voller Energie und Tatkraft.				

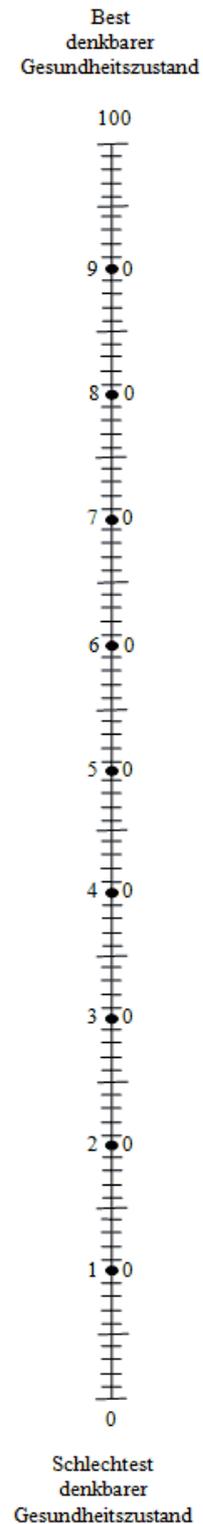
7.8 VAS (Visuelle Analogskala des Gesundheitszustandes des "EuroQOL")



Um Sie bei der Einschätzung, wie gut oder wie schlecht Ihr Gesundheitszustand ist, zu unterstützen, haben wir eine Skala gezeichnet, ähnlich einem Thermometer. Der best denkbare Gesundheitszustand ist mit einer "100" gekennzeichnet, der schlechteste mit "0".

Wir möchten Sie nun bitten, auf dieser Skala zu kennzeichnen, wie gut oder schlecht Ihrer Ansicht nach Ihr persönlicher Gesundheitszustand heute ist. Bitte verbinden Sie dazu den untenstehenden Kasten mit dem Punkt auf der Skala, der Ihren heutigen Gesundheitszustand am besten wiedergibt.

**Ihr heutiger
Gesundheitszustand**



7.9 TV- und Medienkonsum

An einem gewöhnlichen Tag: Wie lange schauen Sie *Fernsehen*?

- 1-2 Stunden 3-4 Std. 5-6 Std. 7-8 Std. 9-10 Std. 11-12 Std. mehr als 12 Std.
 Ich schaue kein Fernsehen

Wie viel Zeit am Tag verbringen Sie durchschnittlich im *Internet* ("surfen")?

- 1-2 Stunden 3-4 Std. 5-6 Std. 7-8 Std. 9-10 Std. 11-12 Std. mehr als 12 Std.
 Ich nutze kein Internet

Wie viel Zeit am Tag verbringen Sie durchschnittlich mit *Computerspielen* (oder *Videospielen wie z.B. XBox, Playstation etc.*)?

- 1-2 Stunden 3-4 Std. 5-6 Std. 7-8 Std. 9-10 Std. 11-12 Std. mehr als 12 Std.
 Ich nutze keine Computer- oder Videospiele

7.10 Zusätzliche Tabellen und Grafiken

Tab. 7.1: Täglicher zeitlicher Konsum für Unterhaltungsmedien zu T1.

		Fernsehen		Internet "surfen"		Computerspiele	
		N	Prozent	N	Prozent	N	Prozent
Täglicher Konsum	0 h.	0	0,0%	59	70,2%	72	85,7%
	1-2 h.	19	22,6%	18	21,4%	10	11,9%
	3-4 h.	39	46,4%	5	6,0%	2	2,4%
	5-6 h.	22	26,2%	2	2,4%	0	0,0%
	7-8 h.	3	3,6%	0	0,0%	0	0,0%
	9-10 h.	1	1,2%	0	0,0%	0	0,0%
	10-12 h.	0	0,0%	0	0,0%	0	0,0%
	> 12 h.	0	0,0%	0	0,0%	0	0,0%
Gesamt		84	100,0%	84	100,0%	84	100,0%

