

# FIT 50+

Evaluation eines modularen Konzeptes zur Verbesserung der physischen und psychischen Gesundheit älterer Langzeitarbeitsloser

Gefördert durch das



**Bundesministerium  
für Bildung  
und Forschung**

Im Rahmen der „Präventionsforschung zur Gesundheitsförderung und Primärprävention von älteren Menschen“ (2007-2010)  
Förderkennzeichen: 01 EL0704

## ABSCHLUSSBERICHT

**Universität  
Rostock**



Traditio et Innovatio



**institut für  
präventivmedizin**  
arbeitsmedizin  
sportmedizin  
sozialmedizin

## Name und Anschrift der wissenschaftlichen Einrichtung



Universität Rostock  
Institut für Präventivmedizin  
St. Georg-Str. 108  
18055 Rostock

## Projektleitung

Prof. Dr. med. habil. Regina Stoll  
Tel.: +49 / 381 - 494 9951  
Fax: +49 / 381 - 494 9952  
Email: [Regina.Stoll@uni-rostock.de](mailto:Regina.Stoll@uni-rostock.de)

## Projektbearbeitung

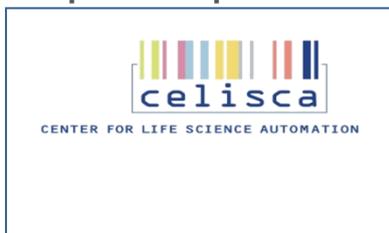
Dr. med. Steffi Kreuzfeld  
Dipl.-Psych. Markus Preuss

## Präventionsanbieter



Dipl.-Ing. Jürgen Brandt  
RegioVision Schwerin GmbH  
Am Margaretenhof 28  
19057 Schwerin  
Tel.: +49 / 385 – 200314-10  
Email: [jbrandtsn@gmx.de](mailto:jbrandtsn@gmx.de)

## Kooperationspartner



Prof. Dr.-Ing. habil. Kerstin Thurow  
Celisca – Center for Life Science Automation  
Universität Rostock  
F.-Barnewitz-Str. 8  
18119 Rostock

## Kontakt/ Korrespondenz:

Email: [Steffi.Kreuzfeld@uni-rostock.de](mailto:Steffi.Kreuzfeld@uni-rostock.de)  
Tel.: +49 / 381 – 494 9957

<b>VORWORT</b>	<b>7</b>
<b>1. EMPIRISCHE BEFUNDE ZUR AUSGANGSLAGE</b>	<b>9</b>
<b>1.1. Psychische Gesundheit und Arbeitslosigkeit</b>	<b>9</b>
1.1.1. Gibt es ein spezifisches psychisches Erwerbslosigkeitssyndrom?	9
1.1.2. Erklärungsansätze	11
1.1.3. Einfluss des Alters	14
1.1.4. Einfluss der Arbeitslosigkeitsdauer	16
<b>1.2. Physische Gesundheit und Arbeitslosigkeit</b>	<b>18</b>
1.2.1. Allgemeine Erkenntnislage	18
1.2.2. Inanspruchnahme medizinischer Versorgungsleistungen	19
1.2.3. Arbeitslosigkeit und Mortalität	20
1.2.4. Arbeitslosigkeit und verändertes Gesundheitsverhalten	20
<b>1.3. Besonderheiten der Langzeitarbeitslosigkeit</b>	<b>21</b>
1.3.1. Relative Zunahme von Selektionseffekten	22
1.3.2. Rückgang des Qualifikationsniveaus	23
1.3.3. Eingeschränktes soziales Netz	23
1.3.4. Effekte der Wiedereinstellung auf die Gesundheit	23
1.3.5. Chancen des Wiedereinstiegs	24
<b>1.4. Maßnahmen zur Förderung von Erwerbslosen</b>	<b>25</b>
1.4.1. Vorstellung von Interventionsstudien	26
1.4.2. Durchschnittlicher Interventionserfolg	30
<b>1.5. Körperliche Aktivität und kardiovaskuläres Risiko</b>	<b>31</b>
<b>2. FRAGESTELLUNGEN UND HYPOTHESEN</b>	<b>34</b>
<b>2.1. Variablen zur Beschreibung der Gesundheit</b>	<b>34</b>
2.1.1. Faktoren der objektiven körperlichen Gesundheit	34
2.1.2. Risikofaktoren des psychischen Wohlergehens	34
2.1.3. Ressourcen des psychischen Wohlergehens	35
2.1.4. Subjektive körperliche Gesundheit	35
<b>2.2. Variable zur Abschätzung des passiven Freizeitverhaltens</b>	<b>35</b>
<b>2.3. Feststellung des Interventionsbedarfes</b>	<b>35</b>
2.3.1. Auswirkung von Antworttendenzen	36
<b>2.4. Korrelative Exploration der Messverfahren</b>	<b>36</b>
<b>2.5. Überprüfung der Gleichheit zwischen den Kohorten</b>	<b>36</b>
<b>2.6. Drop-Out-Analysen</b>	<b>37</b>
<b>2.7. Risikoprädiktion des Sportabbruchs</b>	<b>37</b>
<b>2.8. Längsschnitthypothesen</b>	<b>37</b>
2.8.1. Hypothesen zur Veränderung von Faktoren der körperlichen Gesundheit	37

2.8.2. Hypothesen zur Veränderung der Risikofaktoren der psychischen Gesundheit (L1)	38
2.8.3. Hypothesen zur Veränderung der Ressourcen der psychischen Gesundheit (L2)	39
2.8.4. Hypothesen zur Veränderung der subjektiven körperlichen Gesundheit (L3)	39
2.8.5. Hypothesen zur Veränderung des Fernsehkonsums (L4)	40
2.8.6. Hypothesen zur Fortsetzung des Sportes im follow-up (L5)	41
<b>3. METHODEN</b>	<b>42</b>
<b>3.1. Einleitung</b>	<b>42</b>
<b>3.2. Untersuchungsdesign</b>	<b>42</b>
3.2.1. Stichprobengröße	43
3.2.2. Kontrollgruppe	43
3.2.3. Experimentalgruppen	44
<b>3.3. Rekrutierung der Probanden</b>	<b>46</b>
<b>3.4. Interventionsprogramm</b>	<b>47</b>
3.4.1. Modul „Gesundheit-Basiswissen“	47
3.4.2. Modul „Gesundheit-Praktische Übungen“	48
<b>3.5. Untersuchungsablauf</b>	<b>50</b>
<b>3.6. Methodisches Inventar</b>	<b>50</b>
3.6.1. Erfassung soziodemographischer Merkmale	50
3.6.2. Erfassung von Variablen der körperlichen Gesundheit	51
3.6.3. Erfassung von Variablen der psychischen und psychosozialen Gesundheit	52
3.6.4. Medienkonsum	55
<b>3.7. Statistische Verfahren</b>	<b>56</b>
<b>3.8. Ethikvotum</b>	<b>56</b>
<b>4. ERGEBNISSE</b>	<b>57</b>
<b>4.1. Soziodemographische Beschreibung der Stichprobe</b>	<b>57</b>
4.1.1. Allgemeine soziodemographische Daten	57
4.1.2. Einkommen	59
4.1.3. Häufigkeit sozialer Kontakte	62
4.1.4. Qualität sozialer Unterstützung	63
<b>4.2. Physische Gesundheit</b>	<b>64</b>
<b>4.3. Psychische und psychosoziale Gesundheit</b>	<b>66</b>
4.3.1. Freiburger Persönlichkeitsinventar (FPI)	66
4.3.2. General Health Questionnaire (GHQ)	68
4.3.3. Beck Depressions-Inventar (BDI)	70
4.3.4. Kohärenzsinn – Sense of Coherence (SOC)	72
4.3.5. Allgemeine Selbstwirksamkeitserwartung (SWE)	73
4.3.6. Fragebogen zum Gesundheitszustand (SF12)	74
4.3.7. Visuelle Analogskala zur Einschätzung des Gesundheitsempfindens (VAS)	77
4.3.8. Psychosoziales Wohlbefinden (PSW)	77
4.3.9. Zusammenfassung des Interventionsbedarfs	78
4.3.10. Antworttendenz-Hypothesentest	80
<b>4.4. Fernseh- und Medienkonsum</b>	<b>81</b>

<b>4.5. Zusammenhang der Messinstrumente</b>	<b>82</b>
<b>4.6. Test auf Gleichheit der Kohorten</b>	<b>84</b>
4.7.1. Ebene der physischen Gesundheit	84
4.7.2. Ebene der psychischen und psychosozialen Ebene/ Medienkonsum	85
<b>4.7. Drop-Out-Analyse</b>	<b>85</b>
4.7.1. Teilnehmer- und Drop-Out-Quote T2	85
4.7.2. Teilnehmer- und Drop-Out-Quote T3	86
4.7.3. Drop-Out-Analyse T1-T2	87
4.7.4. Drop-Out-Analyse T1-T3	87
<b>4.8. Sportfortsetzeranalyse</b>	<b>88</b>
4.8.1. T1 Modell	88
4.8.2. Prädiktionsmodell, Odds und Odds-Ratios	90
4.8.3. Kombiniertes Modell	92
<b>4.9. Längsschnittanalysen der Interventionseffekte</b>	<b>94</b>
4.9.1. Statistische Vorüberlegungen zur Zeitstabilität	94
4.9.2. Statistische Vorüberlegungen zu Moderatoreffekten	96
4.9.3. Ergebnisse der Hypothesentests	97
<b>4.10. Effektstärken &amp; Poweranalysen der Längsschnittergebnisse</b>	<b>115</b>
4.10.1. Veränderungssensitivität der Messinstrumente	115
4.10.2. Schätzung der Effektgrößen der Interventionseffekte	120
4.10.3. Moderierender Einfluss der Zwischensubjektfaktoren	121
<b>4.11. Zufriedenheit/ Akzeptanz in der Zielgruppe</b>	<b>121</b>
<b>5. DISKUSSION</b>	<b>123</b>
<b>5.1. Diskussion der Ergebnisse der Erstuntersuchung</b>	<b>123</b>
5.1.1. Diskussion der Fragebogenergebnisse zu T1	123
5.1.2. Diskussion der Drop-Out-Ergebnisse	124
5.1.3. Diskussion der Ergebnisse der Sportfortsetzeranalyse	125
<b>5.2. Diskussion der Längsschnittergebnisse</b>	<b>128</b>
5.2.1. Interventionseffektivität	128
5.2.2. Relative Zeitstabilität der signifikanten Interventionseffekte	128
5.2.3. Diskussion der Interventionseffekte	129
5.2.4. Vergleich mit dem durchschnittlichen publizierten Interventionserfolg	130
5.2.5. Diskussion der nicht signifikanten Interventionserfolge	131
5.2.6. Diskussion von Zwischensubjekteffekten	133
<b>6. Prozessdatenmanagement</b>	<b>135</b>
<b>6.1. Zusammenfassung</b>	<b>135</b>
<b>6.2. Framework-System openLIMS</b>	<b>135</b>
<b>6.3. Prozess-Abbildung unter Nutzung des Frameworksystems</b>	<b>138</b>
<b>6.4. Systemarchitektur</b>	<b>140</b>
<b>6.5. Spezifische Daten-Verarbeitung</b>	<b>141</b>

<b>6.6. Datenanalyse und Exportinterfaces</b>	<b>142</b>
<b>6.7. Zugangsplattform</b>	<b>142</b>
<b>6. LITERATUR</b>	<b>144</b>
<b>7. EIGENE PUBLIKATIONEN</b>	<b>158</b>

Anmerkung: Der Bericht enthält Teile der Dissertationsschrift von Dipl.-Psych. Markus Preuss

## Vorwort

Gesundheitliche Auswirkungen von Arbeitslosigkeit sind seit langem Gegenstand wissenschaftlicher Forschung. Erste grundlegende Untersuchungen in den westlichen Industrienationen gehen auf die 20-iger und 30-iger Jahre des letzten Jahrhunderts zurück (Chapin, 1924; Hofbauer-Flatzek, 1933; M. Jahoda, Lazarsfeld, & Zeisel, 1933; Marsh, Fleming, & Blackler, 1938). In den folgenden Jahrzehnten entwickelte sich in Deutschland eine vor allem durch die Gesundheits- und Sozialwissenschaften vorangetriebene systematische Erwerbslosigkeitsforschung, die in den 90-iger Jahren das Forschungsfeld auf die von Arbeitslosigkeit bedrohten Personen ausdehnte.

Gegenwärtig vollzieht sich im Rahmen der Globalisierung ein Wandel der gesamten Arbeitswelt mit ansteigender Dynamik und Flexibilisierung. In dessen Folge kam es neben der Arbeitsverdichtung auch zu einer Zunahme atypischer und prekärer Arbeitsverhältnisse und zu häufigen Diskontinuitäten in Erwerbsbiographien. Damit überlagerten sich die gesundheitlichen Auswirkungen von Erwerbstätigkeit, bedrohter Erwerbstätigkeit und Erwerbslosigkeit und wurden nun auch vermehrt Forschungsgegenstand der Arbeitswissenschaften und der Arbeitsmedizin. Überwiegend wurden die psychosozialen Auswirkungen o.g. Veränderungen untersucht, zu den physischen Folgen gab es in der Vergangenheit jedoch verhältnismäßig wenige Untersuchungen.

In den letzten Jahren wurde zunehmend deutlich, wie der sozioökonomische Status einer Person deren gesundheitlichen Status determiniert und, dass die gesundheitlichen Chancen in Deutschland in Abhängigkeit von der sozialen Lage ungleich verteilt sind (RKI, 2006a). Auch das Risiko der Arbeitslosigkeit bzw. die Chancen auf Wiedereingliederung in den Arbeitsmarkt stehen in Zusammenhang zur sozialen Lage und zum gesundheitlichen Status einer Person.

Gesundheitliche Einschränkungen gelten als wesentliches Vermittlungshemmnis für die Wiedereingliederung in den Arbeitsmarkt. Eine Analyse von über 300.000 Gutachtenverfahren des Ärztlichen Dienstes der Bundesanstalt für Arbeit aus dem Jahr 2001 ergab, dass der Anteil der Personen mit gesundheitlichen Einschränkungen mit zunehmendem Alter und zunehmender Dauer der Arbeitslosigkeit deutlich ansteigt. Während in der Altersgruppe bis 25 Jahre nur etwa jeder Zehnte gesundheitliche Einschränkungen aufwies, war es in der Gruppe der über 56-Jährigen jeder Dritte (Hollederer, 2003). Neben psychischen Störungen und Verhaltensstörungen wurden bei den Begutachtungen am häufigsten Krankheiten des Muskel-Skelett-Systems und des Bindegewebes diagnostiziert. Gesundheitliche Einschränkungen verschlechtern der Analyse zufolge nicht nur die Wiedereingliederungschancen der Betroffenen, sondern behindern auch die Suche nach einer neuen Stelle. Verminderte Wiedereingliederungschancen in den ersten Arbeitsmarkt für Personen mit gesundheitlichen Einschränkungen sind auch unter Berücksichtigung von Alter, Geschlecht, Ausbildungsgrad und Dauer der Arbeitslosigkeit belegt (Cramer, et al., 2002). Demgegenüber haben Arbeitssuchende ohne gesundheitliche Einschränkungen deutlich höhere Chancen auf Wiedereingliederung (Claussen, 1999; Claussen, Björndal & Hjort, 1993; Taris, 2002). Die Befundlage ist jedoch aufgrund der Komplexität des Geschehens mit zahlreichen Moderatorvariablen nicht einheitlich, es existieren auch Forschungsberichte, die o.g. Ergebnisse nicht bestätigen (Kessler, Turner, & House, 1989).

Das Problem der Arbeitslosigkeit ist für den Einzelnen ein kritisches Lebensereignis, welches mit finanziellen Einbußen, sozialem Abstieg, Perspektivlosigkeit, Verlust von Kompetenzen, Ressourcen, sozialen Kontakten und mit gesundheitlichen Folgen verbunden ist. Für die Gesellschaft bedeutet es eine sowohl soziale als auch ökonomische Herausforderung. Massenarbeitslosigkeit führt zur Verarmung von ganzen Regionen mit Konsequenzen für deren wirtschaftliche Struktur, das Bildungsniveau und die demographische Entwicklung (A. Wiggert, 2009).

Der Arbeitsmarkt ist einerseits durch Fluktuation und andererseits durch Segmentierung geprägt. *„Den Arbeitslosen“* gibt es als Typus genau so wenig, wie es *„den Beschäftigten“* gibt, weil der Bestand an Arbeitslosen lediglich eine Momentaufnahme an einem bestimmten Zähltag darstellt. Erwerbslose bilden keine homogene Gruppe, sondern setzen sich aus verschiedenen Personen mit unterschiedlichen Hintergründen und Verweildauern in der Arbeitslosigkeit zusammen (Hollederer, 2009). Die Lebenssituationen differieren sowohl vor Eintritt in die Arbeitslosigkeit als auch während der Arbeitslosigkeitsperiode und bereits die Ursachen der Arbeitslosigkeit können unterschiedlich sein, z.B. saisonale oder strukturell bedingt (Hradil & Schiener, 2001). Arbeitslosigkeitsrisiken häufen sich bei bestimmten „Problemgruppen“ am Arbeitsmarkt. Im internationalen Vergleich hat Deutschland eine hohe sog. "Sockelarbeitslosigkeit" (Wiggert, 2009). Diese ist geprägt durch einen hohen Prozentsatz an schwierig vermittelbaren und deshalb langfristig Erwerbslosen in höherem Alter mit geringen Qualifikationen und gesundheitlichen Einschränkungen. Nach Eintragungen der Arbeitsvermittler lag der Anteil der gesundheitlich eingeschränkten Erwerbslosen im Jahresdurchschnitt 2005 bei 24% im steuerfinanzierten Rechtskreis des SGB II und bei 19% im SGB III Bereich (Hollederer, 2009). Der „harte Kern“ an Arbeitslosen lässt sich auch in Zeiten eines konjunkturellen Aufschwungs nur begrenzt abbauen, so dass die Sockelarbeitslosigkeit seit 1973, dem Beginn der ersten Ölkrise, kontinuierlich anstieg (Alexander Wiggert, 2009).

Sowohl die Risiken von Arbeitslosigkeit als auch die Chancen auf Wiedereingliederung in den Arbeitsmarkt hängen stark von den konjunkturellen Verläufen ab. In Deutschland gibt es ein starkes Ost-West-Gefälle mit erhöhter Erwerbslosenzahl im Osten, die durch gravierende Strukturbrüche am Arbeitsmarkt und im Sozialgefüge nach dem Wegfall des politischen Systems der DDR bedingt ist. Westdeutschland isoliert betrachtet, hat in Bezug auf die Arbeitslosenquote ein Nord-Süd-Gefälle mit vermehrt auftretender Arbeitslosigkeit im Norden.

## Danksagung

Das Bundesministerium für Bildung und Forschung hat mit seiner Initiative zur Förderung der Präventionsforschung zur Gesundheitsförderung und Primärprävention von älteren Menschen (3. Förderphase) und der Fokussierung auf sozial benachteiligte Gruppen u. a. wissenschaftliche Arbeiten zur Gesundheitsförderung und Primärprävention von älteren Langzeitarbeitslosen unterstützt.

Das hier vorgestellte Projekt Fit 50+ ist eine Kooperation zwischen dem Institut für Präventivmedizin der Universität Rostock und der RegioVision GmbH Schwerin, einem privatem Bildungsträger, der ausschließlich auf dem Gebiet der Qualifizierung und Arbeitsvermittlung älterer Langzeitarbeitsloser tätig ist.

Das Center for Life Science Automation (celisca) hat das Vorhaben durch den Aufbau eines Internet basierten Projektdatenmanagementsystems unterstützt.

Den Förderern und den Kooperationspartnern von Fit50+ sei an dieser Stelle herzlich gedankt! Ohne das große Engagement jedes Einzelnen wäre die erfolgreiche Durchführung der Studie nicht möglich gewesen.

Ganz herzlich danken möchten wir auch allen Menschen, die als Studienteilnehmer dazu beigetragen haben, unser Wissen auf dem Gebiet der Gesundheitsförderung im Kontext von Arbeitsmarktintegration nach langjähriger Erwerbslosigkeit zu erweitern.

## **1. Empirische Befunde zur Ausgangslage**

### **1.1. Psychische Gesundheit und Arbeitslosigkeit**

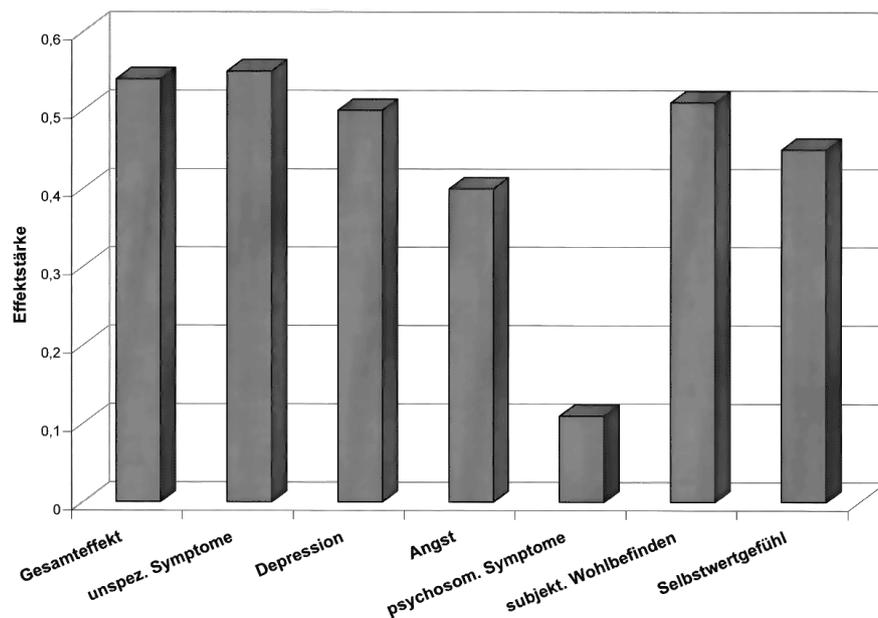
#### **1.1.1. Gibt es ein spezifisches psychisches Erwerbslosigkeitssyndrom?**

Paul et al. führten eine Metaanalyse durch, um die Folgen von Erwerbslosigkeit auf die psychische Gesundheit zusammenfassend quantifizieren zu können. Sie untersuchten 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 (Paul, Hassel & Moser, 2006). Die Autoren resümierten, dass in den Primärstudien eine breite Palette an Symptomen und Syndromen psychischer Beschwerden als Konsequenz der Erwerbslosigkeit angeführt wurde, wie z.B. Depression, Hoffnungslosigkeit, Apathie, psychosomatische Symptome, vermindertes Selbstwertgefühl, verminderte Lebenszufriedenheit, Alkoholismus und Suizid.

Den Depressionssymptomen wird in der Erwerbslosigkeitsforschung eine besondere Bedeutung beigemessen, auch ein geringes Selbstwertgefühl wird als weit verbreitet angesehen, während bei anderen Beanspruchungssymptomen, wie z.B. psychosomatischen Symptomen, Zweifel laut wurden, inwiefern sie im Zusammenhang mit Arbeitslosigkeit eine Rolle spielen (Frese & Mohr, 1978). In ihrer Metaanalyse beschränkten sich Paul et al. auf 6 in der Forschungsliteratur häufig untersuchte Indikatorvariablen der psychischen Gesundheit:

1. Unspezifische psychische Beanspruchungssymptome (z.B. General Health Questionnaire, (Goldberg, 1978))
2. Depressionssymptome (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, (Fahrenberg, Hampel, & Selg, 2010))
6. Selbstwertgefühl (z.B. Self Esteem Scale, (Rosenberg, 1965)).

Paul et al. stellten fest, dass diese Variablen in deutlichem statistischen Zusammenhang stehen und interpretierten, dass sie als Indikatoren für ein abstrakteres und allgemeineres Konstrukt „*Psychische Gesundheit*“ aufgefasst werden können. Die Metaanalyse legte einen klaren Zusammenhang zwischen Erwerbslosigkeit und psychischer Gesundheit offen. Erwerbslose zeigten deutlich mehr Symptome psychischer Beeinträchtigung und geschmäleren Wohlbefindens als Erwerbstätige.



**Abb. 1:** Effektstärken der negativen Auswirkung der Erwerbslosigkeit für psychologische Variablen (Quelle: Paul & Moser, 2006)

Die Unterschiede zwischen Erwerbslosen und Erwerbstätigen geben die Verfasser als Effektstärken (Cohen's *d*, (Jacob Cohen, 1988)) an. Mit  $d=0,54$  war der Gesamteffekt von mittlerer Größe, wie auch die meisten Subskalen (unspezifische psychische Beschwerden, Depression, Angst, eingeschränktes Wohlbefinden, gemindertem Selbstwertgefühl). Die psychosomatischen Symptome zeigten mit  $d=0,11$  den kleinsten aller signifikanten Unterschiedseffekte. Da Effektstärken wie „*d*“ schwierig zu interpretierende Größen darstellen, analysierten die Autoren den Prozentsatz der Fälle von krankheitsrelevanten psychischen Störungen. In dieser Analyse benutzten die Autoren nur etablierte klinische Screening-Tests für die Störungssymptome Angst und Depression. Für Erwerbslose ergab sich ein Anteil von 34%, während für Erwerbstätige ein Anteil von 16% berechnet wurde. Demnach war der Anteil derer, die psychisch deutlich beeinträchtigt waren und möglicherweise psychologische oder psychiatrische bzw. medikamentöse Behandlungen benötigten, unter den Erwerbslosen mehr als doppelt so groß wie unter den Berufstätigen (Paul et al., 2006).

### 1.1.2. Erklärungsansätze

#### Zum Zusammenhang von Arbeitslosigkeit und eingeschränkter psychischer Gesundheit

Um die Zusammenhänge von eingeschränkter psychischer Gesundheit und Arbeitslosigkeit zu erklären, wurden in der psychologischen Forschung verschiedene Theorien herangezogen, z.B. die *Stress-Theorie* nach Lazarus (Lazarus & Alfert, 1964), die *Theorie der erlernten Hilflosigkeit* nach Seligman (Seligman, 1978) oder die *Theorie der kognitiven Dissonanz* nach Festinger (Festinger, 2001). Fünf, mehr spezifische Erklärungsansätze, haben sich als einflussreich bzw. diskussionswürdig erwiesen und sollen kurz vorgestellt werden:

Die "**Deprivationstheorie**" von Jahoda (Jahoda, 1982, 1992) postuliert, dass Erwerbsarbeit mehrere Funktionen erfüllt: Zum Einen die manifeste Funktion des Gelderwerbs, die finanzielle Autonomie bereitstellt und zusätzlich fünf weitere, latente Funktionen: 1.) die Auferlegung einer festen Zeitstruktur, 2.) Sozialkontakt, 3.) Status und Identität, 4.) die Teilhabe an kollektiven Zielen und 5.) regelmäßige Aktivität. Die Zugangsmöglichkeiten zur manifesten und zu den latenten Funktionen von Erwerbsarbeit haben nach dem Modell einen direkten Einfluss auf die psychische Gesundheit. Da Erwerbslose nur eingeschränkt Zugang zu den Funktionen haben, weisen Sie mit höherer Wahrscheinlichkeit psychische Beschwerden auf als Erwerbstätige.

Ähnlich wie in der Deprivationstheorie von Jahoda gehen Warr et al. in ihrem "**Vitamin-Modell**" davon aus, dass externe Faktoren die psychische Instabilität des Erwerbslosen verantworten (Warr & Jackson, 1987). 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. Niedrige Ausprägungen bzw. das Fehlen dieser Funktionen führen zu negativen psychischen Auswirkungen. Höhere Ausprägungen der Faktoren führen zu positiven psychischen Resultaten. Jenseits einer bestimmten Grenze haben höhere Ausprägungen jedoch keine weiteren positiven Effekte und können, wie im Fall von Abwechslung, schädlich sein – analog zur Wirkung bestimmter Vitamine (daher der Name des Modells). Erwerbslose sind normalerweise mit Umweltbedingungen konfrontiert, die ein geringes Maß an „Vitaminen“ bereitstellen, was sich negativ auf das psychische Befinden auswirkt. An dem Modell fällt auf, dass „Verfügbarkeit von Geld“ nur ein Punkt unter vielen ist und keinen Schwerpunkt, wie bei Jahoda, darstellt.

Das "**Handlungsrestriktionsmodell**" von Fryer (1986) distanziert sich von dem deterministischen Menschenbild, welches den beiden vorigen Theorien zugrunde liegt. Fryer geht vom Konzept eines selbständig planenden und handelnden Individuums aus, das intrinsisch motiviert ist. Das Verlangen nach Selbstbestimmung wird durch die Arbeitslosigkeit massiv frustriert: Für Erwerbslose ist die zeitliche Perspektive der Situation unklar, ihre Handlungsmöglichkeiten sind durch relative oder absolute finanzielle Verarmung beschränkt. Diese Formen der Handlungsrestriktion führt zu vermindertem Wohlbefinden und schließlich zu psychischen Problemen. Im Gegensatz zu den anderen theoretischen Überlegungen wird die Rolle der Armut bei Fryer besonders betont.

Die "**Selektionshypothese**" (z.B. Mastekaasa, 1996)) postuliert, dass der Zusammenhang zwischen Arbeitslosigkeit und eingeschränkter psychischer Gesundheit 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 letztlich in die Langzeitarbeitslosigkeit hineinselektiert werden.

Die "**Inkongruenzhypothese**" (Paul & Moser, 2006) fußt auf der empirischen Beobachtung, dass selbst unter Arbeitslosen noch 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 regelmäßig in einer Situation mangelnder Passung ihrer Einstellungen zur Erwerbsarbeit mit ihrer realen Erwerbssituation. Dies erzeugt eine gravierende kognitive Dissonanz (Festinger, 2001). Situationen, die durch eine Inkongruenz von Zielen und Einstellungen einerseits und den realen Gegebenheiten andererseits gekennzeichnet sind, werden in diversen psychologischen Theorien als Ursache psychopathologischer Symptome angesehen (z.B. (Grawe, 2004)). Es wäre demnach möglich, dass bei Erwerbslosen das Festhalten an einer hohen Wertschätzung der Arbeit einer der Gründe für ihre eingeschränkte seelische Gesundheit darstellt.

Die meisten Erklärungsansätze sehen Erwerbslosigkeit als kausale Ursache des negativen seelischen Befindens und gehen nicht von einer umgekehrten Verursachung in dem Sinne aus, dass psychische Beanspruchungssymptome zur Arbeitslosigkeit führen. Prima vista widerspricht der Selektionshypothese bereits die Sachlage, dass die meisten Menschen in Folge von strukturell bedingten Massenentlassungen arbeitslos wurden, wie z.B. in Ostdeutschland nach „der Wende“ durch zahlreiche Werksschließungen geschehen oder in Westdeutschland im strukturellen Wandel des Ruhrgebiets. Dass Massenentlassungen selektiv die psychisch Instabilen traf, entspricht wohl kaum den Tatsachen. Trotzdem können unterschiedliche Verbleibzeiten in der Arbeitslosigkeit mit der Selektionshypothese erklärt werden.

Wenn man nach Beweisen dafür sucht, ob die Erwerbslosigkeit mitsamt ihren Rahmenbedingungen eine (externe) Belastung der Psyche darstellt oder ob schlicht die Personen mit einer hohen psychischen Beanspruchungssymptomatik in die Arbeitslosigkeit hineinselektiert werden, ist man auf die Analyse von Längsschnittdaten angewiesen.

Zur Prüfung der Kausalhypothese wurde in den Metaanalysen von Paul et al. (2006), sowie Paul & Moser (2009) untersucht, ob in Längsschnittstudien ein Wechsel von der Erwerbstätigkeit zur Arbeitslosigkeit mit einer Verschlechterung im psychischen Befinden, oder, ob umgekehrt ein Wechsel von der Arbeitslosigkeit zur Erwerbstätigkeit mit einer Verbesserung des Befindens einherging. Die Autoren analysierten 87 Längsschnittstudien, in denen Personen zwischen dem ersten und dem zweiten Erhebungszeitpunkt einen Wechsel im Erwerbsstatus durchmachten. Es fand sich eine signifikante Verschlechterung der

psychischen Gesundheit für die Personen, die einen Verlust des Arbeitsplatzes ( $d=-0.19$ ) erlebten. Parallel dazu ergab sich eine signifikante Verbesserung der psychischen Gesundheit von Erwerbslosen, die zwischen den Messzeitpunkten einen neuen Arbeitsplatz fanden ( $d=0.35$ ). Diese Veränderungsergebnisse unterstützen die Annahme, dass Erwerbslosigkeit nicht nur mit psychischer Beanspruchung korreliert, sondern sie kausal auslöst.

Für Selektionseffekte erfolgte die Prüfung, ob Erwerbstätige, die im Längsschnitt später arbeitslos wurden, sich in ihrer psychischen Gesundheit bereits während ihrer Berufstätigkeit von den Erwerbstätigen unterschieden, die ihre Anstellung behielten. Oder, ob die Erwerbslosen, die später eine Stelle fanden, sich von kontinuierlich Erwerbslosen in den Werten der Erstdiagnose unterschieden. Paul & Moser führten dazu metaanalytische Querschnittsanalysen des ersten Messzeitpunktes durch und fanden, dass erwerbstätige Personen, die im Verlauf der Studie ihren Arbeitsplatz verloren hatten, bereits zuvor einen schlechteren psychischen Gesundheitszustand ( $d=0.23$ ) als kontinuierlich Erwerbstätige aufwiesen ( Paul & Moser, 2009). Zusätzlich konnten zum ersten Messzeitpunkt bei den kontinuierlichen Arbeitslosen signifikant stärkere psychische Beanspruchungssymptome festgestellt werden, als bei den Erwerbslosen, die zum zweiten Messzeitpunkt eine neue Stelle gefunden hatten ( $d=0.15$ ). Den Autoren zufolge kann festgehalten werden, dass Arbeitsplatzverlust und Wiederbeschäftigung sich in den ursprünglichen Testwerten für die psychische Gesundheit „ankündigen“, was für die Existenz von Selektionseffekten auf dem Arbeitsmarkt spräche. Zu ihrem metaanalytisch ermittelten Selektionseffekt bemerken Paul & Moser allerdings kritisch, dass trotzdem unklar bleibe, ob die gefundenen Unterschiede des ersten Messzeitpunkts tatsächlich Selektionseffekte widerspiegeln. Antizipationseffekte wären ebenfalls eine mögliche Erklärung für die Befunde. Menschen, die demnächst ihre Stelle verlieren werden, könnten bereits im Vorfeld entsprechende Zeichen und Hinweise wahrnehmen und eine Entlassung befürchten. Wenn ein möglicher Arbeitsplatzverlust droht (z.B. bei Insolvenz eines Betriebes oder bei befristeten Arbeitsverträgen), kündigt sich dies normalerweise längere Zeit vor dem Eintreffen des Ereignisses an. Bereits die Angst vor dem Arbeitsplatzverlust wirkt sich negativ auf die psychische Gesundheit aus und könnte zu den gefundenen Unterschieden im Befinden zwischen kontinuierlich Beschäftigten und unsicher Beschäftigten führen. In diese Argumentationslinie reiht sich das Ergebnis von Berth et al. ein, dass bereits ein drohender Arbeitsplatzverlust mit einer Odds-Ratio von 5,3 psychische Beschwerden auslöst (Berth, Förster, Balck, Brähler, & Stöbel-Richter, 2008). Denkbar ist auch, dass diejenigen, die z.B. aufgrund ihrer Qualifikation oder ihres Alters geringe Chancen auf eine schnelle Wiederbeschäftigung am Arbeitsmarkt haben, sich des Umstandes bewusst sind. Diese pessimistische, oder möglicherweise realistische Zukunftseinschätzung könnte eine Beeinträchtigung des Befindens zum ersten Testzeitpunkt auslösen. Paul et al. (2006) resümieren, dass solche Befindensunterschiede also nicht unbedingt Belege für Selektionseffekte sein müssen, sondern auch das Resultat von Antizipationseffekten sein könnten. Von daher empfehlen die Autoren, dass die im Querschnitt ermittelten Effektstärken lediglich als Obergrenzen möglicherweise existierender Selektionseffekte interpretiert werden sollten.

Alles in allem ergaben die metaanalysierten Längsschnittdaten der Autoren das Bild eines Teufelskreises, in dem psychisch beanspruchte Menschen eine etwas höhere Chance haben können, arbeitslos zu werden, sich ihre psychischen Symptome während der Zeit der Belastung durch die Erwerbslosigkeit (und deren finanzielle und sozio-psychologischen

Begleitschäden) verstärken und dadurch ihre Chancen auf Wiedereingliederung in den Arbeitsmarkt abnehmen, was zu längeren Verweildauern in der Erwerbslosigkeit führt.

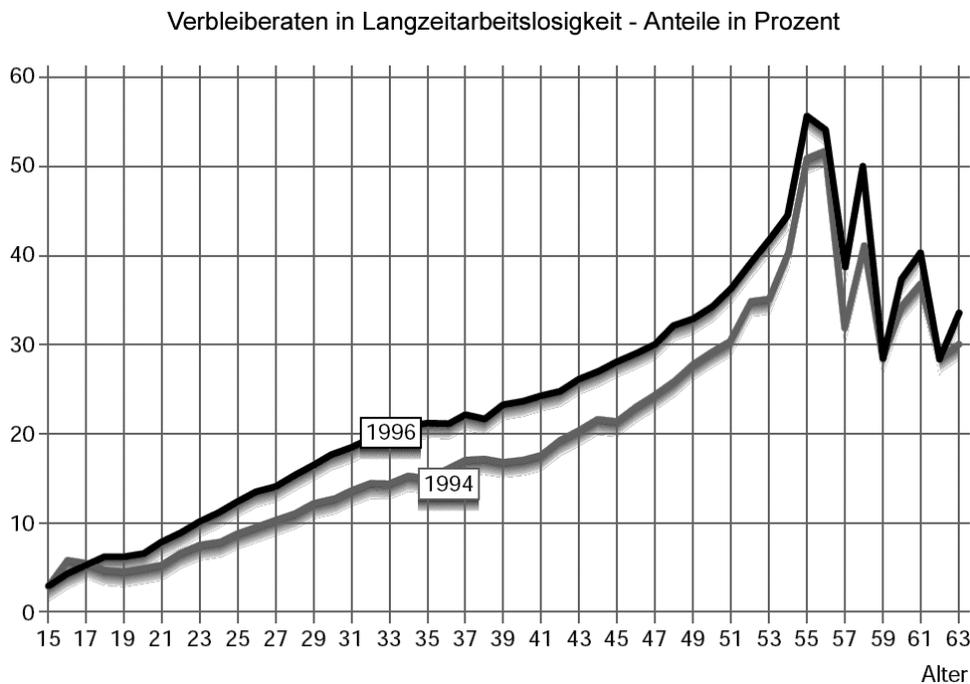
### 1.1.3. Einfluss des Alters

Winefield vertritt die Ansicht, dass im Gegensatz zu jüngeren oder älteren Menschen, diejenigen mittleren Alters stärker unter der Erwerbslosigkeit leiden müssten, weil in dieser Altersperiode die finanziellen familiären Verantwortlichkeiten am größten sind (Winefield, 1995). Die von der Arbeitslosigkeit verursachten finanziellen Verluste treffen bei Erwerbslosen mittleren Alters auf Menschen, die sich gerade ein eigenes Leben mit einer eigenen Familie aufbauen möchten. Junge Leute hätten hingegen bescheidenere Ansprüche, während bei älteren postuliert wird, dass sie die wichtigsten Anschaffungen bereits getätigt und familiäre Ziele schon erreicht hätten. Eine andere Ansicht ist, dass Erwerbsarbeit vor allem für Jugendliche eine wertvolle Unterstützung bei der Bewältigung der zentralen Entwicklungsaufgaben (Herausbildung einer eigenen Identität und Ablösung vom Elternhaus) darstellt (Schober, 1978). Demnach müssten insbesondere junge Erwerbslose Anzeichen eines negativen seelischen Befindens zeigen. Des Weiteren wurde behauptet, dass ältere Menschen häufig ihre Karriereziele hinter sich haben und eine auftretende Erwerbslosigkeit in eine Art vorgezogenen Ruhestand uminterpretieren könnten (Warr & Jackson, 1987). All diese Argumente implizieren, dass ältere Erwerbslose am wenigsten unter ihrem Zustand leiden sollten.

McKee-Ryan et al. (McKee-Ryan, Song, Wanberg, & Kinicki, 2005) stellten einen signifikanten Gruppenunterschied im Vergleich zwischen Schulabgängern und Erwachsenen in Bezug auf die Folgen von Erwerbslosigkeit fest. Dieser zeigte einen stärkeren negativen Effekt auf das Befinden der Jugendlichen (McKee-Ryan, Song, Wanberg & Kinicki, 2005). Murphy & Athansou, die ebenfalls Jugendliche mit Erwachsenen verglichen, fanden keinen signifikanten Effekt des Alters (Murphy & Athanasou, 1999). In der Metaanalyse von Paul & Moser (2009) wurde überprüft, ob es einen linearen oder einen kurvlinearen Zusammenhang zwischen dem Alter der Teilnehmer und der Effektstärke für die negativen Auswirkungen der Arbeitslosigkeit gab. Kein Zusammenhang war signifikant. Auch wenn die Analyse auf Stichproben beschränkt wurde, die ausschließlich aus Männern zusammengesetzt waren, zeigten sich keine signifikanten Zusammenhänge zwischen dem Alter von Erwerbslosen und dem Beschwerdedruck. Demnach ist ein Fazit bezüglich der Existenz eines Moderatoreffektes des Alters nicht möglich. Bemerkenswert ist, dass sich für die häufig vermutete kurvlineare Beziehung, mit schwachen negativen Auswirkungen der Arbeitslosigkeit bei jüngeren und älteren Menschen und starken negativen Effekten bei Menschen mittleren Alters, in keiner Metaanalyse empirische Hinweise fanden.

Dies bedeutet im Umkehrschluss, dass nicht davon auszugehen ist, dass Erwerbslosigkeit im fortgeschrittenen Alter mit signifikant geringerem psychischen Schaden einhergeht. Diese Auffassung ist vermutlich falsch. Bei der Diskussion über die moderierende Wirkung des Alters muss auch Notiz davon genommen werden, dass ältere Menschen (>50 Jahre) beim Auftreten von Erwerbslosigkeit 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 nochmals eine gravierende Verschlechterung der finanziellen Rahmenbedingungen. Die steigende

Zuwachsrate von Langzeiterwerbslosigkeit 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 Erwerbslosen abgetragen.



**Abb. 2:** Risiko der Langzeitarbeitslosigkeit in Abhängigkeit vom Alter für die Jahre 1994 und 96 (Quelle: IAB, 1999)

Deutlich erkennbar ist der sprunghafte Anstieg des Risikos ab dem 50. auf das 55. Lebensjahr. In diesem Alter weisen Menschen beim Auftreten von Arbeitslosigkeit eine über 50-prozentige Wahrscheinlichkeit auf, in dem Status zu verbleiben. Nach dem Gipfel nehmen die Verbleiberaten hauptsächlich aufgrund artifizierlicher Gründe ab, z.B. durch die damals noch forcierte „Frühverrentung“.

Es kann angenommen werden, dass die Bekanntheit des spezifischen Langzeitarbeitslosigkeitsrisikos für Menschen in der Altersphase ab 50 Angst vor der Arbeitslosigkeit auslöst. Zusätzlich dürfte das Wissen um die erheblich geringeren Chancen von älteren Erwerbslosen auf Wiedereingliederung in den Arbeitsmarkt Sorge um einen langen Verbleib in diesem Status auslösen und eine fatalistische Einstellung im Sinne von „Ab 50 wird man nicht mehr gebraucht!“ begünstigen.

#### 1.1.4. Einfluss der Arbeitslosigkeitsdauer

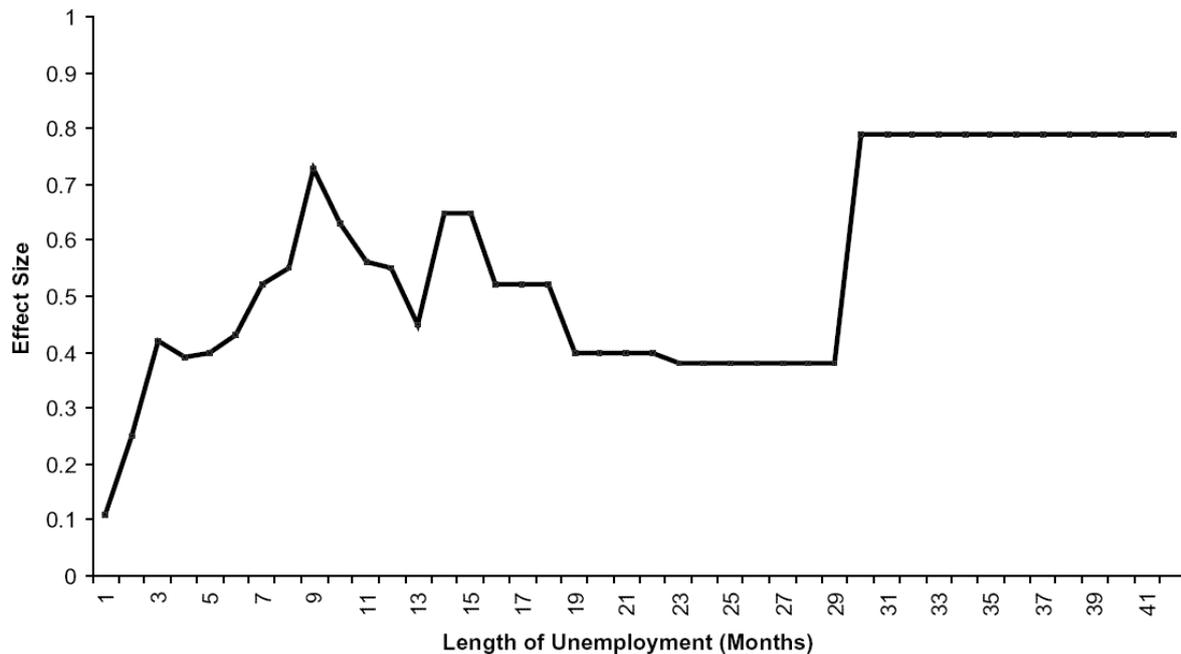
Jones analysierte 28 Längsschnittstudien zur Arbeitslosigkeitsdauer und kam zu der Einschätzung, dass psychische Beeinträchtigungen in der Antizipationsphase, d.h. kurz vor der Erwerbslosigkeit, am deutlichsten sind und zu diesem Zeitpunkt teilweise sogar noch gravierender ausfallen können, als direkt nach dem Arbeitsplatzverlust (Jones, 1991). In den ersten sechs Monaten nach Verlust kommt es demnach zu einer Symptomzunahme und im zweiten Jahr der Arbeitslosigkeit zu einer Abnahme der Symptome.

Im Laufe der wissenschaftlichen Theoriebildung zu den psychologischen Auswirkungen der Erwerbslosigkeitsdauer wurden zahlreiche Phasenmodelle formuliert. Bereits 1935 entwarfen Zawadski & Lazarsfeld ein solches Modell (Zawadski & Lazarsfeld, 1935). Auch in der neueren Literatur finden sich Phasenmodelle, die zum Teil aus Analogien mit anderen, kritischen Lebensereignissen entwickelt wurden. Laut Hill ähnelt der Verlauf der Arbeitslosigkeit den Stufen des Trauerprozesses (Hill, 1978). Winegardner et al. transferieren die von Kübler-Ross an Todkranken gefundenen Sterbeverarbeitungsphasen (Kübler-Ross, 1980) auf die der Arbeitslosigkeit (Winegardner, Simonetti, & Nykodym, 1984). Eine Übereinstimmung in den Phasenmodellen findet sich laut Paul & Moser (2009) in folgenden Phasen zur Verarbeitung der Erwerbslosigkeit: 1. Phase: Gemessen am kritischen Ereignis ein einigermaßen gutes Befinden zu Beginn der Arbeitslosigkeit. 2. Phase: Eine deutliche Befindensverschlechterung. 3. Phase: Nach diversen Zwischenphasen eine resignative Gewöhnung verbunden mit einer Stabilisierung des Befindens auf niedrigem Niveau.

Jackson & Warr formulierten ein "additives Stressmodell" zur Arbeitslosigkeitsdauer. Laut diesem kommt es im Rahmen der Stellensuche zu Zurückweisungen, die zunehmend schmerzhafter werden, da sie endgültiges Versagen zu bedeuten scheinen. Finanzielle Ressourcen werden knapper; die Ersparnisse sind aufgebraucht und die Unterstützungszahlungen werden mit zunehmender Arbeitslosigkeitsdauer geringer. Zudem steigt die soziale Isolation und Stigmatisierung an ( Jackson & Warr, 1984). Nach dem Modell additiver Stressoren müsste sich die psychische Gesundheit kontinuierlich mit steigender Arbeitslosigkeitsdauer verschlechtern und sich ein linearer Zusammenhang zwischen der Zeit- und der Beanspruchungsvariable ergeben. Ergänzend dazu haben Warr & Jackson ein "Adaptionsmodell" formuliert. Gemäß diesem entwickeln Arbeitslose nach einiger Zeit eine bessere Vertrautheit mit ihrer neuen Situation, die es ihnen ermöglicht, hilfreiche Copingstrategien und andere Formen der Anpassung zu entwickeln (Warr & Jackson, 1987). Einer anfänglichen Verschlechterung der psychischen Gesundheit folgt demnach eine Stabilisierung (auf niedrigem Niveau). Das Adaptionsmodell prognostiziert insofern eine kurvlineare Beziehung zwischen der Dauer und dem Befinden.

Metaanalytisch konnten McKee-Ryan et al. belegen, dass die psychische Gesundheit Arbeitsloser, die länger als 6 Monate arbeitslos waren, signifikant schlechter war als die von Kurzzeit-Arbeitslosen (McKee-Ryan, et al., 2005). Paul & Moser (2009) fanden einen hochsignifikanten linearen Zusammenhang zwischen der Dauer der Arbeitslosigkeit und der psychischen Beanspruchung, der das additive Stressmodell stützt. Der statistische Test auf einen kurvlinearen Zusammenhang war ebenfalls signifikant, was darauf hinweist, dass der vermutete Adaptionsprozess existiert. Ihre Analysen zeigen, 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.



**Abb. 3:** Effektstärken psychischer Beschwerden im Zusammenhang mit der Arbeitslosigkeitsdauer (K.I. Paul & K. Moser, 2009), S. 274). Die Werte ab dem 23. Monat repräsentieren keine genauen Veränderungen, da sie nur durch wenige Primärstudien repräsentiert werden.

Insgesamt kann als erwiesen gelten, dass die Arbeitslosigkeitsdauer mit einer zunehmenden Verschlechterung der psychischen Gesundheit zusammenhängt, sich aber immerhin innerhalb der ersten zwei Jahre Hinweise auf eine Stabilisierung der psychischen Gesundheit (auf niedrigem Niveau) finden lassen. Die Entwicklung der psychischen Gesundheit bei extrem langer Arbeitslosigkeit (>3 Jahre) konnten Paul & Moser (2009) nicht mehr differenziert klären, da für diesen Zeitraum nur noch wenige Primärstudien vorhanden waren. Lahelma nimmt an, dass zumindest in Bezug auf die psychische Beeinträchtigung eine weitere Steigerung bei sehr langer Erwerbslosigkeit (>3 Jahre) nicht mehr stattfindet (Lahelma, 1989). In Bezug auf die Begrifflichkeiten stellt sich Mohr jedoch explizit dagegen, den möglichen Stillstand in der Zunahme des psychischen Beschwerdedrucks mit dem eher positiv belegten Begriff der „Adaption“ zu bezeichnen (Mohr, 2001). Der Preis für den Stillstand – so kann der Studie von Zippay entnommen werden – ist der resignative Rückzug vom Arbeitsmarkt und die mit erfolglosen Bewerbungen verbundenen, als frustrierend erlebten, Misserfolgserfahrungen (Zippay, 1995). Die Befunde, wonach bei extrem lang andauernder Erwerbslosigkeit (>3 Jahre) depressive Reaktionen nicht weiter zunehmen, schließen jedoch eine Fluktuation der Depressivität über die Zeit nicht aus. Strehmel & Mayring belegten in ihrer Untersuchung eine Verschlechterung des Befindens bei sehr langfristig erwerbslosen Lehrern in Abhängigkeit von den jährlichen Einstellungsterminen und den damit verbundenen Absagen auf Bewerbungen (Strehmel & Mayring, 1986). Solche Veränderungen nach sehr langer Zeit der Arbeitslosigkeit werden auch von Leana & Feldman) dargestellt, dort in Verbindung mit Veränderungen der finanziellen

Rahmenbedingungen (Leana & Feldman, 1992). Finanzielle Ressourcen spielen selbstverständlich eine zentrale Rolle bei der Bewältigung, insbesondere der Langzeitarbeitslosigkeit. Frese & Mohr wiesen nach, dass die, durch finanzielle Einschränkungen verursachte, Nicht-Kontrolle über die eigene Lebenssituation klinisch relevante Depressionen bei älteren Langzeiterwerbslosen verursacht (Frese & Mohr, 1987).

## **1.2. Physische Gesundheit und Arbeitslosigkeit**

### **1.2.1. Allgemeine Erkenntnislage**

Im Vergleich zur Vielzahl von Befunden einer jahrzehntelangen Forschung zu den psychischen und psychosozialen Auswirkungen von Arbeitslosigkeit existiert nur eine verhältnismäßig kleine Anzahl von Untersuchungen zu den Auswirkungen auf die subjektive und objektive physische Gesundheit.

McKee-Ryan et al. analysierten in einer Meta-Studie 104 empirische Studien zur Gesamtheit der gesundheitlichen Auswirkungen des Arbeitsplatzverlustes. Auch hier beschäftigte sich der größte Teil der Studien mit den Auswirkungen von Arbeitslosigkeit auf die psychische Gesundheit, trotzdem zeigte sich, dass Erwerbslose im Vergleich zu Beschäftigten auch einen ungünstigeren physischen Gesundheitszustand aufwiesen (McKee-Ryan, Song, Wanberg, & Kinicki, 2005). Insgesamt wurden über 100 Variablen gefunden, die bei eingetretener Arbeitslosigkeit im Zusammenhang mit verschiedenen Indikatoren der psychischen oder physischen Gesundheit korrelierten. Erfasst wurde das subjektive körperliche Wohlbefinden meist in Form von Selbstberichten (Gallo, Bradley, Siegel, & Kasl, 2000; Jackson, Stafford, Banks, & Warr, 1983; Schwarzer, Jerusalem, & Hahn, 1994), sehr viel seltener wurden jedoch objektive medizinische Befunde oder Diagnosen erhoben (Brenner & Starrin, 1988; Claussen, 1994; Hall & Johnson, 1988). Gerade auf dem Gebiet der Erfassung physiologischer Parameter und Auswirkungen gesundheitsrelevanter Verhaltensweisen im Zusammenhang mit Arbeitslosigkeit liegt daher noch ein erheblicher Forschungsbedarf vor.

In einer 1999 erschienenen Übersichtsarbeit von Hanisch, in der die englischsprachige, zwischen 1994 und 1998 publizierte Literatur zu den Auswirkungen von Erwerbslosigkeit analysiert wurde, finden sich konkrete Symptome und körperliche Befindlichkeitsstörung, deren Prävalenz bei Arbeitslosen im Vergleich zu Beschäftigten erhöht ist (Hanisch, 1999; Rásky, Stronegger, & Freidl, 1996). Dazu zählen neben Kopfschmerzen und gastrointestinalen Beschwerden auch Schlafstörungen und Kraftlosigkeit. Erwähnung finden ebenso Blutdrucksteigerungen, Magengeschwüre, Bronchialerkrankungen und Fettstoffwechselstörungen, sowie als Todesursachen Schlaganfall und kardiovaskuläre Erkrankungen. Erhöhte Prävalenzen für spezifische Krankheitsbilder wie chronische Bronchitis, obstruktive Lungenerkrankungen und chronisch-ischämische Herzkrankheit bei Erwerbslosen im Vergleich zu Erwerbstätigen konnten auch Cook et al. nachweisen (Cook, Cummins, Bartley, & Shoper, 1982). Klare Ursachen-Wirkungsbeziehungen zur Krankheitsentstehung ließen sich bisher nicht postulieren. Erschwerend kommt hinzu, dass es sich bei der untersuchten Population außer in Bezug auf das Merkmal Erwerbslosigkeit keinesfalls um eine homogene Gruppe handelt.

Anders als bei den psychischen Folgen finden sich in der ersten Phase der Arbeitslosigkeit oft keine Anzeichen einer Verschlechterung der körperlichen Gesundheit. Insbesondere bei vorher körperlich belasteten Beschäftigten kann zunächst sogar eine physische Entlastung

und damit subjektive Verbesserung der körperlichen Gesundheit wahrgenommen werden (Brinkmann & Potthoff, 1983).

Viele moderierende Faktoren, die hinsichtlich des Einflusses auf die psychische Gesundheit bei Arbeitslosen gut untersucht sind wie z. B. finanzielle Absicherung, Copingstrategien und soziale Unterstützung, können bzgl. ihrer Auswirkungen auf die physische Gesundheit aufgrund der geringen Anzahl systematischer Untersuchungen nicht beurteilt werden. Querschnittsstudien belegen allerdings, dass die Dauer der Arbeitslosigkeit einen wesentlichen Einflussfaktor sowohl auf die psychische als auch auf die physische Gesundheit darstellt (Hollederer, 2002; McKee-Ryan, et al., 2005; Warr & Jackson, 1984). Resetka zufolge wird das Auftreten körperlicher Symptome auf dem Boden einer andauernden psychosozialen Stressreaktion v.a. durch das Ausmaß der emotionalen Reaktion auf die Arbeitslosigkeit beeinflusst (Resetka, Liepmann, & Frank, 1996). Betroffene litten daher häufig unter psychosomatischen Beschwerden wie Verdauungsstörungen, Erkrankungen der Atemwege und Muskel- und Skeletterkrankungen.

### **1.2.2. Inanspruchnahme medizinischer Versorgungsleistungen**

Bereits in den 80-iger Jahren wurde ermittelt, dass der relative Anteil von Arztbesuchen bei Arbeitslosen um 4-10% höher liegt als bei Erwerbstätigen (Resetka, et al., 1996). Auch in der durchschnittlichen Bezugsdauer des Krankengeldes (Hollederer, 2002) und der Krankenhausliegezeit gibt es deutliche Unterschiede zwischen Erwerbslosen und Erwerbstätigen. Der Gesundheitsreport 2010 der Techniker Krankenkasse (TK) belegt erneut, dass Arbeitslose in nahezu allen Diagnosen häufiger betroffen sind, als alle anderen Personengruppen (Grobe & Dörning, 2010). Im Jahr 2010 waren Arbeitslose mit durchschnittlich 20,3 Tagen mehr als 5 Tage länger arbeitsunfähig als 10 Jahre vorher. Im Gesundheitsreport 2009 des BKK-Bundesverbandes wurde der mit durchschnittlich 25,8 Krankheitstagen höchste Krankenstand in der Gruppe der Arbeitslosengeld-I-Empfänger registriert (BKK-Bundesverband, 2009). Die häufigsten Ursachen für Arbeitsunfähigkeit waren auch hier Muskel- und Skeletterkrankungen (31,7%) und psychische Störungen (24%), sowie Krankheiten des Atmungs-, Verdauungs-, Herz-Kreislaufsystems und Neubildungen mit jeweils rund 5%.

Betrachtet man die Krankenhausbehandlungen, müssen sich Arbeitslose doppelt so oft einer Krankenhausbehandlung unterziehen wie pflichtversicherte Beschäftigte und verbringen gegenüber diesen die 2, 8-fache Zeit stationär (BKK-Bundesverband, 2009).

Diese Ergebnisse gehen konform mit epidemiologischen Untersuchungen aus Großbritannien, die zeigen konnten, dass sich Erwerbstätige in sicheren Beschäftigungsverhältnissen schneller von Krankheit erholen als Erwerbslose oder geringfügig bezahlte Erwerbstätige (Dorling, 2009).

Obwohl auch bei den Erwerbstätigen psychisch bedingte Krankschreibungen in der letzten Dekade um 40% zugenommen haben, gibt es auffallend große Unterschiede zwischen Arbeitslosen und Erwerbstätigen bei der Verordnung von Antidepressiva. Arbeitslose Frauen erhielten laut TK-Gesundheitsreport 2010 doppelt so viele Antidepressiva wie berufstätige Frauen, arbeitslose Männer liegen sogar um 200% über dem Verordnungsvolumen der Erwerbstätigen.

### **1.2.3. Arbeitslosigkeit und Mortalität**

Für einen Zusammenhang zwischen Arbeitslosigkeit und einer erhöhten Mortalität gibt es relativ viele Belege aus dem europäischen und angloamerikanischen Raum.

Eine aktuelle Studie des Max-Planck-Instituts für demographische Forschung, des Rostocker Zentrums zur Erforschung des Demographischen Wandels und des Forschungsdatenzentrums der Rentenversicherung belegte kürzlich, dass durch den Faktor Arbeitslosigkeit die Lebenserwartung ostdeutscher Männer im Vergleich zu westdeutschen Männern um 1 Jahr sinkt. Bei Frauen wirkt sich hingegen Arbeitslosigkeit nicht auf die Lebenserwartung aus (Scholz, Schulz, & Stegmann, 2010).

Morris et al. konnten bei Untersuchungen in Großbritannien, selbst unter Berücksichtigung verschiedener Einflussfaktoren wie Gesundheitsverhalten, initialer Gesundheitszustand und soziale Schicht signifikant höhere Sterblichkeitsraten (adjustiertes relatives Risiko > 2,0) durch Herz-Kreislauferkrankungen und Krebs bei Personen nachweisen, die im Beobachtungszeitraum Arbeitsloskeitsperioden aufwiesen (Morris, Cook, & Shaper, 1994). Auch Auswertungen von AOK-Daten aus Süddeutschland, erhoben zwischen 1980 und 1990 belegen eine 2,6-fach höhere relative Sterblichkeit von Arbeitslosen gegenüber Erwerbstätigen (Schach, Rister-Mende, Schach, Glimm, & Wille, 1994). Eine große schwedische Studie mit 30 000 untersuchten Personen zwischen 20 und 64 Jahren und einer Nachbeobachtungszeit zwischen 10 und 17 Jahren kam unter Berücksichtigung des initialen Gesundheitszustandes zu vergleichbaren Ergebnissen (Gerdtham & Johannesson, 2003), ähnliche Befunde finden sich auch bei Iversen (Iversen, Andersen, Andersen, Christoffersen, & Kieding, 1987) und in den Arbeiten von Moser et al. (Moser, Fox, & Jones, 1984; Moser, Fox, Jones, & Goldblatt, 1986) ohne Berücksichtigung des initialen Gesundheitsstatus. Gerdtham & Johannesson interpretierten den Zustand der Arbeitslosigkeit als einen Risikofaktor für eine erhöhte Gesamtsterblichkeit, vergleichbar mit dem Rauchen oder einem erhöhten Blutdruck. Einen direkten Zusammenhang zur kardiovaskulären Mortalität konnten sie aber nicht nachweisen. Bekannt ist allerdings, der Einfluss des sozioökonomischen Status auf das kardiovaskuläre Risiko (Kaplan & Keil, 1993; Rogot, Scorlie, Johnson, & al., 1992; Rose & Marmot, 1981; Thurston, Kubzansky, Kawachi, & Berkman, 2005).

Die genauen Mechanismen, die zur Erhöhung der Gesamtsterblichkeit bei Arbeitslosen führen, konnten bislang nicht geklärt werden. Studien, die Veränderungen physiologischer Variablen, z.B. biochemischer oder immunologischer Parameter im Zusammenhang mit Arbeitslosigkeit untersuchten, fanden keine signifikanten Assoziationen (Arnetz, Wasserman & Petrini, 1987; Flemming, Baum, Reddy & Gatchel, 1984; Kasl, Cobb & Brooks, 1968).

### **1.2.4. Arbeitslosigkeit und verändertes Gesundheitsverhalten**

Eine Reihe von Studien untersuchte Änderungen des Gesundheitsverhaltens bei von Arbeitslosigkeit Betroffenen. Gesundheitsschädliche Verhaltensweisen wie Rauchen, Alkoholkonsum und eine gesteigerte Medikamenteneinnahme sind in der Gruppe der Erwerbslosen stärker ausgeprägt als in der Vergleichsgruppe der Erwerbstätigen. Statistisch gesicherte Belege hinsichtlich des unterschiedlichen Rauchverhaltens finden sich für beide Geschlechter in Metaanalysen (McKee-Ryan et al., 2005) ebenso wie in den Nationalen Gesundheitssurveys Deutschlands 1984-86 und 1987-89 des Robert Koch-Institutes (Helmert & Maschewsky-Schneider, 1998) sowie im Gesundheitssurvey Österreichs (Rásky, et al., 1996). Demnach unterscheidet sich nicht nur der Anteil der Raucher, sondern auch die

Intensität des Rauchens, gemessen am Anteil der starken Raucher mit 20 und mehr Zigaretten täglich (Bormann, 1992; Schach, et al., 1994). Dagegen fanden Cook et al. signifikant höhere Raucherquoten nur bei weiblichen Arbeitslosen im Vergleich zu weiblichen Beschäftigten (Cook et al., 1982). Im Bundesgesundheitsurvey 1998 waren die Unterschiede bei Frauen gering ausgeprägt, bei Männern dagegen stark (Grobe & Schwartz, 2003). Auch aktuelle empirische Analysen der Daten des telefonischen Gesundheitssurveys 2003, des Mikrozensus 2005 und des Sozio-ökonomischen Panels (SOEP) 2006 belegen, dass sich hohe Raucherquoten insbesondere bei Personen in dauerhaft benachteiligten Lebenslagen, z.B. bei Langzeitarbeitslosigkeit finden (Lampert, 2010).

