

## Dokumentation zur 9. Auflage des Freiburger Persönlichkeitsinventars (2020)

Im Juli 2020 erschien im Hogrefe Verlag die 9. Auflage des Freiburger Persönlichkeitsinventars mit dem Untertitel „Neue Normierung und Validitätshinweise, Prinzipien der Testkonstruktion und modernen Assessmenttheorie“.

Das Manual zur 9. Auflage hat mehr als 200 Seiten Umfang, weil es den Testautoren nicht allein um die systematische Aktualisierung des FPI-R, sondern auch um die methodenbewusste Anwendung geht. Gerade bei den so einfach erscheinenden Persönlichkeitsfragebogen sind die Prinzipien der Testkonstruktion und der empirischen Validierung wichtig, um kritisch auswählen und interpretieren zu können. Persönlichkeitsfragebogen erfassen Selbstbeurteilungen und Selbstberichte und sie sollen facettenreiche Persönlichkeitseigenschaften repräsentieren. Folglich können die psychometrischen Postulate (Messmodelle) und die Konstruktionsweisen von Intelligenz- und Leistungstests nicht einfach auf Persönlichkeitsfragebogen übertragen werden, und die Nachweise externer Kriterienvalidität haben hier eine noch größere Bedeutung. Diese Grundsätze für Persönlichkeitsfragebogen werden in den meisten Lehrbüchern über Psychologische Diagnostik, Persönlichkeitspsychologie, Statistik und Testkonstruktion nur kurz und unzusammenhängend erläutert. Auch die Prinzipien der modernen Assessmenttheorie, einschließlich der Methodik des ambulanten Assessment zur Evaluation von Persönlichkeitstests, finden erst allmählich Eingang.

Im Manual 2020 konnte nur eine Auswahl der wichtigsten Ergebnisse, Tabellen und Grafiken dargestellt werden. Deswegen wurde dieser zusätzliche Dokumentationsband erstellt, welcher sowohl die im Manual der 9. Auflage ausgeklammerten Tabellen und Abschnitte des Manuals der 7. bzw. 8. Auflage enthält als auch zusätzliche Tabellen und Grafiken aus der Erhebung 2018.

Die Dokumentation umfasst die folgenden Abschnitte:

I.	Textauszüge aus der 8. Auflage des FPI-R (2000) .....	1
II.	Fragebogen Erhebung 2018 .....	67
III.	FPI-R 2018 Datenübersicht .....	79
IV.	FPI-R Skalen: Mittelwert - Vergleiche nach Erhebungsjahr, Geschlecht, Alter .....	80
V.	Itemstatistiken 1982 – 2018 .....	85
VI.	Faktorenanalysen - Ladungsmuster (Erhebungen 1982, 1999, 2018), Patientenstichprobe 2018 .....	98
VII.	Gesundheits- und Belastungsindikatoren 1999 und 2018 .....	108
VIII.	Vergleich Bevölkerung und Patienten 2018 .....	112
IX.	Eigenschaftsprofile aufgrund von Clusteranalysen („Persönlichkeitstypen“) .....	129
X.	Multitrait-Multimethod Hinweise .....	137
XI.	Normenveränderung seit der Erhebung 1999 .....	142
XII.	Zur Qualitätssicherung von Persönlichkeitsfragebogen .....	143

## I. Auszüge aus der 8. Auflage des FPI-Manuals (2000)

### Inhaltsverzeichnis

<b>Vorwort zur 8. Auflage</b> .....	7
<b>Vorwort zur 7. Auflage</b> .....	9
<b>1 Einleitung</b> .....	11
1.1 Entwicklung des FPI-R .....	11
1.2 Zur Testtheorie und Kritik der Persönlichkeitsinventare.....	12
<b>2 Konstruktion des FPI-R</b> .....	19
2.1 Konzeption des FPI und Absichten der Revision 1982.....	19
2.2 Strategien der Skalenkonstruktion.....	21
2.3 Itempool .....	23
2.3.1 Neue Konstrukte im FPI-R.....	23
2.3.2 Aus FPI-G übernommene Konstrukte .....	24
2.4 Repräsentativerhebung .....	25
2.5 Skalenkonstruktion.....	29
2.6 Konfirmatorische Analysen.....	36
2.7 Skaleninterkorrelationen und Sekundäranalysen.....	42
2.8 Beziehungen zu Geschlecht, Alter und Schulabschluss .....	44
2.9 Offenheit oder Soziale Erwünschtheit der Antworten? .....	46
<b>3 Neue Normierung</b> .....	54
3.1 Absichten der Untersuchung.....	54
3.2 Fragebogen und Interview.....	54
3.3 Repräsentativerhebung .....	55
3.4 Item- und Faktorenanalysen, Reproduzierbarkeit der Struktur der FPI-Skalen .....	66
3.5 Beziehungen zu Geschlecht, Alter, Schulabschluss, Einkommen und Ost-/West-Deutschland .....	76
3.6 Stabilität der FPI-Normen 1982/1999 .....	80
3.7 Akzeptanz und Zuverlässigkeit der Erhebung.....	82
<b>4 Beschreibung der Skalen des FPI-R</b> .....	84
<b>5 Anwendung</b> .....	91
5.1 Testdurchführung .....	91
5.2 Testauswertung .....	92
5.3 Normen.....	95
5.4 Auswertungsbogen.....	95
5.5 Interpretierbarkeit von Testwerten .....	96
5.6 Informationsbank FPI-R und Vergleichsgruppen.....	97
5.7 Übersetzungen und Lizenzausgaben .....	98
5.8 Auswertungsbeispiel .....	98
<b>6 Validitätshinweise</b> .....	100
6.1 Faktorielle Validität und Kriterienvalidität .....	100
6.2 Validitätshinweise aus der Repräsentativerhebung 1982 .....	101
6.3 Validitätshinweise aus der Repräsentativerhebung 1999 .....	107
6.4 Weitere Beiträge zur Validierung.....	120

<b>7</b>	<b>Aktuelle Beiträge zur Evaluation des FPI-R .....</b>	<b>132</b>
7.1	Absichten .....	132
7.2	Grundlagenforschung zur biologischen Basis von Extraversion und Emotionalität .....	133
7.3	Klinische Psychologie .....	133
7.4	Forensische Psychologie und Kriminologie .....	136
7.5	Kritik an der Dominanz der Fragebogenmethodik .....	138
7.6	Ambulantes Assessment als innovativer Ansatz der Evaluation .....	139
7.7	Antworttendenzen in Fragebogen: Selbstdarstellung in der Selbstbeurteilung .....	142
<b>8</b>	<b>Konstruktion und methodenbewusste Anwendung von Persönlichkeitsfragebogen .....</b>	<b>152</b>
8.1	Konstruktion und Evaluation von Persönlichkeitsinventaren – kritische Differenzierungen .....	152
8.2	Skalierung: Kontroversen über Messtheorie .....	152
8.3	Fragwürdige Übertragung messtheoretischer Axiome auf Introspektionen und Selbstbeurteilungen .....	155
8.4	Eigenart von Persönlichkeitsfragebogen und Fragebogenmethodik .....	156
8.5	Itementwicklung als Operationalisierung von Persönlichkeitseigenschaften .....	158
8.6	Itemkritik und Aggregationen .....	160
8.7	Konstruktionsprinzipien und Plädoyer für eine multistrategische Heuristik .....	160
8.8	Gibt es für die Konstruktion von Persönlichkeitsfragebogen ein überzeugendes Messmodell? .....	162
8.9	Problematischer Gebrauch der konventionellen Itemanalyse, Faktorenanalyse und IR-Modellierungen .....	166
8.10	Multitrait-Multimethod .....	170
8.11	Multimodale Diagnostik .....	171
8.12	Generalisierbarkeitstheorie .....	173
8.13	Qualitätssicherung .....	174
8.14	Reviews und Dokumentationen .....	176
8.15	Übersetzungen und Adaptationen .....	176
8.16	Diagnostische Urteilsbildung: Datenintegration statt Interpretation? .....	176
8.17	Strategische Konsequenzen – Ausblick von einer mittleren Position .....	179
<b>9</b>	<b>Aspekte der Testgüte .....</b>	<b>183</b>
9.1	Testgrundlage .....	183
9.1.1	Diagnostische Zielsetzung .....	183
9.1.2	Theoretische Grundlage .....	184
9.1.3	Nachvollziehbarkeit der Testkonstruktion .....	185
9.2	Testdurchführung .....	185
9.2.1	Durchführungsobjektivität .....	185
9.2.2	Transparenz .....	186
9.2.3	Zumutbarkeit .....	186
9.2.4	Verfälschbarkeit .....	186
9.2.5	Störanfälligkeit .....	186
9.3	Testverwertung .....	186
9.3.1	Auswertungsobjektivität .....	186
9.3.2	Zuverlässigkeit .....	186
9.3.3	Gültigkeit .....	186
9.3.4	Normierung .....	187
9.3.5	Bandbreite .....	187
9.3.6	Informationsausschöpfung .....	187
9.3.7	Änderungssensitivität .....	187

9.4	Testevaluation .....	188
9.4.1	Ökonomie.....	188
9.4.2	Fairness .....	188
9.4.3	Akzeptanz.....	188
9.4.4	Vergleichbarkeit.....	188
9.4.5	Bewährung .....	188
<b>10</b>	<b>Literaturverzeichnis.....</b>	<b>189</b>
<b>11</b>	<b>Anhang .....</b>	<b>209</b>
	<b>FPI-R Skalen .....</b>	<b>210</b>
	Normstichprobe N=3740 .....	210
	Stanine-Normen.....	211
	<b>FPI-A1 Mittelwerte und Standardabweichungen .....</b>	<b>216</b>
	Statistiken der FPI-A1-Skalen in der Normstichprobe 1982 (N=2035) und in Teilstichproben.....	216
	Normen für FPI-A1.....	217
	<b>Literatur-Recherche .....</b>	<b>220</b>
	Aktuelle Nachweise zum FPI in den Datenbanken PSYINDEX und PSYTKOM.....	220
	<b>Probanden-Gruppen der Informationsbank .....</b>	<b>221</b>

## Vorwort zur 8. Auflage

Zur Qualitätskontrolle des Freiburger Persönlichkeitsinventars wurde 1999, anlässlich der 7. Auflage, eine zweite bevölkerungsrepräsentative Normierung mit 3740 Personen durchgeführt. Die Ergebnisse wichen nur geringfügig von der vorausgegangenen Normierung 1982 ab. Die zusätzlichen Daten der Erhebung lieferten neue Validitätshinweise, außerdem wurde die Skalenkonstruktion überprüft. Wegen der guten Übereinstimmung beider Erhebungen ist eine erneute Normierung gegenwärtig nicht erforderlich.

Die kontinuierliche „Pflege“ eines Tests verlangt, bei neuen Auflagen des Manuals wichtige Ergänzungen und Hinweise einzufügen. Herausragende empirische Untersuchungen und Anwendungserfahrungen mit dem FPI-R werden im Kapitel 7 *Aktuelle Beiträge zur Evaluation* dargestellt. Hier wird auch die Diskussion über *Antworttendenzen* fortgeführt, vor allem in Hinblick auf Selbstdarstellung und *soziale Erwünschtheit*, die je nach Untersuchungsanlass unterschiedliches Gewicht haben. Die persönlichkeitspsychologische Fachliteratur widmet den fragwürdigen Behauptungen, die wichtigsten 3, 4, 5, 6, 7 usw. Grundfaktoren aufgefunden zu haben, einen unverhältnismäßig großen Raum. Statt solcher Kontroversen wäre es ungleich wichtiger, die *Auswahl* geeigneter Eigenschaftskonzepte für bestimmte diagnostische Aufgaben zu begründen sowie an den einzelnen Eigenschaftstheorien und den damit verknüpften -Assessmentstrategien zu arbeiten.

Psychologische Fragebogen sind zurzeit der dominierende Methodentyp in der Persönlichkeitsdiagnostik und Persönlichkeitsforschung – mit weitem Abstand vor allen anderen Methoden. Dennoch gehen die meisten Lehrbücher und die neuen Grundsätze der Qualitätssicherung zu wenig auf die wichtigen Unterschiede zwischen den Persönlichkeitsfragebogen und den Intelligenz- und Leistungstests ein. Die heute propagierten unidimensionalen Messmodelle für intervallskalierte Daten sind jedoch für die Datenebene der Selbstberichte und Selbstbeurteilungen und für die facettenreichen Konstrukte der Persönlichkeitseigenschaften kaum geeignet.

Andere konzeptuelle oder methodische Fortschritte werden in den Lehrbüchern nur unzureichend dargestellt oder sogar ausgeklammert. Dazu gehören das *Dilemma von Reliabilität und Validität*, die für Validierungen wichtigen Prinzipien

der *Datenaggregation*, die nötige *Symmetrie* von Konstrukten und Kriterien, die Prinzipien der *Generalisierbarkeitstheorie* und *ökologischen Validität*. Empirische, aber auch kritische Beiträge zur Validität von Persönlichkeits-tests sind vom *ambulanten Assessment* zu erwarten, denn mit dieser innovativen Methodik können unter Alltagsbedingungen kontextbezogene und zeitgenaue Selbstberichte über Befindlichkeit, Verhalten, Symptome sowie auch Bewegungsmuster und körperliche Reaktionen computerunterstützt erfasst werden (Abschnitt 6.4 und 7.6). Die auffälligsten Lücken bestehen hinsichtlich der *multimodalen Diagnostik*, für die es weiterhin an gründlicher empirischer Forschung und an den notwendigen Standards der Qualitätssicherung mangelt. Wie sind die oft diskrepanten Befunde verschiedener Beschreibungsebenen zu interpretieren? – Diese grundsätzlichen Beiträge zur *Konstruktion und methodenbewussten Anwendung von Persönlichkeitsfragebogen* werden mit Bezug auf das FPI-R in dem neuen Kapitel 8 diskutiert.

In dieser Auflage sind die Adaptationen und Lizenzausgaben des FPI-R in anderen Sprachen aufgeführt; z.T. handelt es sich nicht nur um Übersetzungen, sondern um Rekonstruktionen aufgrund eigener Datenerhebungen (Abschnitte 5.7 und 8.15). Erweitert wurde der Abschnitt über die computerunterstützte Vorgabe und Auswertung des FPI-R (Hogrefe-Testsystem HTS, siehe Abschnitt 5.2). Das Info zur Literaturrecherche wurde aktualisiert. Diese Informationen und Kommentare liegen in der Linie der früheren Auflagen, ausführlich über die Absichten der Konstruktion, über Validierungshinweise, Anwendungen und auch die Kritik an Persönlichkeitsfragebogen zu berichten. Dieses Vorhaben ist heute noch wichtiger, da wegen der Studienreformen wahrscheinlich nur noch ein Teil der Absolventen eines Psychologie-Studiums (mit speziellem Vertiefungsfach) über eine hinreichende diagnostische Ausbildung, einschließlich Testkonstruktion und Gutachtenpraxis, verfügen wird.

Gerade Persönlichkeitsfragebogen verlangen eine methodenbewusste Anwendung und Strategien der multimodalen Diagnostik, d.h. Absicherungen und kritische Interpretationen.

Freiburg, November 2009

*Jochen Fahrenberg*  
*Herbert Selg*  
*Rainer Hampel*

# 1 Einleitung

## 1.1 Entwicklung des FPI-R

In der 7. Auflage des Manuals zum Freiburger Persönlichkeitsinventar 2001 wurden die Ergebnisse einer zweiten Normierung des FPI-R mitgeteilt. Bei dieser Repräsentativerhebung konnte zugleich die Struktur der FPI-Skalen überprüft werden. Die an N=3740 Personen in den alten und neuen Bundesländern durchgeführte Erhebung ergab eine Anzahl weiterer Validierungshinweise für dieses Persönlichkeitsinventar.

Das Freiburger Persönlichkeitsinventar wurde in den sechziger Jahren entwickelt als es noch keine repräsentativ normierten deutschen Persönlichkeitsinventare gab, sondern nur einige Übersetzungen angloamerikanischer Tests vorhanden waren. Anlässlich der vierten Auflage im Jahr 1982 wurde eine große Repräsentativerhebung, welche als Grundlage für die Erweiterung und Teilrevision diente, durchgeführt.

Die Schritte dieser Testentwicklung von der ersten Fassung FPI-G (mit Parallelförm A und B sowie Kurzform K) zur revidierten Form FPI-R (mit der neuen Normierung der Form A als A1) wurden in den früheren Auflagen der Handanweisung dargestellt. Zusätzliche testmethodische Analysen und umfangreiches Tabellenmaterial wurden an anderer Stelle publiziert oder in Arbeitsberichten zugänglich gemacht (Fahrenberg, Selg & Hampel 1983, Fahrenberg, Hampel & Selg 1985, Fahrenberg, Ewert & Maier 1987, Hampel & Fahrenberg 1983).

Seit der 1. Auflage im Jahr 1970 sind zahlreiche Untersuchungen veröffentlicht worden, in denen das FPI angewendet oder auch testmethodisch analysiert und kritisch kommentiert wurde. Diese für die Anwender z.T. wesentlichen Informationen wurden in die 2. und 3. Auflage eingearbeitet, hatten aber den Umfang und Preis der Handanweisung stark anwachsen lassen. Die Testautoren hatten sich deshalb entschlossen, von der 4. Auflage an einen anderen Weg zu gehen. Getrennt von der Handanweisung wurden kommentierte Literaturverzeichnisse hergestellt. Das erste, von Potreck (1983) verfasste Verzeichnis berücksichtigte 149 FPI-bezogene Publikationen im Zeitraum 1978–1983; es wuchs bis 1993 auf 655 Einträge (Jehle, 1988; Jehle & Fahrenberg, 1994). Auch diese Hilfestellung bei der Literatursuche ist heute überholt, da die computerunterstützte Literaturrecherche weithin zugänglich geworden ist.

In den Anhang wurde eine Strategie zur Literaturrecherche in den Datenbanken PSYINDEX und PSYTKOM aufgenommen. Diese Datenbanken enthalten bis Oktober 2009 ca. 1400 Hinweise, davon 354 zum FPI-R. Weitere Publikationen mit Bezug auf das FPI sind in PsycINFO und Medline verzeichnet.

Das Freiburger Persönlichkeitsinventar kann aufgrund der in dieser Handanweisung enthaltenen Informationen korrekt angewendet und interpretiert werden. Zur wissenschaftlichen Fundierung und Kritik dieses Testsystems gehören jedoch viele weitere Informationen. In Rezensionen der 4. Auflage (Hollmann 1988, Jäger 1985, Schwenkmezger 1987, Seidenstücker 1985, Ostendorf, 1997; vgl. Fahrenberg, 1997) ist z.T. hervorgehoben worden, dass die Interpretation von FPI-Befunden eigentlich die Kenntnis der 3. Auflage mit den dort referierten methodischen und empirischen Beiträgen voraussetzt. Tatsächlich liegen Veröffentlichungen zu vielen testmethodischen Fragen und zu vielfältigen Anwendungen mit einem sehr breiten Spektrum von Fragestellungen bei den untersuchten klinischen und nicht klinischen Populationen vor. Gründliche Interpretationen von FPI-Testwerten und wissenschaftliche Publikationen werden solche Informationen einbeziehen müssen. In die Handanweisung können aber statt einer regelmäßigen Fortschreibung nur noch die wichtigsten Arbeiten und Hinweise aufgenommen werden. Zur computergestützten Anwendung und Auswertung des FPI-R gibt es heute mehrere Möglichkeiten (siehe Kapitel 5 über die Testanwendung). Methodenkritische Untersuchungen konnten zeigen, dass kein wesentlicher Unterschied in der Durchführungsobjektivität zwischen dem Fragebogen und der Computerversion besteht.

Seit der Revision des FPI im Jahr 1984 wurden die älteren Parallel- und Kurzformen B und K nicht weitergeführt. Um jedoch eine Vergleichbarkeit mit früheren Untersuchungen nicht völlig auszuschließen, wurde die Form A mit geringen Veränderungen und in neuer Normierung als Form A 1 beibehalten, testmethodisch jedoch nicht mehr „gepflegt“. Das Fragebogen-Formular FPI-A1 ist beim Verlag weiterhin erhältlich, und im Anhang sind auch die Normen für FPI-A1

auf der Basis der Repräsentativerhebung von 1982/84 enthalten.

Außer den Ergebnissen der neuen Normierung (Kapitel 3) und den neuen Validitätshinweisen (Kapitel 6 und 7) wurde ein Abschnitt über die Aspekte der Testgüte in die Handanweisung eingearbeitet.

## 1.2 Zur Testtheorie und Kritik der Persönlichkeitsinventare

In den früheren Auflagen dieser Handanweisung wurden in einleitenden Kapiteln auch die operationale Definition von Eigenschaften durch item-analytierte Fragebogen erläutert und die allgemeinen Vorzüge und Nachteile des Testsystems der Persönlichkeitsinventare kommentiert. Die inzwischen weiter angewachsene Literatur schließt hier eine gründliche Auseinandersetzung mit den hauptsächlichen Behauptungen und Einwänden und eine Darstellung methodenkritischer Studien aus. Einerseits wäre diese Diskussion im Rahmen einer allgemeinen Theorie der psychologischen Diagnostik problemspezifischer zu führen, und in bestimmten Anwendungsgebieten wäre nach möglicher Kriterienvorhersage, nach dem Entscheidungsnutzen des Tests und nach alternativen Strategien zu fragen. Andererseits ergeben sich über diese diagnostische Praxis hinaus grundsätzliche Fragen nach der theoretischen Begründung der verwendeten Eigenschaftsbegriffe und schließlich auch methodologische und wissenschaftstheoretische Kontroversen über die psychologische Messtheorie und die Chancen einer angemessenen Operationalisierung von Dispositionsprädikaten.

Hier sollen zumindest einige, im engeren Sinn testmethodische Fragen behandelt werden, welche sich jedem kritischen Anwender des FPI oder eines ähnlichen Persönlichkeitsinventars stellen werden (siehe auch Mummendey, 1999; Tränkle, 1983). Als Vorzüge dieses Testsystems sind zu nennen:

### *(1) Mehrdimensionale Erfassung von relativ überdauernden Persönlichkeitsmerkmalen*

Persönlichkeitsinventare liefern eine interindividuell vergleichende Beschreibung von Individuen hinsichtlich der Ausprägung wichtiger Persönlichkeitsdimensionen. Die verwendeten Skalen sind in einem empirisch-statistischen und psychologisch-interpretativen Konstruktionsprozess entstanden, indem psychologisch ähnliche Items der Selbstschilderungen zu einem relativ invarianten, konsistenten und zeitlich stabilen Muster gruppiert werden. Es wird eine diesem Muster zugrundeliegende Eigenschaftsdimension angenommen, und die Skalenwerte geben die quantitative Einordnung, zumindest die Rangordnung verschiedener Individuen entlang dieser Eigenschaftsdimension. Die Persönlichkeitseigenschaften bzw. die Skalenwerte des Individuums sind Dispositionsprädikate, welche eine statistische Erwartung formulieren, dass sich das betreffende Individuum in seinen künftigen Selbstschilderungen, aber – je nach Kriterienvalidität der Skalen – auch in seinem Verhalten konsistent mit dieser Eigenschaftsausprägung zeigen wird. Diese Erwartung bezieht sich jedoch nicht auf die Vorhersage einer sehr speziellen, isolierten Selbstschilderung oder isolierten Verhaltensweise. Persönlichkeitseigenschaften sind in der Regel facettenreiche theoretische Konstrukte mit einer entsprechenden Vielfalt von empirischen Referenten (möglichen Bezügen), welche in dem als Skalenbezeichnung verwendeten Eigenschaftsnamen nur sehr grob zusammengefasst sind.

### *(2) Breite Basis der Kriterienvalidierung*

Es gibt viele Belege, dass den individuellen Skalenwerten von Persönlichkeitsinventaren empirische Validität zukommt: systematische Beziehungen zu Lebenslaufdaten, Befindensdaten, Fremdeinstufungen, Statusmerkmalen und bestimmten psychopathologischen Kategorien usw. So existieren vielfältige empirische Beziehungen, wenn auch oft die Höhe der Korrelationskoeffizienten bzw. der Anteil gemeinsamer Varianz nur in Gruppenvergleichen überzeugend, aber für individuelle Vorhersagen noch unbefriedigend ist. Auch wegen der Breite der theoretischen Konstrukte und ihrer z.T. divergenten Subkonstrukte sowie der entsprechenden Kriterienproblematik sind solche Validitätsnachweise nur als Netzwerk vieler Einzelstudien – wie es auch beim FPI der Fall ist – möglich.

### *(3) Durchführungs- und Auswertungs-Objektivität*

Die Durchführung und Auswertung des Persönlichkeitsinventars sind in hohem Maße objektiv, d.h. vom untersuchenden Psychologen unabhängig. Die Testinterpretation ist intersubjektiv gut zu kontrollieren. Die Vorzüge dieser Standardisierung im Vergleich zu einem freien oder halbstrukturierten Interview werden es oft nahelegen, beide Verfahren in einander ergänzender Weise zu nutzen.

#### *(4) Ökonomische Informationsgewinnung*

Mit einem Persönlichkeitsinventar können bei relativ geringem Aufwand viele Informationen über einen Probanden gewonnen werden. Die Untersuchung erfordert nur den Fragebogen und den Auswertungsbogen. Der Test kann vom Probanden selbstständig ausgefüllt werden, und ist von einer Hilfskraft mittels Schablone oder durch EDV, d.h. durch Datenerfassung an einem Terminal und anschließende computergestützte Auszählung, leicht auszuwerten. Erst die Interpretation der einzelnen Skalenwerte und der individuellen Konfiguration im Zusammenhang von biografischen Daten und spezieller Fragestellung verlangt diagnostisch geschulte Psychologen/innen. Diese ökonomische Informationsgewinnung bedeutet einen Vorzug des Persönlichkeitsinventars als „Breitband-Verfahren“, als Bestandteil eines Screening, einer umfangreichen Routinediagnostik oder breiter Datenerhebungen für Forschungszwecke. Schließlich gibt es Untersuchungen, bei denen allein schriftliche Erhebungen möglich sind.

Jeder der hier genannten Vorzüge von Persönlichkeitsinventaren wird methodische und grundsätzliche Einwände finden. Gerade die Durchsichtigkeit dieses Testsystems und der sehr breite und z.T. wohl kritiklose Gebrauch von Fragebogen fordern fachliche Bedenken heraus, Zweifel am Nutzen der Daten und Warnungen vor Missbrauch.

Auch in populären Büchern und Massenmedien ist gelegentlich eine pauschale Kritik an Tests und insbesondere an Persönlichkeitsfragebogen verbreitet worden. Dennoch wird von vielen Befragten der psychologische Nutzen des gerade ausgefüllten Fragebogens nicht geringgeschätzt. In der Normstichprobe hat sich die große Mehrzahl so geäußert, dass man nach diesen Noten eine Augenscheinvalidität annehmen darf. Die Aussage lautet: Um Menschen besser zu verstehen, sind solche Fragebogen meiner Meinung nach ... sehr gut geeignet (14% Zustimmung), gut geeignet (35%), befriedigend geeignet (27%), ausreichend geeignet (10%), nicht geeignet (14% Zustimmung).

#### *Hauptsächliche Einwände gegen Persönlichkeitsinventare*

In der Einleitung zu einem Symposium „Das Dilemma von Persönlichkeitsfragebogen“ hat Janke (1973, S. 44ff.) in sehr pointierter Weise neun hauptsächliche Einwände formuliert. Diese Thesen sollen hier aufgegriffen und kurz kommentiert werden, um darzulegen, dass viele Aspekte in Methodenstudien bisher nur unzureichend geklärt sind, die Bewertung der Gütemerkmale aber von vielen speziellen Bedingungen abhängt und allgemeine Feststellungen auch auf diesem Gebiet psychologischer Methodenlehre sehr schwierig sind. In den späteren Kapiteln zur Testkonstruktion des FPI-R werden einige dieser testmethodischen Einwände noch einmal aufgenommen.

(1) Im Vergleich zu Leistungstests haben Persönlichkeitsfragebogen, so stellt Janke fest, eine wesentlich geringere Stabilität ( $.50 < r_{tt} < .70$ ) und es sei unbekannt, ob diese Stabilitätskoeffizienten konstantes Urteilsverhalten, Gedächtnis oder tatsächliche Konstanz der erfassten Merkmale wiedergeben.

Abgesehen von einer sicher auch für Leistungstests nicht einfachen Varianzaufklärung dieser Art, ist nach einem angemessenen Erwartungswert zu fragen. Ist für alle Dispositionsprädikate, die durch Fragebogen erfasst werden, psychologisch dieselbe zeitliche Stabilität zu fordern oder sind Unterschiede der relativen Veränderlichkeit bzw. Umweltstabilität anzunehmen? Von welchen stichprobentechnischen und versuchsplanerischen Bedingungen hängt die Höhe der Stabilitätskoeffizienten auf Skalen- und Itemebene ab?

Die Übersicht von Schuerger, Tait & Tavernelli (1982; siehe auch Amelang & Bartussek, 1997) erläutert solche Bedingungen und referiert für einige Fragebogenskalen Stabilitätskoeffizienten, welche im Abstand weniger Wochen auch höher als  $r_{tt}=0.8$  sein können, aber im Laufe der Zeit natürlich abnehmen und – soweit untersucht – nach zwanzig Jahren in der Größenordnung von 0.5 liegen. Aus mehreren Studien sind Stabilitätskoeffizienten von Items und Skalen des FPI über den Zeitraum von Wochen bis zu einigen Jahren bekannt; sie liegen in der Größenordnung von 0.7 bis 0.8, d.h. in der Höhe der Konsistenzkoeffizienten (siehe Tabelle 12a der 3. Aufl.).

(2) „Sie haben eine geringe innere Konsistenz ( $.60 < r_{tt} < .80$ ).“ Hier ist anzumerken, dass man die Konsistenz vieler Fragebogenskalen wahrscheinlich auf 0.85 oder 0.90 steigern könnte, wenn es in der Testentwicklung nur darauf ankäme. Dies ist bekanntlich zu erreichen, indem man homogene Items, u.U. auch Formulierungsvarianten wie es in vielen Leis-



tungstests geschieht, hinzufügt oder indem man den Konstruktbereich einengt. Solche extrem homogenen Skalen sind aber als Prädiktoren praktisch interessanter Kriterien u.U. weniger nützlich als facettenreichere Skalen (siehe Cattell 1973, Holling 1981, Wittmann & Schmidt 1983). Allerdings muss noch eine relative Konsistenz der Items bestehen, um überhaupt von einem einheitlichen Konstrukt sprechen und zwischen Individuen mit hoher bzw. niedriger Ausprägung der Disposition unterscheiden zu können. Ein zusätzlicher Gesichtspunkt zur Evaluation solcher Mindestkonsistenzen ist, außer den Item-Interkorrelationen einer Skala, der mittlere Konsistenzkoeffizient einer Serie von zufällig aus dem Itempool des Inventars gebildeter Skalen. Dies ist zugleich eine Schätzung der unspezifischen Varianz des betreffenden Inventars. Dieser Koeffizient ist für FPI-G schon in der 1. Auflage des FPI mitgeteilt worden und wurde auch für FPI-R berechnet (siehe Abschnitt 2.5).

(3) „Ihre Korrelationen mit quantitativ abgestuften Außenkriterien von nennenswerter praktischer Bedeutung liegen zwischen .20 und .30.“ Auch dies ist ein schwieriger Sachverhalt, wie die berechtigte Kritik an Mischels (1968) pauschal behaupteter Validitätsgrenze 0.30 zeigt (siehe Epstein 1977, Eysenck & Eysenck 1980 und auch Mischels 1979 spätere Präzisierungsversuche, Kenrick & Funder, 1988). Nicht die Höhe des Koeffizienten kann hier den allein entscheidenden Maßstab liefern, sondern der inkrementelle Entscheidungsnutzen bei einer bestimmten Problemstellung unter Abwägung der Risiken und methodischen Alternativen. Existieren denn andere Prädiktoren, welche die fraglichen Kriterien bzw. Statusunterschiede besser indizieren?

Die außerordentliche breite Literatur zur Validität von Persönlichkeitsfragebogen kann hier nicht referiert werden, schon die FPI-Literatur ist dafür zu umfangreich (siehe Potreck, 1983; Jehle, 1988; Jehle & Fahrenberg, 1994). Beschränkt man sich auf die größeren Studien über FPI-Skalen und Fremdeinstufungen, so lassen sich aus den Arbeiten von Timm (1971) und von Burisch (1976; siehe 3. Aufl.) und von Amelang und Borkenau (1982) sowie Schmidt und König (1986) zahlreiche Beziehungen oberhalb der Grenze von 0.30 entnehmen. Mit Wittmann & Schmidt (1983) ist von gründlichen Untersuchungen zur Vorhersagbarkeit des Verhaltens zu verlangen, dass auf der Kriterienseite ebenso viel Anstrengungen unternommen werden wie bei der Testkonstruktion.

(4) „Größere Kriterien, z.B. klinisch-diagnostische Gruppen, werden zwar oft gut diskriminiert, allerdings häufig auch solche, die gar nicht laut Gültigkeitsanspruch diskriminiert werden sollen; das heißt: die differenzielle Validität im Sinne der Terminologie der Klassifikationsliteratur ist gering.“

(5) „Korrelationen von Q-Daten mit L- und T-Daten gemeinsamen Konstrukt-Anspruchs (meist nur L-R-Daten) sind im allgemeinen so niedrig, dass man Fragebogen oft nur als Techniken kennzeichnen kann, von denen man nicht weiß, wozu sie dienen sollen“ ... „Relativ bessere Validitäten werden bei ‚breiten‘ Faktoren, d.h. bei ‚second-order‘-Faktoren erzielt. Allerdings ergibt auch eine einfache und ökonomische Rating-Skala unter Umständen dasselbe, da Korrelationen in der Höhe von .60 und mehr erzielt wurden zwischen Fragebogen und direkter Selbsteinstufung.“

(6) „Die volle Misere von Fragebogenvaliditäten zeigt sich bei Erstellung von Interkorrelationsmatrizen verschiedener Testbatterien, in denen die Ergebnisse unterschiedlicher methodischer Techniken gleichen und ungleichen Gültigkeitsanspruches zum Zwecke der Überprüfung der konvergenten und diskriminanten Validität eingesetzt wurden“ ...

In diesen Thesen ist der notorisch mangelnde Nachweis konvergenter-diskriminanter Validität im Sinne von Campbell und Fiske angesprochen. Wenn multiple Operationalisierungen (multi-methodale Diagnostik) eines psychologischen Konstrukts versucht werden, dann sind die Ergebnisse gewöhnlich enttäuschend (siehe Jackson & Paunonen 1980, Fahrenberg 1982, 1987b). Dies ist jedoch ein fundamentales Problem operationaler Definitionen in der empirischen Psychologie und nicht etwa ein spezielles Problem des Systems der Persönlichkeitsinventare. Auch wenn ein (zu) großer Teil der Persönlichkeitsforschung von Persönlichkeitsinventaren ausgeht, können die allgemeinen Schwierigkeiten der Eigenschaftsdefinition nicht allein auf diese Methodik geschoben werden. Statt diskrepant erscheinende Daten dieser Informationsquellen auszuklammern, müssen die verwendeten Konstrukte präzisiert und eventuell in Subkonstrukte aufgegliedert werden (Fiske 1978). Irritierend bleibt, dass innerhalb dieses Testsystems gleichlautend etikettierte Skalen vielfach nur mittlere statt hoher Interkorrelationen aufweisen (siehe Angleitner 1976, Literaturübersicht in der 3. Aufl. sowie Abschnitt 6 dieser Handanweisung). Die üblichen Kurzbezeichnungen der Skalen sind ohne nähere Angaben über die hauptsächlichen Subkonstrukte des Eigenschaftsbegriffs missverständlich. Dies gilt auch für das FPI.

Die Überlegenheit direkter Selbsteinstufungen im Vergleich zu der Beantwortung einer aus vielen Fragen bestehenden Skala ist seit Peterson (1965), Hamilton (1971), John & Keil (1972), Burisch (1984a,b), Rosch, Chassein, Semin & Krolage (1984) wiederholt behauptet, aber nur relativ selten untersucht worden. Hohe Korrelationen würden natürlich gegen das zeitraubende Ausfüllen des Fragebogens sprechen. Entsprechende Analysen zum FPI von Schott & Lass, Burisch, Sack (siehe 3. Aufl.) zeigen, dass einer verbindlichen Prüfung dieser zunächst einfach erscheinenden These manche Schwierigkeiten entgegenstehen. Es gibt „Übersetzungsschwierigkeiten“ der FPI-Eigenschaftsbegriffe, die sich nicht ohne weiteres, also nicht ohne Erläuterungen, als eindimensionale Skalen zur Selbsteinstufung formulieren lassen, und die Stabilitätskoeffizienten der Selbsteinstufungen scheinen niedriger zu sein als die der Fragebogenskalen. Wenn direkte und indirekte Selbstbeschreibungen nur in einigen Eigenschaftsbegriffen ca. 0.7 und in anderen deutlich niedriger übereinstimmen, so kann dies mehrere Gründe haben, die weiter zu untersuchen wären (siehe auch J.U. Schmidt, 1991). Die FPI-Konstrukte sind komplexer als übliche, adjektivische Selbsteinstufungsskalen, und die FPI-Skalen gestatten vermutlich eine bessere interindividuelle Differenzierung bei zugleich höherer Stabilität. Schahn und Amelang (1992) haben auf die Mittelwertunterschiede von Selbst- und Fremdbeurteilungen verwiesen.

(7) Aus den Thesen (5) und (6) sowie aus dem Eindruck, dass lediglich zwei bis drei Faktoren über die gesamte Literatur hinweg als „zweifelsfrei invariant“ zu sehen sind, stellt sich für Janke die Frage, ob „Fragebogendimensionen nicht lediglich die konnotativen Charakteristika unserer Sprache widerspiegeln, sich also überhaupt keine Eigenschaften unserer Probanden ausdrücken“ (1973, S. 45).

Diese Deutung stützt sich u.a. auf Arbeiten wie die von Ertel (1970), welche eine Kongruenz der drei Dimensionen des semantischen Differenzials und der drei Persönlichkeitseigenschaften Extraversion, Emotionalität und Rigidität behauptet. Dies ist zwar eine psycholinguistisch interessante Perspektive, aber auch ein besonderer Reduktionismus. Selbst wenn sich die Items oder Skalen – wie beliebiges anderes Material – in jenen semantischen Raum einordnen lassen, ist damit noch nicht die reale interindividuelle Varianz eines mehrdimensionalen Inventars ausgeschöpft oder als redundant nachgewiesen. Dies würde andere, z.B. kanonische Analysen verlangen. Aber die Sorge ist doch gering, dass sich ein Inventar oder gar die menschliche Persönlichkeit dreidimensional und erschöpfend nach Erregung, Valenz und Potenz konstituieren.

Die Frage, ob es eine kleine Zahl invarianter Faktoren in Persönlichkeitsinventaren gibt, ist durch McCrae und Costa (1982) wiederbelebt worden. Es wird postuliert, dass in Persönlichkeitsinventaren sowie Selbst- und Fremdeinstufungen fünf verhältnismäßig robuste Persönlichkeitsfaktoren nachzuweisen sind: Neurotizismus (Emotionale Labilität), Introversion-Extraversion, Offenheit für Erfahrung, Agreeableness (Verträglichkeit) und Conscientiousness (Gewissenhaftigkeit) – nach den Anfangsbuchstaben der ersten drei Faktoren auch als NEO-Persönlichkeitsinventar bzw. Fünf-Faktoren-Modell der Persönlichkeit bezeichnet (siehe Borkenau & Ostendorf, 1993).

Die aktuelle Auseinandersetzung über Postulate von drei, fünf, sechzehn oder mehr grundlegenden Persönlichkeitseigenschaften (Bartussek 1996; Becker, 1996, 2000; Borkenau & Ostendorf 1993, Costa & McCrae 1992; Eysenck 1992; Eysenck & Eysenck, 1976; John 1990; Zuckerman et al., 1991) muss bei methodenkritischer Sicht ziemlich müßig erscheinen. Einwände sind hinsichtlich der Kriterien der Konvergenz/Divergenz und Robustheit vorzubringen, denn die relative Invarianz müsste hinsichtlich der drei Datenbereiche, aber auch hinsichtlich definierter Subpopulationen, d.h. zunächst einmal durch bevölkerungsrepräsentative Erhebungen und systematische Analysen von Alters-, Geschlechts- und Statusgruppen, gezeigt werden. Daran mangelt es bisher. Der wesentliche Einwand richtet sich jedoch auf die Beliebbarkeit der lexikalischen Abgrenzungen bzw. der Auswahlentscheidung hinsichtlich der Item- bzw. Skaleneinhalte. Der Selektionsbias kann am Beispiel der körperlichen Beschwerden erläutert werden.

Körperliche Beschwerden gelten nach konventioneller Meinung kaum als Bestandteil der „personality sphere“. Da die meisten Persönlichkeitsinventare den Bereich körperlicher Beschwerden ausblenden oder nur wenige Allgemeinbeschwerden und Befindensstörungen berücksichtigen, sind diese sehr konsistenten und habituellen Personenunterschiede und ihre Beziehungen zu anderen Konstrukten im Fünf-Faktoren-Modell überhaupt nicht erfasst. Das FPI enthält außer der Skala „Körperliche Beschwerden“ auch eine Skala „Gesundheitssorgen“. Es wäre eine definitorische Willkür, solche Dispositionen nicht als Persönlichkeitsmerkmale aufzufassen. Im Erleben der Personen und in ihrer möglichen

Relevanz für Gesundheits- und Krankheitsverhalten werden diese Aussagen wahrscheinlich wichtiger sein als viele andere der z.B. im NEO-Inventar abgefragten stilistischen und eher peripheren Inhalte. Das Fünf-Faktoren-Modell kann auch als Ausdruck eines gemeinsamen Selektionsbias verstanden werden (siehe Abschnitt 8.9).

Da zentrale Felder möglicher Selbstschilderungen und Selbstbeurteilungen wie die körperlichen Beschwerden, aber auch Einstellungen, Interessen, Lebensqualität und Zukunftsperspektiven sowie psychopathologische Tendenzen in der Regel ohne hinreichende Begründung ausgeklammert bleiben bzw. extrem unterrepräsentiert sind, können jene unrepräsentativen Abbildungen des Universums möglicher Items grundsätzlich nicht zu verbindlichen Taxonomien führen. Die Annahme, es ginge hier um grundlegende Dimensionen, muss außerdem kritisiert werden, wenn übersehen wird, dass es sich zunächst nur um Dimensionen der Selbstschilderungen handelt und eventuell noch um Fremdbeurteilungen, die semantisch und durch Kommunikationsprozesse mit jenen Selbstschilderungen konfundiert sind. Die Korrelation mit manifestem Verhalten ist auf unabhängige Weise nicht leicht zu zeigen, und die Fundierung durch neuro-psycho-physiologische Systemeigenschaften ist ungeklärt; einige Hypothesen sind als widerlegt anzusehen (siehe Abschnitt 7.2).

Persönlichkeitsinventare mit zwei bis fünf Dimensionen (wie im MPI/EPQ oder NEO-FFI) werden für manche Fragestellungen nützlich sein, in vielen Anwendungsgebieten wird jedoch der Wunsch nach einem mehrdimensionalen, ca. 10 bis 15 Dimensionen umfassenden Beschreibungssystem für individuelle Differenzen bestehen. So gibt es inzwischen auch für das Inventar mit den „großen Fünf“ Rekonstruktionen, welche 13 Sub-Skalen definieren (Saucier, 1998).

(8) „Die Ergebnisse von Fragebogen sind nur unter bestimmten situativen Bedingungen verwertbar. Eignungsdiagnostisch-selektive Situationen sind in der Regel für Fragebogenanwendungen nicht angemessen. Auch bestimmte Personenparameter, vor allem niedrige Intelligenz, schränken die Brauchbarkeit ein.“

Diese Einschränkungen treffen auch für viele andere Assessmentverfahren zu, doch gelten Persönlichkeitsfragebogen als relativ anfällig für die Effekte spezieller Testmotivationen. So hatten die Testautoren von Anfang an die Rolle einer günstigen Testmotivation des Probanden betont und ausgeführt, dass typischer Weise an Personen in Forschungssituationen, an Patienten in Beratungs- oder Therapie-Situation oder an anonym bleibende Probanden gedacht sei und weniger an Personen in Prüfungs-, Bewerbungs- und Personalauslese-Situationen. Die zunehmende Verbreitung und Anwendung auch zu Selektionszwecken hat die Testautoren dann veranlasst, diese Warnung zu unterstreichen. Allerdings stützt sich diese Empfehlung vor allem auf grundsätzliche Risikoabwägungen, denn die empirischen Ergebnisse zum Thema „Soziale Erwünschtheit“ bzw. „Offenheit“ in Persönlichkeitsfragebogen sind nicht so eindeutig und so effektstark wie vielleicht zeitweilig angenommen worden ist (siehe Literaturübersicht 3. Aufl. sowie Abschnitt 2.9 und 7.7 dieser Handanweisung).

(9) „Nachweisbare Antworttendenzen wie Reaktionen im Sinne sozialer Erwünschtheit oder Zustimmung werfen die Frage auf, wie weit wesentliche Varianzanteile von Fragebogen nicht zu Lasten inhaltsunabhängiger Antwortstile gehen“.

Diese Antworttendenzen, insbesondere die Offenheit (oder „Lügenskala“) bzw. Soziale Erwünschtheit der Antworten, bilden ein wesentliches Methodenproblem der Fragebogen (siehe Abschnitt 2.9). Neuere Methodenstudien haben das Bild erheblich kompliziert und schließen einfache Antworten oder schematisch zu handhabende Korrektur- und Kontroll-Skalen aus. In mehreren Arbeiten wurden kognitive Prozesse der Itembeantwortung und individuelle Operationscharakteristiken untersucht (z.B. Cliff 1977, Voyce & Jackson 1977, De Boeck 1978; Übersicht bei Jackson & Paunonen 1980, Fiske 1981, Mummendey, 1999; Tränkle 1983, Angleitner & Wiggins 1986). Die Auffassungen über die „inhaltliche“ individualdiagnostische oder „bloß stilistische“ Bedeutung solcher Urteils- und Beantwortungstendenzen sind kontrovers. Es soll hier nicht bestritten werden, dass der Prozess der Beantwortung, stilistische Aspekte des Antwortverhaltens und die Zusammenhänge von sprachlichen Merkmalen der Items mit Antwortzeiten und testmethodischen Gütekriterien wichtige Forschungsgebiete sind (Amelang, Breit & Schneyer 1985, Angleitner & Wiggins 1986, Schneider-Düker & Schneider 1977, Schmolck 1981, Schwarz & Sudmann, 1992; Tränkle 1983). Es wäre jedoch eine außerordentliche Simplifikation des psychologischen Sachverhalts, die Skalen des MMPI, EPI, FPI usw. mit solchen Stil-Dimensionen gleichzusetzen und die Eigenschaftskonzepte auf Unterschiede des formalen Antworten-

verhaltens zu reduzieren. In den verbalen Selbstschilderungen eines Individuums allein oder doch primär stilistische Informationen sehen zu wollen, wird den empirischen Validitätsbelegen der Persönlichkeitsinventare sicher nicht gerecht.