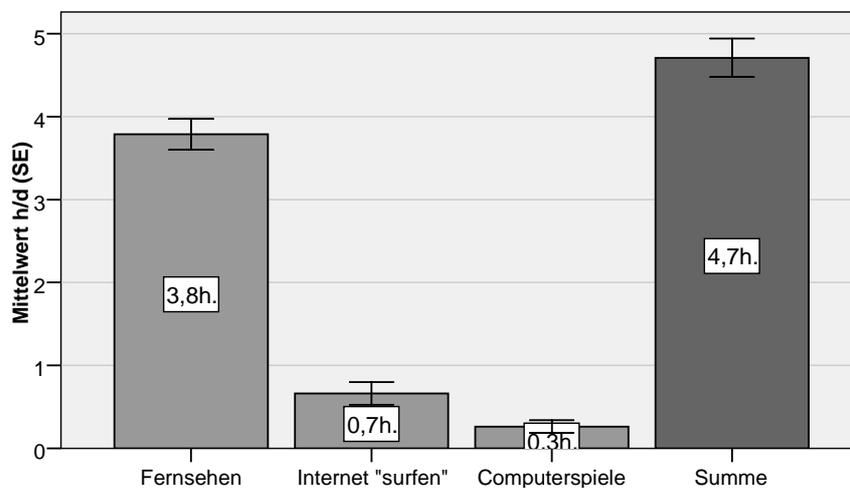


Abb. 7.1: Durchschnittliche tägliche Nutzungsdauer für Unterhaltungsmedien zu T1.

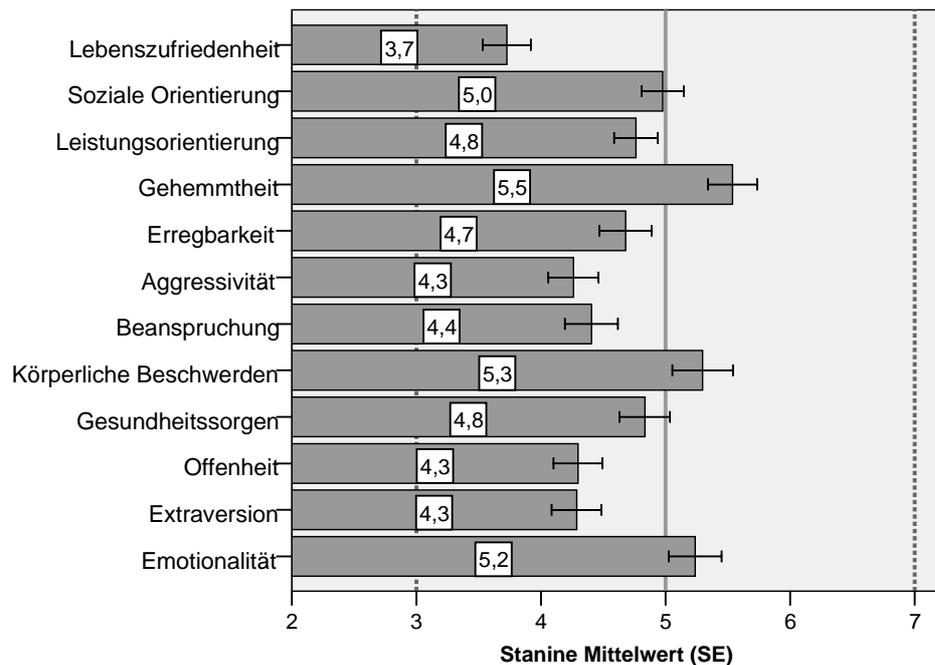


Abb. 7.2: Persönlichkeitsdimensionen des FPI als Stanine-Werte. Der Normbereich (MW=5, SD=1,96) befindet sich zwischen den gestrichelten Markierungen.