Im Vergleich zum Rauchverhalten haben ein riskanter Alkoholkonsum, alkoholbezogene Verhaltensprobleme bzw. Alkoholkrankheit häufig direkte Auswirkungen auf den Arbeitsvermittlungsprozess. Zum Zusammenhang zwischen Arbeitslosigkeit und Alkoholismus gibt es seit langem eine kontroverse Diskussion (Henkel, 1998). Einige Studien deuten daraufhin, dass kein oder sogar ein negativer Bezug zwischen Arbeitslosigkeit und Alkoholgebrauch besteht (Cook & Allan, 1983; Cook et al., 1982; Heather, Laybourn & MacPherson, 1987; Morris, Cook & Shaper, 1992), andererseits belegen eine Vielzahl von empirischen Befunden den positiven Zusammenhang zu Trinkgewohnheiten und Verhaltensauffälligkeiten in deren Folge und zur Alkoholkrankheit (Bloomfield, 1998; Bray, Zarkin, & Dennis, 2000; Brenner, 1987; Crawford, Plant & Kreitman, 1987; Montgomery, Cook, Bartley et al., 1998; Smart, 1979). Zusammenfassend kann von einem äußerst komplexen Prozess ausgegangen werden, der durch zahlreiche Faktoren wie persönliche Ressourcen, soziale Konflikte, chronische finanzielle Probleme u.a.m. beeinflusst wird. Die Reduktion von negativen Emotionen, insbesondere von Depressivität, scheint ebenso wie die individuelle Bewältigungskompetenz von Problemen eine wichtige Rolle zu spielen (Henkel, Dornbusch & Zemlin, 2005; Puls, Inhester, Reinecke & Wienold, 2005).

Rásky et al. analysierten in den 1990-iger Jahren epidemiologische Daten des Gesundheitssurveys in Österreich hinsichtlich möglicher Unterschiede zwischen Erwerbslosen und Beschäftigten in Bezug auf gesundheitsrelevante Verhaltensweisen. Dabei stellten sie fest, dass männliche Arbeitslose seltener körperlich aktiv sind und weniger motiviert, ihr Körpergewicht zu reduzieren oder ihre Ernährungsgewohnheiten zu ändern als männliche Erwerbstätige. Zwischen Frauen und Männern zeigten sich unabhängig vom Erwerbsstatus signifikante Unterschiede im Gesundheitsverhalten. (Rásky et al., 1996). Konsistent zur Studie von Wingard achteten Frauen generell mehr auf einen gesunden Lebensstil (Wingard, 1984).

### **1.3. Besonderheiten der Langzeitarbeitslosigkeit**

In Deutschland ist mehr als die Hälfte der Arbeitslosen ein Jahr oder länger erwerbslos und somit "langzeitarbeitslos". Deutschland gehört mit ca. einer Million zu den vier der (damals) 25 EU-Länder mit dem höchsten Anteil an Langzeiterwerbslosen (STABU, 2006). Der tatsächliche Anteil der Betroffenen liegt allerdings deutlich höher. In den amtlichen Statistiken fehlen jene, die infolge des Arbeitsplatzverlustes frühzeitig in Rente gegangen sind. Auch diejenigen, welche durch die Einkünfte ihres Partners nicht zum Leistungsbezug berechtigt sind und sich von daher nicht arbeitslos melden, fehlen. Des Weiteren gehen Personen, die aufgrund von Erkrankungen nicht mehr ausreichend arbeitsfähig sind, nicht mehr in die Statistiken ein (Hollederer, 2002). Das Risiko der Langzeiterwerbslosigkeit steigt

mit zunehmendem Lebensalter, mit fehlendem Qualifikationsniveau und mit gesundheitlichen Einschränkungen.

Laut Mohr ist evident, dass auch die Langzeitarbeitslosen keine homogene Gruppe bilden (Mohr, 2001). Einige der von ihr revidierten Längsschnittstudien lassen es als angemessen erscheinen, zwischen *lange* Erwerbslosen (>1Jahr) und *sehr lange* Erwerbslosen (>3 Jahre) zu differenzieren. Shlay & Rossi postulieren, dass unter den sehr lange Erwerbslosen die Gruppe der "sozial Desintegrierten" eine relevante Subgruppe darstellt (Shlay & Rossi, 1992). Die Autoren vermuten, dass sich sehr langfristig erwerbslose Personen in ihrem Bewältigungsverhalten und in ihrem Befinden von solchen Langzeitarbeitslosen unterscheiden, die auch nach einjähriger Erwerbslosigkeit noch eine reale Chance auf dem Arbeitsmarkt haben. Mohr gibt zu bedenken, dass ein allgemeingültiges Kriterium zur Trennung der Langzeitarbeitslosen (>1 Jahr) von den sehr langen Langzeitarbeitslosen (>3 Jahre) schwierig zu formulieren ist. Vielmehr seien psychologische Kriterien zu erheben, wie z.B. die Aufgabe der Wiedervermittlungshoffnung. Diese Kriterien unabhängig zu ermitteln, ist jedoch fast unmöglich, da sie stark von externen Merkmalen, wie z.B. der regionalen Arbeitsmarktlage, abhängen.

### 1.3.1. Relative Zunahme von Selektionseffekten

In einer umfassenden, US-amerikanischen epidemiologischen Studie (N=89.267) belegten Brown und Mitarbeiter, dass Langzeiterwerbslose gegenüber Beschäftigten eine annähernd doppelt so hohe Wahrscheinlichkeit aufweisen, innerhalb eines definierten Zeitraumes psychisch zu erkranken (Brown et al., 2003). Zu den gesundheitlichen Folgen vertritt Mohr die Ansicht, dass die negative kausale Wirkung der Arbeitslosigkeit auf das psychische Befinden bei Langzeiterwerbslosen einen Teufelskreis in Verbindung mit Selektionseffekten in Gang setzt (Mohr, 2001). Auch Zempel & Frese konstatieren auf der Grundlage ihrer Längsschnittdaten:

„...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).

Bereits 1994 spricht sich Clausen bei langfristig Erwerbslosen für das Wirken von Selektionseffekten aus, in dem er die psychische Gesundheit als Kriterium für eine Wiedereinstellung beschreibt (Clausen, 1994). Klein dokumentierte 1990 für die alten Bundesländer, dass zwei Drittel aller wieder eingestellten Langzeitarbeitslosen innerhalb von 4 Jahren erneut erwerbslos wurden (Klein, 1990). Auch dies könnte als Signal dafür verstanden werden, dass Teile der langfristig Erwerbslosen mehrfache Selektionsprozesse durchlaufen haben, welche zur weiteren Reduktion von Vermittlungschancen führen. Diese Menschen haben ein besonderes Risiko, Erwerbslose auf Dauer zu bleiben. Selektionsprozesse werden für das Alter, die Gesundheit und das Qualifikationsniveau der Betroffenen angenommen.

### 1.3.2. Rückgang des Qualifikationsniveaus

Mit den stetigen technologischen Weiterentwicklungen findet ein Lernen zunehmend innerhalb der Arbeitstätigkeit („learning on the job“) statt. Vor diesem Hintergrund bedeutet der Verlust einer Anstellung auch einen Verlust von Lern- und Qualifikationsmöglichkeiten. Bedenkt man die Halbwertszeit von IT-gestützten Arbeitsmethoden (Hard- und Software), dann erzeugt eine Erwerbslosigkeit von drei Jahren bereits eine so bedeutsame Lücke, dass sie ein gravierendes Vermittlungshemmnis darstellt. Laut Mohr kommt es bei Langzeitarbeitslosen vermutlich auch zu einem Verfall der „soft-skills“, wie z.B. der Kommunikationsfähigkeit, der Fähigkeit zu lernen oder zu organisieren (Gisela Mohr, 2001). Gerade für ältere Langzeiterwerbslose ist dieser Umstand besonders tückisch, da sie bereits aufgrund ihres Alters eine geringere Wiedereinstellungschance haben, die sich mit fortschreitenden Qualifizierungsdefiziten paart.

### 1.3.3. Eingeschränktes soziales Netz

Klein & Strasser stellten fest, dass die geschlechtsunabhängige Scheidungswahrscheinlichkeit bei Langzeiterwerbslosen fast dreimal höher ist als in der Allgemeinbevölkerung ( Klein & Strasser, 1997). Bei langfristig erwerbslosen Männern ist der Prozentsatz der Ehen, die innerhalb der ersten vier Jahre wieder geschieden werden sogar mehr als dreimal 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 weist darauf hin, dass vor allem die sog. „weak ties“ (lockere bzw. flüchtige Bekanntschaften und Beziehungen) nach längerer Erwerbslosigkeit drastisch abnehmen (Jackson, 1988). Dies ist arbeitsmarkttechnisch eine kritische Veränderung. Wie der Studie von Engelbrecht zu entnehmen ist, bietet vor allem das weitere soziale Netz – d.h. außerhalb des engen Familien- und Freundeskreises – Hilfen bei der Wiedervermittlung (Engelbrecht, 1987). 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.

### 1.3.4. Effekte der Wiedereinstellung auf die Gesundheit „Segensreichheit“ von Arbeit?

Bei der Wiederaufnahme einer Erwerbsarbeit kommt es prinzipiell zu einer Verbesserung des Befindens. Paul et al. berechneten in Meta-Längsschnittanalysen, dass für Erwerbslose nach einer neuen Anstellung eine Verbesserung ihrer psychischen Gesundheit mit einer Effektstärke von  $d=0.35$  zu finden war (Paul et al., 2006). Auch für die Gruppe der Langzeitarbeitslosen stellt die Mehrheit der Studien laut Murphy & Athanasou eine Verbesserung der psychischen Gesundheit nach einer Wiedervermittlung fest (Murphy & Athanasou, 1999). Nach Verkleij sind es sogar vor allem die wiederbeschäftigten Langzeiterwerbslosen (im Kontrast zu den eher kurzfristig Arbeitslosen), die eine deutliche Verbesserung ihrer psychosozialen Befindlichkeit aufweisen (Verkleij, 1989). Clausen berichtet, dass bei langfristig Erwerbslosen nach der Wiedereinstellung neben allgemeinen Verbesserungen des Befindens auch der Alkoholkonsum zurückgeht ( Clausen, 1999).

Er stellte jedoch auch fest, dass eine Erhöhung der psychischen Gesundheit nicht für alle Gruppen der 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. von Wanberg – vermehrt darauf hin, dass eine gesundheitliche Verbesserung bei Wiedereinstellung nicht für jene gilt, die in sog. „bad jobs“ (Tätigkeiten mit schlechter Bezahlung, hoher Repititivität, fehlender Entwicklungsmöglichkeit und unsicherer Zukunftsperspektive) vermittelt werden (Wanberg, 1995). Zu diesen Einschätzungen kommen auch die Vertreter des Konzepts der „Prekarisierung“ (Plöger, 2010). Auch Leana & Feldman berichten, dass Wiederbeschäftigte sich nicht besser fühlten als Erwerbslose, wenn die neue Tätigkeit nicht ihren Vorstellungen entsprach und sie mit der Stelle unzufrieden waren (Leana & Feldman, 1995). Demnach ist es nicht entscheidend, irgendeine Arbeit zu haben, sondern einer Tätigkeit nachzugehen, die den eigenen Bedürfnissen entspricht. Auch Broom und Mitarbeiter fanden in komparativen Berechnungen keine Unterschiede des negativen psychischen Befindens zwischen Berufstätigen mit prekären Berufsbedingungen und Arbeitslosen. Sobald die Stärke der von ihnen ermittelten beruflichen Stressoren einen bestimmten Schwellenwert überschritt, fühlten sich beide Gruppen gleichermaßen schlecht (Broom et al., 2006). Selbst Paul et al. sehen die Unterstellung einer „segensreichen“ Funktion für jegliche Formen der Erwerbsarbeit kritisch. Sie schreiben, dass solche verallgemeinernden Schlüsse eigentlich nicht zulässig seien, ohne dass Erwerbstätige nicht nur mit Erwerbslosen, sondern mit der gesamten Gruppe der sog. „out of labour force“ (zur Definition siehe Rengers (Rengers, 2005)) verglichen werden (Paul et al., 2006). Angehörige der "out of labour force" stellen nämlich nicht nur Erwerbslose, sondern z.B. auch Rentner, Hausfrauen/-männer, Studierende oder auch 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.

### **1.3.5. Chancen des Wiedereinstiegs**

Bereits vor 20 Jahren kam Klein zu dem pessimistischen Schluss, dass für die Zahl der langfristig Erwerbslosen kaum eine Reduzierung durch Wiedervermittlung feststellbar war (Klein, 1990). Hinzu kommt, dass knapp drei Viertel aller wiederbeschäftigten langfristig Erwerbslosen nach 4 Jahren erneut arbeitslos werden, die meisten von Ihnen bereits in den ersten neun Monaten nach der Wiedervermittlung (Mohr, 2001). Es handelt sich bei stabil wiedervermittelten Langzeiterwerbslosen also nur um eine Minderheit. Mehr Bewerbungsaktivitäten bieten offenbar keine Garantie für einen besseren Vermittlungserfolg: Zippay berichtet, dass ältere Erwerbslose sich in ihrem anfänglichen Arbeitssuchverhalten nicht von den jüngeren Wiedervermittelten unterschieden, sondern nur im Erfolg ihres Suchverhaltens (Zippay, 1995). Die Analyse amtlicher Statistiken lässt den Schluss zu, dass eine beeinträchtigte Gesundheit, das Alter, Geschlecht und die regionale Arbeitsmarktlage herausragende Determinanten für die Wiedervermittlungschancen von Langzeitarbeitslosen darstellen (Hollederer, 2003). Außer in Bezug auf die Gesundheit sind dies sämtlich Merkmale, auf die die Betroffenen keinen Einfluss haben.

#### **1.4. Maßnahmen zur Förderung von Erwerbslosen**

Will man den Erfolg von Interventionen für Erwerbslose evaluieren, so stellt sich zuerst die Frage, was als Erfolg definiert werden soll. Die dargestellten Forschungsbefunde legen nahe, nicht allein in einer höheren Wiedervermittlungsquote das Erfolgskriterium zu suchen. Unter dem Druck dieses Kriteriums würde die Gefahr steigen, dass lediglich den Erwerbslosen Förderungsmaßnahmen offeriert würden, für die ohnehin eine bessere Wiedervermittlungschance besteht. Oder, dass Fördermaßnahmen nur noch in Regionen mit starkem Arbeitsmarkt als profitabel angesehen werden könnten. Trube und Luschei fordern eine Erweiterung auf qualitative Zielkriterien und eine Evaluation, die Informationen enthält über Bedingungen von Erfolg und Misserfolg (Trube & Luschei, 2000). Ernst et al. weisen auf die Notwendigkeit hin, für die Bewertung von Bildungs- und Beschäftigungsmaßnahmen für Erwerbslose „...zusätzlich die Erfolge sozialer Integration, verbesserter Gesundheit und sozialer Kompetenzentwicklung zu dokumentieren“ (Ernst, Hartwig, & Pokora, 2008).

Eine Maßnahmenbeschreibung für Erwerbslose, die mindestens zwei Jahre arbeitslos waren, gibt Kieselbach (Kieselbach, 1998). Die Studie zeigt, wie schwierig (und teilweise unethisch) es ist, im Rahmen einer Maßnahme, die sich primär den Praxisanforderungen stellen muss, zusätzlich den wissenschaftlichen Design-Anforderungen des randomisierten kontrollierten Trials gerecht zu werden. Amundson & Borgen sprechen sich bei Langzeitarbeitslosen dafür aus, primär Hilfe bei der Bewältigung der Erwerbslosigkeit anzubieten, bevor Fähigkeiten für die Arbeitsuche aufgebaut werden (Amundson & Borgen, 1987). Der Aufbau einer realistischen Erwartungshaltung ist eine Gratwanderung, die es bei der Unterstützung von Langzeiterwerbslosen zu bewältigen gilt, resümieren Frese und Mohr, weil durch Interventionsmaßnahmen induzierte übersteigerte Hoffnungen zu Enttäuschungen führen, die sich in depressiven Reaktionen äußern können (Frese & Mohr, 1987). Mohr und Otto fordern, dass sich Maßnahmen für Erwerbslose vorwiegend an Risikogruppen richten sollten (Mohr & Otto, 2005). Interventionsangebote sollten den Anspruch haben, langfristig Erwerbslose, ältere Personen, Frauen, Geringqualifizierte und Personen mit physischen oder psychischen Gesundheitsbeeinträchtigungen zu fördern. Mohr beschreibt (Mohr, 2001), dass von älteren langfristig Erwerbslosen speziell supportives Verhalten bei der Arbeitssuche als stabilisierend erlebt wird. Rife & Belcher berichten in ihrer Studie, dass die Mehrheit ihrer befragten Langzeiterwerbslosen vor allem andere Arbeitslose als Quelle dieser Form der Unterstützung nannte (Rife & Belcher, 1993). Freunde und Familienmitglieder, die erwerbstätig oder in Rente waren, wurden von den Erwerbslosen kaum als Quelle für diese Art der Unterstützung genannt. Hieraus folgt, dass Interventionen als Gruppenmaßnahme angeboten werden sollten, damit langfristig Erwerbslose die Gelegenheit erhalten, sich untereinander über ihre Copingstrategien auszutauschen.

Mielck beschreibt, dass sich Arbeitslose zwar subjektiv eher krank fühlen, aber trotz dieses Bewusstseins weitaus weniger Angebote zur Prävention und Gesundheitsförderung wahrnehmen (Mielck, 2005). Dies hat nicht zuletzt finanzielle Gründe: Selbst bei Erstattung der Kursgebühren durch die Krankenkasse müssen die Kosten von den Teilnehmern häufig vorgestreckt werden, welches in Anbetracht der ALG-II-Regelsätze meist nicht leistbar ist. Implizit folgt daraus, dass eine Gesundheitsförderung als Maßnahme direkt von der

Institution, welche die Langzeiterwerbslosen betreut (meist die ARGE<sup>1</sup>) angeboten werden sollte, um die potentiellen Teilnehmer auch wirklich erreichen zu können. Da zu erwarten ist, dass eine verbesserte psychische und physische Gesundheit bei langfristig Erwerbslosen ein relevantes Wiedervermittlungshemmnis abbaut, sollte die Unterstützung von gesundheitsförderlichen Maßnahmen im Interesse der ARGE sein.

#### 1.4.1. Vorstellung von Interventionsstudien

In der Recherche zu Forschungsberichten, in denen psychologische Maße als Evaluationskriterien für einen Interventionseffekt bei Erwerbslosen herangezogen wurden, findet man hauptsächlich motivationsfördernde, sozialtherapeutische, psychologische bzw. "psychologienah" Interventionen, aber keine reine Gesundheitsförderung. Maßnahmen zur Erhöhung der Bewerbungsaktivitäten werden teilweise mit psychologischen Fragebögen evaluiert, allerdings beinhalten diese kein gesundheitlich relevantes Curriculum. Es sind keine primär auf die körperliche Gesundheit ausgerichteten Interventionen für Langzeiterwerbslose bekannt, welche psychologische Maße als Erfolgskriterien verwendeten. In Ermangelung solcher potentieller Vergleichsstudien soll im Folgenden zumindest eine Auswahl der bisherigen Interventionen vorgestellt werden.

Rothländer & Richter beschreiben ihr Programm "*Aktiva*" als psychosoziales Training für Erwerbslose (Rothländer & Richter, 2009). Bewährte kognitiv-behaviorale Techniken kamen zur Anwendung, um die Zielgruppe in ihrer Handlungsfähigkeit zu stärken. Das Programm besteht aus den Modulen „Aktivitätenplanung“ (Zeitmanagement), „Konstruktives Denken“ (Identifizierung und Realitätsüberprüfung destruktiver Denkschemata und Erlernung konstruktiver Gedanken), „soziale Kompetenz“ (in Anlehnung an das Gruppentraining von Hinsch & Pflingsten (Hinsch & Pflingsten, 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 (unmittelbar nach dem Training) und T3 (3-monatiges follow-up) für 199 Langzeiterwerbslose (Experimentalgruppe (EG): 98, Kontrollgruppe (KG):101) berechnet. Die Autoren benutzten Fragebögen für körperliche, psychische, soziale Beschwerden und Selbstwirksamkeitserwartung und einen Fragebogen zur sozialen Unterstützung. Von T1-T2 berichten die Autoren eine signifikante Verringerung der psychischen und körperlichen Beschwerden für die EG, nicht jedoch für die KG. Selbstwirksamkeitserwartung und soziale Unterstützung verbesserten sich für beide Gruppen. Weiterhin geben sie an, dass die Ergebnisse im follow-up stabil geblieben sind, d.h. sich nicht signifikant verschlechtert haben. In ihrem bisherigen Bericht nennen die Autoren noch keine statistischen Kennwerte oder deskriptive Statistiken, weil sich die Ergebnisse noch im Publikationsprozess befinden.

Vastamaki et al. untersuchten in Finnland 74 Langzeitarbeitslose, welche verschiedene Maßnahmen absolvierten (Vastamaki, Moser, & Paul, 2009). Diese bestanden aus beruflicher Fortbildung, Gruppenberatung, subventionierter Arbeit und Netzwerkarbeit. Als Maß wurde hauptsächlich der Kohärenzsinn (Antonovsky, 1987) untersucht. Als Kovariaten wurden work-ability (Arbeitsfähigkeit), Coping-Ressourcen, finanzielle Situation und Bildung erhoben. Die Daten wurden vor der Maßnahme (T1) und nach 6 Monaten (T2) erhoben. Die

---

<sup>1</sup> Arbeitsgemeinschaft nach dem SGB II

Autoren fanden Verbesserungen nur in einzelnen Komponenten des Kohärenzsinn. Nur für die Probanden, die während oder nach der Intervention wiedereingestellt wurden, zeigten sich signifikante positive Effekte im Kohärenz-Gesamtscore.

Nitsche & Richter evaluierten das "TAURIS"-Projekt, in dem Langzeitarbeitslose die Möglichkeit hatten, in ehrenamtlicher gemeinnütziger Arbeit mit 14 Stunden Arbeit/Woche und einem finanziellen Ausgleich von € 78,-/Monat (ohne Anrechnung auf ALG-II) eingesetzt zu werden (Nitsche & Richter, 2003). Die Autoren untersuchten u.a. den Einfluss auf unspezifische psychische Beschwerden mit dem General Health Questionnaire (Goldberg & Hillier, 1979)) und das psychosoziale Wohlbefinden. Ihre Analyse basiert auf Daten von 350 langzeitarbeitslosen Probanden (53 Jahre), die mindestens zweimalig im Abstand von 6 Monaten untersucht wurden. Nitsche & Richter fanden heraus, dass sich das seelische Wohlbefinden der Teilnehmer so verbesserte, dass es sich nicht mehr von altersparallelisierten Erwerbstätigen unterschied und damit signifikant höher lag als das von Nichtteilnehmern, die eine Verschlechterung ihres Befindens erlebten. Die positiven Interventionseffekte erwiesen sich allerdings nicht als zeitstabil, sie waren maximal für die Teilnahmedauer von 6 Monaten nachweisbar. Nach etwa einem Jahr hatten sich die positiven Effekte so abgemildert, dass sie auf das Anfangsniveau zurückfielen. Als mögliche Ursache für die instabilen Zeiteffekte sahen die Autoren an, dass in unserer leistungsorientierten Gesellschaft nur der bezahlten Arbeit entscheidende Bedeutung bei der Stabilisierung von Selbstwert und Gesundheit zukäme.

In der Studie von Proudfoot et al. wurden 289 langzeitarbeitslose Probanden in Experimentalgruppe und Kontrollgruppe randomisiert. Nach einem kognitiv-behavioralen Training waren unspezifische psychologische 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 Trainings noch erwerbslos waren, blieb nur die Veränderung des Attributionsstils stabil, alle anderen Variablen unterschieden sich nicht mehr von der Erstuntersuchung (Proudfoot, Guest, Carson, Dunn, & Gray, 1997).

Vuori & Vesalainen untersuchten in Finnland 745 Probanden, die sich in der Übergangsphase von Kurzzeit- zur Langzeitarbeitslosigkeit (d.h. knapp unter einem Jahr) befanden und von den Sachbearbeitern der dortigen Arbeitsämter so beurteilt wurden, dass sie ein erhöhtes Risiko hätten, in dauerhafter Arbeitslosigkeit zu verbleiben (Vuori & Vesalainen, 1999). Fragebögen wurden zu T1 und ein Jahr später zu T2 verschickt. 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) gemessen. In der Auswertung zeigte sich, dass keine der Interventionsmaßnahmen einen herausragenden Einfluss darauf hatte, wie sich die Probanden zu T2 fühlten. Von den anderen Prädiktoren zeigte eine „gute finanzielle Situation“ den einzigen signifikanten positiven Einfluss auf den GHQ.

Nur falls sich Probanden noch zu T2 in einer Maßnahme – egal, in welcher- befanden, zeigte dies einen positiven Effekt auf den GHQ. Dies wurde von den Autoren so interpretiert, dass

es während der Teilnahme zu positiven Effekten kommt, die sich danach aber rasch abbauen. Subgruppenanalysen demaskierten, dass negative GHQ Ergebnisse zu T2 maßgeblich davon abhingen, ob sich die finanzielle Situation zwischen T1 und T2 erneut verschlechterte. Probanden, die an keiner der Maßnahmen teilgenommen hatten und sowohl zu T1 als auch zu T2 arbeitslos waren, zeigten keinen Unterschied im GHQ, was gegen einen Testeffekt des Instrumentes spricht. Personen, die zu T2 eine neue Anstellung hatten, verbesserten hochsignifikant ihr GHQ Ergebnis. Der Wechsel in den Erwerbsstatus könnte demnach in kausalem Zusammenhang mit der Reduktion von unspezifischen psychischen Beschwerden stehen.

Vinokur und Mitarbeiter untersuchten bei 1801 Probanden die Auswirkungen ihrer „JOBS“-Workshops, die den Probanden verbesserte Bewerbungsstrategien sowie Möglichkeiten des konstruktiven Copings bei enttäuschenden Ergebnissen ihrer Bewerbungsaktivitäten näher brachten (Vinokur, Price, & Schul, 1995). Problemlösetraining, Strategien, um optimale Entscheidungen zu treffen 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. Als Trainer wurden Sozialarbeiter und arbeitslose Lehrer engagiert, die vor ihrem Einsatz eine 240 Stunden umfassende Ausbildung absolvieren mussten. Probanden wurden in EG und KG randomisiert. Die Autoren untersuchten unter anderem die Auswirkungen ihrer "JOBS"-Workshops auf Depression mit einer Risikogruppen-Stratifizierung, die sie aus den Daten der Depressionsscores von T1 bildeten. Hierbei zeigte sich, dass nur Probanden aus der Hochrisikogruppe ihre Depressivität signifikant senkten. Die Höhe der erreichten Reduktionen verschlechterte sich in der EG von T2-T3 nicht signifikant. Die Daten der Versuchspersonen (Vpn) aus der Niedrigrisikogruppe blieben unverändert zwischen allen Messzeitpunkten. Sowohl separat 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 bei der Hochrisikogruppe war und dadurch die Möglichkeit gegeben wäre, sie bevorzugt dieser anzubieten. Allerdings geben sie zu bedenken, dass dann auf Vorbildeffekte, die von nicht-depressiven Teilnehmern ausgehen, verzichtet werden müsste, wodurch die Effektivität des Trainings geschmälert werden könnte. In dieser Studie bewirkte übrigens der Faktor „finanzielles Auskommen“ eine signifikant moderierende Wirkung auf die Depressivität.

Vinokur und Mitarbeiter untersuchten mit dem gleichen Probandenkollektiv die Auswirkungen ihrer "JOBS"-Workshops, im 2-jährigen follow-up (Vinokur, Schul, Vuori, & Price, 2000). Hierbei fanden sie, dass die Risikostratifizierung aus der älteren Publikation keinen Interaktionseffekt mehr mit der Versuchsgruppenvariable (EG vs. KG) einging. Dementsprechend interpretieren sie, dass die Stratifizierung aufgrund des Depressionsscores von T1 nur kurzfristige Interaktionseffekte produzierte. Hingegen ermittelten sie, dass ihre „Mastery“ Variable, die einen mit Strukturgleichungsmodellen berechneten, latenten Faktor aus Kompetenzüberzeugung bei Bewerbungstätigkeiten, Selbstbewusstsein und Kontrollüberzeugung darstellte, die Depressionshöhe im 2-jährigen follow-up signifikant beeinflusste.

Vuori & Silvonon 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 (Vuori & Silvonen, 2005). Die Kontrollgruppe bekam ein Booklet zugeschickt, dass sie im Selbststudium durcharbeiten sollten. Die Vpn wurden zu Beginn (T1), nach dem Workshop (T2), nach einem 6-monatigen (T3) und einem 2-jährigen (T4) follow-up befragt. Als psychologische Maße benennen die Autoren Depression und Selbstbewusstsein. Sie geben an, dass die Probanden zu T2 ihre Depressivität 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 signifikant höhere Wahrscheinlichkeit hatten, eine Arbeit im ersten oder zweiten Arbeitsmarkt zu erhalten. Durch diese vermittelnde Variable wären die positiven psychischen Effekte verspätet (wieder) zum Tragen gekommen.

Creed und Mitarbeiter untersuchten 30 junge (19 Jahre), aber trotzdem langzeitarbeitslose Probanden in einer Studie mit einem „occupational training“, welches berufspraktisches Wissen bei Konstruktionsarbeiten, soziale Fähigkeiten, Erste-Hilfe-Kurse, Umweltschutz und Kurse zur Lebenskunde für eine Zeit von max. 16 Wochen vermittelte ( Creed, Machin, & Hicks, 1996). 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 wurde die Persönlichkeitsstruktur mit dem Eysenck Personality Questionnaire erhoben. Von besonderem Interesse war als erwartete Moderatorvariable des Interventionseffektes der Faktor Neurotizismus. Unterschiede in der Verbesserung der psychologischen Variablen konnten zwischen EG und KG nur für Selbstbewusstsein, nicht aber für den GHQ ermittelt werden. Für Selbstbewusstsein geben die Autoren zwischen T1-T2 eine mittlere Interventionseffektstärke an. Der Faktor Neurotizismus zeigte nicht den erwarteten moderierenden Einfluss.

Ende der 1990-iger Jahre untersuchten Creed und Mitarbeiter erneut ein kognitiv-behaviorales Training für 43 junge (19 Jahre) langzeitarbeitslose Probanden (Creed, Machin, & Hicks, 1999). Die Kontrollgruppe setzte sich aus 22 langzeitarbeitslosen Probanden zusammen, die altersentsprechend zur EG ausgewählt wurden. 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 Selbstbewusstsein (Rosenberg, 1965) zum Einsatz, die vor (T1) und nach dem Training (T2) sowie in einem 4-monatigen follow-up (T3) erhoben wurden. Von T1-T2 wurde für die EG ein Zugewinn im GHQ und dem Selbstbewusstsein signifikant, nicht jedoch für die KG. Von T2-T3 verschlechterte sich jedoch der GHQ signifikant für die EG und kehrte etwa zum ursprünglichen Niveau zurück, während Selbstbewusstsein annähernd (deskriptiver Rückgang für EG aber noch oberhalb der Signifikanzschwelle) stabil blieb.

Maysent & Spera beschrieben die Effektivität eines Stress-Management Trainings für 76 (47 Jahre) zeitnah gekündigte Angestellte, die von einer Outplacement-Firma betreut wurden (Maysent & Spera, 1995). Die Teilnehmer entstammten sämtlich dem bildungsnahen „white-collar“-Segment. Durch den kurzfristigen Beginn der Maßnahme nach der Kündigung bestand der Interventionseffekt hier also darin, einer drohenden psychischen Verschlechterung Einhalt zu gebieten und war damit primärpräventiv orientiert. Das Training fand an zwei Tagen für jeweils 3 Stunden innerhalb einer Woche statt und hatte die

Trainingsinhalte „Erkennen von Stress“, „Stressverarbeitung durch Aufbrechen von Denkmustern“, „Stärken des Selbstbewusstseins“, „Kontrollübernahme“ und „Planung von konkreten Schritten“. Die KG wurde als Wartegruppe realisiert, bestehend aus 37 Probanden, die in Alter und Bildungsstand der EG äquivalent war. Die Interventionseffekte wurden mit einem „Personal Strain Questionnaire“ getestet, der 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 Skala der KG sich verschlechterte. In der genannten Publikation wurden leider keine deskriptiven Werte angegeben. Ein follow-up wurde nicht durchgeführt.

Saam & Wodtke beschreiben ein kognitives Training für kürzlich gekündigte Manager (Saam & Wodtke, 1995). Sie untersuchten 42 Akademiker (46 Jahre), die vormals im mittleren und höherem Management arbeiteten und sich in einem Outplacement-Programm befanden. Diese wurden in EG und KG randomisiert. Die Teilnehmer der EG bekamen Einzeltherapie-Sitzungen, in denen sie strukturiert lernten, irrationale und automatisch ablaufende Gedanken zu durchbrechen und Stress zu reduzieren. Die Teilnehmer der KG erhielten ebenso Einzelsitzungen, in welchen ihnen jedoch unstrukturierte Empfehlungen zur Stressverarbeitung vermittelt wurden. Beide Gruppen erhielten 6 Stunden Einzelberatung über einen Zeitraum von 4 Wochen. Die Untersucher setzten einen Angstfragebogen ein. Die Daten wurden vor (T1) und nach dem Training (T2) erhoben. Sie belegten signifikante Interventionseffekte für die situationsbezogenen Angstvariablen und tendenzielle Unterschiede für Variablen der Veranlagungs-(Trait) Angst. Ein follow-up wurde von den Autoren nicht durchgeführt, weil nach 5 Monaten sämtliche Teilnehmer eine neue Anstellung gefunden hatten.

Frese et al. stellten eine Pilotstudie zur Erhöhung der Eigeninitiative bei Arbeitslosen vor, die sie mit 8 Arbeitslosen (40 Jahre) im einfachen prä-post-Vergleich testeten (Frese et al., 2002). Die Intervention wurde dreitägig angelegt und die Trainingsinhalte reichten von Erhöhung der Eigeninitiative über Selbstsicherheitstraining bis hin zu Übungen von selbstregulativem Verhalten. Neben einem eigens operationalisierten Eigeninitiative-Fragebogen kamen ein Fragebogen zur Messung von Selbstwirksamkeitserwartung, einer zur Erfassung von Depressivität und einer zur Diagnose von psychosomatischen Beschwerden zum Einsatz. Die Autoren berichten eine tendenzielle Reduktion von Depressivität, Verringerung psychosomatischer Beschwerden und Erhöhung der Selbstwirksamkeitserwartung. Da die Autoren kein follow-up durchführten, ist nichts über die Stabilität der Effekte bekannt.

#### **1.4.2. Durchschnittlicher Interventionserfolg**

Prima vista kann man aus den hier vorgestellten Studien bereits das Fazit ableiten, dass die meisten Interventionen positive Effekte in verschiedenen psychologischen Konstrukten zu berichten wissen. Falls mindestens drei Messzeitpunkte erhoben wurden, stellen sich in den follow-up Ergebnissen die Effekte entweder als instabil oder als relativ zeitstabil heraus. Numerische Rückgänge der Interventionseffekte waren in den meisten Studien zu finden, die deskriptive Werte über drei Messzeitpunkte berichteten.

Paul & Moser analysierten in ihrem systematischen Review, inwieweit durchschnittlich positive Effekte durch Interventionsmaßnahmen erzielt wurden, die Erwerbslose in der Reduktion von Beschwerden helfen sollten ( Paul & Moser, 2009). In Ermangelung von Kontrollgruppen in vielen Primärstudien testeten die Autoren die Längsschnittdaten von kontinuierlich Erwerbslosen (ohne Intervention) gegen die von Teilnehmern einer Intervention. Für die 16 Interventionsstudien konnten die Autoren von T1-T2 eine durchschnittliche Reduktion der psychischen Beschwerdesymptomatik von  $d=0.35$  feststellen, die hochsignifikant war. In einer weiteren Analyse wiesen sie nach, dass die gefundenen Verbesserungen der psychischen Gesundheit eher kurzfristig ihre Wirksamkeit entfalteten, also wenig zeitstabil waren. Zur Generalisierung ihrer Ergebnisse merken sie an, dass es beim Metaanalysieren von Primärstudien zu Verzerrungen durch einen nicht kontrollierbaren Publikationsbias kommen kann. Weil Studienergebnisse von gescheiterten Interventionen seltener publiziert werden, als von erfolgreichen, führt die Metaanalyse der publizierten Interventionseffekte wahrscheinlich zu einer Überschätzung des tatsächlichen durchschnittlichen Interventionseffekts von Maßnahmen für Erwerbslose.

### **1.5. Körperliche Aktivität und kardiovaskuläres Risiko**

Zu der speziell von uns untersuchten Population der älteren Langzeitarbeitslosen sind uns keine ähnlich gearteten Forschungsvorhaben bekannt. Die folgenden Erläuterungen beziehen sich meist auf altersunabhängige und selten auf altersabhängige Populationen mit bzw. ohne Beziehung zum sozialen Hintergrund, der sich als wichtiger Einflussfaktor auf die körperliche und seelische Gesundheit und auf die Erreichbarkeit der Zielgruppe für Präventionsmaßnahmen herausgestellt hat.

In den Industrienationen wird seit Jahrzehnten eine Zunahme der Prävalenz kardiovaskulärer Risikofaktoren wie Übergewicht, Bluthochdruck, Fettstoffwechselstörungen sowie Diabetes mellitus beobachtet. Die Risikofaktoren treten häufig in Kombination auf und werden unter dem Begriff des Metabolischen Syndroms zusammengefasst. Für diese Kombination existiert keine medikamentöse Therapieoption mit übergreifendem Effekt, behandelt wird jeder einzelne Risikofaktor separat. Körperliche Aktivität hat dagegen den Vorteil, mehrere der o.g. Risikofaktoren gleichzeitig positiv beeinflussen zu können. Aktuelle Studien zeigen zudem, dass körperliche Aktivität und körperliche Fitness den Risiko steigernden Einfluss des Übergewichtes auf die Mortalität abschwächen (Blair, 2006, Lee, 2009).

Zahlreiche Studien und Projekte beschäftigen sich mit präventivmedizinischen Angeboten bzw. der Therapie von Übergewicht und Adipositas in allen Altersgruppen. Trotz aufwändiger Interventionen sind die Ergebnisse häufig ernüchternd. In den höheren Altersgruppen wurde bisher nur selten untersucht, ob Fitness und Adipositas miteinander in Verbindung stehen oder ob sie das Mortalitätsrisiko separat beeinflussen. Diesbezüglich untersuchte die Arbeitsgruppe um Sui und LaMonte (2007) im Rahmen einer Kohortenstudie 2603 Männer und Frauen über 60 Jahre. Dabei stellte sich die körperliche Fitness als entscheidender Risikofaktor für die Mortalität heraus, unabhängig von BMI, prozentualem Körperfettgehalt und dem initialen Gesundheitszustand.

Für das Ausmaß der notwendigen körperlichen Aktivität gibt es unterschiedliche Empfehlungen, deren Ziel die Reduktion des 10-Jahres-Risikos für Herzinfarkt und koronaren Herztod ist. Seit den 1990er Jahren existiert mit der Health-Enhancing-Physical-

Activity (HEPA)- Empfehlung ein Anhalt für ein gesundheitlich wirksames Mindestmaß, das eine moderat-intensive 30-minütige Aktivität an den meisten Tagen einer Woche empfiehlt (Pate, Pratt & Blair, 1995). Demgegenüber existiert eine Empfehlung zur Steigerung der körperlichen Fitness (Fitness-Empfehlung) des American College of Sports Medicine (ACSM, 1998), in der hochintensive Ausdauer-Aktivitäten von mindestens 20-minütiger Dauer an mindestens 3 Tagen in der Woche propagiert werden. Aktuell wird durch das American College of Sports Medicine und die American Heart Association eine Verknüpfung beider Positionen empfohlen (Haskell, Lee & Pate et al., 2007).

Eine 2005 im Rahmen des Nationalen Gesundheitssurveys durchgeführte Kohortenstudie zeigte, dass die beiden o.g. Empfehlungen geschlechtstypische Effekte bewirken: Die Einhaltung der HEPA-Empfehlung mindert v.a. bei Frauen deutlich das Mortalitätsrisiko, Männer profitieren dagegen mehr von der Fitness-Empfehlung (Buksch, 2005). Eine aktuell publizierte Studie aus 2010 hat jetzt die Frage aufgeworfen, ob die beiden derzeit gängigen Aktivitätsvorgaben für unterschiedliche Gewichtsgruppen in gleicher Weise empfehlenswert sind (Buksch, Schlicht, 2010). Die Autoren kommen zu folgenden Schlüssen:

- Bei Befolgung gängiger Empfehlungen profitieren aktive Personen unabhängig vom Gewichtsstatus. Es resultiert eine Verringerung der Sterblichkeit.
- Es ist kein Trend nachweisbar, ob normalgewichtige Personen stärker von körperlicher Aktivität profitieren als adipöse.

Eine weitere interessante Studie aus dem Jahr 2009 (Von Stengel, Löffler & Kemmler, 2009) überprüfte den Effekt eines 18-monatigen Kraft-Ausdauer-Trainings (2x wöchentlich 60 min angeleitetes Training + 2 x 20 min Heimtraining ) bei Frauen über 65 Jahren mit metabolischem Syndrom und konnte eine Reduktion des 10-Jahres-Risikos für koronare Herzkrankheiten von ca. 25% nachweisen. Im Wesentlichen wird dieser Effekt auf die Verbesserung des Lipidprofils und die Senkung des Blutdruckes zurückgeführt. Diese Effekte wurden bisher nur im Zusammenhang mit Ausdauer orientiertem Training beobachtet, die Mehrzahl der Untersuchungen zum Stellenwert des Krafttrainings konnte keine vergleichbaren Effekte bei postmenopausalen Frauen nachweisen.

In der eben genannten Studie von von Stengel wurde eine Kontrollgruppe mitgeführt, die in 4 Blocks a 10 Wochen mit 10-wöchigen Pausen einmal wöchentlich ein 60-minütiges „Wellness-Programm“ mit verschiedenen Bewegungsangeboten niedriger Intensität durchführten. In dieser Studiengruppe zeigten sich dementsprechend auch nur minimale Effekte auf die Reduktion des 10-Jahres-Risikos für koronare Erkrankungen.

Zuletzt soll noch eine Studie Erwähnung finden, die erneut die Effekte körperlichen Trainings auf die Gewichtsreduktion bei Übergewicht und Adipositas untersuchte. Dabei erhielten alle Studienteilnehmer dieselben Diätanweisungen, eine Gruppe erhielt zusätzlich Bewegungsempfehlungen zur selbständigen Durchführung, die zweite Gruppe dagegen ein angeleitetes körperliches Training über 2x 2 Stunden/Woche). Nach 4 Monaten hatten die Patienten der zweiten Gruppe 8 kg Gewicht verloren, die Patienten der ersten Gruppe lediglich 2,8 kg (Nicolai, Kruidenier & Leffers, 2009). Die Autoren kommen zu dem Schluss, dass Übergewichtige für eine schnelle und effektive Gewichtsreduktion ein Training unter Anleitung benötigen.

Fasst man die oben zitierten Studien zusammen, kann man folgende Empfehlungen für zukünftige präventiv wirkungsvolle Gesundheitsangebote geben:

1. Zur Reduktion von Übergewicht sollten konsequent nur **kombinierte Interventionen aus körperlicher Aktivität und Ernährungsschulung** angeboten werden, da körperliche Aktivität
  - den Risiko steigernden Effekt des Übergewichtes auf die Mortalität abmildert.
  - zum schnelleren Erfolg und damit zur Motivationssteigerung beiträgt.
2. Ein **durch sportwissenschaftlich ausgebildetes Personal angeleitetes Training** ist notwendig, da
  - erforderliche Trainingsintensitäten besser eingehalten werden.
  - Trainingsprogramme adaptiv-intensivierend gestaltet werden können.
  - die Programme dann bessere Erfolge zeigen.
3. **Körperliches Training** sollte als Maßnahme der alleinigen Gewichtsreduktion vorgezogen werden, da
  - eine dauerhafte Gewichtsreduktion selten gelingt
  - aktive Personen bzgl. des kardiovaskulären Risikos unabhängig von ihrem Gewichtsstatus profitieren.
4. Die wöchentliche **Trainingsfrequenz** sollte den allgemeinen Empfehlungen (s.o.) entsprechen, wobei geschlechtsspezifische Unterschiede zu beachten sind. Für Frauen ist bereits eine niedrig intensive Belastung nützlich, während Männer sich intensiver belasten müssen, um vergleichbare Effekte zu erzielen.
5. Ein medizinisch sinnvolles Training gestaltet sich als **kombiniertes Ausdauer-Krafttraining**.
6. Angebote zur Steigerung der körperlichen Aktivität müssen **den Bedürfnissen und Erfordernissen der Zielgruppe entsprechen**, da
  - nur so die erforderliche Akzeptanz erreicht wird.
  - Interventionshemmnisse besser berücksichtigt werden können.
  - Nachhaltigkeit gefördert werden kann.

Anmerkung: Für die Gruppe älterer Langzeitarbeitsloser, ist eine medizinische Begleitung bzw. sportärztliche Voruntersuchung dringend anzuraten, da diese Personen lange Zeit körperlich inaktiv waren und häufig eine Kombination verschiedene kardiovaskulärer Risikofaktoren aufweisen ohne in entsprechender medizinischer Kontrolle zu sein. Das Risiko medizinischer Zwischenfälle während des körperlichen Trainings könnte dadurch erhöht sein.

Die hier erläuterten Empfehlungen für präventivmedizinische Sportangebote wurden bei der Optimierung der im Projekt Fit50+ vorgestellten Maßnahme berücksichtigt und umgesetzt.

## 2. Fragestellungen und Hypothesen

### 2.1. Variablen zur Beschreibung der Gesundheit

Das Konstrukt „Gesundheit“ ist laut Definition der WHO als Zustand des vollständigen körperlichen, geistigen und sozialen Wohlergehens<sup>2</sup> aufzufassen. Einflussfaktoren auf das Konstrukt "Gesundheit" wurden mittels mehrerer Ebenen betrachtet. Die einzelnen Messverfahren hierzu werden detailliert im Abschnitt 3.6. „Methodisches Inventar“ vorgestellt.

#### 2.1.1. Faktoren der objektiven körperlichen Gesundheit

##### 2.1.1.1 Erkrankungen, Medikamenteneinnahme

- Anamnestische Ermittlung von körperlichen Beschwerden, akuten und chronischen Erkrankungen, Schmerzen, aktueller Medikation
- Erhebung des klinischen Status

##### 2.1.1.2. Risikofaktoren für Herz-Kreislaufkrankungen

- Erfassung relevanter Lebensstilfaktoren (Ernährung, Nikotingebrauch, körperliche Aktivität/ Sport)
- Erhebung familienanamnestisch relevanter Vorerkrankungen
- Messung von Ruhe-Blutdruck, Ruhe-Herzschlagfrequenz, Ruhe-EKG,
- Verhalten der Herzschlagfrequenzvariabilität bei 40-minütiger liegender Ruhe,
- Erfassung des Belastungsblutdrucks bei 100 Watt (Fahrradergometrie)
- Bestimmung biochemischer Laborparameter (z.B. Blutfette, HbA1c)
- Anthropometrische Erfassung des Körperbautyps, des Ernährungsstatus, der Körperzusammensetzung

##### 2.1.1.3. Körperliche Leistungsfähigkeit

- Ermittlung etablierter, leistungsbezogener Parameter wie maximale Leistung auf dem Fahrradergometer, maximale Sauerstoffaufnahme (VO<sub>2</sub>max), gewichtsbezogene maximale Sauerstoffaufnahme (relVO<sub>2</sub>max), maximale Ventilation, Laktat u.a..  
Anmerkung: Die körperliche Leistungsfähigkeit (physical fitness) gilt als separater Risikofaktor für Herz-Kreislaufkrankungen).

#### 2.1.2. Risikofaktoren des psychischen Wohlergehens

Ermittlung von pathogenetischen Risikofaktoren und Beschwerden der psychischen Gesundheit 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.

---

<sup>2</sup> "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).

### 2.1.3. Ressourcen des psychischen Wohlergehens

Messung von salutogenetischen bzw. relienzförderlichen psychischen Ressourcen. Hiermit soll der Forderung der WHO nachgekommen werden, Gesundheit nicht ausschließlich als Abwesenheit von Krankheit oder Beschwerden zu definieren. Folgende Messinstrumente dienen der Erhebung dieser Ebene:

- *Sense of Coherence Skala (SOC)*, zur Messung des Kohärenzsinns.
- *Allgemeine Selbstwirksamkeitserwartung (SWE)*, zur Messung der erwarteten Kompetenzüberzeugung und mittelbar auch der Kontrollüberzeugung.
- *Psychosoziales Wohlbefinden (PSW)* als weiteres Maß eines salutogenetischen Beitrages.

### 2.1.4. Subjektive körperliche Gesundheit

Erhebung des subjektiven körperlichen Gesundheitsempfindens mit folgenden Messverfahren:

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

## 2.2. Variable zur Abschätzung des passiven Freizeitverhaltens

Der Fernseh- und der Gesamtmedienkonsum sollte ermittelt werden. Vor allem der Fernsehkonsum wurde zur Abschätzung des sitzenden und passiven Freizeitverhaltens erhoben.

## 2.3. Feststellung des Interventionsbedarfes

Zusätzlich zu den Fragebögen der oben vorgestellten Gesundheitsebenen „Risikofaktoren“, „Ressourcen“ und „Körper“ soll das aus der einmaligen Anwendung des FPI gewonnene Persönlichkeitsmerkmal „Lebenszufriedenheit“ erhoben werden. Dieses Merkmal repräsentiert den Faktor „subjektives Wohlbefinden“ aus der Metaanalyse von Paul et al. (2006). Nach der deskriptiven Beschreibung der Stichprobe wird der Interventionsbedarf zur Gesundheit anhand normierter Messverfahren festgestellt. Anhand der dargestellten Literatur (Hollederer & Brand, 2006) ist davon auszugehen, dass langzeiterwerbslose, ältere Menschen höhere Quoten eingeschränkter Gesundheit aufweisen werden als in der entsprechenden Allgemeinbevölkerung. Dies soll anhand von Überschreitungen der aus den Normstichproben ermittelten erwartbaren Quoten "problematischer Fälle" festgestellt werden. Hierbei werden folgende Prognosen aufgestellt:

- Die Stichprobe weist anteilig mehr Personen mit psychischen Beschwerden auf als in der altersentsprechenden Bevölkerung. Dies wird mit transformierten Standardnormen des BDI, GHQ, PSK und dem Persönlichkeitsmerkmal "Lebenszufriedenheit" festgestellt.

- Die Stichprobe verfügt über anteilig mehr Personen mit unterdurchschnittlichen Ressourcen der psychischen Gesundheit als in der Allgemeinbevölkerung. Festgestellt wird dieser Anteil mit normtransformierten Werten des SOC und des SWE.
- Die Stichprobe weist in Gegenüberstellung zu Vergleichsstudien anteilig mehr Personen mit einer eingeschränkten subjektiven körperlichen Gesundheit auf. Dieser Anteil wird mit transformierten Werten der VAS und der KSK bestimmt.

### **2.3.1. Auswirkung von Antworttendenzen auf den ermittelten Interventionsbedarf**

Die Schätzung des Interventionsbedarfs soll gegen mögliche Aggravationstendenzen abgesichert werden. Es gibt Stichproben, in denen mit solchen Tendenzen gerechnet werden kann, wenn unterstellt werden kann, dass sich die Subjekte Vorteile durch die Aggravation eines eingeschränkten gesundheitlichen Zustandes erhoffen (z.B. bescheinigte Arbeitsunfähigkeit zur Beantragung der Rente oder im Rahmen des sog. „sekundären Krankheitsgewinns“). Im Gegensatz dazu stellt der Effekt einer sozial erwünschten Antworttendenz das Gegenphänomen dar, da dieser zur verkleinerten Darstellung von (gesundheitlichen) Problemen neigt. Falls der Effekt einer sozial erwünschten Antworttendenz nachgewiesen werden könnte, wäre somit eine Aggravationstendenz in der Stichprobe falsifiziert. Hierzu werden Zusammenhangsprüfungen der abhängigen Variablen zur Gesundheit mit dem Merkmal "Offenheit" des FPI durchgeführt.

**Antworttendenz-Hypothese:** Je "unwahrheitsgemäßer" die Antworttendenz, desto eher manifestiert sich ein Effekt der sozialen Erwünschtheit in den abhängigen Variablen zur Gesundheit.

### **2.4. Korrelative Exploration der Messverfahren auf der Ebene der psychischen Gesundheit**

Da innerhalb dieses Berichts auch Aussagen darüber gefällt werden sollen, welche Fragebogenverfahren sich zur Erfassung der Gesundheit von älteren Langzeitarbeitslosen im Längsschnitt eignen, sollen explorativ Korrelationen zwischen den Messverfahren berechnet werden, damit ein Eindruck darüber gewonnen werden kann, wie ähnlich und wie unterschiedlich die Konstrukte miteinander zusammenhängen. Dies dient der Hypothesengenerierung für zukünftige Studien. Des Weiteren soll der Versuch unternommen werden, die Variablen zur Gesundheit auf orthogonale Faktoren einer Hauptkomponentenanalyse zu reduzieren.

### **2.5. Überprüfung der Gleichheit zwischen den Kohorten**

Weil die Zuteilung zu den Trainingsgruppen zu unterschiedlichen Zeitpunkten (Kohorte 1 und Kohorte 2) nicht randomisiert erfolgte, sondern die Probanden nach Verfügbarkeit diesen zugewiesen wurden (Gelegenheitsstichprobe), soll geprüft werden, ob sich die Kohorten in den abhängigen Variablen zur Erfassung der Gesundheit voneinander unterscheiden. Pro

abhängiger Variable wird geprüft, ob die Annahme der Nullhypothese zwischen den Gruppen beibehalten werden kann.

**Kohorten-Hypothese:** Die abhängigen Variablen zur Messung der Gesundheit unterscheiden sich nicht zwischen den Trainingsgruppen (Kohorten).

## 2.6. Drop-Out-Analysen

Des Weiteren werden Drop-Out-Analysen zwischen den Probanden, die an der Wiederholungsmessung teilnahmen und denen, die dieser fernblieben, durchgeführt. Hiermit wird geprüft, inwieweit mit Selektionseffekten während des Verlaufs der Studie zu rechnen war. Implizit lassen diese Analysen Aussagen über die Generalisierbarkeit der Ergebnisse zu. Die Drop-Out-Analysen werden zwischen den Teilnehmern, die an der Erst- und Zweituntersuchung teilnahmen und denen, die nur an der Erstuntersuchung teilnahmen, durchgeführt. Des Weiteren werden sie zwischen den Teilnehmern, die an allen drei Untersuchungsterminen teilnahmen und den restlichen Probanden durchgeführt. Für die abhängigen Variablen wird geprüft, ob die Annahmen der Nullhypothesen zwischen den Gruppen beibehalten werden kann:

**Drop-out Hypothese 1:** Die abhängigen Variablen unterscheiden sich nicht zwischen den Teilnahmegruppen [T1&T2] vs. [T1].

**Drop-Out Hypothese 2:** Die abhängigen Variablen unterscheiden sich nicht zwischen den Teilnahmegruppen [T1&T2&T3] vs. [T1 und T1&T2].

## 2.7. Risikoprädiktion des Sportabbruchs

In dieser Studie wurde als ein Erfolgskriterium definiert, wenn die Versuchsteilnehmer den Sport im follow-up Zeitraum freiwillig fortsetzen. Damit wäre der Nachweis gelungen, dass ein initiiertes aktiverer Lebensstil sich nach der Intervention fortsetzt. Es soll von daher explorativ nach Prädiktoren gesucht werden, die das Risiko eines Sportabbruchs voraussagen können. Hierzu soll ein logistisches Prädiktionsmodell des Risikos erstellt werden.

## 2.8. Längsschnitthypothesen

### 2.8.1. Hypothesen zur Veränderung von Faktoren der körperlichen Gesundheit

#### 2.8.1.1. Innersubjekteffekte:

Dass sich eine sportliche Betätigung positiv auf Faktoren der körperlichen Gesundheit auswirkt, ist durch zahlreiche Längsschnittstudien belegt. Aufgrund der demographischen Entwicklung wurden im vergangenen Jahrzehnt zunehmend auch ältere Personen hinsichtlich der Wirksamkeit körperlichen Trainings, insbesondere unter Berücksichtigung kardiovaskulärer Risikoparameter untersucht (siehe dazu Kapitel 1.5.). Längsschnittstudien, die ein ausdauerorientiertes Interventionsprogramm speziell bei älteren Langzeitarbeitslosen

untersuchten, existieren u. E. nicht. Wir gehen davon aus, dass sich unabhängig vom Merkmal Erwerbstätigkeit positive Effekte auf körperlicher Ebene erwarten lassen. Zu berücksichtigen ist, dass die Höhe der Effekte maßgeblich durch die Intensität des Trainings beeinflusst wird und damit abhängig von der individuellen Motivation ist.

Für jede der nachfolgenden Hypothesen wird überprüft, ob die Intervention unter Einbeziehung des follow-up relativ zeitstabile Effekte aufweist.

→ Diverse Risikofaktoren für kardiovaskuläre Erkrankungen werden positiv beeinflusst:

- Systolischer und diastolischer Ruhe-Blutdruck nehmen ab.
- Der Lipidstatus (Gesamtcholesterol, HDL, LDL, Triglyceride) verbessert sich.
- Der Body Mass Index (BMI) nimmt ab.
- Der prozentuale Körperfettanteil wird geringer.

→ Die körperliche Leistungsfähigkeit nimmt zu:

- Die maximale Ergometerleistung nimmt zu.
- Die maximale Sauerstoffaufnahme ( $VO_2\text{max}$ ) nimmt zu.
- Die relative maximale Sauerstoffaufnahme ( $\text{rel}VO_2\text{max}$ ) nimmt zu.
- Die maximale Ventilation nimmt zu.

## **2.8.2. Hypothesen zur Veränderung der Risikofaktoren der psychischen Gesundheit (L1)**

### **2.8.2.1. Innersubjekteffekte:**

Dass sich eine sportliche Betätigung auch positiv auf die psychische Gesundheit auswirkt, ist in Längsschnittstudien einerseits belegt (van Gool et al., 2007)), andererseits wurden im systematischen Review von Lawlor & Hapker auch Zweifel an der methodischen Qualität der Primärstudien laut (Lawlor & Hopker, 2001). Wie sich ein sportbasiertes Interventionsprogramm für ältere Langzeitarbeitslose auswirken mag, ist für die psychologischen Variablen noch unklar. Es lassen sich einerseits positive Effekte erwarten, andererseits könnte die erheblich eingeschränkte Lebenssituation der Langzeiterwerbslosen (Mohr, 2001) zu abgeschwächten Interventionseffekten führen, weil für die meisten Teilnehmer das zentrale Problem der Erwerbslosigkeit durch das Training ungelöst bleibt. Trotzdem wird vermutet, dass die Intervention sich senkend auf die Risikofaktoren der psychischen Gesundheit auswirkt. Zusätzlich wird für jede der nachfolgenden Hypothesen überprüft, ob die Intervention unter Einbeziehung des follow-up relativ zeitstabile Effekte aufweist.

→ Allgemeine psychische Beschwerden sinken (=Erhöhung im GHQ).

→ Depressivität reduziert sich (BDI).

→ Die Einschränkung der psychischen Lebensqualität sinkt (=Erhöhung in PSK).

### **2.8.2.2. Zwischensubjekteffekte:**

Die Wirksamkeit von Sport als Antidepressivum für ältere Menschen mit einer depressiven Episode wird als belegt angesehen (Blumenthal et al., 1999). Es wird erwartet, dass

Probanden, die eine Depressivität mit klinisch relevanter Intensität aufweisen, von der Intervention besonders, d.h. stärker in der Reduktion von Risikofaktoren der psychischen Gesundheit profitieren, als die restlichen Probanden.

- Depressive Probanden weisen eine stärkere Erhöhung im GHQ auf.
- Depressive Probanden weisen eine stärkere Reduktion im BDI auf.
- Depressive Probanden weisen eine stärkere Erhöhung im PSK auf.

Des Weiteren wird geprüft, ob die Intensivierung des Trainings, wie in Kohorte 2 erfolgt, zu einer stärkeren Reduktion der psychischen Risikofaktoren führt, als in Kohorte 1.

- Kohorte 2 weist stärkere Erhöhung im GHQ auf.
- Kohorte 2 weist stärkere Reduktion im BDI auf.
- Kohorte 2 weist stärkere Erhöhung im PSK auf.

### **2.8.3. Hypothesen zur Veränderung der Ressourcen der psychischen Gesundheit (L2)**

#### 2.8.3.1. Innersubjekteffekte:

Analog zu L1 wird erwartet, dass sich positive Interventionseffekte auf die psychische Gesundheit erhöhend in den Ressourcen der psychischen Gesundheit manifestieren. Zusätzlich wird für jede Hypothese überprüft, ob die Intervention auch unter Einbezug des follow-up relativ zeitstabile Effekte aufweist.

- Der Kohärenzsinn steigt (SOC).
- Die Selbstwirksamkeitserwartung steigt (SWE).
- Das psychosoziale Wohlbefinden steigt (PSW).

#### 2.8.3.2. Zwischensubjekteffekte:

Es wird geprüft, ob die Intensivierung des Trainings, wie in Kohorte 2 erfolgt, zu einer stärkeren Verbesserung der Ressourcen führt, als in Kohorte 1

- Kohorte 2 weist stärkere Erhöhung im SOC auf.
- Kohorte 2 weist stärkere Erhöhung im SWE auf.
- Kohorte 2 weist stärkere Erhöhung im PSW auf.

### **2.8.4. Hypothesen zur Veränderung der subjektiven körperlichen Gesundheit (L3)**

#### 2.8.4.1. Innersubjekteffekte:

Ein angemessenes Ausdauertraining kann sich reduzierend auf Übergewicht, Bluthochdruck und Fettstoffwechselstörungen auswirken (Lee, Sui & Blair, 2009). Diese Parameter stellen relevante Risiken für kardio- und zerebrovaskuläre Erkrankungen dar. Des Weiteren gilt sportliche Aktivität als eine der effektivsten Methoden zur Behandlung von generalisierten

Schmerzen, wie der Fibromyalgie und somatoformen Schmerzstörungen (Busch, Barber, Overend, Peloso, & Schachter, 2007). Von daher kann davon ausgegangen werden, dass eine sportliche Aktivität sowohl objektiv als auch subjektiv messbar die körperliche Gesundheit stärkt. Es wird erwartet, dass der Sport sich positiv auf die subjektive Gesundheit und die damit verbundene gesundheitsbezogene Lebensqualität auswirkt. Zusätzlich wird für die folgenden Hypothesen überprüft, ob die Intervention unter Einbeziehung des follow-up relativ zeitstabile Effekte aufweist.

- Die körperliche gesundheitsbezogene Lebensqualität verbessert sich (KSK erhöht sich).
- Der empfundene körperliche Gesundheitszustand verbessert sich (VAS erhöht sich).

#### 2.8.4.2. Zwischensubjekteffekte:

Oft werden präventive Maßnahmen zur Verhinderung von Krankheiten (z.B. Impfungen, Vorsorgeuntersuchungen, Ernährungsberatung) von Männern schlechter angenommen als von Frauen (McQueen, Vernon, Meissner, Klabunde & Rakowski, 2006). Speziell für den Sport aber postulieren Markland & Ingledew, dass dieser von Männern gut akzeptiert sei, da Männer ein körperbetonteres und kompetitiveres Ziel beim Sporttreiben verfolgen würden als Frauen. Geschlechtsdifferenzen in der Wirksamkeit von Sport auf empfundene Gesundheitsverbesserungen werden hin und wieder publiziert: z.B. untersuchten Foy und Mitarbeiter den moderierenden Einfluss des Geschlechtes auf die Wirksamkeit von Sport bei COPD-Patienten und berichteten, dass die männlichen Patienten stärker in der gesundheitsbezogenen Lebensqualität profitierten als die weiblichen (Foy, Rejeski, Berry, Zaccaro, & Woodard, 2001). Jette et al. zeigten, dass sich für 102 ältere Probanden nach einem Sportprogramm ein gesteigertes körperliches Wohlbefinden nur für die Männer und nicht für die Frauen nachweisen ließ (Jette et al., 1996). Für einen möglichen Geschlechtseffekt sollen daher folgende Hypothesen aufgestellt werden:

- Männer weisen eine stärkere Erhöhung im KSK auf.
- Männer weisen eine stärkere Erhöhung im VAS auf.

Des Weiteren wird geprüft, ob die Intensivierung des Trainings in Kohorte 2 zu einer stärkeren Verbesserung der subjektiven Gesundheit führt als in Kohorte 1.

- Kohorte 2 weist eine stärkere Erhöhung im KSK auf.
- Kohorte 2 weist eine stärkere Erhöhung im VAS auf.

#### **2.8.5. Hypothesen zur Veränderung des Fernsehkonsums (L4)**

##### 2.8.5.1. Innersubjekteffekt:

Es ist zu erwarten, dass das sportbasierte Gesundheitsförderungsprogramm ein aktiveres Freizeitverhalten Langzeiterwerbsloser initiiert. Ein aktiveres Verhalten sollte sich gleichzeitig senkend auf ein passives Freizeitverhalten auswirken. Als prototypisch für das Letztere wird der TV-Konsum angesehen. Es wird erwartet, dass die Intervention sich senkend auf die tägliche Fernsehnutzungszeit auswirkt. Zusätzlich wird überprüft, ob die Intervention auch unter Einbeziehung des follow-up einen relativ zeitstabilen Effekt aufweist.

- Der TV-Konsum senkt sich im Interventionszeitraum.

### 2.8.5.2. Zwischensubjekteffekte:

Depp et al. argumentieren in ihrem Übersichtsartikel zum TV-Konsum bei älteren Menschen, dass diese das Fernsehen auch zur Emotionsregulation gebrauchen (Depp, Schkade, Thompson & Jeste, 2010). Von daher ist vorstellbar, dass die Reduktion des TV-Konsums für Probanden mit Depressivität geschmälert sein könnte, da ihre affektive Balance möglicherweise stärker von dem Medium abhängt, als für Probanden ohne Depressivität.

→ Depressive Probanden weisen eine schwächere Reduktion des TV-Konsums auf.

Des Weiteren wird geprüft, ob die Intensivierung des Trainings, wie in Kohorte 2 erfolgt, zu einer stärkeren Reduktion des TV-Konsums führt, als in Kohorte 1.

→ Kohorte 2 weist eine stärkere Reduktion des TV-Konsums auf.