Von vielen Probanden und Testkritikern immer wieder betont werden die semantischen Probleme der Itembeantwortung. Eine vage formulierte Frage oder sprachliche Eigenarten könnten zusätzlich noch Reaktionstendenzen fördern. Tatsächlich enthalten alle Persönlichkeitsinventare mehr oder minder viele Items mit unnötigen sprachlichen Eigenarten oder Defiziten sowie doppelte Verneinungen und andere Ungeschicklichkeiten (Angleitner 1976, Löhr & Angleitner 1980, Tränkle 1983). Wenn solche Items in den Pool gelangt sind, ist es nachträglich nicht ganz einfach, solche Mängel zu beseitigen. Die mangelnde Eindeutigkeit der Items wird auch ein zentraler Mangel bleiben. Selbst wenn in der Aussage „Haben Sie häufig Kopfschmerzen?“ das Wort „häufig“ durch Quantoren „1 mal im Jahr“ bis „täglich“ ersetzt würde (siehe Fahrenberg 1975, 1994a; Hampel & Fahrenberg 1982) blieben Schwierigkeiten bestehen; solche Quantoren passen nicht systematisch zu beliebigen Sachverhalten und der hier wesentliche Begriff Kopfschmerzen bleibt ohnehin nach Intensität und Qualität prinzipiell undefiniert und nicht objektivierbar. Diese Itembeantwortungen sind Selbstbeschreibungen mit den charakteristischen Vorzügen und Nachteilen dieser Informationsquelle über die menschliche Persönlichkeit.

Erwähnenswert ist an dieser Stelle, wie sich die als Normstichprobe befragten Personen nachträglich zu den 240 Items des für die Erhebung verwendeten Fragebogens geäußert haben. Im Mittel wurden 76% der Items subjektiv sicher mit „stimmt“ oder „stimmt nicht“ beantwortet, 86% der Items wurden als verständlich bezeichnet, 57% als auf die eigenen wesentlichen Charaktereigenschaften zutreffend und 22% der Items wurden als zu persönlich oder zudringlich erlebt.

Bei der Beantwortung einiger Fragen spielt – auch im Fragebogen – der Kontext des Items eine Rolle. In einer Methodenstudie anhand des FPI-A konnten Krampen, Hense und Schneider (1992) zeigen, dass im Vergleich zur Standardreihenfolge der Items eine Blockbildung, d.h. Zusammenfassung der Items jeder Skala, signifikante Auswirkungen auf Mittelwerte, innere Konsistenz sowie Interkorrelationen einiger Skalen haben kann. So ergab die Blockbildung in der Skala Geselligkeit einen um 0.5 höheren und in der Skala Emotionale Labilität einen um 1.2 niedrigeren Mittelwert (je Bedingung etwa 30 Psychologiestudenten). Als Effektstärke geben die Autoren  $w^2=.09$  im Vergleich zu einer ebenfalls variierten Instruktionsbedingung mit einer größeren Effektstärke von  $w^2=.19$  an. Die Generalisierbarkeit dieser Befunde ist nicht ohne weiteres zu beurteilen, doch legen sie Zweifel an der nicht selten geübten Praxis nahe, aus einem Inventar einzelne Skalen herauszunehmen und separat zu verwenden (zu Reihenfolge-Effekten im Persönlichkeitsfragebogen siehe auch Knowles et al., 1992). Rost und Hoberg (1997) und Franke (1997, 1998a,b) fanden nur geringe, praktisch kaum relevante Effekte von Veränderungen der Itemposition.

Über die hier referierten und kommentierten neun Thesen hinaus hat Janke (1973) noch weitere Methodenprobleme der Persönlichkeitsfragebogen genannt: zu geringe Beachtung von set- und style-Variablen, Probleme der Itemsammlung, diskutable testtheoretische Modelle der Itemselektion und Vernachlässigung von Validierungsfragen (siehe auch Angleitner & Wiggins, 1986; Gynther & Green, 1982; von Hagen, 1997; Lösel, 1988; Mummendey, 1999; Tränkle, 1983). Eine hoffnungsvolle neue Perspektive sahen nicht wenige Autoren in der Differenzierung von Reiz- und Reaktions-Modi oder Eigenschafts- und Zustands-Modi und in der entsprechenden Konstruktion sog. S-R-Fragebogen (siehe Endler & Hunt, 1969) oder sog. State-Trait-Fragebogen (wie Laux, Glanzmann, Schaffner & Spielberger 1981). Demgegenüber bauen andere Autoren gerade auf die Aggregation einzelner Daten über Situations-, Reaktions- und Termin-Modi (siehe Epstein 1980, Paunonen 1984, Schweizer 1986, Wittmann 1988). Andererseits wird immer wieder auf die unerlässliche, aber schwierige Hermeneutik der Persönlichkeitsbeschreibung insbesondere bei Fragebogenitems aufmerksam gemacht (siehe z.B. Gergen, Hepburn & Fisher 1986) und eine kritische Analyse der operationalen Definition von Konstrukten verlangt (u.a. Hogan & Nicholson 1988, Nicholls, Licht & Pearl 1982).

Die hier an den prägnanten Thesen Jankes orientierte Einleitung der FPI-Handanweisung konnte nur eine Problemübersicht zu geben versuchen. Neuere Methodenstudien haben in der Regel ergeben, dass die testmethodischen Sachverhalte schwieriger sind als früher häufig angenommen wurde. Einfache Schlussfolgerungen zu Stabilität, Konsistenz, Konstruktvalidität, Antworttendenzen usw. von Persönlichkeitsfragebogen können nicht formuliert werden. Die hauptsächlichsten Einwände treffen auch auf andere persönlichkeitsdiagnostische Verfahren zu, dies gilt ähnlich für die mög-

lichen Auswirkungen unzureichender Testmotivation und für die wesentliche Rolle der jeweiligen Fragestellung und des Anwendungskontexts. Selbst die linguistischen und semantischen Schwierigkeiten und Defizite der Itemformulierung teilen die Fragebogenmethoden mit dem psychologischen Interview und anderen verbalen Methoden. Nach Phasen einseitiger Überschätzung oder Unterschätzung von Persönlichkeitsfragebogen wird die Einsicht, dass es sich in mancher Hinsicht höchstens um relative Vorzüge oder relative Nachteile im Rahmen grundsätzlicher Operationalisierungsschwierigkeiten psychologischer Konstrukte handelt, zu einer ausgewogeneren Würdigung beitragen können (siehe die aktuelle Diskussion in Kapitel 8).

Die Testautoren halten an der Überzeugung fest, dass ein mehrdimensionales Persönlichkeitsinventar wie das FPI für bestimmte psychologische Aufgabenstellungen nützlich sein kann und deshalb in Abständen neu zu normieren und empirisch weiterzuentwickeln ist.

## **2 Konstruktion des FPI-R**

### **2.1 Konzeption des FPI und Absichten der Revision 1982**

Das Freiburger Persönlichkeitsinventar ist aus den theoretischen Interessen der Autoren an bestimmten Persönlichkeitseigenschaften entstanden. Für die Bereiche Emotionalität (Neurotizismus), Extraversion – Introversion, Aggressivität und psychovegetative Labilität wurde 1968 mit dem ALNEV Fragebogen eine Vorform des FPI geschaffen, um für die eigenen Forschungsvorhaben, neben anderen Verfahren, auch über standardisierte Fragebogenskalen verfügen zu können. Jede Skalenkonstruktion war hier zugleich eine Operationalisierungsstudie zu dem theoretischen Konstrukt und dessen Subkonstrukten, die z.T. in weiteren Arbeiten verfolgt wurden (Fahrenberg 1975, 1994a; Hampel & Selg 1975).

In Rezensionen des FPI und in vergleichenden Übersichten über Persönlichkeitsinventare ist gelegentlich die Meinung zu finden, das FPI sei ohne Persönlichkeitstheorie, gleichsam kraft faktorenanalytischer Methodik, a-theoretisch-deskriptiv entstanden. Deshalb soll hier – ohne auf das so missverständliche Wort „Theorie“ in der Psychologie näher einzugehen – noch einmal die Konzeption festgehalten werden: Die FPI-Skalen sind nicht aus einem spekulativen „Gesamtsystem der Persönlichkeit“ (siehe die Übersicht von Bartussek 1996) abgeleitet worden, sondern als einzelne Eigenschaftskonzepte nach den Interessen der Autoren mit Ergänzungen und Erweiterungen im Zuge dieser Forschungsvorhaben (siehe Kapitel 2.3.1 über neue Konstrukte im FPI-R). Es ist ja gerade eine theoretische und kritische Einsicht der Persönlichkeitsforschung, dass der Anspruch solcher Gesamtsysteme nicht mehr zu rechtfertigen ist, sondern wesentlich bescheidenere Konstruktionen im Sinne von Fiske (1978) als construct-operation units entwickelt werden müssen.

Zunehmend wurden in der testmethodischen Entwicklung und Anwendung des FPI die Perspektive der Selbstbeschreibung betont und die anfängliche Erwartung hinsichtlich objektiver Korrelate im manifesten Verhalten und in psychophysiologischen Maßen reduziert. Dies ist ein Ergebnis eigener Validierungsversuche, die sich sehr intensiv vor allem mit dem Bereich Emotionale/Vegative Labilität, Körperliche Beschwerden befasst haben, und ist auch durch den Stand der Literatur – bei kritischer Einstellung – gestützt. Multiple Operationalisierungen im Sinne des Multitrait-Multimethod-Ansatzes oder im Sinne des Drei-(Mehr-)Ebenen-Ansatzes und die Chancen einer Vorhersage relevanter Verhaltenskriterien mit hinreichender Effektstärke haben sich generell als sehr schwierig erwiesen (siehe Kapitel 6). Durch die Konzepte der modernen Assessmenttheorie und durch Nutzenüberlegungen, welche explizite Rechenschaft über den Entscheidungsnutzen eines Verfahrens hinsichtlich präziser Zielsetzungen innerhalb einer Mittel-Ziel-Analyse verlangen, ist das Anspruchsniveau an befriedigende Validitätsstudien gestiegen.

Die theoretischen Absichten der Autoren haben die Themen des Itempools bestimmt und auch die Skalenkonstruktion teilweise beeinflusst (siehe Kapitel 2.2). Die Entscheidung über den psychologischen Umfang der Konstrukte/Skalen wurde jedoch auch durch die faktoren- und clusteranalytische Reduktion dessen, was die Durchschnittsbevölkerung empirisch zur laienpsychologischen Selbstbeschreibung/Selbstbeurteilung leistet, bestimmt. Es handelt sich um relativ robuste, empirisch begründete Schemata bzw. Dimensionen der Selbstbeschreibung mit den praktisch wichtigen Vergleichsmöglichkeiten zwischen Personen auf der Basis einer bevölkerungsrepräsentativen Normierung. Die individu-

ellen Testwerte haben – so ist zumindest einigen der Validierungsstudien zu entnehmen – über die Selbstbeschreibung hinaus eine für Aufgaben der Diagnostik bzw. des Assessment wichtige psychologische Bedeutung. Mit den Begriffen „standardisiertes Beschreibungssystem“ für individuelle Unterschiede unter Bezug auf persönlichkeits-theoretische Konzepte und auf „multireferenzielle Laienkonstrukte“ kann der eine Aspekt dieser Operationalisierung gekennzeichnet werden. Daneben bleibt die andere Zielsetzung bestehen, in diesem Operationalisierungsprozess bei der Explikation von einzelnen theoretischen Konstrukten, u.U. durch Präzisierung kleinerer Muster und durch Reduktion von Überschussbedeutungen, Fortschritte zu erreichen.

Das Freiburger Persönlichkeitsinventar FPI (1. Auflage 1970) war anhand des 1965 bis 1969 gesammelten Materials entwickelt und normiert worden. Die sog. Eichstichprobe (N=630), welche hinsichtlich der Verteilung von Geschlecht, Alter und Schulbildung grob dem Bevölkerungsbild entsprach, wurde ergänzt durch zwei absichtlich sehr heterogen gewählte Analysenstichproben: „Klinische Gruppe“ N=300 Männer und „Junge Männer“ N=350, zuzüglich einer Gruppe „Junge Frauen“ N=175. Selbstverständlich war eine breitere Datenbasis von Anfang an wünschenswert, doch standen dem nicht allein der außerordentliche Aufwand einer bevölkerungsrepräsentativen Normierung entgegen, sondern auch die Skepsis, ob sich ein Fragebogen, speziell ein Persönlichkeitsfragebogen, überhaupt für eine Normierung anlässlich einer Umfrage oder Panel-Erhebung im Stil der Markt- und Meinungsforschung eignet. Das FPI-Material war ja mit Unterstützung von Studenten und Diplomanden im Bekanntenkreis und in vielen verschiedenen Institutionen gesammelt worden, wobei trotz aller Probleme dieser Erhebungsweise wahrscheinlich eine relativ günstige Testmotivation der Befragten angenommen werden durfte.

Die Testautoren hatten mit Unterstützung anderer Testbenutzer im Laufe der Jahre eine umfangreiche Informationsbank (178 Probanden-Gruppen in der 3. Aufl. Fahrenberg & Selg, 1978) geschaffen, um wenigstens Vergleichsdaten für spezielle Populationen und Fragestellungen verfügbar zu machen. Doch der geringe Umfang der Norm-Stichprobe N=630 blieb unbefriedigend. Außerdem legten Reanalysen zur Faktorenstruktur des FPI, allerdings sämtlich nicht an bevölkerungsrepräsentativem Material, eine Reduktion der Faktorenzahl nahe, wobei eine – schon in der 1. Auflage diskutierte – Lösung mit fünf Faktoren infrage kommt (Literatur siehe Fahrenberg, Selg & Hampel, 1978; König & Schmidt, 1982; Potreck, 1983; Jehle, 1988). So ergab auch eine Faktorenanalyse anhand eines aus der Freiburger Informationsbank konstruierten FPI-G-Datensatzes von N=3325 nicht klinischen Personen eine inhaltlich und formal verhältnismäßig befriedigende Lösung mit fünf Faktoren bzw. Skalen zu je 14 Items: Faktor A „Nervosität/Emotionalität“ mit Items aus FPI 1 und 3; Faktor B „Alltägliche Normverletzungen und Ungeniertheit“ mit Items aus FPI 9 und 2; Faktor C „Geselligkeit/Lebhaftigkeit“ mit Items aus FPI 5 und 8; Faktor D „Aggressive Erregbarkeit“ mit Items aus FPI 2, 4 und 7; Faktor E „Gelassenes Selbstvertrauen“ mit Items aus FPI 6 und 4 (siehe Tabellen in Hampel & Fahrenberg, 1983).

Diese Konfiguration hat Ähnlichkeiten mit dem Fünf-Faktoren-Modell, das von Mc Crae und Costa (1987) postuliert wurde (Zur Kritik an dem methodisch und empirisch sehr zweifelhaften Anspruch dieser Konzeption siehe Kapitel 1.2 Punkt 7 sowie Abschnitt 8.9).

In einer Reanalyse konnten Bengel & Wittmann (1983) zeigen, dass in den FPI-Normen für 15- bis 30-jährige Männer der sog. Eichstichprobe im Vergleich zu einer Archivstichprobe aus neuerem Material (N=442) signifikante Unterschiede in fünf Skalen bestehen, welche auch in einem nicht geringen Anteil der Fälle praktisch zur unterschiedlichen Klassifikation von Probanden führen würden. Allerdings bleibt offen, ob diese Effekte auf mangelnde Repräsentativität eines oder beider Datensätze und/oder tatsächliche Veränderungen in den Selbstschilderungen der zwischen 1965/69 und der zwischen 1972/78 Befragten bedingt sind (siehe die Reanalyse zu dieser Frage, Fahrenberg et al., 1983). Neue Befunde zur Veränderung der Normwerte über einen Zeitraum von 17 Jahren werden im Abschnitt 3.6 dargestellt.

Die Testautoren hatten sich 1980 für eine neue Normierung des FPI entschieden. Angesichts des damit verbundenen großen finanziellen und sonstigen Aufwandes lag es nahe, diese Normierungsstudie zugleich zur Revision zu nutzen. Die 1982 abgeschlossene Studie ist aus der Verbindung der folgenden Absichten zustande gekommen:

1. Bevölkerungsrepräsentative Normen sollten an einer Stichprobe von ca. 2000 Personen gewonnen werden, um nach Geschlecht und Alter, eventuell auch Schulabschluss, differenzierte Normen hinreichenden Umfangs zu gewinnen. Es ist möglich, dass die FPI-Normen u.a. wegen veränderter sozialer Bedingungen und Leitbilder, wegen mancher

Veränderungen in Sprachgebrauch und Redewendungen oder wegen veränderter Einstellung zu psychologischer Diagnostik nicht mehr zutreffen.

2. Sprachlich und psychologisch sollten jene Items verändert werden, welche nach eigenen Erfahrungen und nach kritischen Anmerkungen anderer Autoren (Wolfram & Moltz 1974, Löhr & Angleitner 1980) verbesserungsbedürftig sind.
3. Aus praktischen und testkonstruktiven Gründen sollte das Inventar gekürzt werden. Das FPI-G enthielt 212+2 Items, um Parallelformen A und B zu ermöglichen. Inzwischen verwendeten nur noch ca. 19% der Testbenutzer die Form G, dagegen ca. 48% die Form A und sogar ca. 7% die Form K, welche nur für spezielle Zwecke, aber nicht für individualdiagnostische Aufgaben dienen soll. Außer einer Reduktion der Itemzahl wurde auch eine Reduktion der Faktorenzahl angestrebt.
4. Die Chance einer großangelegten Erhebung sollte auch genutzt werden, neue Skalen zu entwickeln. Die Testautoren haben sich hier, primär ihren Interessen in ihren Arbeitsgebieten folgend, für fünf Konstruktbereiche entschieden: Prosozialität, Lebenszufriedenheit, Leistungsorientierung, Beanspruchung, Gesundheitssorgen. Weiterhin sollten außer soziodemografischen Merkmalen der Befragten noch psychologisch interessante Zusatzinformationen als Validitätshinweise erhoben werden.
5. Die Konstruktion eines revidierten Inventars FPI-R sollte jedoch nicht zu einem Abbruch der bisherigen FPI-Diagnostik oder zu einer Aufgabe der vielfältig empirisch verknüpften Konstrukte, welche mit den FPI-Skalen operationalisiert sind, führen. Dem Benutzer soll die Wahl zwischen dem eingeführten, jetzt neu normierten FPI und einer weiterentwickelten, einige zusätzliche Gültigkeitsbereiche umfassenden Form FPI-R bleiben. Das mit jeder Revision einer eingeführten Methode verbundene Dilemma von Kontinuität und Vergleichbarkeit gegenüber Verbesserung durch Revision wurde in den realistischen Grenzen so gelöst, dass damals nicht die Gesamtform FPI-G, sondern die gängigere Form FPI-A als FPI-A1 neu normiert worden ist.

Im Folgenden werden die 1982 durchgeführte und 1984 publizierte Revision und die Normierung des FPI-R dargestellt. In Kapitel 3 wird anschließend die 2001 publizierte Untersuchung zur Überprüfung der Skalen und der Normen beschrieben.

## 2.2 Strategien der Skalenkonstruktion

Die Skalenkonstruktion des Persönlichkeitsinventars kann nicht kriterienbezogen geschehen. Wegen der Breite des angestrebten Geltungs- und Gültigkeits-Bereichs wäre eine konsequente Selektion nur solcher Items bzw. Skalen, für die eine empirische Prädiktorleistung nachgewiesen ist, kaum zu leisten. Die Validierung ist – wie seit der 1. Auflage kontinuierlich dokumentiert – nur als Entwicklung und Verknüpfung einzelner Gültigkeitshinweise möglich. So wird auch für die Revision des FPI die früher befolgte Strategie beibehalten, welche hier skizziert werden soll. Diese Strategie nutzt die konventionelle Itemanalyse und die Faktorenanalyse, ohne sich durch diese Verfahren schematisch einschränken zu lassen.

Die Autoren gehen von bestimmten psychologischen Konstrukten aus, die ihnen aufgrund ihrer Erfahrungen und aufgrund der Literatur interessant und wichtig erscheinen. Es sind relativ weitgefasste Konstrukt-Bereiche, welche sich meist in einen Konstrukt Kern sowie verschiedene Subkonstrukte und Facetten gliedern lassen (siehe Fiske 1971). In bis zu 50 Items wird das Konstrukt erläutert, wobei außer psychologischen auch sprachlich-stilistische Varianten formuliert werden. Die empirische Phase der Testkonstruktion soll klären, ob in den Selbstschilderungen einer repräsentativen Bevölkerungstichprobe diese Konstrukte als relativ invariante Strukturen wiederzuerkennen sind. Dies wäre eine notwendige, aber wegen des Gültigkeitsproblems dieser Fragebogendaten noch keine hinreichende Bedingung, die Itemgruppierungen zur Bestimmung psychologischer Dispositionsprädikate zu verwenden.

Es wird ein breiter Geltungsbereich angestrebt mit Ausnahme von Personen, die eine geringere als Hauptschulbildung haben und mit Ausnahme von psychopathologisch stark behinderten Patienten. Auf Skalenentwicklungen für spezielle Subpopulationen, z.B. für Männer, für Frauen oder für ältere Personen wird verzichtet, auch wenn solche Differenzierungen testkonstruktiv u.U. besonders prägnant wären. Wichtig ist jedoch die Kontrolle der Skalenkonstruktion am selben Datensatz mit einer Aufgliederung nach bzw. einer Auspartialisierung von Effekten des Geschlechts, des Alters und der Schulbildung, um zu erkunden, wie universell die Skalenstruktur ist.

Die Faktorenanalyse dient hier zur Ordnung der Items und als Entscheidungshilfe hinsichtlich der Anzahl der genauer auszuarbeitenden Konstrukte. Der Scree-Test des Eigenwertverlaufs wird als ein Hinweis verstanden, doch sind die psychologische Interpretierbarkeit der Ladungsmuster und die verfügbare Anzahl markanter Items praktisch wichtiger. Nachdem die Anzahl der Faktoren festgelegt ist, werden in vielen Schritten Items ausgewählt, aber auch wieder ausgetauscht, um schließlich Skalen hinreichender Länge für jedes Konstrukt zu erreichen. Die Itemselektion richtet sich erstens nach der Ladung (möglichst hoher Anteil  $a^2/h^2$ , möglichst wenig komplexe oder systematische Nebenladungen auf anderen Faktoren). Zweitens werden aufgrund der begleitenden Itemanalysen Trennschärfe-Indizes berücksichtigt und mittlere Itemschwierigkeiten bevorzugt: einerseits als Voraussetzung geeigneter Randverteilungen bei der Berechnung der übrigen Itemstatistiken, andererseits im Hinblick auf näherungsweise normale Verteilungen der künftigen Skalenwerte.

Eine solche konventionelle Strategie der Itemselektion führt zur Bevorzugung inhaltlich sehr ähnlicher Items, welche sich in Ladung und Trennschärfe wechselseitig maximieren und u.U. statt erwünschter Parallelmessungen nur redundante, psychologisch überflüssige und den Probanden u.U. störende Formulierungs- und Polungs-Varianten bilden. Diese Pseudo-Homogenität kann vermieden werden, indem kompensatorisch bestimmte Items ausdrücklich vorgezogen werden, weil sie wichtige Facetten des gemeinten Konstrukts wiedergeben, auch wenn deswegen Faktorladung, Faktorreinheit, Trennschärfe, Skalenkonsistenz und Testwertverteilung schwächer werden. Dieses dritte Prinzip der Itemselektion ist durch formale Kriterien nicht zu ersetzen. Es handelt sich um die systematische Elaboration von Konstrukt-Indikatoren-Beziehungen in einem psychologischen Interpretationsprozess, welcher zwar immer wieder an dem Datensatz kontrolliert wird, aber einen Ermessensspielraum behält.

Mit der Analyse binärer Itembeantwortungen sind zweifellos Probleme verbunden. So wirken sich die Randverteilungen (selbst wenn Schwierigkeitsindizes zwischen 0.20 und 0.80 eingehalten werden) auf den Korrelationskoeffizienten und dadurch auf die anschließende Faktorenanalyse aus (siehe Gaensslen & Schubö 1976, Gorsuch & Yagel 1981). Dennoch hat sich die Faktorenanalyse als Verfahren zur Exploration von Fragebogenitems und zur Ordnung nach ähnlichen Inhalten bewährt. Die metrischen Probleme dürfen auch nicht einseitig überbewertet werden im Vergleich zu den anderen, viel weitergehenden oder grundsätzlich fragwürdigen Annahmen der Fragebogen-Konstruktion: u.a. die Operationalisierung komplexer Konstrukte durch semantisch nicht eindeutige Fragen, die postulierte Kontextunabhängigkeit der Itembeantwortung, die unterstellte Additivität von Items, die dimensionale Auffassung persönlichkeitspsychologischer Dispositionsprädikate usw.

Statt der faktorenanalytisch und itemmetrisch gestützten Entwicklung der gemeinten Konstrukte könnte man andere Verfahren wie das logistische Grundmodell von Rasch oder nicht parametrische Clusteranalysen erwägen, die in der neueren Literatur von mehreren Autoren auch für Persönlichkeitsfragebogen empfohlen, allerdings nur sehr selten tatsächlich genutzt werden (zur Diskussion siehe u.a. Hehl 1981, Henning 1980, Johann & Bergemann 1985, Klapproth 1981, Nunnally 1978, Walkey & Green 1981). Es ist fraglich, welchen Nutzen multistrategische Studien haben, solange Entscheidungsmaßstäbe angesichts offensichtlicher Widersprüche fehlen. Bereits am FPI-G waren Rasch-Analysen unternommen und die modellunverträglichen Items mitgeteilt worden, ohne plausible inhaltliche oder formale Gründe (Schwierigkeit, Trennschärfe) feststellen zu können (siehe Womes 1977). Diese anderen Strategien werden hier nicht zur eigenständigen Konstruktion von Skalen verwendet, sondern nur um zu erkunden, wie invariant bzw. robust die neuen FPI-R-Skalen gegenüber Reanalysen unter anderen Modellannahmen sind. Wenn also im Abschnitt 2.6 und in ergänzenden Berichten (Hampel & Fahrenberg 1983, Fahrenberg, Ewert & Maier 1987) u.a. (1) Ergebnisse der schiefwinkligen Rotation der Primärfaktoren, (2) Clusteranalysen der 120 Items und (3) Skalenanalysen nach dem logistischen Grundmodell mitgeteilt werden, so dient dies dem Zweck, Konvergenzen und Divergenzen der Strategien aufzuzeigen, die Frage nach möglichen Konformitätsindizes (Gutjahr 1972) zu erneuern und sehr viel genauere Methodenstudien anzuregen. Vorstellbar sind auch Versuche, empirisch über die differenzielle Prädiktorleistung bzw. Validität von Itemgruppen zu entscheiden, welche aufgrund verschiedener Strategien gebildet wurden.

Bei der Revision des FPI wird nicht nur aus Gründen der Kontinuität an der hier skizzierten Strategie festgehalten, sondern auch aus der Überzeugung, dass dieser faktorenanalytisch und itemmetrisch gestützte Interpretationsprozess den facettenreichen Konstrukten der Persönlichkeitstheorie angemessener ist als das logistische Grundmodell.



## 2.3 Itempool

### 2.3.1 Neue Konstrukte im FPI-R

Die Anregungen für die Entwicklung zusätzlicher Skalen stammen vor allem aus dem Arbeits- und Interessenbereich der Autoren: psychophysiologisch orientierte Persönlichkeitsforschung, Evaluation von Psychotherapie und Rehabilitationsmaßnahmen insbesondere bei psychovegetativen Syndromen und Herz-Kreislauf-Krankheiten, Untersuchungen über Belastungs-Beanspruchungs-Prozesse (Stress) und über den Bereich Aggressivität und Prosozialität. Vermutlich wird diese Vorentscheidung über den angestrebten Gültigkeitsbereich eines Persönlichkeitsinventars auch von den individuellen und zeitgenössischen anthropologischen Perspektiven beeinflusst, wie Thomae (1968) in seiner Diskussion von normativen Leitbildern in Cattells und Eysencks Fragebogen hervorgehoben hat. So scheinen die Aspekte Lebenszufriedenheit und Beanspruchung (Stress) auch in der gegenwärtigen Fachliteratur besonderes Interesse zu finden. Die Testautoren sehen in diesen Konstrukten, zusammen mit sozialer Orientierung (Prosozialität, Altruismus), Leistungsorientierung, Gesundheitssorgen, eine psychologisch sinnvolle und praxisbezogene Erweiterung des bisherigen „FPI-Menschenbildes“.

Auf die zeitweilig erwogene Entwicklung stärker psychopathologisch orientierter Skalen, u.a. für spezielle Angstformen, Zwanghaftigkeit, Psychotizismus, wurde verzichtet, ebenso auf sehr spezielle Persönlichkeitsmerkmale wie Kontrollüberzeugungen und sog. Coping-Dispositionen. Für diese Zwecke existieren bereits Fragebogenskalen. Die ausgewählten Konstrukte sollen hier nur mit einigen Stichworten und Literaturhinweisen skizziert werden. Bei der folgenden Skalenentwicklung, die über zwei Prätests und die Repräsentativerhebung zur endgültigen Itemsélection führte, blieben nicht alle der hier genannten, ursprünglich gewünschten Konstrukt-Komponenten erhalten. Die Itempools dieser drei Erhebungen (330 Items, 80 Items, 240 Items, teilweise überlappend) sind dokumentiert (Hampel & Fahrenberg 1983).

#### *Lebenszufriedenheit und Selbstverwirklichung*

Zufriedenheit mit dem eigenen Leben, Partner, Beruf usw.; Zukunftsoptimismus; Selbstverwirklichung, persönliche Wertorientiertheit; Akzeptierung eigener Schwächen; verständnisvolle menschliche Beziehungen. Dieser zweifellos viele Aspekte umfassende Bereich sollte durch explorative Analysen auf die vermutete gemeinsame Varianz dieser Selbstbeurteilungen untersucht werden (siehe u.a. Shostrom 1964, Hogan 1969, Schwebcke, Lück & Jandron 1974, Bottenberg & Keller 1975, Silbereisen & Schulz 1977, Fahrenberg, Myrtek, Wilk & Kreutel 1986, Myrtek, Kreutel, Wilk, Welsch & Herzog 1987). Inzwischen ist mit dem Fragebogen zur Lebenszufriedenheit FLZ ein differenziertes Beschreibungssystem entstanden (Fahrenberg, Myrtek, Schumacher & Brähler, 2000).

#### *Soziale Orientierung*

Hilfsbereitschaft (Prosozialität, Altruismus); Fürsorglichkeit, Mitleid, Nachsicht und Sympathie für Benachteiligte und Schwache; Verständnis für Sorgen und Nöte der Mitmenschen; soziale Verantwortung (Solidarität). Dieser Bereich ist breiter gefasst worden als in vorhandenen Skalen für prosoziales Verhalten (Robinson & Shaver 1973, Bilsky 1980).

#### *Leistungsorientierung*

Erfolgstreben; ehrgeizig-konkurrierendes Verhalten; berufliches Engagement, auch auf Kosten des Privatlebens; Zielsetzungen, welche Zeitdruck und Ungeduld zur Folge haben; Neigung, das Lebens- und Arbeitstempo zu beschleunigen; geistige Beweglichkeit; geringe Ausprägung von Langeweile, Apathie, Desinteresse. Mehrere dieser Merkmale gehören auch zum Konstrukt Leistungsmotivation (Ehlers 1965; Fineman 1977) oder sind wichtige Komponenten des sog. Typ A-Verhaltens (siehe Dembroski et al., 1978; Myrtek 1983, 1995; Myrtek, Schmidt & Schwab 1984).

#### *Beanspruchung*

Erleben von Beanspruchung und Überforderung („Stress“) durch Anforderungen, schwierige Aufgaben und Verpflichtungen, welche Zeitdruck schaffen und zu psychischen und körperlichen Folgen (Stressreaktionen) führen; unzureichende Bewältigung von Anforderungen; Überreizung und Erschöpfung; Mangel an Entspannung und Erholung. Dieses Bild der subjektiven Überforderung und Klagsamkeit kennzeichnet wohl das Lebensgefühl vieler Menschen und ist u.a. für psychosomatische Fragestellungen wesentlich (z.B. Myrtek 1980, Siegrist et al., 1980, Fahrenberg, Myrtek & Trichtinger 1985, Myrtek 1985, 1998a, Kreutel 1989).

### *Gesundheitssorgen*

Sorgen über mögliche Ansteckung und Erkrankung; Tendenz zur hypochondrischen Beachtung möglicher Krankheitszeichen; Zweifel an ärztlicher Kompetenz; Streben nach „gesunder“ Lebensweise, Schonhaltung. Dieses Konstrukt interessiert z.B. in Hinblick auf Patienten mit psychovegetativen Syndromen, Gesundheitserziehung im Rehabilitationsbereich und Attributionsprozesse bei Körperwahrnehmungen (siehe Medert-Dornscheidt 1975, Fahrenberg, Walschburger, Foerster, Myrtek & Müller 1979, Myrtek 1987, Kreutel 1989). Für den Bereich der Körperwahrnehmungen und körperlichen Beschwerden wurde inzwischen ein Fragebogen mit 10 Skalen entwickelt, die Freiburger Beschwerdenliste FBL-R (Fahrenberg, 1994a).

Zu allen Konstruktbereichen wurden aus der Literatur und aus den vorausgegangenen eigenen Untersuchungen und Fragebogenentwürfen vielfältige Anregungen aufgenommen. Bei einer genaueren Systematisierung der Konstruktbereiche zeigte sich jedoch, dass mehrere Komponenten oder Facetten, welche hier für psychologisch wesentlich gehalten werden, unterrepräsentiert waren oder fehlten. Deshalb wurde eine große Anzahl von Items neu formuliert und auch mit Kollegen ausgiebig erörtert. Die Listen von durchschnittlich ca. 50 Items pro Konstruktbereich wurden anschließend in gründlicher Diskussion der gemeinten Konstrukte auf Skalen-Vorformen mit 17 bis 28 Items reduziert. Eine empirische Vorselektion mit item- und faktorenanalytischen Kontrollen geschah in einem ersten Prätest bei einer Mehr-Themen-Befragung durch das Institut für Demoskopie Allensbach (IfD) an einer Stichprobe von N=141 (65 Männer, 74 Frauen, 2 ohne Angabe), ergänzt durch N=77 Bamberger Studenten. In einem zweiten Prätest konnten speziell die z.T. noch einmal revidierten Items für die geplanten neuen Skalen an einer weiteren IfD-Stichprobe von N=113 (57 Männer, 55 Frauen, 1 ohne Angabe) erneut geprüft und schließlich in einigen Itemformulierungen noch einmal revidiert werden.

### **2.3.2 Aus FPI-G übernommene Konstrukte**

Die mit dem bisherigen FPI erfassten Persönlichkeitsdimensionen FPI-1 Nervosität, FPI-2 Spontane Aggressivität, FPI-3 Depressivität, FPI-4 Erregbarkeit, FPI-5 Geselligkeit, FPI-6 Gelassenheit, FPI-7 Reaktive Aggressivität/Dominanzstreben, FPI-8 Gehemmtheit und FPI-9 Offenheit sind im Itempool durch die vollständige Form FPI-A repräsentiert. Insgesamt 31 Items unterscheiden sich in ihrer Formulierung von den Items in der 1. bis 3. Auflage des FPI: kleinere sprachliche Veränderungen und Kürzungen (23 Items), psychologische Abschwächung des Iteminhalts (7 Items), und Umformulierung eines Items zur Vermeidung doppelter Verneinung. Aus der Form FPI-G wurden die folgenden 35 Items, welche sich in der erwähnten 5-Faktoren-Lösung ausgezeichnet hatten, aufgenommen: Nr. 7, 8, 13, 15, 17, 23, 33, 34, 35, 40, 41, 65, 71, 76, 81, 82, 84, 88, 93, 102, 105, 115, 128, 131, 133, 149, 151, 166, 173, 184, 185, 192, 196, 201, 203. Da auch die revidierte Form des FPI eine Skala Offenheit enthalten soll, wurden fünf zusätzliche Items des Bereichs Offenheit/Soziale Erwünschtheit formuliert.

Auch die Weiterführung der FPI-Zusatzskalen E und N zur Erfassung der Persönlichkeitsdimensionen Extraversion-Introversion und Emotionalität im Sinne Eysencks wird für zweckmäßig gehalten. Bei der Revision soll (1) die neuere Diskussion über die Komponenten von E und N (Eysenck 1977, Eysenck & Eysenck 1975, 1977, Guilford 1977) berücksichtigt und (2) weiterhin auf Orthogonalität von E und N geachtet werden. Da es sich um allgemeinere Konstrukte (Superfaktoren bzw. Sekundärfaktoren laut Eysenck) handelt, werden Items verschiedener FPI-Standardskalen in sie eingehen, aber auch zusätzliche Items, um wichtige Komponenten zu berücksichtigen und um die rechnerische Abhängigkeit von den FPI-Standardskalen zu verringern. Außer den E- und N-Items der Form FPI-A wurden noch die Items Nr. 57, 142, 152, 157, 158, 172 aus FPI-G in den Pool aufgenommen; auch einige der oben genannten Items der 5-Faktoren-Lösung kommen für eine E- oder N-Skala infrage.

Die Zusatzskala FPI-M (typisch männliche bzw. typisch weibliche Selbstschilderung) hat selbst unter Diplom-Psychologen manche Missverständnisse und wohl auch Missbrauch hervorgerufen, zumal wenn die Absicht und das Konstruktionsprinzip nicht erkannt, also nicht die Möglichkeit zur empirischen Erfassung, sondern nur die Unerwünschtheit dieser Geschlechter-Stereotype gesehen wurden. Solche Stereotype existieren zweifellos und sie haben weiterhin Einfluss auf Erziehung und Persönlichkeitsentwicklung. Für Psychologen/innen kann es durchaus interessant sein, zu erfahren, ob die Selbstschilderungen eines Individuums dem betreffenden Stereotyp entsprechen oder in welchem Ausmaß sie davon abweichen. Darüber hinaus haben die Testautoren jedoch von groben Fehlinterpretationen solcher Diskrepanzen – bis zur unverantwortlichen Diagnose „Homosexualität“ – erfahren, so dass – solchem Missbrauch vorbeugend – die Skala Maskulinität-Feminität im FPI-R nicht mehr enthalten sein wird. Im Abschnitt 2.5 werden dafür jene Items gekennzeichnet, deren Schwierigkeitsindizes die größten Effekte der drei Statusmerkmale Geschlecht, Alter und Schulbildung aufweisen.

Der Itempool enthielt außer den 114 Items der Form FPI-A noch 41 Items der Form FPI-G (also aus der Form FPI-B) und 85 neue Items – insgesamt 240. Der für die Repräsentativerhebung verwendete Fragebogen bestand aus diesem Itempool, sechs Items zur Beurteilung des Fragebogens aus der Sicht des Probanden sowie 25 Items zu sozialen Einstellungen unter der Überschrift „Ihre Meinung zu aktuellen Themen“. Den Abschluss bildeten mehrere Fragen zur beruflichen Belastung, Zufriedenheit, Gesundheitszustand, Rauchen, Alkoholkonsum, Körperlänge und Gewicht, Krankenhausaufenthalte, Kuren, Arztbesuche, chronische Krankheiten, Psychotherapie, Tablettenkonsum (wiedergegeben im Anhang zu Hampel & Fahrenberg 1983). Die neue Form FPI-R wird aus einer einleitenden Frage und 137 der 240 Items gebildet.

## 2.4 Repräsentativerhebung

In Zusammenarbeit mit dem Institut für Demoskopie Allensbach, dem auch an dieser Stelle für die gute Kooperation gedankt sei, wurden zunächst die zwei Prätests durchgeführt, um den Itempool zu verbessern und um die Durchführung und die erreichbare Rücklaufquote zu erkunden. Die Repräsentativerhebung wurde im Oktober und Dezember 1982 in Verbindung mit der üblichen Mehr-Themen-Umfrage vorgenommen. Am Schluss des Interviews wurde der Fragebogen übergeben. Er trug dieselbe Nummer wie das Interview, so dass beides zwar anonym, aber später verknüpfbar blieb. Grundsätzlich sollte das FPI in Anwesenheit des Interviewers ausgefüllt werden, weil dies der üblichen FPI-Testbedingung eher entspricht. Nur für den Fall, dass dies abgelehnt wurde, war der Fragebogen vom Interviewer später abzuholen oder bei Zeitmangel an das IfD Allensbach zu schicken. Der Interviewbogen enthält einen Vermerk, ob das FPI gleich oder später bearbeitet wurde. Der Interviewer ist für jeden Fragebogen, der in seinem Beisein ausgefüllt oder später von ihm abgeholt wurde, zusätzlich honoriert worden.

Die Erhebung geschah nach dem Quoten-Auswahlverfahren, bezogen auf die vom IfD auf 46,8 Millionen Einwohner (21,6 Millionen Männer und 25,2 Millionen Frauen) geschätzte deutsche Bevölkerung ab 16 Jahre in der Bundesrepublik und West-Berlin. Jeder der ca. 510 Interviewer erhielt eine sogenannte Quotenanweisung, die ihm vorschrieb, wie viel Personen er zu befragen hatte und nach welchen Merkmalen er sie auswählen sollte. Die Befragungsaufträge oder Quoten wurden nach Maßgabe der amtlichen statistischen Unterlagen auf 11 Länder und 26 Regierungsbezirke aufgeteilt und innerhalb dieser regionalen Einheiten auf 6 Gemeindegrößenkategorien, gegliedert nach Geschlecht, 7 Altersgruppen, nach Berufstätigen und Nichtberufstätigen in 5 Berufsgruppen.

Insgesamt wurden 2608 Interviewte um zusätzliches Ausfüllen des FPI gebeten. Von den tatsächlich vorliegenden 2181 (84%) Fragebogen wurden 146 ungeeignete ausgesondert, so dass schließlich  $N=2035$  (78%) für die Auswertung übrigblieben. Gründe für die Aussonderung waren vor allem fehlende Angaben, d.h. mehr als 10 missing data (4% von 240 Items), nicht eindeutig identifizierbare Fragebogensnummern oder die Verneinung des Items 1, welches nach der Bereitschaft zur offenen Beantwortung des FPI fragt. Nicht eliminiert wurden dagegen 68 FPIs, in denen Item 1 unbeantwortet geblieben war, denn eine Kontrolle ergab, dass sich diese Gruppe im Mittelwert der Skala FPI-9 nicht von einer nach Geschlecht, Alter, Schulabschluss aus dem übrigen Material parallelisierten Gruppe unterschied ( $t=0.46$ ,  $p=0.65$ ). Für 2024 Personen der beschriebenen Norm-Stichprobe  $N=2035$  liegen vollständige Statusmerkmale des IfD-Interviews vor.

## 2.5 Skalenkonstruktion

Erwartet wurden etwa 9 bis 13 substanzielle Faktoren, da für den bisherigen FPI-Itempool – wie erwähnt – fünf Faktoren hinreichend schienen und fünf zusätzliche Konstruktbereiche als Skalen-Vorformen aufgenommen waren. Die Hauptkomponentenanalyse (Programm MAXVAR) des Datensatzes N=2035 mit 239 Items ergab den folgenden Eigenwert-Verlauf: 20.41, 13.75, 8.42, 6.03, 4.71, 4.15, 3.49, 3.09, 2.89, 2.62, 2.48, 2.13, 2.06, 1.94, 1.69 usw. Dies gibt einen schwachen Hinweis auf eine 11-Faktoren-Lösung, doch fiel nach eingehenden Interpretationsversuchen von einigen sukzessiven Lösungen die Entscheidung für die 10-Faktoren-Lösung, welche die angestrebten Konstruktbereiche in psychologisch leichter interpretierbarer Weise strukturiert (vollständige Matrix siehe Hampel & Fahrenberg 1983).

In einem langwierigen Interpretationsprozess mit häufig wiederholten Faktorenanalysen (Hauptkomponentenanalyse, d.h. Diagonale 1, SPSS) und Itemanalysen (SPSS) wurde eine Optimierung der psychologischen Konstruktexplikation und der formalen Anforderungen versucht. Die im Kapitel 2.2 beschriebene Strategie, deren vielfältige Schritte, Revisionen und Kompromisse hier nicht im Detail beschrieben werden können, führte schließlich zu 10 Standardskalen des revidierten FPI mit jeweils 12 Items:

- FPI-R 1 Lebenszufriedenheit
- FPI-R 2 Soziale Orientierung
- FPI-R 3 Leistungsorientierung
- FPI-R 4 Gehemmtheit
- FPI-R 5 Erregbarkeit
- FPI-R 6 Aggressivität
- FPI-R 7 Beanspruchung
- FPI-R 8 Körperliche Beschwerden
- FPI-R 9 Gesundheitssorgen
- FPI-R 10 Offenheit

Diese Skalen sind unter denselben Vorbehalten wie im bisherigen FPI mit Eigenschaftsnamen etikettiert worden, da solche Kurzbezeichnungen wohl in der Praxis unvermeidlich sind. Der Benutzer des FPI muss jedoch unbedingt berücksichtigen, dass sich jede Skala auf einen breiten Konstruktbereich bezieht und in der Regel mehrere psychologische Facetten erfasst. Deshalb ist es unerlässlich, dass jede inhaltliche Interpretation von FPI-R-Ergebnissen auf der Grundlage der ausführlichen psychologischen Beschreibung der Skalen (Kapitel 4) erfolgt.

Ursprünglich waren Skalen mit 14 Items beabsichtigt, doch reichte der Itempool hierfür nicht aus. Bei einigen Skalen befinden sich auch unter den 12 Items noch ein oder zwei Items, deren Ladung und/oder Trennschärfe nicht ganz befriedigen. Wegen unerwünschter Pseudo-Homogenität oder wegen zu komplexer oder zu hoher Nebenladungen mussten mehrere, zunächst geeignet erscheinende Items aufgegeben werden. Für eine systematische Anwendung des Suppressorkonzepts mangelt es im Pool – anders als bei Cattells schiefwinklig rotierten Faktoren – an Items mit geeigneten Nebenladungen und Vorzeichen.

Die zusätzlichen Skalen FPI-R E (Extraversion) und FPI-R N (Emotionalität) wurden wie bei der ursprünglichen FPI-Konstruktion getrennt von den Standardskalen entwickelt. Sie umfassen 14 statt 12 Items; von den E-Items sind 6 und von den N-Items sind 5 Items in den 10 FPI-R-Standardskalen enthalten. Zugunsten orthogonaler Skalen mussten einige Items, die für bestimmte Facetten dieser Konstrukte interessant waren, ausgeschieden werden.

Die Skalen Extraversion und Emotionalität (Emotionale Labilität, Neurotizismus) entsprechen inhaltlich den von Eysenck entwickelten Skalen E und N, doch liegen im Vergleich zur neuen deutschen Bearbeitung des EPQ (Ruch, 1999) noch keine Äquivalenzkoeffizienten vor.

## 2.9 Offenheit oder Soziale Erwünschtheit der Antworten?

Antworttendenzen wirken sich zwar nicht allein in Persönlichkeitsfragebogen aus, gelten jedoch als das vielleicht wichtigste Methodenproblem des Fragebogenverfahrens. Die Diskussion konzentriert sich hauptsächlich auf zwei mögliche Antworttendenzen: auf die Offenheit bzw. soziale Erwünschtheit (SE) der Antworten und auf die Bejahungstendenz (Akquieszenz).