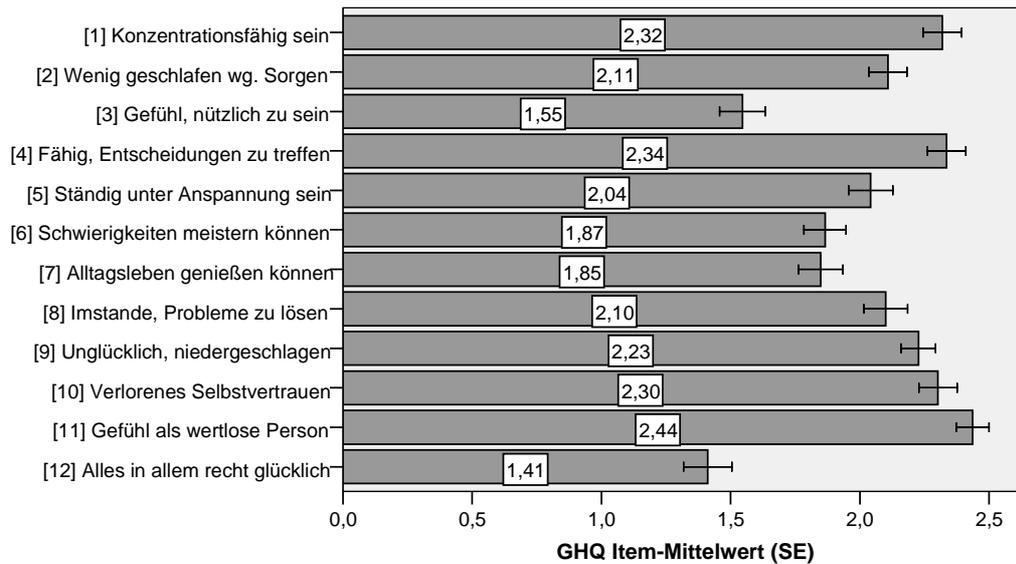


Abb. 7.3: T1-Ergebnisse des GHQ auf Itemebene. Je größer der Wert, desto weniger unspezifische psychische Probleme. Die semantisch negativ formulierten Items 2, 5, 9, 10 und 11 wurden umgepolt.

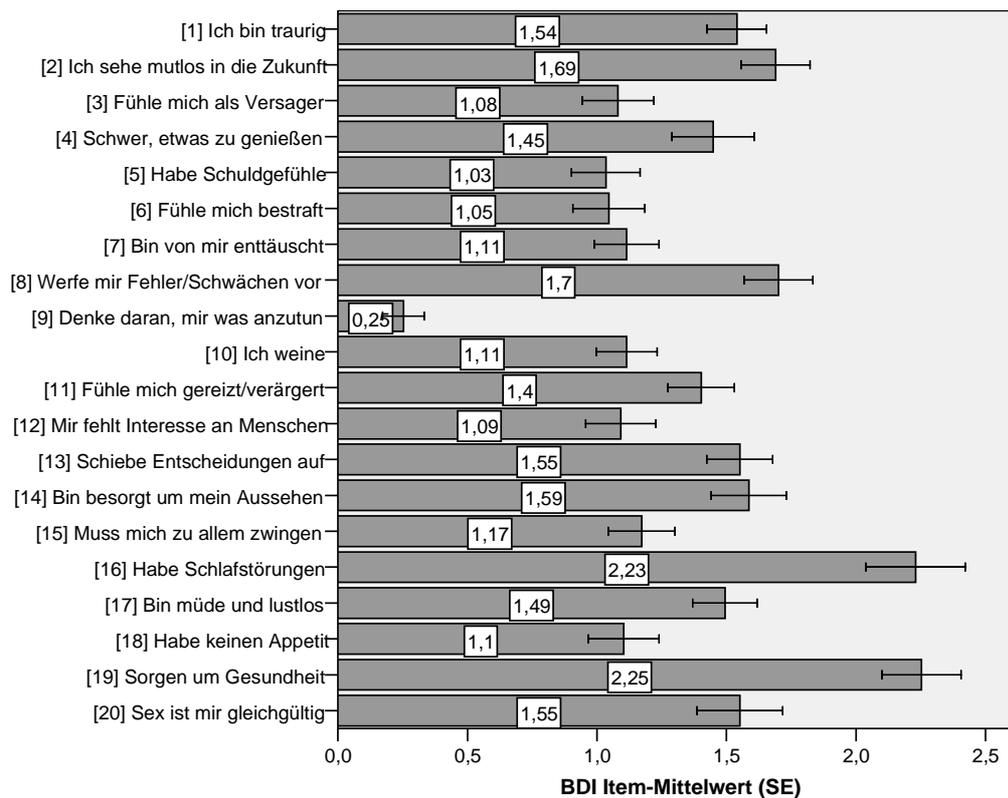


Abb. 7.4: T1-Ergebnisse des BDI auf Itemebene.