### **2.8.6. Hypothesen zur Fortsetzung des Sportes im follow-up (L5)**

Nach der Interventionsphase bekamen die Teilnehmer die Möglichkeit zur Fortsetzung des Sports angeboten. Dieses Angebot ist als niederschwellig zu bezeichnen, da es im selben Fitnesscenter offeriert wurde, in dem auch die Intervention stattfand, so dass sich die Teilnehmer im follow-up nicht umgewöhnen mussten. Zudem wurde es kostenfrei angeboten. Es soll überprüft werden, ob Probanden, die sich im follow-up Zeitraum weiterhin sportlich betätigen, unterschiedliche Veränderungsergebnisse zwischen T2-T3 aufweisen, wenn sie mit denen verglichen werden, die im follow-up den Sport abbrechen. Es wird erwartet, dass die Sportfortsetzer ein hinzugewonnenes Wohlergehen im follow-up Zeitraum entweder besser aufrecht erhalten oder besser ausbauen können, als die Sportabbrecher.

→ Sportfortsetzer zeigen in den Messinstrumenten eine für die Gesundheit günstigere Veränderung im follow-up, als die Sportabbrecher. Diese soll sich für die Sportfortsetzer niederschlagen als höherer Wert im GHQ, niedrigerer Wert im BDI, höherer Wert im PSK, höherer Wert im SOC, höhere SWE, höherer Wert im PSW, höherer Wert in der KSK, höherer Wert in der VAS und niedrigerer TV-Konsum.

## **3. Methoden**

### **3.1. Einleitung**

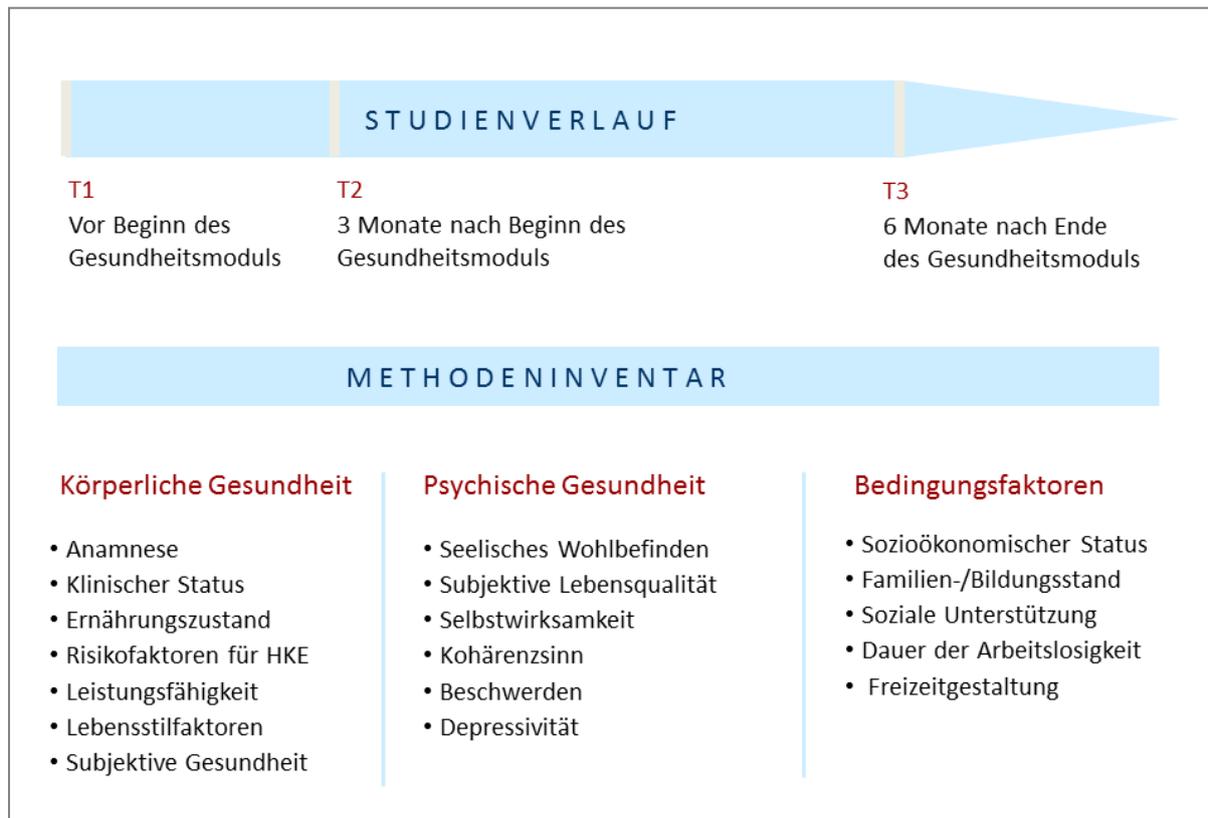
Dass Arbeitslose eine eingeschränkte psychische und körperliche Gesundheit aufweisen, ist unumstritten. Für ältere, v.a. sehr langfristig Erwerbslose (>3 Jahre) kommt hinzu, dass diese wahrscheinlich mehrere Selektionsprozesse aufgrund ihrer eingeschränkten Gesundheit durchlaufen haben, so dass krankheitsrelevante körperliche und psychische Beschwerden in dieser Gruppe 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 bislang ungenutztes Präventionspotential. Eine präventivmedizinische Intervention könnte die körperliche und die psychische Gesundheit stärken und zum Abbau krankheitsrelevanter Vermittlungshemmnisse beitragen. Neben der Hoffnung, für die Zielgruppe die Wiedereinstellungswahrscheinlichkeit zu erhöhen und damit Transferleistungen einzusparen, dürfte eine Gesundheitsförderung darüber hinaus wirksam sein, in dem sie den Krankenkassen hilft, Kosten für die erhöhte Inanspruchnahme von medizinischen Behandlungen durch Langzeiterwerbslose zu reduzieren. Insofern ist anzunehmen, dass eine primär auf die Gesundheit ausgelegte Förderung für ältere, langfristig Erwerbslose in mindestens zwei Kostenbereichen wirksam sein sollte.

In diesem Bericht sollen die psychischen und körperlichen Auswirkungen eines Gesundheitsprogramms für die Zielgruppe evaluiert werden. Veränderte Wiedereinstellungschancen werden nicht evaluiert, da hierfür erstens die Datenbasis zu klein ist und zweitens eine strukturschwache Region wie Mecklenburg-Vorpommern schwierige Rahmenbedingungen für eine Rückkehr in den ersten Arbeitsmarkt bietet. Hierzu wäre eine Stichprobe mit großer Fallzahl unter Einbezug von aktuellen Effektgrößen des lokalen Arbeitsmarktes vonnöten. Das Zielkriterium wird also auf die – zumindest partielle – Wiederherstellung der psychischen und objektiven sowie subjektiven körperlichen Gesundheit beschränkt.

Für die Gesamtstichprobe soll im Querschnitt zuerst der Interventionsbedarf festgestellt werden, bevor die Analysen der Längsschnittdaten berichtet werden. Im nächsten Kapitel wird die hier durchgeführte Studie vorgestellt. Es folgt ein Abriss der Methoden inklusive Design, Ein- und Ausschlusskriterien der Stichprobe, Interventionsbeschreibung, Hypothesen und der Vorstellung der Messinstrumente, mit denen der Interventionsbedarf und der Trainingserfolg gemessen werden soll.

### **3.2. Untersuchungsdesign**

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



### 3.2.1. Stichprobengröße

In jeder Längsschnittstudie ist zu erwarten, dass es im Laufe der Durchführung zu Studienabbruchern kommen wird. In dieser Studie konnten die Probanden z.B. wieder in Arbeit kommen, erkranken oder aus motivationalen Gründen die Intervention beenden. Gerade mit langzeiterwerbslosen älteren Probanden war in den letztgenannten Punkten mit einer erhöhten Abbruchquote zu rechnen. Um trotz der Abbrecher noch eine zufriedenstellende Datenbasis zur Verfügung zu haben, war geplant, 120 Probanden für die Studie zu rekrutieren.

### 3.2.2. Kontrollgruppe

Eine Kontrollgruppe wurde nicht in das Studiendesign aufgenommen.

#### Begründung:

1. Eine Wartegruppe innerhalb eines Crossover-Design wäre organisatorisch nicht realisierbar gewesen, da für die Interventionsteilnahme inklusive follow-up mindestens 9 Monate eingeplant werden mussten. Gleichzeitig hielten sich die Probanden nur 10 Monate in der Weiterbildungsinstitution, in der sie rekrutiert wurden, auf, so dass ein gesicherter Kontakt mit den Teilnehmern nach ihrem Fortbildungszeitraum nicht mehr gewährleistet war. Da die Rekrutierung komplett vom Bildungsträger gewährleistet wurde, hätte die verlängerte Zeitspanne (d.h. über den 10-monatigen Betreuungszeitraum hinaus) vermutlich zu inakzeptablen Ausfallraten geführt.

2. Aus methodischen Gründen wurde es abgelehnt, die Teilnehmer aus den anderen Bildungskursen als Kontrollgruppe einzusetzen. Weil diese Gruppe aus Personen bestand, die explizit nicht an der Gesundheitsförderung interessiert waren, wäre es vermutlich zu systematischen Verzerrungen der Daten gekommen.

Der Verzicht auf eine Kontrollgruppe ist allerdings für Längsschnittuntersuchungen mit Erwerbslosen vertretbar, weil sich in der Forschungserfahrung gezeigt hat, dass sich mit der besonderen Zielgruppe keine sog. "Testeffekte" zwischen den Messzeitpunkten zeigen. Als "Testeffekt" bezeichnet man das Phänomen, dass sich Versuchspersonen bei mehrmaliger Befragung implizit aufgefordert fühlen können, eine Verbesserung ihres Befindens zu beschreiben (Winefield, 1995). Solche unerwünschten Effekte wiederholter Testung sind bei Längsschnittstudien keine Seltenheit und wurden schon von Wohlwill als „ärgerliches Problem“ longitudinaler Forschungsdesigns bezeichnet (Wohlwill, 1977). Paul und Mitarbeiter stellten in ihrer Metaanalyse von Längsschnittdaten fest, dass es bei kontinuierlich Berufstätigen zu einer kleinen ( $d=0,06$ ), aber signifikanten Steigerung ihrer psychischen Gesundheit in der Mehrfachbefragung kam. Dies traf jedoch **nicht** auf die Erwerbslosen zu (Paul et al., 2006). Die Autoren schreiben:

„Bei Personen, die zu beiden Erhebungszeitpunkten arbeitslos waren, zeigt sich keine signifikante Veränderung ihrer psychischen Gesundheit, wenn man Interventionsstudien aus der Datenbasis ausschließt. Nach Ausschluss dieser Studien betrug die Effektstärke für kontinuierlich arbeitslose Personen  $d=-0,03$  und zwar nicht signifikant“. (K.I. Paul, et al., 2006).

Die Autoren erklären den diskrepanten Testeffekt zwischen Berufstätigen und Arbeitslosen damit, dass es bei kontinuierlich Erwerbslosen vermutlich auch zu einem schwachen Testeffekt kommt, aber dieser dadurch neutralisiert wird, dass mit der Zeit ihre psychischen Beschwerden kontinuierlich zunehmen (vgl. Befund zum Zusammenhang von Arbeitslosigkeitsdauer und psychischen Beschwerden, Kapitel 1.1.4.). Durch die Studie von (Paul et al., 2006) ist somit der empirische Meta-Nachweis gelungen, dass Erwerbslose nicht spontan ihre psychische Gesundheit verbessern, welches überspitzt gesagt eine Kontrollgruppe - zumindest in Form einer Wartegruppe - für diese spezielle Zielgruppe redundant macht.

### 3.2.3. Experimentalgruppen

Die Probanden waren ältere ( $54\pm 3$  Jahre) langfristig Erwerbslose ( $>1$  Jahr arbeitslos,  $\emptyset >10$  Jahre arbeitslos), die sämtlich aus dem Raum Schwerin rekrutiert wurden. 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 von Mitarbeitern des Institutes für Präventivmedizin aus medizinischer und sportwissenschaftlicher Sicht optimiert werden sollte. Insgesamt wurden 119 Probanden in zwei Erhebungswellen medizinisch-psychologisch untersucht und anschließend mit dem jeweiligen Interventionsprogramm trainiert. Weil die Gruppen zeitlich disjunkt diagnostiziert und trainiert wurden, werden sie nachfolgend „**Kohorten**“ genannt.

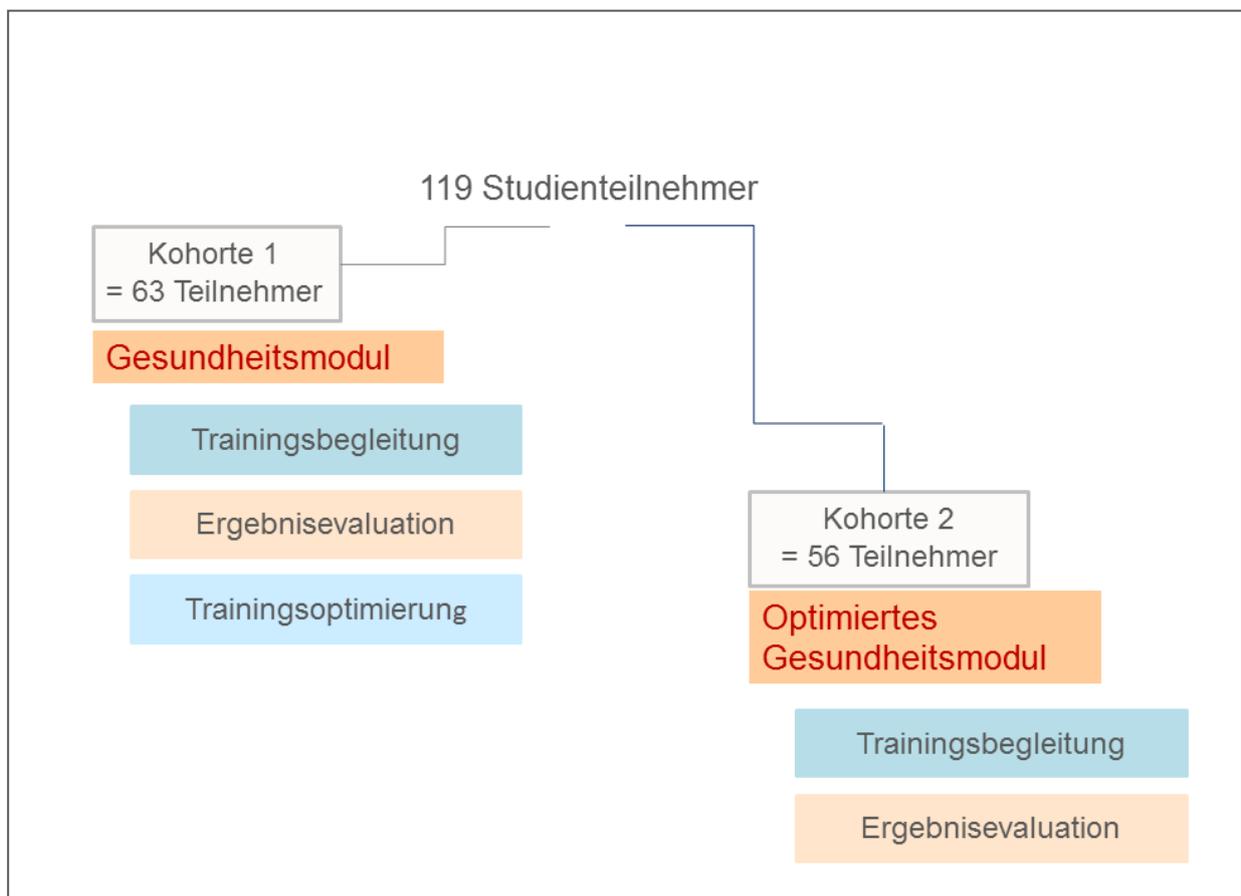
1. Die erste Kohorte (N=63: 25 Männer, 38 Frauen; Alter:  $53,7 \pm 2,5$ ) wurde von September 2007 bis April 2008 mit der Erstuntersuchung in die Studie aufgenommen. Für die letzten Versuchsteilnehmer erfolgte die follow-up Untersuchung im Dezember 2008.
2. Die zweite Kohorte (N=56: 21 Männer, 35 Frauen; Alter:  $53,7 \pm 3,4$ ) fand von März 2009 bis Juli 2009 Eingang in die Studie; die letzten Probanden wurden im April 2010 im follow-up untersucht.

**Einschlusskriterien** zur Aufnahme in die Stichprobe:

- Teilnehmer der Studie sollten das Mindestkriterium von „Langzeiterwerbslosigkeit“ erfüllen, d.h. mindestens ein Jahr arbeitslos sein.
- Sie sollten möglichst 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 bzw. Rückfall bei chronischer Alkoholkrankheit. Die Prüfung dieser Kriterien erfolgte durch eine Ärztin.
- Deutschkenntnisse, die so mangelhaft waren, dass abzusehen war, dass die Teilnehmer die psychologischen Fragebögen sprachlich nicht verstehen konnten. (Migrationshintergrund stellte per se kein Ausschlusskriterium dar).
- Personen über 60 Jahre.



**Abb.4:** Zeitlicher Ablauf der Evaluation

### 3.3. Rekrutierung der Probanden

Ältere Langzeiterwerbslose wurden durch die ARGE Schwerin einem privaten Bildungsträger mit Vermittlungskompetenz (RegioVision GmbH, Schwerin) zugeteilt. Dort wurden die Kunden (offizielle Bezeichnung) für einen Zeitraum von 10 Monaten von Jobmentoren betreut. Durch die Zuweisung durch die ARGE war die Teilnahme an der Qualifizierungsmaßnahme für alle Kunden verpflichtend. Im Verlauf des Betreuungszeitraumes konnten durch die Kunden Bildungsgutscheine für Kursangebote eingelöst werden. Neben der Möglichkeit der Teilnahme am Gesundheitsmodul gab es weitere typische Qualifizierungsangebote wie z.B. Computerkurse, Kommunikationstraining oder die Möglichkeit, betriebswirtschaftliche Kenntnisse zu verbessern. Etwa 70% der Kunden entschieden sich für den Gesundheitskurs.

Die Jobmentoren waren instruiert, die Kunden zu einer Teilnahme am Gesundheitskurs und zu einer Teilnahme an der Studie motivierend anzuregen. Vor allem Kunden, bei denen der Jobmentor ein sog. „gesundheitliches Vermittlungshemmnis“ feststellte, sollten zur Teilnahme motiviert werden. Dadurch kam es möglicherweise zu einer Selektion von mehr „schwierigeren Fällen“ in den Kurs. Eine Teilnahme am Gesundheitskurs implizierte keinen Zwang an der Studie teilzunehmen. Zur Teilnahme an der Studie wurden die Kunden damit motiviert, dass sie am Institut für Präventivmedizin (IPM) die Gelegenheit hatten, umfassend medizinisch untersucht zu werden (inklusive individueller Gesundheitsberatung). Für die Untersuchungstage wurden die Teilnehmer von den Verpflichtungen des Bildungsträgers freigestellt. Sie bekamen das Fahrgeld von Schwerin nach Rostock und zurück im Voraus bezahlt. Die Versuchspersonen erhielten zusätzlich eine Aufwandsentschädigung von € 15,- für jede durchgeführte Untersuchung.

Die Rekrutierung für die Teilnahme an der Studie war so geplant, dass die Probanden vor der Teilnahme am edukativen „Gesundheitsmodul-Basiswissen“ am IPM untersucht werden sollten. Dies ließ sich aus organisatorischen Gründen leider nicht durchgängig gewährleisten, da die Gesundheitskurse des Bildungsträgers mit einer Teilnehmerzahl von jeweils 20 Kunden durchgeführt wurden und alle zugleich mit dem Unterricht beginnen mussten, während die Rekrutierung von Kunden als Probanden für die Studie mit höherem Aufwand und teils terminlichen Herausforderungen verbunden war. Zudem ließen sich einige Teilnehmer erst zu einer Teilnahme an der Studie bewegen, nachdem sie einige Tage durch den Gesundheitskurs mit den Inhalten vertraut waren. So wurden einige Teilnehmer für die Erstuntersuchung erst während der theoretischen Kurse von den Gesundheitsdozenten rekrutiert. Gewährleistet war jedoch stets, dass die Erstuntersuchung vor dem Beginn des 8- bzw. 10-wöchigen Sportprogramms stattfand. Die terminliche Organisation und Absprache für die Untersuchungen erfolgte entweder durch den Jobmentor oder den Sozialpädagogen. Beide standen in intensivem Kontakt zum IPM. Da die Teilnahme an den Untersuchungsterminen freiwillig und zudem für die Probanden recht zeitaufwändig war, wurde bei den Absprachen mit den Versuchsteilnehmern Rücksicht auf deren terminliche Wünsche genommen.

### 3.4. Interventionsprogramm

Um die Integrationschancen gesundheitlich eingeschränkter, älterer Langzeitarbeitsloser zu verbessern, wurde das „Gesundheitsmodul“ vom Kooperationspartner RegioVision GmbH Schwerin unter Einbeziehung von Mitarbeitern und Kunden entwickelt und seit Dezember 2006 durchgeführt. Mitarbeiter des Institutes für Präventivmedizin waren nicht an der Durchführung des Programms beteiligt. Für den wissensvermittelnden Kurs „Gesundheitsmodul-Basiswissen“ erreichte der Bildungsträger eine Zertifizierung durch die CERTQUA, Bonn. Die ARGE Schwerin unterstützte den theoretischen Teil der Maßnahme mit Bildungsgutscheinen. Das körperliche Training wurde zunächst über Mittel des Bildungsträgers, später im Rahmen der „Perspektive 50plus“ durch das Bundesministerium für Arbeit und Soziales (BMAS) finanziert. Das „Gesundheitsmodul-Praktische Übungen“ bestand zunächst aus Circle-Training, Rückenschule, Ausdauertraining und Entspannungsübungen mit wechselndem Schwerpunkt, jeweils als Kombination aus theoretischen Anleitungen im Fitness-Studio und praktischen Übungen. Vor dem „Gesundheitsmodul-Praktische Übungen“ wurde jeweils der edukative Kurs „Gesundheitsmodul-Basiswissen“ durchgeführt.

Ein Projektmitarbeiter mit sportwissenschaftlicher Qualifikation betreute die Probanden zusammen mit einem Fitnesstrainer während des körperlichen Trainings. Vordergründiges Ziel war die Verbesserung der Arbeitsfähigkeit durch Erhöhung der körperlichen und psychischen Gesundheit und Leistungsfähigkeit und der Abbau von Vermittlungshemmnissen mittels sportlicher Aktivierung. Die Fachdozenten für die wissensvermittelnden Kurse stammten aus Gesundheitsberufen. Das Sportprogramm wurde in einem Fitness-Studio in Schwerin durchgeführt, die Gesundheitsedukation in den Seminarräumen des Bildungsträgers.

#### 3.4.1. Modul „Gesundheit-Basiswissen“

Der theoretische Teil des Gesundheitsmoduls umfasste folgende Kernthemen:

Titel	Inhalte (beispielhafte Auswahl)
Gesundheits-Check	<ul style="list-style-type: none"> <li>- Analyse des Ist-Gesundheitszustandes</li> <li>- Motivation zur Bewegung/ Aktivität</li> <li>- Verbesserung des Wohlbefindens durch aktive Lebensweise</li> <li>- Überwindung der negativen Symptomatik von Arbeitslosigkeit</li> </ul>
Gesundheit und Wohlbefinden	<ul style="list-style-type: none"> <li>- Was ist Gesundheit? Zusammenhang Gesundheit und Wohlbefinden</li> <li>- Risiken von Bewegungsmangel, Herz Kreislauferkrankungen, Abhängigkeitserkr.</li> <li>- Einschätzung der eigenen Leistungs- und Belastungsfähigkeit</li> <li>- Ganzheitliche Gesundheitspflege</li> </ul>
Anatomisches/ physiologisches Grundwissen, Rückenschule	<ul style="list-style-type: none"> <li>- Gelenk- Rückenprobleme, Erkrankungen am Bewegungsapparat</li> <li>- Analyse und Korrektur von Fehlhaltungen in untersch. Körperpositionen</li> <li>- Prinzipien belastungsinduzierter Anpassung</li> </ul>
Ernährungsberatung	<ul style="list-style-type: none"> <li>- Stoffwechsel, Nahrungspyramide, Ernährungsplanung, Risiken falscher Ernährung</li> <li>- Vielseitigkeit und Genuss, Kochtipps zur schonenden Zubereitung des Essens</li> <li>- Übungen für eine finanziell sparsame und trotzdem gesunde Ernährungsweise</li> </ul>
Umgang mit psychischen Belastungen	<ul style="list-style-type: none"> <li>- Was bedeutet Stress? Identifikation von potentiellen Stressoren</li> <li>- Unterschied zw. Belastung und Beanspruchung, Möglichkeiten des Copings</li> <li>- Vorstellung von Atemübungen, Yoga, Taiji</li> </ul>

Der Kurs wurde in einem Umfang von 40 Unterrichtsstunden mit einer Gruppengröße von 20 Teilnehmern an 8 Terminen mit jeweils 5 Stunden durchgeführt. Der Unterricht fand an 2 Tagen in der Woche statt, so dass der Kurs unter günstigen Bedingungen (keine Feiertage, keine außerplanmäßigen Termine) innerhalb von 4 Wochen absolviert werden konnte. Die Anwesenheit der Teilnehmer wurde kontrolliert. Ziel des edukativen Teils des Gesundheitsmoduls war die Verbesserung der individuellen Gesundheitskompetenz der Teilnehmer i. S. des Empowerments.

### 3.4.2. Modul „Gesundheit-Praktische Übungen“

Nach dem theoretischen Unterricht folgte das körperliche Training in einem kooperierenden Fitness-Studio unter Anleitung von 2 ausgebildeten Trainern mit folgenden Schwerpunkten:

Titel	Inhalt
Herz-Kreislauf Training	<ul style="list-style-type: none"> <li>- Training an modernen Cardiogeräten (Ergometer, Crosstrainer, Rudermaschinen etc.)</li> <li>- Schulung von Koordination, Kraft, Ausdauer</li> <li>- Anpassungserscheinungen des Herz-Kreislaufsystems</li> <li>- Herzschlagfrequenzgesteuertes Training-</li> </ul>
Circle-Training	<ul style="list-style-type: none"> <li>- Gruppentraining zur Stärkung aller großen Muskelgruppen</li> <li>- „Dr. Wolff“-Trainingsgeräte und „Kraftkreis“</li> <li>- Schulung von Kraft, Ausdauer, Beweglichkeit und Schnelligkeit</li> <li>- Verbesserung der subjektiven Leistungsfähigkeit (empfundene Kondition)</li> </ul>
Rückentraining	<ul style="list-style-type: none"> <li>- gezieltes Kraft-Ausdauertraining zur Festigung des Halte- und Stützapparates</li> <li>- Übungen zum Verinnerlichen rückenschonenden Verhaltens</li> <li>- Übungen für das regelmäßige, eigenständige Training zu Hause</li> </ul>
Haltungs- und Entspannungstraining	<ul style="list-style-type: none"> <li>- Angebote zum Herausfinden individuell passender Übungen (Taiji/ Qigong, Yoga, Atemübungen, Autogenes Training, Progressive Muskelrelaxation, Traumreisen)</li> <li>- Verbesserung der psychophysischen Belastbarkeit, Abbau von Verspannungen</li> <li>- Stretching</li> </ul>

Die Anwesenheit der Teilnehmer wurde bei jedem Termin überprüft. Die Kohorten unterschieden sich im körperlichen Training wie folgt:

#### Kohorte 1:

Das körperliche Training umfasste 40 Trainingsstunden und fand in Gruppen von 12-15 Teilnehmern an 1-2 Terminen pro Woche statt. Der Kurs wurde von einem Trainer des kooperierenden Fitness-Studios durchgeführt. Neben dem körperlichen Training wurde die Zeit auch genutzt, um Trainingswissen zu vermitteln. Pro Trainingstermin wurden Übungen zu einem der o.g. Schwerpunkte durchgeführt. Unter zeitlich bestmöglichen Bedingungen konnte der Kurs innerhalb von 8 Wochen absolviert werden.

## Kohorte 2:

Nach der ersten Erhebungswelle wurde das Modul „Gesundheit-Praktische Übungen“ durch Projektmitarbeiter konzeptionell überarbeitet und bzgl. der Trainingsintensität und –frequenz sowie der Gesamtdauer des Angebotes optimiert. Orientierung gaben dabei u.a. Empfehlungen aus der aktuellen wissenschaftlichen Literatur zur Prävention von Herz-Kreislaufkrankungen.

Natürlich musste berücksichtigt werden, dass die Trainingsintensität adaptiv gestaltet wird und sich nach dem Leistungsvermögen der Teilnehmer richtet. Die meisten Teilnehmer der Gesundheitsmodule hatten über 3 Jahrzehnte keinen Sport getrieben. Entsprechend niedrig war das Leistungsniveau zu Beginn der Kurse. Viele Pausen und eine feinfühlig-motivierende Ausstrahlung des Trainers ermöglichten Vielen aber trotzdem den Zugang zu einem bisher unbekanntem Terrain. Regelmäßigkeit und stabile Trainingsgruppen mit der Möglichkeit, (dabei und/ oder danach) soziale Kontakte zu pflegen, sowie die individuelle Beratung zum Trainingserfolg waren u. E. wesentlich Punkte, die zur Zufriedenheit der Teilnehmer beigetragen haben.

Hinsichtlich der Erhöhung der Wirksamkeit des Trainings sowie der Verbesserung der Nachhaltigkeit empfahlen wir dem Kooperationspartner, das Modul „Gesundheit-Praktische Übungen“ auf eine Gesamtstundenzahl von 60 Stunden, mit einer regelmäßigen Frequenz von 2 Unterrichtseinheiten wöchentlich zu erweitern. Bei der Dauer des Kurses mussten als ein wesentlicher Kostenfaktor die regelmäßig anfallenden Ausgaben für Fahrten mit Öffentlichen Verkehrsmitteln berücksichtigt werden, sodass die Stundenanzahl pro Unterrichtseinheit relativ hoch blieb.

Bezüglich der Trainingsinhalte wurde vereinbart, dass bei jedem Training alle o.g. Schwerpunkte Berücksichtigung finden.



Im Rahmen der geplanten Prozessevaluation wurde mehrfach die Trainingsintensität bei der Durchführung des Circle-Trainings mit mobilen Messgeräten vor Ort überprüft und individuell anhand der bei T1 erhobenen Daten zur kardiopulmonalen Ausdauerleistungsfähigkeit abgeglichen. Dabei konnte anhand des Herzfrequenzverhaltens und der Sauerstoffaufnahme während des Trainings gezeigt werden, dass bei Anleitung durch ausgebildete Trainer hohe Trainingsintensitäten von 60-80% der maximalen Sauerstoffaufnahme/ bzw. Herzschlagfrequenz auch über einen längeren Zeitraum möglich sind.

#### **Follow-up beider Kohorten:**

Im follow-up Zeitraum, 6 Monate nach der Interventionsphase und des zweiten Messzeitpunktes (T2), hatten die Studienteilnehmer die Möglichkeit, freiwillig und kostenlos im selben Fitness-Studio weiter zu trainieren. Das Training konnte eigenständig gestaltet werden, für Trainingsfrequenz und -dauer gab es keine konkreten Vorgaben. An 2 Tagen in der Woche bestand die Möglichkeit, am angeleiteten Circle-Training teilzunehmen. Die durchschnittliche, wöchentliche Trainingsfrequenz und -dauer wurde bei T3 erfasst.

### **3.5. Untersuchungsablauf**

Die Untersuchungen wurden im Institut für Präventivmedizin in Rostock durchgeführt. Dort waren für die Studie zwei Labore eingerichtet: ein klinisches Funktionslabor mit Anthropometriemessplatz, EKG und Fahrrad-Ergospirometrie und ein psychophysiologisches Labor mit Schreibtischen und Untersuchungsfläche für Ruhemessungen.

Die Gesamtuntersuchungszeit im IPM umfasste inklusive Pause etwa 4 Stunden. Je zwei Studienteilnehmer reisten am Untersuchungstag per Zug an und fanden sich gegen 9:00 Uhr im IPM ein. Die Probanden wurden zeitgleich entweder zuerst klinisch-physiologisch oder psychologisch untersucht. Nach ca. 2 Stunden wurde eine 15 minütige Pause mit Tee und Gebäck angeboten, danach wechselten die Probanden die Labore. Im Rahmen der klinischen Funktionsdiagnostik erhielten die Studienteilnehmer eine ausführliche, präventivmedizinische Gesundheitsberatung zu individuellen Ressourcen und Risiken. Alle erhobenen Befunde wurden den Teilnehmern anhand alters- und geschlechtsspezifischer Referenzwerte erklärt und in Kopie mitgegeben. Bei abklärungsbedürftigen Befunden wurde mit Einverständnis der Probanden der betreuende Hausarzt/ die Hausärztin schriftlich informiert.

### **3.6. Methodisches Inventar**

#### **3.6.1. Erfassung soziodemographischer Merkmale**

Während der psychologischen Erstuntersuchung wurden zunächst in Anlehnung an den Fragebogen des TAURIS Projektes (Nitsche & Richter, 2003) alle relevanten soziodemographische Daten erhoben.

## 3.6.2. Erfassung von Variablen der körperlichen Gesundheit

### 3.6.2.1. Anamnestische Daten/ Klinischer Status

Die ausführliche Erhebung der Eigenanamnese mit allen Vorerkrankungen unter besonderer Berücksichtigung kardiovaskulärer und familiärer Risiken sowie die Befragung zu aktuellen Erkrankungen, Beschwerden und regelmäßig eingenommenen Medikamenten ist aus Sicherheitsgründen vor der geplanten Ausbelastung auf dem Fahrrad erforderlich. Daran anschließend erfolgte eine umfassende klinische Untersuchung einschließlich Ruhe-EKG.

### 3.6.2.2. Kardiovaskuläre Risikofaktoren

Ein institutsinterner Fragebogen erfasste das Ausmaß regelmäßiger körperlicher Aktivität. Gefragt wurde explizit nach Ausdauersportarten, die regelmäßig über mindestens 30 Minuten ausgeübt werden, nach Trainingsdauer- und frequenz anderer Sportarten und nach nichtsportlichen, körperlichen Aktivitäten, die regelmäßig oder unregelmäßig ausgeführt werden (z.B. Gartenarbeit, Spaziergänge mit dem Hund u.ä.).

Systolischer und diastolischer Blutdruck wurden mit dem automatischen Blutdruckmessgerät bosotron 2 (Fa. Boso, Deutschland) in liegender Ruhe in 5-minütigen Abständen gemessen. Aus den 3 gemessenen Werten wurde anschließend der Mittelwert berechnet.

Zur Beurteilung einer Grenzwerthypertonie/ Hypertonie diente die Messung des Belastungsblutdrucks im Rahmen der Fahrradergometrie bei 100 Watt.

Die Herzfrequenzvariabilität wurde in 40-minütiger liegender Ruhe mit den Herzfrequenzmessern S810i (Fa. POLAR, Finnland) im Anschluss an die psychologische Untersuchung aufgezeichnet. Für die Auswertung wurden nur artefaktfreie 30-Minuten-Abschnitte verwendet. Die Analyse erfolgte mit der frei verfügbaren Auswertungssoftware der University of Kuopio (Kubios HRV 2.0, Finnland).

Zur Bestimmung diverser Laborparameter (Cholesterol, HDL-Cholesterol, LDL-Cholesterol, Triglyceride, HbA1c, TSH) erfolgte im Anschluss an das Ruhe-EKG eine venöse Blutentnahme in der Ellenbeuge. Die Blutproben wurden im Institut für Klinische Chemie und Laboratoriumsmedizin der Universität Rostock analysiert.

Für die Ermittlung des Ernährungszustandes wurde die Körperhöhe mit einem herkömmlichen Anthropometer und das Körpergewichtes mit der digitalen Personenwaage S20-2760 (Fa. Soenle, Deutschland) gemessen und anschließend der Body Mass Index ( $BMI = \text{Körpermasse} / \text{Körpergröße}^2 \text{ [kg/m}^2 \text{]}$ ) berechnet.

Der prozentuale Körperfettanteil wurde durch Messung der Hautfaltendicken an 10 definierten Körperstellen mit einem nach Stoll (1988, 2002)(Stoll, Sommer, Johnson, & Stoll, 1988; Stoll, Stoll, & Schröder, 2002) modifizierten Skinfold-Caliper Digimac 500 (Fa. Mitutoyo, Japan) und anschließender Berechnung nach Parizkova (1977) sowie nach Durnin & Womersley (1974) ermittelt. Bei der Interpretation wurde der individuelle prozentuale Körperfettanteil mit alters- und geschlechtsspezifischen Referenzwerten verglichen.

### 3.6.2.3. Körperliche Leistungsfähigkeit

Nach Ausschluss aller Kontraindikationen wurde zur Bestimmung der körperlichen Leistungsfähigkeit ein Rampentest (50W/30s/12,5W) auf dem Fahrradergometer ER 900 (Fa. Ergoline, Deutschland) durchgeführt, die ergospirometrische Datenerfassung (z.B. Maximalleistung, maximale Ventilation, relative und absolute maximale Sauerstoffaufnahme

u.v.m.) erfolgte über das Gerät EOS-Sprint (Fa. Jaeger, Deutschland). Häufigstes Abbruchkriterium war die muskuläre Erschöpfung nach Erreichen der individuellen Belastungsgrenze mit einem Respiratorischen Quotienten (RQ)  $>1,16$ . Einige Untersuchungen mussten wegen auftretender Herzrhythmusstörungen bzw. wegen des Überschreitens vorgegebener Blutdruckgrenzen vorzeitig abgebrochen werden. Zwischenfälle aufgrund kardialer Durchblutungsstörungen traten nicht auf. 3 Probanden tolerierten das Tragen der Atemmaske nicht.

### **3.6.3. Erfassung von Variablen der psychischen und psychosozialen Gesundheit**

#### 3.6.3.1. Freiburger Persönlichkeitsinventar – revidierte Fassung – FPI-R (FPI)

Es existiert die Auffassung, dass Depressivität häufig bei Langzeitarbeitslosen zu beobachten ist, weil sie von ihrer Grundpersönlichkeit eher dazu neigen (vgl. Selektionshypothese, Kapitel 1.1.2.). So wird als depressionsrelevant der Persönlichkeitsfaktor „Neurotizismus“ genannt (Hammarstrom & Janlert, 1997). Da Persönlichkeitsmerkmale als relativ zeitstabil gelten und von daher angenommen werden kann, dass sie bereits vor der Arbeitslosigkeit vorlagen, könnten Auffälligkeiten in Richtung der Selektionshypothese gedeutet werden. Allerdings ist zu bedenken, dass sich die Persönlichkeit im Laufe des Lebens ändern kann. 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 persönlichkeitsverändernd wirken. Der Arbeitsplatzverlust zählt laut Filipp & Aymanns dazu (Filipp & Aymanns, 2010).

Zur Aufklärung der Persönlichkeitsmerkmale wurde das Freiburger Persönlichkeitsinventar FPI (Fahrenberg et al., 2010) verwendet. In diesem Persönlichkeitstest werden 138 Fragen, die mit „stimmt“ oder „stimmt nicht“ zu beantworten sind, auf 12 Dimensionen der Persönlichkeit verdichtet. Die Persönlichkeitsfaktoren sind nicht streng unabhängig (orthogonal) voneinander, sondern wurden von den Testautoren nach inhaltlichen Kriterien faktorisiert und zusammengestellt. Der FPI kommt häufig in medizinischen Fragestellungen zur Anwendung, da er körperliche Beschwerden und Gesundheitssorgen als Persönlichkeitsfaktoren beinhaltet. Er wurde in dieser Studie einmalig angewendet.

#### 3.6.3.2. General Health Questionnaire in der 12 Item Version – (GHQ-12)

Der GHQ war in den von Paul & Moser metaanalysierten Primärstudien das am meisten verwendete Instrument, sowohl in den Querschnitts- als auch in den Längsschnittstudien mit Erwerbslosen. Er wurde für sämtliche Erhebungszeitpunkte eingesetzt.

Der GHQ kam in der 12-Item-Version, übersetzt von Linden et al (1996) zum Einsatz. Dieser ursprünglich von Goldberg (Goldberg, 1978) und Hillier (Goldberg & Hillier, 1979) entwickelte Fragebogen wird oft im klinischen Bereich für Veränderungsmessungen eingesetzt. Er gehört zur Gruppe der globalen klinischen Tests und dient zur Identifikation unspezifischer psychischer Beschwerden. Die Items werden innerhalb einer vierstufigen Likert-Skala zwischen 0 und 3 beantwortet.

### 3.6.3.3. Beck Depressions-Inventar in der ökonomisierten Version des „BDI-V“

Studien, die die Interventionseffektivität für Arbeitslose prüfen, setzen häufig Depressivitätsskalen für die Veränderungsmessung ein (z. B. Vinokur et al., 1995; Vinokur, et al., 2000; Vuori & Silvonen, 2005). Die Wirksamkeit von sportlicher Aktivität in der Depressionsbehandlung gilt im allgemeinen als belegt (Kessler et al., 2001) und wird auch als vorbeugende, antidepressive Maßnahme für ältere Menschen empfohlen (Strawbridge, Deleger, Roberts & Kaplan, 2002). Der BDI wurde deshalb in dieser Studie 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. Kuhner, Burger, Keller & Hautzinger, 2007)). Schmitt und Mitarbeiter validierten eine hinsichtlich der Bearbeitungszeit ökonomisierte Form des BDI (Schmitt et al., 2003). In dieser teilt der Proband den Grad seiner Zustimmung zu jeweils 20 negativen Aussagen auf einer sechsstufigen Skala (0-5) mit, anstatt, wie im Original, eine von vier abgestuften, semantisch verschiedenen Antwortmöglichkeiten auszuwählen.

### 3.6.3.4. Kohärenzsinn – Sense of Coherence (SOC)

Vastamaki et al. diskutieren Zugewinne im Kohärenzsinn von Arbeitslosen, die nach einer Wiedereinstellung im ersten Arbeitsmarkt feststellbar waren (Vastamaki et al., 2009). Sie weisen ferner darauf hin, dass die Veränderbarkeit des SOC in ihrer Stichprobe keinen Alterseffekt aufwies, welches die Skala als Instrument für eine Veränderungsmessung interessant macht. In unserer Untersuchung wurde die „Leipziger Kurzsкала“ (Schumacher & Brähler, 2000) angewendet. Der SOC L9 wurde in allen Untersuchungen eingesetzt.

Das Modell der Salutogenese von Antonovsky (Antonovsky, 1987; Antonovsky & Franke, 1997) rückte, in Abkehr vom bis dato dominierenden pathogenetischen Erklärungsansatz zur Entstehung von Krankheiten, die Frage in den Mittelpunkt, was Menschen trotz der vielfältigen und allgegenwärtigen Belastungen gesund erhält bzw. wieder gesund werden lässt (Margraf, 1998). Im salutogenetischen Modell erfolgt die Betrachtung von Gesundheit und Krankheit nicht als dichotome, einander ausschließende Kategorien, sondern als Endpunkte eines Kontinuums (Antonovsky & Franke, 1997). Ein Mensch ist entsprechend dieser Vorstellung nicht entweder gesund oder krank, sondern mehr oder weniger gesund oder krank. Im Mittelpunkt des Salutogenesemodells steht das Konzept des Kohärenzgefühls (sense of coherence, SOC), 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“ (Übersetzung von Becker, 1997, S.10).

Das Kohärenzgefühl ist eine Bewältigungsressource, die Menschen widerstandsfähiger gegenüber Stressoren macht. Es setzt sich nach Antonovsky aus folgenden Komponenten zusammen:

1. Verstehbarkeit: bezeichnet das Ausmaß, in dem Reize, Ereignisse oder Entwicklungen als strukturiert, geordnet und vorhersehbar wahrgenommen werden.

2. Handhabbarkeit: bezieht sich auf das Ausmaß, in dem eine Person geeignete (personale und soziale) Ressourcen wahrnimmt, um interne und externe Anforderungen bewältigen zu können.
3. Sinnhaftigkeit: meint das Ausmaß, in dem eine Person ihr Leben als sinnvoll empfindet und zumindest einige der vom Leben gestellten Anforderungen als Herausforderungen betrachtet, die Engagement und Investitionen verdienen.

Die Leipziger Kurzsкала (Schumacher & Brähler, 2000) des SOC besteht aus 9 Items. Zwei Items reflektieren die Komponente „Verstehbarkeit“, drei Items „Handhabbarkeit“ und vier Items messen „Sinnhaftigkeit“. Im Fragebogen drücken die Probanden ihre Zustimmung oder Ablehnung zu jedem Item mittels einer 7-stufigen (1-7) Skala aus.

### 3.6.3.5. Allgemeine Selbstwirksamkeitserwartung nach Schwarzer und Jerusalem (1999) (SWE)

Die SWE wurde für die Stichprobe als Indikator für psychische Ressourcen erhoben. Die Skala stellt auch ein Indiz für die „*Internale Kontrollüberzeugung*“<sup>3</sup> (Rotter, 1975) dar. Manche Studien (z.B. Rothländer & Richter, 2009; Frese et al., 2002) lassen auf Hinzugewinne der Selbstwirksamkeitserwartung während einer Intervention für Arbeitslose schließen. Die SWE wurde in dieser Studie zu allen Erhebungsterminen eingesetzt.

Das Konzept der Selbstwirksamkeitserwartung („perceived self-efficacy“), auch (optimistische) Kompetenzerwartung genannt, wurde von Bandura als wesentlicher Aspekt seiner sozial-kognitiven Theorie eingeführt (Bandura, 1977). Die Allgemeine Selbstwirksamkeitserwartung (SWE) 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. Die SWE stellt eine personale Bewältigungsressource mit prädictivem Wert für das Wohlbefinden und eine konstruktive Lebensbewältigung dar (Schwarzer, 1994). Die Skala wurde von Jerusalem und Schwarzer in den 1980er Jahren entwickelt (Jerusalem & Schwarzer, 1984). Als offizielle internationale Literatur-Referenz benennen die Autoren die Publikation von (Schwarzer & Jerusalem, 1995) und als entsprechende Quelle der deutschsprachigen Version (Schwarzer & Jerusalem, 1999). Die SWE-Skala umfasst 10 Items, die vierstufig (1-4) zu beantworten sind.

### 3.6.3.6. Psychosoziales Wohlbefinden (PSW)

Die Skala sollte als ökonomischer Kurzfragebogen psychologische Ressourcen im Sinne eines Wohlbefindens messen. Er kam zu allen Untersuchungsterminen zum Einsatz.

Der Fragebogen zum psychosozialen Wohlbefinden ist ein von Riman & Udris vorgestelltes Instrument zur kurzen Erfassung des allgemeinen Wohlbefindens (well being) mithilfe von 6 Items, die vierfach gestuft sind (1-4). Die faktorielle Struktur dieses Screening Instrumentes ist laut der Autoren unidimensional und soll „...als Ressource, die einen salutogenetischen Beitrag für die Gesundheit leisten kann“ betrachtet werden (Rimann & Udris, 1993, S.24). Der Fragebogen wurde innerhalb einer Testbatterie im Rahmen des Forschungsprojektes

---

<sup>3</sup> Internale Kontrollüberzeugung: Inwieweit glaubt ein Mensch, dass er selbst für das Geschick in seinem Leben verantwortlich ist? Im Gegensatz zur externalen Kontrollüberzeugung: Inwieweit glaubt ein Mensch, dass Umweltfaktoren (z.B. „die Gesellschaft“) sein Leben schicksalhaft bestimmen?

SALUTE (Rimann & Udris, 1993) und erneut im Rahmen der Evaluation des TAURIS Projektes (Nitsche & Richter, 2003) eingesetzt.

#### 3.6.3.7. Fragebogen zum Gesundheitszustand (SF12)

Für die vorliegende Studie war die Änderung der gesundheitsbezogenen Lebensqualität interessant, da sich die Intervention vornehmlich auf ein Ausdauertraining stützte. Neben der weiten Verbreitung in der Lebensqualitätsforschung liegt der Vorteil des SF12 darin, dass er explizit *psychische Lebensqualität* und *körperliche Lebensqualität* trennt. Diese Skalen liefern sehr eigenständige Daten, d.h. sie korrelieren kaum miteinander. Der SF12 wurde zu allen Messzeitpunkten angewandt.

Der SF12 entstammt dem SF36, welcher ein etabliertes Messinstrument für die gesundheitsbezogene Lebensqualitätsforschung darstellt (Ravens-Sieberer, 2002). Der Fragebogen wurde international psychometrisch überprüft und normiert. Es sind Versionen in zehn Sprachen adaptiert, so auch in deutsch (Bullinger & Kirchberger, 1998). Die Testautoren beschreiben, dass der SF12 mit dem Ziel entwickelt wurde, 90% der Gesamtvarianz aus der körperlichen und psychischen Summenskala des SF36 mit nur 12 Fragen zu erklären. Es wurden statistische Regressionsmodelle zur Itemselektion und Gewichtung der Antwortmöglichkeiten verwendet, um die körperliche und psychische Summenskala des SF36 einer US-amerikanischen Normstichprobe zu reproduzieren (Bullinger & Kirchberger, 1998). Die Items des SF12 werden mit unterschiedlicher Skalierung beantwortet.

#### 3.6.3.8. Visuelle Analogskala zur Einschätzung des Gesundheitsempfindens (VAS)

Die VAS ist ein sehr einfaches und zeitökonomisch einsetzbares Instrument zur Erfassung des subjektiven, körperlichen Gesundheitszustandes. Die hohe Auflösung (0-100) lässt vermuten, dass sich auch kleinere Veränderungen darin abbilden. Die Skala wurde zu allen Testzeitpunkten eingesetzt.

Die VAS wurde dem EuroQOL Fragebogen entnommen (Greiner & Claes, 2007). In diesem Instrument ist die Einschätzung der eigenen Gesundheit als Skala operationalisiert, die von 0 bis 100 reicht. 0 wird vorgegeben als „schlechtester denkbarer Gesundheitszustand“ und 100 als „bester denkbarer Gesundheitszustand“. Für Vergleichswerte wurde auf die Studie von Hinz et al. zurückgegriffen (Hinz, Klaiberg, Brähler & König, 2006).

#### **3.6.4. Medienkonsum**

Da von Arbeitslosen angenommen wird, dass sie oftmals unstrukturiert den Tag verbringen (Jahoda, 1982; Warr & Jackson, 1987), lag es nahe, den Medienkonsum der Probanden im Längsschnitt 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 Internet-Gebrauch erhoben. Des Weiteren erfolgte die Befragung der Probanden nach ihrer Nutzung von Computerspielen. Die tägliche Mediennutzungshäufigkeit erfolgte pro Kategorie 8-stufig in 2 Stunden Schritten („0 Stunden“, „1-2 Stunden“, „3-4 Stunden“, ..., „über 12 Stunden“).

### 3.7. Statistische Verfahren

Für Häufigkeitsunterschiede erfolgte die statistische Prüfung über den Chi-Quadrat Test. Für paarweise Unterschiede wurden je nach Voraussetzung t-Tests für unabhängige oder abhängige Stichproben durchgeführt. Alternativ dazu wurden nonparametrisch Mann-Whitney-U-Tests oder Wilcoxon-Tests berechnet. Bei mehr als 2 Gruppen erfolgte die Unterschiedsprüfung mit One-Way ANOVAs inklusive Scheffé-Prozedur oder dem nonparametrischen Alternativtest nach Kruskal-Wallis. Für bivariate Zusammenhänge kam der Test für Pearson-Korrelationen, als nonparametrische Alternative Spearman-Rho zur Anwendung. Die Identifizierung von Faktoren erfolgte über Hauptkomponentenanalysen. Die Prädiktion von zweistufigen Variablen wurden über binär-logistische Regressionsanalysen realisiert. Risiken wurden als Odds oder Odds-Ratios berechnet.

Der Test der Zeiteffekte wurde innerhalb multivariater ANOVAs mit Messwiederholung durchgeführt, wobei die Messzeitpunkte prä, post und follow-up die Faktorstufen des Innersubjektfaktors bildeten. Der Test der Interventionshypothesen erfolgte über Innersubjekt Kontraste für das hypothesenrelevante Messzeitpaar oder im Bonferroni adjustierten t-Test für paarweise Vergleiche. Für lineare oder quadratische Trends zwischen den Messzeitpunkten kam die Prüfung von polynominalen Kontrasten zum Einsatz. Mittlere Interventionseffekte zwischen der ersten und allen folgenden Faktorstufen wurden über Helmert-Kontraste des Innersubjektfaktors getestet. Für Moderatoreffekte auf den Interventionseffekt erfolgte der Test über den Effekt der Zwischensubjektfaktor-Innersubjektfaktor-Interaktion oder über den Helmert-Kontrast innerhalb dieser Interaktion. Sphärizität zwischen den Faktorstufen des Innersubjektfaktors wurde mit dem Mauchly-Test (Mauchly, 1940) geprüft. Bei einer Verletzung der Sphärizitätsannahme erfolgte eine Korrektur der Freiheitsgrade nach (Greenhouse & Geisser, 1959). Für sämtliche Hypothesen kam der statistische Test zweiseitig mit einem Signifikanzniveau von  $p=0,05$  ( $p=0,025$  pro Seite) zur Anwendung.

Statistische Berechnungen, Analysen und Tests wurden mit dem Programmpaket SPSS 15.0 durchgeführt. Effektstärkenberechnungen und Poweranalysen erfolgten zusätzlich mit dem Programm G\*Power 3.12.

Per default werden deskriptive Werte in diesem Bericht als Mittelwert $\pm$ Standardabweichung (z.B. 50,00 $\pm$ 10,00) berichtet. Falls andere Maße der zentralen Tendenz (z.B. Median, Modalwert, M-Schätzer) reportiert werden, wird explizit darauf aufmerksam gemacht.

### 3.8. Ethikvotum

Die vorliegende Studie wurde von der Ethikkommission der Universität Rostock im Oktober 2006 geprüft und erhielt ein positives Votum. (Aktenzeichen: A 26/2006).

## 4. Ergebnisse

### 4.1. Soziodemographische Beschreibung der Stichprobe

#### 4.1.1. Allgemeine soziodemographische Daten

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, und war zwischen Männern ( $54,6 \pm 3,2$ ) und Frauen ( $53,2 \pm 2,7$ ) etwa gleich verteilt. Der jüngste Proband war 48, der älteste war 59 Jahre.

In einer Ehe oder festen Lebensgemeinschaft lebten 44 Probanden (37%). 48 Versuchspersonen (Vpn) waren geschieden oder getrennt (40%), 23 waren ledig (19%), 4 verwitwet (3%). Somit ließen sich 75 Probanden (63%) zur Kategorie „ohne Lebensgemeinschaft“ zusammenfassen (67% der Männer und 60% der Frauen). 103 Probanden (87%) hatten Kinder (74% der Männer, 95% der Frauen), die meist erwachsen waren. Lediglich 24 Vpn (20%) gaben an, derzeit mit Kindern im Haushalt zu leben. Dass sich Kinder im Haushalt befanden, war nicht deckungsgleich mit der Aussage, ob die Vpn in Partnerschaft lebte. Von den 24 Vpn mit Kindern im Haushalt lebten 11 allein mit ihrem Kind, während 13 mit Kindern innerhalb einer Partnerschaft lebten. Anders betrachtet lässt sich sehen, dass in den Single Haushalten (N=75) 64 Probanden (85%) keine Kinder im Haushalt hatten. Von den partnerschaftlichen Gemeinschaften (N=44) hatten 30 Probanden (68%) keine Kinder im Haus.

Schulische Abschlüsse lagen in der Mehrzahl der Fälle mit N=68 (57%) beim 10-Klassen POS-Abschluss, der mit dem Realschulabschluss vergleichbar ist, gefolgt vom 8. Klasse POS-Abschluss mit N=36 (30%), der etwa dem Volksschul- bzw. Hauptschulabschluss entspricht. 12 Probanden (10%) hatten das Abitur und 3 (2,5%) gaben einen nicht kategorisierbaren Abschluss an. Nach der Schulzeit erreichten 75 Probanden (63%) den Facharbeiterabschluss (abgeschlossene Berufsausbildung), der die Stichprobe klar dominiert. 11 Personen (9%) verfügten über keinerlei Ausbildung und weitere 11 (9%) hatten ein Hochschulstudium abgeschlossen (restl. Berufskategorien vgl. Abb. 5).

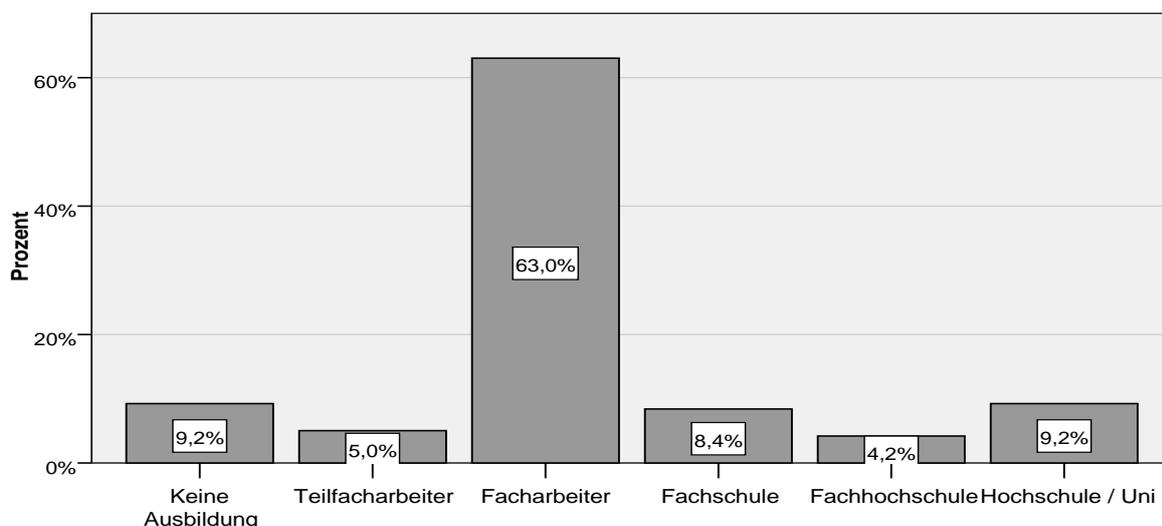


Abb. 5: Berufliche Qualifikation der Stichprobe als prozentuale Häufigkeiten

Fasst man Teilfacharbeiter, Facharbeiter und Personen mit Fachschulabschluss zur Kategorie „abgeschlossene Berufsausbildung“ sowie Fachhochschule und Universität zu „Akademiker“ zusammen und zeigt diese mit den Kategorien der Arbeitslosigkeitsdauer, ergibt sich folgende Aufstellung:

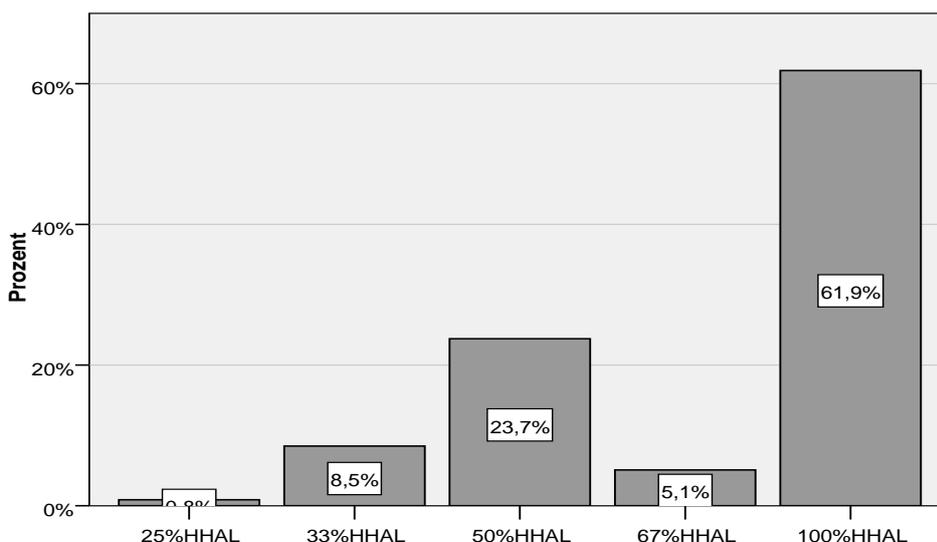
**Tab. 1:** Kreuztabelle Arbeitslosigkeitsdauer mit Ausbildungskategorien

			Ausbildungskategorie			Gesamt
			ohne Ausbildung	Berufsausbildung	Akademiker	
Dauer Arbeitslosigkeit	1 Jahr	Anzahl	1	8	3	12
		% von Beruf	9,1%	8,8%	18,8%	10,2%
	1-2 Jahre	Anzahl	2	11	2	15
		% von Beruf	18,2%	12,1%	12,5%	12,7%
	3-5 Jahre	Anzahl	2	28	7	37
		% von Beruf	18,2%	30,8%	43,8%	31,4%
	> 5 Jahre	Anzahl	6	44	4	54
		% von Beruf	54,5%	48,4%	25,0%	45,8%
Gesamt		Anzahl	11	91	16	118 <sup>4</sup>
		% von Beruf	100,0%	100,0%	100,0%	100,0%

Neben der Tatsache, dass die Stichprobe hauptsächlich aus Probanden mit über 5-jähriger Arbeitslosigkeit besteht (46%), fällt auf, dass bei den Akademikern die Kategorie „3-5 Jahre“ (44%) dominierte. Im Gegensatz dazu überwogen in der Kategorie „>5 Jahre“ die Probanden mit „Berufsausbildung“ (48%) und die „ohne Ausbildung“ (55%).

74 Probanden (62%) bezeichneten sich als Hauptverdiener bevor sie arbeitslos wurden; Männer (80%) signifikant häufiger als Frauen (51%) ( $\chi^2=10,621$   $p=0,001$ ). Da der Beziehungsstand vor der Arbeitslosigkeit nicht abgefragt wurde, lassen sich keine Rückschlüsse ziehen, ob die Angabe „Hauptverdiener“ aufgrund der Dominanz von Single-Haushalten bereits vor der Arbeitslosigkeit zustande kam.

Aus den Angaben „Anzahl der Personen im Haushalt“ und „Anzahl der Arbeitslosen im Haushalt“ wurde eine prozentuale Haushaltsarbeitslosigkeit (HHAL) berechnet.



**Abb. 6:** Häufigkeit prozentualer Haushaltsarbeitslosigkeit

<sup>4</sup> 1 Proband machte keine Angaben zur Berufsausbildung.

Bei dieser Betrachtung dominiert die Kategorie 100%-HHAL mit 73 Probanden (62%), gefolgt von der 50%-HHAL, die für 28 Vpn (24%) zutraf. 33%-HHAL lag für 10 Probanden vor (8,5%); diese entstand, wenn der Proband als Einziger von 3 Haushaltsmitgliedern von Arbeitslosigkeit betroffen war. Die 67%-HHAL traf auf 6 Teilnehmer (5,1%) zu und kam z.B. dann vor, wenn der Proband mit einem (erwerbsfähigen) Kind arbeitslos war, aber der Partner berufstätig war. 1 Proband (0,8%) war innerhalb eines Haushaltes von 4 Personen arbeitslos (25%-HHAL).

**Tab. 2:** Kreuztabelle prozentuale Haushaltsarbeitslosigkeit mit Anzahl der Haushaltsmitglieder

			Anzahl Personen			Gesamt
			1	2	≥3	
HHAL	25%	Anzahl	-	-	1	1
		% von HHAL	-	-	100%	100%
	33%	Anzahl	-	-	10	10
		% von HHAL	-	-	100%	100%
	50%	Anzahl	-	27	1	28
		% von HHAL	-	96,4%	3,6%	100%
67%	Anzahl	-	-	6	6	
	% von HHAL	-	-	100%	100%	
100%	Anzahl	56	15	2	73	
	% von HHAL	76,7%	20,5%	2,7%	100%	
Gesamt	Anzahl	56	42	20	118 <sup>5</sup>	
	% von HHAL	47,5%	35,6%	16,9%	100%	

100%-HHAL kann bedeuten, dass sämtliche Mitglieder im Haushalt von Erwerbslosigkeit betroffen sind oder, dass der (arbeitslose) Proband einen Single-Haushalt bestreitet. Stellt man die 73 Probanden mit 100%-HHAL mit der Anzahl der Personen in ihrem Haushalt dar (vgl. Tab. 2), lässt sich sehen, dass die komplette Arbeitslosigkeit für 56 Probanden (77%) in Single-Haushalten vorkam. 15 Probanden (21%) waren innerhalb einer Lebensgemeinschaft komplett arbeitslos. Dies betraf auch Probanden mit Kindern im erwerbsfähigen Alter. Zwei Vpn (3%) gaben an, dass sie innerhalb eines dreiköpfigen Haushaltes von vollständiger Arbeitslosigkeit betroffen waren.

Ein Proband gab an, innerhalb eines vierköpfigen Haushaltes von 50% Arbeitslosigkeit betroffen zu sein. Die 50%-HHAL kam ansonsten nur im zweiköpfigen Haushalt vor (96%), wenn der (arbeitslose) Proband einen berufstätigen Lebenspartner hatte oder das im Haushalt lebende, erwerbsfähige Kind arbeitete.

#### 4.1.2. Einkommen

Von den 119 Versuchsteilnehmern gingen 27 (23%) (12 Männer, 15 Frauen) einer „stundenweisen Beschäftigung“ nach. Diese stellten meist eine bezuschusste Arbeit oder die in den Medien so bezeichneten „Ein-Euro-Jobs“ dar.