Da die meisten Persönlichkeitsinventare einen asymmetrischen Auswertungsschlüssel haben, also die Mehrzahl der Items in der „stimmt“-Richtung für die betreffende Skala gezählt wird, ist es denkbar, dass sich eine, manchen Personen unterstellte, allgemeine Bejahungstendenz auswirken kann. Zur Kontrolle dieser Antworttendenz wird seit langem die Balancierung der Schlüsselrichtung der Items empfohlen, doch ist es außerordentlich schwierig, dies praktisch zu erreichen, denn die „Spiegelung“ eines Items kann zu unerwünschten Nuancen führen. Auch bei der Entwicklung des FPI-R-Itempools wurde darauf geachtet, Formulierungsvarianten mit gegensätzlicher Polung aufzunehmen. Viele dieser negativ gepolten Items waren aber psychologisch und sprachlich unbefriedigend und haben die verschiedenen Stufen der Itemselektion nicht überstanden. So musste während der Konstruktion auf eine balancierte „stimmt“-„stimmt nicht“-Polung der Items verzichtet werden. Die mögliche Tendenz zur Bejahung solcher Fragebogen-Items und die entsprechende Nein-Sage-Tendenz sind operational nicht leicht zu fassen, und es gibt keinen einfachen Weg zur Kontrolle oder gar Korrektur ihrer möglichen Effekte in Persönlichkeitsinventaren (siehe auch Vagt & Wendt 1978, Amelang & Borkenau 1981, Mittenecker 1982). Voyce & Jackson (1977) versuchten, Erwünschtheit und Akquieszenz als Sensitivität für den zugrundeliegenden SE-Wert eines Items bzw. als Antworten-Schwelle zu interpretieren, doch bleibt die empirische Schätzung dieser Parameter unter Nicht-Ernst-Bedingungen eine fragwürdige Angelegenheit, zumal die pauschale Einstufung nach SE-Werten ohne Rücksicht auf Bezugsgruppen, jeweiligen Interaktionspartner und Kontext kaum gültig sein kann (Dickenberger, Holtz & Gniech 1978).

Die Diskussion über die Tendenz, sozial erwünschte Antworten zu geben oder Antworten sogar bewusst zu verfälschen, ist noch umfangreicher und bewegt sich zwischen sehr kontroversen Äußerungen, welche ausdrücklich vor einer Unterschätzung (z.B. Jackson) oder einer Überschätzung der Effekte (z.B. Rorer) warnen (Literatur siehe Amelang & Bartussek 1997; Borkenau & Amelang 1986; Krahe 1989; Mittenecker 1982; Mummendey 1981, 1999). Auch zum FPI wurden mehrere empirische Studien unter den Gesichtspunkten sozialer Erwünschtheit der Antworten, aber auch anonymer gegenüber nicht anonymer Beantwortung sowie „test-wiseness“ unternommen. Die in der 3. Auflage referierten Befunde zeigen, dass sich durch Zusatzinstruktionen, z.B. einer vorgestellten Bewerbungs- und Auslese-Situation im Vergleich zu einer anonymen Forschungssituation, deutliche Profil- und Struktur-Unterschiede erreichen lassen. Diese Effekte betreffen vor allem den Bereich Aggressivität, Erregbarkeit und Emotionale Labilität. Die aufgrund solcher Analogstudien erwarteten Einflüsse auf Mittelwerte und Struktur der Skalen haben sich in den wenigen Untersuchungen realer Auslese-Situationen allerdings nur teilweise und nicht so eindeutig belegen lassen wie erwartet wurde (siehe Fahrenberg et al., 1978). Die adäquate Versuchsplanung ist auf diesem Gebiet sehr schwierig. Zunächst müsste eine reale Bewerbungs- und Auslesesituation herbeigeführt und dabei durch zusätzliche Erhebungen gesichert sein, dass diese Situation auch subjektiv als relevant angesehen wird. Dann müssten diese Skalenwerte eigentlich in einem intraindividuellen Vergleich an Skalenwerten aus einer weniger involvierenden, anonymen oder nicht anonymen Forschungssituation kontrolliert werden, wobei noch zwischen allgemein (stereotyp) und individuell als erwünscht angesehenen Verhaltensweisen zu unterscheiden wäre.

Dagegen wird man aus den Analogstudien, welche sich meist auf Studenten und auf recht artifizielle Erhebungssituationen stützen, kaum eine direkte Schlussfolgerung für die diagnostische Praxis ziehen können. Dies gilt auch für die Untersuchungen nach der sog. bogus-pipeline-Technik, welche die Studenten angeblich glauben lassen soll, dass der Versuchsleiter durch einen Stimmanalysator oder andere Techniken die Wahrheit der Antworten erkennen werde (Millham & Kellogg, 1980; Mummendey & Bolten, 1981). Dennoch tragen diese Arbeiten dazu bei, mehrere psychologische Facetten zu unterscheiden: Fremdtäuschung aus dem Bedürfnis nach sozialer Anerkennung (other deceptive bzw. impression management) im Unterschied zu Selbst-Täuschung (defensive denial) als hauptsächliche Komponenten (Millham & Kellogg, 1980), außerdem Selbsttäuschung und bewusste Verfälschungsabsicht (Gur & Sackeim, 1979), Selbst-Bewusstsein (Fenigstein, Scheier & Buss, 1975; Heinemann, 1979), positive oder negative Selbstdarstellung mit sicher-bestimmtem oder vorsichtig-defensivem Antwortstil (Schmolck, 1981). Überzeugende Operationalisierungen für

diese begrifflichen Unterscheidungen fehlen noch, so dass typische Lügen- bzw. Soziale Erwünschtheits-Skalen weiterhin als mehrdeutig anzusehen sind (siehe auch Eysenck & Eysenck 1975). Meier (1985) beobachtete differenzielle Effekte, die es nahelegen, das Antwortverhalten nicht so sehr unter dem Gesichtspunkt der Verfälschbarkeit, sondern eher im Hinblick auf die Rezipientenmerkmale zu analysieren. In seiner Studie beantworten Gruppen von Studenten das FPI-A am Videoterminal in Anwesenheit eines gleichgeschlechtlichen oder eines gegengeschlechtlichen Versuchsleiters. Bei 31% der Items wurden bedeutsame Kontraste überwiegend zwischen den beiden Versuchsleitergruppen festgestellt, und anschließende Diskriminanzanalysen anhand von ausgewählten 12 Items ergaben Hinweise, dass das Antwortverhalten im Sinne der stereotyp männlichen/ weiblichen Selbstbeschreibung und Selbstunsicherheit/Selbstsicherheit auch vom Geschlecht des Versuchsleiters abhängt. Die Effektstärke wurde jedoch nicht analysiert.

Zur Aufklärung dieser Antworttendenzen kaum geeignet sind einfache Korrelationsstudien, in denen Ergebnisse von Persönlichkeitsinventaren mit einer speziellen Skala für Soziale Erwünschtheit korreliert werden. Diese sog. SE- oder SD-Skalen enthalten meist eine Anzahl von Items, welche inhaltlich den Bereichen Emotionale Labilität, Ängstlichkeit, Erregbarkeit und Aggressivität zuzuordnen sind. Entsprechende Beziehungen zu jenen Persönlichkeitsskalen und geringere zu Extraversion, Geselligkeit usw. sind deshalb trivial. Weiterhin haben die Bemühungen zur Parallelisierung des SE-Gehalts von Items nicht überzeugen können, da solche Normierungen an den vielfältigen Randbedingungen scheitern und für bestimmte Bereiche wie Aggressivität und Emotionale Labilität wohl aussichtslos sind.

Als sehr verbreitete Auffassung gilt nach wie vor, dass Persönlichkeitsfragebogen zum Zweck der Bewerberauslese wegen des Risikos der Verfälschung ungeeignet sind. Dies gilt ebenso für alle anderen Testsituationen, in denen der Befragte motiviert ist, durch Manipulation seiner Antworten einen bestimmten Zweck, ein günstiges oder ihm in spezieller Hinsicht wichtiges Bild von seiner Person zu liefern. Dies ist eine psychologisch plausible Auffassung, obwohl die empirischen Belege und Analogstudien nicht so eindeutig und pauschal für massive Effekte sprechen (siehe Hossiep, Paschen & Mühlhaus, 2000).

Da eine Kontrolle oder gar eine auswertungstechnisch-rechnerische Kompensation solcher Effekte undurchführbar ist, bleibt nur der Rat, Persönlichkeitsinventare erstens nur dann zu verwenden, wenn das Risiko einer Verfälschungsmotivation gering ist, und zweitens die Testwerte eines Probanden nicht zu berücksichtigen, falls das Testprofil einen relativ niedrigen Offenheits-Wert hat.

Amelang & Borkenau (1981) haben aber betont, dass sich die Empfehlung, bestimmte Tests als uninterpretierbar anzusehen, auf empirische Nachweise stützen müsste, inwieweit SE-Skalen tatsächlich die Validität der anderen Skalen moderieren. Die Befunde hierzu sind jedoch bisher nicht sehr eindrucksvoll (siehe auch Buse 1976, 1980). Aber der von Fiske & Pearson (1970) als chaotisch bezeichnete Zustand der Forschung über Soziale Erwünschtheit und Akquieszenz wird sich nur dann bessern, wenn statt stilistischer Analysen der von College-Studenten in Analogstudien gegebenen Antworten die realen Assessmentprobleme in Kriteriensituationen untersucht werden (siehe Böhme 1997, Janssen & Reitzner 1980, Hampel & Klinkhammer 1978).

Kury (1983) konnte vier randomisierte Gruppen jugendlicher männlicher Straftäter vergleichen, die das FPI-G in der Strafanstalt unter vier verschiedenen Instruktionen ausfüllten: anonym für Forschungszwecke, nicht anonym für Forschungszwecke, als Testunterlage für die Vollzugsakten, als Testunterlage für die Vollzugsakten mit zusätzlicher Antiverfälschungsinstruktion. Varianzanalytisch ergaben sich signifikante Mittelwertunterschiede in den Skalen FPI 1, FPI 2 und FPI 4, während FPI 9 die Signifikanzschwelle knapp verfehlte. Nach Zusammenfassung von Gruppe 1 und 2 sowie 3 und 4 zeigten sich signifikante Effekte in den Skalen 1, 2, 4, 7 und 9. In der angeblichen Ernstsituation zeichneten sich die Selbstbeurteilungen durch relativ geringere Ausprägung von Nervosität, Aggressivität und Erregbarkeit aus. Der Mittelwert der Offenheitsskala ist zwar – wie erwartet – niedriger, wenn die Befunde in die Vollzugsakte kommen sollen, doch liegt er noch im unauffälligen Bereich der Normwerte. Kury weist darauf hin, dass der Diagnostiker in diesem Experiment beim größten Teil der Fälle anhand der Testwerte der Skala FPI 9 keine Möglichkeit hätte, die Testinterpretation infrage zu stellen.

Trotz der fortbestehenden begrifflichen Unklarheiten und der Unsicherheit der praktischen Handhabung wurde auch für

FPI-R eine Offenheitsskala beibehalten. Die Skala FPI-R 10 Offenheit hat 12 Items im mittleren Schwierigkeitsbereich und weist mit 0.74 eine noch akzeptable Konsistenz auf. Die höchsten Beziehungen (Tab. 2.9) hat FPI-R 10 mit FPI-R 6 Aggressivität (gemeinsame Varianz 20%), FPI-R N Emotionalität (18%), FPI-R 5 Erregbarkeit (14%) und FPI-R 1 Lebenszufriedenheit (10%). Dies sind erhebliche Varianzanteile, welche zu genaueren Analysen herausfordern. Keine substanziellen Beziehungen bestehen zu FPI-R 2, 3, 4, 8 und 9, obwohl diese Skalen ebenfalls einige Items enthalten, welche – pauschal betrachtet – ein als mehr oder minder positiv geltendes Verhalten betreffen.

Der zur Repräsentativerhebung verwendete Fragebogen enthielt außerdem sechs Fragen, welche sich als „Meta-Fragen“ auf den Test selbst beziehen: die individuell benötigte Zeit zum Ausfüllen, Fragen nach sicherer Beantwortbarkeit, Verständlichkeit, individuellem Zutreffen, Zudringlichkeit der Fragen sowie nach der Eignung des Fragebogens, ein besseres Verständnis von Menschen zu erreichen (siehe nächster Abschnitt). Zwischen individuellen Antworten zu diesen Meta-Fragen und den FPI-Skalen bestehen nur geringfügige Beziehungen (höchstens 2% gemeinsame Varianz) und die Korrelationskoeffizienten gerade mit der Skala Offenheit sind besonders gering  $r < 0.10$ . Auch für einzelne Items der Skala 10 ergeben sich keine höheren Zusammenhänge. Dagegen sind einige Beziehungen zu Statusmerkmalen festzustellen: FPI-R 10 Offenheit korreliert  $-0.15$  mit Geschlecht,  $-0.31$  mit Alter und  $.01$  mit Schulabschluss, d.h. jüngere Männer haben relativ hohe und ältere Frauen relativ niedrige Werte in dieser Skala, während der Schulabschluss zu vernachlässigen ist. Dieser Sachverhalt konnte erst durch diese Repräsentativerhebung dargestellt werden. Eysenck & Eysenck (1975) berichten dagegen eine niedrige, aber signifikante, negative Beziehung zwischen EPQ-L-Skala und Intelligenztests wie den Progressive Matrices. Die Beziehungen zu den übrigen IfD-Variablen sind gering und außerdem z.T. vermutlich Drittsteuerungen: vier Koeffizienten  $\geq 0.10$  sprechen dafür, dass höhere Erwartungswerte für Offenheit anzunehmen sind, wenn die Probanden berufstätig sind, mit mehreren Personen im Haushalt leben, verheiratet sind, Gewerkschaftsmitglied und nicht CDU-orientiert sind. Darüber hinaus ist den Verteilungen zu entnehmen, dass in Ausbildung befindliche bzw. Berufstätige relativ höhere Werte haben als Hausfrauen oder Rentner und dass Angestellte und Facharbeiter relativ höhere Werte haben als Selbstständige und leitende Angestellte. Während sich im Merkmal Konfession nur ein verhältnismäßig geringer Zusammenhang andeutet, d.h. relativ niedrige Werte bei evangelischer Konfession und relativ hohe Werte bei „anderer“ Konfession, differenziert die Angabe über Kirchenbesuch deutlicher: mit Offenheit hängt die Angabe „selten oder nie“ bzw. mit geringer Offenheit die Angabe „häufig“ zusammen. Wegen der Verknüpfung mit Alter und Geschlecht sind diese Statistiken jedoch sehr zurückhaltend zu interpretieren.

In einem weiteren Schritt wurden die Beziehungen zwischen den Meta-Fragen und der Skala Offenheit untersucht. Als Prädiktoren der Offenheit wurden außer diesen fünf Fragen auch die IfD-Angabe, ob ein Befragter die Auskunft über Mitgliedschaft in der Gewerkschaft oder über seine Parteipräferenz verweigerte, sowie das Geschlecht und das Alter verwendet. Die multiple Regression zeigt, dass Geschlecht und Altersunterschiede 23% der Varianz aufklären, die registrierte Verweigerung jener Auskünfte inkrementell noch 2% und die Meta-Items zusammen nur 1% der Varianz.

Aus den Korrelationsrechnungen mit den übrigen psychologischen Variablen des bei der Repräsentativerhebung verwendeten Fragebogens lassen sich keine weiterführenden Hinweise entnehmen. Bemerkenswert ist nur, dass zu der für viele nicht unbedingt sozial erwünschten bzw. eher privat-vertraulichen Angabe, in einer Psychotherapie zu sein oder gewesen zu sein, und ebenso zu Angaben über Tablettenkonsum praktisch Null-Korrelationen bestehen. Mehr Korrelationen gibt es mit den weltanschaulich-politischen Einstellungen, die als „Aktuelle Themen“ der Repräsentativerhebung formuliert waren, denn 9 dieser 25 Items weisen Beziehungen  $r \geq 0.10$  mit FPI-R 10 auf. Der autoritär-konservative Pol ist mit geringerer Offenheit assoziiert, doch sind erneut alters- und geschlechtsbedingte Effekte zu bedenken.

Schließlich bleibt als interne Analysemöglichkeit noch die Differenzierung nach Teilstichproben, welche sich in der Ausprägung der Skala Offenheit unterscheiden, mit einem Vergleich der Mittelwerte, Standardabweichungen und Konsistenzkoeffizienten aller anderen FPI-R-Skalen. In der Tabelle 2.12 sind die Ergebnisse zusammengestellt. Sie bestätigen das auch in den Korrelationskoeffizienten (Tabelle 2.9) erscheinende Bild deutlicher Effekte, vor allem in den Mittelwerten der Skalen 6, 5, E, 7, 8, oft auch mit Varianzunterschieden und Auswirkungen auf den Konsistenzkoeffizienten der Skala.

Sieht man von der Skala FPI-R 7 ab, so gilt für sechs Skalen, dass in der Gruppe mit geringen Offenheitswerten sowohl

die Varianz als auch die Konsistenz der Skalen systematisch geringer sind als in den Kontrastgruppen mit höheren Offenheitswerten. Diese Varianzrestriktion könnte als ein Hinweis auf die relative Konformität des Antwortenverhaltens dieser Teilstichprobe gesehen werden. Schließlich wurden getrennt für die drei Gruppen die Korrelationen zwischen FPI-R 10 und FPI-R N berechnet, da Eysenck & Eysenck (1975) hohe negative Korrelationskoeffizienten zwischen den Skalen N und L des EPQ als Ausdruck der SE-Motivation ansehen. Demnach müsste die Teilstichprobe mit niedrigen Offenheitswerten im FPI auch niedrigere Skalenwerte der Skala N aufweisen. Die Koeffizienten lauten für die drei in Tabelle 2.12 genannten Gruppen 0.23, 0.22 und 0.20 (nach Partialisierung von Geschlecht, Alter, Schulbildung 0.27, 0.26 und 0.21). Die genannte Hypothese lässt sich also bei bevölkerungsrepräsentativer Erhebung nicht stützen. Ein Effekt der Gruppierung nach Ausprägung von Offenheit deutet sich höchstens für die Skala E an:  $-0.06$ ,  $0.18$ ,  $-0.06$  (nach Partialisierung  $-0.08$ ,  $0.11$ ,  $-0.08$ ).

Zusammenfassend ist festzustellen, dass die Skala FPI-R 10 Offenheit formal ein verhältnismäßig konsistentes Persönlichkeitsmerkmal erfasst, dessen psychologische Interpretation drei Facetten betonen muss: Erstens die Bereitschaft, kleine Fehler und Alltagsschwächen offen einzuräumen und sich ungeniert und unkonventionell zu vielen herkömmlichen Normen zu äußern, wie es junge Männer eher tun als ältere Frauen und wie es sich auch in den erfragten gesellschaftlich-politischen Einstellungen äußert. Zweitens kann sich auch eine allgemeine Antworttendenz abzeichnen, bestimmte Verhaltensweisen zu behaupten, die man für sozial erwünscht hält, um absichtlich einen guten Eindruck auf andere zu machen. Die interne Analyse liefert für diese zweite Interpretation nur schwache und inkonsistente Hinweise. Drittens könnte sich eine Tendenz zur Selbstidealisation äußern. Die empirische Differenzierung und Prüfung dieser vermuteten Effekte ist eine sehr schwierige Aufgabe.

Offenheit als Persönlichkeitsmerkmal (im Sinne von unkonventionellem Verhalten) und zugleich geringe Offenheit als mögliche Verfälschungs-(Lügen-)Tendenz zu sehen, entspricht dem gegenwärtigen Wissensstand am besten. Nur die formale Verfälschungstendenz zu betonen, wäre dagegen eine grobe Überschätzung der empirischen Evidenz für diese Interpretation. Noch einmal ist aber zu betonen, dass weder die üblichen Analogstudien noch die Bedingungen einer Repräsentativerhebung zur Testnormierung wirklich adäquate Datensätze zur Klärung des SE-Problems einschließlich der hypothetischen Moderatorfunktion der SE-Werte (Buse, 1976, 1980) liefern können (siehe die aktuelle Diskussion in Abschnitt 7.7).

Für die Praxis empfehlen die Testautoren weiterhin, die Risiken bestimmter Testsituationen bzw. Testmotivationen zu vermeiden und Tests mit geringen Skalenwerten in Skala FPI-R 10 Offenheit (0 bis 2 Punkte bzw. Stanine 1 bis 3) vorsichtig und möglichst erst nach Einholen zusätzlicher Informationen zu interpretieren (siehe Abschnitt 5.5).

### *Soziale Erwünschtheit aus heutiger Sicht*

Die Diskussion über die Verfälschung von Persönlichkeitsfragebogen ist wiederaufgelebt, u.a. weil auch in Deutschland zunehmend Fragebogen für berufsbezogene Persönlichkeitsdiagnostik eingesetzt bzw. entwickelt wurden, u.a. 16 PF (Schneewind & Graf, 1998) und BIP (Hossiep & Paschen, 2000).

In der Fachliteratur zeichnet sich bei der Diskussion der empirischen Befunde eine pragmatische Haltung ab (Hossiep, Paschen & Mühlhaus, 2000; Mummendey, 1999; Paulhus, 1989; Sarges, 1995). Der Sachverhalt, dass die Antworten, wenn gewünscht bzw. durch eine Instruktion nahegelegt, erfolgreich nach sozialer Erwünschtheit angepasst werden können, wird weithin akzeptiert. Die praktische Relevanz kann jedoch bezweifelt werden, weil in der realen Situation u.U. die gegenläufigen Tendenzen wirksam sind, nicht zu übertreiben bzw. offen und echt zu wirken.

Zwar ist aufgrund solcher Befunde (u.a. Hampel & Klinkhammer, 1978; Hough et al., 1990; Ryan & Sacket, 1987) zunehmend eine pragmatische Tendenz entstanden (siehe Paulhus, 1989), doch darf das Problem der sozialen Erwünschtheit der Antworten nicht verharmlost werden. Im Einzelfall könnte die Anpassung der Antworten an eine Bewerbungssituation – vielleicht noch unterstützt durch die Lektüre eines populären „Testknackers“ – durchaus zu Fehlurteilen führen, wenn Persönlichkeitsfragebogen schematisch ausgewertet werden. Die bisherigen Methodenstudien sind kaum als echte Feldstudien mit einer realitätsnahen Analyse des Entscheidungsnutzens bzw. Schadens des betreffenden Assessments anzusehen.



Das Freiburger Persönlichkeitsinventar FPI-R weist in einigen Skalen deutliche Beziehungen zum Berufsstatus, zur Berufstätigkeit und zur beruflichen Belastung auf (siehe Kapitel 6.3). Das FPI-R ist jedoch nicht für diesen Bereich konstruiert worden, sondern wird vorzugsweise in Anwendungsbereichen eingesetzt, wo die grundsätzlich positive Testmotivation vorausgesetzt werden kann, u.a. im Bereich der Gesundheitspsychologie, klinischen Psychologie und Persönlichkeitsforschung. Insofern stellen sich die Probleme der sozialen Erwünschtheit von Antworten in geringerem Maße als in der Berufseignungsdiagnostik.

## **3 Neue Normierung**

### **3.1 Absichten der Untersuchung**

Die bevölkerungsrepräsentative Normierung des Freiburger Persönlichkeitsinventars konnte sich im Jahr 1982 nur auf die alten Bundesländer beziehen.

Im Sinne der notwendigen Qualitätskontrolle psychologischer Tests war deswegen eine neue Erhebung notwendig, welche die Normen aktualisieren und dabei auch die neuen Bundesländer umfassen sollte. Die Frage, inwieweit sich die Normen in diesem Zeitraum von 1982 bis 1999, d.h. über 17 Jahre geändert haben, kann hier leider nicht an einer Kohorte, sondern nur durch den Vergleich zweier Querschnitte untersucht werden. Welche Effektstärke haben die Veränderungen?

Aufgrund der Repräsentativerhebung von 1982 wurde das 1970 in erster Auflage publizierte FPI in einigen Skalen revidiert bzw. ergänzt (siehe 4. Auflage 1984 bis 6. Auflage 1994). In einer neuen Untersuchung können nicht nur testmethodische Eigenschaften wie die Reliabilitäten der Skalen, sondern auch die Struktur der Skalen durch neue Item- und Faktorenanalysen überprüft werden. Ist die Struktur der FPI-R Skalen reproduzierbar?

Die Repräsentativerhebung wurde wieder genutzt, um über das FPI hinaus soziodemografisch und psychologisch interessante Daten zu gewinnen. Von den Befragten liegen außer den bereits mit dem FPI-Testformular erhobenen Angaben zu Geschlecht, Alter, Schulabschluss, Familienstand und Beruf zahlreiche weitere soziodemografische Daten aus dem Interview vor, z.B. zum Wohnort, Anzahl der Kinder, Einkommen, Konfession, Präferenzen für Parteien.

Der Fragebogen selbst enthält nach dem Block der 138 Items des FPI-R weitere Items hinsichtlich politisch-gesellschaftlicher Einstellungen sowie Fragen zu den Bereichen berufliche Belastung, Lebenszufriedenheit, Risikofaktoren und Gesundheit. Wie in der vorausgegangenen Erhebung sollen die statistischen Analysen dieser Daten Validitätshinweise für bestimmte FPI-R Skalen liefern.

Anlässlich der Erhebung von 1982 wurden mehrere „Meta-Fragen“ zur Verständlichkeit der Items und zu anderen Aspekten gestellt (Kapitel 2.9), welche sich jedoch als nicht sehr aufschlussreich erwiesen. Ergiebiger war der Versuch, die Interviewer in den Validierungsprozess einzubeziehen, indem sie um eine Verhaltenseinstufung gebeten wurden, wie „selbstsicher“ (sehr selbstsicher – unsicher) der/die Befragte auf sie wirkte. Obwohl es sich wegen der fehlenden Standardisierung und wegen der sehr großen Zahl von Beurteilern um eine völlig unübliche Methode handelt, wurde dieser Ansatz jetzt ausgebaut.

### **3.2 Fragebogen und Interview**

Der ausgeteilte Fragebogen enthält:

- 138 Items des FPI-R (stimmt, stimmt nicht);
- 7 Fragen zu Geschlecht, Alter, Schulabschluss, Familienstand, Haushalt, Berufstätigkeit und Berufsgruppe;
- 32 Items zu politisch-gesellschaftlichen Einstellungen (stimmt, stimmt nicht);
- 29 Fragen zu beruflicher Belastung, Lebenszufriedenheit und Gesundheit.

Diese 29 zusätzlichen Fragen lauten (siehe auch Tabelle 6.9):

Üben Sie einen Beruf aus?

- Stufen Sie bitte den Grad Ihrer beruflichen Belastung ein: durch körperliche Arbeit; durch Tempo, Genauigkeit, Konzentration; durch Verantwortung für Menschen und Planung (3 fünfstufige Items: gering – sehr stark).
- Stufen Sie bitte den Grad Ihrer Zufriedenheit ein: mit Ihrer beruflichen Situation; mit Ihrer familiären Situation; mit Ihrer finanziellen Situation; mit Ihrer Wohnsituation (4 fünfstufige Items: sehr zufrieden – unzufrieden).
- Wie schätzen Sie Ihren Gesundheitszustand während der letzten Zeit ein? (fünfstufig: sehr gut – schlecht).
- Rauchen Sie? (ja, nicht mehr, noch nie). Falls ja: Wie viele Zigaretten rauchen Sie täglich?
- Körperlänge (cm), Körpergewicht (kg) und daraus abgeleitet der Body Mass Index (BMI in kg/m<sup>2</sup>) sowie die subjektive Einstufung: Würden Sie sich selbst als übergewichtig bezeichnen? (ja, vielleicht, nein).
- Haben Sie eine chronische Krankheit? (ja, nein).
- Um welche handelt es sich?
- Wie viele Krankenhausaufenthalte hatten Sie bisher (bei Frauen: ohne Entbindungen)?
- Wie oft waren Sie zur Kur?
- Wie oft waren Sie während der letzten 12 Monate beim Arzt (ohne Zahnarzt)?
- Sind oder waren Sie in Psychotherapie (seelischer Behandlung)?
- Haben Sie während der letzten 12 Monate irgendwelche der folgenden Mittel eingenommen? Schlaftabletten, Beruhigungstabletten, Schmerztabletten, Homöopathische Mittel (4 fünfstufige Items: nie, etwa 2-mal im Jahr, etwa 2-mal im Monat, etwa 3-mal pro Woche, fast täglich).

Das Interview, das nach einem Leitfaden mit Merkmalsblatt durchgeführt wurde, führte zu 31 soziodemografischen Variablen (siehe auch Tabelle 3.1), u.a.:

- Bundesland und Gruppierung nach Ost – West,
- Einwohnerzahl des Wohnortes, Stadt – Land,
- Berufstätigkeit, Stellung im Beruf, Einkommen,
- Familienstand bzw. Partnerschaft, Anzahl der Kinder u.a. Personen im Haushalt,
- Mitglied einer Kirche, Konfession, Häufigkeit der Kirchenbesuche,
- Mitgliedschaft in einer Gewerkschaft, Präferenz für eine Partei.

Die Befragten sollten das Interview nach zwei Aspekten beurteilen:

- „Fanden Sie dieses Interview eigentlich interessant oder nicht so interessant?“ (4 stufig: „Sehr interessant bis „gar nicht interessant“). „Wie fanden Sie die Länge?“ (4stufig: „gerade richtig“ bis „viel zu lang“). Der Interviewer notierte außerdem, ob das FPI gleich ausgefüllt (63.9%), später ausgefüllt und abgeholt (28.9%) oder zurückgeschickt wurde (2.1%, ohne Angabe 5.1%).

Die Interviewer sollten schließlich die Befragten hinsichtlich sozialer Schicht und Verhalten einstufen. Die Einstufung der sozialen Schicht wurde anhand eines Merkmalsblattes vorgenommen und berücksichtigte sowohl das Bildungsniveau als auch den Lebensstil und die finanziellen Verhältnisse (Beispiele waren vorgegeben).

Die Verhaltenseinstufungen betrafen drei Aspekte:

- „Der/die Befragte war während des Interviews – (1) sehr selbstsicher, von sich überzeugt, (2) ziemlich selbstsicher, (3) etwas unsicher, (4) ziemlich unsicher.“
- „Der/die Befragte wirkt – (1) sehr energisch, (2) energisch, (3) weniger energisch, (4) gar nicht energisch.“
- „Der/die Befragte ist – (1) sehr freundlich, (2) freundlich, (3) nicht so freundlich, (4) eher unfreundlich.“

### 3.3 Repräsentativerhebung

In der bewährten Zusammenarbeit mit dem Institut für Demoskopie, Allensbach, dem auch an dieser Stelle erneut zu danken ist, wurde die Erhebung im September bis November 1999 in zwei Wellen in Verbindung mit der üblichen Mehr-Themen-Umfrage vorgenommen. Am Schluss des Interviews wurde der Fragebogen übergeben. Er trug dieselbe Nummer wie das Interview, so dass beides zwar anonym, aber später verknüpfbar blieb. Die Erhebung entsprach dem zuvor geschilderten Ablauf im Jahr 1982 (Kapitel 2.4).

Die Erhebung erfolgte repräsentativ für die deutsche Bevölkerung ab 16 Jahren nach dem Quoten-Auswahlverfahren. Dabei werden den Interviewern „Quoten“ in die Hand gegeben, welche vorschreiben, wie viel Personen zu befragen sind und nach welchen Merkmalen diese auszuwählen sind. Für diese Studie wurden etwa 500 Befragungsaufträge oder Quoten für jeweils fünf Interviews auf Bundesländer und Regierungsbezirke nach 6 Ortsgrößenklassen verteilt. Innerhalb dieser regionalen Einheiten wurden dem Interviewer das Geschlecht in Kombination mit dem Alter sowie die Beteiligung am Erwerbsleben in Kombination mit dem Geschlecht und den verschiedenen Berufskreisen vorgegeben. Um auch die neuen Bundesländer in hinreichender Tiefe analysieren zu können, wurden diese in der Stichprobe deutlich stärker berücksichtigt, als es ihrem Anteil an der Gesamtbevölkerung entspricht. Falls bevölkerungsrepräsentative Aussagen für Gesamtdeutschland gewünscht werden, muss diese Disproportionalität über eine Gewichtung aufgehoben werden, bei welcher der Stichprobenanteil der neuen Bundesländer von 44% auf den tatsächlichen Bevölkerungsanteil von 20% heruntergewichtet wird. Von den insgesamt 4097 Personen, die an den beiden Mehr-Themen-Umfragen von September bis November 1999 teilgenommen haben, füllten 3805 oder 93% den ergänzenden schriftlichen Fragebogen aus.

Von diesen Fragebogen wurden 65 (=1.7%) eliminiert: 17 wegen Widersprüchen bei den Angaben zu Geschlecht und Alter zwischen den Fragebogen und dem IfD Datensatz sowie 48 wegen einer zu großen Anzahl fehlender Daten im FPI. In Übereinstimmung mit den im Manual publizierten Empfehlungen sollte ein Fragebogen mit mehr als 7 fehlenden Antworten (5% von 138 Items) nicht ausgewertet werden. Die Normierungsstichprobe umfasste also N=3740 Personen.

Bei der großen Mehrzahl der FPI-Items fehlen weniger als 20 von 3740 Antworten. Am häufigsten bleibt Item 131 „Meine Partnerbeziehung (Ehe) ist gut“ unbeantwortet (n=263). Diese Frage ist wahrscheinlich für viele Alleinlebende unbeantwortbar. Die in dieser Rangliste folgenden Items sind: „Ich bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden“ (Item 119; n=36); „Ich würde kaum zögern, auch alte und schwerbehinderte Menschen zu pflegen“ (Item 14; n=29); „Ich habe manchmal ein Gefühl erstickender Enge in der Brust“ (Item 133; n=28) und „Ich habe (hatte) einen Beruf, der mich voll befriedigt“ (Item 3; n=20). Der Mittelwert fehlender Antworten über die Personen beträgt 0.29 (von 138 Items) mit einem Range von 0 bis 7. Das Eingangs-Item „Ich habe die Anleitung gelesen und bin bereit, jeden Satz offen zu beantworten“ wurde von n=36 verneint. Diese Fragebogen wurden dennoch in die Analyse einbezogen, da hinsichtlich der wichtigsten soziodemografischen Merkmale und der 12 FPI-R Skalen keine signifikanten ( $p < .01$ ) Unterschiede festzustellen waren. Bei den politisch-gesellschaftlichen Items sowie bei Fragen zur beruflichen Belastung und zum Gesundheitsbereich wurde eine größere Anzahl fehlender Daten zugelassen; das valide N wird in den Tabellen angegeben. Bei den Fragen zu politischen/gesellschaftlichen Einstellungen fehlen zwischen n=22 und n=61 Antworten. An der Spitze liegen: „Ich bin für Arbeitszeitverkürzung, auch wenn ich dann weniger verdiene“ (n=61); „Auch die Bürgerkriege und der unaufhörliche Völkermord in vielen Staaten Afrikas sollten durch militärisches Eingreifen der UNO beendet werden“ (n=53); „Ich bin stolz, eine Deutsche/ein Deutscher zu sein“ (n=48). Der Mittelwert fehlender Antworten über die Personen beträgt 0.31 (von 32 Items) mit einem Range von 0 bis 32.

Die Tabelle 3.1 gibt eine Übersicht dieser Repräsentativerhebung von 1999 mit N=1743 Männern und N=1997 Frauen in den alten und neuen Bundesländern. Für die wichtigsten soziodemografischen Merkmale sind die Verteilungen angegeben.

Die soziodemografischen Merkmale der nur im Westen durchgeführten Erhebung von 1982 sind nur sehr eingeschränkt, z.B. hinsichtlich Altersverteilung und Familienstand, zum Vergleich geeignet. Dagegen kann eine andere Repräsentativerhebung, die ebenfalls vom IfD durchgeführt wurde, bessere Vergleichswerte liefern. Diese Erhebung diente der Revision und Normierung der Freiburger Beschwerdenliste FBL-R (Fahrenberg, 1994a). Diese Daten stammen aus dem Jahr 1993 und umfassen alte und neue Bundesländer. Eine weitere Vergleichsmöglichkeit bietet die 1994 durchgeführte Repräsentativerhebung zur Normierung des Fragebogens zur Lebenszufriedenheit (Fahrenberg, Myrtek, Schumacher & Brähler, 2000) mit N=2870 in den alten und neuen Bundesländern. Wegen der z.T. anderen Kategorisierungen der Merkmale wurden die Verteilungen jedoch nicht in die Tabelle 3.1 aufgenommen.

In den wichtigsten soziodemografischen Merkmalen stimmen die Erhebungen FPI-R 1999 und FBL-R 1993 weitgehend überein. Auf eine statistische Prüfung wurde hier verzichtet, da auch relativ kleine Unterschiede angesichts des großen

Stichprobenumfangs sehr signifikant werden (zu Qualitätskriterien der Umfrageforschung siehe auch Kaase, 1999; Sommer, Unholzer & Wiegand, 1999).

#### *Ost-West-Unterschiede*

Beim Vergleich der beiden Datensätze des FPI-R von 1982 und 1999 sind in mehreren soziodemografischen Merkmalen deutliche Unterschiede zu erkennen, die in geringem Maße vielleicht durch allgemeine Trends, aber hauptsächlich durch die Erweiterung auf die Bevölkerung der neuen Bundesländer zu erklären sind.

Die in Tabelle 3.2 dargestellten prozentualen Häufigkeiten und Chi-Quadrat-Tests zeigen keine Unterschiede in der Verteilung von Geschlecht, Altersgruppen, Familienstand, Haushalt und Kindern. Es gibt jedoch hochsignifikante Unterschiede in der Art des Schulabschlusses, in Berufstätigkeit und Berufsgruppe, Einwohnerzahl des Wohnorts und Land – Stadt, Anzahl der Personen im Haushalt, Einkommensgruppe, sozialer Schicht, Mitgliedschaft in Kirchen, Konfession, Häufigkeit des Kirchenbesuchs, Parteipräferenzen. Natürlich hängen mehrere dieser Variablen untereinander zusammen.

In demografischer Hinsicht handelt es sich offensichtlich um zwei heterogene Teil-Stichproben. Dieser Sachverhalt führt zu der Frage wie diese soziodemografischen Unterschiede bei den testmethodischen Analysen, bei der neuen Normierung und bei den Validierungsansätzen angemessen zu berücksichtigen sind. In psychologischer Hinsicht sind hier Geschlechtszugehörigkeit und Lebensalter (sowie Schulabschluss) in der Regel weitaus wichtiger als die anderen soziodemografischen Merkmale, denn sie klären – mit großem Abstand – mehr Varianz der FPI-Testwerte auf. Bevor über die in den Normen-Tabellen zu berücksichtigenden Merkmale entschieden wird, muss folglich die Effektstärke dieser Einflüsse geprüft werden. Ist es sinnvoll, solche Normen-Tabellen nach Geschlecht und Altersgruppen und zusätzlich nach Ost – West zu gliedern?

Eine vorläufige Abschätzung der Effekte ergab, dass die Ost – West Unterschiede in den FPI-Testwerten nur eine geringe Effektstärke aufweisen. Deswegen wurden die Itemanalysen und Faktorenanalysen an der gesamten Stichprobe durchgeführt. Bei den testmethodischen Analysen zur Konstruktion des FPI-R wurde früher auch nicht hinsichtlich Geschlecht und Alter differenziert, da ja ein dimensional einheitliches Beschreibungsinstrument geschaffen werden sollte. Die Ost-West-Unterschiede sollen zwar nicht vernachlässigt, aber in ihren testmethodischen Konsequenzen auch nicht überschätzt werden.

Unter dieser Perspektive wurde auch auf eine Reduktion der Stichprobe, um die proportionale Zusammensetzung nach den Quoten der Bundesländer zu erreichen, verzichtet. Die Untersuchung beabsichtigte nicht, im engeren Sinne „repräsentative“ Aussagen über die Gesamtbevölkerung der Bundesrepublik zu machen. Für die Differenzierung von Persönlichkeitsmerkmalen und für den Vergleich ihrer individuellen Ausprägungen wäre es ein unangemessener Verlust an Informationen, wenn die soziodemografischen Quoten starr eingehalten würden. Für die Validierung und Normierung sind die zusätzlichen Daten nur von Vorteil.

Auf der anderen Seite müssen die möglichen Effekte der disproportionalen Ost-West-Erhebung bei den Kriterienkorrelationen mit FPI-Testwerten beachtet werden. Wenn solche Korrelationen, z.B. mit beruflicher Belastung, Lebenszufriedenheit, Gesundheitsfragen, Verhaltenseinstufungen oder gesellschaftlich-politischen Einstellungen als Validitätshinweise herangezogen werden, sind u.U. solche Zusammenhänge wichtig. Dies ist jedoch ein allgemeines methodisches Problem solcher Analysen, denn die soziodemografischen Merkmale sind in erheblichem Maße konfundiert, z.B. mit Schulbildung, Berufstätigkeit, Einkommensgruppe und sozialer Schicht oder Geschlecht, Lebensalter und Gesundheitsaspekte. Durch Varianzanalyse bzw. Kovarianzanalysen und partielle Korrelationen sind solche Konfundierungen nicht hinreichend zu kontrollieren. Die Methode der Wahl ist hier die Bildung von statistischen Zwillingen (matched pairs), wobei neben anderen Merkmalen auch das Merkmal Ost-West zu berücksichtigen sein wird (Kapitel 6.3).

#### *Zusammensetzung der Normstichprobe*

Die Tabelle 3.3 zeigt die Zusammensetzung der Normstichprobe N=3740, gegliedert nach Geschlechtszugehörigkeit, Altersgruppen (sieben Dekaden), Schulabschluss (drei Gruppen) und Ost – West.

### 3.6 Stabilität der FPI-Normen 1982/1999

Die mittelfristige Veränderung der FPI-Normen wurde bereits durch den Vergleich der Erhebungen von 1965/69 und 1982 untersucht. Es hatten sich Veränderungen in sechs Skalen der Form FPI-A, jedoch mit nur geringer Effektstärke ergeben (Fahrenberg, Selg & Hampel, 1983; Fahrenberg et al., 1984). Die Behauptung, dass die Normen des FPI veraltet sind, war sogar in einem Nachrichtenmagazin (Der Spiegel, 1995) zu lesen.

Die Kontrolle der Normen bzw. eine neue Normierung sind zweifellos für die praktische Anwendung eines Tests wichtig. Darüber hinaus wären Trends in den Mittelwerten der FPI-Skalen ein interessantes Thema. Die Stabilität der FPI-Normen kann jetzt genauer analysiert werden, weil von 1982 und 1999 zwei für Westdeutschland repräsentative Stichproben des gesamten FPI-R vorliegen. Die Analysenstichprobe von 1965/69 war aus verschiedenen Erhebungen zusammengesetzt worden, so dass eine breite Varianz, aber keine Repräsentativität im engeren Sinne erreicht war.

In fünf FPI-R Skalen ergeben sich hoch signifikante ( $p < 0.001$ ) Veränderungen von 1982 zu 1999 (Tabelle 3.11). Im Einzelnen handelt es sich um eine Zunahme der Testwerte in den Skalen Lebenszufriedenheit, Aggressivität, Offenheit und um eine Abnahme in den Skalen Soziale Orientierung und Körperliche Beschwerden. Eine Zunahme des Testwertes Offenheit bedeutet psychologisch, dass alltägliche Schwächen und kleine Fehler eher eingeräumt werden. Dies betrifft, wie die bekannte Korrelation beider Skalen anzeigt, auch die Selbstbeurteilung des aggressiven Verhaltens. Deshalb ist der Anstieg der Testwerte in der Skala Aggressivität mehrdeutig.

Bevor diese Konstellation von erhöhter Lebenszufriedenheit und verbessertem körperlichen Befinden bei verringerter sozialer Orientierung und zunehmender Aggressivität als „säkulare“ Veränderung psychologisch interpretiert wird, ist daran zu erinnern, dass es sich zunächst nur um eine mittelfristige Veränderung der Selbstbeurteilung handelt. Über das manifeste Verhalten sind keine Aussagen möglich. Außerdem wurden zwei unabhängige Querschnitte und nicht eine Kohorte im Längsschnitt untersucht.

Der Mittelwertvergleich ist hier aufgrund der Kontrolle der Einflussgrößen Geschlecht, Alter, Schulabschluss und Einkommensgruppe durch das Verfahren der matched pairs (siehe auch Kapitel 6.3) als besonders zuverlässig anzusehen. Für diesen Vergleich konnten in den beiden Datensätzen  $N = 1651$  Personen von maximal möglichen  $N = 2035$  Paaren gefunden werden. Dieser Datensatz ist zwar nicht mehr bevölkerungsrepräsentativ, erlaubt jedoch gültigere Aussagen zur Fragestellung. Leider existiert in diesem Bereich noch keine Kohortenstudie mit wiederholten Erhebungen im Längsschnitt.

Die Tabelle 3.11 lässt erkennen, dass die Effektstärken dieser mittelfristigen Veränderungen gering sind, d.h. noch nicht einmal als kleine Effektstärken ( $d = 0.20$ ) bezeichnet werden können (Cohen, 1988). Bei einer Effektstärke von 0.10 überlappen sich die beiden Verteilungen weitgehend (zu 92.3%; bei  $d = 0.20$  zu 85.3%). Dieses Ergebnis führt zu der praktischen Frage, ob neue Normen überhaupt erforderlich sind. Für die Testpraxis hatten die Autoren die Stanine-Skala ausgewählt, d.h. relativ grobe Standardwerte, welche dem Range der Testwerte angemessen sind.

Die niedrigen Effektstärken der Veränderungen wirken sich nur geringfügig auf die Normen-Tabellen aus. Die Stichproben-Mittelwerte der FPI-R Testwerte von 1982 würden – wie die Testwerte von 1999 – in den Normen-Tabellen des Anhangs jeweils zu einem Staninewert=5 führen. Beim Nachschlagen des Stanine-Wertes für einen bestimmten Testwert kann es trotz der geringen Effekte, auch wegen der notwendigen Rundungen, zu Abweichungen um einen Stanine-Punkt nach oben oder unten kommen. Deshalb sind die alten und die neuen Normen-Tabellen nicht äquivalent.

Diese Aussagen betreffen nur die Veränderung der Normen für die Bevölkerung der alten Bundesländer. Die Normen-Tabellen des Anhangs gelten jedoch für die gesamte Bevölkerung. Ob sich in den neuen Bundesländern mittelfristige Trends in Testwerten zeigen werden, können erst künftige Studien ergeben.

## 6 Validitätshinweise

### 6.1 Faktorielle Validität und Kriterienvalidität

Die Skalen bzw. Testwerte des FPI-R repräsentieren eine markante psychologische Struktur in der durch 137 Items erfragten Selbstbeschreibung einer repräsentativen Auswahl von Personen der Durchschnittsbevölkerung. Diese Eigenschaftskonstruktionen enthalten jeweils mehrere Facetten und Komponenten, sind jedoch durch innere Konsistenz und Reproduzierbarkeit (relative zeitliche Stabilität) der Antworten gut geeignet, individuelle Differenzen zuverlässig zu erfassen. Aus der Formulierung vieler Items und durch die internen Analysen ergibt sich, dass die FPI-Testwerte relativ überdauernde Dispositionsprädikate indizieren, d.h. sich in der Selbstbeschreibung über breite Klassen von Merkmalen und Kontexten sowie über Zeitspannen von mehreren Jahren als relativ invariant erweisen.