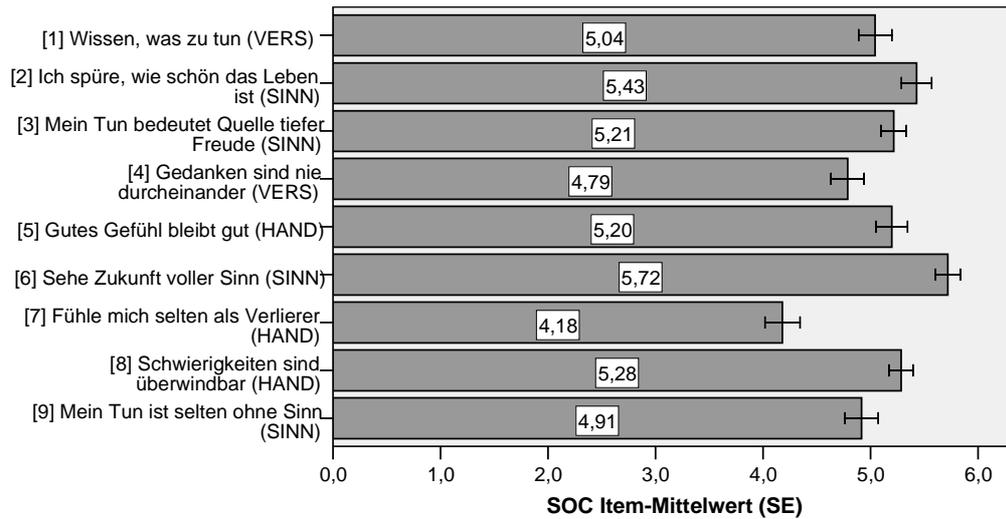


Abb. 7.5: T1-Ergebnisse des SOC auf Itemebene ("Hand": Handhabbarkeit, "Vers": Verstehbarkeit, "Sinn": Sinnhaftigkeit).

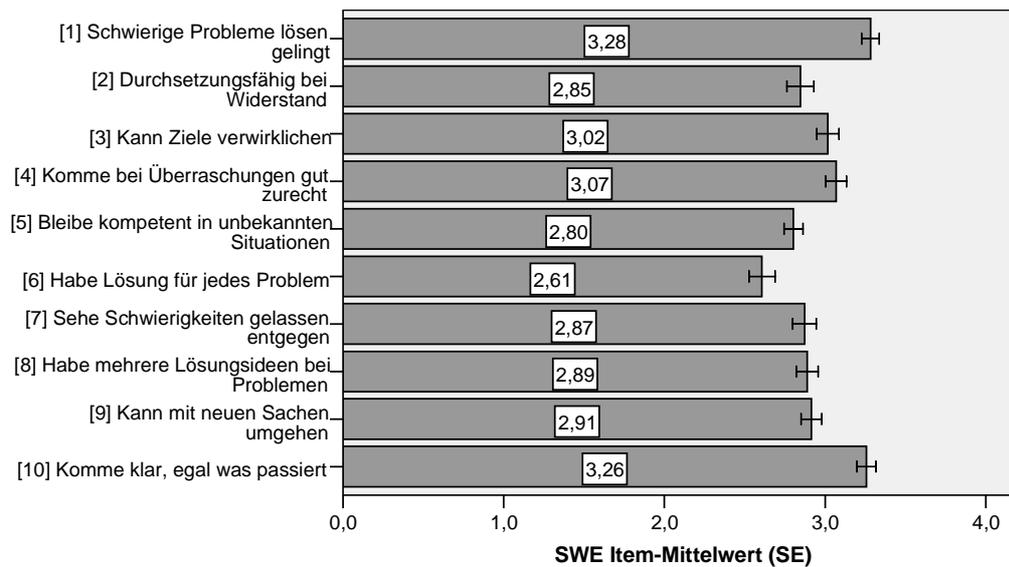


Abb. 7.6: T1-Ergebnisse der SWE auf Itemebene.