Die Vpn wurden gebeten, ihr Einkommen, welches meist ihren Leistungsbezug (ALG II) darstellte, plus evt. Einkommen bzw. Leistungsbezug ihres Partners plus evt. Wohngeld als

<sup>5</sup> 1 Proband machte keine Angaben zur Haushaltsgröße

Gesamthaushaltseinkommen zu schätzen und dieses Haushaltseinkommen in eine von 8 Einkommenskategorien einzuordnen. Folgende Verteilung resultierte für die Stichprobe:

**Tab. 3:** Angegebenes Haushaltseinkommen

Einkommen (€)	N	%
Bis 400	17	14,3%
400-600	21	17,6%
600-800	34	28,6%
800-1000	14	11,8%
1000-1200	15	12,6%
1200-1500	9	7,6%
1500-1800	9	7,6%
1800-2100	0	0%
Gesamt	119	100,0%

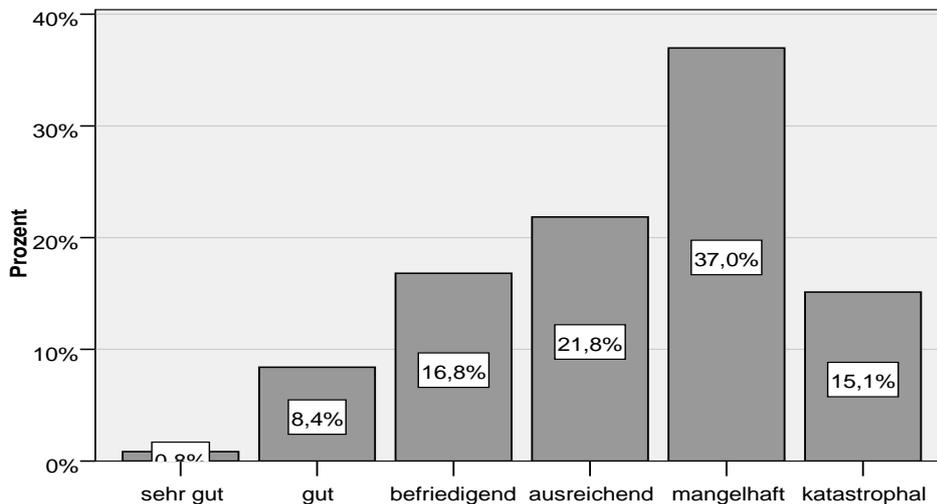
Die Daten waren nicht normalverteilt. Der Median und der Modalwert der Einkommenskategorien beträgt € 700,-. Es fällt auf, dass die Kategorie „Bis 400“ häufig gewählt wurde. Dieser Betrag ist anzuzweifeln, denn laut Regelsatzverordnung des §28 des zwölften Buches SGB betrug das von der ARGE gezahlte ALG II für Einzelpersonen in den Jahren 2007, 2008 und 2009 jeweils € 347,-, € 351,- und € 359,-. Nimmt man dazu noch das unter kommunaler Hoheit verwaltete Mietgeld der Bruttowarmmiete für ALG II-Empfänger von durchschnittlich € 360,-, sollte sich für Einzelpersonen mit 100% der Regelleistung ein Einkommen von € 711,- (im Jahr 2008) ergeben. Für ein Ehepaar (ohne Kind) von zwei ALG II-Empfängern sollte bei 90% der Regelleistung im Jahr 2008 ein Haushaltseinkommen von ALG II: € 630,- + Mietgeld: € 444,- = € 1074,- resultieren.

Eventuell war die Frage nach dem Einkommen nicht klar genug gestellt. Es lässt sich vermuten, dass einige Vpn das Mietgeld nicht mit in ihre Berechnungen aufnahmen, da dieses üblicherweise von der Kommune direkt an den Vermieter gezahlt wird. Manche Vpn gaben mündlich zu verstehen, dass sie nicht genau wüssten, wie viel Mietgeld sie überhaupt bekämen. Leider wurde dieser Punkt nicht systematisch erhoben. Weiterhin wird bei Probanden in fester Lebensgemeinschaft vermutet, dass sie das Haushaltseinkommen schlicht mit ihrem alleinigen Einkommen verwechselten oder nicht wussten, was der Partner verdiente. Beispielsweise stuften sich 2 Versuchsteilnehmer, die in einer Partnerschaft lebten, in die Kategorie „bis 400“ ein, 6 weitere von denen in die Kategorie „400-600“. Diese Angaben sind zu unrealistisch, um davon auszugehen, dass die Probanden die Frage richtig beantwortet haben. Es kann ferner spekuliert werden, dass die Einkommensfrage als zu heikel aufgefasst wurde, so dass die Probanden die Angaben aus Gründen sozialer Erwünschtheit untertrieben. Die sog. „Hartz4-Debatte“ (z.B. Guido Westerwelle: "Es gibt kein Recht auf staatlich bezahlte Faulheit."<sup>6</sup>) war in den Medien besonders im Wahljahr 2009 präsent und die Probanden fühlten sich eventuell implizit aufgefordert, hier einem aktuellen Stigma zu widersprechen. Von daher wird empfohlen, dass der Modalwert für eine Kategorie maßgeblich für die durchschnittliche Haushaltseinkommensschätzung sein sollte. Dieser betrug für die Single-Haushalte 600-800 € und für die Paare 1000-1200 €.

Aufschlussreich – und wahrscheinlich für die psychische Beanspruchung der Probanden relevanter – ist die subjektive Bewertung des finanziellen Auskommens, denn diese wird auch von Faktoren beeinflusst, die sich nicht ohne weiteres erschließen. Ein von

<sup>6</sup> <http://www.welt.de/politik/bundestagswahl/article4469109/Westerwelle-wettert-gegen-bezahlte-Faulheit.html>

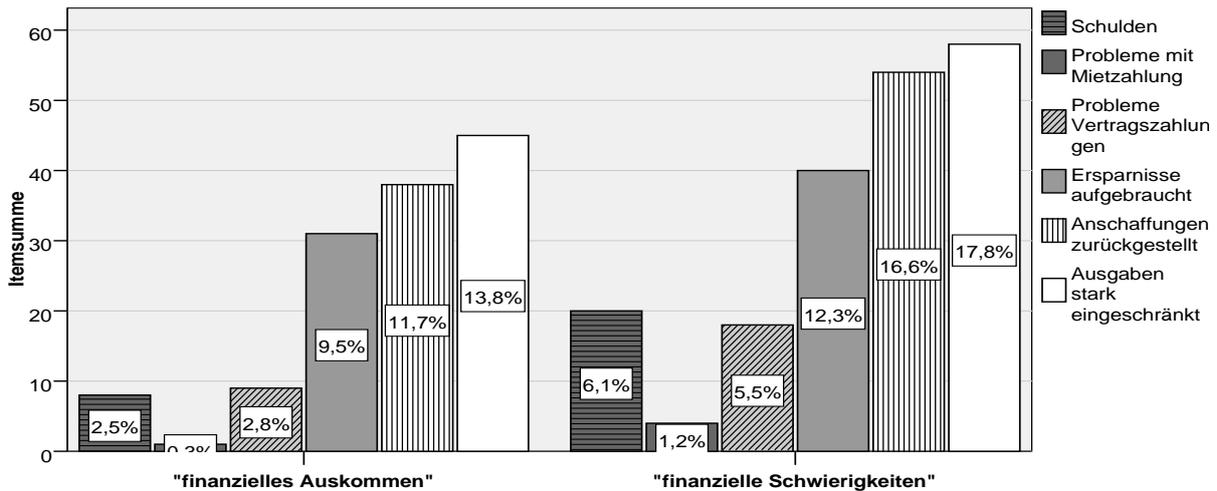
Arbeitslosigkeit betroffener Mensch, der gleichzeitig Schulden hat, empfindet gleichniedrige finanzielle Mittel sicherlich belastender als ein schuldenfreier Erwerbsloser. Die Vpn sollten auf einer Skala von „sehr gut“ bis „katastrophal“ einschätzen, wie sie mit den ihnen verfügbaren Mitteln ihr Auskommen bestreiten können (Abb. 7).



**Abb. 7:** Subjektive Bewertung der zur Verfügung stehenden finanziellen Mittel.

Die Mehrzahl der Probanden bezeichnete ihre finanzielle Situation als „mangelhaft“ (37%), gefolgt von „ausreichend“ (22%), „befriedigend“ (17%) und „katastrophal“ (15%). Da die ersten vier Kategorien semantisch ein finanzielles Auskommen suggerieren, während dies bei den letzten zwei nicht der Fall sein dürfte, lässt sich inhaltlich die Einteilung des Merkmals in zwei annähernd gleich große Gruppen rechtfertigen, in denen „sehr gut“ bis „ausreichend“ zur Kategorie „*finanzielles Auskommen*“ (N=57; 48%) und „mangelhaft“ bis „katastrophal“ zur Kategorie „*finanzielle Schwierigkeiten*“ (N=62; 52%) zusammen gefasst werden.

An einer anderen Stelle wurden die Probanden gebeten, in einem Set von Ja/Nein-Fragen finanzielle Probleme auszuwählen, die auf sie möglicherweise zutreffen. Mehrfachantworten waren möglich. Wie oft diese finanziellen Problemkategorien zutrafen ist gemeinsam mit den zusammengefassten Kategorien zur „*finanziellen Selbsteinschätzung*“ in Abb. 8 dargestellt, um eine inhaltliche Validierung zu ermöglichen.



**Abb. 8:** Finanzielle Problemkategorien (siehe Legende), aufgesplittet nach „finanzielles Auskommen“ vs. „finanzielle Schwierigkeiten“. Die Prozentzahlen beziehen sich auf die Gesamtzahl aller erfassten Problemkategorien. Interessiert der Gesamtprozentsatz, mit dem z.B. das finanzielle Problem „Schulden“ in Relation zu den anderen gewählt wurde, müssen die Prozentzahlen zwischen den Gruppen addiert werden („Schulden“: 2,5% + 6,1% = 8,6%).

Finanzielle Problemkategorien lagen für die, welche in die Gruppe „finanzielle Schwierigkeiten“ eingeordnet wurden, durchweg höher und das in vier Fällen signifikant ( $\chi^2$  nach Pearson): Diese Probanden hatten häufiger Schulden ( $p=0,02$ ), hatten geplante Anschaffungen öfter zurückgestellt ( $p=0,00$ ) und sahen ihre Ausgaben eingeschränkter ( $p=0,01$ ) als die restlichen Vpn. In der Tendenz können ebenso mehr Probleme mit Vertragszahlungen ( $p=0,07$ ) attestiert werden. „Ersparnisse aufgebraucht“ ( $p=0,19$ ) und „Probleme mit Mietzahlung“ ( $p=0,21$ ) war nicht signifikant verschieden zwischen den Gruppen. Da die kategoriale Einteilung in „finanzielles Auskommen“ und „finanzielle Schwierigkeiten“ in den meisten Problembereichen trennscharf zwischen den Subgruppen fungierte, wird die inhaltliche Validität der Gruppierung „finanzielle Selbsteinschätzung“ als bestätigt angesehen. „Probleme mit Mietzahlung“ war möglicherweise durch die kommunale direkte Überweisung der Miete nicht unterschiedlich. Ebenso war „Ersparnisse aufgebraucht“ vermutlich deshalb nicht signifikant unterschiedlich, weil zur Beziehung von ALG II die Sparguthaben bis zur Grenze des sog. "Schonvermögens" verbraucht sein müssen.

#### 4.1.3. Häufigkeit sozialer Kontakte

Die Versuchspersonen sollten abschätzen, wie viele Stunden sie am Tag mit sozialen Kontakten (z.B. Partner, Familie, Freunde, evt. Kollegen) verbringen. Diese Daten waren nicht normalverteilt, sondern rechtsschief. Als Maß der zentralen Tendenz wurde der Maximum-Likelihood-Schätzer nach Huber (Janssen & Laatz, 2010) berechnet. Dieser betrug täglich 3,8 Stunden für die mit sozialen Kontakten verbrachte Zeit.

#### 4.1.4. Qualität sozialer Unterstützung

Mit einem Fragebogen zur sozialen Unterstützung (Nitsche & Richter, 2003) wurden die Probanden nach der Qualität ihrer sozialen Kontakte befragt. Hierbei sollte auf einer Skala von 0-3 (0="gar nicht" bis 3="völlig") angegeben werden, wie man sich jeweils auf die Personengruppen „Partner“, „Kinder & Familie“, „Freunde“, „Kollegen“ und „Vorgesetzte“ 1.) verlassen kann oder ob man 2.) von diesen Personengruppen soziale Unterstützung erhält oder ob man sich 3.) mit ihnen aussprechen kann. Für die Bereiche wurden Durchschnittswerte gebildet. Die Daten waren normalverteilt und erreichten folgende Mittelwerte und Standardabweichungen.

**Tab. 5:** Gesamte und private Qualität der sozialen Unterstützung

	„Ich kann mich auf diese Menschen verlassen“.	"Ich erhalte von diesen Menschen Unterstützung".	"Ich kann mich mit diesen Menschen aussprechen".	Durchschnitt
Gesamte soz. Unterst.	1,15±0,51	1,09±0,52	1,14±0,51	1,13±0,50
„Private“ soz. Unterst.	1,71±0,73	1,63±0,75	1,80±0,73	1,68±0,71

Beim Gesamtmittelwert fällt auf, dass die Probanden in jedem Bereich nur etwas mehr als ein Drittel der Höchstpunktzahl (3) erreichen. Dies ist auch der Tatsache geschuldet, dass die meisten Probanden (d.h. die nicht stundenweise Beschäftigten) über keine „Mitarbeiter“ oder „Vorgesetzte“ verfügten, sodass für diese Personengruppen des Fragebogens der Wert 0 auftreten musste. Bildet man für die Personengruppen „Partner“, „Kinder & Familie“ und „Freunde“ den Durchschnitt für diese „privaten“ sozialen Kontakte, werden verbesserte Mittelwerte erreicht (vgl. zweite Zeile der Tabelle). Die im Vergleich zur Gesamtunterstützung erhöhte soziale Unterstützung durch die privaten Quellen spricht dafür, dass die meisten Probanden von (ehemaligen) Kollegen oder Vorgesetzten kaum soziale Unterstützung erhalten.

**Tab. 6:** Qualität der berufsbedingten sozialen Unterstützung nach Teilzeitbeschäftigung

Stundenweise Beschäftigung	ARBEIT "Verlassen können"	ARBEIT "Unterstützung erhalten"	ARBEIT "Aussprechen können"	Durchschnitt „arbeitsbedingte soziale Unterstützung“
Nein (N=92)	0,15±0,47	0,14±0,48	0,15±0,51	0,14±0,47
Ja (N=27)	0,87±0,98	0,78±0,92	0,76±0,89	0,80±0,89
Gesamt	0,31±0,69	0,28±0,66	0,29±0,67	0,29±0,65

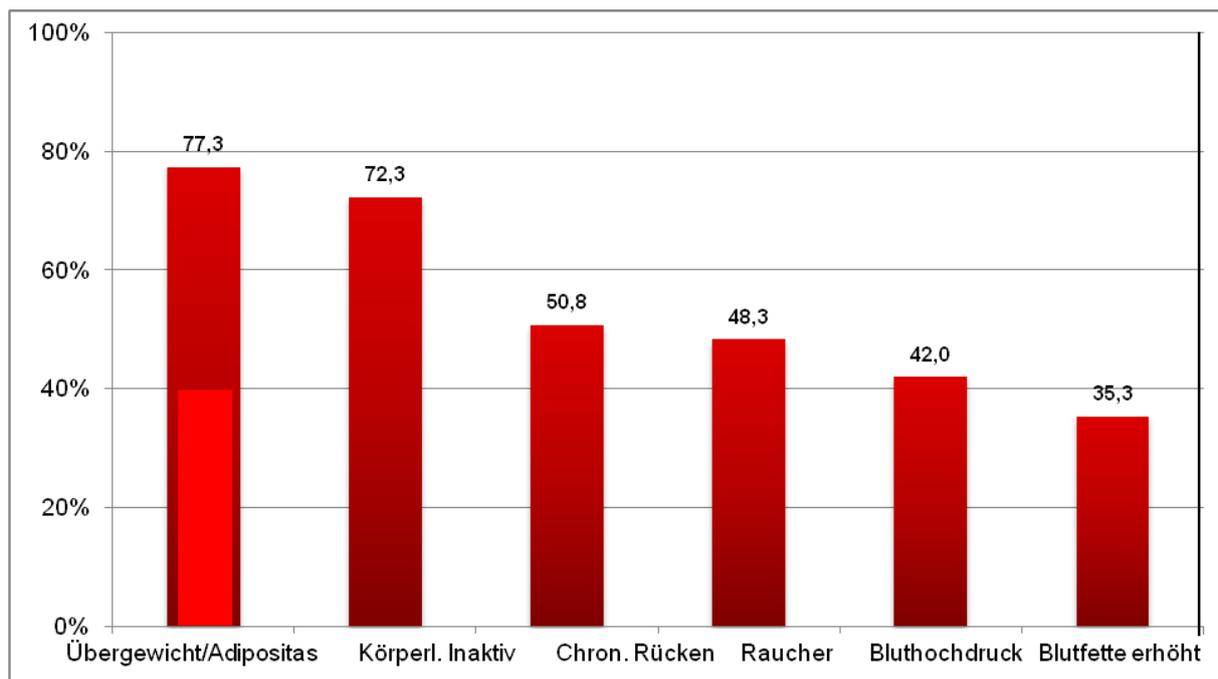
In der Stichprobe befanden sich 27 Probanden, die einer stundenweisen Beschäftigung nachgingen und 92 ohne eine Beschäftigung. Teilt man die Probanden danach ein und prüft die Unterschiede der sozialen Unterstützung durch die Quellen „Mitarbeiter“ und „Vorgesetzte“ (in der Tabelle als „Arbeit“ zusammengefasst), ergeben sich deskriptiv höhere Werte für die Vpn mit Nebenjob. Beim Inspizieren der Durchschnittswerte ist jedoch ersichtlich, dass selbst die Probanden mit einer stundenweisen Beschäftigung kaum eine berufliche soziale Unterstützung erhielten. Ihr Mittelwert erreicht mit 0,8 noch nicht einmal die

Qualität von arbeitsbedingter sozialer Unterstützung=1 (entspricht im Fragebogen: „wenig soz. Unterstützung“). In der Tat gaben 15 (55%) der 27 teilweise beschäftigten Probanden an, sie hätten „0“, also keinerlei soziale Unterstützung, die von ihrer Beschäftigung ausging, weder von ihren Kollegen, noch von ihren Vorgesetzten.

Da diese niedrigen Werte etwas überraschten, wurde explorativ für die Probanden mit stundenweiser Beschäftigung eine Korrelation zwischen „privater Unterstützung“ und „arbeitsbedingter Unterstützung“ durchgeführt, um zu prüfen, wie hoch diese Quellen zusammenhängen. Es ergab sich keinerlei Zusammenhang, noch nicht einmal als Trend ( $r=-0,059$ ;  $p=0,769$ ). Eine Erwartungshaltung im Sinne von: „Wer sich privat eine gute soziale Unterstützung zu sichern vermag, kann dies mit hoher Wahrscheinlichkeit auch in der Arbeitswelt“ wäre für die stundenweisen Beschäftigten nicht belegbar gewesen.

## 4.2. Physische Gesundheit

Dass in der von uns untersuchten Gruppe ein erheblicher Interventionsbedarf besteht, kann man z.B. der Häufigkeit der bei der Erstuntersuchung festgestellten Risikofaktoren für Herz-Kreislauferkrankungen entnehmen, die in der folgenden Abbildung dargestellt ist.



**Abb. 9:** Darstellung der wesentlichen körperlichen Befunde/ Diagnosen bei Erstuntersuchung (N=119) Übergewicht und Adipositas; Körperliche Inaktivität; Chronische Rückenschmerzen; Raucherstatus, Bluthochdruck, erhöhte Blutfette

Übergewicht/Adipositas: In der ersten Säule wurden zur besseren Übersicht übergewichtige und adipöse Probanden zusammengefasst. Üblicherweise gelten Personen mit einem Body Mass Index (BMI [ $\text{kg}/\text{m}^2$ ]) zwischen 25 und 30 als übergewichtig und Personen mit einem BMI oberhalb von 30 als adipös. Zu beachten ist, dass bei Adipösen nicht nur das Körpergewicht per se, sondern der Anteil der Körperfettmasse gesundheitsrelevant erhöht ist. Angaben aus der GEDA-Studie „Gesundheit in Deutschland 2009“ (RKI, 2010) zufolge

liegt der Anteil der übergewichtigen Frauen in der betrachteten Altersgruppe in Deutschland bei 31%, bei den gleichaltrigen Männern bei fast 50%. Adipositas besteht bei knapp 19% der Frauen und 21,1 % der Männer in der altersbezogenen Allgemeinbevölkerung. Werden Übergewichtige und Adipöse zusammengefasst, ergibt sich aus den genannten Zahlen der GEDA-Studie für die Frauen ein Anteil von 50%, für die Männer ein Anteil von 71,1% der altersentsprechenden Allgemeinbevölkerung.

Unter Berücksichtigung der Geschlechterverteilung von 40:60 zugunsten der Frauen war der Anteil Übergewichtiger und Adipöser in der von uns untersuchten Population mit 77% somit deutlich über dem zu erwartenden Anteil. Der mittlere BMI lag bei Erstuntersuchung oberhalb von 30 kg/m<sup>2</sup>, also bereits im Bereich der Adipositas und der Anteil der Adipösen bei 40%, somit doppelt so hoch wie aus der Allgemeinbevölkerung bekannt. Vergleichsdaten aus dem Projekt AmigA belegen einen Anteil Adipöser von 30,2% bei einem Durchschnittsalter von 36,7 Jahren, aus dem Projekt JOB Fit/NRW ist ein Anteil Adipöser von 19,9% bei einem Durchschnittsalter von 40,9 Jahren bekannt (Kirschner, 2009).

Körperliche Inaktivität wurde attestiert, wenn ein Mindestmaß von 30 Minuten körperlicher Aktivität pro Woche unterschritten wurde. Die Schwelle, die den Bewegungsmangel definierte, war also sehr niedrig angesetzt. Laut Empfehlung des US-amerikanischen Centers of Disease Control and Prevention (CDC) sollten sich Erwachsene an mindestens 5 Tagen für mindestens 30 Minuten körperlich betätigen. In der GEDA-Studie gaben 59% der Frauen und 57% der Männer in der Altersgruppe 45-64 Jahre an, weniger als 2,5 Stunden pro Woche körperlich aktiv zu sein. Anzumerken ist, dass die Erfassung der körperlichen Aktivität ein generelles methodisches Problem ist.

Der Anteil der Raucher ist in der hier untersuchten Population mit 48 % vergleichsweise hoch. Der GEDA-Studie zufolge rauchen knapp 30% der altersentsprechenden Frauen und ca. 35% der altersentsprechenden Männer in Deutschland. Die Auswertung der Daten des telefonischen Gesundheitssurveys von 2003 ergab eine starke Beeinflussung des Rauchverhaltens durch den sozialen Status mit einem Anteil von 55% männlicher und 45% weiblicher Raucher in der Gruppe mit niedrigem Sozialstatus. Diese Zahlen lassen sich anhand der aktuellen GEDA-Studie, die den Einfluss des sozialen Status ebenfalls berücksichtigte, aber nicht mehr belegen.

Ein Bluthochdruck bestand in 42% der Fälle und lag damit ebenfalls deutlich über den alters- und geschlechtsspezifischen Angaben für die deutsche Bevölkerung mit einer Prävalenz von ca. 34% für beide Geschlechter in der Altersgruppe 50-59 Jahre (RKI, 2006).

Die Prävalenz von Fettstoffwechselstörungen in der Allgemeinbevölkerung ist schwer zu beziffern, da die Diagnosestellung häufig vom Vorliegen weiterer Risikofaktoren abhängt. Außerdem gibt es verschiedene Formen der Fettstoffwechselstörungen, die in epidemiologischen Untersuchungen häufig unter dem Sammelbegriff subsumiert werden. Man kann aber in etwa von einer Prävalenz von 30 % in der Allgemeinbevölkerung ausgehen. Somit lag der Anteil der Probanden mit Fettstoffwechselstörungen in der hier untersuchten Gruppe etwas darüber.

Grafisch nicht dargestellt ist der Anteil der manifesten Diabetiker in dieser Studie, der mit 10% etwas über dem Bundesdurchschnitt lag.

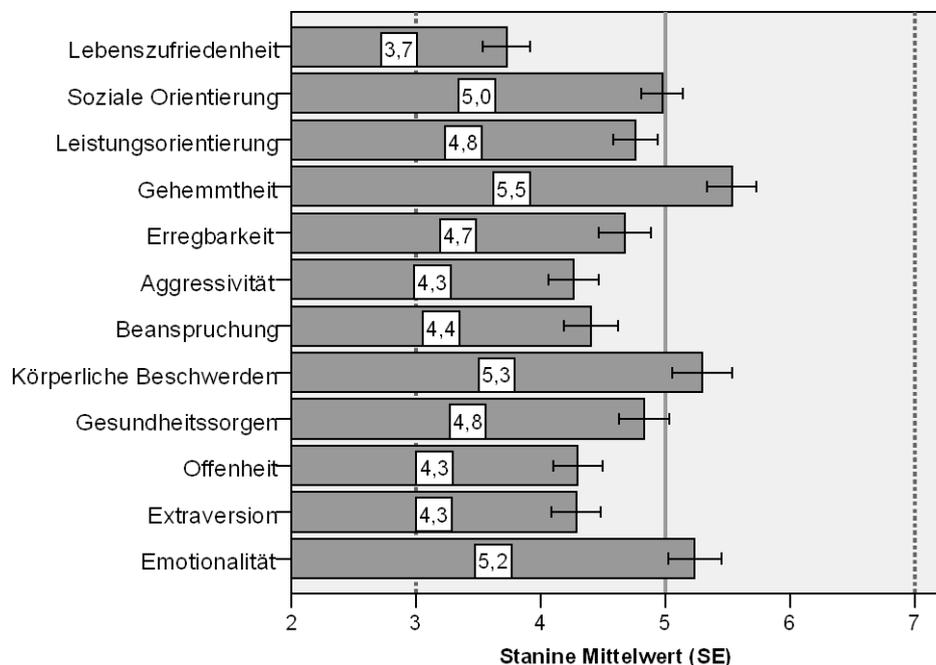
Es kann zusammengefasst werden, dass das Vorliegen einzelner kardiovaskulärer Risikofaktoren deutlich häufiger festgestellt wurde als aus alters- und geschlechtsspezifischen Vergleichsdaten zu erwarten gewesen wäre. Neben dem ermittelten Interventionsbedarf zeigt die Erhebung aber auch, dass über den hier vorgestellten Zugangsweg Personen erreicht werden können, die Präventionsangeboten der Hausärzte und Krankenkassen zur Verminderung des Risikos für Erkrankungen des Herz-Kreislaufsystems (insbesondere Herzinfarkt- und Schlaganfall) schlecht zugänglich sind.

In der Abbildung 9 ist die Häufigkeit von Rückenschmerzen mit einbezogen worden, da chronische Rückenschmerzen erstens ein häufiges Vermittlungshemmnis darstellen, und zweitens, ebenso wie kardiovaskuläre Risikofaktoren einer positiven Beeinflussung durch Sport unterliegen, somit den Interventionsbedarf unterstreichen.

### 4.3. Psychische und psychosoziale Gesundheit

#### 4.3.1. Freiburger Persönlichkeitsinventar (FPI)

Der FPI wurde später in die Studie aufgenommen und konnte für N=86 Vpn angewandt und für 84 Vpn ausgewertet werden. 2 Probanden hatten zu starke sprachliche Verständnisschwierigkeiten, so dass der Test für diese nicht ausgewertet wurde.



**Abb. 10:** Persönlichkeitsdimensionen des Freiburger Persönlichkeitsinventars

Die Rohwerte der Persönlichkeitsfaktoren wurden in alters- und geschlechtsspezifische Stanine-Normen (nach Fahrenberg et al., 2010) transformiert. Stanine kann Werte von 1 bis 9 annehmen, mit einem Mittelwert von 5 und einer Standardabweichung von 1,96. Demnach gilt für Stanine die Spanne von 3,04 bis 6,96 als Normbereich. In der untersuchten

Stichprobe lagen die mittleren Stanine zu den Persönlichkeitsfaktoren im unauffälligen Bereich. Demnach ist nicht davon auszugehen, dass die hier untersuchten älteren Langzeiterwerbslosen durchschnittlich über Persönlichkeitsstrukturen verfügten, die von der entsprechenden Bevölkerung maßgeblich abweichen. Auch das Merkmal „Offenheit“<sup>7</sup> befand sich im Normbereich.

Das Persönlichkeitsmerkmal „Lebenszufriedenheit“ lag am weitesten vom Normstichprobenmittelwert entfernt. Bei der geschlechtsspezifischen Auswertung des Merkmals stellte sich heraus, dass Männer eine signifikant niedrigere Lebenszufriedenheit ( $3,34 \pm 1,53$ ) im Vergleich zu den Frauen ( $4,09 \pm 1,91$ ) aufwiesen ( $t_{(df=82)}=1,994$ ;  $p=0,049$ ). Des Weiteren zeigte sich eine signifikant höhere „soziale Orientierung“ der Männer ( $5,51 \pm 1,27$ ) im Vergleich zu den Frauen ( $4,51 \pm 1,58$ ) ( $t_{(df=82)}=3,151$ ;  $p=0,004$ ).

Da eine eingeschränkte Lebenszufriedenheit als Messgröße zur Bestimmung eines Risikofaktors der psychischen Gesundheit definiert wurde, erfolgte die Berechnung des Anteils von Probanden mit einer eingeschränkten Lebenszufriedenheit. Stanine 1-3 wurde als auffälliger Bereich gezählt.

**Tab. 7:** Kreuztabelle "niedrige Lebenszufriedenheit" und Geschlecht

			LEB ≤ 3 Stanine		Gesamt
			niedrig	o.B.	
Geschlecht	Männlich	Anzahl	24	17	41
		% von Geschlecht	58,5%	41,5%	100,0%
		% von LEB≤3	54,5%	42,5%	48,8%
	Weiblich	Anzahl	20	23	43
		% von Geschlecht	46,5%	53,5%	100,0%
		% von LEB≤3	45,5%	57,5%	51,2%
Gesamt		Anzahl	44	40	84
		% von Geschlecht	52,4%	47,6%	100,0%
		% von LEB≤3	100,0%	100,0%	100,0%

Wie in der Kreuztabelle ersichtlich, befanden sich 52% der Stichprobe (59% der Männer, 47% der Frauen) im Bereich Stanine 1-3 des Merkmals Lebenszufriedenheit. In einer Normalverteilung deckt der Bereich 1-3 von Stanine 23% ab, d.h., in dieser Stichprobe lag der Anteil von Menschen mit einer eingeschränkten Lebenszufriedenheit 2,3 mal so hoch, wie aufgrund der alters- und geschlechtsspezifischen Normwerte zu erwarten gewesen wäre.

In den nächsten Abschnitten werden die Ergebnisse der Eingangsuntersuchungen für die Fragebögen dargestellt, die nicht nur einmalig, sondern auch für die Wiederholungsmessungen T2 und T3 angewandt wurden.

<sup>7</sup> Dieses Merkmal wird in älteren Persönlichkeitstests oft als „Lügenskala“ bezeichnet. Es misst das Ausmaß, wie sehr ein Proband dazu bereit ist, typisch menschliche Schwächen (z.B. „Manchmal verschiebe ich etwas, das ich heute erledigen wollte auf morgen.“) zuzugeben und ist von daher ein Maß für die Offenheit des Untersuchten.

### 4.3.2. General Health Questionnaire (GHQ)

Der GHQ konnte für die gesamte Stichprobe (N=119) angewandt und ausgewertet werden.

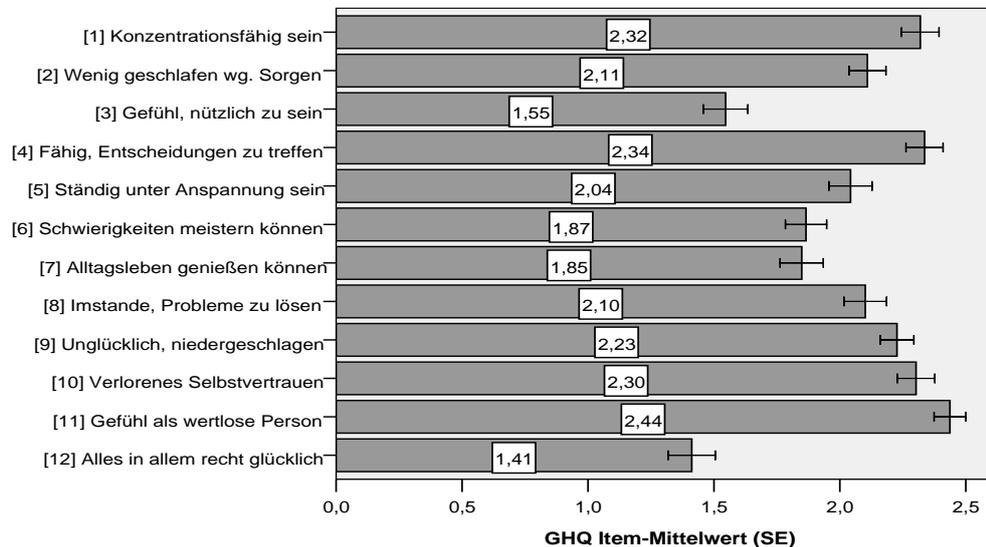


Abb. 11: Ergebnisse des GHQ auf Itemebene

Die in der Grafik dargestellten Werte wurden für negativ formulierte Items (Item 2, 5, 9, 10, 11) umgepolt<sup>8</sup>, damit konsistenter Weise positive Werte für einen positiven psychischen Zustand sprechen. Über alle Items resultierte ein Mittelwert von  $2,05 \pm 0,50$ . Frauen ( $2,00 \pm 0,49$ ) und Männer ( $2,12 \pm 0,52$ ) unterschieden sich nicht ( $t_{(df=117)}=1,242$ ;  $p=0,221$ ), welches dem Ergebnis von (Goldberg et al., 1997) zur Geschlechtsunabhängigkeit des GHQ entspricht.

Laut der Itemwerte hatten die Probanden einerseits Schwierigkeiten beim Empfinden von allgemeinen Glück (Item 12) und dem Gefühl, nützlich zu sein (Item 3). Genussfähigkeit im Alltag (Item 7) und das Vertrauen, Schwierigkeiten meistern können (Item 6), hatten ebenso Verbesserungspotentiale. Auf der anderen Seite hatten die Versuchspersonen relativ wenige Probleme, die absolutistische Aussage „ich fühle mich als wertlose Person“ (Item 11) abzulehnen. Ebenso gaben die Probanden mit hohen Mittelwerten an, Entscheidungen treffen zu können (Item 4) und konzentrationsfähig zu sein (Item 1).

Es existieren unterschiedliche Normen zum GHQ in der 12 Item Version. Die für den deutschen Sprachraum repräsentativsten stammen von Goldberg et al. und greifen auf eine internationale Stichprobe von N=25.916 aus 15 Zentren zurück, anhand derer spezifische „Thresholds“ berichtet werden (Goldberg et al., 1997). Als maßgeblich für den deutschen Sprachraum werden die Analysen für die Zentren aus Mainz und Berlin angesehen. Laut den Normen von Goldberg liegt der für diese Zentren berechnete Threshold jeweils bei 2/3 (Goldberg et al., 1997, S. 194) und korrespondiert mit einer zufriedenstellenden mittleren Spezifität von 73,1% und einer mittleren Sensitivität von 78,1% für die Identifikation von Fällen mit unspezifischen psychischen Beschwerden.

<sup>8</sup> 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.

Um die Normen zum GHQ anwendbar zu machen, musste pro Item die Likert-Skala (0-1-2-3) in ein sog. „GHQ-Scoring“ (0-0-1-1) nach den Vorgaben von Piccinelli et al. umkodiert werden (Piccinelli, Bisoffi, Bon, Cunico & Tansella, 1993). Danach wurde aus den GHQ-Scores pro Versuchsteilnehmer ein individueller Durchschnittswert berechnet, der zwischen 0 und 1 variierte. Der Schwellenwert von 2/3 lieferte für die untersuchten Langzeitarbeitslosen folgende Fallidentifikationen („case“):

**Tab. 8:** Kreuztabelle "GHQ-Case" und Geschlecht

			GHQ Case-Identifikation		Gesamt
			Schwellenwert 2/3		
			o.B.	Case	
Geschlecht	Männlich	Anzahl	31	15	46
		% von Geschlecht	67,4%	32,6%	100,0%
		% von GHQ-Case	42,5%	32,6%	38,7%
	Weiblich	Anzahl	42	31	73
		% von Geschlecht	57,5%	42,5%	100,0%
		% von GHQ-Case	57,5%	67,4%	61,3%
Gesamt	Anzahl	73	46	119	
	% von Geschlecht	61,3%	38,7%	100,0%	
	% von GHQ-Case	100,0%	100,0%	100,0%	

Demnach lassen sich im Vergleich zur Normstichprobe für 39% der Versuchsteilnehmer „Fälle“ von allgemeinen psychischen Beschwerden identifizieren. Da GHQ Werte in der Population normalverteilt sind (Goldberg, et al., 1997), entspricht die hier identifizierte Fallzahl mehr als doppelt so vielen Fällen wie zu erwarten (15,9%) gewesen wäre. Die Fallzahlen waren zwischen den Geschlechtern (33% der Männer, 43% der Frauen) nicht signifikant verschieden ( $\chi^2=1,162$ ;  $p=0,281$ ).

### 4.3.3. Beck Depressions-Inventar (BDI)

Der BDI wurde später in die Studie aufgenommen und konnte für N=89 Vpn angewandt und für N=87 ausgewertet werden; 2 Versuchspersonen brachen den Test aufgrund sprachlicher Verständnisschwierigkeiten ab. Im folgenden Diagramm finden sich die Item-Mittelwerte.

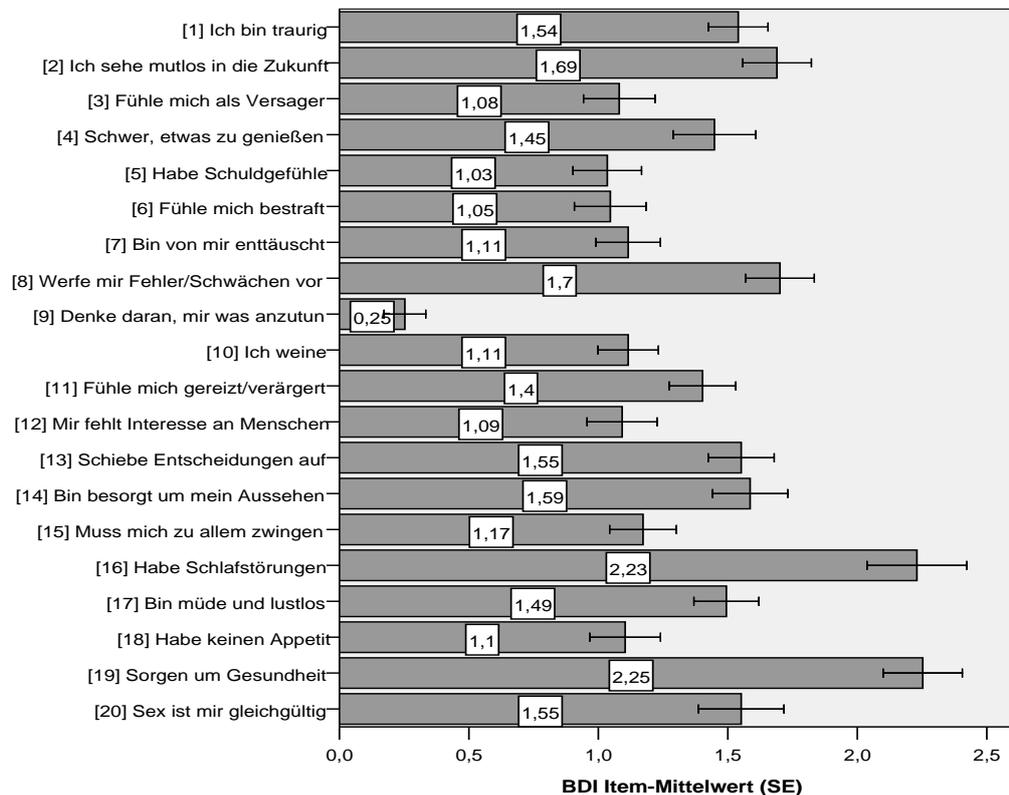
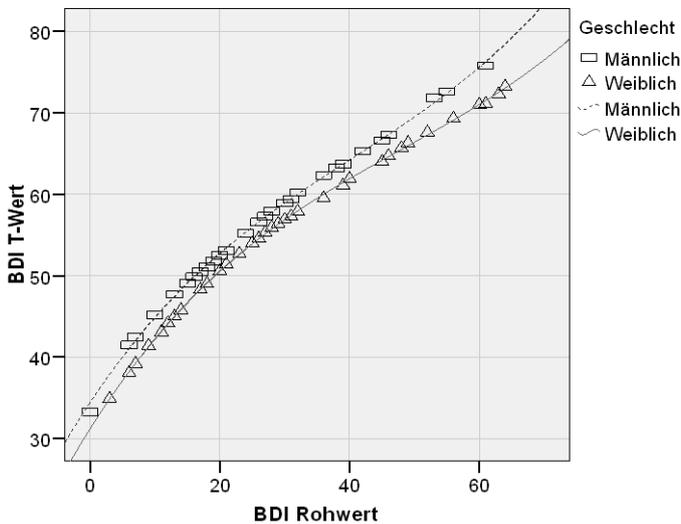


Abb. 12: Ergebnisse des BDI auf Itemebene

Die vergleichsweise höchsten Zustimmungen erhielten das Item zu Gesundheitsorgen (Item 19) und das zu Schlafstörungen (Item 16). Die Langzeiterwerbslosen warfen sich häufig Fehler und Schwächen vor (Item 8), sahen mutlos in die Zukunft (Item 2) und hatten eine negativ gefärbte körperliche Selbstwahrnehmung (Item 14). Selbstbestrafende bzw. suizidale Gedanken (Item 9) erhielt glücklicherweise die kleinste Zustimmung von allen Aussagen.

Zwischen weiblichen und männlichen Versuchsteilnehmern waren im Mann-Whitney U-Test Item 1 „Ich bin traurig“ ( $Z=2,086$ ;  $p=0,037$ ), Item 10 „Ich weine“ ( $Z=2,174$ ;  $p=0,030$ ), Item 17 „Bin müde und lustlos“ ( $Z=2,546$ ;  $p=0,011$ ) und Item 20 „Sex ist mir gleichgültig“ ( $Z=3,061$ ;  $p=0,002$ ) signifikant für die weiblichen Probanden erhöht. Geschlechtsunterschiede sind im BDI ein bekanntes Phänomen, das 1,5% der Gesamtvarianz aufzuklären vermag (Schmitt, Altstotter-Gleich, Hinz, Maes & Brähler, 2006). Diesem Umstand trägt die geschlechtsspezifische Transformation in Normwerte Rechnung, in welcher Frauen einen etwas höheren Rohwert als Männer benötigen, um einen vergleichbaren standardisierten T-Wert zu erhalten (vgl. Abb. 13). Hierbei lässt sich der Zusammenhang zwischen BDI-

Rohwerten und T-Werten pro Geschlecht durch kubische Regressionsgleichungen beschreiben ( $R^2=0,99$ ;  $p=0,000$ ).



**Abb. 13:** Zusammenhang von BDI Rohwerten und BDI T-Werten

T-Werte haben in der Stichprobe der Allgemeinbevölkerung von Schmitt et al. (N=4.494) einen Mittelwert von 50 und eine Standardabweichung von 10 (Schmitt et al., 2006). Nach der Transformation ergab sich für unsere Gesamtstichprobe ein Mittelwert von  $T=54,99 \pm 9,47$ . Zwischen weiblichen ( $T=55,26 \pm 9,55$ ) und männlichen ( $T=54,65 \pm 9,47$ ) Probanden fand sich kein Unterschied ( $t_{(df=85)}=0,300$ ;  $p=0,765$ ). Anhand der geschlechts- und altersspezifischen Normwerte empfehlen Schmitt und Mitarbeiter als Schwellenwert für die Identifikation von klinisch relevanten Depressivitätswerten den T-Wert=60. Dieser geht einher mit einer Sensitivität von 92% und einer Spezifität von 91%.

Nach Anwendung des empfohlenen Schwellenwertes auf die Daten der Stichprobe ergaben sich folgende Fallzahlen für eine klinisch relevante Depressivität:

**Tab. 9:** Kreuztabelle "Depressivität" und Geschlecht

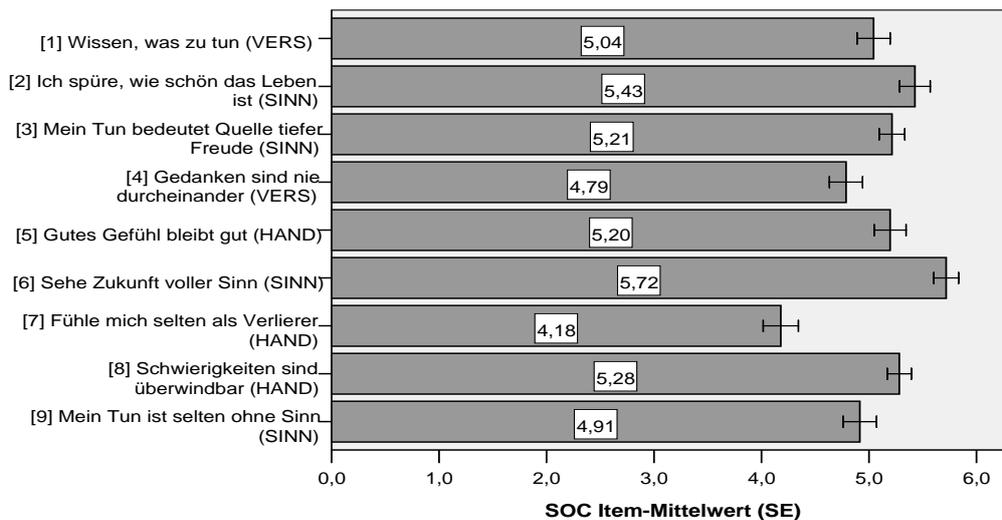
			BDI-T>60		Gesamt
			o.B.	Depressiv	
Geschlecht	Männlich	Anzahl	27	11	38
		% von Geschlecht	71,1%	28,9%	100,0%
		% von BDI-T>60	45,0%	40,7%	43,7%
	Weiblich	Anzahl	33	16	49
		% von Geschlecht	67,3%	32,7%	100,0%
		% von BDI-T>60	55,0%	59,3%	56,3%
Gesamt	Anzahl	60	27	87	
	% von Geschlecht	69,0%	31,0%	100,0%	
	% von BDI-T>60	100,0%	100,0%	100,0%	

Demnach waren insgesamt 31% (Männer: 29%, Frauen: 33%) der hier untersuchten Langzeiterwerbslosen von Depressivität betroffen, d.h. annähernd doppelt so viele Fälle wie

aufgrund der Normdaten zur Allgemeinbevölkerung (15,9%) erwartet werden konnten. Ein Geschlechtsunterschied zeigte sich zwischen den Fallzahlen nicht ( $\chi^2=0,141$ ;  $p=0,709$ ).

#### 4.3.4. Kohärenzsinn – Sense of Coherence (SOC)

Der SOC wurde bei N=119 Probanden angewandt und konnte für N=117 Probanden ausgewertet werden; 2 Versuchspersonen brachen den Test aufgrund von sprachlichen Verständnisschwierigkeiten ab. In folgender Grafik sind die Mittelwerte und Standardfehler der Items dargestellt.



**Abb. 14:** Ergebnisse des SOC auf Itemebene ("Hand": Handhabbarkeit, "Vers": Verstehbarkeit, "Sinn": Sinnhaftigkeit)

Auf der Skala von 1-7 lagen fast alle Items um den Wert 5 herum. Das Item 7 „Ich fühle mich selten als Verlierer“ weicht am deutlichsten von den restlichen ab. Schumacher et al. beschreiben, 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 (Schumacher, Wilz, Gunzelmann & Brähler, 2000). Aufgrund der mangelnden faktoriellen Validität empfehlen die Autoren, dass der Kohärenzsinn – gemessen mit dem SOC – als Generalfaktor analysiert werden sollte.

Für den SOC-L9 existiert eine Normierung von Schumacher (2000). Aus den summierten Items wurden anhand der Normstichprobe alters- und geschlechtsspezifische Prozentränge ermittelt. Um die Vergleichbarkeit zu anderen in dieser Arbeit vorgestellten Skalen zu erleichtern, erfolgte eine Transformation der Prozentränge in T-Werte. Für die Gesamtstichprobe ergab sich ein Mittelwert von  $T=48,51 \pm 9,35$ . Zwischen Frauen ( $T=48,99 \pm 9,38$ ) und Männern ( $T=47,76 \pm 9,35$ ) ergab sich kein Unterschied ( $t_{(df=115)}=0,704$ ;  $p=0,491$ ).  $T < 40$  ergab als Kriterium für ein unterdurchschnittliches Kohärenzgefühl folgende Fallzahlen in Tabelle 10:

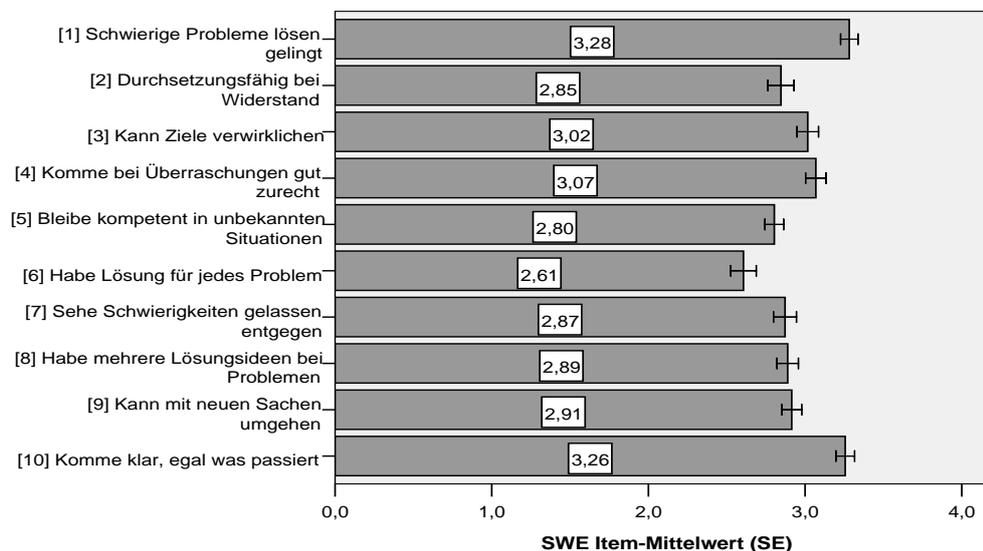
**Tab. 10:** Kreuztabelle "niedriger Kohärenzsinn" und Geschlecht

			SOC T<40		Gesamt
			SOC niedrig	zu o.B.	
Geschlecht	Männlich	Anzahl	9	37	46
		% von Geschlecht	19,6%	80,4%	100,0%
		% von SOC T<40	42,9%	38,5%	39,3%
	Weiblich	Anzahl	12	59	71
		% von Geschlecht	16,9%	83,1%	100,0%
		% von SOC T<40	57,1%	61,5%	60,7%
Gesamt	Anzahl	21	96	117	
	% von Geschlecht	17,9%	82,1%	100,0%	
	% von SOC T<40	100,0%	100,0%	100,0%	

18% der Stichprobe wies einen verminderten Kohärenzsinn auf; Männer zu 20% und Frauen zu 17% ( $\chi^2=0,131$ ;  $p=0,710$ ). Mit 18% ist dieser Anteil gegenüber dem erwarteten der Normalverteilung (15,9%) etwas erhöht.

#### 4.3.5. Allgemeine Selbstwirksamkeitserwartung (SWE)

Die SWE wurde für 119 Probanden angewandt und konnte für N=117 ausgewertet werden; 2 Versuchspersonen brachen den Test aufgrund sprachlicher Verständnisschwierigkeiten ab. In folgender Grafik sind die Durchschnittswerte der Items abgebildet.



**Abb. 15:** Ergebnisse der SWE auf Itemebene

In der SWE lassen sich keine besonders abweichenden Items erkennen. Diese relative Homogenität ist konstruktionsbedingt, da Schwarzer und Jerusalem der Skala eine eindimensionale Struktur unterstellen (Schwarzer & Jerusalem, 1995, 1999). Die summierten

Item-Rohwerte wurden in alters- und geschlechtsspezifische Normwerte anhand der bevölkerungsrepräsentativen Stichprobe (N=2.019) von Hinz et al. transformiert (Hinz, Schumacher, Albani, Schmid & Brähler, 2006). Für die Gesamtstichprobe ergab sich ein Mittelwert von  $T=50,49\pm 10,14$ . Frauen ( $T=50,62\pm 10,52$ ) und Männer ( $T=50,28\pm 9,63$ ) unterschieden sich nicht ( $t_{(df=115)}=0,174$ ;  $p=0,863$ ). Die T-Werte zeigen insgesamt und geschlechtsspezifisch, dass die hier untersuchten Langzeiterwerbslosen über eine Selbstwirksamkeitserwartung verfügen, die mit der Allgemeinbevölkerung gut übereinstimmt. Nach Anwendung des Schwellenwertes von  $T<40$  für die Identifikation von Personen mit unterdurchschnittlicher SWE ergaben sich folgende Fallzahlen:

**Tab. 11:** Kreuztabelle "niedrige Selbstwirksamkeitserwartung" und Geschlecht

			SWE <40?		Gesamt
			SWE niedrig	SWE o.B.	
Geschlecht	Männlich	Anzahl	6	40	46
		% von Geschlecht	13,0%	87,0%	100,0%
		% von SWE <40	31,6%	40,8%	39,3%
	Weiblich	Anzahl	13	58	71
		% von Geschlecht	18,3%	81,7%	100,0%
		% von SWE <40	68,4%	59,2%	60,7%
Gesamt	Anzahl	19	98	117	
	% von Geschlecht	16,2%	83,8%	100,0%	
	% von SWE <40	100,0%	100,0%	100,0%	

16% der Fälle signalisierten eine unterdurchschnittliche Selbstwirksamkeitserwartung. Da die Skala in der Population normalverteilt ist (Hinz et al., 2006), entspricht 16% fast genau dem Erwartungswert (15,9%). Die Fallzahlen unterschieden sich zwischen Männern (13%) und Frauen (18%) statistisch nicht ( $\chi^2=0,57$ ;  $p=0,451$ ).

#### 4.3.6. Fragebogen zum Gesundheitszustand (SF12) Körperliche (KSK) und Psychische Summenskala (PSK)

Der SF12 wurde für 119 Probanden angewandt und ausgewertet. Die 12 Items besitzen eine unterschiedliche Skalierung und zwar entweder 2-stufig (Items 4, 5, 6, 7), 3-stufig (Items 2, 3), 5-stufig (Items 1, 8, 12) oder 6-stufig (Items 9, 10, 11). Eine gemeinsame grafische Darstellung ist bei dieser Skalierung nicht sinnvoll, weshalb die Antwortkategorien der Items im folgenden als prozentuale Häufigkeiten dargestellt sind.

**Tab. 12:** Prozentuale Antworthäufigkeiten auf Itemebene des SF12 (KSK: Item der körperlichen Summenskala; PSK: Item der psychischen Summenskala)

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

Bei der Inspektion der Items fällt auf, dass für die Items 2-8 und 12 die höchst mögliche Antwortkategorie dominant war (fett markiert). Ausgewogener sieht es bei den Items 1, 9, 10, 11 aus, bei denen die häufigste Antwortkategorie keinen Extremwert darstellt. Deskriptiv neigten die Probanden dazu, eher körperliche statt psychische Items als Einschränkung ihrer gesundheitsbezogenen Lebensqualität zu identifizieren. Leider verzichteten Bullinger & Kirchberger in ihrem Manual auf die Angabe einer üblichen Transformation der Rohwerte in Standardnormen (Bullinger & Kirchberger, 1998). Es ist jedoch möglich, empirisch ermittelte Werte mit denen im Manual angegebenen Regressionsgewichten auf eine amerikanische Normstichprobe hochzurechnen. Zur Berechnung der Summenskalen polt man negative Items um und bildet aus den Antwortkategorien Dummy-Variablen. Diese wurden mit den Regressionskoeffizienten der o.g. Autoren spezifisch für die *körperliche- und psychische Summenskala* gewichtet und summiert; abschließend wurde eine Konstante hinzugefügt (Bullinger & Kirchberger, 1998, S. 67). Das Procedere führte dazu, dass die empirisch ermittelten Skalen mit denen der amerikanischen Normstichprobe vergleichbar waren.

Die standardisierten Summenskalen sind laut Bullinger und Kirchberger in der Population normalverteilt und besitzen einen Mittelwert von 50 und eine Standardabweichung von 10, d.h. sie entsprechen der Skalierung einer T-Skala.

Höhere Skalenwerte entsprechen einer höheren Lebensqualität. Als Gesamtdurchschnitt ergab die *körperliche Summenskala (KSK)*  $45,68 \pm 9,33$  (Männer:  $45,98 \pm 9,03$ ; Frauen:  $45,49 \pm 9,56$ ). Für die *psychische Summenskala (PSK)* resultierte  $50,53 \pm 9,32$  (Männer:  $52,97 \pm 7,25$ ; Frauen:  $48,99 \pm 10,17$ ). Der Skalenwert der PSK entsprach damit ca. dem der Normstichprobe, während der Skalenwert der KSK unter den Vergleichsdaten lag.

Männer und Frauen unterschieden sich in der psychischen Summenskala ( $t_{(df=117)}=2,306$ ;  $p=0,023$ ), während sich kein Geschlechtsunterschied in der körperlichen Summenskala

ergab ( $t_{(df=117)}=0,276$ ;  $p=0,783$ ). Es wurde ein Schwellenwert von  $<40$  für die PSK festgelegt, um Probanden mit auffällig psychisch eingeschränkter Lebensqualität zu identifizieren, womit folgende Fallzahlen resultierten:

**Tab. 13:** Kreuztabelle "psychisch eingeschränkte Lebensqualität" und Geschlecht

			PSK <40		Gesamt
			LQ niedrig	LQ normal	
Geschlecht	Männlich	Anzahl	3	43	46
		% von Geschlecht	6,5%	93,5%	100,0%
		% von PSK <40	18,8%	41,7%	38,7%
	Weiblich	Anzahl	13	60	73
		% von Geschlecht	17,8%	82,2%	100,0%
		% von PSK <40	81,3%	58,3%	61,3%
Gesamt	Anzahl	16	103	119	
	% von Geschlecht	13,4%	86,6%	100,0%	
	% von PSK <40	100,0%	100,0%	100,0%	

Demnach litten 13% unter einer psychisch eingeschränkten Lebensqualität, also etwas weniger, als der maximal zu erwartende Anteil der Normalverteilung (15,9%) nahe legte. Die Auftretenshäufigkeit war zwischen den Geschlechtern (18% der Frauen, 7% der Männer) nicht signifikant unterschiedlich ( $\chi^2=3,089$ ;  $p=0,079$ ). Ein anderes Bild ergab sich für den Anteil der Probanden mit körperlich eingeschränkter Lebensqualität, welcher ebenso anhand des Kriteriums  $<40$  ermittelt wurde:

**Tab. 14:** Kreuztabelle "körperlich eingeschränkte Lebensqualität" und Geschlecht

			KSK <40		Gesamt
			LQ niedrig	LQ normal	
Geschlecht	Männlich	Anzahl	12	34	46
		% von Geschlecht	26,1%	73,9%	100,0%
		% von KSK <40	37,5%	39,1%	38,7%
	Weiblich	Anzahl	20	53	73
		% von Geschlecht	27,4%	72,6%	100,0%
		% von KSK <40	62,5%	60,9%	61,3%
Gesamt	Anzahl	32	87	119	
	% von Geschlecht	26,9%	73,1%	100,0%	
	% von KSK <40	100,0%	100,0%	100,0%	

Eine körperlich eingeschränkte Lebensqualität ergab sich für 27% der Probanden und lag damit weitaus höher, als anteilig für die Stichprobe zu erwarten (15,9%) war. Die geschlechtsspezifischen Anteile waren zwischen Frauen (27%) und Männern (26%) nicht verschieden ( $\chi^2=0,025$ ;  $p=0,875$ ). Aus der Diskrepanz der Fallprozente zwischen KSK und PSK kann abgeleitet werden, dass die Probanden sich stärker durch ihr körperliches statt psychisches Funktionsverständnis in ihrer Lebensqualität eingeschränkt fühlten.

#### 4.3.7. Visuelle Analogskala zur Einschätzung des Gesundheitsempfindens (VAS)

Die VAS wurde später in die Studie aufgenommen und konnte für N=76 Probanden angewandt und ausgewertet werden. Je höher die VAS (Skalierung: 0-100) ist, desto besser ist die selbst eingeschätzte Gesundheit. Für die Gesamtgruppe resultierte ein subjektiver Gesundheitszustand von  $72,61 \pm 17,31$ . Männer ( $74,18 \pm 15,85$ ) und Frauen ( $71,69 \pm 18,20$ ) schätzten ihre Gesundheit ähnlich ein ( $t_{(df=74)}=0,601$ ;  $p=0,549$ ).

Für Vergleichswerte wurde auf die Studie von Hinz zurückgegriffen (Hinz et al., 2006). Die Publikation berichtet geschlechtsspezifische Mittelwerte und Standardabweichungen von N=673 (41-60-jährigen) Personen. Die Autoren referieren für Männer einen VAS von  $78,8 \pm 15,2$  und für Frauen einen von  $76,2 \pm 16,9$ .

Probanden, die in unserer Studie mehr als eine Standardabweichung unter denen von Hinz berichteten entfernt lagen (Männer: VAS < 63,6; Frauen: VAS < 59,3), sind in folgender Tabelle als „niedrig“ identifiziert.

**Tab. 15:** Kreuztabelle "niedriges Gesundheitsgefühl" und Geschlecht

			VAS < 1 SD		Gesamt
			niedrig	o.B.	
Geschlecht	Männlich	Anzahl	8	20	28
		% von Geschlecht	28,6%	71,4%	100,0%
		% von VAS	40,0%	35,7%	36,8%
	Weiblich	Anzahl	12	36	48
		% von Geschlecht	25,0%	75,0%	100,0%
		% von VAS	60,0%	64,3%	63,2%
Gesamt	Anzahl	20	56	76	
	% von Geschlecht	26,3%	73,7%	100,0%	
	% von VAS	100,0%	100,0%	100,0%	

Für die Gesamtstichprobe resultierte ein Anteil von 26% der Fälle (29% der Männer, 25% der Frauen), die ihre Gesundheit auffällig niedrig einschätzten; zu erwarten waren 15,9%. Ein Geschlechtseffekt ließ sich zwischen den Fallzahlen nicht nachweisen ( $\chi^2=0,116$ ;  $p=0,733$ ).

#### 4.3.8. Psychosoziales Wohlbefinden (PSW)

Der Fragebogen kam für 119 Vpn zum Einsatz und zur Auswertung. Die von Riman & Udris postulierte unidimensionale Faktorstruktur (Rimann & Udris, 1993) konnte mit einer Faktorenanalyse bestätigt werden. Es wurde ein Faktor (Kriterium: Eigenwerte >1) extrahiert, der 48% der Varianz aufklärte. Die Mittelwerte der Items stellten sich wie folgt dar (Abb. 16):

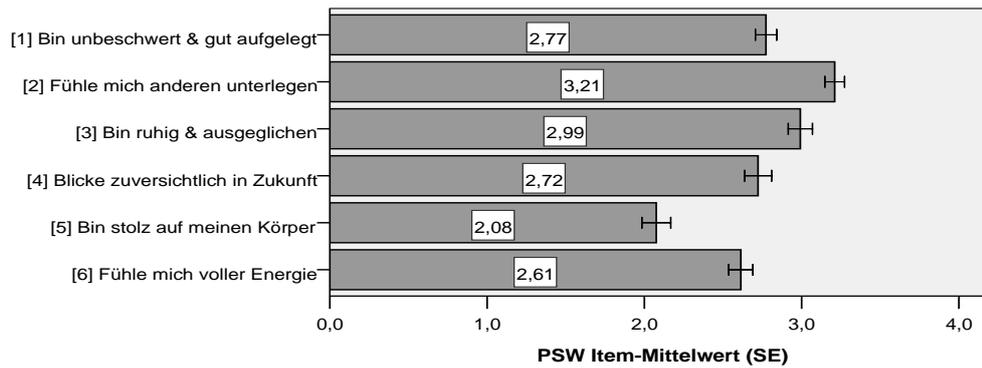


Abb. 16: Ergebnisse des PSW auf Itemebene

Das negativ formulierte Item 2 wurde umgepolt. Am deutlichsten wich das Item 5 („Ich bin stolz auf meinen Körper“) von den anderen ab. Insgesamt ergab sich ein Itemdurchschnitt von  $2,73 \pm 0,58$ . Frauen ( $2,64 \pm 0,57$ ) wiesen demnach ein niedrigeres psychosoziales Wohlbefinden als Männer ( $2,88 \pm 0,58$ ) auf ( $t_{(df=117)}=2,266$ ;  $p=0,025$ ). Durch die fehlende Geschlechtsadjustierung ist nicht klar, ob es sich um einen normalen oder einen relevanten Geschlechtsunterschied handelt. Da für den Fragebogen keine Vergleichswerte vorliegen und somit keine alters- und geschlechtsspezifische Transformation in Standardnormen erfolgen konnte, ließen sich keine Fälle von besonders niedrigem Wohlbefinden im Vergleich zur Allgemeinbevölkerung identifizieren.

#### 4.3.9. Zusammenfassung des Interventionsbedarfs auf der Ebene der psychischen Gesundheit

Zusammenfassend für die Stichprobe sind in der folgenden Grafik die Anteile der Vpn wiedergegeben, die für das jeweilige Testverfahren als „problematisch“ identifiziert wurden (z.B. unterdurchschnittliches Kohärenzgefühl oder überdurchschnittliche Depressivität).

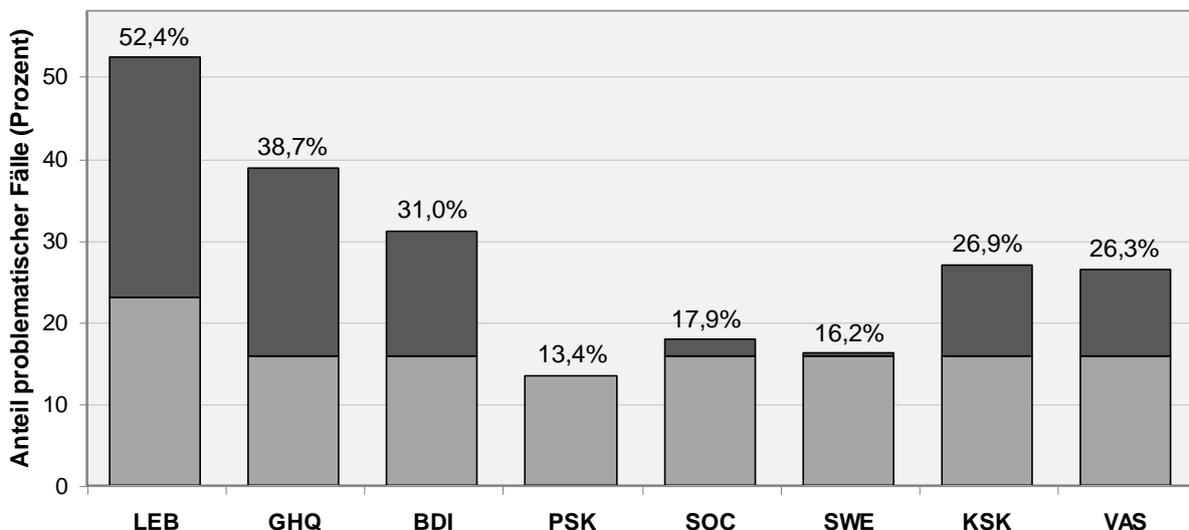


Abb. 17: Als problematisch auffällig identifizierte Anteile der Stichprobe in Prozent für normtransformierte Messverfahren. Die erwarteten Anteile sind hellgrau markiert, die darüber hinaus sind dunkelgrau gefärbt.