Diese Schlussfolgerungen können aus der mehrstufigen Testentwicklung und Skalenkonstruktion sowie aus den multi-strategischen Reanalysen gezogen werden. Die psychologischen Konstruktionen beziehen sich auf einen aus theoretischen Interessen definierten Ausschnitt möglicher Eigenschaftskonzepte. Besonderheiten wichtiger Teilpopulationen der Durchschnittsbevölkerung und des klinischen Anwendungsfeldes sind vernachlässigt, um individuelle Differenzen der Selbstbeschreibung in einem standardisierten und bevölkerungsrepräsentativ normierten, allgemeinen Beschreibungssystem erfassen zu können. Die interne (konstruktbezogene, faktorielle) Gültigkeit des FPI-R kann – auch im Vergleich zu ähnlichen Persönlichkeitsfragebogen und Inventaren – als befriedigend angesehen werden.

Die praktische Anwendung des FPI und anderer Persönlichkeitsinventare wird regelmäßig von der Erwartung geleitet sein, dass über die interne Validität hinaus eine Kriterienvalidität besteht. Die Erfassung individueller Differenzen in Selbstbeschreibungen mit der Reduktion auf bestimmte Dimensionen und normierte Testprofile kann eine deskriptive Funktion haben, doch stellt sich in der Persönlichkeitsforschung und in der diagnostischen Praxis selbstverständlich die Frage nach den objektiven Entsprechungen dieser Testwerte, d.h. nach empirischen Korrelaten, nach Kriterienvorhersage und Entscheidungsnutzen.

Die in der 3. Auflage der FPI-Handanweisung noch kumulativ referierten Untersuchungen belegen, dass eine Vielzahl solcher Validitätsbeiträge, ein heute – angesichts von mehr als 1400 Publikationen – in vollständiger Form nicht mehr darstellbares Netz von empirischen Befunden existiert. In dieser Literatur dominieren jedoch bestimmte Untersuchungsansätze, während zu anderen Aspekten der Validierung wenig oder nichts vorliegt.

In zahlreichen Untersuchungen werden statistisch signifikante Unterschiede in den FPI-Testwerten bestimmter Personengruppen beschrieben, die aufgrund von klinisch-psychologischen, beruflichen und anderen psychosozialen Merkmalen gebildet wurden. Publikationen über die Vorhersage praktisch relevanter Kriterien, über die Evaluation individueller Begutachtungen oder Evaluation von Klassifikations- und Selektionsentscheidungen in bestimmten Anwendungsfeldern sind selten und anspruchsvolle Experimente zur Konstruktvalidität fehlen weitgehend. Meist handelt es sich um kleine oder sehr kleine Personengruppen, oft um eine Gelegenheitsauswahl, d.h. mit unklarer Präselektion, oder es fehlt eine äquivalente Kontrollgruppe. Falls wichtige Befunde mitgeteilt werden, fehlen in der Regel Replikationen.

Häufig anzutreffen sind Korrelationen zwischen FPI-Testwerten und anderen Persönlichkeitsvariablen. Berücksichtigt man die von Cattell vorgeschlagene Taxonomie von L-, Q- und T-Daten, so wird deutlich, dass die Mehrzahl der Publikationen sich nur auf L- und Q-Daten stützt und multimodale Assessmentstudien oder Multitrait-Multimethod-Analysen kaum zu finden sind. Wird außerdem noch zwischen Daten der objektiven Verhaltensmessung (Behavior Observation BO) und der von anderen Personen geleisteten Verhaltenseinstufung (Behavior Rating BR) getrennt, so gibt es kaum Studien mit BO-Daten und nur wenige mit BR-Daten. Auch bei den T-Daten im Sinne Cattells, also Ergebnissen objektiver Tests, neuro- und psychophysiologischen Messungen u.a., muss zwischen solchen Datenquellen unterschieden werden, die (a) in hohem Maß an sprachliche Formulierungen und immanente psychologische Konzepte gebunden sind oder (b) weitgehend sprachfrei, semantisch relativ einfacher, jedenfalls nicht mit psychologischen Alltagstheorien durchsetzt sind. Diese Abgrenzung ist wesentlich, um in Validitätsstudien anhand von Fremdeinstufungen das Risiko zirkulärer Schlüsse zu erkennen und vielleicht zu kontrollieren. Inwieweit sind die gelegentlich als befriedigend angesehenen Korrelationen zwischen Fragebogenergebnissen und Verhaltenseinstufungen (von psychologisch naiven Bekannten oder psychologisch ausgebildeten Einstufern), durch konventionelle und stereotype Anteile der impliziten Persönlichkeitsauffassungen, durch populäre Alltagstheorien, Attributionen usw. zu erklären? Wie ist eine Konfundierung der verbalen Aussagen, die einerseits als standardisierte Selbstbeschreibung im

Fragebogen, andererseits als informelle Selbstbeschreibung in der sozialen Interaktion mit dem Einstufer bzw. Bekannten gegeben werden, zu vermeiden? Die Effekte solcher „Methodenvarianz“ sind noch zu wenig untersucht, obwohl die meist unbefriedigenden Ergebnisse der Multitrait-Multimethod-Analysen in der Persönlichkeitsforschung solche Differenzierungen nahelegen (siehe Amelang & Zielinski, 1994; Fahrenberg, 1982, 1987b; Kenrick & Funder, 1988; McCrae, 1982; Ostendorf, Angleitner & Ruch 1986).

Selbstbeurteilungen spielen in der Persönlichkeitsdiagnostik weiterhin eine große Rolle, und die Verwendung von Fragebogenskalen wird immer wieder Kritik aus verschiedenen theoretischen Richtungen provozieren. Die oft vorgebrachten Hinweise auf die mögliche Wirkung von impliziten Theorien und von psychologischem Alltagswissen (Mummendey, 1999; Semin et al., 1981; Rosch et al., 1984; Westmeyer, 1995) sind doch relativ unergiebig, wenn es nicht gelingt, diese impliziten Konzepte in unabhängiger Weise genauer zu analysieren. Von besonderem Interesse wären daher entwicklungspsychologische Forschungsansätze. Längsschnittstudien über die Formung solcher Selbstbeurteilungen unter dem Einfluss früher Eigenschaftszuschreibungen durch Bezugspersonen fehlen bisher, doch bieten bereits die erinnerten Eigenschaftszuschreibungen eine interessante Informationsquelle, die in mehreren Frankfurter Arbeiten genutzt wurde (zusammenfassend Monden-Engelhardt 1986). So ergaben sich in einer Studie an 335 Personen zwischen den Dimensionen der Selbsteinstufungen und den Dimensionen der vermuteten Einstufungen durch Mutter, Vater und gute Bekannte sowie Dimensionen des FPI-G, insbesondere hinsichtlich der Skala N, signifikante Korrelationskoeffizienten geringer bis mittlerer Höhe.

Mit diesen Vorbemerkungen soll angedeutet werden, dass die Zielsetzung einer umfassenden Validierung eines Persönlichkeitsinventars zu zentralen Themen der Persönlichkeitsforschung führt: Eigenschaftstheorie und Interaktionismus-Kontroverse, Theorie der Selbstbeschreibung und die Entstehung impliziter (naiver) Konzepte, Schwierigkeiten der multiplen Operationalisierung mit dem Nachweis konvergenter und diskriminanter Validität.

## 6.2 Validitätshinweise aus der Repräsentativerhebung 1982

Anlässlich der bevölkerungsrepräsentativen Erhebung zur Entwicklung und Normierung des FPI-R wurden zahlreiche zusätzliche Daten von den Befragten gewonnen. Die Ergebnisse dieser Analysen können mit den zuvor erläuterten Einschränkungen als Validitätshinweise angesehen werden.

Bei den Variablen handelt es sich einerseits um neun der vom IfD-Interviewer notierten Statusmerkmale der Befragten (N=2024) und die Einstufung der Befragten durch den Interviewer hinsichtlich Selbstsicherheit (ziemlich unsicher=1, sehr selbstsicher, von sich überzeugt=4), andererseits um zusätzliche Daten des Fragebogens. Auf in der Regel fünfstufigen Skalen waren Selbsteinstufungen bzw. Angaben erbeten worden, u.a. zur beruflichen Belastung, zur Zufriedenheit mit Lebensbedingungen und Gesundheitszustand, zu Rauchen, Alkoholkonsum, Körperlänge, Gewicht und möglichem Übergewicht, zur Häufigkeit von Krankenhausaufenthalten, Kuren, Arztbesuchen (ohne Zahnarzt), zu chronischer Krankheit, Psychotherapie, relativer Häufigkeit der Einnahme von Beruhigungs-, Schlaf- und Schmerz-Tabletten. Diese Angaben sind nicht immer vollständig, liegen jedoch für die meisten Variablen von mehr als 1900 Befragten vor, zum Tablettenkonsum von 1726 und zur beruflichen Belastung von 1325 Befragten. Für die Gesamtstichprobe sind Koeffizienten  $r \geq 0.04$  sehr signifikant ( $p < 0.01$ ), berichtet werden jedoch hier nur solche Zusammenhänge, welche die folgenden Bedingungen erfüllen: (1) in der Gesamtstichprobe  $r \geq 0.15$ , (2) in jeder der beiden Teilstichproben Männer bzw. Frauen mindestens  $r \geq 0.10$ . Zwar ist der Anteil gemeinsamer Varianz in vielen Fällen minimal, doch sind zumindest einige dieser Koeffizienten als interessante Erläuterungen der Konstrukte aufzufassen.

### FPI-R 1 Lebenszufriedenheit

- höhere Zufriedenheit mit familiärer Situation (.46), mit ausgeübter Berufstätigkeit (.43), mit finanzieller Situation (.36), mit Kollegen und Vorgesetzten (.31), mit Wohnsituation (.28), relativ guter Gesundheitszustand (.28), selten Beruhigungstabletten (.22), nicht in Psychotherapie gewesen (.19), verheiratet oder mit Partner zusammenlebend (.18), selten Schmerztabletten (.17), höhere Selbstsicherheit laut Interviewer-Einstufung (.16), höheres Einkommen (.15), weniger Arztbesuche (.15).

### FPI-R 2 Soziale Orientierung

- relativ häufiger Kirchenbesuch (.22).

### FPI-R 3 Leistungsorientierung

- höhere berufliche Belastung durch Verantwortung (.30), höhere Selbstsicherheit laut Interviewer-Einstufung (.29), höhere Zufriedenheit mit ausgeübter Berufstätigkeit (.25), berufstätig (.23), höhere berufliche Belastung durch Tempo, Genauigkeit, Konzentration (.21), relativ guter Gesundheitszustand (.21), höheres Einkommen (.20), selten Schlafmittel (.18), höhere Schicht (.15), weniger Arztbesuche (.15).

#### FPI-R 4 Gehemmtheit

- geringere Selbstsicherheit laut Interviewer-Einstufung (.27), schlechterer Gesundheitszustand (.22), häufiger Beruhigungstabletten (.15), geringere Zufriedenheit mit Kollegen und Vorgesetzten (.15).

#### FPI-R 5 Erregbarkeit

- häufiger Beruhigungstabletten (.20), geringere Zufriedenheit mit familiärer Situation (.17), mit finanzieller Situation (.17), mit ausgeübter Berufstätigkeit (.17), schlechterer Gesundheitszustand (.17), geringere Zufriedenheit mit Kollegen und Vorgesetzten (.15).

#### FPI-R 6 Aggressivität

- Raucher (.18), geringere Zufriedenheit mit finanzieller Situation (.17), stärkere Belastung durch körperliche Arbeit (.16).

#### FPI-R 7 Beanspruchung

- höhere berufliche Belastung durch Tempo, Genauigkeit, Konzentration (.23), durch körperliche Arbeit (.18), geringere Zufriedenheit mit Kollegen und Vorgesetzten (.18), Raucher (.18), schlechterer Gesundheitszustand (.18), berufstätig (.16), häufiger Beruhigungsmittel (.16), geringere Zufriedenheit mit finanzieller Situation (.15).

#### FPI-R 8 Körperliche Beschwerden

- schlechter Gesundheitszustand (.51), häufiger Schlaftabletten (.46), Beruhigungstabletten (.41) und Schmerztabletten (.38), häufiger Arztbesuche (.35), chronische Krankheit (.25), nicht berufstätig (.23), geringere Selbstsicherheit laut Interviewer-Einstufung (.19), häufigere Krankenhausaufenthalte, Operationen, Kuren (.19), wenige Personen im Haushalt des Befragten (.18), geringere Zufriedenheit mit familiärer Situation (.16), sich selbst als übergewichtig bezeichnend (.15).

#### FPI-R 9 Gesundheitssorgen

- häufigere Arztbesuche (.28), schlechterer Gesundheitszustand (.25), häufiger Beruhigungstabletten (.21) und Schlaftabletten (.19), Nicht-Raucher (.19), relativ häufiger Kirchenbesuch (.18), häufigere Krankenhausaufenthalte, Operationen, Kuren (.16), Tendenz zu Übergewicht laut Broca- und Rohrer-Index (.17), chronische Krankheiten (.15), nicht berufstätig (.15).

#### FPI-R 10 Offenheit

- geringere Zufriedenheit mit familiärer Situation (.20), finanzieller Situation (.19), ausgeübter Berufstätigkeit (.18) und Wohnsituation (.17), relativ selten oder nie Kirchenbesuch (.17), geringere Zufriedenheit mit Kollegen und Vorgesetzten (.15).

#### FPI-R E Extraversion

- höhere Selbstsicherheit laut Interviewer-Einstufung (.26), besserer Gesundheitszustand (.25), selten Schlaftabletten (.19), berufstätig (.16), Raucher (.16).

#### FPI-R N Emotionalität

- relativ schlechter Gesundheitszustand (.38), häufiger Beruhigungstabletten (.32), geringere Zufriedenheit mit familiärer Situation (.31), finanzieller Situation (.25), ausgeübter Berufstätigkeit (.25), Kollegen und Vorgesetzten (.24), häufiger Schlaftabletten (.24) und Schmerztabletten (.24), geringere Zufriedenheit mit Wohnsituation (.19), chronische Krankheit (.17), in Psychotherapie gewesen (.16).

Einige dieser Zusammenhänge, etwa im Bereich Lebenszufriedenheit, Beanspruchung, Körperliche Beschwerden und Gesundheitssorgen, sind naheliegend und lassen die innere Konsistenz von Fragebogenitems und zusätzlichen Selbsteinstufungen erkennen. Bemerkenswerter sind die Zusammenhänge zwischen der von den Interviewern eingestuften



Selbstsicherheit des Befragten und FPI-R Skalen oder die Korrelationen mit Selbsteinstufungen zu Tablettenkonsum, Rauchen und Psychotherapie. Zu Statusmerkmalen zeigen sich – von Geschlecht, Alter und Schulabschluss abgesehen – hinsichtlich Berufstätigkeit, Einkommen, relativer Häufigkeit des Kirchenbesuchs und Familienstand bzw. Personen-zahl im Haushalt einige Beziehungen.

Der aus Fragebogenskalen und sozioökonomischen Informationen bestehende Datensatz wurde noch auf eine zweite Weise analysiert. Die IfD-Variablen dienen als Klassifikationsmerkmale (Prädiktoren), um schrittweise Kontrastgruppen mit den minimalen und maximalen Mittelwerten der abhängigen FPI-R Variablen aufzufinden. Das hier verwendete Programm AID (Sonquist, Baker & Morgan 1973, siehe auch Hampel & Krupp 1978) beruht im Prinzip auf einer varianzanalytisch-regressionsanalytischen Prozedur. Aus der Liste der potenziellen Prädiktoren wird jener gesucht, welcher die Werte der abhängigen Variable am besten vorhersagt. Es werden zwei Extremgruppen gebildet, von denen jede dann wie die Grundgesamtheit behandelt wird, so dass ein Baum von Kontrastgruppen entsteht, bis das Abbruchkriterium erreicht ist (hier: noch 1% aufgeklärte Varianz der FPI-R Skala).

Die Tabelle 6.1 zeigt die resultierenden Extremgruppen der letzten erreichten Stufe und die jeweils erhaltenen Prädiktoren. Zwischen den Skalen gibt es große Unterschiede, welche IfD-Merkmale als wesentliche Prädiktoren dienen und wie viel Prozent der Varianz in den Persönlichkeitsmerkmalen durch sozioökonomische Merkmale aufzuklären ist. Die höchste Lebenszufriedenheit z.B. findet sich in der Untergruppe der Verheirateten, welche entweder Berufstätige oder Rentner sind, die meisten körperlichen Beschwerden in der Untergruppe alleinstehender Hausfrauen und Rentner usw. Alle Befunde sind mit der psychologischen Interpretation der FPI-R-Skalen gut zu vereinbaren.

Die als Korrelationskoeffizienten berichteten Beziehungen zwischen FPI-R und Zusatzdaten wurden durch Mittelwertvergleiche alters- und geschlechtsparallelisierter Gruppen für einige der interessantesten Kriterien überprüft. Die folgenden Informationen stammen zwar ebenfalls aus dem Fragebogen, enthalten aber konkretere Daten als die anderen Items.

1. Angabe des Befragten, ob eine chronische innere Krankheit (ohne rheumatische, allergische und psychiatrische Bilder) besteht, beschränkt auf die Altersgruppen 25–59;
2. Anzahl der Arztbesuche während der letzten 12 Monate: 7 und mehr Arztbesuche;
3. in psychotherapeutischer Behandlung gewesen: ja;
4. Tabakkonsum: 26 oder mehr Zigaretten-Einheiten täglich;
5. Alkoholkonsum: 5,6 und mehr Bier-Einheiten (Liter) wöchentlich (Wein: Faktor 2);
6. Tablettenkonsum (Schlaf-, Beruhigungs-, Schmerz-Tabletten): Index 7 und höher (entspricht Einnahme von jeder Tablettenkategorie „einige Male im Jahr“ bzw. von zwei Tablettenkategorien „einige Male im Monat“ bis „mehrmals in der Woche“.)

Die Gruppen (2), (4) bis (6) sind so gebildet, dass sie ungefähr das oberste Quartil der Verteilung umfassen. Die Tabelle 6.2 enthält die Ergebnisse des Mittelwertvergleichs zwischen diesen Extremgruppen und einer gleich großen Kontrollgruppe, welche nach dem Zufallsprinzip aus der jeweiligen Restpopulation der entsprechenden Altersgruppe der Männer bzw. der Frauen gebildet wurde. Die Ergebnisse können als deutliche Validitätsbelege vor allem für die Skalen Lebenszufriedenheit, Erregbarkeit, Beanspruchung, Körperliche Beschwerden und Emotionale Labilität angesehen werden; spezielle Hinweise gibt es für die Skalen Leistungsorientierung, Gehemmtheit und Gesundheitssorgen, während Soziale Orientierung im erwarteten Sinn unabhängig bleibt.

#### *Einstellung zu gesellschaftlichen und politischen Themen*

Anlässlich der Repräsentativerhebung wurden den Befragten 25 Aussagen zu „Aktuellen Themen“ vorgelegt. Die Items und Itemstatistiken sowie Beziehungen zu ausgewählten IfD-Statusmerkmalen sind im Tabellenband (Hampel & Fahrenberg 1983) enthalten. Aufgrund von Faktoren- und Itemanalysen wurden aus diesem Itempool u.a. thematische Item-Paare und eine 23 Items umfassende Skala konservativ-progressiver Tendenz gebildet. Hervorzuheben sind hier zwei Subskalen zu je 8 Items.

Skala 1 fasst autoritäre Einstellungskomponenten zusammen: zu viel Aufwand für die Betreuung von Verbrechern, Geringschätzung von Wehrdienstverweigerern, gegen Auseinandersetzung über Judenvernichtung und Drittes Reich, gegen Gleichberechtigung von Gastarbeitern, gegen Demonstrationen und Friedensbewegung.

Skala 2 beschreibt eine pessimistische Zukunftsperspektive: zunehmende Verschlechterung der Wirtschaftslage, Umweltverschmutzung, Ablehnung von Kernkraftwerken, deutlicher Wunsch, aus den Lebensbedingungen der Bundesrepublik auszusteigen, Hoffnung auf Friedensbewegung bei grundsätzlich pessimistischer Zukunftssicht.

Mit autoritärer Einstellung korrelieren negativ FPI-R 2 Soziale Orientierung ( $r=-0.33$ , nach Auspartialisierung von Geschlecht und Alter  $r_{\text{part}}=-0.38$ ) und positiv FPI-R 6 Aggressivität ( $r=0.14$ ,  $r_{\text{part}}=0.22$ ). Mit pessimistischer Zukunftsperspektive korrelieren negativ FPI-R 1 Lebenszufriedenheit ( $r=-0.30$ ,  $r_{\text{part}}=-0.29$ ) und positiv FPI-R N Emotionalität ( $r=0.26$ ,  $r_{\text{part}}=0.25$ ).

#### *Multiple Regression*

Die Tabelle 6.3 gibt durch einige ausgewählte multiple Regressionen eine Übersicht über systematische Beziehungen zwischen den genannten demografischen, gesundheitlichen u.a. Variablen und einigen Skalen des FPI-R.

Bemerkenswert ist die Rolle von Lebenszufriedenheit, Beanspruchung und Leistungsorientierung. Im Gegensatz zu allgemeinen Belastungs- und Zufriedenheitsdaten, Gesundheitsverhalten und Einstellungen sind offensichtlich die Parteiidentifikation oder die Häufigkeit des Kirchenbesuchs sehr viel schlechter bzw. nur in einem minimalen Varianzteil vorherzusagen.

### **6.3 Validitätshinweise aus der Repräsentativerhebung 1999**

Bei der neuen Repräsentativerhebung konnten über die soziodemografischen Merkmale hinaus zahlreiche Daten zur beruflichen Belastung, Gesundheit u.a. Bereichen gewonnen werden. Außerdem nahmen die Interviewer Verhaltenseinstufungen der Befragten vor. Der Fragebogen enthielt schließlich Items zu politisch-gesellschaftlichen Einstellungen. Da 10 dieser Items auch 1982 verwendet wurden, können die Mittelwerte auf mögliche Veränderungen geprüft werden.

#### *Politisch-gesellschaftliche Einstellungen*

Der Wortlaut der politisch-gesellschaftlichen Items und die Mittelwerte, welche den Grad der Zustimmung erkennen lassen, sind in Tabelle 6.4 zusammengefasst. Es gibt große Unterschiede zwischen einigen allgemein akzeptierten Aussagen mit mehr als 90% Zustimmung (Frauen auch Vorgesetzte von Männern; Frauen genauso leistungsfähig wie Männer; Notwendigkeit vertrauenswürdiger Politiker) und Aussagen, denen weniger als ein Viertel der Befragten zustimmt: die Bereitschaft, Kontakt mit Holocaust-Überlebenden aufzunehmen; zuviel Geld für die Wissenschaft; Tendenz zur Auswanderung wegen der gegenwärtigen Lebensbedingungen.

Explorative Faktorenanalysen führten zur Gruppierung der 32 Items in fünf Skalen, die als Nationalkonservative, Liberale, Ökologische, Internationalistische und Frauenpolitische Einstellung zu charakterisieren sind. Die Zuordnung der Items zu diesen Skalen, Mittelwerte und Reliabilitätskoeffizienten sind aus Tabelle 6.5 zu entnehmen.

Zwischen Nationalkonservativer und Liberaler Einstellung besteht erwartungsgemäß eine negative Korrelation. Eine frauenpolitisch progressive Einstellung korreliert negativ mit Nationalkonservativer und positiv mit Liberaler Einstellung (Tabelle 6.6). Eine stärkere Nationalkonservative Einstellung findet sich bei älteren Personen und bei geringerer Schulbildung, eine eher liberale Einstellung bei höherer Schulbildung.

Zwischen politisch-gesellschaftlichen Einstellungen und Testwerten des FPI-R bestehen zahlreiche signifikante, aber numerisch meist nur geringe Beziehungen. Hervorzuheben sind die Zusammenhänge zwischen höherer Lebenszufriedenheit und geringerer Ausprägung der ökologischen Einstellung sowie zwischen sozialer Orientierung und geringerer Ausprägung Nationalkonservativer und stärkerer Ausprägung Liberaler Einstellung. Aggressivität scheint mit Nationalkonservativer und Ökologischer Einstellung bzw. mit eher geringer Liberaler und Frauenpolitischer Einstellung assoziiert zu sein, während Beanspruchung und – relativ hoch – Emotionalität mit Ökologischer Einstellung korrelieren.

Für die westdeutsche Bevölkerung können bei zehn der Items Mittelwertvergleiche zur Prüfung von Einstellungsänderungen zwischen 1982 und 1999 durchgeführt werden (Tabelle 6.7). Tatsächlich zeigen sich in drei Items sehr signifikante ( $p<0.01$ ) Veränderungen: eine kritischere Haltung zu Kernkraftwerken und zur Umweltverschmutzung sowie der Wunsch nach einer starken, integrativen Partei. Von diesen Items abgesehen ist die westdeutsche Bevölkerung in ihren Meinungen hinsichtlich der hier erfassten Aspekte politisch-gesellschaftlicher Einstellungen zwischen 1982 und 1999 bemerkenswert konstant geblieben.

Auf einen Vergleich der politisch-gesellschaftlichen Einstellungen zwischen der Bevölkerung der alten und der neuen Bundesländer auf der Basis der neuen Erhebung wird an dieser Stelle verzichtet, weil die Konfundierung mit wichtigen soziodemografischen Merkmalen berücksichtigt werden muss (siehe unten).

### *Verhaltenseinstufungen*

Die Interviewer stuften ein, wie selbstsicher, energisch und freundlich die Befragten auf sie wirkten. Außerdem beurteilten sie die soziale Schichtzugehörigkeit. Die höchste Korrelation besteht zwischen den Einstufungen „selbstsicher“ und „energisch“ sowie „freundlich“ und höherer sozialer Schicht. Der Eindruck der Selbstsicherheit korreliert deutlich mit hoher Schulbildung und höherem Einkommen der Befragten, der Eindruck energischen Verhaltens mit der Schulbildung (Tabelle 6.8).

Unter den zahlreichen hochsignifikanten Zusammenhängen zwischen diesen Verhaltenseinstufungen und den unabhängig erhobenen Testwerten des FPI-R sind die folgenden hervorzuheben: Höhere Lebenszufriedenheit mit der Verhaltenseinstufung selbstsicher, energisch, freundlich; Soziale Orientierung und freundlich; Leistungsorientierung und selbstsicher, energisch; Gehemmtheit und weniger selbstsicher, weniger energisch; Körperliche Beschwerden und weniger selbstsicher; Extraversion und selbstsicher, energisch. Die gemeinsame Varianz von Persönlichkeitseigenschaften und Verhaltenseinstufungen erreicht bei den Skalen Soziale Orientierung, Leistungsorientierung und Extraversion ca. 10%.

Angesichts dieser relativ hohen Kriterienkorrelationen sind methodenkritische Vorbehalte angebracht. Die Verhaltenseinstufungen erfolgten zwar ohne Kenntnis der FPI-Testwerte, aber im Kontext der Informationen, die während des Interviews gewonnen wurden. Die Verhaltenseinstufungen wurden von mehr als 500 Interviewern ohne methodische Vorbereitung durchgeführt. Um so bemerkenswerter sind die beobachteten Zusammenhänge, welche psychologisch ohne Schwierigkeiten mit den betreffenden Persönlichkeitskonzepten in Verbindung zu bringen sind. Es ist auch einsichtig, dass die Bereiche Erregbarkeit, Beanspruchung, Körperliche Beschwerden und Gesundheitssorgen in den drei ausgewählten Verhaltensaspekten kaum erfasst wurden. Die Ergebnisse belegen auch, dass die ausgewählten Konzepte geeignet sind, solche Verhaltenseinstufungen zu ermöglichen. Offensichtlich gibt es eine teilweise Übereinstimmung zwischen den populären psychologischen Konzepten der Selbstbeurteilung und der Fremdbeurteilung.

Auf einem methodisch unüblichen Weg wurden Validierungshinweise für mehrere FPI-Skalen gewonnen, insbesondere für Lebenszufriedenheit, Soziale Orientierung, Leistungsorientierung, Gehemmtheit und Extraversion. Die Verhaltensbeurteilungen durch die Interviewer stimmen mit den Selbstbeurteilungen der Befragten in diesen Persönlichkeitseigenschaften in einem unerwarteten Ausmaß überein.

### *Berufliche Belastung, Lebenszufriedenheit, Gesundheit*

Eine Übersicht über die vorliegenden Informationen d.h. die Selbsteinstufungen der beruflichen Belastung hinsichtlich verschiedener Faktoren, die Lebenszufriedenheit in verschiedenen Bereichen und zahlreiche Aspekte von Gesundheit und Krankheit gibt die Tabelle 6.9.

In vielen Merkmalen bestehen, wie Mittelwerte und Korrelationskoeffizienten anzeigen, deutliche Beziehungen zum Geschlecht und zum Alter. Dies gilt nicht nur für Körperbaumaße, Krankheiten und Medikamente, sondern auch für die Indizes der beruflichen Belastungen und für die Lebenszufriedenheit in verschiedenen Bereichen. Einfache varianzanalytische Auswertungen würden jedoch zu fragwürdigen Ergebnissen führen, da auch hier wichtige soziodemografische Einflussfaktoren als konfundiert anzusehen sind. So sind sehr viel weniger Frauen berufstätig, die Art der beruflichen Belastung hängt u.a. von der Schulbildung ab und viele Aspekte des Bereichs Gesundheit korrelieren sowohl mit dem Geschlecht als auch mit dem Lebensalter und anderen Bedingungen. Durch Partialkorrelationen bzw. Kovarianzanalysen wäre eine befriedigende statistische Kontrolle dieser multiplen Effekte nicht möglich.

### *Statistische Zwillingsbildung*

Bei großen Datensätzen kann die Methode der statistischen Zwillingsbildung (matched pairs) zur Kontrolle der wichtigsten Einflussgrößen verwendet werden. Das hier von Dipl. Math. F. Foerster programmierte Verfahren wurde bereits bei den entsprechenden Analysen im Zusammenhang mit der Entwicklung der Freiburger Beschwerdenliste (Fahrenberg, 1994a) und des Fragebogens zur Lebenszufriedenheit (Fahrenberg et al., 2000) benutzt.

Hinsichtlich eines bestimmten Gruppierungsmerkmals, z.B. ob sich eine Person in psychotherapeutischer Behandlung

befand (befindet) oder nicht, werden zwei Klassen gebildet. Für jede Person der ersten (kleineren) Gruppe wird in der zweiten Gruppe ein „statistischer Zwilling“ gesucht, welcher dieselbe Kombination soziodemografischer Merkmale aufweist, u.a. Geschlecht, Altersgruppe, Schulabschluss usw. In einer schrittweisen hierarchischen Prozedur können noch weitere soziodemografische Merkmale hinzugefügt werden. In den entstandenen Gruppen sind dann die betreffenden soziodemografischen Merkmale für den Mittelwertvergleich der abhängigen Variablen, hier der FPI-R Skalen, kontrolliert.

Für das Matching wurden die folgenden Merkmale ausgewählt (Tabelle 6.10): Geschlecht, Alter, Schulabschluss, Einkommen und Ost-/Westdeutschland. Da diese Auswahl bereits zu  $2 \times 7 \times 3 \times 4 \times 2 = 336$  möglichen, allerdings sehr unterschiedlich besetzten Merkmalskombinationen führt, wurde auf weitere Merkmale verzichtet. Die hier verwendete Auswahl repräsentiert wesentliche soziodemografische Einflussgrößen. Aus persönlichkeitspsychologischer Sicht handelt es sich nicht nur um „unabhängige“ Variablen wie Geschlecht, Alter und Bundesland, sondern auch um Variablen, deren Ausprägung von Persönlichkeitsmerkmalen mitbedingt sein wird: Schulabschluss und Einkommen.

In der Tabelle 6.11 sind für ausgewählte Gruppierungsmerkmale die Umfänge der Teilstichproben angegeben. Das kleinere der beiden  $n$  zeigt an, welche Anzahl von Zwillingspaaren maximal erreicht werden könnte. Die tatsächlich erhaltene Anzahl von Paaren für den Mittelwertvergleich ist in der Regel viel geringer (siehe Tabelle 6.12).

In einem ersten Schritt dieser Analysen wurden die fünf soziodemografischen Merkmale, welche dann im zweiten Abschnitt zu kontrollieren sind, ihrerseits in einer schrittweisen Prozedur analysiert. Dabei wurde das jeweilige Merkmal nicht mehr als unabhängige, sondern als abhängige Variable angesehen. Da von dem Matching-Verfahren nur eine dichotome Gruppierung vorgenommen werden kann, wurden hinsichtlich des Alters zwei Kontraste ausgewählt und hinsichtlich Schulbildung und Einkommensgruppe wurden Extremgruppen verwendet (Zeile 1 bis 6 der Tabelle 6.12). Beim Matching der folgenden Gruppierungen wurden jedoch alle Stufen der fünf soziodemografischen Merkmale berücksichtigt.

Die anschließenden Zeilen der Tabelle 6.12 zeigen, welche Mittelwertunterschiede es in den abhängigen Variablen (FPI-R Skalen u.a.) gibt, wenn die Einflüsse von Geschlecht, Altersgruppen, Schulabschluss, Einkommensgruppe und Ost – West durch multiple Parallelisierung kontrolliert werden. Da viele der Gruppierungsmerkmale wie Familienstand, Berufsgruppen, Parteipräferenzen, Krankenhausaufenthalte usw. mehr als zwei Klassen haben, mussten auch unter Berücksichtigung der Personenzahl jeweils zwei dieser Klassen – oft die unterste und die oberste Klasse – ausgewählt werden.

#### *FPI-Skalen und Kriterien*

Die soziodemografisch kontrollierten Gruppierungen mit Mittelwertvergleichen führten zu zahlreichen sehr signifikanten ( $p < 0.001$ ) Ergebnissen. Die Mittelwerte und  $p$ -Werte sind in Tabelle 6.12 für die 12 FPI-R Skalen und die 4 Einstellungsskalen zusammengestellt. Durch die statistischen Kontrollen der fünf wichtigsten soziodemografischen Merkmale haben diese Befunde eine höhere Aussagekraft als oberflächliche univariate Vergleiche. Es könnten hier Alpha-Adjustierungen vorgenommen werden, doch würde dies angesichts der sehr zahlreichen hochsignifikanten Befunde ( $p < 0.001$ ) kein wesentlich anderes Bild liefern.

Aus den obersten Zeilen dieser Ergebnistabelle sind die Effekte der kontrollierten Variablen zu entnehmen. Die in den folgenden Zeilen beschriebenen Effekte sind wesentliche Validierungshinweise, die z.T. auch als empirische Ergebnisse von eigenständiger Bedeutung aufgefasst werden können. Die meisten Bezüge zu den ausgewählten Variablen wiesen die FPI-R Skalen Emotionalität, Lebenszufriedenheit, Körperliche Beschwerden und Erregbarkeit auf. Keine Beziehungen zwischen FPI-Skalen und Kriterienvariablen ergaben sich hinsichtlich Nord/Süd (welches als Kontrollgruppierung für die Ost-West-Fragestellung diente), Land/Stadt, mit/ohne Partner lebend, Kinder im Haushalt, Konfession evangelisch/katholisch, Mitgliedschaft in einer Gewerkschaft.

Die berichteten Befunde, welche als gut abgesichert gelten können, lassen sich ohne Schwierigkeiten mit den Eigenschaftskonzepten der FPI-R Skalen vereinen und sie wurden in die Beschreibung der FPI-R Skalen in Kapitel 4 aufgenommen. Solche Zusammenhänge wurden bereits in der vorausgegangenen Analyse an dem Datensatz von 1982 durch bivariate Korrelationen, durch Vergleiche von Kontrastgruppen und durch Regressionsanalysen (Kapitel 6.2) herausgearbeitet. Die jetzt verwendete Methodik hat jedoch Vorzüge.

Die vielfältigen empirischen Zusammenhänge zwischen FPI-R Skalen und soziodemografischen Variablen und Daten

aus dem Gesundheitsbereich sind als wichtige Kriterienkorrelationen und empirische Validierungshinweise des Tests anzusehen.

### *Einstellungen*

Auch zwischen den für diese Untersuchung abgeleiteten Einstellungsskalen zeigen sich zahlreiche Beziehungen zu soziodemografischen und anderen Variablen, wenn die wichtigen Einflüsse anderer Merkmale kontrolliert werden. Eine nationalkonservative Einstellung ist eher bei Männern, u.a. bei den hier verglichenen Altersgruppen der 20- bis 29-Jährigen im Vergleich zu den 17- bis 19-Jährigen, bei Personen mit geringerer Schulbildung, in Ostdeutschland, bei Verheirateten, bei Personen mit Präferenz für die CDU/CSU im Vergleich zu SPD, Grünen, PDS, und Präferenz für SPD im Vergleich zu PDS zu finden. Höhere Testwerte zeigen sich außerdem bei Rauchern und bei objektivem Übergewicht (Body Mass Index).

Eine liberale Einstellung ist vergleichsweise eher bei Frauen, bei den 17- bis 19-Jährigen im Vergleich zu den 20- bis 29-Jährigen, bei Personen mit höherer Schulbildung, bei relativ häufigem Kirchenbesuch, bei Parteipräferenz für SPD im Vergleich zur CDU/CSU, für Grüne im Vergleich zur CDU/CSU sowie bei Präferenz für PDS oder Republikaner im Vergleich zur CDU, sowie bei Nichtrauchern zu erkennen.

Eine ökologische Einstellung zeigt sich eher bei Personen mit Hauptschulabschluss, in Ostdeutschland, bei Nicht-Verheirateten, bei Arbeitslosen/Nicht-Berufstätigen, unterer sozialer Schichtzugehörigkeit, Parteipräferenz PDS im Vergleich zu CDU/CSU und SPD, bei schlechterem Gesundheitszustand, Rauchern, bei Personen mit mehr Kuren, psychotherapeutischer Behandlung und häufigerer Einnahme von Tabletten. Die internationalistische Einstellung tritt eher in Westdeutschland hervor, bei Personen, die Kirchenmitglieder sind, eine Präferenz für die CDU/CSU im Vergleich zur SPD und PDS haben. Die frauenpolitische Einstellung zeigt sich bei Frauen, Personen mit höherer Schulbildung, Parteipräferenz für Grüne, PDS im Vergleich zu CDU/CSU und SPD bzw. CDU im Vergleich zu Republikanern, sowie bei Nichtrauchern.

Mit den wenigen Items dieser Einstellungsskalen sind sicher nur einige Facetten der Konstrukte erfasst. Dies mag einige der Zusammenhänge, die auf den ersten Blick überraschend sein mögen, verständlich machen. So treffen die Items der als ökologisch bezeichneten Einstellung offensichtlich nicht das Profil der Wähler von Bündnis 90/Die Grünen. Die internationalistische Einstellung ist hier u.a. durch Items einer interventionistischen Auffassung (Kosovo, Afrika) bestimmt, so dass andere Aspekte des Internationalismus u.a. im Sinne gerechter Güterverteilung und Partizipation zu kurz kamen. Die Mehrzahl der Kriterienbeziehungen ist jedoch mit den Konzepten nationalkonservativer, liberaler und frauenpolitischer Einstellung gut zu vereinbaren.

### *Weitere Kriterienbeziehungen zum beruflichen und gesundheitlichen Bereich*

Die kontrollierten Mittelwertvergleiche können hinsichtlich weiterer Gruppierungen vorgenommen werden, um Aspekte der diskriminanten Validität weiter herauszuarbeiten. Die Tabelle 6.13 ist trotz des absichtlich anders gewählten Aufbaus als Fortsetzung der Tabelle 6.12 zu sehen.

Leitende Angestellte haben im Vergleich zu nicht-leitenden Angestellten höhere Werte in Leistungsorientierung und Extraversion sowie niedrigere in Gehemmtheit. Beim Vergleich zu Facharbeitern kommen außerdem höhere Testwerte in Sozialer Orientierung hinzu. Arbeitslose haben im Vergleich zu Berufstätigen eine geringere Lebenszufriedenheit, Leistungsorientierung und Beanspruchung sowie höhere Gehemmtheit, Erregbarkeit, körperliche Beschwerden und Emotionalität. Bei höherer beruflicher Belastung (Index BERBEL>11) im Vergleich zu geringerer beruflicher Belastung (Index BERBEL<8) zeigen sich höhere Leistungsorientierung und Beanspruchung. Angehörige der beiden obersten sozialen Schichten haben im Vergleich zu Angehörigen der untersten Schicht, mit Ausnahme von Beanspruchung, Gesundheitssorgen und Offenheit, in fast allen Skalen abweichende Mittelwerte.

Die Gruppierung nach chronischer Krankheit umfasst hier nicht alle 460 Personen, welche diese Fragen bejahten, sondern 416 Personen, welche anschließend eine bestimmte Krankheit aufgeschrieben hatten: Bluthochdruck n=40; Herzinfarkt, Schlaganfall n=38; Rheuma, Arthrose, Bandscheiben n=74; Asthma, Bronchitis n=72; Diabetes n=62; Allergie n=55; Sonstige n=43; Multimorbide n=32.

Personen mit chronischen Krankheiten hatten generell höhere Testwerte in Sozialer Orientierung, Erregbarkeit, Beanspruchung, Körperlichen Beschwerden und Emotionalität, dagegen gab es hier u.a. hinsichtlich Lebenszufriedenheit und auch bei den Gesundheitssorgen keine Mittelwertunterschiede. Dieses Muster zeigte sich auch beim Vergleich der

einzelnen Krankheitsgruppen mit ihren gesunden Kontrollgruppen.

Wer häufig (mehr als 10-mal im Jahr) zum Arzt gegangen ist, unterschied sich von anderen Personen, die nicht beim Arzt waren, durch höhere Testwerte der Erregbarkeit, Beanspruchung, Körperliche Beschwerden, Gesundheitssorgen und Emotionalität. Dieses Muster kehrt bei Personen, die in Psychotherapie sind oder waren, mit Ausnahme der Gesundheitssorgen wieder. Psychotherapiepatienten äußern jedoch eine ausgeprägtere Soziale Orientierung, die hier vielleicht als größere Aufgeschlossenheit für soziale Zuwendung bzw. als Hilfesuche und Hilfsbereitschaft zu interpretieren ist.

## 8.16 Diagnostische Urteilsbildung: Datenintegration statt Interpretation?

Das diagnostische Gutachten (Amelang & Schmidt-Atzert, 2006; Fisseni, 2004) verknüpft die einzelnen Befunde zu einem diagnostischen Urteil. Hier muss erneut grundsätzlich zwischen Intelligenz- und Leistungstests und Persönlichkeitsfragebogen unterschieden werden. Von den einzelnen Validitätshinweisen ist die *empirische Validität der Testinterpretation* zu unterscheiden. Bei Persönlichkeitsfragebogen sind, noch kritischer als z.B. bei Intelligenz- und Leistungstests, der Kontext der Anwendung und die Möglichkeit von Antworttendenzen zu berücksichtigen. Im Interpretationsprozess der diagnostischen Urteilsbildung muss grundsätzlich zwischen der Selbstbeurteilung des Verhaltens und dem manifesten Verhalten unterschieden werden. Das Wissen über die methodischen Schwierigkeiten der Selbstbeurteilungen in Fragebogen (und Interviews) und die bekannten Methodenprobleme von Persönlichkeitsfragebogen machen deren fachlich adäquate Auswertung und Anwendung zu einer herausfordernden Aufgabe. Darüber hinaus existieren die kritischen Einsichten aus der multimodalen Diagnostik, die viele wichtige Praxisbereiche betreffen. Nach welchen Regeln sollen diese Daten verknüpft und für die Urteilsbildung interpretiert werden?

In seinen *Essentials of psychological testing* hatte Cronbach (1970) drei typische Strategien der Interpretation von Fragebogenergebnissen unterschieden: Der Persönlichkeitsfragebogen wird als Selbstbeschreibung angesehen, die Inhalte dienen einer psychologisch bzw. psychoanalytisch orientierten, inhaltlichen Interpretation oder werden aktuarisch, d.h. aufgrund von Kriterienkorrelationen, verwendet. Als Anwendungsmöglichkeiten beschrieb er hauptsächlich die Unterscheidung zwischen Patienten und Gesunden, die Suche nach auffälligen Personen für eine genauere Untersuchung, die Klassifikation von Patienten. Außerhalb des klinischen Bereichs dominieren die Vorhersage des Berufserfolgs bzw. des akademischen Erfolgs und die psychologische Begutachtung für institutionelle Entscheidungen. – Die Aufgaben der Klassifikation, die Selektion bzw. das Screening, die Vorhersage und Begutachtung werden auch heute unterschieden. Doch in der differenziellen Psychologie wurden seit Cattell zahlreiche typische Assessmentstrategien entwickelt, um die inter- und intraindividuelle Variabilität der multivariaten Datenbox perspektivisch zu differenzieren und die speziellen Konstrukte, Aggregate und Indizes zu unterscheiden. In den meisten Lehrbuchtexten zur psychologischen Diagnostik fehlen diese fortgeschrittenen Konzepte (und der Begriff Assessmentstrategie).

Praktische Regeln der Kombinatorik, die in der scheinbar als veraltet geltenden diagnostischen Psychologie (Heiß, 1982; vgl. Fahrenberg, 2002) intensiv trainiert wurden, werden heute kaum noch erwähnt. Nur als Hinweis sind zu nennen u.a. *Aspekte* wie Interpretationsebenen und Interpretationstiefe, Interpretationsdivergenz (Widerspruchsanalyse), Vermeidung von Vorurteilen und Reflexion der Abhängigkeiten, Kontrolle durch eine Interpretationsgemeinschaft, *Prinzipien* der Folgerichtigkeit, Ebenen des Kontextbezugs, Einpassung in Muster, Bedeutung von sog. Dominanten, *technische Regeln* über typische Marker für individuelle Auffälligkeiten, über Gewichtung und Kombinatorik usw. Demgegenüber werden heute oft nur die logischen Regeln, d.h. die konjunktive, additive und disjunktive Verknüpfung erläutert, oder allgemeine Unterschiede zwischen hypothesengeleiteter und explorativer Diagnostik. Oft wird der Informationsverarbeitungs-Prozess mit ausgedehnten Flussdiagrammen illustriert (z.B. Westhoff, Hagemeyer & Strobel, 2006; Schmitt & Gschwendner, 2006). Viele Lehrbuchbeiträge über die Methodik psychologischer Gutachten bleiben sehr allgemein oder schildern primär die äußere Gestaltung und Kommunikation der Ergebnisse. In der Literatur gäbe es „verstreute empirisch gesicherte Regeln“ (Westhoff et al., 2006, S. 398) und in der entscheidungsorientierten Diagnostik wären diese Regeln zusammengetragen und stünden in Form von Checklisten zu Verfügung (Kubinger, 2003a; Westhoff & Kluck, 2003). Diese Checklisten enthalten eine große Anzahl von Gesichtspunkten mit sehr knappen Erläuterungen ohne übergreifende strategische Konzeption. Nur sehr selten werden,

wie von Fisseni (2004), mehr Hinweise auf die notwendige Kombinatorik gegeben. Demgegenüber schreibt Westhoff (2004) von der notwendigen Komplexitätsreduktion, die vorab geplant werden und nachvollziehbar sein müsse, und weist allgemein darauf hin, dass Aussagen zu kombinieren sind: „Sollten sich Informationen widersprechen, so ist diese Tatsache zu berichten ...“ – Aber sollen die Experten ggf. die Deutung der Widersprüche den Auftraggebern überlassen? Aus dieser Sicht scheinen psychologische Widersprüche zwischen Befunden nur Ausnahmen zu sein, für deren Interpretation keine methodischen Regeln oder Standards entwickelt werden müssen.