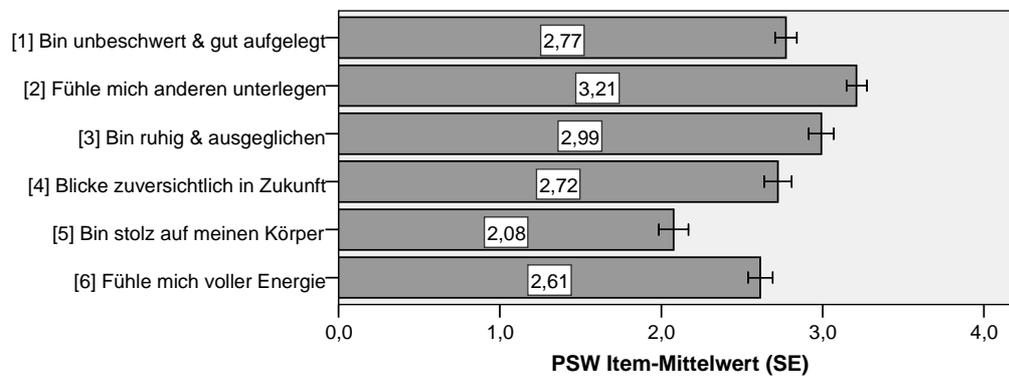
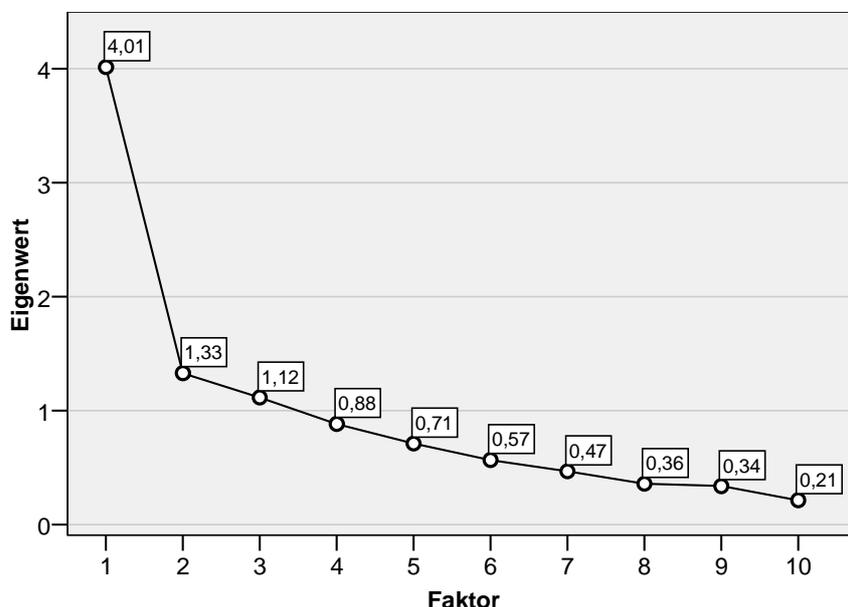


Abb. 7.7: T1-Ergebnisse des PSW auf Itemebene. Das semantisch negativ formulierte Item 2 wurde umgepolt.

Tab. 7.2: T1: Prozentuale Antworthäufigkeiten auf Itemebene des SF12 (KSK: Item der körperlichen Summenskala; PSK: Item der psychischen Summenskala). Pro Item sind die höchsten Antworthäufigkeiten fett markiert.