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 unterdurchschnittlichem Bereich maximal erwartet werden darf; dies sind 15,9%. Dieser Anteil ist in den Säulen für die Messverfahren GHQ bis VAS hellgrau markiert. Alle Anteile über 15,9% sind als auffällig zu werten und wurden in den Säulen dunkelgrau gefärbt. Diese von der entsprechenden Allgemeinbevölkerung abweichenden dunkelgrauen Bereiche stellen somit den Interventionsbedarf dar. Der Anteil des Merkmals "Lebenszufriedenheit" (LEB) wurde anhand von Stanine-Normen berechnet. In Stanine entspricht 23% dem Anteil, der laut alters- und geschlechtsspezifischer Normierung für das Merkmal im einseitigen Bereich (hier: unterdurchschnittliche Lebenszufriedenheit) erwartet werden darf.

Es ist darauf hinzuweisen, dass der eigentliche Interventionsbedarf vermutlich konservativ erhoben wurde. In den jeweiligen bevölkerungsrepräsentativen Stichproben sind neben Berufstätigen natürlich auch Erwerbslose und weitere Angehörige der sog. "out of labour force" (Rengers, 2005) erfasst (z.B. Hausfrauen/-männer, Frührentner, krankheitsbedingt berufsunfähige Menschen). Für eine realistischere Einschätzung des Interventionsbedarfs hätte die Stichprobe ausschließlich mit erwerbstätigen 50-60-jährigen Menschen verglichen werden dürfen. Derart differenzierte Vergleichsdaten waren für die Messverfahren jedoch nicht verfügbar.

Im Vergleich mit der Allgemeinbevölkerung indizierten LEB, GHQ und BDI den deutlichsten Interventionsbedarf. Deren hohe Anteile problematischer Fälle reflektieren sehr deutliche psychische Beschwerden und damit hohe Risikofaktoren der psychischen Gesundheit für die Stichprobe. In der Höhe gefolgt werden diese von der KSK und der VAS, die ein stark unterdurchschnittliches Gesundheitsempfinden sowie eine stark körperlich eingeschränkte Lebensqualität signalisierten. Bemerkenswert ist, dass auffällig eingeschränkte Ressourcen der psychischen Gesundheit (SOC und SWE) kaum Abweichungen zum erwarteten Anteil der Allgemeinbevölkerung zeigten. In Relation zu den anderen Konzepten können sowohl die Selbstwirksamkeitserwartung als auch der Kohärenzsinn als einigermaßen unauffällig bezeichnet werden, so dass hier nicht von einem deutlichen Interventionsbedarf auszugehen war. Die PSK, welche eine psychisch eingeschränkte Lebensqualität reflektiert und inhaltlich (wie LEB, GHQ und BDI) als Risikofaktor der psychischen Gesundheit definiert wurde, wich als einzige Skala nicht problematisch von ihren Normwerten ab.

Für sämtliche Messverfahren ergab sich ein durchschnittlicher Anteil von 28% problematischer Fälle. In Anlehnung an die hier definierten Konzepte zur Erfassung der Gesundheit betrug für die als erhöhte Beschwerden (Risikofaktoren der psychischen Gesundheit) definierten Merkmale der mittlere Anteil 34%. Für die als unterdurchschnittliche Ressourcen definierten Skalen 17% und für die als körperliche Einschränkungen der Gesundheit definierten Merkmale 27%.

#### 4.3.10. Antworttendenz-Hypothesentest

##### Übertreibung des Interventionsbedarfs durch Antworttendenzen?

Für die Versuchspersonen erfolgte vor jeder Untersuchung eine schriftliche Aufklärung darüber, dass ihre Angaben dem Datenschutz unterliegen, die Daten in anonymisierter Form gespeichert und lediglich in Form von Gruppenmittelwerten veröffentlicht werden. Auch mündlich wurde ihnen verdeutlicht, dass sie durch ihre Beantwortung weder Vorteile erwarten können noch, dass ihnen Nachteile bei wahrheitsgemäßer Beantwortung drohen. Bei Fragebogendaten können sich je nach Stichprobe verschiedene Biases ergeben. Problematisch für die Interpretation von psychologischen und medizinischen Daten sind Aggravations- bzw. Simulationstendenzen, d.h. wenn Probanden mit weniger wahrheitsgemäßem Antwortstil eher zur Übertreibung ihrer Beschwerdesymptomatik neigen. Demgegenüber bedeutet eine sozial erwünschte Antworttendenz, dass Probanden sich an gesellschaftlichen Werten orientieren und ihre Angaben dahingehend modifizieren, die eigene Persönlichkeit in ein positives Licht zu rücken. Ein gewisser sozial erwünschter Bias ist völlig normal. Diese Antworttendenz neigt aber zur verkleinerten Darstellung von (psychischen) Problemen und muss deshalb als Gegenphänomen der Aggravationstendenz aufgefasst werden. Somit wäre beim Nachweis des Wirkens eines sozial erwünschten Bias eine Aggravationstendenz falsifiziert.

Das FPI Merkmal "Offenheit" stellt eine alters- und geschlechtsnormierte 9-stufige Variable zur Einschätzung des Wahrheitsgehalts von Antworten bereit. Je höher ihr Wert, desto offener bzw. wahrhaftiger die Antworttendenz. Zwischen "Offenheit" und den Messinstrumenten zur Erfassung der Gesundheit wurden Korrelationen zur Aufdeckung von Zusammenhängen berechnet.

**Tab. 16:** Korrelative Zusammenhänge der Messinstrumente mit dem FPI Merkmal "Offenheit"

	LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV <sup>9</sup>
N	84	84	82	84	82	82	84	84	76	84
r	-0,035	-0,257	0,265	-0,255	-0,282	-0,265	-0,084	-0,062	-0,220	-0,032
p	0,754	<b>0,018</b>	<b>0,027</b>	<b>0,019</b>	<b>0,010</b>	<b>0,016</b>	0,449	0,578	0,085	0,789

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. Umgekehrt betrachtet, lautet die inhaltliche Relation, dass eine verschlossene Antworttendenz zur Verkleinerung von potentiellen Problemen im Zusammenhang mit der psychischen Gesundheit führte. Diese signifikanten Zusammenhänge sind als Wirkung eines sozial erwünschten Bias zu interpretieren. Dass die Variable "Offenheit" durchschnittlich einen Wert von  $4,3 \pm 1,80$  aufwies (Normstichprobe  $5 \pm 1,96$ ) und demnach die Probanden eher verschlossen als offen antworteten, ist als Hinweis dafür zu deuten, dass die Antworttendenz bei der Ermittlung des Interventionsbedarfs eher zu einer Unter- statt Überschätzung geführt haben sollte.

<sup>9</sup> Rangkorrelationskoeffizient Spearman-Rho

Dass der Interventionsbedarf zur Gesundheit anhand „unwahrscheinlicher“ Antworttendenzen der Probanden übertrieben ermittelt sein könnte, gilt durchschnittlich als falsifiziert. In 5 von 9 Indikatoren zur Gesundheit zeigte sich der umgekehrte Zusammenhang. Je weniger „wahrscheinlich“ die Antworttendenz, desto eher neigten die Probanden dazu, psychische Probleme abgemildert darzustellen.

#### 4.4. Fernseh- und Medienkonsum

Ein selbst erstellter Fragebogen zur Mediennutzung wurde später in die Studie aufgenommen und konnte für N=84 Probanden angewandt und ausgewertet werden. Es ergaben sich folgende Häufigkeiten der durchschnittlichen, täglichen zugebrachten Zeit in folgenden Mediennutzungskategorien:

Tab. 17:: Täglicher zeitlicher Konsum für verschiedene Medien

Täglicher Konsum	Fernsehen		Internet		Computerspiele	
	N	Prozent	N	Prozent	N	Prozent
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%

Für die Altersgruppe der Probanden war zu erwarten, dass das Fernsehen das Hauptmedium darstellen würde. Niemand gab an, kein Fernsehen zu schauen. Dieses Medium wurde zu 100% von der Stichprobe genutzt. 30% der Probanden „surften“ im Internet und immerhin noch 14% der Vpn verbrachten einen Teil ihrer Freizeit mit Computerspielen. Nach der Überführung der Zeitkategorien in numerische Werte (z.B. „1-2h.“=1,5h. etc.) ergaben sich für die Medienkategorien folgende Mittelwerte (Abb.18.):

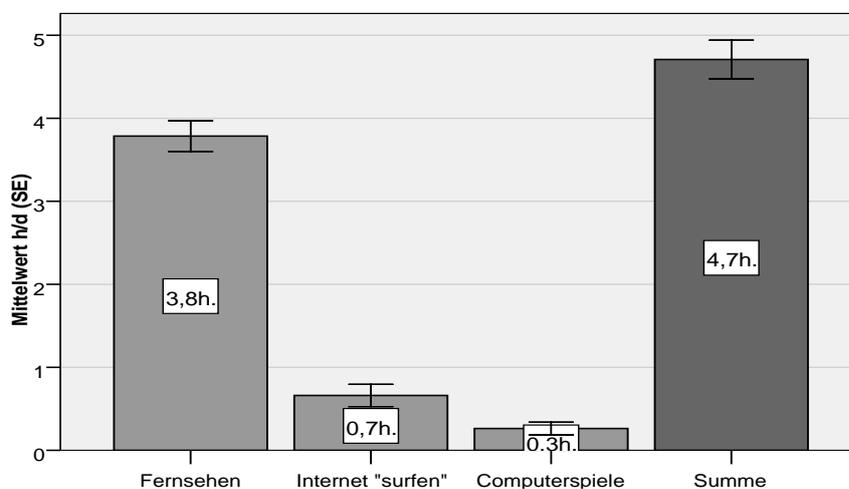


Abb. 18: Durchschnittliche tägliche Nutzungsdauer für verschiedene Medien

Demnach betrug die durchschnittliche gesamte Mediennutzungsdauer für die Stichprobe  $4,7 \pm 2,1$  Stunden täglich. Diese setzte sich aus den gemittelten täglichen 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. Die Daten waren unimodal, aber rechtsschief verteilt. Der Mann-Whitney U-Test zeigte keine Geschlechtseffekte an, weder für die Gesamtmediennutzung ( $Z=0,005$ ;  $p=0,996$ ) noch für Fernsehen ( $Z=0,778$ ;  $p=0,437$ ) noch für Internet ( $Z=1,007$ ;  $p=0,314$ ) oder Computerspiele ( $Z=0,508$ ;  $p=0,611$ ). Die "Arbeitsgemeinschaft Fernsehforschung" (AGF, 2010) berichtet für die Jahre 2007, 2008, 2009 für den durchschnittlichen Bundesbürger einen täglichen Fernsehkonsum von 3,48h. Während die AGF registriert, zu welchen Zeiten ein Fernsehgerät auf Empfang war, gaben die Studienteilnehmer an, wie viel Fernsehen sie tatsächlich täglich konsumierten.

#### 4.5. Zusammenhang der Messinstrumente zur psychischen und psychosozialen Gesundheit/ zum Medienkonsum

Explorativ wurden Korrelationen für die Messinstrumente durchgeführt, um einen Eindruck für die Zusammenhänge der Daten zum ersten Messzeitpunkt zu erhalten.

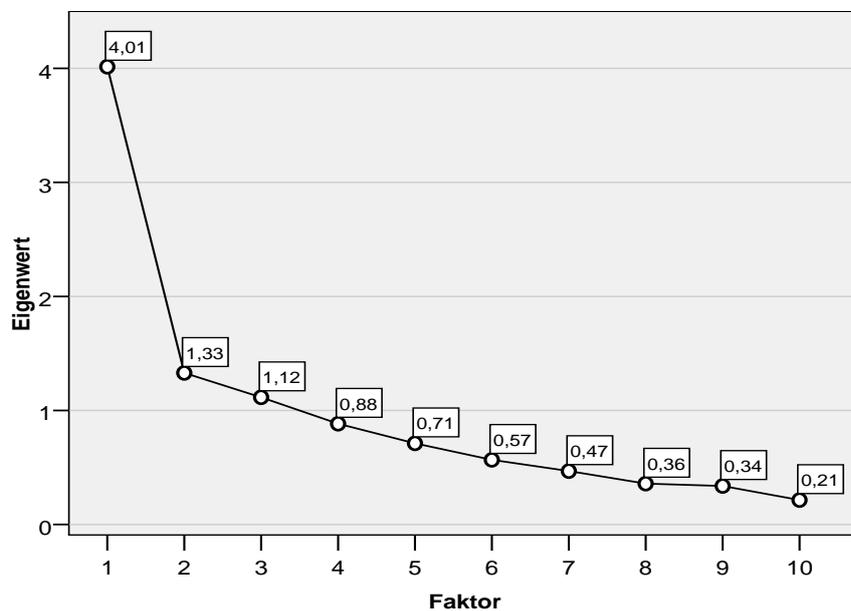
**Tab. 18:** Korrelative Zusammenhänge der Messinstrumente zur Erfassung der psychischen Gesundheit, der subjektiv körperlichen Gesundheit und des TV-Konsums. (\*  $p < 0,05$ ; \*\*  $p < 0,01$ )

	LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV <sup>10</sup>
LEB	<i>r</i>	0,327**	-0,267*	0,119	0,370**	0,426**	0,390**	0,259*	0,163	-0,073
	<i>N</i>	84	70	84	82	82	84	84	62	68
GHQ	<i>r</i>	0,327**	-0,633**	0,538**	0,627**	0,404**	0,585**	0,216*	0,377**	0,012
	<i>N</i>	84	87	119	117	117	119	119	76	84
BDI	<i>r</i>	-0,267*	-0,633**	-0,528**	-0,812**	-0,478**	-0,632**	-0,236*	-0,277*	-0,084
	<i>N</i>	70	87	87	87	87	87	87	74	82
PSK	<i>r</i>	0,119	0,538**	-0,528**		0,559**	0,406**	-0,116	0,299**	0,023
	<i>N</i>	84	119	87		117	119	119	76	84
SOC	<i>r</i>	0,370**	0,627**	-0,812**			0,642**	0,263**	0,446**	0,208
	<i>N</i>	82	117	87			117	117	74	82
SWE	<i>r</i>	0,426**	0,404**	-0,478**	0,288**		0,524**	0,188*	0,251*	0,097
	<i>N</i>	82	117	87	117		117	117	74	82
PSW	<i>r</i>	0,390**	0,585**	-0,632**	0,406**	0,642**		0,270**	0,181	-0,046
	<i>N</i>	84	119	87	119	117		119	76	84
KSK	<i>r</i>	0,259*	0,216*	-0,236*	-0,116	0,263**	0,188*		0,463**	0,097
	<i>N</i>	84	119	87	119	117	119		76	84
VAS	<i>r</i>	0,163	0,377**	-0,277*	0,299**	0,446**	0,251*	0,181		0,223
	<i>N</i>	62	76	74	76	74	74	76		76
TV <sup>12</sup>	<i>Rho</i>	-0,073	0,012	-0,084	0,023	0,208	0,097	-0,046	0,097	
	<i>N</i>	68	84	82	84	82	82	84	84	

Die Variablen der Erstuntersuchung hingen oftmals mit mittlerer Stärke zusammen. Dies signalisiert, dass sie einerseits Varianz teilen und andererseits recht unterschiedliche Konstrukte messen. Zwischen BDI und SOC war die Korrelation ( $r=0,81$ ) am stärksten, obwohl deren Normwerte stark unterschiedliche Anteile auffälliger Fälle indizierten. Am

<sup>10</sup> Für die Zusammenhänge mit „TV“ wurden Rangkorrelationskoeffizienten nach Spearman (Rho) berechnet.

wenigsten hängen die körperbezogenen Skalen KSK und VAS mit den Indikatoren der psychischen Gesundheit zusammen. Dies ist methodisch positiv zu sehen, weil es signalisiert, dass die Skalen zur körperlichen Gesundheit tatsächlich etwas substantiell anderes erfassen als die Variablen zur Psyche. Da TV-Konsum zwar unipolar, aber rechtsschief verteilt war, wurden Spearman-Rangkorrelationen mit den restlichen Variablen durchgeführt. Zum Fernsehverhalten ergaben sich keine signifikanten Korrelationen. Dies überrascht, da Studien, gerade mit älteren Menschen, teils Zusammenhänge zwischen erhöhtem TV-Konsum und negativem psychischen Befinden nachweisen konnten (Depp et al., 2010). Die Variablen wurden in eine Analyse zur Identifizierung von orthogonalen Hauptkomponenten eingegeben. Es ergaben sich laut Scree-Kriterium zwei unabhängige Faktoren.

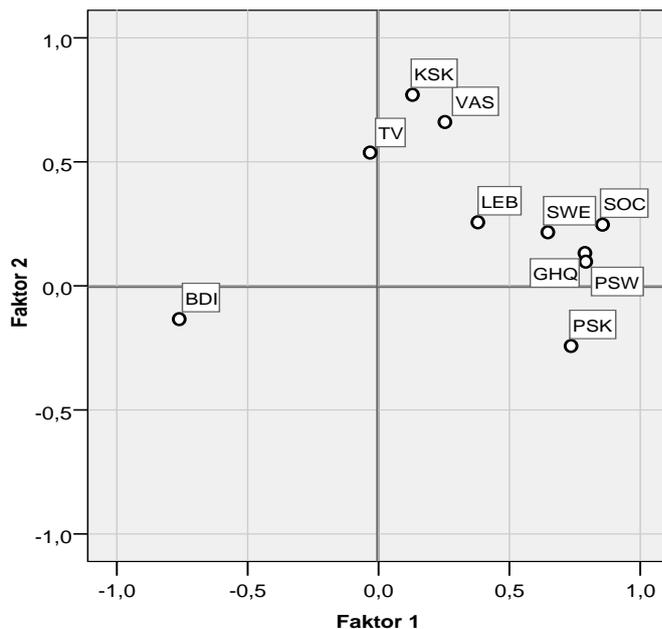


**Tab. 19:** Rotierte Faktormatrix

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

**Abb. 19:** Scree-Plot der Faktorenanalyse

Der erste Faktorklärte 37,5% der Varianz auf, der zweite 17,0%. Zusammen entschlüsselten sie 54,5% der Gesamtvarianz. Wie anhand der rotierten Faktormatrix ersichtlich (Abb. 20), liegen die Variablen SOC, PSW, GHQ, BDI, PSK, SWE und LEB vorwiegend auf dem ersten Faktor, während KSK, VAS und TV-Konsum auf dem zweiten Faktor abgebildet sind. Inhaltlich ist zu interpretieren, dass der erste Faktor die psychische Gesundheit repräsentiert, während der zweite Faktor am stärksten die körperliche Gesundheit beschreibt.



**Abb. 20:** Darstellung der Komponenten im rotierten Raum

Die orthogonalen Faktoren unterstreichen den prinzipiellen Unterschied zwischen der Messung der körperbezogenen und der psychischen Gesundheit. Im Modell erweist sich der TV-Konsum verwandter mit den körperlichen Variablen als mit den psychischen. Dies gibt einen Hinweis darauf, dass für die Stichprobe der Fernsehkonsum eher mit dem körperlichen Freizeitverhalten zusammenhängt, als Relationen mit der psychischen Gesundheit einzugehen (z.B. TV-Konsum als Copingverhalten zur affektiven Regulation). Es bildeten sich keine separaten Faktoren, die einerseits die Risikofaktoren und andererseits die Ressourcen der psychischen Gesundheit abbilden; allerdings wäre es unlogisch gewesen, zwischen diesen eine (orthogonale) Unabhängigkeit vorauszusetzen.

#### 4.6. Test auf Gleichheit der Kohorten

Für die Analyse der Längsschnitthypothesen zur differenzierten Betrachtung des Kohorten-Interventionseffektes war die Überprüfung einer Zufallsverteilung in den abhängigen Variablen zwischen den Interventionsgruppen von Interesse, weil die Teilnehmer diesen nicht randomisiert zugeteilt wurden sondern per Gelegenheit Eingang in diese fanden.

##### 4.7.1. Ebene der physischen Gesundheit

**Tab. 20:** Tests der Gruppenunterschiede des Zwischensubjektfaktors „Kohorte“ bei T1

(RR<sub>syst</sub> = Systolischer Blutdruck; RR<sub>diast</sub>= Diastolischer Blutdruck; HSF= Herzschlagfrequenz; Chol= Gesamtcholesterin; HDL=HDL-Cholesterin; LDL= LDL-Cholesterin; Triglyc= Triglyzeride; HbA1c= Glykohämoglobin)

	RR <sub>syst</sub>	RR <sub>diast</sub>	HSF	Chol	HDL	LDL	Triglyc	HbA1c
df	116	116	116	114	114	115	113	117
<i>t</i>	1,426	0,061	0,701	-1,533	-0,880	-0,066	-0,784	0,380
<i>p</i>	0,157	0,952	0,485	0,128	0,381	0,948	0,435	0,705

**Tab. 21:** Tests der Gruppenunterschiede des Zwischensubjektfaktors „Kohorte“ bei T1 (Leistung= Ergometerleistung; VO<sub>2</sub>max= maximale Sauerstoffaufnahme bei Ausbelastung; relVO<sub>2</sub>max= gewichtsadaptierte maximale Sauerstoffaufnahme; Ventil= maximale Ventilation; Größe= Körperhöhe; Gewicht; BMI= Body Mass Index; KF% = prozentualer Körperfettanteil; KF= Gesamtmasse an Körperfett)

	Leistung	VO <sub>2</sub> max	relVO <sub>2</sub> max	Ventil	Größe	Gewicht	BMI	KF%	KF
df	115	110	110	110	117	117	117	117	117
t	1,324	0,188	-0,616	-0,215	0,026	1,213	1,562	1,669	1,487
p	0,188	0,851	0,539	0,830	0,979	0,228	0,121	0,098	0,140

Den Tabellen 20 und 21 ist zu entnehmen, dass die Probanden auf der körperlichen Ebene in den abhängigen Variablen keine signifikanten Unterschiede zwischen den Kohorten aufwiesen. Die Annahmen der Nullhypothesen konnten stets beibehalten werden.

#### 4.7.2. Ebene der psychischen und psychosozialen Ebene/ Medienkonsum

**Tab. 22:** Tests der Gruppenunterschiede des Zwischensubjektfaktors "Kohorte"

	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum <sup>11</sup>
df	117	85	117	115	115	117	117	74	N=84
t	-0,255	0,092	-1,637	-1,220	0,726	-1,449	1,212	0,643	Z=1,068
p	0,799	0,927	0,104	0,225	0,469	0,150	0,228	0,522	0,286

Die Probanden zeigten in den abhängigen Variablen keine signifikanten Unterschiede zwischen den Kohorten. Die Annahmen der Nullhypothesen konnten stets beibehalten werden.

### 4.7. Drop-Out-Analyse

Longitudinalen Forschungsdesigns ist immanent, dass nicht alle Probanden an Wiederholungsmessungen teilnehmen. In der vorliegenden Studie konnten die Versuchspersonen z.B. in Arbeit kommen, erkranken oder aus persönlichen Gründen nicht mehr an der Studie teilnehmen.

#### 4.7.1. Teilnehmer- und Drop-Out-Quote T2

- An T2 nahmen 94 der 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 sind in Arbeit gekommen (5% der Stichprobe),
- 8 gaben an, erkrankt zu sein (6,7% der Stichprobe),
- 11 hatten das Interesse an der weiteren Teilnahme verloren (9,3% der Stichprobe).

<sup>11</sup> U-Test nach Mann-Whitney

#### 4.7.2. Teilnehmer- und Drop-Out-Quote T3

- An T3 nahmen 60 Probanden teil (50% der Stichprobe, 67% der Männer, 40% der Frauen). Weitere 14 Probanden willigten ein, trotz physischer Nichtteilnahme an der Untersuchung in Rostock, postalisch zugeschickte Fragebögen zu beantworten. Insofern liegen für die Fragebogeninstrumente Daten von maximal 74 Probanden vor (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 nahmen eine Arbeit auf (6,7% der Stichprobe),
  - 13 gaben an, erkrankt zu sein (10,9 der Stichprobe%),
  - 23 hatten das Interesse an der Teilnahme verloren (19,3 der Stichprobe%).

Die Drop-Out-Gründe der 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 meist auch aus denen von T2.

Folgendes ist zu den Abbruch-Ursachen anzumerken:

- Arbeitsaufnahme: Insgesamt nahmen 17 Vpn während des gesamten Studienzeitraums eine Erwerbsarbeit auf (14% der Stichprobe). Von diesen 17 nahmen jedoch 7 an sämtlichen Untersuchungen teil und 5 weitere an der Wiederholungsmessung T2. Die Aufnahme von Erwerbsarbeit machte demnach einen Studienabbruch zwar wahrscheinlicher, stellte jedoch keine hinreichende Bedingung für eine Nichtteilnahme dar.
- Erkrankung: Ob eine Versuchsperson erkrankt war, wurde für die Drop-Out-Statistik nur dann erhoben, wenn der Proband den Wiederholungsmessungen fernblieb. Der Abbruchgrund „Erkrankung“ fußt auf der mündlichen Selbstauskunft des Probanden. Medizinische Belege für eine Erkrankung wurden nicht angefordert und waren auch nicht forderbar, da die weitere Studienteilnahme selbstverständlich auf Freiwilligkeit (Ethik) beruhte. Zusätzlich wurde eine mögliche Erkrankung in der Drop-Out-Analyse nicht ausgewertet, wenn die Vpn trotzdem weiter an der Studie (oder dem Sport) teilnahm. Der Hauptbefund der Salutogeneseforschung (Antonovsky, 1987; Antonovsky & Franke, 1997), dass die Psyche einen erheblichen moderierenden Einfluss darauf hat, ob sich Menschen (bei gleicher Symptomatik) noch gesund oder bereits krank fühlen, wird als weiteres Argument dafür herangezogen, dass in dieser Studie die mündliche Selbstauskunft einer Erkrankung nicht als hinreichender Abbruchgrund gewertet werden sollte.
- Verlorenes Interesse: Die Gruppe für diesen Abbruchgrund setzt sich aus den Probanden zusammen, die ein mangelndes Interesse ehrlich zugaben und denen, die sich trotz mehrfacher Kontaktaufnahme nicht auf Termine für die Wiederholungsmessung festlegen bzw. Untersuchungstermine unangemessen oft verschieben wollten. Auch Vpn, die mehrfach die vereinbarten Untersuchungstermine sehr kurzfristig absagten bzw. einfach nicht erschienen, wurden in dieser Gruppe zusammengefasst. Hier wird ein „verlorenes Interesse“ der Einfachheit halber vorausgesetzt, obwohl möglicherweise relevante private Gründe vorlagen, die sich jedoch nicht reliabel erheben ließen.

Da eine Arbeitsaufnahme sich heterogen auf die weitere Studienteilnahme auswirkte und dies wahrscheinlich auch für den Abbruchgrund „Erkrankung“ zutraf, sollen die folgenden Analysen nicht differenziert für diese Abbruchursachen ausgewertet werden.

#### 4.7.3. Drop-Out-Analyse T1-T2

Zur Generalisierung der Studienergebnisse ist eine Aufklärung darüber wichtig, ob in Bezug auf die abhängigen Variablen selektiv bestimmte Probanden den Wiederholungsmessungen fernblieben. Es ließe sich z.B. spekulieren, dass eine erhöhte Depressivität zu einer Antriebsschwäche führt, welche die weitere Teilnahmewahrscheinlichkeit an der Studie senkt. Zu diesem Zweck wurde getestet, ob Probanden, die an der Wiederholungsmessung T2 teilnahmen, sich von den Nichtteilnehmern bereits aufgrund der Ergebnisse der Eingangsuntersuchung unterschieden. Im Folgenden finden sich die statistischen Tests für die abhängigen Variablen der Erstuntersuchung.

**Tab. 23:** Tests der Gruppenunterschiede des Zwischensubjektfaktors "T2-Teilnahme"

	LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum <sup>12</sup>
df	82	117	85	117	115	115	117	117	74	N=84
<i>t</i>	-1,629	-0,433	0,107	-1,053	-0,137	-0,151	-1,208	-1,918	0,361	Z=-0,949
<i>p</i>	0,107	0,666	0,915	0,295	0,891	0,881	0,230	0,058	0,719	0,343

Die Nullhypothese konnte für alle Gruppenunterschiede beibehalten werden. Es fand sich lediglich eine tendenziell niedrigere körperliche Lebensqualität (KSK) der Nichtteilnehmer (42,54±10,18) gegenüber den Teilnehmern (46,52±8,96). Signifikante Anzeichen für a priori unterschiedliche Ausprägungen konnten jedoch nicht ermittelt werden.

#### 4.7.4. Drop-Out-Analyse T1-T3

Im Folgenden finden sich die Tests der abhängigen Variablen zwischen den 74 Probanden, die an allen Untersuchungen teilnahmen (inklusive derer, die für T3 die Fragebögen postalisch zur Verfügung stellten) im Vergleich mit den 44 restlichen<sup>13</sup> Probanden (Teilnahme ausschließlich an T1 oder an T1&T2).

**Tab. 24:** Tests der Gruppenunterschiede des Zwischensubjektfaktors "T3-Teilnahme"

	LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum <sup>14</sup>
df	82	116	85	116	114	114	116	116	74	N=84
<i>t</i>	-1,606	-0,817	-0,621	-1,367	-0,122	-0,975	-2,528	-0,703	1,477	Z=-2,649
<i>p</i>	0,112	0,415	0,536	0,174	0,903	0,332	<b>0,013</b>	0,483	0,144	<b>0,008</b>

Es wurde ein signifikanter Unterschied für ein schwächeres psychosoziales Wohlbefinden (PSW) der Nichtteilnehmer (2,55±0,54) gegenüber den Studienkomplettierern (2,82±0,58) sichtbar. Fernsehkonsum war mit +53 Minuten hochsignifikant für die Nichtteilnehmer

<sup>12</sup> U-Test nach Mann-Whitney

<sup>13</sup> Ein Versuchsteilnehmer, der zwischen T2 und T3 verstorben war, wurde nicht in die Analysen einbezogen

<sup>14</sup> U-Test nach Mann-Whitney

(263min±90min) gegenüber den Studienkomplettierern (210min±104min) erhöht. Das psychosoziale Wohlbefinden und der Fernsehkonsum spielten demnach eine Rolle, ob Probanden bis zum Schluss an der Studie teilnahmen.

#### **4.8. Sportfortsetzeranalyse**

Das Ziel der Intervention war die Förderung der Gesundheit der Teilnehmer. Da die Gesundheit primär mit dem ca. 10-wöchigen Ausdauertraining gefördert wurde, wurde bereits als ein Interventionserfolg definiert, wenn die Teilnehmer ihre sportliche Aktivität im follow-up Zeitraum fortsetzten. Sie hatten die Gelegenheit, im selben Fitness-Studio, in dem zuvor auch die Intervention stattfand kostenlos ein eigenständiges Training fortzusetzen. Das Angebot war von daher als niederschwellig zu bezeichnen.

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 abgebrochen hatten (64% der Stichprobe, 59% der Männer, 67% der Frauen), kommen dieselben Ursachen infrage, die bereits zu den Gründen des Studienabbruchs diskutiert wurden (vgl. Drop-Out-Analyse, Kap. 4.7.2.): Ein Proband war verstorben, 12 Vpn gaben an, in Arbeit gekommen zu sein, 14 gaben an, erkrankt gewesen zu sein und 49 Probanden hatten vermutlich kein Interesse, den Sport fortzusetzen. Umgekehrt betrachtet hatten von den 43 Sportfortsetzern 5 eine Arbeit gefunden und somit trotz ihrer neuen Anstellung weiter trainiert. Ebenso ist es möglich, dass Probanden trotz einer Erkrankung weiter trainierten. Beide Gründe stellen durchaus Wahrscheinlichkeitserhöhungen, aber keine hinreichenden Ursachen dafür dar, den Sport im 6-monatigen follow-up gänzlich abzuberechen. Insofern soll auch in dieser Analyse eine differenzierte Betrachtung der Abbruchgründe "Arbeit" und "Erkrankung" unberücksichtigt bleiben.

Im Folgenden wurden Analysen zur Beantwortung der Fragestellung durchgeführt, ob sich mit einer bestimmten Konfiguration an Daten retrospektiv das Risiko "vorhersagen" lässt, dass ein begonnener Sport nicht von den Studienteilnehmern aufrecht erhalten wurde. Hierzu erfolgten Berechnungen von Prädiktionsmodellen auf der Basis von logistischen Regressionen sowie Odds- und Odds-Ratio Kalkulationen.

##### **4.8.1. T1 Modell**

###### **Sportfortsetzeranalyse aufgrund der Daten der Erstuntersuchung**

Hier lautete die Fragestellung, 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; also analog zum Verfahren für die Drop-Out-Analysen. Zum Auffinden von potentiell geeigneten Prädiktoren für das Prädiktionsmodell wurden für die abhängigen Variablen zwischen den Gruppen „Sportfortsetzer“ und „Sportabbrecher“ Unterschiedstests berechnet<sup>15</sup>.

---

<sup>15</sup> Die Prüfungen erfolgten unter Ausschluss des nach T2 verstorbenen Versuchsteilnehmers

**Tab. 25:** Tests der Gruppenunterschiede des Zwischensubjektfaktors "Sportfortsetzung" für T1-Variablen

		LEB	GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum <sup>16</sup>
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
	<i>t</i>	-1,329	-0,996	-0,214	-1,023	-0,463	-1,227	-1,983	-1,598	-0,326	Z=-2,911
	<i>p</i>	0,187	0,321	0,831	0,308	0,644	0,222	0,050	0,113	0,746	<b>0,004</b>

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. Weitere Analysen, auch bezüglich der Persönlichkeitsfaktoren (hier nicht dargestellt), erbrachten keine Unterschiede.

Auf Itemebene stimmten die Sportfortsetzer im Mann-Whitney-U-Test folgenden Items signifikant stärker zu:

- 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 Sportabbrecher für:

- SF12 Item 12: „Meine sozialen Kontakte sind aufgrund meiner seelischen Probleme eingeschränkt“ (Z=1,810 p=0,070).

Um zu entscheiden, welche Daten den größten Einfluss darauf hatten, ob ein Proband, der zu T1 befragt wurde, später dem Risiko ausgesetzt war, das Training nicht fortzusetzen, wurde eine schrittweise binäre logistische Regressionsanalyse berechnet, in welche die trennscharfen (Kriterium:  $p < 0,1$ ) Messverfahren und Items als Prädiktoren für den Regressand "Sportfortsetzung" eingesetzt wurden. Vollständige Daten ergaben sich zwischen allen Prädiktoren für 82 Versuchspersonen<sup>17</sup>.

**Tab. 26:** Ergebnisse der binären logistischen Regressionsanalyse für die Prädiktoren aus T1

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

<sup>16</sup> Mann-Whitney-U-Test

<sup>17</sup> Hauptsächliche N-limitierende Variable zwischen den Prädiktoren stellte "TV-Konsum" dar.

In den Tests 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 für die Regression. 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 (-2LL Wert) benutzt. Dieser betrug 96,11 und war mit Hinzunahme der signifikanten Prädiktoren kleiner als der Ursprungswert, welcher nur die Konstante enthielt. Dieser Unterschied war hochsignifikant mit  $ChiSq=12,64$   $p=0,002$ ; Nagelkerkes  $R^2$  betrug 0,194.

#### 4.8.2. Prädiktionsmodell, Odds und Odds-Ratios

Mit den Regressionskoeffizienten wurde für jede Versuchsperson die Wahrscheinlichkeit berechnet, mit der das Ereignis "Sport abbrechen" 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 deren Konstante;  $e$  repräsentiert die Eulersche Zahl<sup>18</sup>. Ergibt sich für  $p$  ein Wert  $<0,5$ , nimmt man an, dass das Ereignis nicht eintritt. Im anderen Fall nimmt man das Eintreffen des Ereignisses an. Im Prädiktionsmodell resultierte für den Exponenten  $z$  nach Einsetzen der Koeffizienten

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

Inhaltlich bedeutet die Beziehung: Je höher der Fernsehkonsum und je weniger die Aussage "Ich fühle mich selten als Verlierer" abgelehnt wurde, desto größer der Exponent  $z$ , welches nach Einsatz in die Formel zu einer gesteigerten 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, resultierte

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

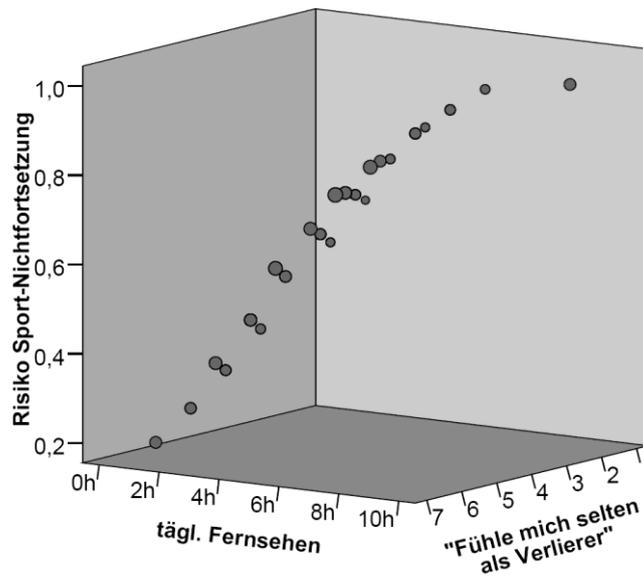
und hiermit

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

Da im Beispiel die Wahrscheinlichkeit des Sportabbruchs 77% beträgt, wäre das Eintreffen des Ereignisses angenommen. Der Zusammenhang der Prädiktoren mit dem Risiko ist in folgender Abbildung dargestellt.

---

<sup>18</sup>  $e=2,718281828459\dots$



**Abb. 21:** Zusammenhang des berechneten Risikos mit den Prädiktoren

Im Folgenden findet sich eine Klassifizierungstabelle, in der für die 82 Fälle die beobachtete Gruppenzugehörigkeit der aufgrund des berechneten Modells vorhergesagten gegenübergestellt wird:

**Tab.26:** Beobachtete und vorhergesagte Häufigkeiten der Sportfortsetzung und des -abbruchs aufgrund des T1-Modells

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

Demnach sind von den 51 Sportabbrechern 41 richtig klassifiziert worden (80%) und von den 31 Sportfortsetzern mit 16 etwas mehr als die Hälfte (52%). Insgesamt wurden 57 von 82 Fällen korrekt beurteilt, das sind 69,5%.

Die Berechnungen von Odds und Odds-Ratios zum Vergleich der Daten des Prädiktionsmodells mit den tatsächlich eintreffenden Ereignissen erfolgten nach den Empfehlungen von Bland & Altman (2000). Für die Prädiktion „Sport abbrechen“ ergab sich eine Odds von 1,661 (CI=1,127 – 2,449), dass der Sport tatsächlich abgebrochen wurde. Die Odds, dass sich trotz dieser 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 Prädiktion „Sport abbrechen“ dieses Ereignis eintrat als die Gegenwahrscheinlichkeit, wurde die Odds-Ratio berechnet. Diese betrug

$$OR=4,373 \text{ (CI=1,630 – 11,734)}.$$

Das 95%-Konfidenzintervall des Odds-Ratios umschließt nicht 1 und ist demzufolge statistisch signifikant. Erhält ein Fall im Prädiktionsmodell die retrospektive „Vorhersage“,

dass das Training nicht fortgesetzt wird, hat dieser eine um das 4,4-fache erhöhte Wahrscheinlichkeit, den Sport tatsächlich abgebrochen zu haben gegenüber der Alternative, den Sport fortgesetzt zu haben.

#### 4.8.3. Kombiniertes Modell

##### Sportfortsetzeranalyse aufgrund der Veränderungsdaten von T1-T2 in Kombination mit den signifikanten Prädiktoren des T1-Modells

In diesen Analysen 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 einen BDI=55 und nach dem Training einen BDI=50, resultierte ein Differenzwert von  $Dif_{(T2-T1)}BDI=-5$ , also eine Reduktion der Depressivität um 5 Standardwerte.

**Tab. 27:** Tests der Gruppenunterschiede des Zwischensubjektfaktors "Sportfortsetzung" für Differenzwerte T2-T1

		Dif <sub>(T2-T1)</sub> GHQ	Dif <sub>(T2-T1)</sub> BDI	Dif <sub>(T2-T1)</sub> PSK	Dif <sub>(T2-T1)</sub> SOC	Dif <sub>(T2-T1)</sub> SWE	Dif <sub>(T2-T1)</sub> PSW	Dif <sub>(T2-T1)</sub> KSK	Dif <sub>(T2-T1)</sub> VAS	Dif <sub>(T2-T1)</sub> TV <sup>19</sup>
Forts. Sport	Nein	51	40	46	50	50	51	46	34	37
	Ja	42	30	42	42	42	42	42	26	30
	df	91	68	86	90	90	91	86	58	N=67
	t	0,398	0,468	-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	0,671	0,353

Wie man anhand der Tabelle 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 im Mann-Whitney-U-Test gefunden werden:

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

In den ersten 4 Items hatten die Sportfortsetzer signifikant positivere Veränderungen als die Sportabbrecher, also weniger Enttäuschung (BDI-7), weniger Selbstvorwürfe (BDI-8), weniger selbstbestrafende Gedanken (BDI-9) und mehr Ruhe und Gelassenheit (SF12-9) als zum ersten Untersuchungszeitpunkt. Im PSW Item 2 hatten die Sportabbrecher eine höhere positive Veränderung, also weniger Unterlegenheitsgefühle.

Die trennscharfen Items wurden wieder als Prädiktoren benutzt, um das Risiko "Sport abbrechen" mittels einer logistischen Regressionsanalyse vorherzusagen. Sie wurden zusätzlich mit den signifikanten Koeffizienten aus dem T1-Modell (Fernsehkonsument und SOC-7: „Fühle mich selten als Verlierer“) innerhalb der Analyse kombiniert. Vollständige Daten zwischen den Prädiktoren lagen für 60 Versuchspersonen vor.

<sup>19</sup> Mann-Whitney-U-Test

**Tab. 28:** Ergebnisse der binären logistischen Regressionsanalyse für die Prädiktoren der Differenzwerte und der signifikanten Koeffizienten des T1-Modells

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

Der Fernsehkonsum zu T1, die Veränderungen des BDI-07 („Ich bin von mir enttäuscht“) und die des SF12-9 („Fühle mich ruhig und gelassen“) erwiesen sich als signifikante Koeffizienten für die Regression. Der für das T1-Modell signifikante Koeffizient SOC-7 stellte sich im kombinierten Regressionsmodell als entbehrlich heraus; die Variable erreichte lediglich eine tendenzielle Verbesserung ( $p=0,080$ ). Das Item 2 des PSW und die BDI Items 8 und 9 wurden mangels Signifikanz ebenso nicht in das finale Modell aufgenommen.

Die Anpassungsgüte des Regressionsmodells (-2LL) betrug 58,16 und war mit  $ChiSq=24,95$  ( $p=0,000$ ) hochsignifikant verschieden von den Werten, die durch die Nullhypothese angenommen wurden; Nagelkerkes  $R^2$  betrug 0,454. Mithilfe der Regressionskoeffizienten wurde erneut für jeden Probanden das Risiko berechnet, den Sport abzubrechen. Die Gleichung für den Exponenten  $z$  der Formel  $p = 1 / (1 + e^{-z})$  lautete

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

Inhaltlich lautet die Beziehung: Je weniger Ruhe und Gelassenheit der Proband nach dem Training hinzugewinnen konnte, je höher sein ursprünglicher Fernsehkonsum war und je weniger er Gefühle eigener Enttäuschung im Interventionszeitraum abbauen konnte, desto größer war der Exponent  $z$ , der nach Einsatz in die Formel zu einem höheren berechneten Risiko des Sportabbruchs führte. Im Folgenden findet sich die Klassifizierungstabelle, in der die beobachtete Gruppenzugehörigkeit der -aufgrund des berechneten Modells vorhergesagten- gegenübergestellt ist:

**Tab. 29:** Beobachtete und vorhergesagte Häufigkeiten der Sportfortsetzung und des -abbruchs aufgrund des kombinierten Modells

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

Demnach wurden von den 31 Sportabbrechern 25 richtig klassifiziert (81%) und von den 29 Sportfortsetzern 20 (69%). Insgesamt waren mit dem Modell 45 von 60 Fällen korrekt beurteilt, das sind 75%. Für die Prädiktion des Risikos ergab sich eine Odds von 2,599 (CI=1,471 – 4,592), dass der Sport tatsächlich abgebrochen wurde. Die Odds, dass sich trotz der Risikoprädiktion der Sport fortsetzte, betrug 0,281 (CI=0,131 – 0,599). Die Odds-Ratio des Risikos erreichte den Wert

OR=9,259 (CI=2,821 – 30,394).

Weil das Konfidenzintervall 1 nicht umschließt, ist die Odds-Ratio signifikant. Erhält ein Proband die Klassifizierung „Sport wird abgebrochen“, hatte er eine um den Faktor 9,3 erhöhte Wahrscheinlichkeit, den Sport tatsächlich abzubrechen, als den Sport – entgegen der Vorhersage – fortgesetzt zu haben.

## 4.9. Längsschnittanalysen der Interventionseffekte

### 4.9.1. Statistische Vorüberlegungen zur Zeitstabilität

Um die Stabilität von möglichen Interventionseffekten im follow-up beurteilen zu können, erfolgten die Tests der Zeiteffekte für die Probanden, die an allen Messzeitpunkten teilgenommen hatten.

Zur Ermittlung des Interventionseffektes wurde für jedes Fragebogeninstrument eine multivariate Varianzanalyse mit Messwiederholung für die Untersuchungszeitpunkte prä, post und follow-up (Faktorstufen des Innersubjektfaktors) berechnet. Vor der Analyse wurde mit dem Mauchly-Test (Mauchly, 1940) geprüft, ob Sphärizität zwischen den Faktorstufen des Innersubjektfaktors angenommen werden konnte. War dies nicht der Fall, erfolgte in der Analyse eine Korrektur der Freiheitsgrade nach (Greenhouse & Geisser, 1959), welches zu einem adjustierten  $p$  führte.

In Bezug auf eine mutmaßliche Zeitstabilität der Interventionseffekte wurde in den Hypothesen keine absolute Zeitstabilität (z.B. signifikante Erhöhung des SOC mit keinerlei Rückgang dieser im follow-up) erwartet. Die im Theorieteil vorgestellten Interventionsstudien (vgl. Kapitel 1.4.1.) zur Verbesserung der psychischen Gesundheit von Erwerbslosen zeigten deskriptiv fast alle einen Rückgang des Interventionserfolges. Auch der metaanalytische durchschnittliche Interventionseffekt von Paul & Moser (2009) wurde von den Autoren als eher kurzfristig wirksam eingeschätzt. Von daher wurde nach einem statistischen Test gesucht, 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 zu beobachten ist, dieser aber nicht bis zum Ausgangsniveau zurückkehrt. Im Gegensatz zur absoluten Zeitstabilität sollte eine relative Zeitstabilität prüfbar sein. Zur Beurteilung der relativen Zeitstabilität des Innersubjekteffektes wurde der Helmert-Kontrast nach den Empfehlungen von Rudolf & Müller 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 für 3 Messzeitpunkte somit prüfbar, ob ein Interventionseffekt selbst unter Einbeziehung seiner (zu erwartenden) Reduktion im follow-up noch signifikant messbar ist. Die Beurteilungen der Zeiteffekte erfolgten nach folgendem Schema:

Anmerkung: Geprüfte Innersubjekteffekte werden im Folgenden mit (A)-(D), geprüfte Zwischensubjekteffekte mit (a)-(d) bezeichnet.

#### 4.9.1.1. Prüfung

- (A) Für den Innersubjekteffekt (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 der Zeit bestätigt, womit eine Veränderung der Daten zwischen allen Erhebungszeitpunkten als belegt galt.
- (B) Zur differenzierten Evaluation der Veränderung zwischen den Messzeitpunkten wurden innersubjektive Kontraste für die wiederholte Messung im Untersuchungszeitpaar T1-T2 und im Messzeitpaar T2-T3 geprüft. Es erfolgte jeweils der Test der Nullhypothese, dass die Daten zwischen den Zeitpunkten unverändert waren. Konnte die jeweilige Nullhypothese signifikant abgelehnt werden, galt ein Unterschied im betreffenden Zeitpaar als belegt.
- (C) Zur Beurteilung der relativen Zeitstabilität des Interventionseffektes wurde auf der Ebene der psychischen und psychosozialen Gesundheit/ Medienkonsum zusätzlich der Helmert Kontrast nach den Empfehlungen von Rudolf & Müller (2004) berechnet. Es wurde die Nullhypothese geprüft, dass sich T1 nicht vom mittleren Effekt von T2 unter Hinzunahme von T3 unterscheidet. Konnte die Nullhypothese falsifiziert werden, galt der mittlere Unterschied als belegt.
- (D) Falls sich die Nullhypothese entweder für (A) oder (B) oder (C) als falsifiziert erwies, wurden zur Beurteilung der Veränderungsrichtung deskriptiv die geschätzten Randmittel der Faktorstufen des Innersubjektfaktors (T1, T2, T3) inspiziert. Hierbei erfolgte die Kontrolle, ob der Unterschied zwischen T1-T2 deskriptiv die vorausgesagte Änderung der Hypothese zum Interventionseffekt signalisierte (z.B. Erhöhung des Kohärenzsinn im SOC von T1-T2). Des Weiteren wurde examiniert, in welche Richtung sich die Daten während des follow-up (T2-T3) entwickelten. Falls weder in (A) noch in (B) oder in (C) die Nullhypothesen zurückgewiesen werden konnte, entfiel die Überprüfung der deskriptiven Werte, da von einer statistischen Gleichheit der Daten zwischen den Messzeitpunkten ausgegangen werden musste.

#### 4.9.1.2. Interpretation

- (1) Lag ein Haupteffekt des Innersubjektfaktors (A) vor, galt lediglich eine Veränderungssensitivität des entsprechenden Instrumentes als belegt (die drei Messzeitpunkte unterscheiden sich).
- (2) Falls der innersubjektive Kontrast zwischen T1-T2 signifikant war (B) und gleichzeitig die Veränderung zwischen den Messzeitpunkten T1-T2 deskriptiv in die hypothetisierte Richtung wies (D), galt die entsprechende Hypothese zum Interventionseffekt (z.B. „Kohärenzgefühl steigt während des Interventionszeitraums“) als bestätigt.
- (3) War der geprüfte Kontrast zwischen T2-T3 signifikant (B), galt die deskriptiv ermittelte Veränderungsrichtung (D) des follow-ups als bestätigt (z.B. „Kohärenzgefühl verringert sich im follow-up“).
- (4) Falls die Hypothese zum Interventionseffekt bestätigt war (2) und gleichzeitig eine Signifikanz des Helmert-Kontrastes belegt werden konnte (C), galt eine relative Zeitstabilität des Interventionseffekts als erwiesen (z.B. „Die Verbesserung des Kohärenzgefühls war auch unter Einbezug des follow-up nachweisbar“).

- (5) Falls in diesem Setting der Helmert-Kontrast eine Signifikanz vermissen ließ (C), wurde der Effekt der bestätigten Hypothese (2) als zeitlich instabil interpretiert (z.B. „Die Verbesserung des Kohärenzgefühls löste sich im follow-up wieder auf“).

#### 4.9.2. Statistische Vorüberlegungen zu Moderatoreffekten

Für die Variablen, denen hypothetisch ein interventionsmoderierender Einfluss unterstellt wurde, erfolgten Tests als Interaktionseffekt zwischen dem Zwischensubjektfaktor und dem Innersubjektfaktor (Zeiteffekt). Mit der Prüfung innerhalb desselben multivariaten varianzanalytischen Modells war gewährleistet, dass es beim Hypothesentesten nicht zur Alphafehlerkumulierung kommt, weil das Signifikanzniveau von  $p=0,05$  auf das gesamte Modell seine Anwendung fand. Als maßgeblich moderierend wurde ein differenzierter Interventionseffekt inklusive des follow-up Ergebnisses definiert. Von daher erfolgte die Prüfung der Hypothesen zu den Zwischensubjektfaktoren über den Helmert-Kontrast des Interaktionseffektes. Die Beurteilung der Interaktionseffekte lehnte sich an das Schema der Zeiteffekte an, wurde aber weniger komplex abgearbeitet.

##### 4.9.2.1. Prüfung

- (a) Die Kombination *Zwischensubjektfaktor* \* *Innersubjektfaktor* wurde auf Interaktion getestet. Es wurde die Annahme der Nullhypothese geprüft, dass der Zwischensubjektfaktor keinen Einfluss auf die Änderung der Faktorstufen (T1, T2, T3) des Innersubjektfaktors hat. Konnte die Nullhypothese abgelehnt werden galt ein differentieller Einfluss des Zwischensubjektfaktors als belegt.
- (b) Zum differenzierten Test wurde die Annahme der Nullhypothese des Helmert-Kontrastes geprüft, die besagt, dass der Zwischensubjektfaktor keinen Einfluss auf die Veränderung des Innersubjektfaktors von T1 vs. allen folgenden hat. Konnte die Nullhypothese abgelehnt werden, galt ein differentieller Einfluss des Zwischensubjektfaktors von T1 nach T2 inklusive T3 als belegt. Innerhalb der gleichen Prozedur wurde die Nullhypothese getestet, dass der Zwischensubjektfaktor keinen Unterschied zwischen den Faktorstufen T2 und T3 des Innersubjektfaktors bedingt. Bei Ablehnung der Nullhypothese galt ein Interaktionseffekt des Zwischensubjektfaktors isoliert für das follow-up als belegt.
- (c) Falls die Nullhypothese in (a) oder in (b) erfolgreich falsifizierbar war, wurden die geschätzten Randmittel der Faktorstufen des Innersubjektfaktors (T1, T2, T3) deskriptiv erfasst. Es wurde kontrolliert, ob die durch den Zwischensubjektfaktor bedingte unterschiedliche Veränderungsrichtung in die hypothetische Richtung wies.

##### 4.9.2.2. Interpretation

- (1) Bei (a) wurde davon ausgegangen, dass der Zwischensubjektfaktor den Zeiteffekt signifikant zu moderieren vermochte.
- (2) Falls der Helmert-Kontrast (b) eine signifikante Interaktion für T1 vs. T1&T2 anzeigte, wurde interpretiert, dass der Zwischensubjektfaktor für einen differentiellen mittleren Interventionseffekt sorgte. Wenn dieser mit der hypothetisch erwarteten in Übereinstimmung stand (c), galt die entsprechende Hypothese zum Zwischensubjektfaktor als bestätigt (z.B. „Kohorte 2 erzielt im mittleren Interventionseffekt einen signifikant höheren Zuwachs des Kohärenzsinn im Vergleich zur Kohorte 1“).

(3) War der Interaktionseffekt ausschließlich signifikant für T2-T3 (b), wurde dies als isolierte Wirkung des Zwischensubjektfaktors während der Veränderung im follow-up Zeitraum interpretiert (z.B. „Der Kohärenzsinn verschlechterte sich im follow-up selektiv für Kohorte 1“).

### 4.9.3. Ergebnisse der Hypothesentests

#### 4.9.3.1. Hypothesentests zur Veränderung der physischen Gesundheit

##### 4.9.3.2.1. Innersubjektfaktor „Messzeitpunkte“

Tabelle 30 fasst ausschließlich die Parameter zusammen, an denen sich signifikante Interventionseffekte auf körperlicher Ebene nachweisen ließen (der innersubjektive Kontrast zwischen T1-T2 war signifikant)

**Tab. 30:** Ergebnisse der Interventionseffekte zwischen T1 und T2, körperliche Variablen

p: Signifikanzniveau des Innersubjektcontrastes; Q: Anzahl der Messungen;  $\eta^2$ : Effektstärke Eta ;  $\eta$  Cohen Class: Bewertung der Effektstärke nach Cohen; N: Anzahl der Probanden

Innersubjektcontrasten T1-T2	p	Q	$\eta^2$	$\eta$ Cohen Class.	N
Systolischer Ruheblutdruck [mm Hg]	0,016	3	0,096	Mittel	59
Diastolischer Ruheblutdruck [mm Hg]	0,000	3	0,195	Groß	59
Maximal erreichte Leistung [W]	0,000	3	0,205	Groß	56
Maximale Ventilation [l/min]	0,014	3	0,111	Mittel	53
Maximale Sauerstoffaufnahme [l/min]	0,011	3	0,118	Mittel	53
Relative max. Sauerstoffaufnahme [l/min*kg]	0,002	3	0,171	Groß	53
Gesamtkörperfett [kg]	0,040	3	0,072	Mittel	59
Prozentualer Körperfettanteil [%]	0,017	3	0,096	Mittel	59

In allen Fällen entsprach die Richtung der Veränderung der Messwerte der erwarteten Veränderungsrichtung, insofern können folgende Hypothesen bestätigt werden:

- Der systolische und der diastolische Ruhe-Blutdruck nehmen ab.
- Die absolute Körperfettmasse verringert sich.
- Der prozentuale Körperfettanteil wird geringer.
- Die maximale Ergometerleistung nimmt zu.
- Die maximale Sauerstoffaufnahme ( $VO_2\text{max}$ ) nimmt zu.
- Die gewichtsbezogene maximale Sauerstoffaufnahme ( $\text{rel}VO_2\text{max}$ ) nimmt zu.
- Die maximale Ventilation nimmt zu.

Somit wurden durch das Interventionsprogramm einerseits wesentliche Risikofaktoren für Herz-Kreislaufkrankungen positiv beeinflusst, andererseits auch die körperliche Leistungsfähigkeit gesteigert, die zusätzlich einen separaten Einfluss auf das kardiovaskuläre Risiko besitzt.

Entgegen unserer Erwartung konnten keinerlei Effekte auf das Profil der Blutfette nachgewiesen werden, somit wurde die Hypothese:

- Der Lipidstatus (Gesamtcholesterol, HDL, LDL, Triglyceride) verbessert sich, nicht bestätigt.

Auch bezüglich des Körpergewichtes konnten im Interventionszeitraum keine signifikanten Veränderungen nachgewiesen werden. Damit konnte auch die Hypothese:

- Der Body Mass Index (BMI) nimmt ab, nicht bestätigt werden.

Man muss berücksichtigen, dass Körpergewicht und damit BMI stark von konstitutionellen Voraussetzungen, insbesondere vom Umfang der Muskelmasse abhängen. Die Körperzusammensetzung, speziell die gesundheitsrelevanten Anteile an Muskelmasse und Körperfett werden durch den Body Mass Index nicht berücksichtigt.

In unserer Studie sank der prozentuale Körperfettanteil der Teilnehmer durchschnittlich um 0,62%. Dies entsprach einem Absolutgewicht von durchschnittlich 0,5 kg reiner Fettmasse. Durch das körperliche Training wurde im Interventionszeitraum durchschnittlich im gleichen Umfang Muskelmasse aufgebaut. Die Ergebnisse fallen in der 2. Kohorte deutlicher aus und erreichen hier statistische Signifikanz. Bei der Erhöhung der fettfreien Körpermasse erreichen die Männer im Vergleich zu den Frauen höhere und statistisch signifikante Werte. Diese äußerst positiven Veränderungen der Körperzusammensetzung spiegeln sich jedoch nicht in einer signifikanten Veränderung des Körpergewichtes und damit auch nicht in einer Veränderung des BMI wider. Gerade Interventionsstudien zum Thema Übergewicht sollten daher nicht nur auf die letztgenannten Parameter fokussieren, sondern den gesundheitsrelevanten prozentualen Körperfettanteil mit berücksichtigen!

Im Folgenden wurden zur Bewertung der Zeitstabilität der erreichten Interventionseffekte die Innersubjektkontraste zwischen T2 und T3 geprüft. War der geprüfte Kontrast zwischen T2-T3 nicht signifikant, wurde dies als Zeitstabilität gewertet.

**Tab. 31:** Ergebnisse der Interventionseffekte zwischen T2 und T3, körperliche Variablen

p: Signifikanzniveau des Innersubjektcontrastes; Q: Anzahl der Messungen;  $\eta^2$ : Effektstärke Eta ;  $\eta$  Cohen Class: Bewertung der Effektstärke nach Cohen; N: Anzahl der Probanden

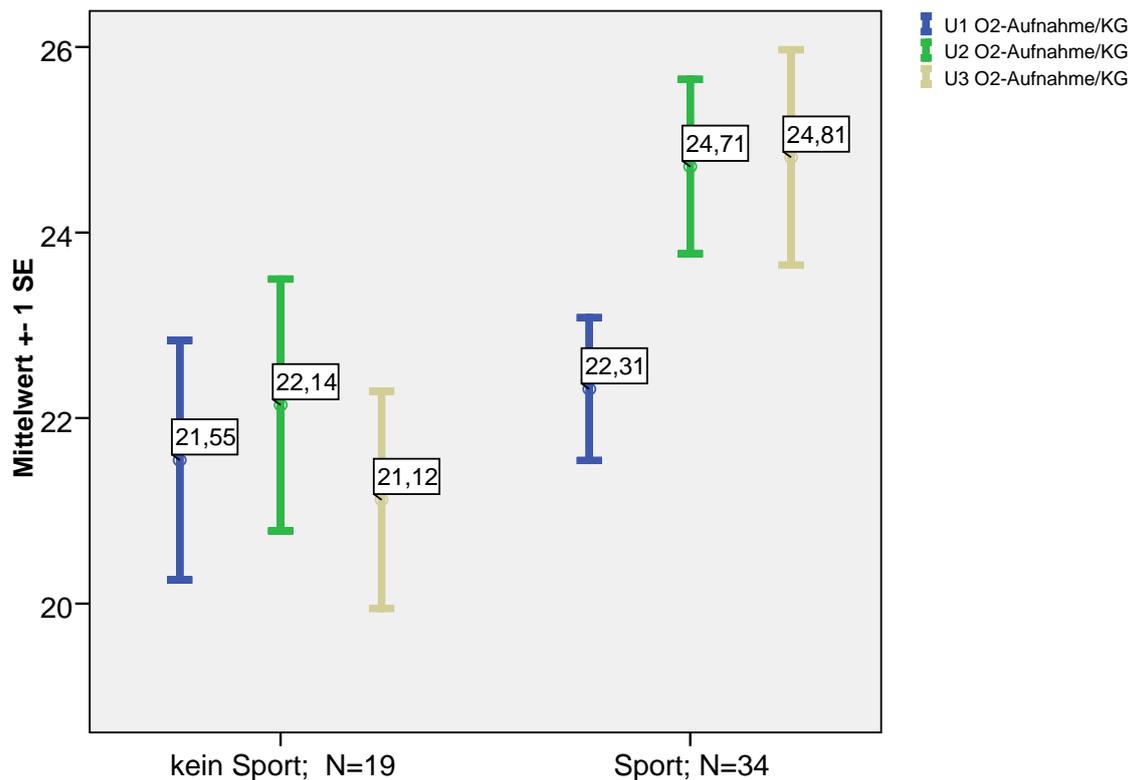
Innersubjektkontraste T2-T3	p	Q	$\eta^2$	$\eta$ Cohen Class.	N
Systolischer Ruheblutdruck [mm Hg]	0,414	3	0,010	Klein	59
Diastolischer Ruheblutdruck [mm Hg]	0,866	3	0,000	Klein	59
Maximal erreichte Leistung [W]	0,380	3	0,014	Klein	56
Maximale Ventilation [l/min]	0,512	3	0,008	Klein	53
Maximale Sauerstoffaufnahme [l/min]	0,707	3	0,003	Klein	53
Relative max. Sauerstoffaufnahme [l/min*kg]	0,629	3	0,005	Klein	53
Gesamtkörperfett [kg]	0,391	3	0,013	Klein	59
Prozentualer Körperfettanteil [%]	0,146	3	0,036	Klein	59

Es zeigte sich, dass alle zwischen T1 und T2 nachgewiesenen signifikanten Interventionseffekte für die Gesamtgruppe auch zeitstabil sind, obwohl nur etwa 64% der Teilnehmer das Angebot einer kostenlosen Trainingsfortsetzung nutzten.

#### 4.9.3.2.2. Zwischensubjekteffekt „Sportfortsetzer“

Die folgende Abbildung stellt den Interventionseffekt bezogen auf die Verbesserung der körperlichen Leistungsfähigkeit differenziert nach Personen dar, die den Sport für weitere 6 Monate fortsetzten oder nicht. Erstaunlicherweise zeigten sich bereits nach Teilnahme am angeleiteten Training deutliche Unterschiede zwischen den zukünftigen Sportfortsetzern und denen, die später nicht weiter trainierten. Während in der Gruppe der Sportfortsetzer ein statistisch signifikanter Anstieg der gewichtsbezogenen maximalen Sauerstoffaufnahme zwischen T1 und T2 nachweisbar war und das verbesserte Leistungsniveau auch bei selbständigem Training gehalten wurde, konnte in der Gruppe der zukünftigen Abbrecher kein signifikanter Trainingserfolg zwischen T1 und T2 nachgewiesen werden und die körperliche Leistungsfähigkeit fiel nach Beendigung des Trainings sogar unter das Ausgangsniveau ab.

Dies belegt u. E. die Abhängigkeit des Interventionseffektes durch die individuelle Motivation.