Das bekannte Schema der richtigen und falschen Diagnosen ist oft dargestellt, jedoch wird selten erläutert, weshalb für Cronbach neben der Nutzenfunktion auch die *Schadensfunktion* solcher professionellen Entscheidungen wichtig war. Die rationale Bewertung des Schadens verlangt natürlich das zu tun, was meistens fehlt, d.h. ein unverzerrtes follow-up auch der abgewiesenen, letztlich vielleicht viel geeigneteren Bewerber bzw. der falsch diagnostizierten oder falsch behandelten Patienten.

Zwar sind, hauptsächlich in der älteren Literatur, einige Hinweise auf die Kompetenzen des Diagnostikers zu lesen, doch bleiben diese sehr allgemein: „Der kompetente Psychodiagnostiker ist sich der verschiedenen diagnostischen Perspektiven und ihrer konzeptionellen und methodologischen Herausforderung bewusst: Neben dem bislang diskutierten ‚Datenverarbeitungsmodell‘ und der ‚psycho-sozio-ökologischen Perspektive‘ des diagnostischen Prozesses gibt es noch andere diagnostische Modell-Perspektiven ...“ (Booth, 1995, S. 144). Fiedler (1984) schreibt über den Diagnostiker: „Anstatt mit formalen Methoden zu konkurrieren, sollte er seine Kräfte und seine Arbeitszeit für jene Probleme reservieren, die statt formaler Methoden den menschlichen Verstand benötigen, d.h. für die es weniger auf Präzision und absolute Reliabilität ankommt als auf Kreativität, Flexibilität, soziale Intelligenz, Improvisation, komplexe Mustererkennung und nicht zuletzt ‚Sprachgefühl‘“ (S. 309).

Die Kontroverse zwischen statistischer und „klinischer“ Urteilsbildung (Meehl, 1954; Wiggins, 1973) und die mögliche *Kombination* beider Strategien haben zeitweilig großes Interesse gefunden. Ein typischer Untersuchungsansatz war damals die statistische Auswertung bzw. die klinisch-diagnostische Interpretation von MMPI-Profilen im Vergleich zur „richtigen“ psychiatrischen Diagnose – ein Evaluationsverfahren, das heute auf *beiden* Seiten überholt ist. Diese Auseinandersetzung verlangt Differenzierungen, u.a. zwischen den statistisch evaluierbaren Prognosen und den diagnostischen Urteilen im Einzelfall, und vor allem eine gründlichere Evaluation der Kriterienvalidität, der Konstruktoperationalisierung und der Datenaggregationen. Die Kontroverse ist keineswegs befriedigend geklärt. Dennoch taucht dieses Thema in den Lehrbüchern nur noch am Rande auf, entweder als ältere Kontroverse in einem historischen Rückblick oder pauschal als der Gegensatz nomothetischer und idiografischer Auffassungen (Fisseni, 2004; Heil, 1995; Petermann, 1995), statt die Chancen der strategischen Kombination darzustellen. Diese Debatte müsste auf dem heutigen Stand fortgesetzt und an realistischen Lehrbeispielen vertieft werden (z.B. Dahle, 2005).

Deutet heute vielleicht der technisch klingende Begriff *Datenintegration* darauf hin, dass von vornherein Homogenes, sehr Ähnliches zu kombinieren ist wie bei einer mathematischen Funktion? Kann es auch auf die psychologische Interpretation tiefreichender Widersprüche von Befunden ankommen und um den psychologischen Kontext gehen? Konkret bleibt die heuristisch-beziehungsstiftende Aufgabe der Interpretation oft unerwähnt. Dementsprechend wird der Begriff Interpretation (auch in den Sachregistern erscheint höchstens die Interpretationsobjektivität) vermieden und damit das umfangreiche System von traditionellen Strategien und die Regeln der psychologischen Interpretation ausgeklammert, wenn nicht vergessen. Die Flussdiagramme und langen Checklisten deuten an, dass auch nach einer Debatte von ca. 40 Jahren die Hoffnung auf eine weitgehende Algorithmisierung und ein intelligentes Computerprogramm zur diagnostischen Urteilsbildung fortbesteht. Auch das Aufzählen von möglichen Fehlern und Verzerrungen im Prozess der diagnostischen Urteilsbildung (u.a. Westhoff & Kluck, 2003) demonstriert, wie wichtig methodenkritische Reflexionen sind.

Aus dem Nachweis von möglichen Denkfehlern der Urteilsbildung oder der „Entmystifizierung“ des Diagnostikers, folgt natürlich nicht, dass die traditionellen Strategien der Interpretation insgesamt falsch oder irrational waren. So hielten es die Herausgeber des neuen *Handbuchs der Psychologischen Methoden und Evaluation* nicht für wichtig, ein Kapitel über die Prinzipien Psychologischer Interpretation aufnehmen. Damit vertieft sich der Eindruck, dass ein Teil der Methodik aufgegeben wird, die bereits für Wilhelm Wundt (1921) und dessen psychologische Interpretationslehre wesentlich war.

Die einseitig kognitionspsychologisch-formale Perspektive der Informationsverarbeitung berücksichtigt zu wenig, dass ein persönlichkeits-theoretisches Bezugssystem vorhanden sein muss. Die diagnostische Urteilsbildung kann ja nicht

allein als *theoriefreie* Datenverarbeitung für eine bestimmte Aufgabenstellung ablaufen. Die Urteilsbildung des Diagnostikers findet unvermeidlich – ausgesprochen oder unausgesprochen – zugleich in einem persönlichkeits-theoretischen Bezugssystem statt. Dieser Interpretationsrahmen wird zu selten diskutiert als ob die diagnostischen Fragestellungen psychologisch isoliert zu beantworten wären. Wie könnten die berufsbezogenen und die pädagogischen Beurteilungen oder die ätiologischen und die therapeutischen Konzeptionen *ohne theoretischen Bezug* auf zugrundeliegende Annahmen-gefüge oder „Funktionsmodelle“ von Einstellungen, Motiven und Persönlichkeitseigenschaften getroffen werden? Die Implikationen solcher Vorentscheidungen für die psychologische Diagnostik sind noch kein Thema der Lehrbücher und wohl auch kaum der akademischen Lehre.

Die skizzierten Tendenzen wären dann als Fehlentwicklung anzusehen, wenn sie – in einer ohnehin reduzierten fachlichen Ausbildung – generell zu einer Unterbewertung der grundlegenden Prinzipien der differenziellen Psychologie und Persönlichkeitstheorie führen und zu einer Kürzung des gründlichen Trainings in psychologischer Diagnostik.

## 8.17 Strategische Konsequenzen – Ausblick von einer mittleren Position

Selbstberichte in Fragebogen und Interviews sind zweifellos geeignet – und unersetzlich – wenn die subjektive (mentale) Repräsentation des Erlebens, der Einstellungen und des Verhaltens erfasst werden sollen. Solche Selbstbeurteilungen sind am leichtesten zu erhalten, standardisiert und testökonomisch, sie haben eine Augenscheinvalidität. Wer auf Persönlichkeitsfragebogen verzichtet, verliert viele – auch durch ein langes Interview nur bedingt zu ersetzende – Informationen. Aber diese Selbstbeurteilungen können die Untersuchung des aktuellen, manifesten Verhaltens im Alltag nicht ersetzen, sondern die Verhaltensunterschiede nur zu erläutern helfen. Grundsätzlich kann zwischen verschiedenen Absichten und Strategien unterschieden werden.

### *Strategien und Evaluationen*

Cronbach (1970, S. 555f.) argumentiert, dass ein Persönlichkeitsfragebogen deskriptiv – wie ein Spiegel – dem Untersuchten hilft, von sich selbst ein Bild zu machen, auch im Vergleich zu anderen Menschen. Ein psychologischer Berater kann die Testwerte als Ausgangsinformationen für das weitere Kennenlernen verwenden und dabei Irrtümer der Testinterpretationen verringern. Inwieweit die Beschreibung aufgrund der Testwerte mit dem Selbstbild übereinstimmt, kann mittels anderer Informationen oder eventuell durch weitere Fragen, wie jemand eigentlich sein möchte, vertieft werden. Die Eigenschaftsbezeichnungen der Skalen zu verwenden oder das Testprofil zu zeigen, sei nicht ratsam.

Diese *explorative Strategie* eignet sich für einen biografisch-diagnostischen Ansatz in der psychologischen Beratung und Fallarbeit. Auch die gegenüber Messmodellen und statistischen Auswertungen skeptischen Psychologen werden einen Persönlichkeitsfragebogen als einen besonders strukturierten Teil einer breiteren psychologischen Untersuchung akzeptieren können. In diesem Zusammenhang könnten am ehesten auch die möglichen Antworttendenzen erkannt und in das Gesamtbild einbezogen werden (vgl. Cronbach, 1970).

Die Vorzüge der Persönlichkeitsfragebogen werden jedoch *strategisch* erst durch den *Vergleich* der individuellen Testprofile mit den Normwerten genutzt. Entsprechen die individuellen Eigenschaftsschilderungen der Durchschnittsbevölkerung oder weichen sie deutlich ab? Deshalb sind sehr große bevölkerungsrepräsentative Stichproben gerade für die Konstruktion und die differenzierte Normierung von Persönlichkeitsfragebogen unverzichtbar. (Außerdem ermöglichen große Stichproben die zur Absicherung von Forschungsarbeiten notwendige, simultane statistische Kontrolle einer Anzahl soziodemografischer Merkmale, wie es mit der Methode statistischer Zwillinge geschieht.) Die Wiederholung der repräsentativen Normierung des FPI-R hatte das überraschende Ergebnis, dass die Itemmuster und die Testwert-Verteilungen (Normen) nach 17 Jahren nur geringfügig voneinander abwichen. Die FPI-Skalen repräsentieren also psychologische Konstrukte, die offensichtlich in den Selbstbeschreibungen der Durchschnittsbevölkerung einen herausragenden und überdauernden Platz haben. Diese Konzepte überlappen sich, wie sich zeigen liess, deutlich mit den Konzepten von direkten Selbsteinstufungen und Fremdbeurteilungen. Es handelt sich um robuste Dimensionen eines differenziell-psychologischen Beschreibungssystems. Aus der Einsicht in die *strukturelle Subjektivität* dieser vielschichtigen Selbstbeurteilungen und Selbstberichte folgt noch nicht, dass alle Antworten wirklichkeitsfern sind, d.h. die gesamte oder ein überwiegender Teil der Varianz ohne Bezug zum manifesten und künftigen Verhalten ist. Die extremen sozialpsychologisch-konstruktivistischen Erklärungsversuche, dass es sich grundsätzlich *nur* um Stereotypen der Alltagspsychologie handele, reichen hier nicht aus. Es gibt systematische empirische Zusammenhänge zwischen den Testwerten von Persönlichkeitsinventaren und Statusmerkmalen soziodemografischer, klinischer und beruflicher Art, zu



anderen selbstberichteten Inhalten, zu Fremdeinstufungen und zu vielen empirisch prüfbareren Daten und Verhaltensweisen (Kapitel 6 und 7).

Ein standardisierter Fragebogen ist unentbehrlich, um solche Fragen zu beantworten, wie typisch in der Bevölkerung bestimmte Selbsteinschätzungen sind, z.B. der Lebenszufriedenheit, der Beanspruchung (des Stresserlebens), der körperlichen Beschwerden und Gesundheitssorgen, der Lebenszufriedenheit oder der Extraversion-Introversion, und wie ausgeprägt dieses Bild vergleichsweise bei einer einzelnen Person oder in Personengruppen ist. Zur psychologischen Diskrimination von klinischen und anderen Gruppen und zur Unterscheidung von Gesunden liegen wohl die meisten Untersuchungen vor. Die *Strategie des Screening* ist eine typische Anwendung eines Persönlichkeitsfragebogens, z.B. um in klinischen Einrichtungen auf testökonomische, und deshalb organisatorisch einfache Weise auf solche Personen aufmerksam zu werden, die in einem anschließenden Schritt genauer untersucht werden sollten. Ein solches Screening in einer mehrstufigen Strategie kann in vielen Institutionen nützlich sein, beispielsweise in der Psychosomatischen Klinik, in der Neurologischen Reha-Klinik oder in einer allgemeinen Rehabilitationseinrichtung für chronisch Kranke. Über die im Inventar erfassten Persönlichkeitseigenschaften hinaus können jedoch störungsspezifische diagnostische Aussagen nicht erwartet werden. Mehrere Untersuchungen zeigten in Katamnesen, dass der subjektiv erfahrene – und teils auch der medizinisch beurteilte – Erfolg mit den Testwerten der Erstuntersuchung, u.a. der Emotionalität, Beanspruchung, Neigung zu körperlichen Beschwerden und Lebenszufriedenheit deutlich zusammenhängt. – Die zahlreichen Befunde und auch einschränkende Feststellungen in sehr vielen Anwendungsfeldern lassen sich inzwischen nicht mehr in wenigen Aussagen bilanzieren. Die Testautoren versuchen mit den bei jeder Auflage eingearbeiteten ausführlichen Hinweisen zur fortgesetzten Evaluation beizutragen.

Die Grenzen der Fragebogenmethodik werden deutlicher, wenn nach den direkten Verhaltenskorrelationen gefragt wird. Wer auf die psychologischen Informationen aus Persönlichkeitsfragebogen nicht verzichten will, muss multimodal vorgehen und dabei die strukturelle Subjektivität von Selbstberichten und Selbstbeurteilungen akzeptieren.

Selbstbeurteilungen und verbale Auskünfte über eigenes Verhalten sind eben keine objektiven Verhaltensdaten. Die deutlich begrenzten und z.T. widersprüchlichen Resultate im Hinblick auf objektivierbare Kriterien sind zwar seit langer Zeit bekannt, scheinen allerdings in manchen Anwendungssituationen unterschätzt oder ausgeklammert zu werden. Die mono-methodisch angewendeten Fragebogen sind für einige psychologische Fragestellungen zweckmäßig, sollten aber sonst nur ein Bestandteil einer multimodalen Diagnostik sein, um krasse Fehlbeurteilungen zu vermeiden. – Diese Beschränkungen bewusst zu halten, bleibt wichtig, um zur vorsichtigen Interpretation zu motivieren. Weder eine unkritische behaviorale Interpretation noch die ausschließlich selbsttheoretische oder alltagspsychologische Deutung treffen die Eigenart von Persönlichkeitsfragebogen.

*Strategische Schwierigkeiten* grundsätzlicher Art bestehen in der Persönlichkeitsforschung, die sich ausschließlich auf Fragebogen stützt, wie es noch häufig der Fall ist. Offensichtlich steht die Persönlichkeitsforschung weiterhin vor der von Fiske erkannten konzeptuellen und methodischen Herausforderung, anstelle der globalen Eigenschaftskonstrukte wesentlich kleinere Beschreibungseinheiten zu entwickeln. Das eigene Arbeitsprogramm zur psychophysiologischen Persönlichkeitsforschung, dessen Ergebnisse in diesem Manual häufiger zitiert sind (u.a. Abschnitt 7.3), hat die Divergenz zwischen Selbstbeurteilungen der emotionalen Labilität und Messungen der vegetativen Labilität aufgezeigt statt der erwarteten Korrelate im Sinne Eysencks. Auch die neue Methodik des ambulanten Assessment führt zu der Erfahrung, wie häufig Diskrepanzen zwischen den Beschreibungsebenen von Emotionen, körperlicher Aktivität, Symptomen und Verhaltensstörungen sind. Diese Befunde sind jedoch keineswegs spezifisch für die Persönlichkeitsfragebogen oder die Einstufungen des Befindens und Erlebens. Sie sind typisch für eine psychologische Diagnostik, die subjektiv-verbale, behaviorale und physiologische Methoden kombiniert. Die notorischen Diskrepanzen sind aus der behavioralen und psychophysiologischen Diagnostik chronischer Angststörungen und Phobien seit langem bekannt. Gelegentlich wird vorgeschlagen, in der psychologischen Diagnostik nicht uni-methodisch vorzugehen, sondern multimethodisch verschiedene Verfahren zur Absicherung zu verwenden. Diese Empfehlung müsste noch entschiedener lauten: nicht *multi-methodisch* (z.B. mit einer Sammlung verschiedenster Fragebogen), sondern *multi-modal* auf kategorial verschiedenen Datenebenen. Die *multimodale Strategie* ist jedoch gegenwärtig weder testtheoretisch noch in hinsichtlich praktisch-diagnostischer Standards hinreichend ausgearbeitet (vgl. Abschnitte 8.10 und 8.11).

Die Fragebogenantworten konsequent als Selbstbeurteilungen anzusehen, ändert nichts an der Möglichkeit, solche Aussagen sinnvoll mit den Daten von Bezugsgruppen und Normen sowie Verhaltensinformationen zu vergleichen, verlangt jedoch eine andere Interpretationsweise und – wo immer möglich – eine *multimodale Strategie*, die möglichst

viele Kontextinformationen einzubeziehen versucht. Eine mittlere Position einzunehmen, bedeutet auch, bei pauschaler Kritik an dieser Methodik nach der wissenschaftlichen Qualität der Argumente zu fragen und vor allem danach: Welche Alternative für die Persönlichkeitsforschung und Persönlichkeitsdiagnostik schlagen die Kritiker beim Verzicht auf die Fragebogenmethodik vor?

#### *Zur Testkonstruktion*

Die allgemeinen Aspekte der *Testgüte* des Freiburger Persönlichkeitsinventars FPI-R werden in dem abschließenden Kapitel 9 zusammengefasst. Hinsichtlich der *Testkonstruktion* ist eine über neue Normierungen hinausgehende „Pflege“ in der Hauptsache nicht möglich, weil jede eingreifende Revision, die Kontinuität und die vorhandenen Vergleichsmöglichkeiten, insbesondere innerhalb von Institutionen und Längsschnittstudien, zerstören würde. Darüber hinaus bestehen die ausführlich begründeten Zweifel, ob die propagierten Latent Trait-Modelle für die facettenreichen und deshalb nicht perfekt „homogenisierbaren“ Persönlichkeitskonstrukte gerechtfertigt sind. Aus einer mittleren Position, welche die Argumente der Kritiker und der Befürworter aufzunehmen sucht, werden solche Konzepte nicht grundsätzlich abgelehnt, sondern als *heuristische Strategien der Testanalyse* angesehen. Künftig werden die Latent Class- und mixed IR-Modelle und fortgeschrittene Verfahren der Musteranalyse für nominale Daten wahrscheinlich eine größere Bedeutung bekommen (vgl. Abschnitt 8.8). Solche Analysen könnten u.U. zu überzeugenderen Strukturaussagen und entsprechenden Item-Aggregaten in den Persönlichkeitsfragebogen führen. In diesem Zusammenhang wäre auch das Konzept der Indexmessung im Sinne von Dawes (1977) aufzunehmen und in pragmatischen Assessmentstrategien weiterzuführen.

Die allgemeine Kritik an Fragebogen dauert, seit Wilhelm Wundt, nun mehr als 100 Jahre an. Dabei rückt das Thema der Antworttendenzen in Persönlichkeitsfragebogen oft so sehr in den Vordergrund der Methodenkritik, dass andere Perspektiven zu kurz kommen. Die möglichen Effekte formaler Antworttendenzen oder der sozialen Erwünschtheit dürfen nicht verharmlost werden. Es ist jedoch immer deutlicher geworden, dass die operationale Definition und Differenzierung solcher Tendenzen oder die Abgrenzung von den typischen Persönlichkeitsmerkmalen nicht erreicht wurde. Abgesehen von der Anwendung von Persönlichkeitsfragebogen zum Zweck eines ersten Screenings, gehören deshalb zur Interpretation eines Testprofils der psychologische Kontext und eine *zu trainierende Kompetenz in der psychologischen Interpretation* von konvergenten und von divergenten diagnostischen Informationen (Abschnitt 8.16).

#### *Ausblick*

Strategisch folgt aus diesen Überlegungen, primär an der Umsetzung der methodischen Fortschritte in dieser Richtung zu arbeiten, u.a. unter den Stichwörtern: Generalisierbarkeit und ökologische Validität, repräsentative, symmetrische Validierungsforschung und multimodale Diagnostik. Wichtiger als konventionelle Gruppenvergleiche, Korrelationen mit anderen Persönlichkeitsfragebogen oder sonstigen Informationen sind *anspruchsvollere Assessmentstrategien* und die Prüfung des *Entscheidungsnutzens* solcher Persönlichkeitsdaten im Hinblick auf praktisch relevante Kriterien. Dazu sollten auch überzeugende Replikationen wichtiger Befunde mit relativ großer Personenzahl zunächst innerhalb und dann zwischen den Arbeitsgruppen gehören. Da solche systematischen Replikationen in der Psychologie (im Unterschied zu den Naturwissenschaften) weithin fehlen bzw. als nicht besonders wichtig oder als nicht kreativ zu gelten scheinen, ist ein Seitenblick auf die Medizin angebracht. Hier führen die international und multizentrisch durchgeführten Studien, z.B. über bereits eingeführte Medikamente, oft zu überraschenden Ergebnissen und gelegentlich sogar zur dramatischen Beendigung solcher Studien. Weshalb sollte die Evaluation psychologischer Prädiktoren und Interventionen einfacher sein?

Überzeugende Validierungsstudien haben eine höhere Priorität als die Fokussierung auf Messmodelle oder die Person--Situation-Debatte anhand dafür ungeeigneter Fragebogendaten oder die essentialistischen Behauptungen über die richtige Anzahl von hauptsächlichen Persönlichkeitsfaktoren.

Die eigenen Bemühungen und problematischen Erfahrungen führten zu der Schlussfolgerung und dem wiederholten Plädoyer, dass inhaltlich und statistisch überzeugende Evaluationen nur *innerhalb* großer Institutionen zweckmäßig und aussichtsreich sind. Nur dort sind auch die notwendigen follow-up Informationen bzw. Katamnesen zu gewinnen, die zur Beurteilung des Entscheidungsnutzens notwendig sind. Die bisherigen Untersuchungen sind in ihrer Überzeugungskraft beeinträchtigt, da solche Informationen in der Regel fehlen und deswegen neben der Nutzenfunktion die zugehörige Schadensfunktion der diagnostischen Entscheidungen bzw. des Screenings unbekannt bleibt.

Die Absicht der Qualitätssicherung stößt hier an Grenzen, die gewöhnlich nicht erörtert werden. Falls einige Institutio-

nen, Kliniken, Rehabilitationseinrichtungen, Organisationen, Betriebe, staatliche Verwaltungen tatsächlich Persönlichkeitsdaten und Persönlichkeitsfragebogen *intern* evaluieren, was anzunehmen ist, bleiben die Ergebnisse in der Regel unzugänglich. Dies kann mit dem Datenschutz zusammenhängen, obwohl es geeignete Wege der Anonymisierung gibt, oder mit der geringen Bereitschaft, solche Erfahrungen weiterzugeben. – Ob das zunehmende Bewusstsein von Qualitätssicherung diese Grenzen im gemeinsamen Interesse überwinden wird?

In Anlehnung an den ersten Kriterienkatalog des *Testkuratoriums* (1986a; siehe auch Häcker, Leutner & Amelang, 1998) werden hier die wichtigsten Aspekte der Testgüte des Freiburger Persönlichkeitsinventars aus der Sicht der Testautoren kommentiert. Diese Darstellung folgt noch nicht dem neueren *Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen* TBS-TK im Zusammenhang mit DIN 33430 (Testkuratorium 2007; Westhoff et al., 2004), denn diese Richtlinien sind gegenwärtig primär auf die Eignungsdiagnostik und den Berufserfolg ausgerichtet. Zwar sind viele allgemeingültige Gesichtspunkte enthalten, doch fehlt noch eine adäquate Erweiterung für Persönlichkeitsdiagnostik, für Stimmungs-/Emotions-Skalen sowie für Klinische Skalen.

## 10 Literaturverzeichnis

- Albani, C., Blaser, G., Rietz, U., Villmann, T. & Geyer, M. (2006). Die geschlechtsspezifische Erfassung körperlicher Beschwerden bei Psychotherapiepatientinnen mit dem „Gießener Beschwerdebogen“ (GBB) von Brähler & Scheer. In A. Hinz & O. Decker (Hrsg.), *Gesundheit im gesellschaftlichen Wandel. Altersspezifik und Geschlechterrollen. Festschrift zum 60. Geburtstag von Elmar Brähler* (S. 124–143). Gießen: Psychosozial-Verlag.
- Allbutt, J., Ling, J., Heffernan, T.M. & Shafiullah, M. (2008). Self-Report Imagery Questionnaire scores and subtypes of social-desirable responding auditory imagery, visual imagery, and thinking style. *Journal of Individual Differences*, 29 (4), 181–188.
- Amelang, M. (1987). Fragebogen-Tests und experimentalpsychologische Variablen als Korrelate der Persönlichkeitsdimensionen Extraversion/Introversion (E/I) und Neurotizismus (N). In M. Amelang (Hrsg.), *Bericht über den 35. Kongreß der Deutschen Gesellschaft für Psychologie in Heidelberg 1986* (S. 403–416). Göttingen: Hogrefe.
- Amelang, M. (2006). Testinformation. Becker, P. (2003). Trierer Integriertes Persönlichkeitsinventar TIPI. *Diagnostica*, 52 (3), 154–157.
- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. Aufl.). Stuttgart: Kohlhammer.
- Amelang, M. & Borkenau, P. (1981). Untersuchungen zur Validität von Kontroll-Skalen für Soziale Erwünschtheit und Akquieszenz. *Diagnostica*, 27, 295–312.
- Amelang, M., Bartussek, D., Stemmler, G. & Hagemann, D. (2006). *Differentielle Psychologie und Persönlichkeitsforschung* (6. Aufl.). Stuttgart: Kohlhammer.
- Amelang, M. & Borkenau, P. (1982). Über die faktorielle Struktur und externe Validität einiger Fragebogen-Skalen zur Erfassung von Dimensionen der Extraversion und emotionalen Labilität. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 119–145.
- Amelang, M., Breit, C. & Schneyer, T. (1985). Bearbeitungszeiten bei Fragebogenitems und Ratingskalen verschiedener Dimensionszugehörigkeit. *Psychologische Beiträge*, 27, 318–331.
- Amelang, M., Schäfer, A. & Yousfi, S. (2002). Comparing verbal and non-verbal personality scales: Investigating the reliability and validity, the influence of social desirability, and the effect of fake good instructions. *Psychologische Beiträge*, 44, 24–41.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention* (4. Aufl.). Heidelberg: Springer.
- Amelang, M. & Ullwer, U. (1990). Untersuchungen zur experimentellen Bewährung von Eysencks Extraversionstheorie. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 11, 127–148.
- Amelang, M. & Ullwer, U. (1991). Correlations between psychometric measures and psychophysiological as well as experimental variables in studies on extraversion and neuroticism. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament* (pp. 287–315). New York: Plenum Press.
- Amelang, M. & Zielinski, W. (1994). *Psychologische Diagnostik und Intervention*. Berlin: Springer.
- Andresen, B. (1986). *Reanalysen am FPI-R*. Unveröff. Forschungsbericht. Psychiatrische und Nervenklinik, Universität Hamburg.

- Andresen, B. (1987). *Differentielle Psychophysiologie valenzkonträrer Aktivierungsdimensionen*. Diss. Frankfurt a.M.: Lang.
- Andresen, B. & Beauducel, A. (2008). TBS-TK Rezension: NEO-Persönlichkeitsinventar nach Costa und McCrae, revidierte Fassung (NEO-PI-R). *Report Psychologie*, 11/12, 543–544.
- Andresen, B. & Stemmler, G. (1982). Eine Dimensionsanalyse von 61 Persönlichkeitsskalen unter besonderer Berücksichtigung des Extraversions- und Emotionalitätskonzepts. *Diagnostica*, 28, 340–347.
- Angleitner, A. (1976). *Methodische und theoretische Probleme bei Persönlichkeitsfragebogen unter besonderer Berücksichtigung neuerer deutschsprachiger Fragebogen*. Unveröffentl. Habilitationsschrift. Universität Bonn.
- Angleitner, A. & Wiggins, J.S. (Eds.). (1986). *Personality assessment via questionnaire. Current issues in theory and measurement*. Berlin: Springer.
- Angst, J. & Clayton, P. (1986). Premorbid personality of depressive, bipolar, and schizophrenic patients with special reference to suicidal issues. *Comprehensive Psychiatry*, 27, 511–532.
- Angst, L. & Clayton, P. (1998). Personality, smoking and suicide: A prospective study. *Journal of Affective Disorders*, 51, 55–62.
- Angst, J., Degonda, M. & Ernst, C. (1992). The Zurich Study: XV. Suicide attempts in a cohort from age 20 to 30. *European Archives of Psychiatry & Clinical Neuroscience*, 242, 135–141.
- Angst, J. & Wicki, W. (1991). The Zurich Study. XI. Is dysthymia a separate form of depression? Results from the Zurich Cohort Study. *European Archives of Psychiatry and Clinical Neuroscience*, 240, 349–354.
- Asendorpf, J.B. (2007). *Psychologie der Persönlichkeit* (5. Aufl.). Berlin: Springer.
- Assmann, B., Dähne, A., Hinz, A. & Ettrich, C. (2003). Essstörungsspezifische Psychodiagnostik bei Anorexia nervosa und Bulimia nervosa. *Suchtmedizin in Forschung und Praxis*, 5 (3), 185–191.
- Austin, E.J., Deary, I.J., Gibson, G.J., McGregor, M.J. & Dent, J.B. (2006). Individual differences in response to scale use: Mixed Rasch modelling of response to NEO-FFI items. *Personality and Individual Differences*, 40, 1235–1245.
- Bacher, J. (2002). Statistisches Matching: Anwendungsmöglichkeiten, Verfahren und ihre praktische Umsetzung in SPSS. *Z4-Informationen (Zentralstelle für Empirische Sozialforschung an der Universität zu Köln)*, 51, 38–66.  
[http://www.soz.jku.at/Portale/Institute/SOWI\\_Institute/Soziologie/aes/content/e50/e1512/e2128/StatisticalMatching\\_ger.pdf](http://www.soz.jku.at/Portale/Institute/SOWI_Institute/Soziologie/aes/content/e50/e1512/e2128/StatisticalMatching_ger.pdf)
- Baltissen, R. & Boucsein, W. (2005). Vegetatives System und Persönlichkeit. In J. Hennig, & P. Netter (Hrsg.), *Biopsychologische Grundlagen der Persönlichkeit* (S. 397–510). Heidelberg: Spektrum.
- Bartussek, D. (1996). Faktorenanalytische Gesamtsysteme der Persönlichkeit. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie. Differentielle Psychologie und Persönlichkeitsforschung. Bd. 3. Temperaments- und Persönlichkeitsunterschiede* (S. 51–105). Göttingen: Hogrefe.
- Baumann, U. & Stieglitz, R.D. (2008). Multimodale Diagnostik – 30 Jahre später. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 56, 191–202.
- Baumann, U., Thiele, C. & Laireiter, A.R. (2003). Felddiagnostik (ambulantes Assessment) – insbesondere mittels computerunterstützter Verfahren – als Methode der psychopathologischen Verlaufsforschung. In M. Soyka, H.-J. Möller & H.U. Wittchen (Hrsg.), *Psychopathologie im Längsschnitt. Methoden, Analyse, Bewertung* (S. 65–87). Landsberg/Lech: Ecomed.
- Baumeister, R.F., Vohs, K.D. & Funder, D.C. (2007). Psychology as the science of selfreport and finger movements. Whatever happened to actual behavior? *Perspectives on Psychological Science*, 2, 396–403.
- Beauducel, A., Biehl, B., Bosnjak, M., Conrad, W., Schönberger, G. & Wagener, D. (Hrsg.). (2005). *Multivariate research strategies. Festschrift in Honor of Werner W. Wittmann*. Aachen: Shaker.
- Becker, P. (1988). Seelische Gesundheit und Verhaltenskontrolle: Zwei replizierbare, varianzstarke Persönlichkeitsfaktoren. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 9, 13–38.
- Becker, P. (1989). *Der Trierer Persönlichkeitsfragebogen TPF. Handanweisung*. Göttingen: Hogrefe.
- Becker, P. (1996). Wie big sind die Big Five? *Zeitschrift für Differentielle und Diagnostische Psychologie*, 17, 209–221.
- Becker, P. (2000). Die „Big Two“. Seelische Gesundheit und Verhaltenskontrolle: zwei orthogonale Superfaktoren höherer Ordnung? *Zeitschrift für Differentielle und Diagnostische Psychologie*, 21, 113–124.
- Becker, P. (2003a). *Trierer Integriertes Persönlichkeitsinventar TIPI*. Göttingen: Hogrefe.
- Becker, P. (2003b). Persönlichkeitsfragebogen. In K.D. Kubinger & S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 332–337). Weinheim: Beltz.

- Beckmann, D., Brähler, E. & Richter, H.E. (1977). Neustandardisierung des Gießen-Test (GT). *Diagnostica*, 23, 287–297.
- Bengel, J. & Wittmann, W.W. (1983). Urteilsfehler aufgrund zeitlicher Veränderungen von Testnormen. *Diagnostica*, 29, 101–117.
- Bilke, S. (2002). Freiburger Persönlichkeitsinventar (FPI-R). In U.-P. Kanning & H. Holling (Hrsg.), *Handbuch personaldiagnostischer Instrumente* (S. 356–401). Göttingen: Hogrefe.
- Bilsky, W. (1980). *Fragebogen zu Altruismus, Hilfsbereitschaft und Sozialer Verantwortung. Untersuchung ihrer faktoriellen Struktur und Aufweis korrelativer Zusammenhänge*. Braunschweiger Psychologische Arbeiten, Nr. 2.
- Bilsky, W. & Hosser, D. (1998). Soziale Unterstützung und Einsamkeit: Psychometrischer Vergleich zweier Skalen auf der Basis einer bundesweiten Repräsentativbefragung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19, 130–144.
- Bilsky, W. & Schwartz, S.H. (1994). Values and personality. *European Journal of Personality*, 8, 163–181.
- Bishop, G.D., Ngau, F. & Pek, J. (2008). Domain-specific assessment of anger expression and ambulatory blood pressure. *Personality and Individual Differences*, 44 (8), 1726–1737.
- Bland, J.M. & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 8 (1), 307–310.
- Böhme, H. (1997). Abwehr in Rating-Skalen und Therapieerfolg. *Nervenarzt*, 68, 896–902.
- Booth, J.F. (1995). Kompetenz. In S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik*. (3. Aufl.). (S. 138–146). Weinheim: Beltz, Psychologie-Verlags-Union.
- Borg, I. & Staufenbiel, T. (2007). *Lehrbuch Theorien und Methoden der Skalierung* (4. Aufl.). Bern: Huber.
- Borkenau, P. (2006). Selbstbericht. In F. Petermann & M. Eid, M. (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 135–143). Göttingen: Hogrefe.
- Borkenau, P. & Amelang, M. (1986). Zur faktorenanalytischen Kontrolle Sozialer Erwünschtheitstendenzen. Eine Untersuchung anhand des Freiburger Persönlichkeits-Inventars. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 7, 17–28.
- Borkenau, P. & Ostendorf, F. (1989). Untersuchungen zum Fünf-Faktoren-Modell der Persönlichkeit und seiner diagnostischen Erfassung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 10, 239–251.
- Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar nach Costa und McCrae. Handanweisung*. Göttingen: Hogrefe.
- Bortz, J., Lienert, G.A. & Boehncke, K. (2000). *Verteilungsfreie Methoden in der Biostatistik* (2. Aufl.). Berlin: Springer.
- Bottenberg, E.H. & Keller, J.A. (1975). Beitrag zur empirischen Erfassung von Selbst-Aktualisierung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 23, 21–53.
- Brähler, E. & Richter, H.-E. (1995). Deutsche Befindlichkeiten im Ost-West-Vergleich. In H.J. Wirth (Hrsg.), *Ossis und Wessis: Psychogramm deutscher Befindlichkeiten* (Psychosozial Bd. 59, S. 7–20). Gießen: Psychosozial-Verlag.
- Brähler, E., Schumacher, J. & Brähler, C. (1999). Erste gesamtdeutsche Normierung und spezifische Validitätsaspekte des Gießen-Test. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 20, 231–243.
- Brambring, M. (1983). Spezielle Eignungsdiagnostik. In K.J. Groffmann & L. Michel (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich B. Methodologie und Methoden. Serie II. Psychologische Diagnostik. Band 2 Intelligenz- und Leistungsdiagnostik* (S. 414–481). Göttingen: Hogrefe.
- Brielmaier, I. (1995). *Bewirkt der Computer einen Unterschied? Eine experimentelle Studie von Papier-Bleistift- und Computerversion des FPI-R*. Unveröffentlichte Diplomarbeit. Fachbereich Psychologie. Universität Marburg.
- Brodner, G. (1983). *Beiträge zur Prognose des Rehabilitationserfolgs bei „jugendlichen Herzinfarktpatienten“*. Multivariate Analyse psychophysiologischer Befunde. München: Minerva.
- Brügner, G. (1988). MONITOR: Ein flexibles Programm zur Datenerhebung mittels Pocket-PC. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19, 145–147.
- Brunsch, M. (1987). *Über Clusteranalysen am FPI-R*. Unveröff. Arbeitsbericht. Psychologisches Institut, Universität Freiburg.
- Brunswik, E. (1956). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193–217.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Aufl.). München: Pearson Studium.
- Burisch, M. (1976). *Konstruktionsstrategien für multi-dimensionale Persönlichkeitsfragebogen*. Unveröff. Diss. Universität Hamburg.
- Burisch, M. (1984a). Approaches to personality inventory construction. A comparison of merits. *American Psychologist*, 39, 214–

- Burisch, M. (1984b). You don't always get what you pay for: Measuring depression with short and simple versus long and sophisticated scales. *Journal of Research in Personality*, 18, 81–98.
- Burisch, M. (1997). Testlänge und Testvalidität. *European Journal of Personality*, 11, 303–315.
- Buse, L. (1976). Zur Interpretation einer Lügenskala. *Diagnostica*, 22, 34–43.
- Buse, L. (1980). Kritik am Moderatoransatz in der Akquieszenz-Forschung. *Psychologische Beiträge*, 22, 119–127.
- Buse, L. & Pawlik, K. (1984). Inter-Setting-Korrelationen und Setting-Persönlichkeit-Wechselwirkungen: Ergebnisse einer Felduntersuchung zur Konsistenz von Verhalten und Erleben. *Zeitschrift für Sozialpsychologie*, 15, 44–59.
- Buse, L. & Pawlik, K. (1994). Differenzierung zwischen Tages-, Setting- und Situationskonsistenz ausgewählter Verhaltensmerkmale, Maßen der Aktivierung, des Befindens und der Stimmung in Alltagssituationen. *Diagnostica*, 40, 2–26.
- Bussmann, J.B.J., Ebner-Priemer, U.W. & Fahrenberg, J. (2009). Ambulatory behavior monitoring: Progress in measurement of activity, posture, and specific motion patterns in daily life. *European Psychologist*, 14, 142–152.
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Cattell, R.B. (1957). *Personality and motivation. Structure and measurement*. New York: World Book.
- Cattell, R.B. (1973). *Personality and mood by questionnaire*. San Francisco: Jossey-Bass.
- Cattell, R.B. & Warburton, F.W. (1967). *Objective personality and motivation tests: a theoretical introduction and practical compendium*. Urbana: University of Illinois Press.
- Cima, M., Hollnack, S., Kremer, K., Knauer, E., Schellbach-Matties, R., Klein, B. & Merckelbach, H. (2003). Strukturierter Fragebogen Simulierter Symptome. Die deutsche Version des „Structured Inventory of Malingered Symptomatology: SIMS“. *Der Nervenarzt*, 74, 977–986.
- Cliff, N. (1977). Further study of cognitive processing models for inventory response. *Applied Psychological Measurement*, 1, 41–49.
- Cloninger, C.R., Przybeck, T.R., Svrakic, D.M. & Wetzel, R.D. (2000). *Das Temperament- und Charakter-Inventar TCI*. Frankfurt a.M.: Harcourt Swets Services.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, New Jersey: Lawrence Erlbaum.
- Costa, P.T. & McCrae, R.R. (1992). Four ways five factors are basic. *Personality & Individual Differences*, 13, 653–665.
- Cronbach, L.J. (1970). *Essentials of psychological testing* (3<sup>rd</sup> ed.). New York: Harper and Row.
- Cronbach, L.J., Gleser, G.C., Nanda, H., Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Dahle, K.P. (2005). *Psychologische Kriminalprognose. Wege zu einer integrativen Methodik für die Beurteilung der Rückfallwahrscheinlichkeit bei Strafgefangenen*. Herbolzheim: Centaurus.
- Daum, I. & Reitz, E. (1998). Personality correlates of delinquent behaviour in juveniles and adolescents: evidence from German studies. In D.F. Saklofski & S.B.G. Eysenck (Eds.), *Individual differences in children and adolescents* (pp. 38–49). New Brunswick, NJ, USA: Transaction Publishers.
- Davier, M. von (2005). *A general diagnostic model applied to language testing data* (Research Report No. RR-05–16). Princeton, NJ: Educational Testing Service.
- Dawes, R.M. (1977). Grundlagen der Einstellungsmessung. Weinheim: Beltz.
- Dawes, R.M., Faust, D. & Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668–1674.
- DeBoeck, P. (1978). Validity of a cognitive processing model for responses to adjective and sentence type inventories. *Applied Psychological Measurement*, 2, 371–378.
- Deinzer, R., Steyer, R., Eid, M. & Notz, P. (1995). Situational effects in trait assessment: The FPI, NEOFFI, and EPI questionnaires. *European Journal of Personality*, 9, 1–23.
- Dembroski, T.M., Weiss, S.M., Shields, J.L., Haynes, S.G. & Feinleib, M. (1978). *Coronaryprone behavior*. New York: Springer.
- Demisch, L., Georgi, K., Pathke, B., Demisch, K. & Bochnik, H.J. (1982). Correlation of platelet MAO activity with introversion: A study on a German rural population. *Psychiatry Research*, 6, 303–311.

- Deneke, F.W., Hilgenstock, B. & Müller, R. (1989). *Das Narzismusinventar. Handbuch*. Bern: Huber.
- Der Spiegel (1995). *Alte Hütte*. Heft 35. S. 166–168.
- Deter, H.-C. & Schepank, H. (1991). Patterns of self-definition of asthma patients and normal persons in the Freiburg Personality Inventory. *Psychotherapy and Psychosomatics*, 55, 47–56.
- Dickenberger, D., Holtz, S. & Gniech, G. (1978). Bedürfnis nach sozialer Anerkennung: Validierung der „Marlowe-Crowne Social Desirability Scale“ über ein Konzept individuell relevanter Gruppen. *Diagnostica*, 24, 24–38.
- Digman, J.M. & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50, 116–123.
- Dilchert, S., Ones, D.S., Viswesvaran, C. & Deller, J. (2006). Response distortion in personality measurement: Born to deceive, yet capable of providing valid self-assessments? *Psychology Science*, 48 (3), 209–225.
- Driessen, M. & Balck, F. (1991). Chronische Niereninsuffizienz: Prädiktoren für eine günstige Adaptation an Krankheit und Behandlung. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 41, 362–371.
- Eberwein, M. (2008). Verzeichnis psychologischer und pädagogischer Testverfahren aus dem Datenbanksegment PSYNDEX Tests (15. Aufl.). Trier: Zentrum für Psychologische Information und Dokumentation (ZPID). Universität Trier. Zugriff am 1.3.2009. [http://www.zpid.de/pub/tests/verz\\_teil5.pdf](http://www.zpid.de/pub/tests/verz_teil5.pdf) [pdf 858 KB].
- Ehlers, T. (1965). Über persönlichkeitsbedingte Unfallgefährdung. *Archiv für die gesamte Psychologie*, 117, 252–279.
- Eid, M. & E. Diener, E. (Eds.). (2005). *Handbook of multimethod measurement in psychology*. Washington, D.C.: American Psychological Association.
- Eid, M., Nussbeck, F.W. & Lischetzke, T. (2006). Multitrait-Multimethod-Analyse. In F. Petermann & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 332–345). Göttingen: Hogrefe.
- Endler, N.S. & Hunt, J. McV. (1969). Generalizability of contributions from sources of variance in the S-R-Inventories of anxiousness. *Journal of Personality*, 37, 1–24.
- Epstein, S. (1977). Traits are alive and well. In D. Magnusson & N.S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 83–98). Hillsdale, N.J.: Erlbaum.
- Epstein, S. (1980). The stability of behavior: II. Implications for psychological research. *American Psychologist*, 35, 790–806.
- Ertel, S. (1970). Selbstbeurteilung, Semantik und Persönlichkeit. *Psychologische Forschung*, 33, 254–276.
- Esser, A. (1991). *Rückfallbedingungen bei Alkoholikern nach stationärer Entwöhnungsbehandlung*. Unveröff. Diss. Universität Marburg.
- Eysenck, H.J. (1967). *The biological basis of personality*. Springfield IL: Thomas.
- Eysenck, H.J. (1977). Personality and factor analysis: A reply to Guilford. *Psychological Bulletin*, 84, 405–411.
- Eysenck, H.J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667–673.
- Eysenck, H.J. & Eysenck, M.W. (1985). *Personality and individual differences*. New York: Plenum.
- Eysenck, H.J. & Eysenck, S.B.G. (1969). *Personality structure and measurement*. London: Routledge and Kegan Paul.
- Eysenck, H.J. & Eysenck, S.B.G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder & Stoughton.
- Eysenck, H.J. & Eysenck, S.B.G. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.
- Eysenck, M.W. & Eysenck, H.J. (1980). Mischel and the concept of personality. *British Journal of Psychology*, 71, 191–204.
- Eysenck, S.B.G. & Eysenck, H.J. (1977). The place of impulsiveness in a dimensional system of personality description. *British Journal of Social and Clinical Psychology*, 16, 57–68.
- Fahrenberg, J. (1964). Objektive Tests zur Messung der Persönlichkeit. In Heiß, R. (Hrsg.): *Psychologische Diagnostik. Handbuch der Psychologie*. Band 6 (S. 488–532). Göttingen: Hogrefe.
- Fahrenberg, J. (1975). Die Freiburger Beschwerdenliste FBL. *Zeitschrift für Klinische Psychologie*, 4, 79–100.
- Fahrenberg, J. (1977). Physiological concepts in personality research. In R.B. Cattell & R.M. Dreger (Eds.), *Handbook of modern personality theory* (pp. 585–611). Washington, D.C.: Hemisphere.
- Fahrenberg, J. (1982). *Probleme der Mehrebenen-Beschreibung und Prozeß-Forschung*. Forschungsberichte des Psychologischen Instituts der Universität Freiburg, Nr. 8.