[1] (KSK) Beschreibung allgemeiner Gesundheitszustand?	ausgezeichnet	5%	[9] (PSK) Fühlen Sie sich ruhig und gelassen?	Immer	9%
	sehr gut	11%		Meistens	55%
	Gut	58%		Ziemlich oft	16%
	weniger gut	24%		Manchmal	15%
[2] (KSK) Einschränkung bei mittelschweren Tätigkeiten?	schlecht	2%	Selten	4%	
	stark eingeschr.	2%	Nie	1%	
	etwas eingeschr.	40%	[10] (KSK) Fühlen Sie sich voller Energie?	Immer	8%
	nicht eingeschr.	58%		Meistens	26%
stark eingeschr.	8%	Ziemlich oft		27%	
etw. eingeschr.	39%	Manchmal		25%	
[3] (KSK) Einschränkung beim steigen von mehreren Treppenabsätzen?	nicht eingeschr.	53%	Selten	13%	
	ja	31%	Nie	1%	
[4] (KSK) Körperlich: Weniger geschafft als Sie wollten?	nein	69%	[11] (PSK) Fühlen Sie sich entmutigt und traurig?	Immer	0%
	ja	26%		Meistens	6%
[5] (KSK) Körperlich: Konnten Sie nur bestimmte Dinge tun?	nein	74%		Ziemlich oft	10%
	ja	19%		Manchmal	28%
[6] (PSK) Seelische Probleme: Weniger geschafft?	nein	81%	Selten	34%	
	ja	18%	Nie	22%	
[7] (PSK) Seelische Probleme: Nicht so sorgfältig arbeiten?	nein	82%	[12] (PSK) Einschränkung sozialer Kontakte durch körperliche oder seelische Probleme?	Immer	0%
	Überhaupt nicht	40%		Meistens	8%
[8] (KSK) Haben Sie Schmerzen?	Ein bisschen	29%		Manchmal	17%
	Mäßig	18%		Selten	28%
	Ziemlich	12%	Nie	47%	
	Sehr	1%			



Tab. 7.3: Rotierte Faktormatrix.

	Faktor	
	1	2
SOC	0,856	0,247
PSW	0,792	0,098
GHQ	0,788	0,132
BDI	-0,762	-0,134
PSK	0,736	-0,242
SWE	0,647	0,216
LEB	0,379	0,256
KSK	0,130	0,770
VAS	0,254	0,661
TV	-0,032	0,537

Abb. 7.8: Scree-Plot der Faktorenanalyse (Methode: PCA) sämtlicher Variablen zu T1. Laut Scree-Kriterium bilden sich 2 orthogonale Faktoren: Variablen zur psychischen Gesundheit laden sämtlich am höchsten auf den 1. Faktor während Variablen zur subjektiv körperlichen Gesundheit und zum TV-Konsum auf dem 2. Faktor die höchsten Ladungskoeffizienten aufweisen (vgl. nebenstehende Tab. 7.3). Die Faktorenanalyse unterstreicht damit den prinzipiellen Unterschied zwischen psychischen und subjektiv-körperlichen Gesundheitsindikatoren.

8 Zusammenfassung der Dissertation

Die bisherige Forschung berichtet übereinstimmend, dass Arbeitslose deutlich stärker gesundheitlich beeinträchtigt sind als Berufstätige, wobei den psychischen Beeinträchtigungen meist eine besondere Rolle zugemessen wird. Meta-Längsschnittanalysen konnten belegen, dass eine Erwerbslosigkeit nicht nur mit höheren psychischen Beschwerden zusammenhängt, sondern sie auch kausal auslöst. Aus der Forschungsliteratur lässt sich ferner ableiten, dass die negativen Folgen der Arbeitslosigkeit vom finanziellen Auskommen, von der Arbeitslosigkeitsdauer, vom Geschlecht und von der sozialen Unterstützung moderiert werden, nicht aber vom Alter der Betroffenen. Über 50-jährige Langzeiterwerbslose haben eine besonders schlechte Prognose auf dem ersten Arbeitsmarkt wiedervermittelt zu werden. Neben dem höheren Alter und Qualifikationsdefiziten stellt eine eingeschränkte psychische oder körperliche Gesundheit eine herausragende wahrscheinlichkeitmindernde Determinante für eine Rückkehr zur Erwerbstätigkeit dar.