**Abb.22:** Vergleichende Darstellung der körperlichen Leistungsfähigkeit anhand der gewichtsbezogenen maximalen Sauerstoffaufnahme (relVO<sub>2</sub>max [ml/min\*kg]) bei Ausbelastung auf dem Fahrradergometer zwischen Personen, die nach T2 den Sport fortsetzten oder nicht (U1 = Untersuchung zu T1, U2 und U3 entsprechend)

#### 4.9.3.2. Hypothesentests zur Reduktion von Risikofaktoren der psychischen Gesundheit (L1)

Eine Reduktion der Risikofaktoren käme einer Erhöhung der Werte des *General Health Questionnaires (GHQ)*, einer Senkung der Werte des *Beck-Depression-Inventars (BDI)* und einer Erhöhung des psychischen Funktionsverständnisses in der *Psychischen Summenskala (PSK)* des SF12 gleich. Die Messzeitpunkte der jeweiligen Fragebögen bildeten den Innersubjektfaktor und wurden jeweils zusammen mit den Zwischensubjektfaktoren „Kohorte“ und „Depressivität“ innerhalb einer multivariaten ANOVA berechnet; diese enthielt 4 Gruppen (Kohorte1, Kohorte2, Depressiv, nicht-Depressiv).

##### 4.9.3.2.1. Innersubjektfaktor „Messzeitpunkte“

###### GHQ (N=56<sup>20</sup>)

- (A) Die Veränderung zwischen den Untersuchungsterminen drückte sich als hochsignifikanter Haupteffekt der Messzeitpunkte aus ( $F_{(df=2;104)}=9,058$ ;  $p=0,000$ ).
- (B) Der innersubjektive Kontrast im Zeitpaar T1-T2 war hochsignifikant ( $F_{(df=1;52)}=17,750$ ;  $p=0,000$ ), ebenso wie der Kontrast zwischen T2-T3 ( $F_{(df=1;52)}=11,485$ ;  $p=0,001$ ).
- (C) Der Helmert-Kontrast war hochsignifikant zwischen T1 vs. T2&T3 ( $F_{(df=1;52)}=7,647$ ;  $p=0,008$ ).

<sup>20</sup> Da innerhalb der ANOVA der Zwischensubjektfaktor "Depressivität" mit getestet wurde verringerte sich das ursprüngliche N=69 im GHQ um die fehlenden Werte der Depressivitätsdiagnose (BDI zu T1) auf N=56.

- (D) Für den GHQ bestätigte sich die hypothetisch erwartete Steigerung von T1 (1,890±0,488) nach T2 (2,131±0,434), jedoch mit anschließender Absenkung nach T3 (1,962±0,546).

Die Aussage von (D) in Kombination mit dem signifikanten Kontrast zw. T1-T2 (B) ergibt die Bestätigung der Hypothese „*Allgemeine psychische Beschwerden senken sich im Interventionszeitraum*“. Im follow-up verkleinert sich der Interventionseffekt signifikant (B), allerdings konnte durch den signifikanten Helmert-Kontrast (C) belegt werden, dass der ursprüngliche Interventionseffekt selbst unter Hinzunahme des follow-up hochsignifikant messbar und somit relativ zeitstabil war.

### **BDI (N=57)**

- (A) Die Daten des BDI erfüllten nicht die Sphäritätsannahme, weshalb die Freiheitsgrade nach Greenhouse & Geisser (1959) korrigiert wurden. Es zeigte sich kein signifikanter Haupteffekt ( $F_{(df=1,783;94,34)}=2,029$ ;  $p=0,142$ ) für den Zeitfaktor; demnach wichen die Werte zwischen den drei Zeitpunkten nicht signifikant von der Nullhypothese ab.
- (B) Die Kontraste im Messzeitpaar T1-T2 erreichten einen signifikanten Unterschied ( $F_{(df=1;53)}=5,136$ ;  $p=0,028$ ), die im Messzeitpaar T2-T3 nicht ( $F_{(df=1;53)}=1,904$ ;  $p=0,173$ ).
- (C) Der Helmert-Kontrast war nicht signifikant zwischen T1 vs. T2&T3 ( $F_{(df=1;53)}=2,257$ ;  $p=0,139$ ).
- (D) Die Messzeitwerte des BDI gingen in die hypothetische Richtung mit einer Reduktion der Depressivität von T1 (58,50±6,69) nach T2 (56,45±9,22), jedoch mit anschließender Erhöhung dieser nach T3 (58,16±9,20).

Die Aussage von (D) in Kombination mit dem signifikanten Kontrast zwischen T1-T2 (B) ergibt die Bestätigung der Hypothese „*Depressivität senkt sich im Interventionszeitraum*“. Zwischen T2-T3 senkt sich der Interventionseffekt zwar nicht signifikant (B), allerdings konnte mangels signifikanten Helmert-Kontrastes (C) keine relative Zeitstabilität des Trainings attestiert werden.

### **PSK (N=56<sup>21</sup>)**

- (A) Es ergab sich ein signifikanter Haupteffekt der Zeit ( $F_{(df=2;104)}=3,762$ ;  $p=0,026$ ).
- (B) Sowohl das Messzeitpaar T1-T2 ( $F_{(df=1;52)}=5,697$ ;  $p=0,021$ ) als auch das von T2-T3 ( $F_{(df=1;52)}=6,563$ ;  $p=0,013$ ) wiesen signifikante Kontraste auf.
- (C) Der Helmert-Kontrast zwischen T1 vs. T2&T3 erwies sich als nicht signifikant ( $F_{(df=1;52)}=0,991$ ;  $p=0,324$ ).
- (D) Für den PSK zeigten die Teilnehmer die hypothetisch angenommene Erhöhung ihres psychischen Funktionsverständnisses von T1 (49,11±7,7) nach T2 (51,92±7,27), jedoch mit anschließender Senkung dieser nach T3 (48,57±9,64).

Die Aussage von (D) in Kombination mit dem signifikanten Kontrast zwischen T1-T2 (B) ergibt die Bestätigung der Hypothese „*Psychische Lebensqualität erhöht sich im Interventionszeitraum*“. Zwischen T2-T3 senkt sich der Interventionseffekt signifikant (B) und auch der nicht signifikante Helmert-Kontrast (C) weist auf eine zeitliche Instabilität des Interventionseffektes hin.

<sup>21</sup> Da innerhalb der ANOVA der Zwischensubjektfaktor "Depressivität" mit getestet wurde verringerte sich das ursprüngliche N=70 in der PSK um die fehlenden Werte der Depressivitätsdiagnose (BDI zu T1) auf N=56.

#### 4.9.3.2.2. Zwischensubjekteffekt „Kohorte“

- (a) Im **GHQ** zeigte sich für den Zwischensubjektfaktor keine signifikante Interaktion mit dem Innersubjektfaktor ( $F_{(df=2;104)}=2,283$ ;  $p=0,107$ ).
- (b) Es wurde keine Interaktion im Helmert-Kontrast ausgelöst ( $F_{(df=1;52)}=0,991$ ;  $p=0,324$ ). Es fand sich keine Interaktion im Kontrast zwischen T2-T3 ( $F_{(df=1;52)}=0,156$ ;  $p=0,695$ ).
- (c) Da die Daten statistisch unverändert waren, entfiel deren deskriptive Erfassung.

Die Hypothese „Kohorte 2 löst eine stärkere Reduktion allgemeiner psychischer Beschwerden als Kohorte 1 aus“ ist nicht bestätigt.

- (a) Im **BDI** ging der Kohorten-Zwischensubjektfaktor nicht in Interaktion mit dem Innersubjektfaktor ( $F_{(df=1,783;94,34)}=0,552$ ;  $p=0,558$ ).
- (b) Eine Interaktion innerhalb des Helmert-Kontrastes war nicht nachweisbar ( $F_{(df=1;53)}=0,375$ ;  $p=0,543$ ). Ein Interaktionseffekt im follow-up blieb unbelegt ( $F_{(df=1;53)}=0,649$ ;  $p=0,424$ ).
- (c) Deskriptive Statistiken wurden mangels Signifikanz nicht examiniert.

Die Hypothese „Kohorte 2 führt zu einer stärkeren Reduktion von Depressivität als Kohorte 1“ ist nicht bestätigt.

- (a) Im **PSK** zeigte sich kein Interaktionseffekt des Zwischen- mit dem Innersubjektfaktor ( $F_{(df=2;104)}=1,835$ ;  $p=0,165$ ).
- (b) Innerhalb des Helmert-Kontrastes blieb eine Interaktion unbelegt ( $F_{(df=1;52)}=2,075$ ;  $p=0,156$ ). Zwischen T2-T3 konnte ebenso keine Interaktion objektiviert werden ( $F_{(df=1;52)}=2,075$ ;  $p=0,156$ ).
- (c) Deskriptive Statistiken wurden mangels signifikanter Verschiedenheit nicht erfasst.

Die Hypothese „Kohorte 2 vermindert das eingeschränkte psychische Rollenverständnis stärker als Kohorte 1“ ist nicht bestätigt.

#### 4.9.3.2.3. Zwischensubjekteffekt „Depressivität“

- (a) Im **GHQ** ging der Zwischensubjektfaktor eine hochsignifikante Interaktion mit dem Innersubjektfaktor ein ( $F_{(df=2;104)}=5,049$ ;  $p=0,008$ ).
- (b) Der Helmert-Kontrast war signifikant ( $F_{(df=1;52)}=5,036$ ;  $p=0,029$ ), was als Beleg dafür zu werten ist, dass die Depressivitätsdiagnose einen signifikant verschiedenen mittleren Therapieerfolg produzierte. Der Kontrast zwischen T2-T3 war ebenso signifikant ( $F_{(df=1;52)}=5,070$ ;  $p=0,029$ ), welches belegt, dass der Zwischensubjektfaktor auch isoliert für das follow-up für differentielle Zeiteffekte sorgte.
- (c) Deskriptiv zeigte sich, dass Probanden mit einer Depressivitätsdiagnose die stärksten Erhöhungen im GHQ aufwiesen (T1:  $1,455\pm 0,816$ ; T2:  $1,879\pm 0,733$ ), die sich im follow-up abschwächten ( $1,598\pm 0,913$ ), während die restlichen Probanden nur sehr schwache Erhöhungen zwischen T1-T2 aufwiesen (T1:  $2,325\pm 0,531$ ; T2:  $2,383\pm 0,479$ ), die wieder auf das Ausgangsniveau zurück kehrten (T3:  $2,326\pm 0,591$ ) (vgl. Abb. 23).

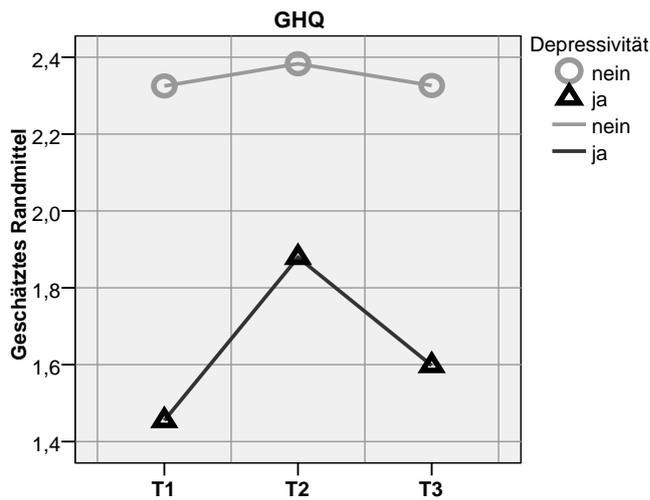


Abb. 23: Interaktionseffekt Depressivität \* Innersubjektfaktor im GHQ

Die Hypothese "Probanden mit klinisch relevanter Depressivität profitieren im mittleren Trainingserfolg mehr von der Reduktion allgemeiner psychischer Probleme als die restlichen Probanden" ist bestätigt. Des Weiteren war der Einfluss des Zwischensubjektfaktors auch isoliert für den follow-up Zeitraum aktiv. Dies sprach für eine selektive Verschlechterung der Probanden mit Depressivitätsdiagnose nach T2 .

- (a) Im **BDI** ging der Faktor „Depressivität“ eine signifikante Interaktion mit dem Zeitfaktor ein ( $F_{(df=1,783;94,34)}=3,574$ ;  $p=0,037$ ).
- (b) Ein signifikanter Helmert-Kontrast belegte einen unterschiedlichen mittleren Interventionserfolg ( $F_{(df=1;53)}=5,610$ ;  $p=0,022$ ). Der nicht-signifikante Kontrast zwischen T2-T3 signalisierte, dass der Zwischensubjektfaktors keinen selektiven Einfluss für das follow-up hatte ( $F_{(df=1;53)}=2,457$ ;  $p=0,123$ ).
- (c) Auch hier profitierten hauptsächlich die Probanden mit Depressivität von der Intervention. Von diesen Probanden senkte sich die Depressivität zuerst (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 Nicht-Depressiven relativ gleich blieben (T1:  $50,59 \pm 7,07$ ; T2:  $51,40 \pm 9,74$ ; T3:  $51,17 \pm 9,72$ ) (vgl. Abb. 24).

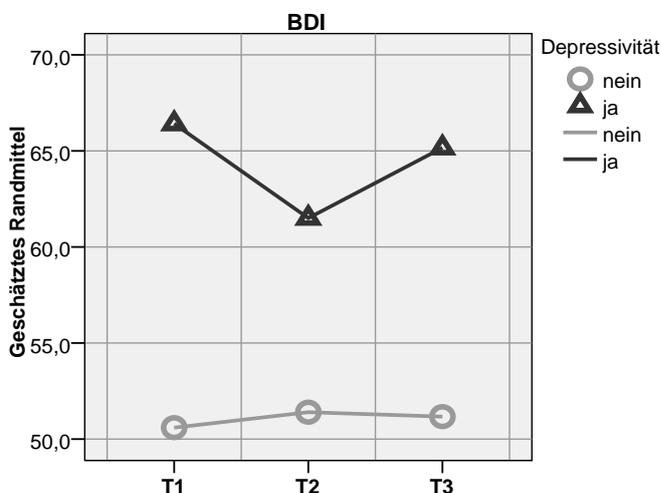


Abb. 24: Interaktionseffekt Depressivität \* Innersubjektfaktor im BDI

Die Hypothese "Vor allem Probanden mit klinisch relevanter Depressivität können diese durch das Training senken" ist bestätigt.

- (a) Der Zeitfaktor der **PSK** wurde vom Faktor Depressivität tendenziell beeinflusst ( $F_{(df=2;104)}=3,063$ ;  $p=0,051$ ).
- (b) Der Helmert-Kontrast zeigte einen signifikanten differenziellen mittleren Therapieerfolg zwischen den Gruppen ( $F_{(df=1;52)}=4,457$ ;  $p=0,040$ ). Der nicht-signifikante Kontrast für T2-T3 ( $F_{(df=1;52)}=1,593$ ;  $p=0,213$ ) spricht gegen eine isolierte differentielle Wirkung des Zwischensubjektfaktors im follow-up.
- (c) Hier bestätigte sich erneut das Phänomen, dass hauptsächlich Probanden mit der Depressivitätsdiagnose eine positive Veränderung im Interventionszeitraum aufwiesen (T1:  $42,81 \pm 13,29$ ; T2:  $48,86 \pm 12,44$ ; T3:  $43,83 \pm 16,51$ ), während die restlichen Probanden kaum Änderung erfuhren (T1:  $55,42 \pm 8,05$ ; T2:  $54,98 \pm 7,54$ ; T3:  $53,32 \pm 10,00$ ) (vgl. Abb. 25).

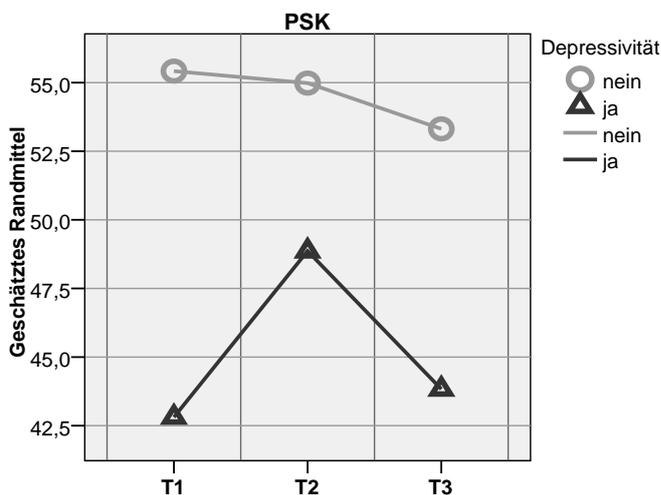


Abb. 25: Interaktionseffekt Depressivität \* Innersubjektfaktor in der PSK

Die Hypothese "Vor allem Probanden mit klinisch relevanter Depressivität reduzieren psychische Einschränkungen ihrer Lebensqualität" ist bestätigt.

#### 4.9.3.3. Hypothesentests zur Erhöhung der Ressourcen der psychischen Gesundheit (L2)

Eine Erhöhung der Resilienz sollte sich in einem Anstieg des Kohärenzsинns im entsprechenden Fragebogen *Sense-of-Coherence (SOC)*, in einer Steigerung der Selbstwirksamkeitserwartung des Fragebogens *Allgemeine Selbstwirksamkeitserwartung (SWE)* und im Zuwachs der Skala *Psychosoziales Wohlbefinden (PSW)* manifestieren. Die Messzeitpunkte eines Testverfahrens stellten die Faktorstufen des Innersubjektfaktors dar und wurden innerhalb der multivariaten ANOVAs zusammen mit dem Zwischensubjektfaktor „Kohorte“ (2 Gruppen) berechnet.

#### 4.9.3.3.1. Innersubjektfaktor „Messzeitpunkte“

##### **SOC (N=72)**

- (A) Die Freiheitsgrade mussten nach Greenhouse & Geisser (1959) korrigiert werden. Es zeigte sich kein Haupteffekt des Innersubjektfaktors ( $F_{(df=1,82;127,4)}=1,220$ ;  $p=0,296$ ).
- (B) Innersubjektive Kontraste zeigten sich nicht zwischen den Messzeitpaaren, weder für T1-T2 ( $F_{(df=1;70)}=0,051$ ;  $p=0,823$ ) noch für T2-T3 ( $F_{(df=1;70)}=2,236$ ;  $p=0,139$ ).
- (C) Der Helmert-Kontrast war nicht signifikant ( $F_{(df=1;70)}=0,289$ ;  $p=0,593$ ).
- (D) Da die Daten als statistisch unverändert zu werten waren, entfiel die deskriptive Erfassung der Messzeitpunkte.

Die Hypothese "*Kohärenzsinn steigt im Interventionszeitraum*" ist nicht bestätigt.

##### **SWE (N=72)**

- (A) Es konnte kein signifikanter Haupteffekt der Zeit ermittelt werden ( $F_{(df=2;140)}=0,986$ ;  $p=0,376$ ).
- (B) Innersubjektive Kontraste konnten weder für T1-T2 ( $F_{(df=1;70)}=0,284$ ;  $p=0,596$ ), noch für T2-T3 gefunden werden ( $F_{(df=1;70)}=1,611$ ;  $p=0,209$ ).
- (C) Der Helmert-Kontrast zeigte kein signifikantes Ergebnis ( $F_{(df=1)}=0,039$ ;  $p=0,843$ ).
- (D) Mangels Signifikanz entfiel eine deskriptive Erfassung der Messzeitpunkte.

Die Hypothese "*Selbstwirksamkeitserwartung steigt im Interventionszeitraum*" ist nicht bestätigt.

##### **PSW (N=69)**

- (A) Das Signifikanzkriterium wurde für den Haupteffekt der Zeit nicht erfüllt ( $F_{(df=2;134)}=1,531$ ;  $p=0,220$ ).
- (B) Innersubjektive Kontraste in den Messzeitpaaren T1-T2 ( $F_{(df=1;67)}=2,055$ ;  $p=0,156$ ) und T2-T3 ( $F_{(df=1;67)}=2,393$ ;  $p=0,127$ ) waren nicht belegbar.
- (C) Es fand sich kein Helmert Kontrast ( $F_{(df=1;67)}=0,351$ ;  $p=0,555$ ).
- (D) Wegen unzureichender Signifikanz wurden die deskriptiven Werte der Messzeitpunkte nicht inspiziert.

Die Hypothese "*Psychosoziales Wohlbefinden steigt im Interventionszeitraum*" ist nicht bestätigt.

#### 4.9.3.3.2. Zwischensubjekteffekt „Kohorte“

- (a) Im **SOC** zeigte sich keine signifikante Interaktionen des Zwischensubjektfaktors mit dem Innersubjektfaktor ( $F_{(df=1,82;127,4)}=0,447$ ;  $p=0,641$ ).
- (b) Es wurde kein Helmert-Kontrast ausgelöst ( $F_{(df=1;70)}=0,298$ ;  $p=0,593$ ). Es fand sich kein Kontrast zwischen T2-T3 ( $F_{(df=1;70)}=2,236$ ;  $p=0,133$ ).
- (c) Da die Daten statistisch als gleich zu deuten waren, entfiel die Inspektion der deskriptiven Werte.

Die Hypothese „In Kohorte 2 kommt es zur stärkeren Erhöhung des Kohärenzsинns als in Kohorte 1“ ist nicht bestätigt.

- (a) Im **PSW** zeigte sich kein Interaktionseffekt des Zwischensubjektfaktors "Kohorte" mit dem Innersubjektfaktor ( $F_{(df=2;134)}=1,835$ ;  $p=0,165$ ).
- (b) Innerhalb des Helmert-Kontrastes blieb eine Interaktion unbelegt ( $F_{(df=1;67)}=0,606$ ;  $p=0,439$ ). Zwischen T2-T3 fand keine Interaktion mit dem Gruppenfaktor statt ( $F_{(df=1;67)}=2,767$ ;  $p=0,102$ ).
- (c) Deskriptive Statistiken blieben durch mangelnde Verschiedenheit der Daten ungeprüft.

Die Hypothese „In Kohorte 2 erhöht sich das psychosoziale Wohlbefinden stärker als in Kohorte 1“ ist nicht bestätigt.

- (a) In der **SWE** konnte ein hochsignifikanter Interaktionseffekt des Zwischensubjektfaktors „Kohorte“ mit dem Zeiteffekt gefunden werden ( $F_{(df=2;140)}=4,936$ ;  $p=0,008$ ).
- (b) Ein hochsignifikanter Helmert-Kontrast belegte einen differentiellen mittleren Interventionserfolg ( $F_{(df=1;70)}=9,758$ ;  $p=0,003$ ). Der Kontrast zwischen T2-T3 war nicht signifikant ( $F_{(df=1;70)}=1,750$ ;  $p=0,190$ ).
- (c) Die Inspektion der Daten ergab, dass in Kohorte 1, ausgehend 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 wurde, während in Kohorte 2 der gegenteilige Effekt zu beobachten war. Hier verbesserten sich die Probanden von T1 ( $49,31\pm 10,35$ ) nach T2 ( $51,26\pm 9,41$ ), um in Richtung T3 stabil zu bleiben ( $51,31\pm 12,12$ ) (vgl. Abb. 26).

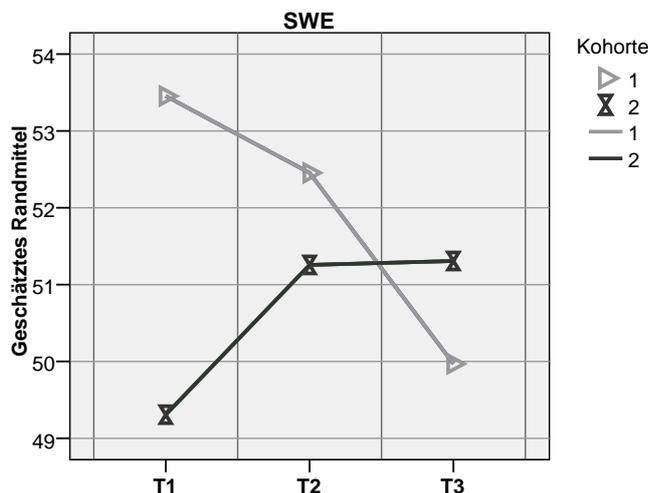


Abb. 26: Interaktionseffekt Kohorte \* Innersubjektfaktor im SWE

Während der Kontrast zwischen den Gruppen in die hypothetische Richtung verlief (Kohorte 2 profitierte stärker als Kohorte 1), war jedoch unerwartet, dass sich Kohorte 1 über die Zeit deskriptiv scheinbar verschlechterte.

Während der hochsignifikante Helmert-Kontrast für einen unterschiedlichen mittleren Therapieerfolg spricht, bedeutet der nicht-signifikante Kontrast zwischen T2-T3, dass

zwischen den Kohorten etwaige Veränderungsunterschiede im follow-up Zeitraum im Rahmen von Zufallsschwankungen zu interpretieren sind. Demnach wäre der Interaktionseffekt vor allem den differentiellen Veränderungen zwischen T1-T2 geschuldet.

Um aufzuklären, ob der Interaktionseffekt evt. als Phänomen einer "Regression zur Mitte" erklärt werden könnte, wurden die verschiedenen anmutenden Ausgangsbedingungen zwischen den Gruppen (Kohorte1:  $53,46 \pm 9,68$ ; Kohorte2:  $49,31 \pm 10,35$ ) mit dem t-Test für unabhängige Stichproben zwischen den Kohorten geprüft. Dieser zeigte einen tendenziellen Unterschied zwischen den Gruppen an ( $t_{(df=70)}=1,745$   $p=0,085$ ).

Um zu prüfen, ob innerhalb der Kohorten Unterschiede zwischen T1-T2 auffindbar sind, wurden paarweise Vergleiche dieser Messzeitpunkte kohortenspezifisch getestet. In den t-Tests für abhängige Stichproben erfolgte hierbei eine Anpassung der Signifikanzschwelle für die zweifache Testung als Šidák-Korrektur nach den Empfehlungen von Abdi (Abdi, 2007)<sup>22</sup>. Für Kohorte 2 zeigte sich keine signifikante Erhöhung der SWE zwischen T1-T2 ( $t_{(df=38)}=1,469$   $p=0,150$ ). Auch für Kohorte 1 war die Verringerung der SWE zwischen T1-T2 nicht signifikant ( $t_{(df=32)}=-0,884$   $p=0,383$ ).

Dass Kohorte 2 für sich genommen im Messzeitpaar T1-T2 keine signifikante Erhöhung der SWE zeigte und sich in Kohorte 1 keine signifikante Verschlechterung offenbarte, liefert einen Hinweis darauf, dass der obige signifikante Interaktionseffekt des Zwischensubjektfaktors "Kohorte" sich nicht gezeigt hätte, wenn die numerische Erhöhung der SWE in Kohorte 2 nicht gleichzeitig mit einer numerischen Verringerung von Kohorte 1 einhergegangen wäre.

Die Hypothese „In Kohorte 2 kommt es zur stärkeren Erhöhung der Selbstwirksamkeitserwartung als in Kohorte 1“ ist zu bestätigen. Allerdings war der Interaktionseffekt nicht eindeutig auf das gesteigerte Training attributierbar. Da die Senkung der Werte in Kohorte 1 nicht als signifikante „Verschlechterung“ belegt werden konnte und ebenso die Erhöhung der Werte in Kohorte 2 keine signifikante „Verbesserung“ demaskierte, soll geschlossen werden, dass es sich bei dem in der SWE gefundenen Interaktionseffekt um einen nicht erklärbaren Befund handelt.

#### 4.9.3.4. Hypothesentests zur Erhöhung der subjektiven körperlichen Gesundheit (L3)

Eine Erhöhung der empfundenen körperlichen Gesundheit sollte in einem Anstieg der körperlichen Lebensqualität in der *Körperlichen Summenskala (KSK)* des SF12 und in einem Zuwachs in der *Visuellen Analogskala* zum Gesundheitszustand (**VAS**) identifizierbar sein. Die Messzeitpunkte der Testverfahren bildeten die Innersubjektfaktoren. „Kohorte“ und „Geschlecht“ bildeten die Zwischensubjektfaktoren der multivariaten ANOVAs. Somit wurden jeweils 4 Gruppen getestet.

---

<sup>22</sup> Die Anpassung für eine einfache Signifikanz erfolgte auf  $p=0,02532$ . Für einen hochsignifikanten Unterschied wurde die Schwelle auf  $p=0,00501$  adjustiert. Lediglich ein Trend wäre mit dem Unterschreiten von  $p=0,05132$  erreicht.

#### 4.9.3.4.1. Innersubjektfaktor „Messzeitpunkte“

##### **KSK (N=70)**

- (A) Es ließ sich kein Haupteffekt der Messzeitpunkte objektivieren ( $F_{(df=2;132)}=0,457$ ;  $p=0,634$ ).
- (B) Innersubjektkontraste zeigten sich nicht für das Messzeitpaar T1-T2 ( $F_{(df=1;66)}=0,615$ ;  $p=0,436$ ) und auch nicht für das von T2-T3 ( $F_{(df=1;66)}=0,010$ ;  $p=0,920$ ).
- (C) Ein Helmert-Kontrast war nicht zu finden ( $F_{(df=1;66)}=0,746$ ;  $p=0,391$ ).
- (D) Da eine statistische Gleichheit zwischen den Messzeitpunkten angenommen werden musste, entfiel die Prüfung der deskriptiven Werte.

Die Hypothese *"Die körperliche Lebensqualität steigt im Interventionszeitraum"* ist nicht bestätigt.

##### **VAS (N=50)**

- (A) Es konnte kein signifikanter Haupteffekt für den Innersubjektfaktor belegt werden ( $F_{(df=2;92)}=1,659$ ;  $p=0,196$ ).
- (B) Für das Messzeitpaar T1-T2 zeigte sich kein Innersubjektkontrast ( $F_{(df=1;46)}=2,443$ ;  $p=0,125$ ). Auch der von T2-T3 zeigte keine Signifikanz ( $F_{(df=1;46)}=0,218$ ;  $p=0,643$ ).
- (C) Der Helmert-Kontrast zeigte sich als nicht verschieden von den Annahmen der Nullhypothese ( $F_{(df=1;46)}=2,391$ ;  $p=0,129$ ).
- (D) Wegen mangelnder Signifikanz zwischen den Messzeitpunkten entfiel die deskriptive Prüfung.

Die Hypothese *"Das subjektive Gesundheitsempfinden steigt im Interventionszeitraum"* ist nicht belegt.

#### 4.9.3.4.2. Zwischensubjekteffekt „Kohorte“

- (a) Im **KSK** zeigte sich keine signifikante Interaktionen von „Kohorte“ mit dem Innersubjektfaktor ( $F_{(df=2;132)}=0,421$ ;  $p=0,657$ ).
- (b) Es wurde kein Helmert-Kontrast ausgelöst ( $F_{(df=1;66)}=0,566$ ;  $p=0,455$ ). Es fand sich kein Kontrast zwischen T2-T3 ( $F_{(df=1;66)}=0,197$ ;  $p=0,659$ ).
- (c) Da die Daten als statistisch unverändert zu interpretieren waren, entfiel die deskriptive Inspektion der Daten.

Die Hypothese *"Die körperliche Lebensqualität im Interventionszeitraum steigt stärker für Kohorte 2 im Vergleich mit Kohorte 1"* ist nicht belegt.

- (a) In der **VAS** ging der Zwischensubjektfaktor nicht in Interaktion mit dem Innersubjektfaktor ( $F_{(df=2;92)}=0,762$ ;  $p=0,470$ ).
- (b) Es wurde kein Helmert-Kontrast ausgelöst ( $F_{(df=1;46)}=0,001$ ;  $p=0,972$ ). Es fand sich kein Kontrast zwischen T2-T3 ( $F_{(df=1;46)}=2,260$ ;  $p=0,140$ ).
- (c) Da die Daten statistisch als gleich zu deuten waren, entfiel die deskriptive Inspektion der Daten.

Die Hypothese "Das subjektive Gesundheitsempfinden steigt im Interventionszeitraum stärker in Kohorte 2 im Vergleich mit Kohorte 1" ist nicht bestätigt.

#### 4.9.3.4.3. Zwischensubjekteffekt „Geschlecht“

- In der **VAS** ging der Zwischensubjektfaktor „Geschlecht“ keine signifikante Interaktion mit dem Innersubjektfaktor ein ( $F_{(df=2;92)}=0,397$ ;  $p=0,674$ ).
- Es wurde kein signifikanter Helmert-Kontrast gefunden ( $F_{(df=1;46)}=0,408$ ;  $p=0,526$ ). Es fand sich kein Kontrast zwischen T2-T3 ( $F_{(df=1;46)}=0,375$ ;  $p=0,543$ ).
- Da sich kein differenzierter Interventionserfolg für den Geschlechtssfaktor zeigte, unterblieb die deskriptive Prüfung der Daten.

Die Hypothese "Das Gesundheitsempfinden steigt stärker für die männlichen Versuchsteilnehmer in Vergleich zu den weiblichen" ist nicht belegt.

- Im **KSK** war der durch das Geschlecht ausgelöste Interaktionseffekt nicht signifikant ( $F_{(df=2;132)}=1,887$ ;  $p=0,156$ ).
- Es wurde ein statistischer Trend im Helmert-Kontrast gefunden ( $F_{(df=1;66)}=2,951$ ;  $p=0,091$ ), welcher tendenziell einen differentiellen mittleren Interventionserfolgs für den Geschlechtssfaktor andeutete. Der Kontrast zwischen T2-T3 war nicht signifikant ( $F_{(df=1;66)}=0,244$ ;  $p=0,623$ ), welches Geschlechtsunterschiede isoliert für den follow-up Zeitraum negierte.
- Da der Helmert-Kontrast zumindest tendenziell einen differentiellen Interventionseffekt für den Geschlechtssfaktor andeutete, wurde dies zum Anlass genommen, die Daten deskriptiv zu examinieren. Deskriptiv zeigte sich hypothesenkonform ein stärkerer Zuwachs von körperlicher Lebensqualität für die Männer (T1:  $46,32\pm 13,20$ ; T2:  $48,70\pm 12,68$ ; T3:  $49,26\pm 13,34$ ) im Vergleich zu den Frauen (T1:  $47,11\pm 12,17$ ; T2:  $46,42\pm 11,67$ ; T3:  $46,05\pm 12,28$ ) (vgl. Abb. 27).

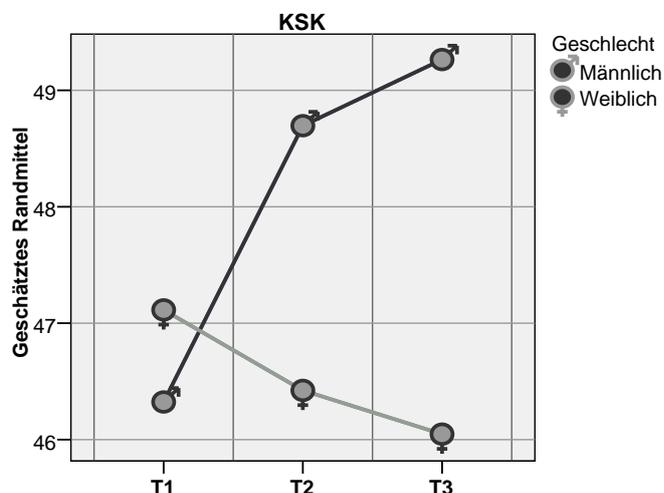


Abb. 27: Interaktionseffekt Geschlecht \* Innersubjektfaktor in der KSK

Grafisch dargestellt lässt sich rasch erkennen, dass die Geschlechter ähnliche Ausgangsbedingungen aufwiesen und sich stetig in verschiedene Richtungen entwickelten. Ebenfalls ist erkennbar, dass der tendenzielle Helmert-Kontrast hauptsächlich dem selektiven Zuwachs der männlichen Probanden geschuldet ist, denn die Werte der weiblichen Versuchsteilnehmer 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 offen legen würden, wurde ein t-Test für abhängige Stichproben für die Subgruppe der männlichen Probanden berechnet. Es zeigte sich eine tendenzielle Erhöhung der körperlichen Lebensqualität ( $t_{(df=31)}=1,891$   $p=0,068$ ) für die Männer im Interventionszeitraum.

Die Hypothese *"Die körperliche Lebensqualität steigt stärker für die männlichen Versuchsteilnehmer im Vergleich zu den weiblichen"* kann nicht belegt werden. Die Ergebnisse weisen aber im Trend in die hypothetische Richtung.

#### 4.9.3.5. Hypothesentest zur Senkung des passiven Freizeitverhaltens (L4)

Der Fernsehkonsum wurde zur Abschätzung eines passiven Freizeitverhaltens erhoben. Gleichzeitig wird Fernsehen vor allem von älteren Menschen auch zur Regulierung des emotionalen Zustandes gebraucht (Depp et al., 2010). Es wurde die Hypothese formuliert, dass die Intervention ein aktiveres Freizeitverhalten initiieren könnte, welches sich in einer Verringerung des täglichen Fernsehkonsums niederschlagen sollte.

Die Variable Fernsehkonsum war unimodal, aber nicht symmetrisch normalverteilt, sondern zeigte eine Rechtsschiefe. Da es im multivariaten Fall mit der Aufnahme von Zwischensubjektfaktoren kein entsprechendes nichtparametrisches Verfahren gibt, wurde dennoch die Varianzanalyse mit Messwiederholung herangezogen, da Varianzanalysen gegen Abweichungen von der Normalverteilung als relativ robust eingeschätzt werden (Bortz, 2005; Zöfel, 2003). Bei schiefen Populationsverteilungen sind Abweichungen von der Normalverteilung nach Bortz auch zu vernachlässigen (Bortz, 2005).

##### 4.9.3.5.1. Innersubjektfaktor „Messzeitpunkte“

- (A) Im **TV-Konsum** ( $N=52^{23}$ ) zeigte sich ein signifikanter Haupteffekt des Innersubjektfaktors ( $F_{(df=2;96)}=4,334$ ;  $p=0,016$ ), womit die Unterschiede zwischen den Messzeitpunkten als statistisch abgesichert gelten.
- (B) Ein hochsignifikanter Innersubjekt Kontrast fand sich für das Messzeitpaar T1-T2 ( $F_{(df=1;48)}=7,532$ ;  $p=0,008$ ). Es ergab sich ein Trend für den Kontrast des Messzeitpaares T2-T3 ( $F_{(df=1;48)}=2,888$ ;  $p=0,096$ ).
- (C) Der Helmert-Kontrast war signifikant ( $F_{(df=1;48)}=5,420$ ;  $p=0,024$ ).
- (D) In Bezug auf die drei Messzeitpunkte zeigte sich die hypothetisch erwartete Reduktion des täglichen TV-Konsums von T1 ( $214,6\pm 106,7$ min) um 36min nach T2 ( $178,7\pm 98,3$ min), jedoch mit anschließender Korrektur nach T3 ( $197,9\pm 129,3$  min).

<sup>23</sup> Da innerhalb der ANOVA der Zwischensubjektfaktor "Depressivität" mit getestet wurde verringerte sich das ursprüngliche  $N=53$  im TV-Konsum um den fehlenden Wert der Depressivitätsdiagnose (BDI zu T1) auf  $N=52$ .

Die Aussage von (D) in Kombination mit dem signifikanten Kontrast zwischen T1-T2 (B) ergibt die Bestätigung der Hypothese „TV-Konsum verringert sich im Interventionszeitraum“. Zwischen T2-T3 geht der Effekt nicht signifikant zurück (B) und der signifikante Helmert-Kontrast (C) belegt die relative zeitliche Stabilität des Interventionseffekts.

#### 4.9.3.5.2. Zwischensubjekteffekt „Kohorte“

- (a) Die Versuchsgruppen erzeugten als Zwischensubjektfaktor keine signifikante Interaktion mit den Messzeitpunkten des Fernsehkonsums ( $F_{(df=2;96)}=0,989$ ;  $p=0,376$ ).
- (b) Der Helmert-Kontrast ( $F_{(df=1;48)}=0,060$ ;  $p=0,807$ ) war nicht signifikant. Auch zwischen T2-T3 fand sich kein Kontrast ( $F_{(df=1;48)}=2,314$ ;  $p=0,135$ ).
- (c) Da statistische Gleichheit angenommen werden musste, entfiel die deskriptive Prüfung der Daten.

Die Hypothese "Der TV-Konsum reduziert sich im Interventionszeitraum stärker in Kohorte 2 als in Kohorte 1" ist nicht belegt.

#### 4.9.3.5.3. Zwischensubjekteffekt „Depressivität“

- (a) Es wurde ein signifikanter Interaktionseffekt für den Zwischensubjektfaktor „Depressivität“ belegt ( $F_{(df=2;96)}=3,134$ ;  $p=0,048$ ).
- (b) Der Helmert-Kontrast zwischen T1 und den folgenden Messzeitpunkten T2&T3 deutete nicht auf einen differentiellen durchschnittlichen Interventionseffekt des Faktors „Depressivität hin“ ( $F_{(df=1;48)}=0,473$ ;  $p=0,495$ ). Im Gegensatz dazu wies der innersubjektive Kontrast zwischen T2 und T3 ein signifikantes Ergebnis auf ( $F_{(df=1;48)}=6,677$ ;  $p=0,013$ ). Die Diskrepanz zwischen den Kontrasten zeigt, dass der Zwischensubjektfaktor keinen differentiellen mittleren Interventionseffekt, sondern einen unterschiedlichen Effekt im follow-up Zeitraum bewirkte.
- (c) Deskriptiv wiesen die depressiven Probanden einen Rückgang ihrer zuvor gewonnenen Reduktion des TV-Konsums (T1: 236±176min; T2: 177±162min; T3: 226±204min) im follow-up auf, während sich die Reduktion des Fernsehkonsums bei den restlichen Teilnehmern (T1: 194±121min; T2: 180±111min; T3: 170±140min) im follow-up fortsetzte bzw. stabil blieb (vgl. Abb. 28).

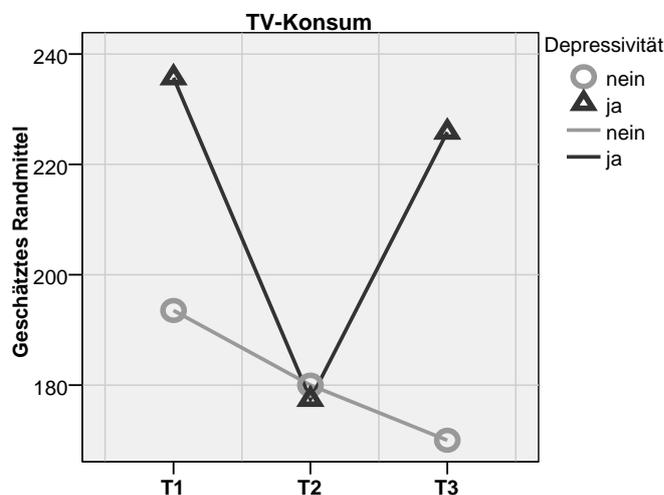


Abb. 28: Interaktionseffekt Depressivität \* Innersubjektfaktor im TV-Konsum

Es wurde kein differentieller Interventionseffekt durch den Faktor „Depressivität“ belegt. Allerdings zeigte sich eine hochsignifikante Wirkung des Zwischensubjektfaktors isoliert für den follow-up Zeitraum. Demnach kehrten Depressive wieder zu ihrem ursprünglichen Fernsehnutzungsverhalten zurück.

#### 4.9.3.6. Hypothesentests zur gesundheitlich positiven Wirkung von sportlicher Aktivität im follow-up Zeitraum (L5)

Positive Effekte einer Fortsetzung des Sports nach T2 könnten sich z.B. dadurch auszeichnen, dass die abhängigen Variablen eine selektive Verbesserung der Sportfortsetzer oder eine selektive Verschlechterung der Sportabbrecher im follow-up Zeitraum aufzeigen. Demnach sollten sich durch die Sportfortsetzung ausgelöste positive follow-up-Effekte entweder durch eine weitere Erhöhung der Gesundheit (bzw. des aktiven Freizeitverhaltens) oder einen besseren Erhalt der Gesundheit manifestieren.

Für die Hypothesen erfolgte der Test des follow-up-Effekts mit ANOVAs für wiederholte Messungen. T2 und T3 wurden für die jeweiligen Messverfahren als Faktorstufen des Innersubjektfaktors und „Sportfortsetzung“ als Zwischensubjektfaktor aufgenommen. Um möglichen divergierenden Ausgangsbedingungen der Erstuntersuchung zwischen den Gruppen Rechnung zu tragen, wurde das jeweilige Modell für dessen Daten von T1 kontrolliert. Die Tests erfolgten für dieselben Probandenkollektive, wie zu den Interventionseffekten.

Da für die Hypothesentests der durch den Zwischensubjektfaktor ausgelöste Unterschied zwischen T2 und T3 von Interesse war, wurde der Interaktionseffekt von „Sportfortsetzung“ mit dem Innersubjektfaktor auf Signifikanz geprüft<sup>24</sup>. Zusätzlich erfolgte der Signifikanztest der Kontrollvariable (T1), um ihren Einfluss auf die Veränderungsergebnisse von T2-T3 abschätzen zu können.

##### 4.9.3.6.1. Interaktionseffekt „Sportfortsetzung“ mit Innersubjekteffekt T2-T3

**Tab. 32:** Tests der follow-up Unterschiede des Zwischensubjektfaktors "Sportfortsetzung"

		GHQ	BDI	PSK	SOC	SWE	PSW	KSK	VAS	TV-Konsum
	df2	53	54	53	69	69	66	67	47	49
Interaktion T2-T3 * Sportfortsetzung	$F_{(df1=1)}$	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	11,396
	$p$	<b>0,031</b>	0,819	0,341	0,090	<b>0,009</b>	0,936	0,326	0,082	<b>0,001</b>

Wie in der Tabelle 32 ersichtlich, zeigte sich in keinem Messinstrument ein signifikanter Interaktionseffekt des Zwischensubjektfaktors „Sportfortsetzung“ mit dem Messzeitpaar T2-T3. Für die Selbstwirksamkeitserwartung (SWE) und die körperliche Lebensqualität (KSK)

<sup>24</sup> Kontraste waren im Modell obsolet, da der Innersubjektfaktor nur zwei Messzeitpunkte enthielt.

deutete sich lediglich eine tendenzielle Beeinflussung des follow-up in der hypothetischen Richtung an. Probanden, die den Sport fortsetzten, erzielten in diesen Messverfahren ein tendenziell gesünderes Ergebnis als die Sportabbrecher.

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 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 z.B. einem starken Aufwärtstrend (T1-T2) eher eine Abwärtskorrektur (T2-T3) folgt, als das dies nach relativ unveränderten Werten (T1-T2) der Fall ist (vgl. z.B. "JoJo-Effekt" in der Gewichtsreduktion oder auch die sog. "Elliott-Wellen" bei Aktienkursen).

Im Falle des follow-up-Effektes der SWE, der dem tendenziellen Einfluss von „Sportfortsetzung“ ( $p=0,089$ ) und gleichzeitig einem hochsignifikanten Einfluss der T1 Variable ( $p=0,009$ ) ausgesetzt war, ist zu interpretieren, dass dessen follow-up-Veränderung stärker von den Werten der Erstuntersuchung abhing, als von der Fortsetzung des Sports. Im Gegensatz dazu ließ sich in der KSK kein Einfluss der Kontrollvariable finden ( $p=0,326$ ), welches den tendenziellen Einfluss der Sportfortsetzung für den follow-up-Effekt der KSK ( $p=0,084$ ) etwas bedeutsamer erscheinen lässt als in der SWE.

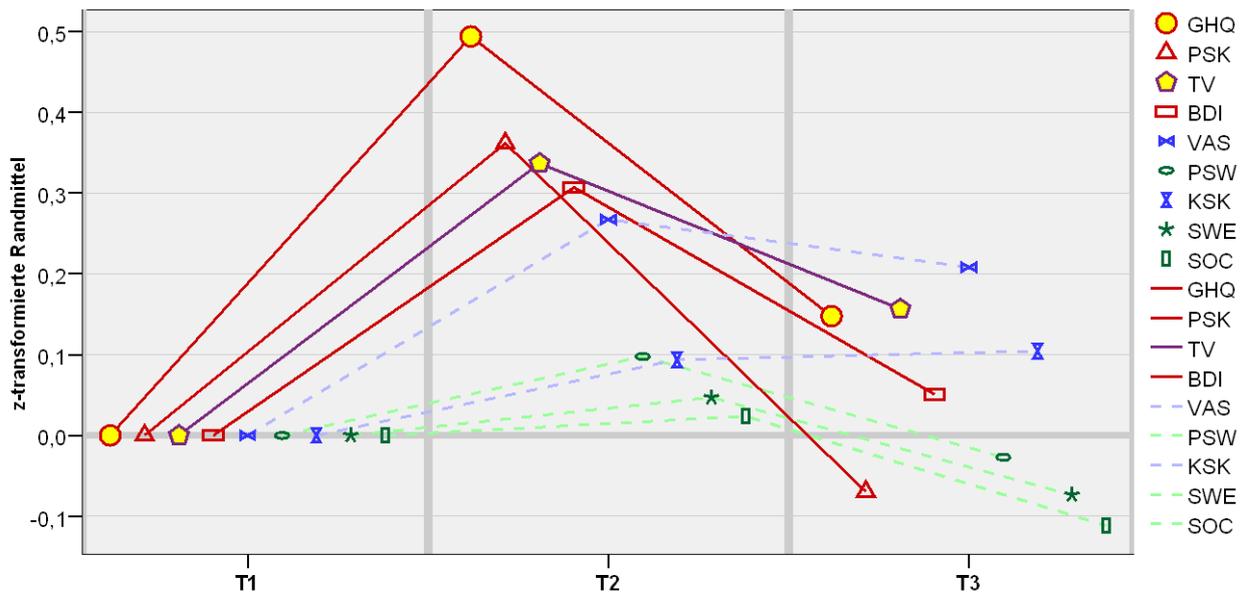
Die Hypothesen zu L5 konnten sämtlich nicht bestätigt werden. Probanden, die nach T2 den Sport fortsetzten, zeigten keine signifikant positiveren Ergebnisse im follow-up Zeitraum als die Sportabbrecher. In der KSK und der SWE deutete sich die hypothetische Richtung lediglich als Trend an.

#### 4.9.3.7. Vergleichende Zusammenfassung der Interventionseffekte

Um die unterschiedlichen Veränderungsergebnisse vergleichend innerhalb einer Übersichtsgrafik darstellen zu können, wurden die Daten der Messverfahren jeweils pro Erhebungszeitpunkt anhand der zu T1 geschätzten Randmittel und Standardabweichungen z-transformiert. Damit beginnen alle Messverfahren mit dem Ursprung Null im Zeithorizont, womit das Ausmaß ihrer Veränderungen vergleichend erfasst werden kann<sup>25</sup>. Die z-Werte des BDI und des TV-Konsums wurden umgepolt, damit für alle Maße nach der Standardisierung eine numerisch positive Veränderungsrichtung gleichzeitig eine Verbesserung darstellt.

---

<sup>25</sup> Verzerrend an dieser Darstellung ist allerdings, dass die tatsächlich vergangene Zeit keine Berücksichtigung fand, die zwischen T1-T2 ca. 3 Monate und zwischen T2-T3 ca. 6 Monate betrug. Insofern wären bei einer zeitlich korrekten Darstellung die negativen Korrekturen von T2-T3 weniger steil als in der Darstellung.



**Abb. 29:** Zusammenfassende Darstellung der z-standardisierten Messzeitpunkte sämtlicher Längsschnittvariablen. Die Werte des BDI und des TV-Konsums wurden umgepolt (weitere Erläuterung siehe Text)

Variablen des Hypothesekomplexes **L1** (psychische Beschwerden/Risikofaktoren) sind rot, die von **L2** (psychische Ressourcen) sind grün und die von **L3** (Körper) sind blau gefärbt. **L4** (TV-Konsum) ist lila dargestellt. Längsschnittdaten, deren Hypothesen zum Interventionseffekt in der ANOVA bestätigt wurden, 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 und damit eine relative Zeitstabilität signalisierten, haben eine gelbe Füllung, die restlichen sind ungefüllt dargestellt.

Deutlich ist erkennbar, dass im GHQ die mit Abstand höchsten standardisierten Mittelwertunterschiede zwischen T1 und T2 gefunden wurden, gefolgt von der PSK, TV-Konsum und BDI. Des Weiteren kristallisiert sich visuell heraus, dass die Messverfahren zu psychischen Beschwerden veränderungssensitiv waren, während die Messverfahren zu psychischen Ressourcen über die Messzeitpunkte relativ unverändert blieben. Die deskriptive Veränderungshöhe der körperlichen Skalen (VAS & KSK) suggeriert, dass diese durchschnittlich über der der Resilienzfaktoren lag.

Des Weiteren zeigt sich, dass die Veränderungsrichtung zwischen T1 und T2 in allen Verfahren zumindest in die hypothetisierte Richtung deutet; Ausschläge unterhalb des Startwertes (Null) fanden im Zeitraum T1-T2 nicht statt.

Auch kristallisiert sich nun auf einen Blick heraus, wie sich die Messverfahren mit einem signifikanten Helmert-Kontrast (GHQ und TV-Konsum) von den restlichen unterscheiden. Nach dem signifikanten Interventionseffekt kommt es für diese zwar zu einer Abwärtskorrektur, allerdings bewegt sich diese weniger stark auf das Ursprungsniveau zurück als bei den Messverfahren, die neben dem bestätigten Interventionseffekt keinen signifikanten Helmert-Kontrast aufwiesen (BDI und PSK). Somit konnte der Helmert-Kontrast sensitiv anhand des Merkmals "relative Zeitstabilität" zwischen den signifikanten Interventionseffekten trennen.

Anhand der Veränderungshöhe wundert man sich fast, warum der VAS keine Signifikanz des Innersubjektfaktors, nebst Helmert-Kontrast offenbarte. Dies ist hauptsächlich der vergleichsweise niedrigen Korrelation zwischen den Messzeitpunkten geschuldet und wird später noch diskutiert.

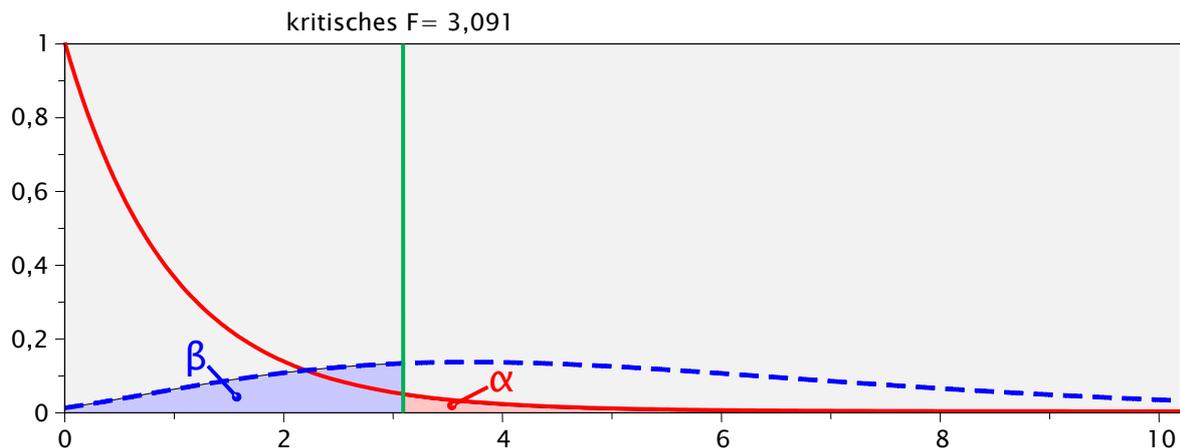
## 4.10. Effektstärken & Poweranalysen der Längsschnittergebnisse

### 4.10.1. Veränderungssensitivität der Messinstrumente

Ziel der folgenden Analysen ist das Erreichen einer Vergleichbarkeit der Sensitivität der Messinstrumente im hier besprochenen Forschungsdesign. Als Sensitivität soll definiert werden, inwieweit der Haupteffekt des Innersubjektfaktors (d.h. sämtliche Messzeitpunkte) eines Messinstrumentes Signifikanz erreichen konnte. Falls für den Innersubjektfaktor die Annahme der Nullhypothese abgelehnt werden konnte, sprach dies für eine Unterschiedlichkeit der Messzeitpunkte. Ein signifikanter Innersubjektfaktor bedeutet nach dieser Argumentationslinie eine nachgewiesene Veränderungssensitivität des entsprechenden Messinstrumentes für die hier verwendete Forschungsfrage.

Eine Möglichkeit, die Sensitivität der Messinstrumente zu vergleichen, könnte darin bestehen, die F-Werte oder die dazugehörigen  $p$ -Werte der Innersubjektfaktoren vergleichend gegenüber zu stellen. Je höher der F-Wert bzw. je niedriger das Signifikanzniveau des Innersubjektfaktors, desto sensibler konnte das betreffende Messinstrument Unterschiede zwischen T1, T2 und T3 quantifizieren. Diese Form des Vergleichs ist bei unterschiedlichem (oder gar inadäquatem) N problematisch und wurde in der Psychologie immer wieder kritisch reflektiert (z.B. Bredenkamp, 1969; Hager, 2006). Da BDI, VAS und TV-Konsum zu einem späteren Zeitpunkt in der Studie eingesetzt wurden, sind bei diesen die N-Zahlen gegenüber den restlichen Testverfahren verringert. Des Weiteren wurde die Depressivitätsdiagnose als Zwischensubjektfaktor in die ANOVA für den GHQ, PSK und TV-Konsum aufgenommen, wodurch deren N-Zahl um die nicht vorhandenen T1-Werte für den BDI (der zur Stellung der Depressivitätsdiagnose vonnöten war) verkleinerte. Da F-Werte und deren  $p$ -Werte vom N beeinflusst werden, ist auf diesem Wege eine „faire“ vergleichende Betrachtung zwischen den Messverfahren nicht gewährleistet.

Eine etablierte Möglichkeit zur Prüfung der Teststärke ist die Poweranalyse nach Cohen (1988). Die Power ( $1-\beta$ ) eines Testverfahrens stellt die Gegenwahrscheinlichkeit des  $\beta$ -Fehlers dar. Beim Hypothesentesten reflektiert der  $\beta$ -Fehler die Wahrscheinlichkeit, 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 fällen.



**Abb. 30:** Beispielhaft dargestellt sind die konkurrierenden Hypothesen zum Innersubjektfaktor der Variable „TV-Konsum“. Die Verteilung der  $H_0$  ("Messzeitpunkte unterscheiden sich nicht") ist rot, die der  $H_1$  ("Messzeitpunkte unterscheiden sich") ist blau gestrichelt dargestellt. Der  $\beta$ -Fehler entspricht der Fläche von  $H_1$ , die beim Unterschreiten des kritischen F-Wertes entsteht. Der  $\alpha$ -Fehler ist auf 0,05 festgelegt. Für den Innersubjekteffekt der Variable „TV-Konsum“ ergibt sich ein  $\beta$ -Fehler=0,26. Damit hat die  $H_1$  eine Power von  $1-\beta=0,74$ .

Innerhalb der multivariaten Varianzanalyse mit Messwiederholung hängt die  $\beta$ -Fehlergröße des Innersubjektfaktors hauptsächlich von folgenden Parametern ab: Gewähltes  $\alpha$ -Fehlerniveau, Innersubjektfaktor-Effektstärke  $f$ , mittlere Korrelation<sup>26</sup> zwischen den Messzeitpunkten  $r_{(T1,T2,T3)}$  und Stichprobengröße  $N$ . Hierbei gilt:

→ Je größer  $\alpha$  und je größer  $f$  und je höher  $r_{(T1,T2,T3)}$  und je umfangreicher  $N$ , desto kleiner wird  $\beta$  und umso größer wird  $1-\beta$ . Da  $\alpha=0,05$  für jedes Messverfahren festgelegt war und die Anzahl der Messungen stets 3 betrug, konnten nur die restlichen Parameter variieren. Die Effektstärken und Poweranalysen für den Innersubjektfaktor wurden nach den Vorgaben von Faul et al. und den Empfehlungen von Rasch et al. berechnet (Faul, Erdfelder, Lang & Buchner, 2007; Rasch, Friese, Hofmann, & Naumann, 2010). Hierbei ergaben sich folgende Werte:

**Tab. 33:** Ergebnisse der Poweranalyse

$p$ : Signifikanzniveau des Innersubjekteffektes;  $Q$ : Anzahl der Messungen;  $K$ : Anzahl der Zwischensubjektfaktorgruppen;  $f$ : Effektstärke  $f$  des Innersubjektfaktors;  $r_{(T1,T2,T3)}$ : mittlere Korrelation zwischen den Messzeitpunkten;  $\eta$ : Effektstärke Eta des Innersubjektfaktors;  $\varepsilon$ : Nicht-Sphärizitätskorrektur;  $1-\beta$ : Power

Innersubjekt Effekt	$p$	$Q$	$K$	$f$	$f$ Cohen Class.	$r_{(T1,T2,T3)}$	$\eta^2$	$\eta$	$\eta$ Cohen Class.	$N$	$\varepsilon$	$1-\beta$
GHQ	0,000	3	4	0,167	Klein	0,745	0,148	0,385	Groß	56	1	0,972
BDI	0,142	3	4	0,079	-	0,715	0,037	0,192	Klein	57	0,89	0,402
PSK	0,026	3	4	0,153	Klein	0,475	0,068	0,260	Mittel	56	1	0,676
SOC	0,296	3	2	0,062	-	0,670	0,017	0,131	Klein	72	0,91	0,251
SWE	0,376	3	2	0,049	-	0,740	0,014	0,118	Klein	72	1	0,219
PSW	0,220	3	2	0,056	-	0,795	0,022	0,150	Klein	69	1	0,321
KSK	0,634	3	4	0,047	-	0,495	0,007	0,083	-	70	1	0,123
VAS	0,196	3	4	0,125	Klein	0,277	0,035	0,187	Klein	50	1	0,342
TV	0,016	3	4	0,128	Klein	0,705	0,083	0,288	Mittel	52	1	0,740

<sup>26</sup> Da Skalenwerte von Korrelationen nicht äquidistant (und damit nicht intervallskaliert) sind, wurden die Korrelationskoeffizienten jedes Messverfahrens 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 erfolgten die Durchschnittsbildungen mit anschließender Rücktransformation derer als jeweiliger durchschnittlicher Korrelationskoeffizient.

Die Effektstärke  $f$  repräsentiert die Effektstärke der Varianzanalyse für Effekte bei unabhängigen Tests (Rasch et al., 2010). Sie wurde berechnet, weil sie zur Hochrechnung der Power bei Variation von  $N$  vonnöten ist. Nach den Konventionen von Cohen (1988, 1992) gelten für  $f$  folgende Klassifikationen zur Beurteilung von unabhängigen Effektstärken:

$f \geq 0,1$ : kleiner Effekt

$f \geq 0,25$ : mittlerer Effekt

$f \geq 0,40$ : großer Effekt

Demnach erreichten die Innersubjektfaktor-Effektstärken  $f$  der Messverfahren GHQ, PSK, VAS und TV-Konsum das Prädikat „kleiner Effekt“, während dies für die restlichen Verfahren nicht mehr attestiert werden konnte. Vor diesem Hintergrund lässt sich anhand der Tabelle erkennen, dass die Power in der ANOVA mit Messwiederholung nicht nur hauptsächlich von der Effektstärke  $f$  und der Stichprobengröße abhängt. Vergleicht man z.B. das  $f$  von VAS und TV-Konsum, sieht man, dass diese (wie deren  $N$ ) zwar von ähnlichem Wert sind, aber für die Variable TV-Konsum in eine mehr als doppelt so hohe Power im Vergleich mit der von VAS resultiert. Das Wirken der Innersubjektfaktor-Effektstärke auf dessen Power wird in der ANOVA mit Messwiederholung entscheidend von der Korrelation der Messzeitpunkte moderiert (Rasch et al., 2010, S. 130). Diese ist im TV-Konsum vergleichsweise hoch (0,705) und im VAS relativ niedrig (0,277), wodurch sich die wesentlich höhere Power der Variable TV-Konsum im Vergleich mit der des VAS erklären lässt.

Im Vergleich zur Effektgröße  $f$  beinhaltet die Effektstärke  $\eta^2$  (Eta-Quadrat) bereits die Korrelation der Messzeitpunkte, womit sie mit Effektstärken aus unabhängigen Prüfmethode nicht mehr vergleichbar ist. Das  $\eta^2$  spiegelt den Anteil der durch den Innersubjektfaktor aufgeklärten Variabilität in der Stichprobe wider. Diese erreicht z.B. im GHQ eine Aufklärung von 14,8%. Diese häufig berichtete Effektstärke überschätzt laut Rasch et al. (Rasch et al., 2010, S.114) allerdings die Größe des Effektes in der Population; der wahre Populationseffekt ist somit stets etwas kleiner als der hier berechnete. Zu Klassifikationszwecken wurde in der Tabelle  $\eta^2$  in  $\eta$  umgerechnet und mit folgenden Konventionen von Cohen (1988, 1992) für die Beurteilung von  $\eta$  bewertet:

$\eta \geq 0,1$ : kleiner Effekt

$\eta \geq 0,24$ : mittlerer Effekt

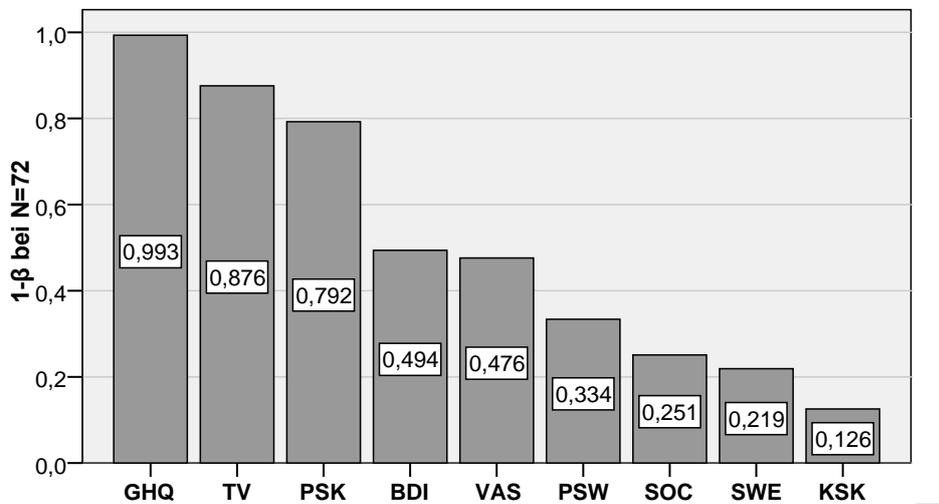
$\eta \geq 0,37$ : großer Effekt

Demnach erreicht der Innersubjekteffekt des GHQ als einziger einen "großen Effekt", der von PSK und TV-Konsum einen "mittleren Effekt" und der von BDI, SOC, SWE, PSW und VAS einen "kleinen Effekt". Der Innersubjektfaktor von KSK liegt unterhalb der Schwelle für einen kleinen Effekt.