- Fahrenberg, J. (1987a). Concepts of activation and arousal in the theory of emotionality (neuroticism): A multivariate conceptualization. In J. Strelau & H.J. Eysenck (Eds.), *Personality dimensions and arousal* (pp. 99–120). New York: Plenum.
- Fahrenberg, J. (Hrsg.). (1987b). Multimodale Diagnostik [Themenheft]. *Diagnostica*, 33 (4).
- Fahrenberg, J. (1991). Differential psychophysiology and the diagnosis of temperament. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament* (pp. 317–333). London: Plenum Press.
- Fahrenberg, J. (1992). Psychophysiology of neuroticism and anxiety. In A. Gale & M.W. Eysenck (Eds.), *Handbook of individual differences: Biological perspectives* (pp. 179–226). Chichester: Wiley.
- Fahrenberg, J. (1994a). *Die Freiburger Beschwerdenliste FBL. Form FBL-G und revidierte Form FBL-R*. Handanweisung. Göttingen: Hogrefe.
- Fahrenberg, J. (1994b). Ambulantes Assessment. *Diagnostica*, 40, 195–216.
- Fahrenberg, J. (1995). Biopsychologische Unterschiede. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie. Serie Differentielle Psychologie und Persönlichkeitsforschung. Band 2. Verhaltens- und Leistungsunterschiede* (S. 139–193). Göttingen: Hogrefe.
- Fahrenberg, J. (1997). Replik zur Rezension des FPI. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, 85–86.
- Fahrenberg, J. (2002). *Psychologische Interpretation. Biographien – Texte – Tests*. Bern: Huber.
- Fahrenberg, J. (2005). Representative design and the laboratory field-issue. In A. Beauducel, B. Biehl, M. Bosniak, W. Conrad, G. Schönberger & D. Wagener (Eds.), *Multivariate research strategies. Festschrift in Honor of Werner W. Wittmann* (S. 237–260). Aachen: Shaker.
- Fahrenberg, J. (2006a). Assessment in daily life. A Review of Computer-assisted Methodologies and Applications in Psychology and Psychophysiology, years 2000–2005. Zugriff am 1.3.2009. <http://www.ambulatory-assessment.org/> [92 pages, 447 references].
- Fahrenberg, J. (2006b). Ambulantes Assessment von Befinden, Stimmungen, Emotionen – Zur Methodik von Selbsteinstufungen (Selbstberichten). Zugriff 1.3.2009. <http://www.jochen-fahrenberg.de/> [pdf 39 Seiten, 134 Literaturhinweise, 764 KB].
- Fahrenberg, J. (2008a). Die Wissenschaftskonzeption der Psychologie bei Kant und Wundt. In e-Journal Philosophie der Psychologie -Online10, 1–22. Zugriff am 1.3.2009. [http://www.jp.philo.at/texte/Fahrenberg\\_J2.pdf](http://www.jp.philo.at/texte/Fahrenberg_J2.pdf) [35 Seiten, 115 Literaturhinweise, 199 KB].
- Fahrenberg, J. (2008b). Wilhelm WUNDTs Interpretationslehre [43 Absätze]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, Online 9(3), Art. 29. Zugriff am 1–3–2009. <http://nbnresolving.de/urn:nbn:de:0114-fqs0803291>.
- Fahrenberg, J. (2008c). Die Wissenschaftskonzeptionen der Psychologie bei Kant und Wundt als Hintergrund heutiger Kontroversen. Struktureller Pluralismus der Psychologie und das Komplementaritätsprinzip. Defizite der Philosophischen und Psychologischen Anthropologie und ein Plädoyer für eine interdisziplinäre Anthropologie. Internet-Dokument. (Zugriff am 1.3.2009) URL: <http://psydok.sulb.uni-saarland.de/volltexte/2008/1557> [90 Seiten, 289 Literaturhinweise, 788 KB].
- Fahrenberg, J. (2009a). Open Access – nur Texte oder auch Primärdaten? (revidierte Fassung 5.10.2009). Verfügbar unter <http://www.jochen-fahrenberg.de/index.php?id=27>
- Fahrenberg, J. (2009b, im Druck). Ambulantes Assessment. In H. Holling & B. Schmitz (Hrsg.), *Handbuch der Psychologischen Methoden und Evaluation*. Göttingen: Hogrefe.
- Fahrenberg, J., Bolkenius, K., Maier, S., Schmidt, M., Foerster, F., Hüttner, P., Käßler, C. & Leonhart, R. (2002a). Evaluation des negativen Retrospektionseffektes. Untersuchungen mit Monitor. Forschungsbericht des Psychologischen Instituts der Universität Freiburg, Nr. 156. Zugriff am 1.3.2009. <http://psydok.sulb.uni-saarland.de/volltexte/2004/426/pdf/156.pdf> [97 Seiten, 60 Literaturhinweise, 6,7 MB].
- Fahrenberg, J., Brügger, G., Foerster, F. & Käßler, C. (1999). Ambulatory assessment of diurnal changes with a hand-held computer: Mood, attention and morningness-eveningness. *Personality and Individual Differences*, 26 (4), 641–656.
- Fahrenberg, J., Ewert, U. & Maier, N. (1987). *Reanalysen des Freiburger Persönlichkeitsinventars FPI*. Unveröff. Forschungsbericht. Psychologisches Institut, Universität Freiburg i.Br.
- Fahrenberg, J., Foerster, F., Schneider, H.J., Müller, W. & Myrtek, M. (1984). *Aktivierungsforschung im Labor-Feld-Vergleich*. München: Minerva.
- Fahrenberg, J., Hampel, R. & Selg, H. (1984–2001). *Das Freiburger Persönlichkeitsinventar* (revidierte Fassung FPI-R und teilweise geänderte Fassung FPI-A1 (4. Aufl. 1984, ergänzte 5. Aufl. 1989, ergänzte 6. Aufl. 1994, revidierte 7. Auflage 2001; 1. bis 3. Aufl. siehe Fahrenberg & Selg). Göttingen: Hogrefe.



- Fahrenberg, J., Hampel, R. & Selg, H. (1985). Die revidierte Form des Freiburger Persönlichkeitsinventars FPI-R. *Diagnostica*, 31, 1–21.
- Fahrenberg, J., Hampel, R. & Selg, H. (1994). *Inventário de Personalidade de Freiburg – FPI*. 6ª ed. Lisboa: Infoteste.
- Fahrenberg, J., Hampel, R. & Selg, H. (2007). *Inventarul de Personalitate Freiburg (FPI)*. Adaptat în România de Pitariu, H.D. & Iliescu, D. Cluj-Napoca: Odyseea.
- Fahrenberg, J., Hampel, R. & Selg, H. (2008). *FPI-R. Priručnik. Za Freiburski inventar licnosti*. Urednik hrvatskog izdanja Krunoslav Matesić. Jastrebarsko: Naklada Slap.
- Fahrenberg, J., Heger, R., Foerster, F. & Müller, W. (1991). Differentielle Psychophysiologie von Befinden, Blutdruck und Herzfrequenz im Labor-Feld-Vergleich. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 12, 1–25.
- Fahrenberg, J., Kuhn, M., Kulick, B. & Myrtek, M. (1977). Methodenentwicklung für psychologische Zeitreihenstudien. *Diagnostica*, 23, 15–36.
- Fahrenberg, J., Leonhart, R. & Foerster, F. (2002b). *Alltagsnahe Psychologie mit hand-held PC und physiologischem Mess-System*. Bern: Huber.
- Fahrenberg, J. & Myrtek, M. (Eds.). (1996). *Ambulatory assessment. Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Seattle, WA: Hogrefe & Huber.
- Fahrenberg, J. & Myrtek, M. (2001). Ambulantes Monitoring und Assessment. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Serie Biologische Psychologie. Band 1: Grundlagen und Methoden der Psychophysiologie* (S. 657–798). Göttingen: Hogrefe.
- Fahrenberg, J. & Myrtek, M. (2005). *Psychophysiologie in Labor, Klinik und Alltag. 40 Jahre Projektarbeit der Freiburger Forschungsgruppe Psychophysiologie – Kommentare und Neue Perspektiven*. Frankfurt a.M.: Lang.
- Fahrenberg, J., Myrtek, M., Kulick, B. & Frommelt, P. (1977). Eine psychophysiologische Zeitreihenstudie an 20 Studenten über 8 Wochen. *Archiv für Psychologie*, 128, 242–264.
- Fahrenberg, J., Myrtek, M., Pawlik, K. & Perez, M. (2007). Ambulantes Assessment – Verhalten im Alltagskontext erfassen. Eine verhaltenswissenschaftliche Herausforderung an die Psychologie. *Psychologische Rundschau*, 58, 12–23.
- Fahrenberg, J., Myrtek, M., Schumacher, J. & Brähler, E. (2000). *Fragebogen zur Lebenszufriedenheit (FLZ)*. Handanweisung. Göttingen: Hogrefe.
- Fahrenberg, J., Myrtek, M. & Trichtinger, I. (1985). Die Krankheitsursache aus der Sicht des Koronarpatienten. In W. Langosch (Hrsg.), *Psychische Bewältigung der chronischen Herzerkrankung* (S. 32–100). Berlin: Springer.
- Fahrenberg, J., Myrtek, M., Wilk, D., Kreutel, K. (1986). Multimodale Erfassung der Lebenszufriedenheit: Eine Untersuchung an Koronarkranken. *Psychotherapie Psychosomatik Medizinische Psychologie*, 36, 347–354.
- Fahrenberg, J. & Selg, H. (1970). *Das Freiburger Persönlichkeitsinventar FPI* (1. Auflage 1970, 2. erweiterte Aufl. 1973, ergänzte 3. Aufl. 1978). Göttingen: Hogrefe.
- Fahrenberg, J., Selg, H. & Hampel, R. (1983). Die bevölkerungsrepräsentative Normierung des Freiburger Persönlichkeitsinventars FPI-A. *Diagnostica*, 29, 336–343.
- Fahrenberg, J., Walschburger, P., Foerster, F., Myrtek, M. & Müller, W. (1979). *Psychophysiologische Aktivierungsforschung*. München: Minerva.
- Fahrenberg, J. & Wilhelm, F.H. (2009). Psychophysiologie und Verhaltenstherapie. In J. Margraf & S. Schneider (Hrsg.), *Lehrbuch der Verhaltenstherapie* (3. Aufl.). (S. 163–179). Berlin: Springer.
- Fastje, B. (2008). Angst und Persönlichkeit: FPI-R-Persönlichkeitsprofile bei Cluster C Persönlichkeitsstörungen, generalisierter Angststörung und sozialer Phobie. Unveröffentlichte Diplomarbeit. Universität Koblenz-Landau, Abt. Klinische Psychologie. Zugriff am 1.3.2009. [http://www.helios-kliniken.de/fileadmin/user\\_upload/Helios-Klinken.de/Bad\\_Groenenbach/PDF\\_Downloads/Diplomarbeit\\_Bodo\\_Fastje.pdf](http://www.helios-kliniken.de/fileadmin/user_upload/Helios-Klinken.de/Bad_Groenenbach/PDF_Downloads/Diplomarbeit_Bodo_Fastje.pdf) [112 Seiten, 3 MB].
- Fehr, T. (2002). Die modifizierende Wirkung sozialer Erwünschtheit in der psychologischen Selbstbeschreibung Praktizierender spiritueller Techniken am Beispiel der Transzendentalen Meditation. *Report Psychologie*, 27 (1), 22–31.
- Fenigstein, A., Scheier, M.F. & Buss, A.H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43, 522–527.
- Fersching, A. (1981). *Zu Bedingungen und Wirkungen allgemeiner Rehabilitationsmaßnahmen und besonderer psychologischer Behandlungsverfahren, vor allem für Patienten mit psychosomatisch-psychovegetativen Störungen*. Unveröff. Diss., Universität Freiburg.

- Feuerlein, W. & Küfner, H. (1989). A prospective multicenter study of in-patient treatment for alcoholics: 18- and 48-month follow-up. *European Archives of Psychiatry and Neurological Sciences*, 239, 144–157.
- Fichter, M.M. & Frick, U. (1992). *Therapie und Verlauf von Alkoholabhängigkeit. Auswirkungen auf Patienten und Angehörige*. Berlin: Springer.
- Fichter, M.M. & Quadflieg, N. (2000). Veränderung von Persönlichkeitsmerkmalen im Verlauf einer stationären Therapie. *Verhaltenstherapie*, 10 (3), 166–176.
- Fiedler, K. (1984). Diagnostische Fähigkeiten und diagnostische Erfahrung. In R.S. Jäger, A. Mattenklott & R.D. Schröder (Hrsg.), *Diagnostische Urteilsbildung in der Psychologie. Grundlagen und Anwendungen* (S. 303–327). Göttingen: Hogrefe.
- Finegan, J.E. & Allen, N.J. (1994). Computerized and written questionnaires: Are they equivalent? *Computers in Human Behavior*, 10, 483–496.
- Fineman, S. (1977). The achievement motive construct and its measurement: Where are we now? *British Journal of Psychology*, 68, 1–22.
- Fishbein, M. & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, 81, 59–74.
- Fiske, D.W. (1971). *Measuring the concepts of personality*. Chicago: Aldine.
- Fiske, D.W. (1978). *Strategies for personality research*. San Francisco: Jossey-Bass.
- Fiske, D.W. (1981). *Problems with language imprecision*. San Francisco: Jossey-Bass.
- Fiske, D.W. (1987). On understanding our methods and their effects. *Diagnostica*, 33, 188–194.
- Fiske, D.W. & Pearson, P.H. (1970). Theory and techniques of personality measurement. *Annual Review of Psychology*, 21, 49–86.
- Fisseni, H.-J. (2004). *Lehrbuch der Psychologischen Diagnostik* (3. Aufl.). Göttingen: Hogrefe.
- Flick, U. (2008). *Triangulation. Eine Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Flick, U., von Kardorff, E. & Steinke, I. (Hrsg.). (2000). *Qualitative Forschung – Ein Handbuch*. Hamburg: Rowohlt.
- Food and Drug Administration (2007). Guidance for Industry – Patient-reported outcome measures: Use in medical product development to support labeling claims. Zugriff am 1.3.2009. <http://www.fda.gov/cder/guidance/5460dft.pdf> [36 pages, 756 KB].
- Forgas, J.P. (1999). Soziale Kommunikation und Interaktion. *Eine Einführung in die Sozialpsychologie*. (4. Aufl.). Weinheim: Psychologie Verlags Union.
- Forman, A.K. (2002). Identifying types, response errors, and unscalable respondents from personality questionnaires. *Psychologische Beiträge*, 44, 78–93.
- Franke, G.H. (1997). Über die Möglichkeit der computerunterstützten Darbietung beim revidierten Freiburger Persönlichkeitsinventar. Zwei experimentelle Studien. *Zeitschrift für Experimentelle Psychologie*, 44, 332–356.
- Franke, G.H. (1998a). Auswirkungen der Computerapplikation beim Freiburger Persönlichkeitsinventar unter besonderer Berücksichtigung der Itemreihenfolge. Eine experimentelle Studie. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19, 187–199.
- Franke, G.H. (1998b). *Computerunterstützte klinisch-diagnostische Selbstbeurteilungsverfahren im Äquivalenztest. Experimentelle Studien*. Lengerich: Pabst Science Publishers.
- Franke, G.H. (2000). *Brief Symptom Inventory von L.R. Derogatis (Kurzform der SCL-90-R) – Deutsche Version BSI*. Manual. Göttingen: Beltz Test GmbH.
- Franke, G.H. (2002). Faking bad in personality inventories: Consequences for the clinical context. *Psychologische Beiträge*, 44, 50–61.
- Fricke, W. & Grauer, G. (1994). *Hochschulsozialisation im Sozialwesen. Entwicklung von Persönlichkeit. Studienbezogene Einstellungen. Berufliche Orientierungen*. Hannover: HIS GmbH.
- Friedberg, F. & Quick, J.M. S. (2007). Alexithymia in Chronic Fatigue Syndrome: Associations with momentary, recall, and retrospective measures of somatic complaints and emotions. *Psychosomatic Medicine*, 69 (1), 54–60.
- Funder, D.C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, 40, 21–34.

- Funke, J., Klein, M. & Scheller, R. (1981). Zur Klassifikation von Alkoholikern durch Persönlichkeitsmerkmale. *Psychologische Beiträge*, 23, 146–158.
- Gaensslen, H. & Schubö, W. (1976). *Einfache und komplexe statistische Analyse*. 2. Aufl. Basel: Reinhardt.
- Gergen, K.J., Hepburn, A. & Fisher, D.C. (1986). Hermeneutics of personality description. *Journal of Personality & Social Psychology*, 50, 1261–1270.
- Gerin, W., Christenfeld, N., Pieper, C., Su, O., Stroessner, S.J., Deich, J. & Pickering, T.G. (1998). The generalizability of cardiovascular responses across settings. *Journal of Psychosomatic Research*, 44, 209–218.
- Gitzinger, I. (1990). Akzeptanz der Darbietung eines Tests auf dem Personalcomputer von stationären Patienten/innen. *Psychotherapie – Psychosomatik – Medizinische Psychologie*, 40, 143–145.
- Gollwitzer, M., Eid, M. & Juergensen, R. (2005). Response styles in the assessment of anger expression. *Psychological Assessment*, 17 (1), 56–69.
- Gorin, A.A. & Stone, A.A. (2001). Recall biases and cognitive errors in retrospective self-reports: A call for momentary assessments. In A. Baum, T.A. Revenson & J.E. Singer (Eds.). *Handbook of health psychology* (pp. 405–413). New Jersey: Erlbaum.
- Gorsuch, R.L. & Yagel, J.C. (1981). *Item factor analysis*. Paper presented at the 1981 Annual Meeting of the American Educational Research Association, Los Angeles.
- Grawe, K. (1976). *Differentielle Psychotherapie I. Indikation und spezifische Wirkung von Verhaltenstherapie und Gesprächstherapie. Eine Untersuchung an phobischen Patienten*. Bern: Huber.
- Guilford, J.P. (1959). *Personality*. New York: McGraw Hill.
- Guilford, J.P. (1977). Will the real factor of extraversion-introversion please stand up? A reply to Eysenck. *Psychological Bulletin*, 84, 412–416.
- Gur, R.C. & Sackeim, H.A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, 37, 147–169.
- Gutjahr, W. (1972). *Die Messung psychischer Eigenschaften*. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Gynther, M.D. & Green, S.B. (1982). Methodological problems in research with self-report inventories. In P.C. Kendall & J.N. Butcher (Eds.). *Handbook of research methods in clinical psychology* (pp. 355–386). New York: Wiley.
- Häcker, H. (1982). Objektive Tests zu Messung der Persönlichkeit. In K.J. Groffmann & L. Michel (Hrsg.), *Persönlichkeitsdiagnostik. Enzyklopädie der Psychologie. Psychologische Diagnostik Band 3* (S. 132–185). Göttingen: Hogrefe.
- Häcker, H., Leutner, D. & Amelang, M. (Hrsg.). (1998). Standards für pädagogisches und psychologisches Testen. *Supplementum 1/1998 der Diagnostica und der Zeitschrift für Differentielle und Diagnostische Psychologie*.
- Häcker, H., Schwenkmezger, P. & Utz, E. (1979). Über die Verfälschbarkeit von Persönlichkeitsfragebogen und Objektiven Persönlichkeitstests unter SD-Instruktion und in einer Auslesesituation. *Diagnostica*, 25, 7–23.
- Hänsgen, K.D. (1991). *Berliner Verfahren zur Neurosendiagnostik BVND* (2. Aufl.). Göttingen: Hogrefe.
- Hageböck, J. (1994). *Computerunterstützte Diagnostik in der Psychologie*. Göttingen: Hogrefe.
- Hagen, H.P. von (1997). *Validierung von Persönlichkeitskonstrukten am Beispiel des Freiburger Persönlichkeitsinventars*. Frankfurt a.M.: Peter Lang.
- Hambleton, R.K., Merenda, P.F. & Spielberger, C.D. (Eds.). (2005). *Adapting psychological and educational tests for cross-cultural assessment*. New York: Routledge.
- Hambleton, R.K. & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and Guidelines for improving adaptation practices. Zugriff am 1.3.2009. <http://www.testpublishers.org/journal01.htm> [86 KB].
- Hambros, K. (2002). On reasonableness of personality inventories with dichotomous item response format. *Psychologische Beiträge*, 44 (1), 126–135.
- Hamilton, D.L. (1971). A comparative study of five methods of assessing self-esteem, dominance, and dogmatism. *Educational and Psychological Measurement*, 31, 441–452.
- Hampel, R. & Fahrenberg, J. (1982). *Die Freiburger Beschwerdenliste FBL. Gruppenvergleiche und andere Studien zur Validität*. Forschungsberichte des Freiburger Psychologischen Instituts, Nr. 7.
- Hampel, R. & Fahrenberg, J. (1983). *Zur Revision des FPI: Weitere Analysen und ergänzende Tabellen zur 4. Auflage*. Forschungsberichte des Psychologischen Instituts der Universität Freiburg, Nr. 12. Zugriff 1.3.2009 <http://www.jochen->

- Hampel, R. & Klinkhammer, F. (1978). Verfälschungstendenzen beim Freiburger Persönlichkeitsinventar in Bewerbungssituationen. *Psychologie und Praxis*, 22, 58–69.
- Hampel, R. & Krupp, B. (1978). Die Bedeutung kultureller und politischer Rahmenbedingungen für das Rassenvorurteil. Ein Vergleich von Rassenvorurteilen in Südafrika und Großbritannien. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 30, 116–135.
- Hampel, R. & Selg, H. (1975). *F-A-F. Fragebogen zur Erfassung von Aggressivitätsfaktoren*. Göttingen: Hogrefe.
- Hehl, F.J. (1981). Eine Entgegnung auf „Eine exemplarische Kritik am Persönlichkeitsskalen-System 25“ (PSS 25). *Diagnostica*, 27, 175–184.
- Hehl, F.J. & Hehl, R. (1975). *Persönlichkeits-Skalen System PSS 25*. Weinheim: Beltz.
- Heil, F.H. (1995). Klinische versus statistische Urteilbildung. In S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik* (3. Aufl.). (S. 39–42). Weinheim: Beltz, Psychologie-Verlags-Union.
- Heinemann, W. (1979). The assessment of private and public self-consciousness: A German replication. *European Journal of Social Psychology*, 9, 331–337.
- Heiß, R. (1982). Technik, Methodik und Problematik des psychologischen Gutachtens. In R. Heiß (Hrsg.), *Handbuch der Psychologie. Band 6. Psychologische Diagnostik* (S. 975–995). Göttingen: Hogrefe.
- Hektner, J.M. & Csikszentmihalyi, M. (2002). The Experience Sampling Method: Measuring the context and content of lives. In R.B. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 233–243). New York: Wiley.
- Hennig, J. & Netter, P. (Hrsg.). (2005). *Biopsychologische Grundlagen der Persönlichkeit*. München: Elsevier.
- Henning, H.J. (1980). Zur Testkonstruktion und Testanwendung von Persönlichkeitsfragebögen. Eine exemplarische Kritik am Persönlichkeitsskalen-System 25 (PSS 25). *Diagnostica*, 26, 119–134.
- Herrmann, C., Buss, U., Lingen, R. & Kreuzer, H. (1998). Persönlichkeitsfaktoren und Beschwerdepersistenz bei Patienten mit thorakalen Beschwerden und angiographisch freien Koronarien. *Zeitschrift für Psychosomatische Medizin und Psychoanalyse*, 44 (1), 37–53.
- Heubrock, D. & Petermann, F. (2008) *K-FAF. Kurzfragebogen zur Erfassung von Aggressivitätsfaktoren*. Göttingen: Hogrefe-Verlag.
- Heymans, G. & Wiersma, E. (1906). Beiträge zu einer speziellen Psychologie auf Grund einer Massenuntersuchung. *Zeitschrift für Psychologie und Physiologie des Sinnesorgane*, 42, 81–127.
- Hill, O. (1979). *Die Untersuchung der Homogenität zweier Persönlichkeitsfragebögen*. Unveröff. Diss. Universität Wien.
- Hobi, V. & Gerhard, U. (1983). Mitteilung über faktorielle und clusteranalytische Ergebnisse von Fragebogendaten des MMPI, 16 PF und FPI. *Zeitschrift für experimentelle und angewandte Psychologie*, 30, 588–609.
- Hobi, V. & Klär, A. (1973). Eine gemeinsame Faktorenanalyse von MMPI, FPI und 16 PF. *Zeitschrift für Klinische Psychologie*, 2, 27–48.
- Hodapp, V., Gableske, K., Riedemann, P. & Bongard, S. (2004). Konstruktion und Validierung eines Feindseligkeitsfragebogens auf der Basis des Handlungs-Häufigkeits-Ansatzes. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie*, 52 (4), 346–358.
- Hofstede, G.H. (2006). *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations* (2<sup>nd</sup> ed.). Thousand Oaks, Calif.: Sage.
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33, 306–316.
- Hogan, R. & Nicholson, R.A. (1988). The meaning of personality test scores. *American Psychologist*, 43, 621–626.
- Holling, H. (1981). Homogenität und faktorielle Validität von Skalen. *Diagnostica*, 27, 97–106.
- Hollmann, H. (1988). Das Freiburger Persönlichkeitsinventar. *Diagnostica*, 34, 277–285.
- Holtermann, M. (1989). Unterschiede bei schriftlicher und mündlicher Darbietung von Persönlichkeitsfragebogen. Zur Durchführungsobjektivität des Freiburger Persönlichkeitsinventars. *Diagnostica*, 35, 180–183.
- Hosser, D. Lauterbach, O. & Camehn, K. (2008). Validität und Reliabilität des FPI-R beim Einsatz im Strafvollzug. *Diagnostica*, 54 (3), 129–137.

- Hossiep, R. & Paschen, M. (2000). *Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung*. Göttingen: Hogrefe.
- Hossiep, R., Paschen, M. & Mülhhaus, O. (2000). *Persönlichkeitstests im Personalmanagement*. Göttingen: Verlag für angewandte Psychologie.
- Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D. & McCloy, R.A. (1990). Criterion-related validities of personality constructs and the effects of response distortion on those validities. *Journal of Applied Psychology*, 75, 581–595.
- Hoyer, J., Beauducel, A. & Franke, G.H. (2002). Kriterien der Angstdiagnostik – Woran der Anwender auch denken muss. In Hoyer, J. & Margraf, J. (Hrsg.), *Angstdiagnostik – Grundlagen und Testverfahren* (S. 77–90). Berlin: Springer.
- Hoyer, J. & Helbig, S. (2005). Diagnostische Verfahren bei Angststörungen. *Psychotherapie im Dialog*, 6 (4), 425–430.
- Ille, R., Lahousen, T., Rous, F., Hofmann, P. & Kapfhammer, H.P. (2005). Persönlichkeitsprofile und psychische Abweichungen bei psychiatrisch-forensisch begutachteten Straftätern. *Der Nervenarzt*, 76 (1), 52–60.
- International Test Commission (2008). Guidelines on Computer-Based and Internet Delivered Testing. (Zugriff 1.1.2009) [http://www.intestcom.org/guide lines/computer+based+testing](http://www.intestcom.org/guide%20lines/computer+based+testing)
- International Test Commission (2008) Guidelines test adaptation (April 21, 2000 Version). (Zugriff 1.3. 2009). [http://www.intestcom.org/Guidelines/test+ adaptation.php](http://www.intestcom.org/Guidelines/test+adaptation.php)
- International Test Commission (2008). International Guidelines for Test Use. (siehe auch Trier: ZPID. Internet-Dokument). (Zugriff am 10.3.2009). [http:// www.intestcom.org/Guidelines/guidelines+for+test +use.php](http://www.intestcom.org/Guidelines/guidelines+for+test+use.php)
- Jackson, D.N. & Paunonen, S.V. (1980). Personality structure and assessment. *Annual Review of Psychology*, 31, 503–551.
- Jäger, R.S. (1985). Testinformationen. Das Freiburger Persönlichkeitsinventar (FPI). *Diagnostica*, 31, 246–250.
- Jäger, R., Lischer, S., Münster, B. & Ritz, B. (1976). *Biographisches Inventar zur Diagnostik von Verhaltensstörungen (BIV)*. Göttingen: Hogrefe.
- Jäger, S. & Petermann, F. (Hrsg.). (1995). *Psychologische Diagnostik*. (3. Aufl.). Weinheim: Beltz, Psychologie-Verlags-Union.
- Janke, W. (1973). Das Dilemma von Persönlichkeitsfragebogen. Einleitung des Symposiums über Konstruktion von Fragebogen. In G. Reinert (Hrsg.), *Bericht über den 27. Kongress der Deutschen Gesellschaft für Psychologie in Kiel 1970* (S. 44–48). Göttingen: Hogrefe.
- Janssen, J.P. & Reitzner, C. (1980). Zur Validität und Reliabilität des KFP-30, der FPI-Skalen „Aggressivität, Erregbarkeit und Dominanz“ in einer echten Prüfungssituation. *Psychologie und Praxis*, 24, 148–161.
- Jehle, M. (1988). *Literaturverzeichnis zum Freiburger Persönlichkeitsinventar FPI. Veröffentlichungen von 1983 bis 1988* (mit Nachträgen). Psychologisches Institut der Universität Freiburg.
- Jehle, M. & Fahrenberg, J. (1994). *Literaturverzeichnis zum Freiburger Persönlichkeitsinventar FPI 1983 – 1993*. (Forschungsberichte des Psychologischen Instituts der Albert-Ludwigs-Universität Nr. 105). Freiburg i.Br., Psychologisches Institut.
- Johann, G.K. & Bergemann, N. (1985). Alternativen multivariater Forschungsstrategien. Die Analyse von -Intelligenz- und Persönlichkeitsdaten mittels Faktorenanalyse, Clusteranalyse und NMDS. In D. Albert (Hrsg.), *Bericht über den 34. Kongress der Deutschen Gesellschaft für Psychologie in Wien 1984*, Band 1 (S. 164–165). Göttingen: Hogrefe.
- John, D. & Keil, W. (1972). Selbsteinschätzung und Verhaltensbeurteilung. *Psychologische Rundschau*, 23, 10–29.
- John, O.P. (1990). The search for basic dimensions of personality. A review and critique. In P. McReynolds, J.C. Rosen & G.J. Chelune (Eds.), *Advances in psychological assessment*. (Vol. 7, pp. 1–7). New York: Plenum Press.
- Jüttemann, G. (1991). Systemimmanenz als Ursache der Dauerkrise „wissenschaftlicher“ Psychologie. In G. Jüttemann, M. Sonntag & C. Wulf (Hrsg.), *Die Seele. Ihre Geschichte im Abendland* (S. 340–363). Weinheim: Psychologie Verlags Union.
- Jüttemann, G. (2004). *Psychologie als Humanwissenschaft. Ein Handbuch*. Göttingen: Vandenhoeck & Ruprecht.
- Jüttemann, G. (Hrsg.). (2006). *Wilhelm Wundts anderes Erbe. Ein Missverständnis löst sich auf* (S. 13–30). Göttingen: Vandenhoeck & Ruprecht.
- Kaase, M. (Hrsg.). (1999). *Qualitätskriterien der Umfrageforschung. Denkschrift der Deutschen Forschungsgemeinschaft*. Berlin: Akademie Verlag.
- Käppler, C. (1994). *Psychophysiologische Bedingungsanalyse von Blutdruckveränderungen im alltäglichen Lebenskontext*. Phil. Diss., Albert-Ludwigs-Universität, Freiburg i.Br. Frankfurt a.M.: Peter Lang.
- Käppler, C., Brügger, G. & Fahrenberg, J. (2001). Pocketcomputer-unterstütztes Assessment mit MONITOR: Befindlichkeit im

- Alltag, Methodenakzeptanz und die Replikation des Retrospektionseffektes. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 249–266.
- Kenrick, D.T. & Funder, D.C. (1988). Profiting from controversy: Lessons from the person-situation-debate. *American Psychologist*, 43, 23–34.
- Kersting, M. (2006). Zur Beurteilung der Qualität von Tests: Resümee und Neubeginn. *Psychologische Rundschau*, 57, 243–253.
- Klages, U. (1985). Ein Vergleich von Alkoholismus-Patienten, längerfristig abstinenten Alkoholikern und Normalpersonen auf dem Freiburger Persönlichkeitsinventar. *Zeitschrift für Klinische Psychologie*, 4, 53–59.
- Klapproth, H.E. (1981). Kritik und Weiterentwicklung der Testtheorie-Realisierungsmöglichkeiten moderner Psychodiagnostik. *Psychologie und Praxis*, 25, 139–150.
- Knowles, E.S., Coker, M.C., Cook, D.A., Diercks, S.R., Irwin, M.E., Lundeen, E.J., Neville, J.W. & Sibicky, M.E. (1992). Order effects within personality measures. In N. Schwarz & S. Sudman (Eds.), *Context effects in social and psychological research* (pp. 221–236). Berlin: Springer.
- Koch, U. & Laschinsky, D. (1979). Ein Fragebogen zur Erfassung der Situation am Arbeitsplatz und in der Familie (KOLA). *Psychologie und Praxis*, 4, 165–173.
- Köhler, D., Müller, S., Kernbichler, A., van den Boogaart, H. & Hinrichs, G. (2007). Tathergangsanalyse in der forensischen Praxis? Die Beziehung zwischen Täterverhalten und Persönlichkeit bei Gewalt- und Sexualstraftätern. *Monatsschrift für Kriminologie und Strafrechtsreform*, 90 (5), 360–373.
- König, F. & Schmidt, J.U. (1982). Skalen- und Strukturanalyse des Freiburger Persönlichkeitsinventars (FPI). *Diagnostica*, 28, 33–48.
- Kopp, M., Bonatti, H., Haller, C., Rumpold, G., Soellner, W., Holzner, B., Schweigkofler, H., Aigner, F., Hinterhuber, H. & Guenther, V. (2003). Life satisfaction and active coping style are important predictors of recovery from surgery. *Journal of Psychosomatic Research*, 55, 371–377.
- Krahe, B. (1989). Faking personality profiles on a standard personality inventory. *Personality and Individual Differences*, 10, 437–443.
- Krampen, G. (1989). Zur Validität der deutschen Form des 16 PF: Faktorielle Validität und Beziehungen zum FPI. *Diagnostica*, 32, 91–99.
- Krampen, G. (1991). *Fragebogen zu Kompetenz- und Kontrollüberzeugungen*. Handanweisung. Göttingen: Hogrefe.
- Krampen, G., Hense, H. & Schneider, J.F. (1992). Reliabilität und Validität von Fragebogenskalen bei Standardreihenfolge versus inhaltshomogener Blockbildung ihrer Items. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 229–248.
- Krauth, J. (1995). *Testkonstruktion und Testtheorie*. Weinheim: Beltz, Psychologie-Verlags-Union.
- Kreim, G. (2006). Belastungsreaktion und Trauma im Vollzug polizeilicher und militärischer Spezialaufgaben. *Polizei & Wissenschaft*, 2, 93–101.
- Kreutel, K. (1989). *Erfassung des Krankheitsverhaltens. Eine Untersuchung an Herz-Kreislauf-Patienten zu Beginn und zu Ende eines stationären Heilverfahrens*. Frankfurt a.M.: Peter Lang.
- Kröber, H.-L., Scheurer, H. & Richter, P. (1993). *Ätiologie und Prognose von Gewaltdelinquenz*. Regensburg: Roderer.
- Kröner-Herwig, B. & Weich, K.W. (1988). Untersuchung zur Prädiktion des Erfolgs verhaltensmedizinischer Intervention bei chronischem Kopfschmerz. *Zeitschrift für Klinische Psychologie*, 17, 55–69.
- Kubinger, K.D. (1995) *Einführung in die psychologische Diagnostik*. Weinheim: Psychologie Verlags-Union.
- Kubinger, K.D. (2008). Einige Ansätze einer experimentalpsychologischen Verhaltensdiagnostik. In W. Sarges & N. Scheffer (Hrsg.), *Innovative Ansätze in der Eignungsdiagnostik* (S. 109–120). Göttingen: Hogrefe.
- Kubinger, K.D. (Ed.). (2002a). Special issue: Personality questionnaires: Some critical points of view. *Psychologische Beiträge*, 44, 3–158.
- Kubinger, K.D. (2002b). On faking personality inventories. *Psychologische Beiträge*, 44, 10–16.
- Kubinger, K.D. (2003a). Gutachten, psychologisches. In K.D. Kubinger & S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 187–194). Weinheim: Beltz.
- Kubinger, K.D. (2003b). Gütekriterien. In K.D. Kubinger & S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 195–201). Weinheim: Beltz.

- Kubinger, K.D. (2003c). Objektiver Persönlichkeitstest. In K.D. Kubinger & S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 304–309). Weinheim: Beltz.
- Kubinger, K.D. & Draxler, C. (2007). Probleme bei der Testkonstruktion nach dem Rasch-Modell. *Diagnostica*, 53, 131–143.
- Kubinger, K.D. & Jäger, S. (Hrsg.). (2003). *Schlüsselbegriffe der psychologischen Diagnostik*. Weinheim: Beltz.
- Külpe, O. (1920). *Vorlesungen über Psychologie* (hrsg. von Karl Bühler). Leipzig: Hirzel.
- Kury, H. (1983). Zur Verfälschbarkeit von Persönlichkeitsfragebogen bei jungen Strafgefangenen. *Zeitschrift für Strafvollzug und Straffälligenhilfe*, 32, 323–332.
- Kury, H. (1986). Problems in the treatment of offenders and in treatment research. *The German Journal of Psychology*, 10, 61–80.
- Kury, H. (2002). Das Freiburger Persönlichkeitsinventar und sein Einsatz bei kriminologischen Fragestellungen. Das Problem der Verfälschungstendenzen. In M. Myrtek (Hrsg.), *Person im biologischen und sozialen Kontext* (S. 249–270). Göttingen: Hogrefe.
- Lamberti, G. (1991). Persönlichkeitsdiagnostik: Computergestützte oder Paper-Pencil-Vorgabe. *Zeitschrift für Neuropsychologie*, 2, 125–129.
- Langosch, W. (1989). *Psychosomatik der koronaren Herzkrankheit*. Weinheim: Verlag Chemie, Edition Medizin.
- Langosch, W., Borchering, H., Brodner, G. & Wytibul, K. (1989). Zur Vorhersage einiger Aspekte von Lebensqualität nach Myokardinfarkt. In H. Speidel & B. Strauss (Hrsg.), *Zukunftsaufgaben der psychosomatischen Medizin* (S. 252–266). Berlin: Springer.
- Lankes, W. (1915). Perseveration. *British Journal of Psychology*, 7, 387–419.
- Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C.D. (1981). *State-Trait-Angstinventar* (STAI). Weinheim: Beltz.
- Lawyer, S.R. & Smitherman, T.A. (2004). Trends in anxiety assessment. *Journal of Psychopathology and Behavioral Assessment*, 26, 101–106.
- Leung, K., Bond, M.H., de Carrasquel, S.R., Munoz, C., Hernandez, M., Murakami, F., Yamaguchi, S., Bierbrauer, G. & Singelis, T.M. (2002). Social Axioms. The search for universal dimensions of general beliefs about how the world functions. *Journal of Cross-Cultural Psychology*, 33, 286–302.
- Lienert, G.A. (1961). *Testaufbau und Testanalyse* (2. Aufl.). Weinheim: Beltz.
- Löhr, F.J. & Angleitner, A. (1980). Eine Untersuchung zu sprachlichen Formulierungen der Items in deutschen Persönlichkeitsfragebogen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 1, 217–236.
- Lösel, F.J. (1988). Persönlichkeitsdaten (Tests). In R.S. Jäger (Hrsg.), *Psychologische Diagnostik* (S. 303–320). München: Psychologie Verlags Union.
- Lösel, F. (1995). Persönlichkeitsdaten (Tests). In S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik* (3. Aufl.). (S. 362–380). Weinheim: Beltz, Psychologie-Verlags-Union.
- Lucas, R.E. & Baird, B.M. (2005). Global self-assessment. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 29–42). Washington, D.C.: American Psychological Association.
- Marsella, A.J., Dubanoski, J., Hamada, W.C. & Morse, H. (2000). The measurement of personality across cultures. *American Behavioral Scientist*, 44, 41–62.
- Massoubre, C., Jaeger, B., Milos, G., Schmidt, U., Soares, I., Papezova, H., Denia, M., Faragalli, G., Westerlund, A.M., Pellet, J. & Lang, F. (2005). FPI Profiles in a European Sample of 1068 female patients suffering from Anorexia or Bulimia Nervosa. *European Eating Disorders Review*, 13 (3), 201–210.
- Massoubre, C., Lang, F., Jaeger, B., Jullien, M. & Pellet, J. (2002). La traduction des questionnaires et des tests: techniques et problèmes. Translating Questionnaires and Tests: Techniques and Problems. *The Canadian Journal of Psychiatry/La Revue canadienne de psychiatrie*, 47 (1), 63–69.
- Matesanz, A. & Hampel, R. (1978). Intercultural comparative study using the FPI (Freiburg Personality Inventory) of German and Spanish subjects. *Zeitschrift für experimentelle und angewandte Psychologie*, 25, 218–229.
- Matthews, G. & Amelang, M. (1993). Extraversion, arousal theory and performance: A study of individual differences in the EEG. *Personality and Individual Differences*, 14, 347–363.
- McCrae, R.R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293–303.

- McCrae, R.R. & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81–90.
- McCrae, R.R. & Costa, P.T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509–516.
- McCrae, R.R. & Terracciano, A. (2005). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology*, 89, 407–425.
- Medert-Dornscheidt, G. (1975). *Psychophysiologische Korrelationen bei kardiovaskulären Erkrankungen und ihre Bedeutung für den Rehabilitationsverlauf*. Phil. Diss. Freiburg.
- Meehl, P.E. (1954). *Clinical versus statistical prediction. A theoretical analysis and a review of evidence*. Minneapolis: University of Minnesota Press.
- Mehl, M.R. & Holleran, S.E. (2007). An unobtrusive observation sampling approach to psychological ambulatory assessment: The Electronically Activated Recorder (EAR). *European Journal of Psychological Assessment*, 23, 248–257.
- Meier, F. (1985). Sozial erwünschtes Antwortverhalten: ein fiktives Forschungsproblem? *Diagnostica*, 31, 289–299.
- Merikangas, K.R., Stevens, D.E. & Angst, J. (1993). Headache and personality: Results of a community sample of young adults. *Journal of Psychiatric Research*, 27, 187–196.
- Merten, T., Friedel, E., Mehren, G. & Stevens, A. (2007). Über die Validität von Persönlichkeitsprofilen in der nervenärztlichen Begutachtung. *Der Nervenarzt*, 78, 511–520.
- Meyer, A.E. (1981). The Hamburg Short Psychotherapy Comparison Experiment. *Psychotherapy and Psychosomatics*, 35, 81–107.
- Meyer, A.E., Stuhr, U., Wirth, U. & Rüster, P. (1988). 12-year follow-up study of the Hamburg short psycho-therapy experiment: An overview. *Psychotherapy and Psychosomatics*, 50, 192–200.
- Michel, G. (2006). The influence of neuroticism on concurrent symptom reporting: A multilevel modelling approach. *Personality and Individual Differences*, 41 (3), 549–560.
- Michel, L. & Conrad, W. (1982). In K.J. Groffmann & L. Michel (Hrsg.), *Persönlichkeitsdiagnostik. Enzyklopädie der Psychologie. Psychologische Diagnostik Band 3* (S. 1–129). Göttingen: Hogrefe.
- Millham, J. & Kellogg, R.W. (1980). Need for social approval: Impression management or self-deception? *Journal of Research in Personality*, 14, 445–457.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W. (1979). On the interface of cognition and personality. Beyond the person-situation debate. *American Psychologist*, 34, 740–754.
- Mittenecker, E. (1982). Subjektive Tests zur Messung der Persönlichkeit. In K.J. Groffmann & L. Michel (Hrsg.), *Persönlichkeitsdiagnostik. Enzyklopädie der Psychologie. Psychologische Diagnostik. Band 3 Persönlichkeitsdiagnostik* (S. 57–131). Göttingen: Hogrefe.
- Monden-Engelhardt, C. (1986). *Attribuierungserinnerungen und Selbstbeurteilungen in der Persönlichkeitsdiagnostik*. Diss. Universität Frankfurt. Eschborn: Fachbuchhandlung für Psychologie.
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Moosbrugger, H., Stemmler, G. & Kersting, M. (2008). Qualitätssicherung und -optimierung im Aufbruch: Die ersten Testrezensionen nach dem neuen TBS-TK System. *Psychologische Rundschau*, 59, 182–184.
- Moosbrugger, H., Zwingmann, C. & Frank, D. (Hrsg.). (1996). *Religiosität, Persönlichkeit und Verhalten. Beiträge zur Religionspsychologie*. Münster: Waxmann.
- Mueller-Hanson, R.A., Heggestad, E.D. & Thornton, G.C. (2006). Individual differences in impression management: An exploration of the psychological processes underlying faking. *Psychology Science*, 48 (3), 288–312.
- Mühlig, S. & Petermann, F. (2006). Grundprinzipien multimethodaler Diagnostik. In F. Petermann & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 99–108). Göttingen: Hogrefe.
- Mummendey, H.D. (1981). Methoden und Probleme der Kontrolle sozialer Erwünschtheit (Social Desirability). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 2, 199–218.
- Mummendey, H.D. (1995). *Psychologie der Selbstdarstellung* (2. Aufl.). Göttingen: Hogrefe.
- Mummendey, H.D. (1999). *Die Fragebogen-Methode* (3. Aufl.). Göttingen: Hogrefe.