Herkömmliche Fördermaßnahmen für Arbeitslose konzentrieren sich auf den Abbau von Qualifikationsdefiziten oder auf die Motivation zu mehr Bewerbungsaktivität. Nur wenige Interventionen fördern die psychische Gesundheit und eine primär körperliche Gesundheitsförderung, die zudem psychologisch evaluiert wurde, ist für Langzeiterwerbslose bislang unbekannt. In dieser Dissertation wurde für die Zielgruppe erstmalig ein Präventionsprogramm getestet, das im Kern aus der Erhöhung des Bewegungsverhaltens durch ein körperliches Ausdauertraining besteht und somit explizit das gesundheitsbezogene Präventionspotential von langjährig Arbeitslosen ausnutzen sollte.

Die Probanden der Studie wurden über einen Bildungsträger im Raum Schwerin rekrutiert, der von der örtlichen ARGE Arbeitslose für Qualifizierungsmaßnahmen zugeteilt bekam. 70% der älteren Langzeiterwerbslosen konnten vom Bildungsträger für eine Teilnahme am Gesundheitsförderungsprogramm gewonnen werden. Für N=119 (61% Frauen) ältere (54 Jahre) langzeiterwerbslose (Median: 5,1 Jahre arbeitslos) Probanden erfolgten drei Untersuchungen im Längsschnitt: Prä, post und 6-monatiges follow-up im Zwei-Gruppen-Design mit 2 unterschiedlichen Intensitätsstufen des Trainings.

Für die psychologische Untersuchung wurden Messgrößen zur Bestimmung von Risikofaktoren sowie Ressourcen der psychischen Gesundheit und zur Erhebung des subjektiven körperlichen Gesundheitszustands verwendet. Des Weiteren wurde der Fernsehkonsum als Schätzmaß eines prototypischen passiven Freizeitverhaltens erhoben. Mit den Daten der Erstuntersuchung erfolgte die Ermittlung des Gesundheitszustands im Vergleich zu Studien aus der alters- und ggf. geschlechtsparallelisierten Allgemeinbevölkerung.

Die Auswertung der Erstuntersuchung zeigte, dass die Stichprobe in den Fragebögen, welche die Risikofaktoren der psychischen Gesundheit erfassten, etwa doppelt so hohe Quoten von psychischen Beschwerden im Vergleich zu den Normstichproben aufwies. Demgegenüber waren in den Messungen, welche die Ressourcen der psychischen Gesundheit abbil-

deten, kaum Abweichungen von der Allgemeinbevölkerung feststellbar. Des Weiteren fühlte sich die Stichprobe zu zwei Drittel häufiger durch körperliche Gesundheitsprobleme beeinträchtigt, als man es aufgrund der Normen hätte erwarten können. Der Interventionsbedarf wurde in Bezug auf die Senkung von psychischen und körperlichen Beschwerden deswegen als hoch, in Bezug auf Steigerungen von psychischen Ressourcen als niedrig eingeschätzt. Ferner ließ sich die Feststellung des Interventionsbedarfs gegen problematische Antworttendenzen absichern.

Querschnittsanalysen zum moderierenden Einfluss des Geschlechts ergaben zwischen langzeiterwerbslosen Männern und Frauen nur wenige und zudem inkonsistente Unterschiede. Demgegenüber konnte deutlich nachgewiesen werden, dass ausschließlich erwerbslose Frauen vom Vorhandensein einer Partnerschaft seelisch profitierten, während bei arbeitslosen Männern der Beziehungsstatus keinerlei Unterschiede in ihrem Befinden offenbarte.

Drop-out-Analysen ließen für die Variablen zur Erfassung der psychischen und körperlichen Gesundheit des ersten Messzeitpunktes zwischen späteren Studienabbrechern und -fortsetzern kaum Unterschiede objektivieren, welches für die Generalisierbarkeit der sich anschließenden Längsschnittauswertung sprach. Im Gegensatz dazu war für die Studienabbrecher der ursprüngliche TV-Konsum signifikant gegenüber den Studienkomplettierern erhöht.