#### 4.10.1.1. Analyse für Power bei $N=72$

In einer Hochrechnung der Power nach der Methode von Faul (Faul et al., 2007) wurden die in Tabelle 33 genannten Einflussgrößen außer der Stichprobengröße konstant gehalten und die Power jedes Messverfahren für ein  $N=72$  berechnet. 72 entspricht der maximalen Stichprobengröße, die in dieser Studie für einen Vergleich aller 3 Messzeitpunkte vorhanden

war. Dieses maximale N traf auf den SWE und den SOC zu (vgl. Tab. 33). Nach der Hochrechnung der Power für die übrigen Messverfahren resultierten folgende Teststärken:



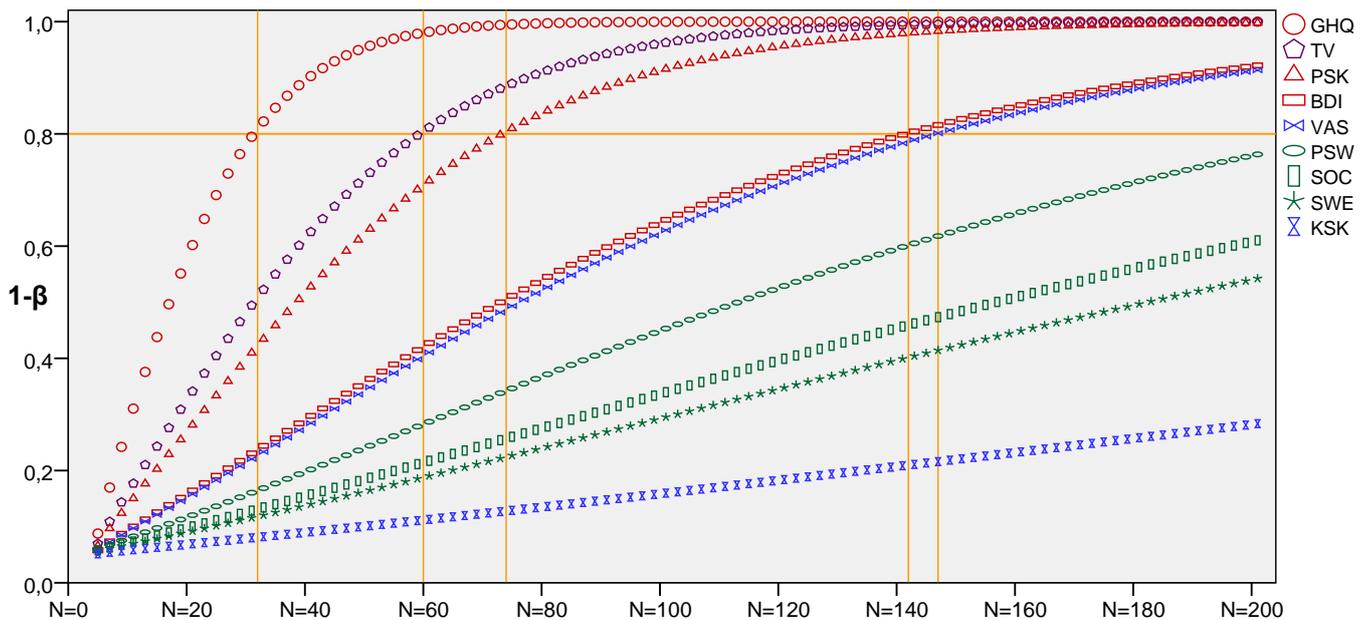
**Abb. 31:** Vergleichende Darstellung der errechneten Power bei einem N=72

Mit der Hochrechnung auf ein identisches N sind die Powerwerte der Messverfahren erstmals miteinander vergleichbar und deswegen ist deren Reihenfolge in der Grafik absteigend geordnet, beginnend mit dem stärksten Messverfahren und endend mit dem am wenigsten sensitiven Verfahren. Bei einem N=72 hätte der Innersubjektfaktor des GHQ mit 99% die größte Wahrscheinlichkeit erzielt, eine Veränderung zwischen allen Messzeitpunkten signifikant werden zu lassen. Die Chance auf Signifikanz des Innersubjektfaktors des TV-Konsums hätte 88% und die der PSK hätte 79% betragen. Am wenigsten sensitiv erwiesen sich die Testverfahren KSK (13%), SWE (22%) und SOC (25%)

#### 4.10.1.2. Optimale Stichprobenumfänge

Nach den Empfehlungen von Cohen (1992) sowie Westermann (2000) sollte der Stichprobenumfang für den statistischen Nachweis eines Effektes so gewählt werden, dass  $1-\beta$  einen Wert von 0,8 nicht unterschreitet. Dieses Kriterium kann in der hiesigen Studie zu den Schlussfolgerungen führen, dass entweder für die Mehrzahl der hier eingesetzten Testverfahren die Stichprobengröße nicht ausreichte oder die „unter-powerten“ Tests den Interventionseffekt nicht ausreichend sensitiv abbildeten, oder, dass der Interventionseffekt für das betreffende Messverfahren schlicht zu klein war, um statistisch gesicherte Unterschiede zwischen den Zeitpunkten T1, T2, T3 zu erzeugen.

Gemäß  $1-\beta \geq 0,8$  sollten Messverfahren, die das Kriterium unterschreiten, in vergleichbaren Untersuchungsdesigns mit identischem N nicht eingesetzt werden, da ihre Wahrscheinlichkeit, ein richtig-positives Ergebnis zu liefern zu unsicher wäre. Von daher wurde für die Instrumente berechnet, wie umfangreich bei einem vergleichbaren Interventionssetting mit identischer Auswertungsmethodik die Stichprobengröße sein müsste, um eine Power von mindestens 0,8 zu erzielen.



**Abb. 32:** Visualisierung der Bestimmung optimaler Stichprobenumfänge für die Innersubjektfaktoren 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 jeweiligen optimalen Stichprobenumfang zum Erreichen des Power-Kriteriums (vgl. Tab. 34 für konkrete Ergebnisse).

Während in der Grafik aus Darstellungsgründen nur die Hochrechnung von  $N=5$  bis  $N=201$  visualisiert ist, wurden die zugrunde liegenden Berechnungen von  $N=5$  in 1er Schritten bis zu einer Stichprobengröße von  $N=1000$  kalkuliert. Hierbei ergaben sich für die Messverfahren folgende optimalen Stichprobenumfänge:

**Tab. 34:** Identifizierte optimale Stichprobenumfänge zum Erreichen des Kriteriums  $1-\beta \geq 0,8$

	GHQ	TV-Konsum	PSK	BDI	VAS	PSW	SOC	SWE	KSK
$1-\beta \geq 0,80$	N=32	N=60	N=74	N=142	N=147	N=219	N=306	N=354	N=740

Entsprechend dieser Hochrechnung hätte der GHQ bereits mit  $N=32$  eine Power von 0,8 erzielt. Fernsehkonsum erreicht ab  $N=60$  dieses Niveau, PSK hätte die Powergrenze ab  $N=74$  erreicht. Wie man sieht, liegen die optimalen Stichprobenumfänge für den TV-Konsum und den PSK höher, als das tatsächliche  $N$  der Studie, in welcher - trotz suboptimaler Stichprobenumfänge - sowohl die PSK ( $N=56$ ) als auch der TV-Konsum ( $N=52$ ) das Signifikanzkriterium erreichte. Ferner resultierten als optimale Stichprobengröße zum Erreichen einer Power von 0,8 für den BDI  $N=142$ , für den VAS  $N=147$ , für den PSW  $N=219$ , für den SOC  $N=306$ , für die SWE  $N=354$  und der KSK hätte laut Hochrechnung 740 Probanden benötigt, um den Haupteffekt des Innersubjektfaktors mit einer Wahrscheinlichkeit von 0,8 signifikant werden zu lassen.

Trotz der erfolgten Ausführungen soll daran erinnert werden, dass sich der jeweilige optimale Stichprobenumfang ausschließlich auf die Sensitivität des Messinstruments in Bezug auf den Haupteffekt des Innersubjektfaktors bezieht, d.h., es wäre laut Hochrechnung für den

optimalen Stichprobenumfang lediglich gewährleistet, dass sich die Messzeitpunkte im Innersubjektfaktor signifikant voneinander unterscheiden. In welche Richtung sich diese Unterschiedlichkeit zwischen den Messzeitpunkten auswirken würde, kann für die Messverfahren, deren Hypothese über die Richtung des Interventionseffektes nicht bestätigt werden konnte (SWE, SOC, PSW, KSK und VAS), eigentlich nur spekuliert werden. Zu Spekulationszwecken sollte man sich die Abbildungen zur Zusammenfassung des Interventionseffektes noch einmal anschauen und die Richtungen nebst Effektstärke der Interventionseffekte examinieren. Anhand der Abbildung und der Effektstärken der Hypothesentests wird klar, dass es vor allem für die Testverfahren SWE, SOC und PSW zweifelhaft wäre, ob ein - durch Erhöhung des Stichprobenumfangs erzielter - signifikanter Haupteffekt des Innersubjektfaktors gleichzeitig die hypothetisierte Richtung bestätigen würde. Im Vergleich dazu verbreitet der ausgeprägtere Kurs der VAS in die erwartete Richtung nebst einer annähernd mittleren Effektstärke eine wesentlich optimistischere Stimmung in dem Sinne, dass - bei Erhöhung des N - ein signifikanter Innersubjekteffekt gleichzeitig die aufgestellte Hypothese zum Interventionseffekt bestätigen könnte.

#### 4.10.2. Schätzung der Effektgrößen der Interventionseffekte

Anhand der Innersubjektkontraste der Messwiederholung T1-T2 wurden die Hypothesen der Interventionseffekte getestet. Obwohl für den Innersubjektkontrast der ANOVA  $\eta$  die maßgebliche Effektgröße darstellt, soll die Effektstärke des Unterschiedes zwischen T1-T2 auch als *Cohen's d* berichtet werden, welche die klassische Effektstärke für paarweise Vergleiche darstellt. Publierte Studien und Metaanalysen verwenden oft *Cohen's d*, wenn eine Effektstärke berichtet wird. *Cohen's d* wird benötigt, um im Diskussionsteil die Effektivität des Gesundheitsförderungsprogramms vergleichend gegenüberstellen zu können. *d* wurde nach den Vorgaben von Faul (2007) berechnet und anhand der Cohenschen Konventionen bewertet.

**Tab. 35:** Effektstärken von Innersubjektkontrasten ( $\eta$ ) und paarweisen Vergleichen (*d*) der Interventionshypothesen;  $p$ : Signifikanzniveau des Innersubjektkontrastes;  $r_{(T1,T2)}$ : Korrelation zwischen den Messzeitpunkten;  $\eta$ : Effektstärke Eta des Innersubjektkontrastes;  $1-\beta$ : Power; T1, T2: geschätzte Randmittel der Untersuchungszeitpunkte, *d*: Effektstärke *Cohen's d* für paarweise Vergleiche

	$p$	$r_{(T1,T2)}$	$\eta_p^2$	$\eta_p$	$\eta_p$ Cohen Class.	N	$1-\beta$	T1	T2	<i>d</i>	<i>d</i> Cohen Class.
GHQ	0,000	0,706	0,255	0,505	Groß	56	0,985	1,890±0,49	2,131±0,43	0,675	Mittel
BDI	0,028	0,766	0,088	0,297	Mittel	57	0,605	58,50±6,69	56,45±9,22	0,345	Klein
PSK	0,021	0,474	0,099	0,314	Mittel	56	0,649	49,11±7,7	51,92±7,27	0,366	Klein
SOC	0,823	0,575	0,001	0,027	-	72	0,056	48,61±8,96	48,83±8,33	0,028	-
SWE	0,596	0,699	0,004	0,063	-	72	0,082	51,38±10,1	51,86±9,40	0,063	-
PSW	0,156	0,830	0,030	0,173	Klein	69	0,293	2,822±0,58	2,878±0,54	0,171	-
KSK	0,436	0,482	0,001	0,096	-	70	0,121	46,72±8,98	47,56±8,61	0,094	-
VAS	0,125	0,077	0,051	0,226	Klein	50	0,334	70,94±22,1	76,84±16,6	0,222	Klein
TV	0,008	0,591	0,136	0,369	Mittel	52	0,767	214,6±106	178,7±98,3	0,386	Klein

In den Effektstärken  $\eta$  nach Cohen (1988,1992) erreichte der Interventionshypothesentest im Innersubjektkontrast des GHQ einen "großen Effekt", die Kontraste von BDI, PSK und TV ergaben jeweils einen "mittleren Effekt" und die von PSW und VAS jeweils einen "kleinen Effekt". Die restlichen Verfahren (SOC, SWE, KSK) lagen noch unterhalb der Schwelle von kleinen Effekten. 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 absicherbar.

Während  $\eta$  höhere Effektstärkeklassen für die Interventionseffekte indizierte, entstehen für die Effekte der paarweisen Vergleiche (*Cohen's d*) niedrigere Effektstärkeklassifikationen<sup>27</sup>. Demnach resultierte für  $d$  im GHQ und TV-Konsum jeweils ein "mittlerer Effekt"; für BDI, PSK und VAS ein "kleiner Effekt" und die restlichen Messverfahren (SOC, SWE, PSW, KSK) lagen noch unterhalb der Schwelle eines kleinen Effektes.

#### 4.10.3. Moderierender Einfluss der Zwischensubjektfaktoren

Die Hypothesen zur moderierenden Wirkung von Zwischensubjektfaktoren auf die Interventionseffekte oder auf die follow-up Effekte wurden als Interaktionen geprüft. In folgender Tabelle sind die Interaktionseffekte der Hypothesen aufgeführt, die mindestens tendenziell ( $p < 0,1$ ) die Längsschnittergebnisse unterschiedlich moderierten.

**Tab. 35:** Moderierender Einfluss von Zwischensubjektfaktoren

$p$ : Signifikanzniveau des Kontrastes im Interaktionseffekt; Q: Anzahl der Messungen; K: Anzahl der Zwischensubjektfaktorgruppen;  $\eta$ : Effektstärke Eta des Innersubjektfaktors;  $1-\beta$ : Power

Kontrast-Interaktionseffekt Innersubjektfaktor * Zwischensubjektfaktor	$p$	Q	K	$\eta^2$	$\eta$	$\eta$ Cohen Class.	N	$1-\beta$
GHQ * Depressivität	0,029	3	4	0,088	0,297	Mittel	56	0,596
BDI * Depressivität	0,022	3	4	0,096	0,310	Mittel	57	0,643
PSK * Depressivität	0,040	3	4	0,079	0,281	Mittel	56	0,545
SWE * Kohorte	0,003	3	2	0,122	0,349	Mittel	72	0,869
KSK * Geschlecht	0,091	3	4	0,043	0,207	Klein	70	0,395
KSK * Sportfortsetzung	0,084	2	2	0,044	0,210	Klein	70	0,409
SWE * Sportfortsetzung	0,089	2	2	0,041	0,202	Klein	72	0,399

Auch in dieser Betrachtung stellt sich heraus, dass nur die Moderatoreffekte mit mindestens mittlerer Effektstärke ihre Wirkung signifikant unter Beweis stellen konnten. Kleine Effektstärken verfehlten durchweg das Signifikanzkriterium.

#### 4.11. Zufriedenheit/ Akzeptanz in der Zielgruppe

Im Rahmen der Prozessevaluation wurden Zufriedenheit und Akzeptanz der Teilnehmer der Studie anhand eines programmbezogenen Fragebogens erhoben. Dabei konnten durch die Studienteilnehmer zunächst alle Einzelinhalte des edukativen Teils („Gesundheit-Basiswissen“) und des praktischen Teils („Gesundheit-Praktische Übungen“) über eine 6-stufige Skala (1= sehr gut; 6=mangelhaft) beurteilt werden. Die überwiegende Mehrheit der Studienteilnehmer bewertete die Kursinhalte mit sehr gut und gut.

<sup>27</sup> Nach Cohen (1988, 1992) indiziert  $d \geq 0,20$  einen kleinen Effekt,  $d \geq 0,50$  einen mittleren und  $d \geq 0,80$  einen großen Effekt.

Erfragt wurden aber auch die individuellen Auswirkungen der Teilnahme auf Aspekte der physischen und psychosozialen Gesundheit. Die subjektive Einschätzung der Teilnehmer auf einer 6-stufigen Skala (1= stimme voll zu; 6= stimme überhaupt nicht zu) fiel auch hier überwiegend positiv aus: 80% der Teilnehmer meinten, dass sie mehr Kraft und Ausdauer hätten, ebenso viele, dass sie jetzt mehr auf ihre Gesundheit achten würden. Bemerkenswert ist auch, dass jeder zweite Teilnehmer zustimmte (oder voll zustimmte), dass sich seine Rückenproblematik verbessert habe. Knapp 70% stimmten z.B. der Aussage zu, dass sie zuversichtlicher geworden sind. Zu den sozialen Aspekten der Teilnahme befragt, äußerten über 75%, dass sie sich aussprechen konnten, 65% der Teilnehmer konnten neue Bekanntschaften knüpfen.

**Tab. 36:** Ergebnisse der Teilnehmerbefragung, N=92

Bewertungskriterium	Bereich	Ausprägung	%
Durch die Maßnahme habe ich mehr auf meine Gesundheit geachtet.	physisch	Stimme voll zu	46,2
		Stimme zu	38,5
		Stimme etwas zu	11,5
		Stimme eher nicht zu	3,8
Ich ernähre mich gesünder.	physisch	Stimme voll zu	25
		Stimme zu	36,5
		Stimme etwas zu	30,8
		Stimme eher nicht zu	5,8
Ich habe mehr Kraft und eine bessere Ausdauer.	physisch	Stimme voll zu	38,5
		Stimme zu	40,4
		Stimme etwas zu	13,5
		Stimme eher nicht zu	7,7
Meine Rückenproblematik hat sich verbessert.	physisch	Stimme voll zu	20,5
		Stimme zu	31,8
		Stimme etwas zu	31,8
		Stimme eher nicht zu	4,5
Ich bin entspannter geworden.	psychisch	Stimme voll zu	36,5
		Stimme zu	32,7
		Stimme etwas zu	28,8
		Stimme eher nicht zu	1,9
Ich bin zuversichtlicher geworden.	psychisch	Stimme voll zu	26,9
		Stimme zu	40,4
		Stimme etwas zu	23,1
		Stimme eher nicht zu	9,6
Ich bin selbstbewusster geworden.	psychisch	Stimme voll zu	17,3
		Stimme zu	42,3
		Stimme etwas zu	23,1
		Stimme eher nicht zu	9,6
Ich habe neue Leute kennengelernt.	sozial	Stimme voll zu	59,6
		Stimme zu	32,7
		Stimme etwas zu	5,8
		Stimme eher nicht zu	1,9
Der Teilnehmerkreis war für mich angenehm.	sozial	Stimme voll zu	48,1
		Stimme zu	44,2
		Stimme etwas zu	5,8
		Stimme eher nicht zu	1,9
Ich habe nette Bekanntschaften knüpfen können.	sozial	Stimme voll zu	23,1
		Stimme zu	42,3
		Stimme etwas zu	26,9
		Stimme eher nicht zu	3,8
Man konnte sich aussprechen.	sozial	Stimme voll zu	38,5
		Stimme zu	36,5
		Stimme etwas zu	17,3
		Stimme eher nicht zu	5,8

## 5. Diskussion

### 5.1. Diskussion der Ergebnisse der Erstuntersuchung

Der hohe Anteil derer ohne feste Lebensgemeinschaft (67% der Männer und 60% der Frauen) reflektiert vermutlich die Ergebnisse von Klein & Strasser (1997), nachdem die Scheidungswahrscheinlichkeit von langfristig Erwerbslosen dreimal höher als in der Allgemeinbevölkerung liegt. Für die Berufskategorien wurde evident, dass geringer Qualifizierte die Stichprobe überwiegen und gleichzeitig den Löwenanteil derer bildeten, die über 5 Jahre arbeitslos waren. Akademiker waren in der Stichprobe unterrepräsentiert und im Vergleich für kürzere Dauer (3-5 Jahre) arbeitslos. Dies steht in Übereinstimmung mit den Befunden von Claussen sowie Klein, die in Bezug auf langfristig Erwerbslose Selektionseffekte im Qualifikationsniveau postulieren (Claussen, 1994; Klein, 1990).

Die Probanden berichteten mehrheitlich, mangelhafte finanzielle Mittel zur Verfügung zu haben; fast alle bezogen die Transferleistung ALG-II. Die Ersparnisse waren aufgebraucht, Ausgaben stark eingeschränkt und jeder Elfte berichtete über Probleme mit Schulden. Die Probanden verbringen durchschnittlich wenig Freizeit mit anderen und die erlebte soziale Unterstützung hatte, weil ihnen berufsbedingte Kontakte fehlten, ein niedriges Niveau. Interessanterweise berichteten sogar die Hälfte der Vpn, die einer stundenweisen Beschäftigung nachgingen („Ein Euro-Job“), dass sie keinerlei berufsbedingte soziale Unterstützung erhielten. Selbst durchschnittlich erreichten die teilzeitbeschäftigten Probanden nur wenig berufsbedingte soziale Unterstützung. Eine derartige Anstellung erzeugt offenbar keine relevanten sozialen Zugewinne.

#### 5.1.1. Diskussion der Fragebogenergebnisse zu T1

Der FPI wies kein besonders auffälliges Persönlichkeitsprofil für die Stichprobe auf. Allerdings lag „Lebenszufriedenheit“ niedrig (Stanine=3,7) und lässt Zweifel am Vorhandensein dieser Einschränkung bei Erwerbslosen verblasen. Diese Abweichung von der entsprechenden Allgemeinbevölkerung war vor allem der signifikant niedrigeren Lebenszufriedenheit der Männer geschuldet. Gegen die Selektionshypothese spricht der Faktor „Emotionalität“, der als Traitvariable als „Neigung zur Depressivität“ (z.B. Hammarstrom & Janlert, 1997) aufgefasst werden könnte. Das Emotionalitäts- bzw. Neurotizismusmerkmal lag mit einem Stanine-Wert von 5,2 recht zentral im Normbereich, so dass sich aus diesem keine erhöhte Depressions-Prädisposition für die Stichprobe ableiten ließ, obwohl mit dem BDI ein fast doppelt so hoher Anteil für unter Depressivität leidende Probanden im Vergleich zur Allgemeinbevölkerung objektiviert wurde.

Der gesamte Interventionsbedarf, der anhand repräsentativer Normstichproben objektiviert wurde, lag vor allem für die Fragebögen, welche psychische Risikofaktoren abbildeten (LEB, GHQ, BDI) mehr als doppelt so hoch wie in der entsprechenden Allgemeinbevölkerung. Für den BDI als etabliertes klinisches Screening-Instrument wurde das metaanalytische Ergebnis von Paul et al. annähernd bestätigt, das über ca. doppelt so viele Personen mit Depressivität in der Gruppe der Arbeitslosen berichtete (Paul et al., 2006). Ressourcenorientierte Fragebögen (SWE, SOC) zeigten dagegen recht unauffällige Werte an.

Warum manifestieren sich psychische Beschwerden so deutlich trotz normaler Ressourcen? In Analogie zur Gesundheitsdefinition der WHO lässt sich folgendes Postulat aufstellen:

„*Psychisches Wohlergehen ist mehr als die Abwesenheit von Risikofaktoren der psychischen Gesundheit.*“, woraus im Umkehrschluss resultieren müsste: „*Psychische Beschwerden sind mehr als die Abwesenheit von Ressourcen der psychischen Gesundheit*“. Man kann unterstellen, dass eine Abwesenheit von psychischen Ressourcen in Form eines eingeschränkten Kohärenzsinn und einer unterdurchschnittlich ausgeprägten Selbstwirksamkeitserwartung mögliche *interne Ursachen* von psychischen Beschwerden repräsentieren. Können in den Ressourcen jedoch keine substantiellen Abweichungen von der Allgemeinbevölkerung - wie in dieser Studie - festgestellt werden, sind zusätzliche, auslösende Faktoren von psychischen Beschwerden als *externe Ursachen* zu suchen. Es ist anzunehmen, dass die externe Ursache, die hierbei am ehesten infrage kommt, in der Erwerbslosigkeit und deren Rahmenbedingungen (Unkontrollierbarkeit der Lebenssituation, Frustration der Selbstbestimmung, gesellschaftliches Stigma etc.) zu suchen ist.

Innerhalb der explorativen Faktorenanalyse konnte gezeigt werden, dass VAS und KSK gemeinsam das körperbezogene Gesundheitsempfinden reflektieren. Körperlich fühlten sich die Probanden im Vergleich zu den Normstichproben in beiden Maßen ca. 70% häufiger als erwartet, eingeschränkt. Der hohe Anteil derer, die sich körperlich in ihrer gesundheitsbezogenen Lebensqualität eingeschränkt fühlen und ein auffällig vermindertes Gesundheitsgefühl aufweisen, muss als Hinweis für einen erhöhten Bedarf an körperlichen gesundheitsförderlichen Interventionen für die Zielgruppe interpretiert werden. Die Stichprobe ist also nicht nur psychisch gravierend beansprucht, sondern fühlt sich auch durch ihre körperliche Gesundheit stark beeinträchtigt.

Auch wenn durch die meisten Risikofragebögen evident wurde, dass die Probanden durchschnittlich viele psychische Beschwerden aufwiesen, fühlten sich die Probanden nicht häufiger durch ihr psychisches Rollenverständnis in ihrer Lebensqualität (PSK) eingeschränkt als die Allgemeinbevölkerung. Dies traf vor allem auf die Männer zu. Eventuell war die Hürde im SF12 Fragebogen für die Items der psychischen Summenskala für die spezielle Gruppe der Versuchsteilnehmer zu hoch. Für die Items der PSK wird z.B. explizit gefragt, ob man aufgrund „seelischer Probleme“ weniger geschafft hätte oder ob man durch seine "seelischen Probleme" seine sozialen Kontakte nicht pflegen könne. Eine zumindest tendenzielle Bejahung dieser Fragen setzt natürlich voraus, dass der betreffende Proband bereits für sich akzeptiert hat, überhaupt psychische Probleme zu haben, welches eine Hürde darstellen könnte. In der PSK könnte sich eventuell aber auch eine verminderte Anforderung des Alltags von Erwerbslosen widerspiegeln; ähnliche Ergebnisse deuteten sich auch bei Sullivan & Karlsson (1998) an.

### **5.1.2. Diskussion der Drop-Out-Ergebnisse**

Die Drop-out Analysen ergaben, dass von T1 ausgehend 79% an T2 und noch 62% an T3 teilnahmen. Die Probandenzahl an T3 setzte 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%). Bedenkt man die besondere Zielgruppe, die neunmonatige Dauer der Studie und den zeitlichen Aufwand, den die Probanden erfüllen mussten, um an allen Wiederholungsmessungen teilzunehmen, sind dies insgesamt noch zufrieden stellende Quoten, wenngleich diese niedriger waren, als erhofft (für T3 wurde ursprünglich mit einer Teilnehmerquote von 66% physisch anwesender Versuchsteilnehmer gerechnet).

Zwischen den Teilnehmern und den Nichtteilnehmern am weiteren Studienverlauf konnten zwischen den T2-Teilnehmern und T2-Abbrechern keine signifikanten Abweichungen in den Daten der Erstuntersuchung ermittelt werden. Teilnehmer an allen 3 Untersuchungsterminen zeigten jedoch hochsignifikant weniger Fernsehkonsum zu T1, wenn sie mit denen verglichen wurden, die nicht an allen Untersuchungsterminen teilnahmen. Dass es zwischen den Studienkomplettierern und den Abbrechern außer in der PSW zu sonst keinen signifikanten Unterschieden in den gesundheitsbezogenen Variablen kam, spricht für die Generalisierbarkeit der Ergebnisse der Längsschnittauswertung. So konnte zusammenfassend nicht festgestellt werden, dass selektiv die Probanden mit ursprünglich stark erhöhtem Beschwerdedruck den Untersuchungen fernblieben. Allerdings ist nicht auszuschließen, dass solche Selektionseffekte bereits vorab entstanden sind, in dem z.B. sehr Depressive erst gar nicht an der Studie teilnahmen. Ob dies der Fall war, lässt sich natürlich nicht ermitteln.

### 5.1.3. Diskussion der Ergebnisse der Sportfortsetzeranalyse

Vor allem die Rolle des ursprünglichen (T1) Fernsehkonsums zeigte in der Sportfortsetzeranalyse einen bedeutenden Einfluss auf die Wahrscheinlichkeit, dass Probanden nach der Interventionsmaßnahme den Sport abbrachen. TV-Konsum erhöhte sowohl im *T1-Modell* (Kapitel 4.8.1.) als auch im *kombinierten Modell* (Kapitel 4.8.3.) die Abbruchwahrscheinlichkeit des Sportes in der hiesigen Untersuchung.

Es ließ sich keine vergleichbare Publikation finden, die sich mit der Aufrechterhaltungswahrscheinlichkeit sportlicher Aktivität bei Erwerbslosen beschäftigt. Insofern wurde Literatur gesichtet, die sich mit den Unterschieden von älteren Menschen, die Sport treiben oder nicht, befasste. Diese stammte hauptsächlich aus dem Kreis der Präventivmedizin, Sportmedizin oder der medizinischen Altersforschung. In den Publikationen wurden fast ausschließlich körperliche Unterschiede geprüft. Wenn die Veröffentlichungen überhaupt psychometrische Maße berichten, dann die SWE oder spezifischere Formen davon (z.B. "sportbezogene SWE"). Studien, die sich explizit mit der Fortsetzungswahrscheinlichkeit eines zu T1 begonnenen Sportes (d.h. prospektiv) für ältere Probanden befassen, sind von der Anzahl her noch stärker überschaubar. Deren Ergebnisse weisen zudem in verschiedene Richtungen.

Resnick analysierte retrospektiv die Daten von älteren Menschen, die angaben, Sport zu treiben oder nicht (Resnick, 2001). Er konnte eine stark ausgeprägte SWE als trennscharfen Faktor für eine höhere Wahrscheinlichkeit, dass ältere Menschen Sport treiben, identifizieren. Ebenso retrospektiv identifizierten Rhodes et al. in ihrem Review, dass eine stärkere SWE, eine „positive Einstellung zum Sport“ und soziale Unterstützung einen erhöhenden Einfluss darauf hatten, ob Personen im Alter noch sportlich aktiv waren (Rhodes et al., 1999). Andere sozio-psychologische Einflussfaktoren werden in ihrer Übersichtsarbeit nicht untersucht; das Review konzentriert sich hauptsächlich auf die körperbezogenen Unterschiede zwischen den Gruppen.

In Bezug auf die prospektive Aufrechterhaltungswahrscheinlichkeit eines begonnenen Sportprogramms für ältere Menschen betont McAuley die Rolle einer initial höheren SWE (McAuley, 1993). McAuley et al. analysierten in einer separaten Publikation für ältere Probanden die Veränderungen von körperlichen Variablen und der SWE nach einem 5-monatigen Training (McAuley, Lox & Duncan, 1993). Sie fanden, dass eine verbesserte SWE

im Gegensatz zu den Veränderungen der körperlichen Variablen den einzigen Faktor darstellte, welcher das Aufrechterhalten des Bewegungsverhaltens im follow-up vorhersagen konnte. Zu einem anderen Ergebnis kommen Stigjelbout und Mitarbeiter (Stigjelbout, Hopman-Rock, Crone, Lechner, & Van Mechelen, 2006). Sie testeten die Aufrechterhaltungswahrscheinlichkeit von mehreren Sportprogrammen für ältere Menschen mit  $N=1.725$  im follow-up und konnten eine initial höhere SWE nur als signifikanten Prädiktor für die Absicht, den Sport fortzusetzen, belegen. In der tatsächlichen Aufrechterhaltung des Sports spielte SWE keine signifikante Rolle. Klonoff et al. fanden eher einen umgekehrten Zusammenhang zwischen ursprünglich negativem Befinden und höherer Aufrechterhaltungswahrscheinlichkeit bei Frauen, denen die Gelegenheit geboten wurde, kostenlos an einem Aerobic Programm teilzunehmen (Klonoff, Annechild & Landrine, 1994). Es stellte sich heraus, dass zu T1 die Variablen „hoher BMI“, „körperliche Beschwerden“ und „Ängstlichkeit“ die Teilnahmewahrscheinlichkeit und Frequenz der Inanspruchnahme des Programms erhöhten. Williams & Lord untersuchten ältere Frauen, nachdem diese an einem 12-monatigen Fitnessprogramm teilgenommen hatten, im 6-monatigen follow-up und belegten, dass eine höhere Depressivität ein signifikanter Faktor für die Fortsetzungswahrscheinlichkeit des Sportes darstellte (Williams & Lord, 1995). Für Emery et al. wiederum stellte die durch eine sportliche Intervention herbeigeführte „reduzierte Ängstlichkeit“ eine signifikante Vorhersagevariable zur erhöhten Aufrechterhaltungswahrscheinlichkeit des Sports dar (Emery, Hauck & Blumenthal, 1992).

Bisherige Ergebnisse zur Wahrscheinlichkeit der Aufrechterhaltung eines begonnenen Sportprogramms sind also keineswegs konsistent und auch die Anzahl der psychologischen Prädiktoren zur Erfassung des Phänomens ist überschaubar und wird zudem von der Verwendung der SWE dominiert. In keiner Studie, die sich explizit mit der Analyse der prospektiven Aufrechterhaltungswahrscheinlichkeit beschäftigte, wurde Fernsehkonsum als prädiktive Variable berichtet. Es ist möglich, dass diese Variable in der bisher publizierten Forschung zur Thematik noch nicht untersucht wurde.

In der hier vorgestellten Studie konnte im Gegensatz zu manchen der obigen Publikationen kein prädiktiver Nutzen der SWE festgestellt werden, weder in der ursprünglichen Messung, noch in den Veränderungsdaten dieses Messinstrumentes. Selbst auf dessen Itemebene fanden sich keine Unterschiede zwischen Sportfortsetzern und -abbrechern. Insofern muss zumindest für ältere Langzeiterwerbslose die SWE als bedeutender Faktor für die Aufrechterhaltungswahrscheinlichkeit eines begonnenen Sports angezweifelt werden. Diese mangelnde Trennschärfe traf aber auf sämtliche psychometrischen Tests in den Untersuchungen zu. Neben Fernsehkonsum ließen sich lediglich mit einzelnen Items und Veränderungsitems der Testverfahren SOC, BDI und SF12 Logit-Modelle zur Vorhersage des Sportabbruch-Risikos konstruieren.

#### 5.1.3.1. Diskussion zur Nützlichkeit der Modelle zur Vorhersage des Risikos Sportabbruch

Das *kombinierte Modell* aus Veränderungsdaten und den Daten der Erstuntersuchung, ist dem *T1-Modell*, das nur die Variablen der ersten Messung enthält, vorzuziehen, da es eine höhere Treffsicherheit (75% vs. 69,5%) und eine größere Varianzaufklärung ( $R^2=0,45$  vs.  $R^2=0,19$ ) bereitstellt. Des Weiteren erzeugt das *kombinierte Modell* eine doppelt so hohe Odds-Ratio für die Risikowahrscheinlichkeit des Sportabbruchs wie das *T1-Modell* ( $OR=9,3$

vs.  $OR=4,4$ ). Inhaltlich bedeutet dies, dass die Kombination der Erstuntersuchung mit den Veränderungsvariablen für die Risikovorhersage einen deutlichen Mehrwert darstellt. Für die Entscheidung des Probanden, den Sport aufrecht zu erhalten oder nicht, kam es offensichtlich darauf an, in welche Richtung die Intervention ihre Wirksamkeit entfaltete. Im kombinierten Vorhersagemodell mit den zwei Veränderungs- und der ursprünglichen Fernsehkonsum-Variable, lässt sich die Reihenfolge des Einflusses anhand der  $p$ -Werte der Koeffizienten ableiten, welches folgende Zusammenhangsreihenfolge produziert: Je weniger sich der Teilnehmende nach dem Training entspannter fühlte und je höher der ursprüngliche Fernsehkonsum war und je weniger er Gefühle eigener Enttäuschung im Laufe der Intervention abbauen konnte, desto größer war das Risiko, dass das Training im follow-up nicht aufrechterhalten wurde.

Dieses Modell ist nützlich, weil es zukünftigen Interventionen einen Hinweis darauf gibt, welche psychologischen Variablen durch die Intervention verstärkt beeinflusst werden sollten, wenn man die Chance erhöhen will, dass Teilnehmer eine initiierte sportliche Aktivität eigenständig fortsetzen. Es lässt sich entnehmen, dass gewonnene Entspannung und Reduktion von eigener Enttäuschung verstärkt werden sollten. Zusätzliche Vermittlung von Entspannungstechniken und kognitive-behaviorale Techniken, welche die Probanden darin anregen, Erfolgserlebnisse durch das Training besser wahrzunehmen bzw. stärker für sich anzuerkennen, könnten hier hilfreich sein.

Aus den Koeffizienten des T1-Modells hätte man sicherlich die Schlussfolgerung gezogen, primär den Fernsehkonsum zu senken und sekundär, dem Gefühl, dass sich die Probanden „als Verlierer“ zu fühlen, entgegenzuwirken. Wie anhand der Suche nach Prädiktoren für das Veränderungsmodell ermittelt wurde, unterschieden sich Sportfortsetzer jedoch nicht von den -abbrechern in der erzielten Reduktion des TV-Konsums. Ebenso wenig führte die Veränderung des Items „Ich fühle mich als Verlierer“ zu Gruppenunterschieden. 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 ist trotz wesentlich kleinerer Odds-Ratio und Varianzaufklärung nützlich. Durch seine alleinige Verwendung von T1 Daten bietet es die Vorteile, dass die dazu notwendigen Variablen einfacher zu erheben sind und es für die Risikoprädiktion auf eine größere Fallzahl zurückgreifen kann. Zudem gibt das Modell durch die ausschließliche Verwendung von T1-Daten wertvolle Hinweise darauf, womit das Risiko des Sportabbruchs von vorneherein erhöht ist und dies war vor allem für den höheren Fernsehkonsum belegbar.

Es kann somit geschlossen werden, dass ein hoher TV-Konsum ältere langfristig Erwerbslose gewissermaßen daran "hindert", ein initiiertes aktiveres Freizeitverhalten aufrechtzuerhalten. Wie die Drop-Out-Analyse in dieser Studie ergab, wirkte der Fernsehkonsum aber nicht nur der Fortsetzung des Sportes entgegen, sondern "hinderte" die Probanden auch an der Komplettierung der Studie. Insofern unterstreicht auch die Drop-Out-Analyse die Wichtigkeit der Erhebung des Fernsehnutzungsverhalten für künftige Interventionsstudien mit Langzeiterwerbslosen.

## 5.2. Diskussion der Längsschnittergebnisse

### 5.2.1. Interventionseffektivität

Der Hypothesenkomplex zu **L1**, der die Erwartung aufstellte, dass die Intervention zu einer Reduktion von Risikofaktoren der psychischen Gesundheit führt, konnte für sämtliche Messverfahren bestätigt werden. Im Kontrast zwischen T1-T2 zeigte sich eine hochsignifikante Abnahme allgemeiner psychischer Beschwerden im GHQ, die sich als großer Effekt manifestierte ( $\eta=0,51$ ), gefolgt von der signifikanten Einschränkungsminderung psychischer Lebensqualität im PSK mit einem mittleren Effekt ( $\eta=0,31$ ). Die signifikante Reduktion von Depressivität im BDI indizierte ebenso einen mittleren Effekt ( $\eta=0,30$ ). Die *durchschnittliche Effektstärke* für die Reduktion von Risikofaktoren der psychischen Gesundheit erreichte  $\eta_p=0,37$  und überschritt sehr knapp die Schwelle zu einem großen Effekt.

Im Gegensatz dazu konnten die Hypothesen von **L2**, in denen erwartet wurde, dass die Intervention sich erhöhend auf die Ressourcen der psychischen Gesundheit auswirkt, sämtlich nicht bestätigt werden. Den besten Kontrast zwischen T1-T2 erreichte hier der (nicht signifikante) Aufbau des psychosozialen Wohlbefindens im PSK mit einem kleinen Effekt ( $\eta=0,17$ ). Der nicht signifikante Anstieg der Selbstwirksamkeitserwartung, gemessen mit der SWE, lag noch unterhalb eines kleinen Effektes ( $\eta=0,06$ ), ebenso, wie der des Kohärenzsinnens im SOC ( $\eta=0,03$ ). Die *durchschnittliche Effektstärke* für den Aufbau von Ressourcen der psychischen Gesundheit lag mit einem  $\eta=0,09$  unterhalb der Schwelle für einen kleinen Effekt.

Der Hypothesenkomplex zu **L3**, welcher die Erhöhung der körperlich wahrgenommenen Gesundheit prognostizierte, konnte ebenso nicht bestätigt werden. Die VAS signalisierte zwar eine Erhöhung des subjektiven Gesundheitsempfindens mit einem kleinen, knapp unter der Grenze eines mittleren, Effektes ( $\eta=0,23$ ), war aber nicht signifikant. Der Anstieg der körperlichen gesundheitsbezogenen Lebensqualität ( $\eta=0,10$ ), gemessen mit der KSK, verfehlte knapp das Prädikat „kleiner Effekt“ und war ebenso nicht signifikant. Es zeigte sich demnach eine *durchschnittliche Effektstärke* von  $\eta=0,17$ , welches einen kleinen Effekt für den Aufbau der subjektiven körperlichen Gesundheit indiziert.

Die Hypothese **L4**, die eine Abnahme des passiven Lebensstils anhand der Reduzierung des Fernsehkonsums vorhersagte, konnte hochsignifikant bestätigt werden. Die dazugehörige Effektstärke ( $\eta_p=0,37$ ) lag ganz knapp unter der Schwelle eines großen Effektes.

### 5.2.2. Relative Zeitstabilität der signifikanten Interventionseffekte

Deskriptiv zeigte sich in allen signifikanten Interventionseffekten im follow-up ein Rückgang der zuvor gewonnenen Verbesserungen. Numerisch betrachtet manifestierte sich demnach in keinem Messverfahren ein absoluter zeitstabiler Effekt. Ein solcher wurde allerdings auch nicht erwartet, da fast alle Interventionsstudien zur psychischen Stabilisierung von Erwerbslosen numerische Rückgänge der signifikanten Trainingseffekte im follow-up berichten, wenn deskriptive Werte angeführt werden.

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

Im Gegensatz dazu war für den BDI und die PSK der Helmert-Kontrast jeweils nicht signifikant; in diesen Messverfahren war demnach unter Einbezug des follow-up der Interventionseffekt nicht mehr nachweisbar. Insofern ist für die Reduktion von Depressivität (BDI) und der psychischen Einschränkungsminderung der Lebensqualität (PSK) von lediglich kurzfristig wirksamen Effekten auszugehen; d.h. sie lösen sich im follow-up wieder auf.

Methodisch gesehen war der Einsatz des Helmert-Kontrasts sehr erfolgreich, das Kriterium "relative Zeitstabilität" zu operationalisieren. Mit ihm gelang es, die signifikanten Interventionseffekte sensitiv in zwei gleich große Gruppen anhand des hier verwendeten Zeitstabilitätskriteriums einzuteilen.

### 5.2.3. Diskussion der Interventionseffekte

Für die Einordnung der Interventionseffekte existieren keine Vergleichsstudien, die ältere Langzeiterwerbslose sportlich trainierten und deren Interventionseffekte mit psychologischen 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 positiv auf das gesundheitliche Wohlergehen von älteren Menschen auswirkt, sondern, ob das Wohlergehen speziell von älteren, langfristig Erwerbslosen auch mit einer primär körperlichen Gesundheitsförderung (statt einer primär psychologischen Förderung) gesteigert werden kann. Insofern muss sich diese Studie mit Interventionen für Erwerbslose messen lassen, welche die Aufgabe hatten, die psychische Gesundheit zu fördern. Für einen direkten Vergleich wäre die Studie „AktivA“ von Rothländer & Richter sehr geeignet. In dieser Studie wurde ebenso in Ostdeutschland (Dresden) dieselbe Zielgruppe (ältere Langzeiterwerbslose) mit einer kognitiv-behavioralen Intervention behandelt und diverse psychologische Maße zu drei Messzeitpunkten erhoben (Rothländer & Richter, 2009). In Bezug zu *AktivA* könnte die hier vorgestellte Studie sozusagen als "alternative Behandlungsmethode" verglichen werden. Leider wurden die Ergebnisse von *AktivA* noch nicht wissenschaftlich publiziert; in den bisherigen Berichten fehlen daher Mittelwerte und Standardabweichungen, so dass sich keine Effektstärken berechnen lassen, die für einen Vergleich vonnöten wären. Die Studie von Proudfoot et al., die eine kognitiv-behaviorale Intervention mit 112, allerdings beruflich hoch qualifizierten, Langzeiterwerbslosen (43 Jahre) durchführten, berichten folgende Werte zwischen den Messzeitpunkten (Proudfoot, Gray, Carson, Guest & Dunn, 1999): Self-esteem: Prä:  $3,67 \pm 1,07$ , Post:  $4,31 \pm 0,89$ . Dies ergibt eine Effektstärke (Methode nach Faul (2007)) von  $d=0,650$ . Für den GHQ (Prä:  $10,02 \pm 7,76$ , Post:  $3,72 \pm 5,81$ ) ergibt sich eine Effektstärke von  $d=0,919$ . Insofern konnten Proudfoot et al. einen mittleren Effekt für self-esteem und einen großen Effekt im GHQ in Ihrer Studie nachweisen.

### 5.2.4. Vergleich mit dem durchschnittlichen publizierten Interventionserfolg

Der für die Vergleichbarkeit womöglich repräsentativste Befund zur Interventionseffektivität stammt von Paul & Moser (2009), die in ihrer Publikation den durchschnittlichen Interventionserfolg für 16 Interventionsstudien, die mit Erwerbslosen durchgeführt wurden, metaanalysierten und als Cohen's  $d$  für abhängige Stichproben angaben:

“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, 2009, S. 276), Hervorhebung durch Verfasser)

Die Autoren verwendeten in ihren Analysen ausschließlich Interventionsstudien für Erwerbslose, in denen mindestens eine von 6 häufig untersuchten Indikatorvariablen zu psychischen Beschwerden im Längsschnitt getestet wurde (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 der Allgemeinheit gewährleisten zu können.

Ein Vergleich unserer Ergebnisse mit dem von Paul und Moser identifizierten durchschnittlichen Meta-Interventionserfolg von  $d=0,35$  sind mit den  $d$ -Werten aus Tab. 35 (Kapitel 4.10.2.) möglich. Grafisch dargestellt verhalten sich die hier vorgestellten Interventionseffekte folgendermaßen mit dem metaanalysierten (publizierten) Interventionserfolg für die Reduktion von psychischen Beschwerden bei Erwerbslosen:

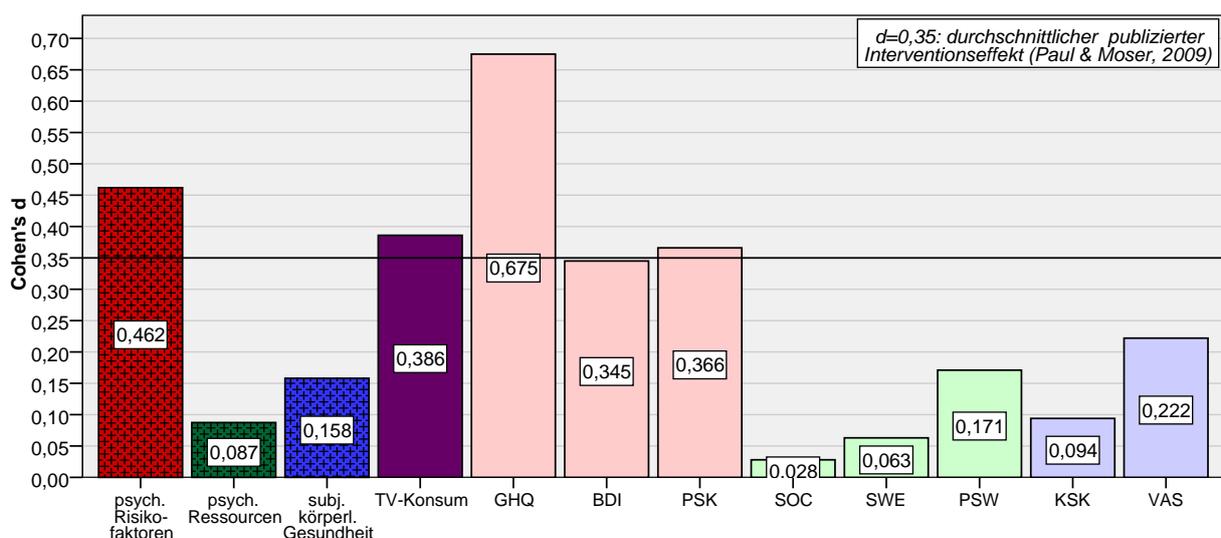


Abb. 33: Fit 50+ Interventionseffekte im Vergleich zum publizierten Meta-Interventionseffekt

Grafik von links nach rechts: Der Meta-Interventionseffekt von Paul und Moser (2009) ist als Linie bei  $d=0,35$  dargestellt. Die durchschnittliche Cohen's  $d$  Interventionseffektstärke der Hypothese **L1**, also von GHQ, BDI und PSK ist in der Grafik als „psych. Risikofaktoren“ (rot,+) dargestellt. Die von Hypothese **L2**: SOC, SWE und PSW ist als Durchschnittswert in „psych. Ressourcen“ (grün,+) zu finden und die von Hypothese **L3**: KSK und VAS ist als mittlere Interventionseffektstärke in „subj. körperl. Gesundheit“ (blau,+) repräsentiert. "TV-Konsum" steht als einzelne Variable für die Hypothese **L4**, welche die Interventionseffektstärke zur Reduktion des passiven Freizeitverhaltens repräsentiert. Im Anschluss an diese 4 Hauptbereiche befinden sich die  $d$ -Effektstärken der einzelnen Messinstrumente der Hypothesenkomplexe L1, L2 und L3 (entsprechende Gruppen-Farbe abgeschwächt dargestellt). Inhaltlich direkt vergleichbar mit dem Ergebnis des Meta-Interventionseffektes von Paul und Moser sind die „psychischen Risikofaktoren“ der Hypothese L1 und deren dazugehörige Messverfahren GHQ, BDI, PSK, weil die Autoren explizit Interventionseffekte zur Reduktion von psychischen Beschwerden metaanalysierten. Die restlichen Werte sind zum Überblick der Effektstärken dargestellt.

Die gemittelte Effektstärke zur Reduktion von psychischen Beschwerden (Risikofaktoren) liegt mit  $d=0,46$  über dem des Meta-Interventionseffektes. Der Interventionseffekt zur Reduktion von unspezifischen psychischen Beschwerden (GHQ) liegt mit  $d=0,68$  annähernd doppelt so hoch wie der Meta-Effekt. Der Interventionseffekt zur Reduktion von Depressivität (BDI) liegt mit  $d=0,35$  auf gleichem Niveau mit der Meta-Effektstärke und ist etwa vergleichbar mit dem Effekt zur psychischen Einschränkungsminderung der Lebensqualität (PSK) von  $d=0,37$ . Bis auf Fernsehkonsum liegen die restlichen Effektstärken darunter, sind aber, wie gesagt inhaltlich nicht mehr mit dem Meta-Effekt von Paul und Moser vergleichbar.

Im Übrigen betonen Paul & Moser, dass ihr metaanalysierter publizierter Interventionserfolg den "wahren" mittleren Interventionserfolg vermutlich überschätzt. Dieser liegt durch den sog. "Publikationsbias" wahrscheinlich unter einem  $d=0,35$ . Weil erfolgreiche Interventionen eher publiziert werden, als gescheiterte Interventionsversuche, erhöht sich der metaanalytisch berechnete Interventionseffekt also um einen Fehler, von dem niemand weiß, wie groß er ist, eben weil die Ergebnisse von missglückten Interventionen oft nicht verfügbar gemacht werden und von daher nicht in die Meta-Effektstärke einfließen konnten.

Insofern bleibt festzuhalten, dass die Studie Fit50+ im Vergleich zum publizierten Meta-Interventionseffekt im BDI und der PSK sehr gut mithalten kann und der GHQ den Meta-Effekt fast um das Doppelte übersteigt. Da der gemittelte Interventionseffekt unserer Studie zur Reduktion von psychischen Beschwerden ( $d=0,46$ ) den Meta-Effekt übertrifft, ist im Vergleich zu den publizierten Interventionsstudien die hier durchgeführte Studie durchschnittlich als etwas erfolgreicher zu bewerten.

### **5.2.5. Diskussion der nicht signifikanten Interventionserfolge**

Es ist schwierig und auch zum Teil fragwürdig, im Nachhinein nach Gründen zu suchen, warum manche erwarteten Interventionseffekte sich nicht zeigten. Trotzdem sollen einige Erklärungsversuche unternommen werden, wobei man um gewisse Spekulationen nicht herum kommt.

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 „Psychosoziales Wohlbefinden“ konnte keine signifikante Änderung zwischen den Zeitpunkten T1-T2 für sich verbuchen. Es lässt sich schließen, dass das Sportprogramm in der Steigerung der Ressourcen zur psychischen Gesundheit wirkungslos war. Warum die Ressourcen kaum eine 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 (16,2%) und unterdurchschnittlichem Kohärenzsinn (17,9%), lag nur etwas über dem Erwartungswert der Normen (15,9%), so dass nicht von einem relevanten Interventionsbedarf auszugehen war. Zusätzlich ist 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. weist zwar darauf hin, dass sich eine Erhöhung des SOC, selbst für ältere Arbeitslose zeigt, wenn diese auf dem ersten Arbeitsmarkt wiedereingestellt werden (Vastamaki et al., 2009). Im Vergleich dazu war der Erfolg der sportbasierten Intervention aber offensichtlich zu schwach, um einen Effekt vergleichbarer Größe herbeizuführen.

Schwieriger ist die Interpretation der nicht bestätigten Hypothesen zum körperbezogenen Gesundheitsempfinden (VAS) und dessen Lebensqualität (KSK). Bereits der gesunde Menschenverstand vermutet für ein primär auf die körperliche Aktivierung aufbauendes Gesundheitsförderungsprogramm auf der subjektiv-körperlichen Ebene gut messbare Effekte. Auch die Anteile von Probanden mit einer unterdurchschnittlichen körperlichen Lebensqualität, gemessen mit dem KSK (26,9%) und der Anteil von Probanden mit unterdurchschnittlichem Gesundheitsempfinden, gemessen mit der VAS (26,3%), lagen relativ zum Erwartungswert (15,9%) hoch, 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, trotz niedriger Korrelation der Messzeitpunkte, die Effektstärke des VAS ( $\eta=0,23$ ) knapp unter dem eines „mittleren Effektes“<sup>28</sup> lag. Ein mittlerer Effekt wird in der Interventionsforschung meist als „relevanter Effekt“ gehandelt. Wenn ein knapp relevanter Effekt in einer Analyse nicht signifikant wird, kann angenommen werden, dass er "unterpowered" erhoben wurde, also die Stichprobengröße in diesem Fall (N=50) nicht ausreichte, um den (durchaus vorhandenen) Effekt der VAS statistisch abzusichern.

Für die KSK lässt sich argumentieren, dass dadurch, dass sie ausschließlich anhand der körperlichen Einschränkungen die dazugehörige Lebensqualität misst, nur ein kleines Veränderungsergebnis resultiert. Dafür kommen folgende Gründe infrage: An anderer Stelle<sup>29</sup> konnte für ein Sub-Sample unserer Stichprobe gezeigt werden, dass die Anzahl der an den Probanden erhobenen medizinischen Befunde mit der KSK zu  $r=-0,49$  ( $p<0,01$ ) zusammenhing. Krankheiten bzw. Gebrechen oder Schmerzen senkten demnach substantiell die körperbezogene Lebensqualität. Während manche der erhobenen medizinischen Diagnosen (z.B. Adipositas, Hypertonie) durchaus einer Linderung durch Sport zugänglich sind, traf dies jedoch nicht auf alle zu. Des Weiteren war der

<sup>28</sup> Mittlerer Effekt für partielles Eta nach Cohen (1988, 1992): Ab  $\eta \geq 0,24$

<sup>29</sup> Preuss M., Kreuzfeld S., Arndt D., Kumar M., Neubert S., Weippert M., Stoll R. *Körperbezogene Lebensqualität und psychologische Faktoren bei älteren Langzeitarbeitslosen*. Arbeitsmed.Sozialmed.Umweltmed. 44, 3 (2009): p.

Interventionszeitraum für eine relevante Verbesserung von bestimmten medizinischen Diagnosen (v.a. Adipositas) möglicherweise zu kurz. Für andere, im Rahmen des Sportprogramms signifikant verbesserte Diagnosen, wie z.B. arterielle Hypertonie, war eine objektiv feststellbare Senkung des Blutdrucks für die Probanden eventuell nicht ausreichend subjektiv erlebbar mit der Reduktion von körperlichen Einschränkungen verknüpft und konnte von daher nicht zu einer Erhöhung der gesundheitsbezogenen Lebensqualität beitragen. Eventuell war die KSK also nicht sensitiv genug für das körperliche Wohlergehen ausgerichtet, welches durch Sport steigerbar ist, oder der Interventionszeitraum für die Erhöhung der körperlichen Gesundheit schlicht zu kurz. Auch Nakagawa et al. fanden für ihr Sportprogramm für ältere Menschen (N=65) keine signifikanten Zugewinne in der körperlichen Summenskala des SF36, lediglich deren Subskala „Körperfunktion“ zeigte einen Hinzugewinn (Nakagawa, Inomata, Nakazawa, & Sakamoto, 2007).

### 5.2.6. Diskussion von Zwischensubjekteffekten

Als bedeutender Zwischensubjektfaktor erwies sich "**Depressivität**" in der Interaktion mit den Innersubjektfaktoren der Instrumente GHQ ( $\eta=0,30$ ), BDI ( $\eta=0,31$ ) und PSK ( $\eta=0,28$ ) in jeweils mittlerer Effektstärke. Die Hypothesen, dass die Depressivitätsdiagnose zu stärkeren mittleren Interventionserfolgen führen wird, bestätigten sich für sämtliche Risikofaktoren der psychischen Gesundheit. Versuchsteilnehmer, die eine Depressivität von klinisch 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 bestätigte sich auch für ältere, depressive Langzeiterwerbslose eine psychisch positive und antidepressive Wirkung von Sport, eine Erkenntnis, die für ältere, nicht zwingend erwerbslose, depressive Menschen ja bereits belegt ist (z.B. van Gool et al., 2007; Kessler et al., 2001; Blumenthal et al., 1999). Des Weiteren können auch die Ergebnisse von Vinokur et al. bestätigt werden, die ebenfalls belegen, dass vor allem die depressiven, erwerbslosen Teilnehmer von der Intervention profitieren (Vinokur et al., 1995).

Im TV-Konsum führte die Depressivitätsdiagnose nicht zu einem differenzierten mittleren Interventionseffekt, aber dafür signifikant zu einem selektiv ungünstigeren follow-up-Ergebnis mit mittlerer Effektstärke ( $\eta_p=0,35$ ). Inhaltlich ist zu interpretieren, dass eine Depressivität nicht, wie prognostiziert, zu einer verminderten Reduktion des TV-Konsums führte. Stattdessen zeigte sich selektiv für die depressiven Probanden, 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 Depressiven im follow-up einen "Rückfall". Vor dem Hintergrund, dass sich der Interventionseffekt zur Verminderung der Depressivität (BDI) als zeitlich instabil erwies, könnte spekuliert werden, dass der Rückfall zu einem passiven Freizeitverhalten als Ausgleich der verlorenen (weil nur kurzzeitig wirksamen) Linderung der Depressivität diene. Eventuell spiegelt der follow-up Effekt auch eine stärkere Abhängigkeit der Depressiven zum Medium TV wider.

Während der Zwischensubjektfaktor "**Geschlecht**" in der VAS nicht zu differenzierten mittleren Interventionseffekten führte, deutete er eine geschlechtsspezifische Wirkung des Ausdauertrainings in der KSK als Trend an ( $p=0,091$ ), der sich als kleiner Effekt ( $\eta=0,21$ ) manifestierte. Die deskriptive Inspektion der Daten ergab, dass isoliert die Männer in der

Erhöhung ihrer gesundheitsbezogenen Lebensqualität profitierten. Ein gruppenspezifischer Test für die Männer zwischen T1-T2 deutete ebenfalls eine tendenzielle ( $p=0,068$ ) Stärkung ihrer körperlichen Lebensqualität im Interventionszeitraum an. Mangels Signifikanz des Interaktionseffektes dürfen die Unterschiede zwischen Männern 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., die für ein Ausdauertraining männerspezifische Steigerungen der körperlichen Lebensqualität bei älteren Menschen nachwiesen (Jette et al., 1996).

Falls sich der Befund in künftigen Studien mit Erwerbslosen bestätigen würde, wäre dies für die präventive Praxis eine relevante Erkenntnis aus folgendem Grund: Männer nehmen Präventionsangebote generell weniger gut als Frauen an (z.B. McQueen et al., 2006)). Das trifft auch auf erwerbslose Männer zu (Mielck, 2005; Berth et al., 2008). Obwohl erwerbslose Männer einen ebenso hohen gesundheitsbezogenen Interventionsbedarf wie erwerbslose Frauen aufweisen, nehmen sie herkömmliche Präventionsangebote zur Stärkung ihrer psychischen (z.B. psychosoziale Beratung) und körperlichen Gesundheit (z.B. Ernährungsberatung) signifikant weniger gut an, als erwerbslose Frauen (Berth, et al., 2008). Wie sich in der Studie Fit50+ zeigte, setzten männliche Versuchsteilnehmer aber häufiger als weibliche die Studie bis zum Ende fort (67% der Männer, 40% der Frauen) und auch im selbständigen Training (follow-up) zeigten sie höhere Fortsetzungsquoten (41% der Männer, 33% der Frauen), so dass zu schließen ist, dass mit dem sportbasierten Gesundheitsprogramm eine Intervention bereitgestellt wurde, die - im Gegensatz zu herkömmlichen Präventionsangeboten - sehr gut von Männern angenommen wird. Vor diesem Hintergrund könnte eine *mögliche* männerspezifische Effektivität auf die körperliche gesundheitsbezogene Lebensqualität besonders interessant für die präventivmedizinische Praxis zur Erhöhung der Gesundheit von Langzeiterwerbslosen sein.

## 6. Prozessdatenmanagement

### 6.1. Zusammenfassung

Das finale Projektergebnis des IT-Fachanteils in Fit50+ kennzeichnet ein medizinisches Probanden- und Prozessdatenmanagement für den Einsatz in der Funktions- und Leistungsdiagnostik.

Das Center for Life Science Automation (celisca) hat im Zusammenwirken mit dem IPM u.a. eine für das Projekt relevante Expertise auf dem Gebiet der Datenerfassung, -speicherung und -visualisierung von Messdaten aus dem Life Science Bereich. Durch Nutzung von Vorarbeiten eines in weiten Teilen applikationsneutralen Laborinformationsmanagementsystems (Kürzel openLIMS) konzentrierten sich die IT-Infrastrukturarbeiten auf die prozess- und methodenspezifische Datenverarbeitung und -visualisierung (AP 0 – Dateninterfacegenerierung).

Eine Web-gestützte Datenbank mit implementiertem Zugriffsrechtmanagement ermöglicht einen Zugriff auf die Erfassungs- bzw. Messergebnisse in 2 Modi: Individualmodus zur Visualisierung der persönlichen Daten des Probanden in verständlicher Form und der Expertenmodus zur Datenvisualisierung für den untersuchenden bzw. den beratenden Arzt. Das geschaffene IT-Instrument ist von allgemeingültiger Bedeutung für die Erfolgskontrolle von Präventionsmaßnahmen dieser Art.

Die notwendigen initialen IT-Arbeiten in AP0 konnten frist- und anwendungsgerecht umgesetzt werden. Unter Einbeziehungen von Vorarbeiten gelang dem Projektteam die rasche Erschließung einer leistungsfähigen webbasierten IT-Infrastruktur mit teils dezentraler medizinischer Informationserfassung und teilautomatisierter Messdatenakquise.

Die dem Projekt angemessenen Mittel- und Methoden des Datenschutzes und der Datensicherheit folgten dem fortgeschrittenen relevanten Stand der Technik. Die der IT-Lösung zugrundeliegende Client-Serverarchitektur berücksichtigt u.a. ein automatisiertes Pre- und Postprocessing, auch unter Einbeziehung verfügbarer KI-Algorithmen. Nachfolgend werden auszugswise konzeptionelle Ansätze sowie entstandene IT-Infrastrukturlösungen für die Projektumsetzung Fit50+ hervorgehoben.

### 6.2. Framework-System openLIMS

Aufgrund der sehr unterschiedlichen Aufgabenstellungen und Struktur von Forschungsprojekten und ihren spezifischen Anforderungen, muss ein entsprechend universelles Abbildungssystem für die Datenerfassung den Forderung nach Generierbarkeit und Flexibilität folgen. Die während einer präventivmedizinischen Untersuchung erhobenen Daten reichen von physiologischen Parametern und Messwerten bis hin zu subjektiven Einschätzungen und objektiven Bewertungen. Diese Verschiedenartigkeit der Daten bedingt projektspezifische Informationsmanagementsysteme, die einen nicht unbeträchtlichen Mehraufwand erfordern. In der Regel ist in IT-Lösungen fest vorgegeben, welche Daten erfasst werden können. Die Hinzunahme neuer Informationsklassen ist mit großem Aufwand verbunden und bedingt oft die Mitwirkung des Systemanbieters. Defizite bestehen dabei vor

allem in der Ausnutzung vorhandener IT- und Kommunikationstechnologien hinsichtlich generischer Systemansätze.

Das am Center for Life Science Automation entwickelte Informationssystem openLIMS erfüllt die Forderungen nach höchster Flexibilität, Skalierbarkeit und Anpassbarkeit mit einem dem Informationssystem zugrundeliegenden Frameworkkonzept, das auf einem Parametermodell mit weit reichenden Datentypen und einer neunstufigen Prozessabbildungsstruktur basiert. Es erlaubt, Projekte und ihre vernetzten Prozesse in einem beliebigen Detaillierungsgrad abzubilden. Der Anwender erschließt sich das skalierbare System bedarfsgerecht über eine geeignete Prozessabstraktion und eine anpassbare Systemkonfiguration. Das Integrationskonzept führt verteilte automatisierte und manuelle Prozesse ohne Medienbruch in beliebiger Struktur und Abstraktion so zusammen, dass stets von einer konsistenten Datenspeicherung gesprochen werden kann.