- Mummendey, H.D. & Bolten, H.G. (1981). Die Veränderung von Social-Desirability-Antworten bei erwarteter Wahrheitskontrolle (Bogus-Pipeline-Paradigma). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 2, 151–156.
- Myrtek, M. (1980). *Psychophysiologische Konstitutionsforschung*. Göttingen: Hogrefe.
- Myrtek, M. (1983). *Typ-A-Verhalten. Untersuchungen und Literaturanalysen unter besonderer Berücksichtigung der psychophysiologischen Grundlagen*. München: Minerva.
- Myrtek, M. (1984). *Constitutional psychophysiology*. New York: Academic Press.
- Myrtek, M. (1985). Experimentelle Untersuchungen zum Allgemeinen Psychovegetativen Syndrom. Versuch einer Zusammenfassung. In D. Vaitl, T.W. Knapp & N. Birbaumer (Hrsg.), *Psychophysiologische Merkmale klinischer Symptome*, Band 1: Psychophysiologische Dysfunktionen (S. 36–43). Weinheim: Beltz.
- Myrtek, M. (1987). Life satisfaction, illness behaviour, and rehabilitation outcome: Results of a one year follow-up study with cardiac patients. *International Journal of Rehabilitation Research*, 10, 373–382.
- Myrtek, M. (1995). Type A behavior pattern, personality factors, disease, and physiological reactivity: A meta-analytic update. *Personality and Individual Differences*, 18, 491–502.
- Myrtek, M. (1998a). *Gesunde Kranke – kranke Gesunde. Psychophysiologie des Krankheitsverhaltens*. Bern: Huber.
- Myrtek, M. (1998b). Metaanalysen zur psychophysiologischen Persönlichkeitsforschung. In F. Rösler (Hrsg.), *Enzyklopädie der Psychologie. Serie I Biologische Psychologie. Bereich Theorie und Forschung. Bd 5. Ergebnisse und Anwendungen der Psychophysiologie* (S. 285–344). Göttingen: Hogrefe.
- Myrtek, M. (1999). Psychophysiologische Reaktivität, Streß, Typ-A-Verhalten und Feindseligkeit als Risikofaktoren der koronaren Herzkrankheit. *Verhaltenstherapie und Verhaltensmedizin*, 20, 89–119.
- Myrtek, M. (2002). Psychophysiologische Untersuchungen zum gestörten Schlaf. In M. Myrtek (Hrsg.), *Person im biologischen und sozialen Kontext* (S. 61–83). Göttingen: Hogrefe.
- Myrtek, M. (2004). *Heart and emotion. Ambulatory monitoring studies in everyday life*. Cambridge, MA: Hogrefe & Huber.
- Myrtek, M. & Fahrenberg, J. (1998). Somatoforme Störungen: Konzeptuelle und methodologische Kritik und ein Plädoyer für die funktionale Analyse des Krankheitsverhaltens. In J. Margraf, S. Neumer, & W. Rief (Hrsg.), *Somatoforme Störungen: Ätiologie, Diagnose und Therapie* (S. 191–211). Berlin: Springer.
- Myrtek, M., Kreutel, K., Wilk, D., Welsch, M. & Herzog, M. (1987). Lebenszufriedenheit und Rehabilitationsverlauf. Eine Untersuchung an Herz-Kreislauf-Patienten. *Rehabilitation*, 26, 11–19.
- Myrtek, M., Schmidt, T.H. & Schwab, G. (1984). Untersuchungen zur Reliabilität und Validität der deutschen Version des Jenkins Activity Survey (JAS). *Zeitschrift für Klinische Psychologie*, 13, 322–337.
- Myrtek, M. & Welsch, M. (1990). Determinants of rehabilitation outcome – Results of follow-up studies of different patient groups with special reference to cardiac patients. *Psychology and Health*, 5, 25–37.
- Nicholls, J.G., Licht, B.G. & Pearl, R.A. (1982). Some dangers of using personality questionnaires to study personality. *Psychological Bulletin*, 92, 572–580.
- Nübling, R. (1992). *Psychotherapiemotivation und Krankheitskonzept. Zur Evaluation psychosomatischer Heilverfahren*. Phil. Diss. Albert-Ludwigs-Universität Freiburg i.Br., Frankfurt a.M.: Verlag für Akademische Schriften.
- Nunnally, J.C. (1978). *Psychometric theory*. (2<sup>nd</sup> edition). New York: McGraw Hill.
- Orth, B. (1983). Grundlagen des Messens. In H. Feger & J. Bredenkamp (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich B. Methodologie und Methoden. Serie I Forschungsmethoden der Psychologie. Band 3 Messen und Testen* (S. 136–180). Göttingen: Hogrefe.
- Orth, B. (1995). Meßtheoretische Grundlagen der Diagnostik. In S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik* (3. Aufl.). (S. 286–295). Weinheim: Beltz, Psychologie-Verlags-Union.
- Ortmann, R. (1987). *Resozialisierung im Strafvollzug. Theoretischer Bezugsrahmen und empirische Ergebnisse einer Längsschnittstudie zu den Wirkungen von Strafvollzugsmaßnahmen. Kriminologischer Forschungsbericht aus dem MPI ausländisches und internationales Strafrecht*. Freiburg: MPI.
- Ortner, T., Proyer, R.T. & Kubinger, K.D. (2006). *Theorie und Praxis objektiver Persönlichkeitstests*. Bern: Huber.
- Ostendorf, F. (1997). Freiburger Persönlichkeitsinventar – Rev. Fassung (FPI-R). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, 81–85.

- Ostendorf, F., Angleitner, A. & Ruch, W. (1986). *Die Multitrait-Multimethod Analyse. Konvergente und diskriminante Validität der Personality Research Form*. Göttingen: Hogrefe.
- Palmblad, M. & Tiplady, B. (2004). Electronic diaries and questionnaires: Designing user interfaces that are easy for all patients to use. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, 13, 1199–1207.
- Paulhus, D.L. (1989). Socially desirable responding: Some new solutions to old problems. In D.M. Buss & N. Cantor (Eds.). *Personality psychology. Recent trends and emerging directions* (pp. 201–209). Berlin: Springer.
- Paulhus, D.L. (2002). Social desirable responding: The evolution of a construct. In H.I. Braun, D.N. Jackson & D.E. Wiley (Eds.). *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paunonen, S.V. (1984). The reliability of aggregated measurements: Lessons to be learned from psychometric theory. *Journal of Research in Personality*, 18, 383–394.
- Paunonen, S.V. & Ashton, M.C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, 81, 524–539.
- Pawlik, K. & Buse, L. (1982). Rechnergestützte Verhaltensregistrierung im Feld: Beschreibung und erste psychometrische Überprüfung einer neuen Erhebungsmethode. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 101–118.
- Pawlik, K. & Buse, L. (1992). Felduntersuchungen zur transsituativen Konsistenz individueller Unterschiede im Erleben und Verhalten. In K. Pawlik & K.H. Stapf (Hrsg.), *Umwelt und Verhalten* (S. 25–69). Bern: Huber.
- Pawlik, K. & Buse, L. (1996). Verhaltensbeobachtung in Labor und Feld. In K. Pawlik (Hrsg.), *Enzyklopädie der Psychologie. Differentielle Psychologie und Persönlichkeitsforschung. Band 1. Grundlagen und Methoden der Differentiellen Psychologie* (S. 359–394). Göttingen: Hogrefe.
- Pawlik, K. & Buse, L. (2008). Umweltpsychologische Methoden der Beobachtung und Datenerhebung. In E.-D. Lantermann & V. Linneweber (Hrsg.). (2008). *Enzyklopädie der Psychologie. Umweltpsychologie. Band 1. Grundlagen, Paradigmen und Methoden der Umweltpsychologie* (S. 751–783). Göttingen: Hogrefe.
- Perrez, M., Schoebi, D. & Wilhelm, P. (2000). How to assess social regulation of stress and emotions in daily family life? A computer-assisted family self-monitoring system (FASEM-C). *Clinical Psychology and Psychotherapy*, 7, 326–339.
- Perrez, M., Watzek, D., Michel, G., Schoebi, D., Wilhelm, P. & Hänggi, Y. (2005). Facets of Emotion Regulation in Families with Adolescents: A New Research Approach. In H. Kriesi, P. Farago, M. Kohli & M. Zarin-Nejadan (Eds.). *Contemporary Switzerland. Revisiting the Special Case* (pp. 61–80). New York: Palgrave Macmillan.
- Petermann, F. (1995). Historische Kontroversen. In S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik (3. Aufl.)*. (S. 36–39). Weinheim: Beltz, Psychologie-Verlags-Union.
- Petermann, F. & Eid, M. (2006). *Handbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- Peterson, D.R. (1965). Scope and generality of verbally defined personality factors. *Psychological Review*, 72, 48–109.
- Pohl, R.F. (Ed.). (2004). *Cognitive illusions. A handbook on fallacies and biases in thinking, judgment and memory*. New York: Psychology Press.
- Ponocny, I. & Klauer, K.C. (2002). Towards identification of unsalable personality questionnaire respondents: The use of person fit indices. *Psychologische Beiträge*, 44, 94–107.
- Potreck, F. (1983). *Kommentiertes Literaturverzeichnis zum Freiburger Persönlichkeitsinventar FPI. Veröffentlichungen seit 1978*. Forschungsberichte des Psychologischen Instituts der Universität Freiburg, Nr. 11.
- Prill, T. & Fahrenberg, J. (2007). New methods in ambulatory BP monitoring: Interactive monitoring and detection of posture and movement patterns. *Behavior Research Methods, Instruments & Computers*, 39 (3), 390–398.
- Propping, P., Rey, E.-R., Friedl, W. & Beckmann, H. (1981). Platelet monoamine oxidase in healthy subjects. The „biochemical high-risk paradigm“ revisited. *Archiv für Psychiatrie und Nervenkrankheiten*, 230, 209–219.
- Rammsayer, T. & Troche, S. (Hrsg.). (2005). *Reflexionen der Psychologie. 100 Jahre Deutsche Gesellschaft für Psychologie. Bericht über den 44. Kongress der Deutschen Gesellschaft für Psychologie in Göttingen 2004*. Göttingen: Hogrefe.
- Rammstedt, B. (2006). Fragebogen. In F. Petermann & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 109–134). Göttingen: Hogrefe.
- Raselli, C. & Broderick, J.E. (2007). The association of depression and neuroticism with pain reports: A comparison of momentary and recalled pain assessment. *Journal of Psychosomatic Research*, 62 (3), 313–320.

- Rau, R. (1992). *Handlungssicherheit bei der Dispatchertätigkeit im Elektroenergieversorgungssystem. Eine psychophysiologische Untersuchung*. Unveröff. Diss., TU Dresden.
- Remschmidt, H., Höhner, G. & Walter, R. (1983). Die spätere Entwicklung delinquenter Kinder. In M.H. Schmidt & H. Remschmidt (Eds.), *Epidemiological approaches in child psychiatry 11. International symposium Mannheim 1981* (pp. 43–56). Stuttgart: Thieme.
- Richman, W.L., Kiesler, S., Weisband, S. & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84, 754–775.
- Rink, J. (1987). Überprüfung und Versuch einer Erklärung der Wirksamkeit eines regional konzipierten Behandlungskonzepts für Abhängige in einem psychiatrischen Landeskrankenhaus. In D. Kleiner (Hrsg.), *Langzeitverläufe bei Suchtkrankheiten* (S. 310–318). Berlin: Springer.
- Robinson, J. P. & Shaver, P.R. (1973). *Measures of social psychological attitudes*. rev. ed. Ann Arbor, Mich.: Univ. of Michigan.
- Rohrmann, B. (1978). Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9, 222–245.
- Rosch, E., Chassein, J., Semin, G.R. & Krolage, J. (1984). Psychometrische Modelle der Persönlichkeit als soziale Konstruktionen: Experimentelle Untersuchungen zu Gütekriterien. *Schweizerische Zeitschrift für Psychologie*, 43, 91–110.
- Rost, J. (1999). Was ist aus dem Rasch-Modell geworden? *Psychologische Rundschau*, 50 (3), 140–156.
- Rost, J. (2002). When personality questionnaires fail to be unidimensional. *Psychologische Beiträge*, 44, 108–125.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Aufl.). Bern: Huber.
- Rost, J. (2006). Latent-Class-Analyse. In F. Petermann & M. Eid, M. (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 275–287). Göttingen: Hogrefe.
- Rost, D.H. & Hoberg, K. (1997). Itempositionsveränderungen im Persönlichkeitsfragebogen: Methodischer Kunstfehler oder tolerierbare Praxis? *Diagnostica*, 43, 97–112.
- Ruch, W. (1999). Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK. *Zeitschrift für Diagnostische und Differentielle Psychologie*, 20, 1–24.
- Ryan, A.M. & Sackett, P.R. (1987). Pre-employment honesty testing: Fakability, reactions of test takers, and company image. *Journal of Business and Psychology*, 1, 248–256.
- Sarges, W. (Hrsg.). (1995). *Management-Diagnostik* (2. Aufl.). Göttingen: Hogrefe.
- SAS (1996). *SAS/STAT User's Guide. Version 6.12*. (4<sup>th</sup> ed.), Volume 1–2. Cary, NC: SAS Institute Inc.
- Saucier, G. (1998). Replicable item-cluster subcomponents in the NEO Five-Factor Inventory. *Journal of Personality Assessment*, 70, 263–276.
- Schahn, J. & Amelang, M. (1992). Mittelwertsunterschiede zwischen Selbst- und Fremdbeurteilungen: Eine vernachlässigte Größe? *Diagnostica*, 38, 187–208.
- Scheibehenne, B., Saller, T. & Fahrenberg, J. (2000). *Befinden im Tageslauf: Zwei Untersuchungen mit MO-NITOR. Forschungsbericht des Psychologischen Instituts*. (Forschungsbericht Nr. 147). Freiburg i.Br.: Universität, Psychologisches Institut, Forschungsgruppe Psychophysiologie.
- Scheller, R. & Klein, M. (1982). Persönlichkeitspsychologische Determinanten des Therapieerfolgs bei Alkoholabhängigen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 47–54.
- Scheller, R. & Klein, M. (1986). Stationäre Suchttherapie unter dem Aspekt der Behandlungsdauer. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 7, 69–76.
- Schepank, H. (1987). *Psychogene Erkrankungen der Stadtbevölkerung. Eine epidemiologisch-tiefenpsychologische Feldstudie in Mannheim*. Berlin: Springer.
- Schepank, H. (Hrsg.). (1990). *Seelische Gesundheit und psychogene Erkrankungen heute*. Berlin: Springer.
- Schermelleh-Engel, K. & Schweizer, K. (2006). Multitrait-Multimethod-Analyse. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie: Test- und Fragebogenkonstruktion* (3. Aufl.). (S. 325–341). Heidelberg: Springer.
- Scheurer, H. (1993). *Persönlichkeit und Kriminalität. Eine theoretische und empirische Analyse*. Phil. Diss., Universität Regensburg. Regensburg: Roderer.
- Schmidt, J. (1991). *Evaluation einer Psychosomatischen Klinik*. Phil. Diss., Albert-Ludwigs-Universität Freiburg i.Br., Frankfurt

a.M.: Verlag für Akademische Schriften.

- Schmidt, J.U. (1991). Selbstkonzeptskalen für die Diagnose von Temperamentsmerkmalen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 12, 215–229.
- Schmidt, J.U. & König, F. (1986). Untersuchungen zur Validität der revidierten Form des Freiburger Persönlichkeitsinventars FPI-R. *Diagnostica*, 3, 197–208.
- Schmidt-Atzert, L. (1989). Ein Fragebogen zur Erfassung emotional relevanter Alltagsereignisse. *Diagnostica*, 35, 354–358.
- Schmitt, T. & Gschwendner, T. (2006). Regeln der Datenintegration. In F. Petermann & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 383–395). Göttingen: Hogrefe.
- Schmitz, K. (2004). Standardisierung psychometrischer Diagnostik in der neurologischen Rehabilitation. Eine computergestützte neuropsychologische Testbatterie. Dissertation. Universität Konstanz, Mathematisch-Naturwissenschaftliche Sektion. Elektronische Publikation im Internet Promotion Date: 26.11. 2004. <http://kops.ub.uni-konstanz.de/volltexte/2005/1441/>
- Schmolck, P. (1981). Zwei allgemeine Faktoren des Antwortverhaltens bei Selbstbeschreibungsinventaren: Social Desirability und Acquiescence – Neu interpretiert. In W. Michaelis (Hrsg.), *Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich 1980*, (Band 2, S. 507–509). Göttingen: Hogrefe.
- Schneewind, K.A. & Graf, J. (1998). *Der 16-Persönlichkeits-Faktoren-Test. Revidierte Fassung*. Testmanual. Bern: Huber.
- Schneider, K. (1985). *Wissenschaftshistorische Perspektiven und Grundzüge einer biochemischen Persönlichkeitsforschung, dargestellt im Rahmen einer empirischen Untersuchung zum Zusammenhang zwischen Katecholaminstoffwechsel und Emotionalität*. Unveröff. Diss., Universität Hamburg.
- Schneider-Düker, M. & Schneider, J.F. (1977). Untersuchungen zum Beantwortungsprozeß bei psychodiagnostischen Fragebogen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 24, 282–302.
- Schorr, A. (1995). Stand und Perspektiven diagnostischer Verfahren in der Praxis. Ergebnisse einer repräsentativen Befragung westdeutscher Psychologen. *Diagnostica*, 41, 3–20.
- Schuerger, J.M., Tait, E. & Tavernelli, M. (1982). Temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology*, 43, 176–182.
- Schuhfried Ges.m.b.H. (1993). *Das Freiburger Persönlichkeitsinventar* (Version 3.00). Mödling: Dr.G. Schuhfried Ges.m.b.H.
- Schuler, H. (1992). Das multimodale Einstellungsinterview. *Diagnostica*, 38, 281–300.
- Schuler, H. & Funke, U. (Hrsg.). (1991). *Eignungsdiagnostik in Forschung und Praxis*. Göttingen: Verlag für Angewandte Psychologie.
- Schuler, H. & Schmitt, N. (1987). Multimodale Messung in der Personalpsychologie. *Diagnostica*, 33, 259–271.
- Schulz, W. (1978). *Untersuchung zur Klassifikation neurotischer Klienten und Zusammenhänge zu Prozeß- und Effektivvariablen in der klientenzentrierten Gesprächspsychotherapie. Ein Beitrag zur Indikationsfrage*. Unveröff. Diss., Technische Universität Berlin.
- Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In C. Hendrick & M.S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 98–119). Newbury Park, CA: Sage.
- Schwarz, N. (2007). Retrospective and concurrent self reports. The rationale for real-time data capture. In A.A. Stone, S. Shiffman, A.A. Atienza & L. Nebeling (Eds.), *The science of real time data capture. Self reports in health research* (pp. 11–26). New York: Oxford University Press.
- Schwarz, N. & Scheuring, B. (1992). Selbstberichtete Verhaltens- und Symptommhäufigkeiten: Was Befragte aus Antwortvorgaben des Fragebogens lernen. *Zeitschrift für Klinische Psychologie*, 21, 197–208.
- Schwarz, N. & Sudman, S. (Eds.). (1992). *Context effects in social and psychological research*. New York: Springer.
- Schwebcke, A., Lück, H.E. & Jandron, E.L. (1974). Einige Anmerkungen zu Shostrom's Personal Orientation Inventory (POI). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 22, 318–323.
- Schweizer, K. (1986). *Aggregation und Vorhersagbarkeit. Vom Nutzen des Aggregationsansatzes für die Persönlichkeitsforschung*. Rheinfelden: Schäuble.
- Schweizer, K. (1989). Eine Analyse der Konzepte, Bedingungen und Zielsetzungen von Replikationen. *Archiv für Psychologie*, 141,

- Schweizer, K. (1990). *Der Aggregationseffekt. Konsequenzen der Aggregation und der Disaggregation von Daten*. Frankfurt a.M.: Verlag Peter Lang.
- Schweizer, K. (Hrsg.). (1999). *Methoden für die Analyse von Fragebogendaten*. Göttingen: Hogrefe.
- Schwenkmezger, P. (1987). Freiburger Persönlichkeitsinventar (FPI). Revidierte Fassung FPI-R. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 8, 152–153.
- Schwenkmezger, P. & Hank, P. (1993). Papier-Bleistift versus computerunterstützte Darbietung von State-Trait Fragebogen: Eine Äquivalenzprüfung. *Diagnostica*, 39, 189–210.
- Schwenkmezger, P., Hodapp, V. & Spielberger, C.D. (2000). *Das State-Trait-Ärgerausdrucks-Inventar STAXI*. Göttingen: Hogrefe.
- Seidenstücker, G. (1985). Testbesprechung. Das Freiburger Persönlichkeitsinventar in der vierten Auflage. Was ist neu? Was bringt es dem Klinischen Psychologen. *Zeitschrift für Klinische Psychologie*, 14, 242–246.
- Seidenstücker, G. & Baumann, U. (1978). Multimodale Diagnostik. In U. Baumann, H. Berbalk & G. Seidenstücker (Hrsg.), *Klinische Psychologie: Trends in Forschung und Praxis*. Band 1 (S. 134–182). Bern: Huber.
- Seidenstücker, G. & Baumann, U. (1987). Multimodale Diagnostik als Standard in der Klinischen Psychologie. *Diagnostica*, 33, 243–258.
- Seiwald, B.B. (2002). Replicability and generalizability of Kubinger's results: Some more studies on faking personality inventories. *Psychologische Beiträge*, 44, 17–23.
- Seiwald, B.B. (2003). Antworttendenzen (response sets). In K.D. Kubinger & S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 29–32). Weinheim: Beltz.
- Selg, H. (1968). *Diagnostik der Aggressivität*. Göttingen: Hogrefe.
- Selg, H., Mees, U. & Berg, D. (1997). *Psychologie der Aggressivität* (2. Aufl.). Göttingen: Hogrefe.
- Semin, G.R., Rosch, E., Krolage, J. & Chassein, J. (1981). Alltagswissen als implizite Basis für „wissenschaftliche“ Persönlichkeitstheorien: Eine sozialpsychologische Untersuchung. *Zeitschrift für Sozialpsychologie*, 12, 233–242.
- Shostrom, E.F. (1964). An inventory for the measurement of self-actualization. *Educational Psychological Measurement*, 24, 207–217.
- Sieber, M.F. (1986). Persönlichkeitskorrelate des Drogenkonsums – Eine Kausalanalyse mit dem LISREL-Verfahren. In C. Möbus & W. Schneider (Hrsg.), *Strukturmodelle für Längsschnittdaten und Zeitreihen. LISREL, Pfad- und Varianzanalysen. Methoden der Psychologie*, (Band 5, S. 198–208). Bern: Huber.
- Sieber, M. & Bentler, P. (1982). Zusammenhänge zwischen Persönlichkeitsmerkmalen und späterem Konsum legaler und illegaler Drogen bei jungen Männern. Eine Längsschnittuntersuchung. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 29, 649–668.
- Siegrist, J., Dittmann, K., Rittner, K. & Weber, I. (1980). *Soziale Belastungen und Herzinfarkt*. Stuttgart: Enke.
- Silbereisen, R.K. & Schulz, W. (1977). Prüfung der Testgüte einer „Empathie-Skala“. *Diagnostica*, 23, 179–187.
- Soares, I., Machado, P.P., Dias, P., Pinho, A. & Klein, J. (2005). Inventário de Personalidade de Freiburg-Revisto-FPI-R: Estudo de validação junto de amostra de estudantes universitários. *International Journal of Clinical and Health Psychology*, 5, 319–333.
- Sokal, R.R. & Sneath, P.H. (1968). *Principles of numerical taxonomy*. San Francisco, CA: Freeman.
- Sommer, R., Unholzer, G. & Wiegand, E. (1999). *Standards zur Qualitätssicherung in der Markt- und Sozialforschung*. Frankfurt a.M.: ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.
- Sonquist, J.A., Baker, E.L. & Morgan, J.N. (1973). *Searching for structure*. Ann Arbor, Mich.: Univ. of Michigan.
- Statistisches Bundesamt (Hrsg.). (1982). *Statistisches Jahrbuch 1982 für die Bundesrepublik Deutschland*. Stuttgart: Kohlhammer.
- Steck, P. (1997). Psychologische Testverfahren in der Praxis. Ergebnisse einer Umfrage unter Testanwendern. *Diagnostica*, 43, 267–284.
- Stegmüller, W. (1973). *Aufgaben und Ziele der Wissenschaftstheorie. Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*. Band IV. Berlin: Springer.
- Steller, M. & Hunze, D. (1984). Zur Selbstbeschreibung von Delinquenten im Freiburger Persönlichkeitsinventar (FPI). Eine Sekundäranalyse empirischer Untersuchungen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 5, 87–109.

- Stemmler, G. (1992). *Differential psychophysiology: Persons in situations*. Berlin: Springer.
- Stemmler, G. (1996). Strategies and designs in ambulatory assessment. In J. Fahrenberg & M. Myrtek (Eds.). *Ambulatory Assessment: computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 257–268). Seattle, WA: Hogrefe & Huber.
- Stemmler, G. (2001). Grundlagen psychophysiologischer Methodik. In F. Rösler (Hg.). *Ergebnisse und Methoden der Psychophysiologie. Enzyklopädie der Psychologie*, Band 4, Serie I, Themenbereich C (S. 1–84). Göttingen: Hogrefe.
- Stemmler, G. & Meinhardt, E. (1991). Personality, situation and arousability. *Personality and Individual Differences*, 11, 293–308.
- Stieglitz, R.D. (2000). *Diagnostik und Klassifikation psychischer Störungen*. Göttingen: Hogrefe.
- Strack, F. & Schwarz, N. (2007). Asking questions: Measurement in the social sciences. In: M. Ash & T. Sturm (Eds.). *Psychology's territories. Historical and contemporary perspectives from different disciplines* (pp. 225–250). Mahwah, N.J.: Lawrence Erlbaum.
- Strecker, A. (1985). *Interpersonelle Wahrnehmung von Partnern*. Unveröff. Diplomarbeit mit Anhang. Universität Freiburg.
- Stumpf, H., Angleitner, A., Wieck, T., Jackson, D.N. & Beloch-Till, H. (1985). *Deutsche Personality Research Form (PRF)*. Göttingen: Hogrefe.
- Tauss, R. (1992). *Die Veränderung von Selbstkonzeptkomponenten im Inhaftierungsverlauf jugendlicher Strafgefangener. Ergebnisse einer wissenschaftlichen Begleitung des sozialtherapeutischen Modells in der Jugendstrafanstalt Berlin-Plötzensee*. Phil. Diss. Albert-Ludwigs-Universität Freiburg i.Br.
- Technical Standards for Translating and Adapting Tests and Establishing Test Score Equivalence 2000 (Zugriff 1.3.2009)  
<http://www.testpublishers.org/journal1.htm>
- Testkuratorium (2007). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen. *Psychologische Rundschau*, 58, 25–30.
- Testkuratorium der Föderation Deutscher Psychologenvereinigungen (1986a). Richtlinien für den Einsatz und den Vertrieb psychologischer Testverfahren. *Psychologische Rundschau*, 37, 162–163.
- Testkuratorium der Föderation Deutscher Psychologenvereinigungen (1986b). Richtlinien für den Einsatz elektronischer Datenverarbeitung in der psychologischen Diagnostik. *Psychologische Rundschau*, 37, 163–165.
- Teusch, L. & Böhme, H. (1991). Was bewirkt ein stationäres Behandlungsprogramm mit gesprächspsychotherapeutischem Schwerpunkt bei Patienten mit Agoraphobie und/oder Panik? Ergebnis einer 1Jahres-Katamnese. *Psychotherapie, Psychosomatik und medizinische Psychologie*, 41, 68–76.
- Teusch, L., Böhme, H. & Finke, J. (2001). Konfliktzentrierte Monotherapie oder Methodenintegration? Veränderungsprozesse von Gesprächspsychotherapie mit und ohne verhaltenstherapeutische Reizkonfrontation bei Agoraphobie mit Panikstörung. *Der Nervenarzt*, 72 (1), 31–39.
- Thomae, H. (1968). *Das Individuum und seine Welt*. Göttingen: Hogrefe.
- Timm, U. (1968). Reliabilität und Faktorenstruktur von Cattell's 16 PF in einer deutschen Stichprobe. *Zeitschrift für experimentelle und angewandte Psychologie*, 15, 354–373.
- Timm, U. (1971). Multiple Regression von Verhaltenseinstufungen (Bekanntenratings) auf FPI-Skalen. *Diagnostica*, 17, 26–45.
- Tong E.M. W, Bishop G.D. & Enkelmann H.C. (2006). The role of the Big Five in appraisals. *Personality and Individual Differences*, 41 (3), 513–523.
- Tränkle, U. (1983). Fragebogenkonstruktion. In H. Feger & J. Bredenkamp (Hrsg.), *Datenerhebung, Enzyklopädie der Psychologie. Forschungsmethoden*, (Band 2, S. 222–301). Göttingen: Hogrefe.
- Ullrich, S., Draba, S., Pillmann, F., Sannemüller, U. & Marneros, A. (1999). Täterpersönlichkeit und soziobiographischer Hintergrund. Eine empirische Analyse begutachteter Angeklagter. *Monatsschrift für Kriminologie und Strafrechtsreform*, 82 (4), 291–298.
- Vagt, G. & Wendt, W. (1978). Akquieszenz und die Validität von Fragebogenskalen. *Psychologische Beiträge*, 20, 428–439.
- Van de Vijver, F. & Hambleton, R.K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89–99.
- Vazire, S. & Mehl, M.R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95, 1202–1216.

- Viswesvaran, C., Deller, J. & Ones, D. (2007). Personality measures in personnel selection: Some new contributions. *International Journal of Selection and Assessment*, 15 (3), 354–358.
- Volkers, A.C., Tullen, J.H.M. et al. (2002). Effect of personality dimensions on the diurnal pattern of motor activity. *Journal of Personality*, 70, 233–247.
- Vollmer, H.C., Seydel, F. & Ferstl, R. (1991). Psychischer Befund und Therapieverlauf bei HIV-positiven Drogenabhängigen. *Sucht*, 37, 361–368.
- Voyce, C.D. & Jackson, D.N. (1977). An evaluation of a threshold theory for personality assessment. *Educational and Psychological Measurement*, 37, 383–408.
- Walkey, F.H. & Green, D.E. (1981). The structure of the Eysenck personality inventory: A comparison between simple and more complex analyses of a multiple scale questionnaire. *Multivariate Behavioral Research*, 16, 361–372.
- Walter, P. (1999). Die „Vermessung des Menschen“: Meßtheoretische und methodologische Grundlagen psychologischen Testens. In S. Grubitzsch (Hrsg.), *Testtheorie – Testpraxis: psychologische Tests und Prüfverfahren im kritischen Überblick* (2. Aufl.). (S. 98–127). Eschborn: Klotz.
- Warncke, P. & Fahrenberg, J. (1966). Eine Itemanalyse am E.N. NR-Fragebogen von Breugelman und Breugelman. *Diagnostica*, 12, 105–115.
- Watzl, H. & Rist, F. (1982). Katamnestic Erfahrungen mit einem Therapieprogramm für alkoholabhängige Frauen. In D. Ladwig (Hrsg.), *Drogen und Alkohol 2. Erfahrungen und Ergebnisse in der Behandlung Drogen- und Alkoholabhängiger* (S. 42–54). Basel: Karger.
- Weber, H. & Rammsayer, T. (Hrsg.). (2005). *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie*. Göttingen: Hogrefe.
- Welsch, M. (1986). *Multivariate Erfassung von Rehabilitationsverläufen*. Unveröff. Diss., Universität Freiburg.
- Westhoff, G. (1993). *Handbuch psychosozialer Meßinstrumente. Ein Kompendium für epidemiologische und klinische Forschung zu chronischer Krankheit*. Göttingen: Hogrefe.
- Westhoff, K. (2004). Die Eignungsbeurteilung. In K. Westhoff et al. (Hrsg.), *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430* (S. 201–206). Lengerich: Pabst.
- Westhoff, K., Hagemester, C. & Strobel, A. (2006). Psychologische Begutachtung. In F. Petermann & M. Eid (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 396–406). Göttingen: Hogrefe.
- Westhoff, K., Hellfritsch, L.J., Hornke, L.F., Kubinger, K.D., Lang, F., Moosbrugger, H., Püschel, A. & Reimann, G. (Hrsg.). (2004). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. Lengerich: Pabst.
- Westhoff, K. & Kluck, M.L. (2003). *Psychologische Gutachten schreiben und beurteilen* (4. Aufl.). Berlin: Springer.
- Westmeyer, H. (1995). Persönlichkeitspsychologie zwischen Realismus und Konstruktivismus. In K. Pawlik (Hrsg.), *Bericht über den 39. Kongreß der Deutschen Gesellschaft für Psychologie in Hamburg 1994* (S. 748–753). Göttingen: Hogrefe.
- Wicki, W. & Angst, J. (1992). Funktionelle Magen- und Darmbeschwerden bei jungen Erwachsenen: Vorkommen, Verlauf, Persönlichkeit und psychosoziale Faktoren. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 42, 371–380.
- Wietersheim, J. von, Malewski, W., Jäger, B., Köpp, W., Gitzinger, I., Köhler, P. & Grabhorn, R. (2001). Der Einfluß von stationärer dynamischer Psychotherapie auf Persönlichkeitsmerkmale von Patientinnen mit Anorexia nervosa und Bulimia nervosa – Ergebnisse der multizentrischen Eßstörungsstudie. *Zeitschrift für psychosomatische Medizin und Psychotherapie*, 47, 366–379.
- Wiggins, J.S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, Mass.: Addison-Wesley.
- Wilhelm, P. (2004). *Empathie im Alltag von Paaren. Akkuratheit und Projektion bei der Einschätzung des Befindens des Partners*. Bern: Huber.
- Wilhelm, P. & Perez, M. (2008). Ambulantes Assessment in der Klinischen Psychologie und Psychiatrie. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 56 (3), 169–179.
- Wissenschaftlicher Beirat Psychotherapie WBP (2007). Methodenpapier des WBP – Verfahrensregeln zur Beurteilung der wissenschaftlichen Anerkennung von Methoden und Verfahren der Psychotherapie. Version 2.6 vom 21.11.2007. <http://www.wbpscho-therapie.de/page.asp?his=0.1.78>
- Wittmann, W.W. (1977). *Faktorenanalytische Modelle, Methodenstudien und Probleme der Reproduzierbarkeit*. Phil. Diss., Freiburg.

- Wittmann, W.W. (1987). Grundlagen erfolgreicher Forschung in der Psychologie: Multimodale Diagnostik, Multiplismus, multivariate Reliabilitäts- und Validitätstheorie. *Diagnostica*, 33, 209–226.
- Wittmann, W.W. (1988). Multivariate reliability theory. Principles of symmetry and successful validation strategies. In J.R. Nesselroade & R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2<sup>nd</sup> ed.). (pp. 505–560). New York: Plenum.
- Wittmann, W.W. (2002). Brunswik-Symmetrie: Ein Schlüsselkonzept für erfolgreiche psychologische Forschung. In M. Myrtek (Hrsg.), *Die Person im biologischen und sozialen Kontext* (S. 163–186). Göttingen: Hogrefe.
- Wittmann, W.W. & Hampel, R. (1976). Faktorenanalyse, Rotationsproblem und Implikationen für die Testkonstruktion – Bemerkungen zu einem optimalen Rotationsverfahren anhand eines empirischen Datensatzes (FPI). *Archiv für Psychologie*, 128, 219–239.
- Wittmann, W.W. & Klumb, P.L. (2006). How to fool yourself with experiments in testing theories in psychological research. In R.R. Bootzin & P.E. McKnight (Eds.), *Strengthening research methodology: Psychological measurement and evaluation* (pp. 185–211). Washington, DC: American Psychological Association.
- Wittmann, W.W. & Schmidt, J. (1983). *Die Vorhersagbarkeit des Verhaltens aus Traitinventaren*. Forschungsberichte des Psychologischen Instituts der Universität Freiburg, Nr. 10.
- Wolfram, M. & Moltz, A. (1974). Zur Brauchbarkeit des Freiburger Persönlichkeitsinventars (FPI) für die Neurosendiagnostik. In J. Helm, E. Kasielke & J. Mehl (Hrsg.), *Neurosendiagnostik*. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Womes, E. (1977). *Eine Analyse des Konzeptes der Aggressivität im FPI mit dem Rasch-Modell*. Unveröff. Diplomarbeit. Bonn: Psychologisches Institut.
- Woodworth, R.S. (1918). *Personal Data Sheet*. Chicago: Stoelting.
- Wundt, W. (1902–1903). *Grundzüge der Physiologischen Psychologie*. Band 1–3 (5. Aufl.). Leipzig: Engelmann.
- Wundt, W. (1907). Über Ausfrageexperimente und über die Methoden zur Psychologie des Denkens. *Psychologische Studien*, 3, 301–360.
- Wundt, W. (1908). Kritische Nachlese zur Ausfragemethode: *Archiv für die gesamte Psychologie*, 11, 445–459.
- Wundt, W. (1921). *Logik. Eine Untersuchung der Prinzipien der Erkenntnis und der Methoden Wissenschaftlicher Forschung*. Band 3. Logik der Geisteswissenschaften (4. Aufl.). Stuttgart: Ferdinand Enke.
- Yousfi, S. & Steyer, R. (2006). Messtheoretische Grundlagen der Psychologischen Diagnostik. In F. Petermann & M. Eid, M. (Hrsg.), *Handbuch der psychologischen Diagnostik* (S. 46–56). Göttingen: Hogrefe.
- Zielke, M. (1993). *Wirksamkeit stationärer Verhaltens-therapie*. Weinheim: Beltz, Psychologie Verlags-Union.
- Zuckerman, M., Kuhiman, D.M., Thornquist, M. & Kiers, H. (1991). Five (or three) robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, 12, 929–941.



# Literatur-Recherche

## Aktuelle Nachweise zum FPI in den Datenbanken PSYINDEX und PSYTKOM

(mit freundlicher Unterstützung von Herrn Dipl.-Psych. Jürgen Wiesenhütter, ZPID, Trier)

Die Datenbank PSYINDEX (herausgegeben vom Zentrum für Psychologische Information und Dokumentation ZPID) enthält ca. 1400 Nachweise zu Untersuchungen, die das FPI angewendet haben. Sie wird monatlich um aktuelle Nachweise ergänzt.

Neben der Literaturdatenbank PSYINDEX gibt die ZPID auch eine Testdatenbank PSYINDEX Tests heraus, in der ausführliche (und immer wieder aktualisierte) Einträge zum FPI und zum FPI-R zusammengefasst werden.

PSYINDEX und PSYINDEX Tests sind an nahezu allen Fachbibliotheken im deutschsprachigen Raum zugänglich. Überdies werden sie (meist zusammengefasst zu einer einzigen Datenbank) bei verschiedenen fachspezifischen und fachübergreifenden Datenbank-Hosts geführt. Einen individuellen Zugang über das Internet ermöglicht das ZPID mit „PSYINDEX Direct“ für einen Jahresabonnementspreis von 50,- Euro bzw. für eine Tagespauschale von 5,- Euro (Mitglieder vieler Fachverbände erhalten Rabatt).

Die Nachweise zum Freiburger Persönlichkeitsinventar werden in PSYINDEX aufgefunden durch Eingabe von FPI. Gesucht wird ohne Feldeinschränkung (Expertensuche oder einfache Suche) bzw. in „Alle Felder“ (Feldsuche). Versionen wie FPI-R werden automatisch mitgesucht, können aber auch speziell angesprochen werden (Je nach Suchsystem entweder mit FPI R oder FPI Leerzeichen R).

Vor allem die „Expertensuche“ (Advanced Search) bietet viele Möglichkeiten zur Spezifizierung der Suche. Über die Operatoren AND, OR oder NOT kann der Kernbegriff mit inhaltlichen Begriffen verknüpft werden. Dem Operator AND kommt eine besondere Bedeutung zu, denn er bildet die Schnittmenge zwischen zwei Begriffskomplexen, so dass eine Selektion auf bestimmte Anwendungsbereiche oder Teilthemen erfolgen kann. Der Operator OR bildet die Vereinigungsmenge, mit dem Operator NOT können Bereiche ausgeschlossen werden.

Bei der Eingabe der Begriffe kann der Asterisk (\*) als „Joker“ eingesetzt werden, um verschiedene Wortvarianten gleichzeitig zu erfassen. Die Begriffe können in deutscher oder in englischer Sprache eingegeben werden.

Beispiele:

- FPI and psychosomati\*
- FPI and (anxiety or fear or phobia\*)
- FPI and (sucht\* or drogenabh\* or alkohol\*)

[Gesucht werden Anwendungen des FPI in verschiedenen Bereichen]

oder auch:

- FPI and (test construction OR test validity)

[Gesucht werden Untersuchungen, in denen der FPI zur Konstruktion und Validierung von anderen Tests eingesetzt wurde, sowie zur Validität des FPI selber]

Die verschiedenen Varianten der Datenbanken enthalten vielfältige weitere Möglichkeiten, eine Suche noch zu präzisieren. Näheres ist den jeweiligen Hilfefunktionen der Datenbank zu entnehmen oder bei der ZPID selbst zu erfragen. Die Adresse lautet:

- ZPID (Zentrum für Psychologische Information und Dokumentation)  
Universität Trier, Universitätsring 15  
54286 Trier  
E-Mail: [info@zpid.de](mailto:info@zpid.de)  
Tel: 0651/2012876  
FAX: 0651/2012871  
Internet: [www.zpid.de](http://www.zpid.de)

## Probanden-Gruppen der Informationsbank<sup>1</sup>

Bis einschließlich der 3. Auflage des FPI wurde in der Handanweisung eine Liste der Probandengruppen der Informationsbank fortgeschrieben, so dass die Testbenutzer eine Übersicht über mögliche Vergleichsdaten hinsichtlich der Formen G, A, B und K hatten. Die folgende Liste enthält nur die Probanden-Gruppen, von denen FPI-R Daten vorliegen und aufgrund ausdrücklicher Autorisierung für wissenschaftliche Zwecke in anonymisierter Form weitergegeben werden können. Diese Datensätze sind auf einer CD erhältlich.<sup>1</sup>

<sup>1</sup> Zuschriften an Prof. Dr. Jochen Fahrenberg, Freiburg: jochen.fahrenberg@psychologie.uni-freiburg.de

Die Probanden-Gruppen R sind nach folgendem Schlüssel gekennzeichnet: Geschlecht (1=männlich, 2=weiblich); laufende Nr. der Probanden-Gruppe.

FPI-R Männer			
<b>R 1.01</b>	Normstichprobe	16–24 Jahre	N = 174
<b>R 1.02</b>	Normstichprobe	25–44 Jahre	N = 384
<b>R 1.03</b>	Normstichprobe	45–59 Jahre	N = 215
<b>R 1.04</b>	Normstichprobe	60 Jahre und älter	N = 180
<b>R 1.05</b>	Studenten		N = 79
<b>R 1.06</b>	Studenten		N = 31
<b>R 1.07</b>	Studenten		N = 54
<b>R 1.08</b>	Patienten mit Herzinfarkt u.a. HK-Krankheiten (1.Termin)		N = 123
	(2.Termin)		N = 109
<b>R 1.09</b>	Durchschnittsbevölkerung		N = 59
<b>R 1.10</b>	Durchschnittsbevölkerung		N = 42
<b>R 1.11</b>	Patienten mit Herzinfarkt u.a. HK-Krankheiten		N = 51
<b>R 1.12</b>	Alkohol- und Medikamentenabhängige		N = 71
<b>R 1.13</b>	Durchschnittsbevölkerung		N = 164
<b>R 1.14</b>	Patienten in Psychotherapie		N = 18
<b>R 1.15</b>	Patienten mit Herzinfarkt		N = 63
<b>R 1.16</b>	Durchschnittsbevölkerung		N = 110
<b>R 1.17</b>	Patienten einer Psychosomatischen Ambulanz		N = 228
<b>R 1.18</b>	Hypertoniker/Normotoniker		N = 55
<b>R 1.19</b>	Patienten mit Morbus Crohn		N = 76
<b>R 1.20</b>	Patienten einer Psychosomatischen Ambulanz		N = 147
<b>R 1.21</b>	Depressive Patienten		N = 109
<b>R 1.22</b>	Teilnehmer einer Arzneimittelprüfung		N = 40
<b>R 1.23</b>	Kurzsichtige		N = 51
<b>R 1.24</b>	Patienten einer Rehaklinik		N = 55
<b>R 1.25</b>	Patienten einer Psychiatrischen Praxis		N = 48
<b>R 1.26</b>	Normstichprobe 1999		N = 1743

FPI-R Frauen			
<b>R 2.01</b>	Normstichprobe	16–24 Jahre	N = 169
<b>R 2.02</b>	Normstichprobe	25–44 Jahre	N = 382
<b>R 2.03</b>	Normstichprobe	45–59 Jahre	N = 233

<b>R 2.04</b>	Normstichprobe	60 Jahre und älter	N = 298
<b>R 2.05</b>	Studentinnen		N = 156
<b>R 2.06</b>	Studentinnen		N = 63
<b>R 2.07</b>	Studentinnen		N = 65
<b>R 2.08</b>	Alkohol- und Medikamentenabhängige		N = 33
<b>R 2.09</b>	Durchschnittsbevölkerung		N = 58
<b>R 2.10</b>	Durchschnittsbevölkerung		N = 76
<b>R 2.11</b>	Durchschnittsbevölkerung		N = 279
<b>R 2.12</b>	Patientinnen mit Essstörungen		N = 60
<b>R 2.13</b>	Patientinnen in Psychotherapie		N = 42
<b>R 2.14</b>	Durchschnittsbevölkerung		N = 121
<b>R 2.15</b>	Patientinnen einer Psychosomatischen Ambulanz		N = 319
<b>R 2.16</b>	Patientinnen mit Morbus Crohn		N = 88
<b>R 2.17</b>	Patientinnen einer Psychosomatischen Ambulanz		N = 292
<b>R 2.18</b>	Depressive Patientinnen		N = 236
<b>R 2.19</b>	Teilnehmer einer Arzneimittelprüfung		N = 37
<b>R 2.20</b>	Kurzsichtige		N = 58
<b>R 2.21</b>	Patientinnen einer Rehaklinik		N = 50
<b>R 2.22</b>	Patientinnen einer psychiatrischen Praxis		N = 22
<b>R 2.23</b>	Normstichprobe 1999		N = 1997

# INSTITUT FÜR DEMOSKOPIE ALLENSBACH



Sehr geehrte Dame, sehr geehrter Herr,

annähernd 75.000 Interviews führt das Institut für Demoskopie Allensbach jährlich durch; in über 70 Jahren wurden mehrere Millionen Fragen beantwortet, Fragen aus allen Bereichen des Lebens: Politik und Wirtschaft, Fragen zu kulturellen Ereignissen, über das Berufsleben, über Moralvorstellungen, Erziehungsfragen und vieles andere mehr.

All diese Umfragen wären sicher nicht möglich, wenn unsere Interviewer nicht überall aufgeschlossene, auskunftsbereite Menschen finden würden, die unsere Arbeit unterstützen. Im Vertrauen darauf möchten wir Sie jetzt noch bitten, einen Zusatzbogen selbst auszufüllen, der von uns in Zusammenarbeit mit der Universität Freiburg ausgewertet wird.

Wir haben uns bemüht, Themen zu finden, die auch Sie interessieren dürften. So kann diese kleine Mühe, die Sie auf sich nehmen, Ihnen sicherlich auch Freude bereiten, dort nämlich, wo Sie Ihre Meinung äußern können zu Themen, bei denen Sie vielleicht noch nie zuvor die Gelegenheit dazu hatten.

Wir danken Ihnen für Ihre Mitarbeit, die zum Gelingen dieser Untersuchung beiträgt.

Mit freundlichen Grüßen  
INSTITUT FÜR DEMOSKOPIE ALLENSBACH

## FPI-R

--	--	--	--

Sie werden auf den folgenden Seiten eine Reihe von Aussagen über bestimmte Verhaltensweisen, Einstellungen und Gewohnheiten finden. Sie können jede entweder mit „s t i m m t“ oder mit „s t i m m t n i c h t“ beantworten. Setzen Sie bitte ein Kreuz (X) in den dafür vorgesehenen Kreis. Es gibt keine richtigen oder falschen Antworten, weil jeder Mensch das Recht zu eigenen Anschauungen hat. Antworten Sie bitte so, wie es für sie zutrifft.