Die Analyse zur Aufrechterhaltung des Bewegungsverhaltens im follow-up ergab, dass ein Drittel der Versuchsteilnehmer den Sport fortsetzte, Männer deskriptiv etwas häufiger als Frauen. Statistische Modelle zur retrospektiven "Vorhersage" des Sportabbruch-Risikos offenbarten, dass dieser zum einen durch mangelnde psychisch förderliche Interventionseffekte wahrscheinlicher wurde und zum anderen durch einen a priori hohen TV-Konsum determiniert war.

Die Berechnung der Interventionseffekte zeigte, dass die Gesundheitsförderung signifikant vor allem 1.) unspezifische psychische Beschwerden senkte, 2.) das passive Freizeitverhalten minderte, 3.) psychische Einschränkungen der Lebensqualität reduzierte und 4.) Depressionssymptome linderte. Die zwei erstgenannten Behandlungseffekte erreichten zudem eine relative Zeitstabilität im 6-monatigen follow-up, die beiden letztgenannten nicht. Sämtliche Variablen zu psychischen Risikofaktoren indizierten demnach signifikante Interventionseffekte, während sowohl die psychischen Ressourcen als auch die subjektive Wahrnehmung des körperlichen Gesundheitszustandes durch das Training eine signifikante Steigerung vermissen ließen.

Interventionsmoderierende Analysen ergaben die Interpretation, dass die Variation der Trainingsintensität keinen Einfluss auf die hier gemessenen Interventionseffekte hatte. Demgegenüber konnte deutlich nachgewiesen werden, dass vor allem Probanden mit a priori relevanter erhöhter Depressivität vom Programm profitierten. Geschlechtseffekte waren zwar nicht signifikant nachweisbar, aber ergaben für die körperliche Lebensqualität einen Trend für einen stärkeren Profit der männlichen Probanden im Vergleich zu den weiblichen. Eine Ana-

lyse von follow-up-Veränderungen der hier getesteten Längsschnittvariablen demaskierte, dass die freiwillige Fortsetzung des Sports allenfalls eine kleine gesundheitsförderliche Wirkung im follow-up andeutete, aber sich keine signifikant unterschiedlichen Effekte im Kontrast zum Sportabbruch objektivieren ließen.

Die Schätzung von Kontrollgruppeneffekten aus einer im Setting gut vergleichbaren externen Interventionsstudie ließ die Interpretation zu, dass sich die hiesigen Interventionseffekte zur Linderung von psychischen Beschwerden keinesfalls zufriedenstellend mit "Placeboeffekten" von herkömmlichen Qualifizierungsmaßnahmen für Erwerbslose erklären lassen. In der Gegenüberstellung der hiesigen Interventionseffekte mit anderen Studien ließ sich schätzen, dass das hier getestete Programm ca. 30% effektiver war als eine aktuelle Metaanalyse von 16 internationalen Interventionsstudien für Erwerbslose indizierte. Dies führte zu der Interpretation, dass die Wirksamkeit der hier evaluierten Gesundheitsförderung zu bejahen ist und aus psychologischer Sicht empfohlen werden kann.

Des Weiteren wurden mithilfe von Poweranalysen die Veränderungssensitivitäten der hier verwendeten Messinstrumente quantifiziert, welches vor allem für die Planung einer Replikationsstudie von Interesse wäre. Aus den Ergebnissen der Sportabbruchvorhersage wird abgeleitet, wie die Aufrechterhaltungswahrscheinlichkeit des Sports in künftigen Programmen gesteigert werden könnte. Schließlich werden Empfehlungen zum verstärkten Angebot der Gesundheitsförderung für bestimmte Zielgruppen von älteren Langzeiterwerbslosen (Männer und Personen mit depressiver Symptomatik) diskutiert und Vorschläge für die weitere Forschung angeregt.