Der Begriff des Parameters dient in diesem Frameworksystem der Beschreibung von beliebigen Hierarchiestufen der abgebildeten Prozesse durch Daten über Inhalt und Eigenschaften, wobei alle relevanten Datentypen und auch vernetzte Informationsklassen Berücksichtigung finden (vergl. 12 Parametertypen in Abbildung 1). Parameter sind dabei nicht nur die Container für Stammdaten, Prozessgrößen, Prozessvariablen oder Dokumentationselemente von Prozessen, sondern auch für Ein- und Ausgangsbedingungen oder charakterisierende Prozesszustände, die z. B. eine schnelle Prozessinterpretation durch Automaten oder durch manuelle Prozessbeobachtung erlauben.

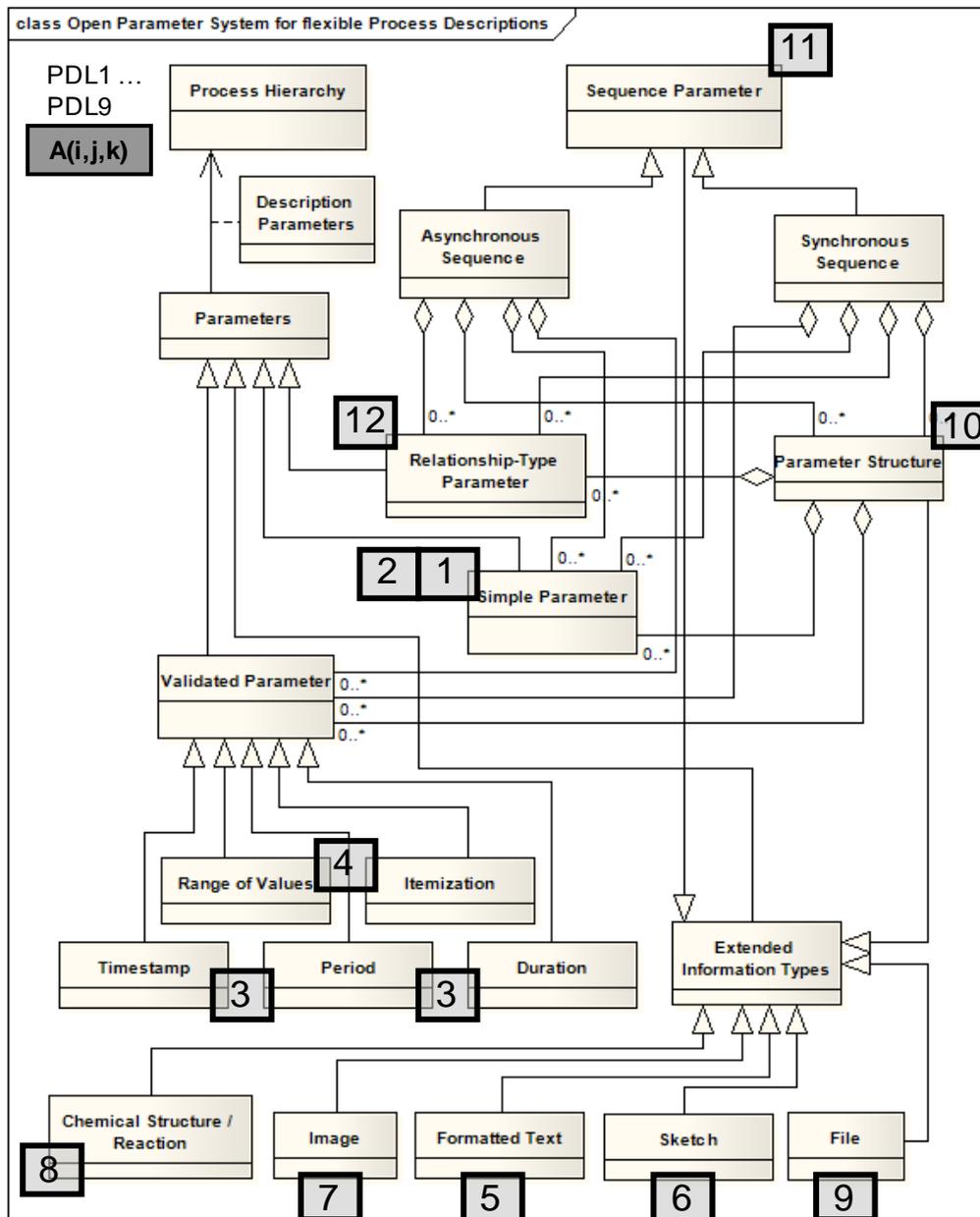


Abb. 34: UML-Klassendiagramm des Parametersystems zur Aufnahme beliebiger Untersuchungsdaten

Die freie Bereitstellung "strukturierter Parameter" ermöglicht es, verschiedene Parameter zu einem zusammenhängenden Datencluster frei zu bündeln (beispielsweise für die Kombination von einem Messwert und der verbalen Bewertung "erhöht"), sowie die Unterstützung von "Zeitreihen" für die Aufnahme von zeitbezogenen Daten, z.B. für den Verlauf von Herzschlag- und Atemfrequenz während einer definierten Belastung.

Das Informationssystem gestattet durch entsprechende Dateninterface-Generierung die flexible, zentrale Ablage von Stammdaten der Probanden und weiterer Bestandteile einer elektronischen Probandenakte für variable, sich fortentwickelnde Forschungsziele sowie die Verknüpfung mit Beschreibungen aller durchgeführten Untersuchungen und ihrer Ergebnisse. Insbesondere für die ad-hoc-Erfassung standardisierter oder methodenindividueller Fragebögen wurde die IT-Lösung openLIMS um einen leistungsfähigen Formulargenerator ergänzt. Daten von drei stationären Untersuchungsgeräten können auf

verschiedenen konfigurierbaren Wegen über standardisierte Schnittstellen (XML Excel, csv, Datenbanken) importiert bzw. die vorhandenen Prozessdatenbanken direkt angebunden werden (Intranet-Online-Kopplung).

Entwickelte Interfaces zu Datenfiles mobiler Sensorik (upload-Lösung) unterstützten die Datenerfassung am Standort Schwerin. Die Erfassung zeitbasierter physiologischer Parameter in das Informationssystem stellt nicht nur wegen des damit erfassten umfangreichen Datenvolumens eine Herausforderung dar. Die angebundenen Rohdaten müssen überwacht, akquiriert, selektiert, visualisiert, verarbeitet und über komfortable Schnittstellen den Datenanalyse- und Statistiktools bereitgestellt werden.

Einen besonderen Adaptionsbedarf erforderten die nach Berechtigungen und Funktionen zu differenzierenden User-Interfaces der kollaborativen IT-Plattform.

### **6.3. Prozess-Abbildung unter Nutzung des Frameworksystems**

Beschreibungs-, Erfassungs- bzw. Messparameter können durch den Anwender bedarfsgerecht definiert und in so genannten *"Untersuchungsschritten"* (Prozessaktivitäten) und *"Untersuchungsabläufen"* in einer Methodenbibliothek zusammengestellt werden, z.B. auch zur Abbildung von umfangreichen Fragebögen, die sich aus einzelnen Abschnitten zusammensetzen. Abbildung 35 zeigt die Gesamtstruktur des generierten projektspezifischen Prozessabbildungssystems für Fit50+. Die Prozessabbildungsgenerierung umfasste mehr als 200 Beschreibungsparameter für mehr als 150 Prozessschritte für alle Teiluntersuchungen. Die oberste Strukturierungsebene des Frameworkansatzes openLIMS ist das *"Projekt"*, das in *"Studien"* unterteilt werden kann. Einer Studie werden die *"Untersuchungen"* zugeordnet. Die *"Teiluntersuchungen"* werden mit den *"Probanden"* gemäß der definierten *"Untersuchungsabläufe"* durchgeführt. *"Untersuchungsabläufe"* werden durch *"Untersuchungsschritte"* (beschriebene Aktivitätenmenge A in Abbildung 34) unterteilt, die wiederum durch die Interfacegenerierung *"Untersuchungsparameter"* gemäß der Parametertypen in Abbildung 34 definiert werden können.

Auch die Interfacelösung Datenakquise konnte weitgehend auf der Grundlage der generischen Systemlösung openLIMS realisiert werden. Hervorzuheben sind die Arbeiten zu den Kommunikationsinterfaces der eingesetzten stationären und mobilen medizinischen Untersuchungstechnik.

Die automatisierte /teilautomatisierte Übernahme von verteilten Daten externer Systeme (z.B. beliebige Parameter/ Datensätze von Systemen der stationären sowie mobilen Datenerfassung) erfolgt über ein Kommunikationsframework, das typische Schnittstellen wie Text-, XML- und Excel-Dateien sowie Prozessdatenbanken und Web Services unterstützt. Durch Anwendung eines Syntax- und Semantikkonverters können die eingelesenen Datensätze den Untersuchungsparametern korrekt in der Zielcodierung zugeordnet werden. Ausgehend von der projektspezifischen Prozessabbildung entstanden Userinterfaces der selektiven Prozessdatenvisualisierung, Datenergänzung sowie zur Datenauskopplung.

Zielstellung des Projekts fit 50+ ist die Wirksamkeitsanalyse eines Programms zur Förderung der physischen und psychischen Gesundheit/ Leistungsfähigkeit älterer Langzeitarbeitsloser. Im Laufe des Projekts wurden daher die 120 Probanden des Förderprogramms dreimal untersucht und deren Gesundheits- und Ernährungszustand bestimmt, z.B. durch:

- die Befragung zu bestehenden und früheren Erkrankungen (Anamnese, diverse Fragebögen in Papierform sowie digital erfasst)
- die Bestimmung von Ernährungszustand und Körperzusammensetzung (Anthropometrie, Messwerte automatisiert digital erfasst)
- die Bestimmung der körperlichen Leistungsfähigkeit durch Belastungsuntersuchung auf dem Fahrrad mit Aufzeichnung von Blutdruck, Herzschlagfrequenz und Herzrhythmus sowie verschiedenen Kenngrößen der Atmung, Bestimmung von Laktat (Spiroergometrie, Messwerte automatisiert digital erfasst).

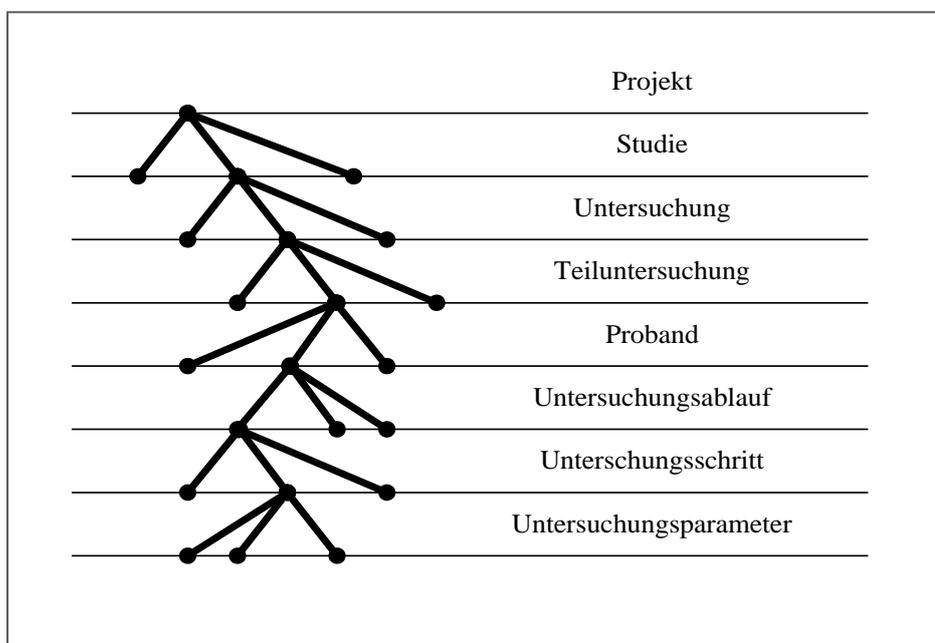


Abb. 35: Abbildungshierarchie der Untersuchungen

Alle Untersuchungsprozesse sind vollständig mittels der IT-Infrastrukturlösung dokumentiert. Für jeden der drei Untersuchungsblöcke (Erstuntersuchung sowie zwei Nachuntersuchungen) wurde eine separate Untersuchung angelegt und jeweils beispielsweise den Teiluntersuchungen "Anamnese", "Anthropometrie" und "Spiroergometrie" zugeordnet. Grundlage hierfür bilden die in der Methodenbibliothek zunächst erstellten Untersuchungsschritte und Untersuchungsabläufe. Abbildung 36 zeigt beispielhaft einen Auszug der Teiluntersuchung "Anamnese". Die dezentral gewonnenen Messwerte u.a. der Anthropometrie und Spiroergometrie werden durch den Einsatz des Kommunikationsframeworks automatisiert den Teiluntersuchungen zugeordnet.

## 6.4. Systemarchitektur

Die Architektur des verteilten Systems aus Messdatengewinnung, Rohdatenspeicherung, Preprocessing und Prozessdateneinbindung in die Prozessabbildung openLIMS geht aus Abbildung 37 hervor. Drei Daten- und Kommunikationsinterfaces existieren zur Projektumsetzung:

1. Die eingangsseitige Prozesskommunikation zur Datenerfassung (mobile Datenerfassung mit Pulsuhren, Interfaces zu stationärer Messtechnik mit zeitweisem Online- bzw. Offline-Betrieb)
2. Die Infrastruktur zum variablen Preprocessing (im Beispiel über verteilte Standard Webservices)
3. Die Prozessdatenkommunikation mit dem medizinischen Informationssystem (Applikationsentwicklung mit generischem LIMS)

Die verteilten Rechnerressourcen mit Servereigenschaft (Netzwerkknoten) gliedern sich demnach in einen oder mehreren Rohdatenservern, die Prozessdatenbanken als logische Speicherschicht nutzen, dedizierte Netzwerkknoten, die auf Anwendungsebene Protokolle für Fern-Prozeduren bereithalten sowie den Web-Applikationsserver **openLIMS**.

**Fit50+U1 Fragebogen**

Fit50+U1 Fragebogen (Fit50+U1 Fragebogen)

Basierend auf Ablauf: → 14 (Fit50+U1)

→ alle Schrittparameter anzeigen

← 0) Plattenpositionsdetails

Proband: (1) → KIHE26.07.1956M  
Untersuchungstag-ID: (1) 070904KH  
Untersuchungstag: (1) 04.09.2007

← 1) Blutdruck/Herzfrequenz

**RR Ruhe systolisch (1)**

RR Ruhe systolisch 1. Wert (1)	122	mmHg
RR Ruhe systolisch 2. Wert (1)	121	mmHg
RR Ruhe systolisch 3. Wert (1)	129	mmHg
RR Ruhe systolisch Durchschnittswert (1)	124	mmHg

**RR Ruhe diastolisch (1)**

RR Ruhe diastolisch 1. Wert (1)	86	mmHg
RR Ruhe diastolisch 2. Wert (1)	81	mmHg
RR Ruhe diastolisch 3. Wert (1)	87	mmHg
RR Ruhe diastolisch Durchschnittswert (1)	84	mmHg

← 2) Sportanamnese

regelmässiger Sport (1) ja

Regelmaessiger Sport (=> 30 min/d Ausdauersport) (1) 1-2x/Wo

Gartenarbeit (1) nein

Andere Sportart (1) 1-2x/Wo

Andere Sportart (Kommentar) (1) Radfahren, Spazieren

→ 3) Eigenanamnese

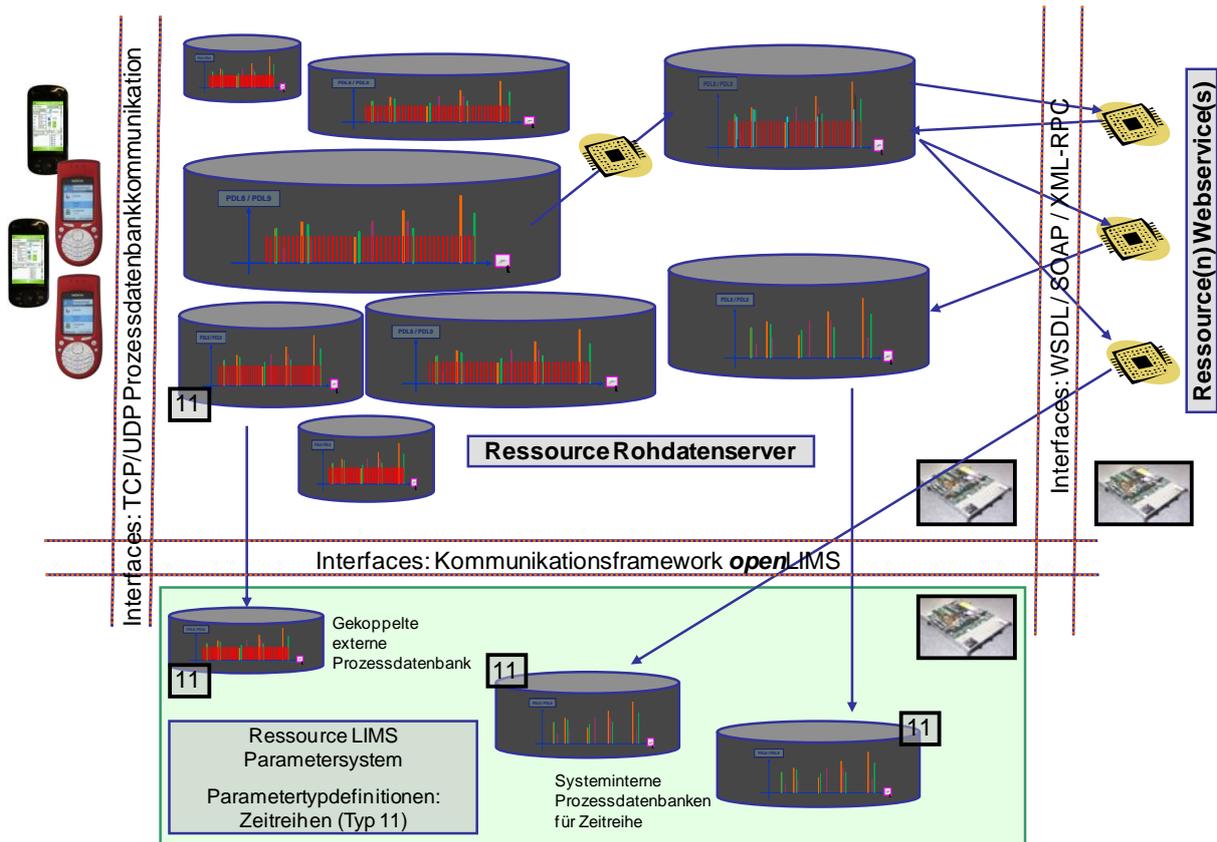
→ 4) Familienanamnese

→ 5) Medikamente

Abb. 36: Auszug aus dem Anamnese-Online-Fragebogen

## 6.5. Spezifische Daten-Verarbeitung

Die aufgezeichneten sensorspezifischen Prozessgrößen können als Rohdaten mit Messfehlern behaftet sein. Um diese Daten weiterverarbeiten und daraus neue wissenschaftliche Erkenntnisse gewinnen zu können, müssen sie fehlerbereinigt, vervollständigt und eventuell gefiltert werden. Dabei werden spezielle Algorithmen zur Bereinigung sensorspezifischer Fehler und Fehlerkennungen benötigt.



**Abb. 37:** Architektur des verteilten Systems zeitbasierter Prozessdaten

Des Weiteren werden in den präventivmedizinischen Studien verschiedene Algorithmen und statistische Funktionen eingesetzt, um neue Prozessgrößen aus den Sensor-Rohdaten zu gewinnen oder die Daten zu verdichten (z. B. Reduktion der Auflösung durch Mittelwertbildung). Diese aggregierten Daten ergänzen dynamisch die Prozessdokumentation (Probandenakte). Auch umfangreichere Methoden, wie die Modellierung individueller oder gruppenspezifischer Profile und deren Nutzung zur Vorhersage von Parametern sind denkbar. Ein effizienter Weg, beliebige Algorithmen in das webbasierte Informationssystem einzubinden, ist deren Realisierung als Webservice.

Werden Algorithmen in einer Studie routinemäßig durchgeführt, ist eine Automatisierung und Auslagerung des jeweiligen Preprocessings sinnvoll, so dass das Informationsmanagementsystem nur noch zur Aufrechterhaltung der Datenkonsistenz auf die Rohdaten zugreifen muss, zur Auswertung, Analyse und Exploration aber nur bereinigte und berechnete Daten verwendet werden.

## 6.6. Datenanalyse und Exportinterfaces

Verschiedene in das Informationssystem integrierte Visualisierungsmethoden erlauben einen schnellen Überblick über nominale, numerische und Multiparameterdaten. So können interessante Aspekte, Ausreißer oder Tendenzen innerhalb der Probandengruppe schnell identifiziert und gegebenenfalls darauf reagiert werden.

Das eingerichtete Informationssystem fasst alle zur Durchführung und Auswertung der relevanten Daten „Fit50+“ aus den verschiedenen Quellen zusammen. Für eine Auswertung können die Daten selektiert und über standardisierte Schnittstellen (csv, Excel, XML) dem bevorzugten Analysewerkzeug wie z. B. SPSS oder Spotfire DecisionSite zur weiteren Bearbeitung bereitgestellt werden.

## 6.7. Zugangsplattform

Das finale Ergebnis von AP0 spiegelt sich in den selektiven Dateninterfaces für Probanden und Experten wider. Hier werden zugangsbegrenzt Auszüge der dynamischen Untersuchungsprozessdokumentation, einschließlich generierter Daten des Postprocessing webbasiert bereitgestellt. Die nachfolgenden beiden Abbildungen 38 und 39 vermitteln exemplarisch einen Eindruck zu den Anwendungsfällen und den potenziellen Anwendungsmöglichkeiten.

Der Expertenzugang und der Probandenzugang erfolgen über eine gemeinsame Applikationsschnittstelle. Das Generierungsmodul zum Dateninterface berücksichtigt einen passwortgeschützten Zugang sowie SSL/HTTPS-Übertragungsprotokolle. Der Expertenzugang ermöglicht grundsätzlich die Sichtung und Auswertung aller Untersuchungsprozesse/ Teilprozesse aller Probanden. Dazu wird die Prozessabbildungsebene „Untersuchungen“ freigegeben. Abbildung 38 zeigt das Zugangsszenario eines ausgewählten Probanden mit insgesamt 15 differenzierten Fragebögen. Demgegenüber besteht das Dateninterface des Probandenzugangs gemäß der Abbildung 39 aus einem selektiven Lesezugang zu Untersuchungsinformationen der betreffenden Anmeldung. Datenexporte sind in dieser Sicht nicht vorgesehen. Allerdings eröffnen die Realisierungsbeispiele interessante Entwicklungspotenziale der Online-Probandendialoge unter Einbeziehung von Experten.

celisca - health guard <Ft 50+ Experte> <Verwaltung von Projekten, Studien etc.> <virtuelles Labor> Mittwoch, 30.06.2010

Projekte  
Studien  
Versuchsreihen  
→ Untersuchungen  
Probanden

Versuche mit folgendem Status auswählen: in Planung in Bearbeitung beendet **alle Versuche**

Versuch-ID: 15  
 Nummer: Ff50+U1 Fragebogen  
 Name: Ff50+U1 Fragebogen  
 Kurzbeschreibung:  
 Ziel:  
 Kommentar:  
 angelegt von: → Administrator,  
 angelegt am: 05.12.07  
 durchgeführt am: 05.12.07  
 Status: in Planung

Direktsuche (\* ist Wildcard):

Projekt: ff50+ (ff50+)  
 Studie: ff50+ (ff50+)  
 Versuchsreihe: ff50+U1 (ff50+U1)

Versuche: 3 Suchergebnisse  
 Ff50+U1 Anthropometrie (Ff50+U1 Anthropometrie)  
 Ff50+U1 Ergospirometrie  
 → Ff50+U1 Fragebogen (Ff50+U1 Fragebogen)

Planungsdaten  
 Versuch: Ff50+U1 Fragebogen (Ff50+U1 Fragebogen) / Platte Nr:33 / Plattenposition 49  
 Basierend auf Ablauf: → 14 (ff50+U1)

→ alle Schrittparameter anzeigen

1) Blutdruckmesswert  
 RR Ruhe systolisch (49) → strukturierter Parameter  
 RR Ruhe diastolisch (49) → strukturierter Parameter  
 HF Ruhe (49) → strukturierter Parameter  
 fertig (49) ja

2) EKG-Befund  
 → 3) Eigenanamnese  
 → 4) Familienanamnese

5) Medikamente  
 MED1 (ACE-Hemmer) (49) nein  
 MED2 (Blöcker) (49) nein  
 MED3 (Diuretika) (49) nein  
 MED4 (Ca-Antagonisten) (49) nein  
 MED5 (Lipidsenker) (49) nein  
 MED6 (Insulin) (49) nein  
 MED7 (Orale Antidiabetika) (49) nein  
 MED8 (SD-Hormone) (49) nein  
 MED9 (Thyreostatika) (49) nein  
 MED10 (inhalative Cortikosteroide) (49) nein  
 MED11 (Beta-Sympathomimetika) (49) ja  
 MED12 (Antidepressiva) (49) nein  
 MED13 (Düchvorbeugung) (49) nein  
 MED14 (Nitrat) (49) nein  
 MED15 (Andere) (49) nein  
 MED15\_KOMM (Andere Kommentare) (49) fertig (fertig) (49) ja

6) Sportanamnese  
 7) sozialer Status  
 8) wirtschaftliche Situation 1  
 9) Fragebogen GNG 12  
 10) Psychosoziales Wohlbefinden  
 11) Soziale Unterstützung (StreuVohr)  
 12) wirtschaftliche Situation 2  
 13) Fragebogen Freizeit  
 14) Fragebogen Ernährung  
 15) Fragebogen DME

weitere Daten  
 berechtigter Kontakt: → Aclr09051955W, Proband, Ft 50+ (Tabelle: berechtigter Kontakt)  
 berechtigter Kontakt: → BalMo20061951W, Proband, Ft 50+ (Tabelle: berechtigter Kontakt)

Versuchsdurchführung  
 120 Probanden (Nummer: 33 [Platte Nr:33]) (in Planung)  
 → Administrator, I in Bearbeitung

verwendete Abläufe:  
 1) Ff50+U1 (blau) (→ Dateien indizieren)

(1)	(11)	(21)	(31)	(41)	(51)	(61)	(71)	(81)	(91)	(101)	(111)
(2)	(12)	(22)	(32)	(42)	(52)	(62)	(72)	(82)	(92)	(102)	(112)
(3)	(13)	(23)	(33)	(43)	(53)	(63)	(73)	(83)	(93)	(103)	(113)
(4)	(14)	(24)	(34)	(44)	(54)	(64)	(74)	(84)	(94)	(104)	(114)

Expertenzugang

Abb. 38: Dateninterface Expertenzugang (untersuchender / beratender Arzt)

celisca - health guard celisca - health guard - Probandenzugang - Login

celisca - health guard celisca - health guard - Probandenzugang - Account: Proband BalMo20061951W

→ Untersuchungen

Login-Name Passwort GO

Untersuchungsdaten 30.06.2010

Projekt: ff50+ (12.09.07)  
 Studie: ff50+ (12.09.07)  
 Versuchsreihe: Nr: ff50+U1 / Name: ff50+U1 (12.09.07)  
 Untersuchung: Nr: Ff50+U1 Anthropometrie / Name: Ff50+U1 Anthropometrie (21.01.08)  
 Basierend auf Ablauf: Durchführung der Anthropometrie mit HRV  
 Durchführer: Administrator,  
 Status: beendet  
 Proband: BalMo20061951W - Pos: 18

alle verfügbaren Untersuchungen:  
 Ff50+U1 Anthropometrie (Ff50+U1 Anthropometrie)  
 Ff50+U1 Ergospirometrie (Ff50+U1 Ergospirometrie)  
 Ff50+U1 Fragebogen (Ff50+U1 Fragebogen)  
 Ff50+U2 Anthropometrie (Ff50+U2 Anthropometrie)  
 Ff50+U2 Ergospirometrie (Ff50+U2 Ergospirometrie)

verwendete Abläufe in der Untersuchung:  
 Durchführung der Anthropometrie mit HRV

1) Durchföhrung der Anthropometrie  
 Datenbank (Anthropometrie): 070921BM  
 TESTDAT: 21.09.2007  
 Oresse: 152 cm  
 Gewicht: 73.4 kg  
 SB: 34.45 cm  
 BB: 20.07 cm  
 BSB: 27.55 cm  
 BT: 20.64 cm  
 UAU: 25.2 cm  
 HU: 18.4 cm  
 Waaft: 64.8 cm  
 Hjp: 112.1 cm  
 OSU: 60.4 cm  
 Tragus: 13.6 mm  
 Kinn: 16.9 mm  
 Achael: 33.7 mm  
 Trizeps: 31.8 mm  
 Biceps: 24.4 mm  
 Rücken: 27.6 mm  
 Brust: 30.1 mm  
 Bauch: 30.4 mm  
 Hüfte: 49.0 mm  
 Femur: 37.6 mm  
 Knie: 29.5 mm  
 Abw. IM-AM\_Opt gew. (LBFI%): -21.39 %  
 Wade: 20.3 mm  
 Summe HFD (P): 290.9 mm  
 Summe HFD (DW): 132.8 mm  
 KF %: 36.25 %  
 KF kg (P): 26.81 kg  
 LBM (P): 46.79 kg  
 KF % (M): 36.22 %  
 LBM (M): 46.81 kg  
 Abw. IM-AM\_idealgew. (BMI%): -29.5 %  
 KF kg (DW): 33.56 kg  
 LBM (DW): 39.84 kg  
 VWR: 0.8  
 BMI: 31.77 kg/m²

Verfügbare Untersuchungen

Informationen zur ausgewählten Untersuchung

Zugang zu eigenen Untersuchungsdaten

Probandenzugang

Abb. 39: Probandenzugang

## 6. Literatur

Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. In N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*, CA, Thousand Oaks, 111-119.

ACSM. (1998). American College of Sports Medicine Position Stand. The recommend quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness and flexibility in healthy adults. *Med Sci Sports Exerc* 30, 975-991.

AGF. (2010). Entwicklung der durchschnittlichen Sehdauer pro Tag/ Person in Minuten. Retrieved 25.10.2010, from <http://agf.de/daten/zuschauermarkt/sehdauer/>

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). *Qual Life Res* 19(3), 333-338.

Arnetz, B. B., Wasserman, J. & Petrini, B. (1987). Immune function in unemployed women. *Psychosom Med*, 49, 3-12.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*, 84(2), 191-215.

Beck, A. T., Steer, R. A. & Hautzinger, M. (2005). *Beck-Depressions-Inventar (BDI). Testhandbuch*. Bern [u.a.]: Huber.

Becker, P. (1997). *Psychologie der seelischen Gesundheit. Theorien, Modelle, Diagnostik* Göttingen [u.a.]: 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-Bundesverband. (2009). Gesundheitsreport 2009: Gesundheit in Zeiten der Krise. Essen: BKK.

Blair, S. N., LaMonte, S., J. (2006). Commentary current perspectives on obesity and health: black and white or shadows of grey? *Int J Epidemiol* 35, 69-72

Bland, J. M. & Altman, D. G. (2000). Statistics notes. The odds ratio. *BMJ*, 320(7247), 1468.

Bloomfield, K. (1998). West german drinking patterns in 1984 and 1990. *Eur addict Res* 4, 163-171.

- Blumenthal, J. A., Babyak, M. A., Moore, K. A., Craighead, W. E., Herman, S., Khatri, P. (1999). Effects of exercise training on older patients with major depression. *Arch Intern Med*, 159(19), 2349-2356.
- Bormann, C. (1992). Arbeitslosigkeit und Gesundheit. *Sozialer Fortschritt*, 41(3), 63-71.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Berlin , Heidelberg , New York: Springer.
- Bray, J. W., Zarkin, G. A. & Dennis, M. L. (2000). Symptoms of dependence, multiple substance use, and labor market outcomes. *American Journal of drug and alcohol abuse*, 26, 77-95.
- Bredenkamp, J. (1969). The application of significance tests in theory-testing experiments. *Psychologische Beiträge*, 11(3), 275-285.
- Brenner, M. H. (1987). Economic change, alcohol consumption and heart disease mortality in nine industrialized countries. *Soc Sci Med*, 25(2), 119-132.
- Brenner, S. O. & Starrin, B. (1988). Unemployment and health in Sweden: Public issues and private troubles. *Journal of Social Issue*, 44, 125-140.
- 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. *Soc Sci Med*, 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. *J Occupation Environ Med* 45(11), 1159-1166.
- Buksch, L., Schlicht, W. (2010) Reduziert sich das Mortalitätsrisiko sowohl für normal- als auch für übergewichtige Personen durch körperliche Aktivität? *Deut Z Sportmed* 61, 72-77.
- Buksch, J. (2005). Physical activity of moderate intensity in leisure time and the risk of all cause mortality. *Br J Sports Med* 39, 632-638.
- Bullinger, M. & Kirchberger, I. (1998). *SF-36. Fragebogen zum Gesundheitszustand - Handanweisung*. Göttingen [u.a.]: Hogrefe.
- Busch, A. J., Barber, K. A. R., Overend, T. J., Peloso, P. M. J. & Schachter, C. L. (2007). Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews*(4).
- Chapin, C. V. (1924). Death among taxpayers and non-taxpayers income tax, providence, 1865. *Am J Public Health*, XIV(8), 647-651.
- 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. *Add* 94(1), 133-138.
- Claussen, B., Björndal, A. & Hjort, P. F. (1993). Health and re-employment in a two year follow up of long term unemployment. *J Epidemiol Commun Health*, 47, 14-18.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ [u.a.]: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychol Bull*, 112(1), 155-159.
- Cook, D. G., & Allan, C. A. (1983). Self reported alcohol consumption and dissimulation in a Scottish urban sample. *Journal of studies on alcohol*, 44(4), 617-629.
- Cook, D. G., Cummins, R. O., Bartley, M. J. & Shaper, A. G. (1982). Health of unemployed middle-aged women in Great Britain. *Lancet*, 5(1), 1290-1294.
- Cramer, R., Gilberg, R., Hess, D., Marwinski, K., Schröder, H. & Smid, M. (2002) Suchintensität und Einstellungen Arbeitsloser IAB, *Beiträge zur Arbeitsmarkt- und Berufsforschung*, 261, Nürnberg.
- Crawford, A., Plant, M. A. & Kreitman, N. (1987). Unemployment and drinking behaviour: some data from a general population survey of alcohol use. *Br J addict*, 82, 1007-1016.
- Creed, P., Machin, M. & Hicks, R. (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. *Am J Prev Med*, 39(2), 173-179.
- Derogatis, L. R. & Cleary, P. A. (1977). Factorial invariance across gender for the primary symptom dimensions of the SCL-90. *Br J Soc Clin Psychol*, 16(4), 347-356.
- Dorling, D. (2009). Unemployment and health. *BMJ*, 338, 829.
- Durnin, J.V.G.A., Womersley, J. (1974). Body fat assessed from total body density and its estimation from body skinfold thickness: measurements on 481 men and women aged from 16-72 years. *Br J Nutr* 32: 77-97.
- 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*, 5, 267-273.
- Fahrenberg, J., Hampel, R. & Selg, H. (2010). *Freiburger Persönlichkeitsinventar (FPI-R); Manual*. Göttingen [u.a.]: Hogrefe.
- 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. *Behav Res Meth*, 39(2), 175-191.

- Festinger, L. (2001). A theory of cognitive dissonance. Stanford: Stanford Univ. Press.
- Filipp, S.-H. & Aymanns, P. (2010). Kritische Lebensereignisse und Lebenskrisen. Vom Umgang mit den Schattenseiten des Lebens. Stuttgart: Kohlhammer.
- Flemming, R., Baum, A., Reddy, D. & Gatchel, R. J. (1984). Behavioral and biochemical effects of job loss and unemployment stress. *J Hum Stress*, 10, 12-17.
- 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., 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 (Hrsg.), *Industrielle Psychopathologie*. Bern [u.a.]: Huber, 343.
- Frese, M. & Mohr, G. (1987). Prolonged unemployment and depression in older workers: A longitudinal study of intervening variables. *Soc Sci Med*, 25(2), 173-178.
- Gallo, W. T., Bradley, E. T., Siegel, M. & Kasl, S. (2000). Health Effects of Involuntary Job Loss Among Older Workers: Findings From the Health and Retirement Survey. *J Gerontol Soc Sci* 55B(3), 131-140.
- Gerdtham, U.-G. & Johannesson, M. (2003). A note on the effect of unemployment on mortality. *J Health Econom*, 22(3), 505-518.
- 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. *Psychol Med*, 27(1), 191-198.
- Goldberg, D. P. & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychol Med*, 9(1), 139-145.
- Grawe, K. (2004). Psychological therapy. Cambridge, Mass. [u.a.]: 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*, Berlin [u.a.]: Springer, 403-414.
- Grobe, T. & Dörning, H. (2010) Gesundheitsreport 2010. Gesundheitliche Veränderungen bei Berufstätigen und Arbeitslosen von 2000-2009. Hamburg: Techniker Krankenkasse.
- Grobe, T. & Schwartz, F. W. (2003) Arbeitslosigkeit und Gesundheit. "Gesundheitsberichterstattung des Bundes". Berlin: Robert Koch-Institut.

- Hager, W. (2006). Die Fallibilität empirischer Daten und die Notwendigkeit der Kontrolle der Wahrscheinlichkeiten falscher Entscheidungen *Z Psychol*, 214(1), 10-23.
- Hall, E. M. & Johnson, J. V. (1988). Depression in unemployed Swedish women. *Soc Sci Med*, 27, 1349-1355.
- Hammarstrom, A. & Janlert, U. (1997). Nervous and depressive symptoms in a longitudinal study of youth unemployment--selection or exposure? *J Adolesc*, 20(3), 293-305.
- Hanisch, K. A. (1999). Job Loss and Unemployment Research from 1994 to 1998: A Review and Recommendations for Research and Intervention. *J Voc Behav*, 55(2), 188-220.
- Haskell, W. L., Lee, I. M., Pate, R. R. et al. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc* 39, 1423-1434.
- Heather, N., Laybourn, P. & MacPherson, B. (1987). A prospective study of the effects on drinking behaviour. *Soc Psych*, 22, 226-233.
- Helmert, U. & Maschewsky-Schneider, U. (1998) Zur Prävalenz des Tabakrauchens bei Arbeitslosen und Armen. In D. Henkel & I. Vogt (Hrsg.). Sucht und Armut. Alkohol, Tabak, Medikamente, illegale Drogen. Opladen: Westdeutscher Verlag, 153-165.
- Henkel, D. (1998) Die Trunksucht ist die Mutter der Armut - zum immer wieder fehlgedeuteten Zusammenhang von Alkohol und Armut vom Beginn des 19. Jahrhunderts bis zur Gegenwart. In D. Henkel (Hrsg.). Sucht und Armut. Alkohol, Tabak, illegale Drogen. Opladen: Westdeutscher Verlag, 13-79.
- Henkel, D., Dornbusch, P. & Zemlin, U. (2005). Prädiktoren der Alkoholrückfälligkeit bei Arbeitslosen 6 Monate nach Behandlung: Empirische Ergebnisse und Schlussfolgerungen für die Suchtrehabilitation. *Suchttherapie*, 6, 165-175.
- Hill, J. (1978). The psychological impact of unemployment. *New Society*, 19, 118-120.
- Hirsch, R. & Pflingsten, U. (2002). Gruppentraining sozialer Kompetenzen. GSK : Grundlagen, Durchführung, Anwendungsbeispiele. Weinheim: Beltz.
- 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.
- Hofbauer-Flatzek, A. (1933). Beitrag zur Frage des Einflusses der Wirtschaftsnot, insbesondere der Erwerbslosigkeit (Arbeitslosigkeit ) auf die Gesundheit. *Int Arch Occup Environ Health*, 4(2), 362-380.
- 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. (2003). Arbeitslos - Gesundheit los - chancenlos? *IAB-Kurzbericht*, 4, 1-8.

- Hollederer, A. (2009). Gesundheit und Krankheit von Arbeitslosen sowie Chancen und Grenzen arbeitsmarktintegrativer Gesundheitsförderung. In Hollederer, A. (Hrsg.). *Gesundheit von Arbeitslosen fördern!*. Frankfurt am Main: Fachhochschulverl., 12-38.
- Hollederer, A. & Brand, H. (Hrsg.) (2006). *Arbeitslosigkeit, Gesundheit und Krankheit*. Bern: Huber.
- Hradil, S. & Schiener, J. (2001). *Soziale Ungleichheit in Deutschland*. Opladen: Leske & Budrich.
- Iversen, L., Andersen, O., Andersen, P. K., Christoffersen, K. & Kieding, N. (1987). Unemployment and mortality in Denmark, 1970-1980. *BMJ*, 295, 879-884.
- Jackson, P. R. (1988). Personal networks, support mobilization and unemployment. *Psychol Med*, 18(2), 397-404.
- Jackson, P. R., Stafford, E. M., Banks, M. H. & Warr, P. B. (1983). Unemployment and psychological distress in young people: The moderating role of employment commitment. *J Appl Physiol*, 68, 525-535.
- Jackson, P. R. & Warr, P. B. (1984). Unemployment and psychological ill-health: The moderating role of duration and age. *Psychol Med*, 14(3), 605-614.
- Jahoda, M. (1982). *Employment and unemployment. A socialpsychological analysis*. Cambridge M.: Cambridge Univ. Press.
- Jahoda, M. (1992). Reflections on Marienthal and after. *J Occup Organ Psychol*, 65(4), 355-358.
- Jahoda, M., Lazarsfeld, P. & Zeisel, H. (1933). *Die Arbeitslosen von Marienthal*. Frankfurt: Suhrkamp Verlag.
- Janssen, J. & Laatz, W. (2010). *Statistische Datenanalyse mit SPSS. Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*, Berlin [u.a.]: Springer.
- Jerusalem, M. & Schwarzer, R. (1984). Selbstwirksamkeit. In R. Schwarzer (Hrsg.), *Skalen zur Befindlichkeit und Persönlichkeit*. Berlin: Institut für Psychologie, Freie Universität Berlin, 15-28.
- 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. *J Am Geriatr Soc*, 44(6), 644-649.
- Kaplan, G. A. & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation*, 88, 1973-1998.
- Kasl, S., Cobb, S. & Brooks, G. W. (1968). Changes in serum uric acid and cholesterol levels in men undergoing job loss. *JAMA*, 206, 1500-1507.
- Kessler, R. C., Soukup, J., Davis, R. B., Foster, D. F., Wilkey, S. A., Van Rompay, M. I. (2001). The use of complementary and alternative therapies to treat anxiety and depression in the United States. *Am J Psychiatry*, 158(2), 289-294.

- Kessler, R. C., Turner, J. B. & House, J. S. (1989). Unemployment, reemployment and emotional functioning in a community sample. *Am Soc Rev*, 54, 648-657.
- Kieselbach, T. (1998). "Ich wäre ja sonst nie mehr an Arbeit rangekommen!". Evaluation einer Reintegrationsmaßnahme für Langzeitarbeitslose. Weinheim: Deutscher Studien-Verl.
- Kirschner, W. (2009). Möglichkeiten und Grenzen der Kursangebote zur Gesundheitsförderung bei Arbeitslosen nach §20 SGB V. In A. Holleder (Hrsg.) *Gesundheit von Arbeitslosen fördern! Ein Handbuch für Wissenschaft und Praxis*. Band 22, Frankfurt a.M.: Fachhochschulverlag, 101.
- Klein, G. & Strasser, H. (Hrsg.). (1997). *Schwer vermittelbar. Zur Theorie und Empirie der Langzeitarbeitslosigkeit*. Opladen: Westdeutscher Verl.
- 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. *Prev Med*, 23(2), 257-262.
- Kübler-Ross, E. (1980). *Interviews mit Sterbenden*. Stuttgart: Kreuz-Verl.
- 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. *Scand J Soc Med Suppl*, 43, 1-170.
- Lampert, T. (2010). Soziale Determinanten des Tabakkonsums bei Erwachsenen in Deutschland. *Bundesgesundheitsbl*, 53(2/3), 108-116.
- 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. *J Abnorm Psychol*, 69, 195-205.
- Leana, C. R. & Feldman, D. C. (1992). *Coping with job loss. How individuals, organizations, and communities respond to layoffs*. New York [u.a.]: 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? *Br J Sports Med*, 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.

- Margraf, J. (Hrsg.) (1998). Gesundheits- oder Krankheitstheorie?. Saluto- versus pathogenetische Ansätze im Gesundheitswesen. Berlin [u.a.]: Springer.
- Marsh, L. C., Fleming, A. G. & Blackler, C. F. (1938) Health and Unemployment. Oxford Univ. Press.
- 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. *Ann Math 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. *Job stress interventions*, 159-170.
- McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *J Behav Med*, 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. *J Gerontol*, 48(4).
- McKee-Ryan, F. M., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and Physical Well-Being During Unemployment: A Meta-Analytic Study. *Journal of Applied Physiology*, 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. (2001). Langzeiterwerbslosigkeit. In J. Zempel, J. Bacher & K. Moser (Hrsg.), *Erwerbslosigkeit. Ursachen, Auswirkungen und Interventionen*. Opladen: Leske & Budrich, 111-131.
- Mohr, G. & Otto, K. (2005). Langzeiterwerbslosigkeit: Welche Interventionen machen aus psychologischer Sicht Sinn?. *Zeitschrift für Psychotraumatologie und Psychologische Medizin*, 3(4), 45-63.
- Montgomery, S. M., Cook, D. G., Bartley, M. J. et al.(1998). Unemployment, cigarette smoking, alcohol consumption and body weight in young British men. *Eur J Public Health*, 8, 21-27.
- Morris, J. K., Cook, D. G. & Shaper, A. G. (1994). Loss of employment and mortality. *BMJ* 308, 1135-1139.
- Morris, J. K., Cook, D. J. & Shaper, A. G. (1992). Non-employment and changes in smoking, drinking, and body weight. *BMJ* 304, 536-541.
- Moser, K. A., Fox, A. J. & Jones, D. R. (1984). Unemployment and mortality in the OPCS longitudinal study. *Lancet*, 1324-1329.
- Moser, K. A., Fox, A. J., Jones, D. R. & Goldblatt, P. O. (1986). Unemployment and mortality: further evidence from the OPCS longitudinal study 1971-81. *Lancet*, 365-366.

- Murphy, G. C. & Athanasou, J. A. (1999). The effect of unemployment on mental health. *J Occup Organ Psychol*, 72(1), 83-99.
- Nakagawa, K., Inomata, N., Nakazawa, R. & Sakamoto, M. (2007). The effect of a health promotion program consisting of easy and simple exercises for community living elderly people. *J Phys Ther Sci*, 19(4), 235-242.
- Nicolai, S. P. A., Kruidenier, L. M., Leffers, P. et al. (2009). Supervised exercise versus non-supervised exercise for reducing weight in obese adults. *J Sports Med Phys Fitness* 49, 85-90.
- Nitsche, I. & Richter, P. (2003). Tätigkeiten außerhalb der Erwerbsarbeit. Evaluation des TAURIS-Projektes. Münster: Lit Verlag.
- Paul, K. I., Hassel, A. & Moser, K. (2006). Die Auswirkungen von Arbeitslosigkeit auf die psychische Gesundheit – Befunde einer quantitativen Forschungsintegration. *Arbeitslosigkeit, Gesundheit und Krankheit*. Bern: Hans Huber.
- Paul, K. I. & Moser, K. (2006). Incongruence as an explanation for the negative mental health effects of unemployment: Meta-analytic evidence. *J Occup Organ Psychol*, 79(4), 595-622.
- Paul, K. I. & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *J Vocational Behav*, 74(3), 264-282.
- Paul, K. I. & Moser, K. (2009). Wie wirken sich Interventionsmaßnahmen auf die psychische Gesundheit Arbeitsloser aus, wenn die Stellensuche erfolglos bleibt?. Erste Befunde einer Metaanalyse. In A. Holleder (Hrsg.) *Gesundheit von Arbeitslosen fördern!* Frankfurt am Main: Fachhochschulverl., 124-134.
- Parizkova, J. (1977). *Body fat and physical fitness*. Martinus Nijhoff b.v., The Hague, The Netherlands.
- Pate, R. R., Pratt, M., Blair, S. N. et al. (1995). Physical activity and public health. A recommendation from the Center of Disease Control and Prevention and the American College of Sports Medicine. *JAMA* 273, 402-407
- 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. *Compr Psychiatry*, 34(3), 198-205.
- Plöger, P. (2010). *Arbeitssammler, Jobnomaden und Berufsartisten. Viel gelernt und nichts gewonnen? Das Paradox der neuen Arbeitswelt*. München: Hanser.
- Proudfoot, J., Gray, J., Carson, J., Guest, D. & Dunn, G. (1999). Psychological training improves mental health and job-finding among unemployed people. *Int Arch Occup Environ 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.
- Puls, W., Inhester, M.-L., Reinecke, J. & Wienold, H. (2005). Arbeitslosigkeit, riskanter Alkoholkonsum und alkoholbezogene Verhaltensprobleme: eine Längsschnittstudie bei Metallarbeitern. *Suchttherapie*, 6, 70-77.

Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2010). Quantitative Methoden Band 2. Einführung in die Statistik für Psychologen und Sozialwissenschaftler.

Rásky, É., Stronegger, W.-J. & Freidl, W. (1996). Employment Status and Its Health-Related Effects in Rural Styria, Austria. *Prev Med*, 25(6), 757-763.

Ravens-Sieberer, U. (2002). Measuring and monitoring quality-of-life in population surveys: still a challenge for public health research. *Soz Präz Med*, 47(4), 203-204.

Rengers, M. (2005). Das international vereinbarte Labour-Force-Konzept. Wiesbaden: Statistisches Bundesamt, 1369-1383.

Resetka, H.-J., Liepmann, D. & Frank, G. (Hrsg.) (1996). Qualifizierungsmaßnahmen und psychosoziale Befindlichkeit bei Arbeitslosen. Frankfurt a. M.: Peter Lang.

Resnick, B. (2001). A prediction model of aerobic exercise in older adults living in a continuing-care retirement community. *J Aging 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 Med*, 28(6), 397-411.

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.

Robert Koch-Institut. (2006a) Gesundheit in Deutschland. *Beiträge zur Gesundheitsberichterstattung des Bundes*. Berlin: Robert Koch-Institut.

Robert Koch-Institut. (2006b) Telefonischer Gesundheitssurvey des Robert Koch-Institutes (2.Welle). *Beiträge zur Gesundheitsberichterstattung des Bundes*. Berlin: Robert Koch-Institut.

Robert Koch-Institut (2010) Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2009". *Beiträge zur Gesundheitsberichterstattung des Bundes*. Berlin: Robert Koch-Institut

Rogot, E., Scorlie, P. D., Johnson, N. J. et al. (1992). A mortality study of 1.3 million persons by demographic, social and economic factors: 1979-1985 follow-up: National Institute of Health.

Rose, G. & Marmot, M. G. (1981). Social class and coronary heart disease. *Br Heart J* 45, 13-19.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton Univer. Press.

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 [u.a.]: Hogrefe.
- Saam, R. H. & Wodtke, K. H. (1995). A cognitive stress reduction program for recently unemployed managers. *Career Development Quarterly*, 44(1), 43.
- Schach, E., Rister-Mende, S., Schach, S., Glimm, E. & Wille, L. (1994) Die Gesundheit von Arbeitslosen und Erwerbstätigen im Vergleich. Eine Analyse anhand von AOK- und Befragungsdaten. *Schriftenreihe der Bundesanstalt für Arbeitsschutz*. Dortmund.
- 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. & Brähler, 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.
- Scholz, R. D., Schulz, A. & Stegmann, M. (2010) Die ostdeutsche Übersterblichkeit der Männer im arbeitsfähigen Alter: ein Analyse auf Grundlage der "Aktiv Versicherten" der Deutschen Rentenversicherung. *DRV-Schriften*. Berlin: Deutsche Rentenversicherung Bund.
- Schumacher, J., Wilz, G., Gunzelmann, T. & Brähler, E. (2000) Die Sense of Coherence Scale von Antonovsky. Teststatistische Überprüfung in einer repräsentativen Bevölkerungsstichprobe und Konstruktion einer Kurzsкала. *Psychother Psychosom Med Psychol*, 50, 472-482.
- 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*. Windsor: UK: NFER-Nelson, 35-37.
- 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.
- Schwarzer, R., Jerusalem, M. & Hahn, A. (1994). Unemployment, social support and health complaints: A longitudinal study of stress in East German refugees. *Journal of Community & Applied Social Psychology*, 4(1), 31-45.
- Seligman, M. E. (1978). Learned helplessness as a model of depression. Comment and integration. *J Abnorm Psychol*, 87(1), 165-179.
- Seligman, M. P. (1991). *Learned optimism*. New York: A. A. Knopf.

- Shlay, A. B. & Rossi, P. H. (1992). Social science research and contemporary studies of homelessness. *Ann Rev Sociol*, 18, 129-160.
- Smart, R. G. (1979). Drinking problems among employed, unemployed and shift workers. *J Occup Med*, 21, 731-736.
- 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.
- Stoll, R., Sommer, K., Johnson, D. & Stoll, N. (1988). Instrument zur Hautfaltendickenmessung. *Ärztl Jugendkunde*, 80, 269-271.
- Stoll, R., Stoll, N. & Schröder, R. (2002). Anthropometrisches on-line Mess- und Expertensystem zur individuellen Einschätzung des Körperbaus und des Ernährungszustandes. *Arbeitsmed. Sozialmed. Umweltmed.*, 37, 28-33.
- Strawbridge, W. J., Deleger, S., Roberts, R. E. & Kaplan, G. A. (2002). Physical activity reduces the risk of subsequent depression for older adults. *Am J Epidemiol*, 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* (Vol. 2). Augsburg: Forschungsstelle.
- Sui, X., LaMonte, M. J., Laditka, J. N. et al. (2007). Cardiorespiratory fitness and adiposity as mortality predictors in older adults. *JAMA* 298, 2507-2516.
- Sullivan, M. & Karlsson, J. (1998). The Swedish SF-36 Health Survey III. Evaluation of criterion-based validity: results from normative population. *J Clin Epidemiol*, 51(11), 1105-1113.
- Taris, T. W. (2002). Unemployment and mental health: A longitudinal perspective. *Int J Stress*, 9, 43-57.
- Thurston, R. C., Kubzansky, I. K., Kawachi, I. & Berkman, L. F. (2005). Is the Association between Socioeconomic Position and Coronary Heart Disease Stronger in Women than in Men? *Am J Epidemiol*, 162(1), 57-65.
- Trube, A. & Luschei, F. (2000). Entwicklungs- und Vermittlungs-Assistenz (EVA) - ein Instrument zur Wiedereingliederung Langzeitarbeitsloser. Abschlussbericht über die Evaluation des Projekts "Entwicklungs- und Vermittlungs-Assistenz (EVA). Siegen: Zentrum für Planung und Evaluation Sozialer Dienste.
- van Gool, C. H., Kempen, G. I., Bosma, H., van Bortel, M. P., Jolles, J. & van Eijk, J. T. (2007). Associations between lifestyle and depressed mood: longitudinal results from the Maastricht Aging Study. *Am J 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. *Scand J Psychol*, 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*. Berlin: Ed. Sigma., 366.
- Vinokur, A. D., Price, R. H. & Schul, Y. (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. *Am J Community Psychol*, 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. *J Occup Health Psychol*, 5(1), 32-47.
- Von Stengel, S., Löffler, V., Kemmler, W. (2009). Körperliches Training und das 10-Jahres CHD-Risiko bei Frauen über 65 Jahren mit metabolischem Syndrom. *Deut Z Sportmed* 60, 281-287.
- Vuori, J. & Silvonon, J. (2005). The benefits of a preventive job search program on re-employment and mental health at 2-year follow-up. *J Occup Organ Psychol*, 78(1), 43-52.
- Vuori, J., & Vesalainen, J. (1999). Labour market interventions as predictors of re-employment, job seeking activity and psychological distress among the unemployed. *J Occup Organ Psychol*, 72(4), 523-538.
- Wanberg, C. R. (1995). A longitudinal study of the effects of unemployment and quality of reemployment. *J Voc Behav*, 46(1), 40-54.
- Warr, P. & Jackson, P. (1984). Men without jobs: Some correlates of age and length of unemployment. *J Occup Psychol*, 57, 77-85.
- Warr, P. & Jackson, P. (1987). Adapting to the unemployed role: a longitudinal investigation. *Soc Sci Med*, 25(11), 1219-1224.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur psychologischen Methodenlehre*. Göttingen: Hogrefe.
- Wiggert, A. (Hrsg.) (2009). *Kosten von Arbeitslosigkeit*. Hamburg: Igel Verlag.
- Williams, P. & Lord, S. R. (1995). Predictors of adherence to a structured exercise program for older women. *Psychol 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, England: Wiley & Sons.
- Winegardner, D., Simonetti, J. L. & Nykodym, N. (1984). Unemployment: The living death? *Journal of Employment Counseling*, 21(4), 149-155.
- Wingard, D. L. (1984). The sex differences in morbidity, mortality, and lifestyle. *Ann Rev Pub Health*, 5, 433-458.
- Wohlwill, J. F. (1977). *Strategien entwicklungspsychologischer Forschung*. Stuttgart: Klett-Cotta.

Zawadski, B., & Lazarsfeld, P. (1935). The psychological consequences of unemployment. *J Soc Psychol* 6, 224-251.

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. Eigene Publikationen

### Buchbeiträge:

**Kreuzfeld S, Weippert M, Preuss M, Kumar M, Stoll R**

Effekte eines kombinierten Ausdauer-Kraft-Trainings auf die sympathovagale Balance von Personen mit erhöhtem kardiovaskulären Risiko. In: Hottenrott K, Hoos O and Esperer HD (ed) Herzfrequenzvariabilität: Gesundheitsförderung- Trainingssteuerung- Biofeedback, Czwalina, Hamburg 2011, pp 175-183

**Weippert, M., Kumar, M., Kreuzfeld, S., Arndt, D., Stoll, R.:** HRV-basierte VO<sub>2</sub>max-Schätzung unter submaximaler Belastung. In: Hottenrott, K., Hoos, O. & Esperer, H. D. (Hrsg.): Herzfrequenzvariabilität: Risikodiagnostik, Stressanalyse, Belastungssteuerung. ISBN 978-3-88020-534-5 dvs-Schriften, Bd. 192, Czwalina, Hamburg 2009, pp. 94-103

### Peer Reviewed Proceedings:

**Kreuzfeld, S., Preuss, M., Weippert, M., Stoll, R.**

Health effects and acceptance of a physical activity intervention for older long-term unemployment people. Int Arch Occ Environ Health (under submission)

**Kreuzfeld, S., Stoll, R., Preuss, M.:** Abbau von Vermittlungshemmnissen bei älteren Langzeitarbeitslosen durch gezieltes körperliches Training. In Gesundheit Berlin-Brandenburg (Hrsg.): Dokumentation 16. Kongress Armut und Gesundheit-Verwirklichungschancen für Gesundheit und Satellit „Gesund und aktiv älter werden“, CD-ROM, Gesundheit Berlin-Brandenburg e.V. (2011), ISBN 978-3-939012-12-2

**Preuss, M., Kreuzfeld, S., Weippert, M., Kumar, M., Rieger, A., Stoll, R.:** Prädiktoren zur Aufrechterhaltung eines Bewegungsprogramms für ältere Langzeitarbeitslose. . Arbeitsmed.Sozialmed.Umweltmed. 46, 3 (2011): p. 202

**Kreuzfeld, S., Preuss, M., Weippert, M., Kumar, M., Rieger, A., Stoll, R.:** Gender-Aspekte in der Gesundheitsförderung von Erwerbslosen. . Arbeitsmed.Sozialmed.Umweltmed. 46, 3 (2011): p. 204

**Kreuzfeld, S., Preuss, M., Arndt, D., Weippert, M., Kumar, M. Stoll, R.:** Gesundheitsförderung verbessert Perspektiven für Langzeitarbeitslose. Arbeitsmed.Sozialmed.Umweltmed. 45, 6 (2010): p. 298

**Preuss, M., Kreuzfeld, S., Arndt, D., Weippert, M., Kumar, M., Neubert, S., Stoll, R.:** Körperbezogene Lebensqualität und psychologische Faktoren bei älteren Langzeitarbeitslosen. . Arbeitsmed.Sozialmed.Umweltmed. 44, 3 (2009): p. 156

**Preuss, M., Arndt, D., Kumar, M., Neubert, S., Kreuzfeld, S., Weippert, Kaber, D., Stoll, R.:** Experiments to assess work related emotional strain. Proc., 7<sup>th</sup> International Forum Life Science Automation ISBN 978-3-938042-23-6, 16.-18.09.2009 Hohe Duene (D), p. 45

**Kreuzfeld, S., Preuss, M., Arndt, D., Weippert, Stoll, R.:** Wieder fit mit 50plus? Ergebnisse eines präventivmedizinischen Angebotes für Langzeitarbeitslose. Dtsch Z Sportmed 60 (7/8), p. 217

**Weippert, M., Kumar, M., Kreuzfeld, S., Arndt, D., Stoll, R.:** HRV-basierte Schätzung der Ausdauerleistungsfähigkeit. Proceeding, 12. Symposium Arbeitsphysiologie für Nachwuchswissenschaftler der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V., 14.-16.11.2008. Schwerte(D), ISBN 978-3-98038342-6-1, p. 64-65

**Rimane, K., Neubert, S., Holzmüller-Laue, S., Kreuzfeld, S., Arndt, D., Göde, B., Thurow, K., Stoll, R.:** Flexible R&D Integration Platform of Process Informatics for Automated Medical Application and Mobile Data Acquisition. Proceedings, 4<sup>th</sup> Annual IEEE Conference on Automation Science and Engineering (IEEE Case 2008), Washington DC (USA), 23.-26.08.2008, IEEE Catalog Number: CFP08ASE-CDR, ISBN: 978-1-4244-2023-0 Library of Congress: 2007943845 Omnipress, 621-626

**Kreuzfeld, S., Arndt, D., Preuss, M., Weippert, M., Stoll, R.:** Internet Based Data Management: Applied Research in the Field of Occupational Medicine and Health Promotion. Proc, 6<sup>th</sup> International Forum of Life Science Automation. ISBN 978-3-938042-17-5, 10.-12.09.2008 Hohe Düne (D)

**Rimane, K., Neubert S., Holzmüller-Laue, S., Kreuzfeld, S., Göde, B.; Thurow, K., Stoll, R.:** Flexibles medizinisches Probanden- und Prozessdatenmanagement für den Einsatz in der Leistungsdiagnostik. Proceedings, eHealth 2008- Medical Informatics Meets eHealth, Wien(A), 29.05.-30.05.2008, pp.185-190

**Kreuzfeld, S., Arndt, D., Kumar, M., Weippert, M., Stoll, R.:** Internet basiertes Datenmanagement beim psychophysiologischen Screening älterer Langzeitarbeitsloser. Dokumentation Deutsche Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V. 48. Jahrestagung 12.-15.03.2008, Hamburg(D), ISBN 978-3-9811784-1-8, pp 615

#### Vorträge:

**Kreuzfeld, S., Preuss, M., Stoll, R.:** Fit 50+: Eine Intervention zur Verbesserung der physischen und psychischen Gesundheit älterer Langzeitarbeitsloser. Vortrag i.R. der BMBF geförderten Kooperation für nachhaltige Prävention (KNP), 21.09.2010, Berlin

**Kreuzfeld, S., Stoll, R.:** Fit50+: Ergebnisse des BMBF-Forschungsprojektes. Vortrag i.R. der Tagung der Arbeitsgruppe „Perspektiven für Langzeitarbeitslose“ des Landes Mecklenburg-Vorpommern, Landesamt für Gesundheit und Soziales Mecklenburg Vorpommern, 09.03.2010, Rostock

**Kreuzfeld, S., Preuss, M., Stoll, R.:** Fit50plus- Erste Ergebnisse des BMBF-Forschungsprojektes. Internationale Fachtagung Integration in Arbeit/ Gesundheit/Chancen am Arbeitsmarkt, Schwerin, 30.09.2009

**Kreuzfeld, S., Stoll, R.:** Gesundheitliche Prävention und ihr Einfluss auf die Arbeitsfähigkeit älterer Langzeitarbeitsloser. Intern. Fachtagung „ Ganzheitliche Integrationskonzepte für Migranten und ältere Arbeitslose“, Schwerin, 17.04.2008

**Stoll, R.:** Präventivmedizin an der Schnittstelle Wissenschaft- Gesundheitswirtschaft. Workshop „Fokus 2008“, Rostock, 19.08.2008

**Stoll, R., Thurow, K., Berndt, R.:** Präventivmedizin- Personal Health Monitoring - Telemedizin. 11. Technologieabend “Wertschöpfung aus Mecklenburg-Vorpommern-Assistenztechnologien aus dem Gesundheitsland” Max Planck Institut für Demographie Rostock, 04.12.2008

### Sonstige Publikationen:

**Brandt, J.** (RegioVision GmbH Schwerin): FIT 50plus-aktiv – Modulares Konzept zur Verbesserung der Arbeitsfähigkeit von Langzeitarbeitslosen durch Gesundheitsförderung. Darstellung der praktischen Umsetzung und Ergebnisse (2010)

**Brandt, J.:** FIT 50plus – Ein Konzept mit Zukunft. Jobjournal 2010, RegioVision GmbH Schwerin (Hrsg.) (2010)

Einträge in der Datenbank „Gesundheitliche Chancengleichheit“ der Bundeszentrale für gesundheitliche Aufklärung (BZgA) unter:

- Fit50+
- Fit50plus-aktiv (Folgeprojekt, praktische Umsetzung)

### Pressebeiträge:

**Hennings, A.:** „Ohne Sport wär' ich schon abgestürzt...“, Langzeitarbeitslose in Bewegung. In: Nachspiel, Deutschlandradio Kultur, 09.05.2010.

**gsub – Gesellschaft für soziale Unternehmensberatung:** Aktivität und Bewegung fördern ein selbstbestimmtes Leben. In: Bundesministerium für Arbeit und Soziales (Hrsg.): Reisende Reporter 2010, Perspektive 50plus, Beschäftigungspakte in den Regionen.