Beachten Sie bitte folgende Punkte:

- Überlegen Sie bitte nicht erst, welche Antwort vielleicht den „besten Eindruck“ machen könnte, sondern antworten Sie so, wie es für Sie persönlich gilt. Manche Fragen kommen Ihnen vielleicht sehr persönlich vor. Bedenken Sie aber, dass Ihre Antworten unbedingt vertraulich behandelt werden.
- Denken Sie nicht lange über einen Satz nach, sondern geben Sie die Antwort, die Ihnen unmittelbar in den Sinn kommt. Natürlich können mit diesen kurzen Fragen nicht alle Besonderheiten berücksichtigt werden. Vielleicht passen deshalb einige nicht gut auf Sie. **Kreuzen Sie aber trotzdem immer eine Antwort an**, und zwar die, welche noch am ehesten für Sie zutrifft.

	stimmt	stimmt nicht
5. Ich glaube, dass ich mir beim Arbeiten mehr Mühe gebe als die meisten anderen Menschen.....	<input type="radio"/>	<input type="radio"/>
6. Ich scheue mich, allein in einen Raum zu gehen, in dem andere Leute bereits zusammensitzen und sich unterhalten.....	<input type="radio"/>	<input type="radio"/>
7. Manchmal bin ich zu spät zu einer Verabredung oder zur Schule gekommen.....	<input type="radio"/>	<input type="radio"/>
8. Ich würde mich beim Kellner oder Geschäftsführer eines Restaurants beschweren, wenn ein schlechtes Essen serviert wird.....	<input type="radio"/>	<input type="radio"/>
9. Ich habe manchmal hässliche Bemerkungen über andere Menschen gemacht.....	<input type="radio"/>	<input type="radio"/>
.....		
.....		
.....		

	stimmt	stimmt nicht
134. Ich habe schon unbezahlt in meiner Gemeinde oder in sozialen Einrichtungen wie dem Roten Kreuz geholfen.....	<input type="radio"/>	<input type="radio"/>
135. Ich bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde.....	<input type="radio"/>	<input type="radio"/>
136. Ich nehme mir viel Zeit, anderen Menschen geduldig zuzuhören, wenn sie von ihren Sorgen erzählen.....	<input type="radio"/>	<input type="radio"/>
137. Es gab Leute, die mich so ärgerten, dass es zu einer tätlichen Auseinandersetzung kam.....	<input type="radio"/>	<input type="radio"/>
138. Meistens blicke ich voller Zuversicht in die Zukunft.....	<input type="radio"/>	<input type="radio"/>

Beim Beantworten dieses Fragebogens haben Sie vielleicht einige der Fragen zunächst zurückgestellt. Bitte überprüfen Sie deshalb noch einmal, ob Sie wirklich alle Fragen beantwortet haben.

**Bei den folgenden Fragen geht es um Ihre Meinung zu aktuellen Themen:**

	stimmt	stimmt nicht
1. Ich bin überzeugt, dass wir für die Umweltprobleme unserer Zeit vernünftige Lösungen finden können .....	<input type="radio"/>	<input type="radio"/>
2. Umweltverschmutzung, Kohlekraftwerke und Atommüll werden unser Leben immer weiter vergiften und ruinieren .....	<input type="radio"/>	<input type="radio"/>
3. Es gibt Personen außerhalb meines/unseres Haushalts, für die ich regelmäßig spende oder denen ich aktiv helfe, z.B. bei Besorgungen, kleineren Arbeiten oder der Betreuung von Kindern oder Kranken .....	<input type="radio"/>	<input type="radio"/>
4. Ich finde, dass bei gleicher Qualifikation genauso viele Frauen wie Männer in Führungspositionen kommen sollen .....	<input type="radio"/>	<input type="radio"/>
5. Jungen und Mädchen sollen nach denselben Grundsätzen erzogen werden .....	<input type="radio"/>	<input type="radio"/>
6. Ich denke, in Deutschland werden immer mehr Menschen an den Rand der Gesellschaft gedrängt .....	<input type="radio"/>	<input type="radio"/>
7. Es beschäftigt mich sehr, wenn andere ungerecht behandelt werden .....	<input type="radio"/>	<input type="radio"/>
8. Wenn jemand weint, nimmt mich das sehr mit und ich werde ebenfalls traurig .....	<input type="radio"/>	<input type="radio"/>
9. Was der heutigen Jugend am meisten fehlt, sind Disziplin, Fleiß und Pflichtgefühl .....	<input type="radio"/>	<input type="radio"/>
10. Ich bin stolz, eine Deutsche/ein Deutscher zu sein .....	<input type="radio"/>	<input type="radio"/>
11. Ich finde, wir sollten uns als Deutsche mit der Ermordung der Juden im Holocaust auseinandersetzen .....	<input type="radio"/>	<input type="radio"/>
12. Wir sollten wieder eine starke Partei haben, die wirklich die Interessen aller Schichten vertritt .....	<input type="radio"/>	<input type="radio"/>
13. In Krisenzeiten sollen zuerst die ausländischen Arbeiter entlassen werden .....	<input type="radio"/>	<input type="radio"/>
14. Ich hoffe auf einen starken Regierungschef, der die Flüchtlingsprobleme energisch löst und die Rückführung in weitgehend sichere Länder durchsetzt .....	<input type="radio"/>	<input type="radio"/>
15. Ich bin für die Entwicklung einer multikulturellen Gesellschaft in Deutschland (verschiedene Lebensformen, Religionen usw.) .....	<input type="radio"/>	<input type="radio"/>
16. Eine multikulturelle Gesellschaft wird dazu führen, dass die Deutschen ihre kulturelle und religiöse Identität verlieren .....	<input type="radio"/>	<input type="radio"/>
17. Ich habe bereits an Demonstrationen teilgenommen mit negativen Folgen für mich .....	<input type="radio"/>	<input type="radio"/>
18. Ich lehne Gewalt aus Überzeugung ab .....	<input type="radio"/>	<input type="radio"/>
19. Ich gehe gern auf Reisen, um an anderen Lebensweisen und Kulturen teilzunehmen und diese kennenzulernen .....	<input type="radio"/>	<input type="radio"/>
20. In meiner Freizeit lebe ich spontan und plane nicht meinen Tageslauf .....	<input type="radio"/>	<input type="radio"/>
21. Ich gehe lieber allein meinen Interessen und Hobbys nach als mit anderen Menschen zusammen zusein .....	<input type="radio"/>	<input type="radio"/>
22. Ich habe bereits erfahren, dass Gott mir in einer konkreten Situation geholfen hat .....	<input type="radio"/>	<input type="radio"/>
23. Ich glaube nicht an Gott .....	<input type="radio"/>	<input type="radio"/>
24. Echte Wunderheilungen wie in Lourdes (oder in anderen Wallfahrtsorten) können auch bei schweren und chronischen Krankheiten tatsächlich vorkommen .....	<input type="radio"/>	<input type="radio"/>
25. Aus fachkundig gestellten Horoskopen können zutreffende Aussagen über die Eigenart eines Menschen und gelegentlich sogar Vorhersagen über künftige Lebensereignisse (Schicksal) gewonnen werden .....	<input type="radio"/>	<input type="radio"/>
26. In extremen Fällen kann es sinnvoll sein, dass ein erfahrender Geistlicher einen Exorzismus (Ritual zur Unterbindung dämonischer Einflüsse/Teufelsaustreibung) vornimmt .....	<input type="radio"/>	<input type="radio"/>
27. Das Leben hat keinen tieferen philosophischen oder religiösen Sinn. Der Sinn des Lebens besteht darin, das Leben zu genießen .....	<input type="radio"/>	<input type="radio"/>
28. Ich glaube an eine Auferstehung und ewiges Leben nach dem Tod .....	<input type="radio"/>	<input type="radio"/>

29. Stufen Sie sich bitte auf der folgenden 10-stufigen Skala ein !

Nicht religiös 1	2	3	4	5	6	7	8	9	Religiös 10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

30. Sind Sie an der Frage nach dem Sinn des Lebens interessiert?

Nicht interessiert 1	2	3	4	5	6	7	8	9	Interessiert 10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

31. Viele Deutsche haben gegenwärtig verschiedene Sorgen und Ängste. Wie ist das bei Ihnen?

Ich habe Angst vor ...	Stimme gar nicht zu 1	2	3	4	Stimme völlig zu 5
eigener schwerer Erkrankung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
eigener Altersarmut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schadstoffen in Nahrungsmitteln	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Überfremdung des Landes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
steigender Kriminalität	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anstieg des Rechtsradikalismus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Terroranschlägen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

32. Wie häufig haben Sie persönlich seit dem Alter von 16 Jahren Situationen erlebt, in denen jemand Sie körperlich angegriffen hat, Sie zum Beispiel geschlagen, geohrfeigt, an den Haaren gezogen, getreten oder mit einer Waffe oder einem Gegenstand bedroht hat?

Häufig.....☐

Gelegentlich .....☐

Nur einmal .....☐

Nie.....☐

33. Wie häufig haben Sie in Ihrer Freizeit Kontakt mit Freunden? Gemeint ist damit persönlicher, telefonischer und schriftlicher/Email Kontakt. Ist das ...

täglich oder mehrmals täglich .....☐

wöchentlich .....☐

einige Male im Monat.....☐

seltener .....☐

34. Treiben Sie leistungsorientierten Sport, also dass Sie auch an Wettkämpfen teilnehmen, Sportabzeichen erwerben usw., oder treiben Sie Sport eher ohne Leistungsgedanken, z.B. für Ihr Wohlbefinden, oder treiben Sie gar keinen Sport?"

Leistungsorientierten Sport.....☐

Sport ohne Leistungsgedanken.....☐

Gar kein Sport.....☐

35. Wie viel Zeit verbringen Sie im Durchschnitt pro Woche für Ihr wichtigstes Hobby? .....Stunden pro Woche

Habe kein Hobby.....☐

36. Welchen Betrag geben Sie als Spenden für soziale Zwecke durchschnittlich im Jahr aus? .....Euro

Spende nichts für soziale Zwecke .....☐

**Abschließend noch einige ergänzende Fragen zur beruflichen Belastung und zum Gesundheitszustand!**

- A. Üben Sie einen Beruf aus? Ja ..... ☐  
Nein..... ☐

Falls Sie einen Beruf ausüben: Stufen Sie bitte den **Grad ihrer beruflichen Belastung** ein –

durch körperliche Arbeit	gering	1	2	3	4	5	sehr stark
durch Tempo, Genauigkeit, Konzentration	gering	1	2	3	4	5	sehr stark
durch Verantwortung für Menschen und Planung	gering	1	2	3	4	5	sehr stark

- B. Stufen Sie bitte den **Grad Ihrer Zufriedenheit** ein –

mit Ihrer beruflichen Situation	sehr zufrieden	1	2	3	4	5	unzufrieden
mit Ihrer familiären Situation	sehr zufrieden	1	2	3	4	5	unzufrieden
mit Ihrer finanziellen Situation	sehr zufrieden	1	2	3	4	5	unzufrieden
mit Ihrer Wohnsituation	sehr zufrieden	1	2	3	4	5	unzufrieden

- C. Würden Sie, falls es finanziell machbar wäre, vorzeitig in Rente gehen? Ja ..... ☐  
Nein..... ☐

- D. Wie schätzen Sie Ihren Gesundheitszustand während der letzten Zeit ein?

Mein Gesundheitszustand ist	sehr gut	1	2	3	4	5	schlecht
-----------------------------	----------	---	---	---	---	---	----------

- E. Rauchen Sie? Ja ..... ☐  
Nicht mehr..... ☐  
Noch nie..... ☐

Falls Sie „Ja“ angekreuzt haben:  
Wie viele Zigaretten rauchen Sie täglich?

.....

- F. Körpergröße: ..... cm

- G. Körpergewicht: ..... kg

- H. Würden Sie sich selbst als übergewichtig bezeichnen? Ja ..... ☐  
Vielleicht..... ☐  
Nein..... ☐

- I. Haben Sie eine chronische Krankheit? Ja ..... ☐  
Nein..... ☐

Um welche handelt es sich?



J. Wieviel Krankenhausaufenthalte hatten Sie bisher (bei Frauen: ohne Entbindung)? ..... Wie

--	--

viele Operationen? ..... Wie

--	--

oft waren Sie zu einer Kur? ..... Wie

--	--

oft waren Sie während der letzten 12 Monate beim Arzt (ohne Zahnarzt)? .....

--	--

K. Sind oder waren Sie in Psychotherapie (seelischer Behandlung)?

Ja .....

☐

Nein.....

☐

L. Haben Sie während der letzten 12 Monate irgendwelche der folgenden Mittel eingenommen?

	Nie	Etwa 2-mal im Jahr	Etwa 2-mal im Monat	Etwa 3-mal pro Woche	Fast täglich
Schlaftabletten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beruhigungstabletten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schmerztabletten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Homöopathische Mittel	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Bitte überprüfen Sie noch einmal, ob Sie wirklich alle Fragen beantwortet haben. Vielen Dank für Ihre Mühe!**

## Auszug aus den statistischen Ermittlungen des IFD

STATISTIK: "Zum Schluss möchte ich Sie um einige statistische Angaben bitten!"

1.	Geschlecht: MÄNNLICH.....1 WEIBLICH .....2	15. T	Berufsstellung des Befragten: bei Nichtberufstätigen: Beruf des Hauptverdieners; bei Rentnern: ehemalige Berufsstellung:
2. ALTER: T	..... JAHRE		Facharbeiter mit abgelegter Prüfung.....1 Sonstiger Arbeiter.....2 Landarbeiter.....3 Landwirt.....4 Leitender Angestellter.....5 Nichtleitender Angestellter.....6 Beamter des höheren od. gehobenen Dienstes..7 Beamter des mittleren od. einfachen Dienstes.....8 Inhaber und Geschäftsführer von größeren Unternehmen, Direktoren.....9 Mittlere und kleine selbständige Geschäftsleute Selbständige Handwerker.....0 Freier Beruf.....1 OHNE BERUF.....2
3.	Eine Frage zum Schulabschluss: Könnten Sie nach dieser Liste sagen, was auf Sie zutrifft, welche Nummer? (INTERVIEWER überreicht <b>grüne</b> Liste S!)  / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 /		
4. T	Sind Sie berufstätig? JA.....1 JA, mithelfend im eigenen Betrieb.....2 In Berufsausbildung.....3 Arbeitslos.....4 Rentner, (Vor-)Ruhestand.....5 Hausfrau, Hausmann.....6 Schüler.....7 Student.....8 OHNE BERUF.....9		

6. T    Leben Sie in einem Mehrpersonenhaushalt oder allein?	MEHRPERSONENHAUSHALT .....1 ALLEIN .....2
7. T    Familienstand:	VERHEIRATET zusammenlebend.....1 VERHEIRATET getrenntlebend.....2*  LEDIG .....3* VERWITWET .....4* GESCHIEDEN.....5*  JA.....6 NEIN .....7
*    Leben Sie mit einem Partner, einem Lebensgefährten zusammen?	

L I S T E S

**Welchen Schulabschluss haben Sie?**

(Schüler geben bitte den angestrebten Abschluss an!)

- (1) Ich bin von der Schule abgegangen ohne Hauptschulabschluss/Volksschulabschluss
- (2) Ich habe den Hauptschulabschluss/Volksschulabschluss
- (3) Ich bin von der Realschule, polytechnischen Oberschule oder einer vergleichbaren Schule abgegangen ohne Realschulabschluss, ohne Mittlere Reife
- (4) Ich habe den Realschulabschluss (Mittlere Reife, Abschluss der 10-klassigen polytechnischen Oberschule)
- (5) Ich habe die Fachhochschulreife
- (6) Ich habe die allgemeine oder fachgebundene Hochschulreife (Abitur, Fachabitur, Abschluss der 12-klassigen EOS)
- (7) Ich habe ein Studium an einer Universität, Fachhochschule oder Pädagogischen Hochschule begonnen, aber nicht abgeschlossen
- (8) Ich habe ein Studium an einer Universität, Fachhochschule oder Pädagogischen Hochschule abgeschlossen

### Verhaltensratings durch den IfD-Interviewer nach dem mündlichen Interview:

1.	Soziale Schicht des/der Befragten –	A-SCHICHT (OBERSTE SCHICHT)..... 1 B-SCHICHT (OBERE MITTELSCHICHT) ..... 2 C-SCHICHT (BREITE MITTELSCHICHT) ..... 3 D-SCHICHT (UNTERSTE SCHICHT) ..... 4
2.	Der/die Befragte war während des Interviews –	SEHR SELBSTSICHER, VON SICH ÜBERZEUGT..... 1 ZIEMLICH SELBSTSICHER ..... 2 ETWAS UNSICHER ..... 3 ZIEMLICH UNSICHER ..... 4
3.	Der/die Befragte wirkt –	SEHR ENERGISCH ..... 1 ENERGISCH ..... 2 WENIGER ENERGISCH ..... 3 GAR NICHT ENERGISCH ..... 4
4.	Der/die Befragte ist –	SEHR FREUNDLICH ..... 1 FREUNDLICH ..... 2 NICHT SO FREUNDLICH ..... 3 EHER UNFREUNDLICH ..... 4
5.	Der/die Befragte wirkt auf mich –	SEHR UNGEZWUNGEN, KONTAKTBEREIT ..... 1 UNGEZWUNGEN, KONTAKTBEREIT ..... 2 EHER GEHEMT, KONTAKTSCHAU ..... 3 SEHR GEHEMT, KONTAKTSCHAU ..... 4
6.	Der/die Befragte scheint nach meinem Eindruck –	SEHR RUHIG, GELASSEN ..... 1 RUHIG, GELASSEN ..... 3 EHER ERREGBAR, EMPFINDLICH ..... 3 SEHR ERREGBAR, EMPFINDLICH ..... 4
7.	Der/die Befragte scheint nach meinem Eindruck –	SEHR DURCHSETZUNGSBEREIT ..... 1 DURCHSETZUNGSBEREIT ..... 2 EHER NICHT DURCHSETZUNGSBEREIT ..... 3 ÜBERHAUPT NICHT DURCHSETZUNGSBEREIT ..... 4
8.	Der persönliche Stil des/der Befragten (Äußeres, Frisur, Schmuck, Kleidung) wirkt nach meinem Eindruck –	SEHR AUFFÄLLIG ..... 1 EHER AUFFÄLLIG ..... 2 EHER UNAUFFÄLLIG ..... 3 UNAUFFÄLLIG ..... 4
9.	Das Ausdrucksverhalten (Sprechen, Gesichtsausdruck, Gestik) wirkt –	SEHR LEBHAFT, AUSDRUCKSSTARK ..... 1 EHER LEBHAFT ..... 2 EHER AUSDRUCKSLOS ..... 3 AUSDRUCKSLOS ..... 4

### Wohnung der/des Befragten – Einstufung durch den Interviewer

1.	Eindruck von der Wohnungs-Umgebung (z.B. Lage im Gebäude, Straße, Umgebung)	GUT, GÜNSTIG GELEGEN.....	1
		EHER GUT, GÜNSTIG GELEGEN .....	2
		EHER SCHLECHT, UNGÜNSTIG GELEGEN.....	3
		SCHLECHT, UNGÜNSTIG GELEGEN .....	4
2.	Eindruck von der Wohnungs-Einrichtung (Möbel, Bilder, "Wohnaccessoires")	SEHR UNPERSÖNLICH.....	1
		EHER UNPERSÖNLICH.....	2
		EHER PERSÖNLICH .....	3
		PERSÖNLICH.....	4
3.	Wohnungs-Zustand (Ordnung, Sauberkeit)	SEHR AUFGERÄUMT, GEPFLEGT.....	1
		EHER AUFGERÄUMT, GEPFLEGT .....	2
		EHER UNAUFGERÄUMT, UNGEPFLEGT.....	3
		SEHR UNAUFGERÄUMT, UNGEPFLEGT .....	4
4.	Wohnungs-Helligkeit	SEHR HELL .....	1
		EHER HELL.....	2
		EHER DUNKEL.....	3
		SEHR DUNKEL .....	4
5.	Wohnungs-Farbgebung	EINFARBIG.....	1
		EHER DEZENT GETÖNT .....	2
		FARBENFROH (HARMONISCH) .....	3
		KNALLIGE FARBEN (NICHT HARMONISCH) .....	4

<b>Zusätzliche Informationen, die das IfD regelmäßig bei jeder Umfrage erfasst:</b>		
1.	Statistik 1	Geschlecht
2.	Statistik 2	Alter (absolut und in Kategorien)
3.	Statistik 3	Schulabschluss (Listenvorlage)
4.	Statistik 4	Berufstätigkeit/Berufsstellung (Listenvorlage)
5.	Statistik 5	Sind Sie im öffentlichen Dienst beschäftigt?
6.	Statistik 6a	Monatsnettoeinkommen des Hauptverdieners (Listenvorlage)
7.	Statistik 6b	Der Hauptverdiener ist: Befragter selbst, andere Person im Haushalt
8.	Statistik 7a	Haben Sie Kinder? Falls ja, wie viele Kinder?
9.	Statistik 7b	Könnten Sie mir noch sagen, in welchem Alter die Kinder sind?
10.	Statistik 7c	Wie viele Kinder leben zurzeit im Haushalt?
11.	Statistik 8a	Leben Sie in einem Mehrpersonenhaushalt oder allein?
12.	Statistik 8b	Wie viele Personen – Kinder und Erwachsene zusammen – leben im Haushalt, Sie selbst bitte mitgezählt?
13.	Statistik 9	Wie viele Personen im Haushalt sind berufstätig oder haben sonst Einkommen irgendwelcher Art, wie Rente, Mieteinkommen usw.?
14.	Statistik 10	Monatliches Haushaltsnettoeinkommen (Listenvorlage)
15.	Statistik 11a	Familienstand: Ledige, Verwitwete, Geschiedene bzw. getrenntlebende Verheiratete
16.	Statistik 11b	Leben Sie mit einem Partner, einem Lebensgefährten zusammen?
17.	Statistik 12a	Mitglieder einer Kirche oder Religionsgemeinschaft, ehemalige Mitglieder
18.	Statistik 12b	Welcher Religion bzw. Glaubensgemeinschaft gehören Sie an?
19.	Statistik 13	Häufigkeit des Gottesdienstbesuchs: 5 Kategorien
20.	Statistik 14	Parteisympathie Platz 1: 6 politische Parteien zur Auswahl
21.	Statistik 15a	Geburtsjahrgang: Sind Sie vor 1990 geboren, oder im Jahr 1990 oder später?
22.	Statistik 15b	Vor 1990 Geborene: Können Sie mir sagen, in welchem Teil Deutschlands Sie im Herbst 1989 gewohnt haben, also vor der Wende in Ostdeutschland?
23.	Statistik 16	Bundesland
24.	Statistik 17	Einwohnerzahl des Wohnortes: 8 Kategorien
25.	Statistik 18	Charakter des Wohnortes: (1) Großstadtatmosphäre .... (5) ländliches Milieu mit kleineren Städten in der Nähe
26.	Statistik 19	Soziale Schicht des Befragten: (1) Oberschicht (Schicht A), (2) Obere Mittelschicht (Schicht B), (3) Untere Mittelschicht (Schicht C), (4) Unterschicht (Schicht D)

Der nach dem mündlichen Interview ausgeteilte Fragebogen umfasste:

- 138 Items des FPI-R (stimmt/stimmt nicht);
- 16 Items (Aussagen) zu politisch-gesellschaftlichen Themen (stimmt/stimmt nicht);
- 7 Items zu religiös-weltanschaulichen Einstellungen (stimmt/stimmt nicht) plus zwei 10-stufige Rating- Skalen zu Religiosität und Interesse für die Frage nach dem Sinn des Lebens;
- 7 Items (stimmt/stimmt nicht) mit vermutetem Bezug zu FPI-Konstrukten;
- 7 Items zu verbreiteten Sorgen und Ängsten der Deutschen (jeweils 5-stufige Skala);
- 25 Fragen zu beruflicher Belastung, Lebenszufriedenheit und Gesundheit (identisch mit den 1999 verwendeten Fragen);
- 6 verhaltensnahe Fragen mit Bezug zu FPI-Aggressivität, Extraversion, Leistungsorientierung, Beanspruchung und Soziale Orientierung:
- Erlebte körperliche Angriffe seit dem Alter von 16 (4-stufig: häufig, gelegentlich, nur einmalig, nie)
- Kontakt mit Freunden in der Freizeit (4-stufig: täglich, wöchentlich, monatlich, seltener)
- Leistungsorientierter Sport (3-stufig: leistungsorientierter Sport, Sport ohne Leistungsorientierung, gar kein Sport)
- Zeit für wichtigstes Hobby in Stunden (× Stunden pro Woche, habe kein Hobby)
- Arbeitszeit pro Woche laut Arbeitsvertrag (4-stufig: soviel wie vorgesehen, etwas mehr, deutlich mehr, bin nicht berufstätig).
- Spenden für soziale Zwecke pro Jahr (Betrag in × Euro, spende nichts)

Das mündliche Interview des IfD, welches jeweils nach einem Leitfaden mit Merkmalsblatt durchgeführt wird, dauerte in der Erhebung 2018 durchschnittlich 60 Minuten, einschließlich der Zusatzfragen. Der Interviewer erfasste auch 24 soziodemografische Merkmale, u. a.:

- Geschlecht, Alter, Schulabschluss,
- Bundesland und Gebiet nach Ost – West,
- Einwohnerzahl des Wohnortes, Stadt – Land,
- Berufstätigkeit, Stellung im Beruf, Einkommen,
- Familienstand bzw. Partnerschaft, Anzahl der Kinder und anderer Personen im Haushalt,
- Mitglied einer Kirche, Konfession, Häufigkeit der Kirchenbesuche,
- Präferenz für eine politische Partei,
- Einstufung der Schicht durch Interviewer.

### III. FPI-R 2018 Datenübersicht

Der Datensatz ist heterogen hinsichtlich psychologischer und metrischer Merkmale

**Personen-Merkmale und sozioökonomische Daten (IfD)**  
**Beurteilungen und Beobachtungen durch Interviewer**  
**Selbstbeurteilungen (FPI-Items, Einstellungen)**  
**Selbstberichte (im Prinzip objektivierbar)**  
**Indizes** (außer FPI-Testwerten, Anzahl MD): Häufigkeiten (nominal),  
Item-Summen (zweistufige Items), Item-Summen (mehrstufige Items),  
weitere Indizes

#### **Personen-Merkmale und sozioökonomische Daten**

Geschlecht , Alter, Altersgruppen, Schulabschluss, Berufstätigkeit, Einkommen, Kinder,  
Haushaltsgröße,  
Familienstand, Zusammenleben mit Partner;  
Kirchenmitglied, Konfession, Gottesdienstbesuch (Einstellung), Parteisympathie (Einstellung)  
Geburtsjahrgang  
Wohnort, Bundesland, Ost-West, Wohnortgröße

#### **Beurteilungen und Beobachtungen durch Interviewer**

Charakter des Wohnortes, Verhalten der Befragten während des Interviews,  
Wohnung  
Tätowierungen, Piercings, Ringe

#### **Selbstbeurteilungen (Items)**

Selbstbeurteilungen (FPI-Items)  
Politisch-soziale Einstellungen  
Religiös-weltanschauliche Einstellungen  
Ängste und Sorgen  
Zufriedenheit (Beruf, Familie, Finanzen, Wohnung), vorzeitige Rente  
Gesundheitszustand

#### **Selbstberichte**

Erfahrungen und Gewohnheiten  
Beruf, Arbeitszeit  
Berufliche Belastung  
Rauchen, Zigaretten, Körpergröße, Körpergewicht  
Chronische Krankheiten, Nennungen  
Krankenhaus, Operationen, Kuren, Arztbesuche  
Psychotherapie  
Tabletten

#### **Indizes / Aggregate**

Selbstbeurteilungen (FPI-Items)  
Itemgruppen (Miniskalen, Aggregate)  
Politisch-soziale Einstellungen  
Religiös-weltanschauliche Einstellungen  
Ängste und Sorgen  
Berufliche Zufriedenheit  
Berufliche Belastung  
Chronische Krankheiten  
Krankenhaus, Operationen, Kuren, Arztbesuche  
Tabletten  
Body Mass Index



#### IV. FPI-R Skalen: Vergleiche nach Erhebungsjahr, Geschlecht, Alter

FPI-Skalenwerte für Erhebungen 1982, 1999, 2018 (nur West-Deutschland, repräsentative Gesamtstichproben)

Jahr der Normierung	1982			1999			2018			ES
ungewichtete Stichproben	M	Mdn	s	M	Mdn	s	M	Mdn	s	$\eta^2$
FPIR1 - LEBENSZUFRIEDENHEIT	7,0	8	3,1	7,6	8	3,0	7,6	8	3,0	0,007
FPIR2 - SOZ. ORIENTIERUNG	6,5	7	2,8	6,4	7	2,9	6,5	7	2,8	0,000
FPIR3 - LEISTUNGSORIENTIERUNG	6,9	7	3,1	6,9	7	3,2	6,8	7	3,1	0,000
FPIR4 - GEHEMMTHEIT	5,2	5	3,1	5,0	5	3,1	4,7	4	3,0	0,005
FPIR5 - ERREGBARKEIT	5,5	5	3,0	5,2	5	3,0	4,9	5	3,1	0,005
FPIR6 - AGGRESSIVITÄT	4,2	4	2,8	4,4	4	2,9	3,8	3	2,8	0,008
FPIR7 - BEANSPRUCHUNG	5,7	6	3,5	5,5	5	3,5	5,2	5	3,4	0,003
FPIR8 - KÖRP. BESCHWERDEN	4,0	4	3,0	3,4	3	3,0	3,3	3	2,8	0,013
FPIR9 - GESUNDHEITSSORGEN	5,9	6	3,2	6,1	6	3,1	6,2	6	3,0	0,001
FPIR10 - OFFENHEIT	5,9	6	2,9	6,2	6	2,9	6,0	6	3,0	0,001
FPIRE - EXTRAVERSION	6,9	7	3,6	6,9	7	3,6	6,8	7	3,5	0,000
FPIRN - EMOTIONALITÄT	6,2	6	3,5	5,7	5	3,7	5,3	5	3,7	0,010
Durchschnittsalter	43	42	17	46	45	18	50	51	19	
N =	2035			2086			2579			

Jahr der Normierung	1982			1999			2018			ES
gewichtete Stichproben nach Geschlecht und Altersgruppe	M	Mdn	s	M	Mdn	s	M	Mdn	s	$\eta^2$
FPIR1 - LEBENSZUFRIEDENHEIT	7,0	8	3,1	7,6	8	3,0	7,6	8	3,0	0,020
FPIR2 - SOZ. ORIENTIERUNG	6,5	7	2,8	6,3	7	2,9	6,4	7	2,8	0,058
FPIR3 - LEISTUNGSORIENTIERUNG	6,9	7	3,1	6,9	7	3,1	6,9	7	3,0	0,031
FPIR4 - GEHEMMTHEIT	5,2	5	3,1	4,9	5	3,1	4,6	4	3,0	0,032
FPIR5 - ERREGBARKEIT	5,5	5	3,0	5,3	5	3,0	5,1	5	3,1	0,025
FPIR6 - AGGRESSIVITÄT	4,2	4	2,8	4,5	4	2,9	4,0	3	2,9	0,072
FPIR7 - BEANSPRUCHUNG	5,7	6	3,5	5,5	5	3,5	5,4	5	3,4	0,014
FPIR8 - KÖRP. BESCHWERDEN	4,0	4	3,0	3,3	3	2,9	3,1	2	2,8	0,148
FPIR9 - GESUNDHEITSSORGEN	5,9	6	3,2	6,0	6	3,1	5,9	6	3,0	0,155
FPIR10 - OFFENHEIT	5,9	6	2,9	6,3	6	2,9	6,3	6	3,0	0,096
FPIRE - EXTRAVERSION	6,9	7	3,6	7,1	7	3,6	7,0	7	3,5	0,079
FPIRN - EMOTIONALITÄT	6,2	6	3,5	5,8	6	3,7	5,3	5	3,7	0,034
Durchschnittsalter	43	42	17	44	43	17	43	43	18	
N	2035			2035			2035			

Gesamtstichproben	1982			2018				Cohen
altersgewichtete Stichprobe 2018	M	Mdn	s	M	Mdn	s	$p \leq .01$	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,0	8	3,1	7,6	8	3,0	**	0,19
FPIR2 - SOZ. ORIENTIERUNG	6,5	7	2,8	6,4	7	2,8		
FPIR3 - LEISTUNGSORIENTIERUNG	6,9	7	3,1	7,0	7	3,0		
FPIR4 - GEHEMMTHEIT	5,2	5	3,1	4,6	4	3,0	**	-0,19
FPIR5 - ERREGBARKEIT	5,5	5	3,0	5,1	5	3,1	**	-0,14
FPIR6 - AGGRESSIVITÄT	4,2	4	2,8	4,0	3	2,9		
FPIR7 - BEANSPRUCHUNG	5,7	6	3,5	5,4	5	3,4		
FPIR8 - KÖRP. BESCHWERDEN	4,0	4	3,0	3,0	2	2,7	**	-0,34
FPIR9 - GESUNDHEITSSORGEN	5,9	6	3,2	5,9	6	3,0		
FPIR10 - OFFENHEIT	5,9	6	2,9	6,3	6	3,0	**	0,11
FPIRE - EXTRAVERSION	6,9	7	3,6	7,0	7	3,5		
FPIRN - EMOTIONALITÄT	6,2	6	3,5	5,3	5	3,7		
N =	2035			2035				

	1982			2018				Cohen
<b>Männer (nach Alter gewichtet)</b>	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,4	8	3,0	7,7	8	3,1		
FPIR2 - SOZIALE ORIENTIERUNG	5,9	6	2,8	5,8	6	2,9		
FPIR3 - LEISTUNGSORIENTIERUNG	7,5	8	2,9	7,2	8	3,0		
FPIR4 - GEHEMMTHEIT	4,6	4	2,9	4,3	4	2,9		
FPIR5 - ERREGBARKEIT	4,9	5	2,9	4,8	4	3,0		
FPIR6 - AGGRESSIVITÄT	4,7	4	3,0	4,5	4	3,1		
FPIR7 - BEANSPRUCHUNG	5,3	5	3,4	5,1	5	3,3		
<b>FPIR8 - KÖRPERL. BESCHWERDEN</b>	<b>3,1</b>	<b>2</b>	<b>2,8</b>	<b>2,3</b>	<b>2</b>	<b>2,4</b>	<b>**</b>	<b>-0,30</b>
FPIR9 - GESUNDHEITSSORGEN	5,4	5	3,2	5,1	5	2,9		
FPIR10 - OFFENHEIT	6,4	6	2,8	6,7	7	3,0		
FPIRE - EXTRAVERSION	7,4	8	3,5	7,3	7	3,5		
FPIRN - EMOTIONALITÄT	5,4	5	3,5	4,8	4	3,7	<b>**</b>	<b>-0,17</b>
N ungewichtet	<b>953</b>			<b>1240</b>				
N gewichtet nach Alter	<b>953</b>			<b>953</b>				
Durchschnittsalter vor Gewichtung	<b>42</b>			<b>49</b>				
Durchschnittsalter nach Gewichtung	<b>42</b>			<b>42</b>				

	1982			2018				Cohen
<b>Frauen (nach Alter gewichtet)</b>	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	6,7	7	3,2	7,5	8	3,0	<b>**</b>	<b>-0,27</b>
FPIR2 - SOZIALE ORIENTIERUNG	7,0	7	2,7	7,0	7	2,7		
FPIR3 - LEISTUNGSORIENTIERUNG	6,3	6	3,1	6,7	7	3,0	<b>**</b>	<b>-0,13</b>
FPIR4 - GEHEMMTHEIT	5,7	6	3,1	5,0	5	3,1	<b>**</b>	<b>0,24</b>
FPIR5 - ERREGBARKEIT	6,0	6	2,9	5,3	5	3,2	<b>**</b>	<b>0,22</b>
FPIR6 - AGGRESSIVITÄT	3,7	3	2,6	3,5	3	2,6		
FPIR7 - BEANSPRUCHUNG	6,0	6	3,6	5,6	5	3,4		
<b>FPIR8 - KÖRPERL. BESCHWERDEN</b>	<b>4,9</b>	<b>5</b>	<b>3,0</b>	<b>3,8</b>	<b>3</b>	<b>2,8</b>	<b>**</b>	<b>0,37</b>
FPIR9 - GESUNDHEITSSORGEN	6,4	6	3,1	6,6	7	2,9		
FPIR10 - OFFENHEIT	5,5	5	3,0	5,9	6	2,9	<b>**</b>	<b>-0,12</b>
FPIRE - EXTRAVERSION	6,4	6	3,5	6,8	7	3,5		
<b>FPIRN - EMOTIONALITÄT</b>	<b>6,8</b>	<b>7</b>	<b>3,5</b>	<b>5,7</b>	<b>5</b>	<b>3,7</b>	<b>**</b>	<b>0,31</b>
N ungewichtet	<b>1082</b>			<b>1339</b>				
N gewichtet nach Alter	<b>1082</b>			<b>1082</b>				
Durchschnittsalter vor Gewichtung	<b>45</b>			<b>52</b>				
Durchschnittsalter nach Gewichtung	<b>45</b>			<b>45</b>				

	1982			2018				Cohen
<b>16-30-Jährige</b>	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
<b>FPIR1 - LEBENSZUFRIEDENHEIT</b>	<b>6,5</b>	<b>7</b>	<b>3,2</b>	<b>7,3</b>	<b>8</b>	<b>3,0</b>	<b>**</b>	<b>0,27</b>
FPIR2 - SOZIALE ORIENTIERUNG	6,3	6	2,8	6,3	6	2,9		
FPIR3 - LEISTUNGSORIENTIERUNG	6,8	7	2,9	6,9	7	2,9		
FPIR4 - GEHEMMTHEIT	4,9	5	3,1	4,7	4	3,2		
FPIR5 - ERREGBARKEIT	5,7	6	3,1	5,2	5	3,1	<b>**</b>	<b>-0,17</b>
FPIR6 - AGGRESSIVITÄT	4,8	4	3,1	4,3	4	3,0	<b>**</b>	<b>-0,16</b>
FPIR7 - BEANSPRUCHUNG	5,6	5	3,4	5,3	5	3,3		
FPIR8 - KÖRPERL. BESCHWERDEN	3,0	3	2,5	2,6	2	2,6	<b>**</b>	<b>-0,16</b>
FPIR9 - GESUNDHEITSSORGEN	4,4	4	2,8	4,9	5	2,8	<b>**</b>	<b>0,19</b>
FPIR10 - OFFENHEIT	7,1	7	2,7	7,0	7	2,9		
FPIRE - EXTRAVERSION	8,0	8	3,4	7,7	8	3,6		
<b>FPIRN - EMOTIONALITÄT</b>	<b>6,4</b>	<b>6</b>	<b>3,5</b>	<b>5,1</b>	<b>5</b>	<b>3,6</b>	<b>**</b>	<b>-0,36</b>
N ungewichtet	<b>619</b>			<b>526</b>				
N gewichtet nach Alter	<b>619</b>			<b>619</b>				
Durchschnittsalter vor Gewichtung	<b>24</b>			<b>23</b>				
Durchschnittsalter nach Gewichtung	<b>24</b>			<b>24</b>				

	1982			2018				Cohen
60-70-Jährige	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,4	8	3,0	7,9	9	3,0		
FPIR2 - SOZIALE ORIENTIERUNG	6,8	7	2,9	6,5	7	2,8		
FPIR3 - LEISTUNGSORIENTIERUNG	6,6	7	3,3	6,6	7	3,1		
FPIR4 - GEHEMMTHEIT	5,5	5	3,0	4,6	4	2,9	**	-0,30
FPIR5 - ERREGBARKEIT	5,0	5	2,8	4,6	4	3,1		
FPIR6 - AGGRESSIVITAET	3,5	3	2,5	3,4	3	2,5		
FPIR7 - BEANSPRUCHUNG	4,9	5	3,5	4,5	4	3,4		
FPIR8 - KÖRPERL. BESCHWERDEN	5,5	5	3,2	3,5	3	2,9	**	-0,62
FPIR9 - GESUNDHEITSSORGEN	7,4	8	2,9	6,7	7	2,9	**	-0,26
FPIR10 - OFFENHEIT	4,7	5	2,8	5,4	5	2,9	**	0,23
FPIRE - EXTRAVERSION	5,6	6	3,4	6,2	6	3,5		
FPIRN - EMOTIONALITAET	5,9	6	3,5	5,1	5	3,7	**	-0,23
N ungewichtet	345			370				
N gewichtet nach Alter	345			345				
Durchschnittsalter vor Gewichtung	64			64				
Durchschnittsalter nach Gewichtung	64			64				

**OST-Deutschland - Erhebungen 1999 und 2018, repräsentative Stichproben**  
**FPI-R Skalen; Mittelwertvergleich mit Altersgewichtung; Signifikanzen und Effektstärken**

	Erhebung 1999			Erhebung 2018				
Gesamtstichprobe (Ost)	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,5	8	3,0	7,7	8	2,9		
FPIR2 - SOZIALE ORIENTIERUNG	6,2	6	2,8	6,0	6	2,9		
FPIR3 - LEISTUNGSORIENTIERUNG	7,2	7	3,1	7,0	7	3,1		
FPIR4 - GEHEMMTHEIT	5,1	5	3,0	4,8	4	3,1		
FPIR5 - ERREGBARKEIT	5,2	5	3,0	4,9	5	3,0		
FPIR6 - AGGRESSIVITAET	4,6	4	2,9	3,9	3	2,8	**	-0,24
FPIR7 - BEANSPRUCHUNG	5,3	5	3,4	5,7	6	3,5		
FPIR8 - KÖRPERLICHE BESCHWERDEN	3,6	3	2,9	2,9	2	2,5	**	-0,23
FPIR9 - GESUNDHEITSSORGEN	6,5	7	3,1	6,1	6	2,9	**	-0,13
FPIR10 - OFFENHEIT	5,7	6	2,9	6,1	6	2,9	**	0,13
FPIRE - EXTRAVERSION	6,8	7	3,6	6,9	7	3,5		
FPIRN - EMOTIONALITAET	5,7	5	3,6	5,3	5	3,5		
N	871 (Zufall aus N = 1574)			871 (original)				
Durchschnittsalter nach Erhebung	47			53				
Durchschnittsalter nach Gewichtung	47			47				

	Erhebung 1999			Erhebung 2018				
Männer (Ost)	M	Mdn	s	M	Mdn	s	p ≤ .01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,8	8	2,8	7,7	8	2,9		
FPIR2 - SOZIALE ORIENTIERUNG	5,4	6	2,8	5,1	5	2,9		
FPIR3 - LEISTUNGSORIENTIERUNG	7,7	8	2,9	7,2	8	3,1	**	-0,19
FPIR4 - GEHEMMTHEIT	4,7	5	2,7	4,5	4	2,9		
FPIR5 - ERREGBARKEIT	4,8	5	2,8	4,6	4	2,9		
FPIR6 - AGGRESSIVITAET	5,1	5	3,1	4,4	4	3,1	**	-0,22
FPIR7 - BEANSPRUCHUNG	5,2	5	3,3	5,4	5	3,3		
FPIR8 - KÖRPERLICHE BESCHWERDEN	2,6	2	2,6	2,3	2	2,3		
FPIR9 - GESUNDHEITSSORGEN	5,8	6	3,2	5,3	5	2,9		
FPIR10 - OFFENHEIT	6,3	6	2,8	6,6	7	2,8		
FPIRE - EXTRAVERSION	7,0	7	3,5	7,1	7	3,4		
FPIRN - EMOTIONALITAET	5,0	5	3,5	4,9	5	3,4		
N	412			414				
Durchschnittsalter nach Erhebung	45			52				
Durchschnittsalter nach Gewichtung	45			45				

	Erhebung 1999			Erhebung 2018				
<b>Frauen (Ost)</b>	M	Mdn	s	M	Mdn	s	p≤.01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,3	8	3,2	7,7	8	2,8		
FPIR2 - SOZIALE ORIENTIERUNG	6,9	7	2,7	6,7	7	2,6		
FPIR3 - LEISTUNGSORIENTIERUNG	6,6	7	3,2	6,9	7	3,1		
FPIR4 - GEHEMMTHEIT	5,5	5	3,2	5,0	5	3,3		
FPIR5 - ERREGBARKEIT	5,5	5	3,1	5,1	5	3,1		
<b>FPIR6 - AGGRESSIVITAET</b>	4,0	4	2,6	3,4	3	2,5	**	-0,26
<b>FPIR7 - BEANSPRUCHUNG</b>	5,4	5	3,4	6,0	6	3,6	**	0,19
<b>FPIR8 - KÖRPERLICHE BESCHWERDEN</b>	4,4	4	3,0	3,5	3	2,6	**	-0,32
FPIR9 - GESUNDHEITSSORGEN	7,1	7	3,0	6,8	7	2,8		
FPIR10 - OFFENHEIT	5,3	5	2,9	5,6	6	2,9		
FPIRE - EXTRAVERSION	6,6	7	3,7	6,8	7	3,5		
FPIRN - EMOTIONALITAET	6,3	6	3,7	5,7	5	3,6		
N	459			456				
Durchschnittsalter nach Erhebung	<b>49</b>			<b>53</b>				
Durchschnittsalter nach Gewichtung	49			49				

	Erhebung 1999			Erhebung 2018				
<b>16-30-Jährige (Ost - M+F)</b>	M	Mdn	s	M	Mdn	s	p≤.01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	7,3	8	3,0	7,7	8	2,6		
FPIR2 - SOZIALE ORIENTIERUNG	5,1	5	2,7	5,7	6	3,0		
FPIR3 - LEISTUNGSORIENTIERUNG	7,5	8	2,7	7,1	7	2,9		
FPIR4 - GEHEMMTHEIT	4,9	5	3,0	5,2	5	3,3		
FPIR5 - ERREGBARKEIT	5,5	5	3,0	5,0	5	2,8		
<b>FPIR6 - AGGRESSIVITAET</b>	5,7	6	3,1	4,3	4	2,8	**	-0,49
FPIR7 - BEANSPRUCHUNG	5,3	5	3,3	5,6	5	3,5		
FPIR8 - KÖRPERLICHE BESCHWERDEN	2,5	2	2,6	2,4	2	2,3		
FPIR9 - GESUNDHEITSSORGEN	4,9	5	2,6	5,0	5	2,7		
FPIR10 - OFFENHEIT	7,2	8	2,6	7,1	7	2,8		
<b>FPIRE - EXTRAVERSION</b>	8,3	8	3,3	7,3	7	3,5	**	-0,29
FPIRN - EMOTIONALITAET	5,6	6	3,6	5,2	5	3,4		
N	190			150				
Durchschnittsalter nach Erhebung	<b>23</b>			<b>24</b>				

	Erhebung 1999			Erhebung 2018				
<b>60-70-Jährige (Ost - M+F)</b>	M	Mdn	s	M	Mdn	s	p≤.01	ES
FPIR1 - LEBENSZUFRIEDENHEIT	8,1	9	2,7	8,0	8	2,8		
<b>FPIR2 - SOZIALE ORIENTIERUNG</b>	7,2	7	2,4	6,3	7	2,7	**	-0,37
FPIR3 - LEISTUNGSORIENTIERUNG	6,7	7	3,3	6,5	7	3,1		
FPIR4 - GEHEMMTHEIT	5,5	5	2,9	5,2	5	3,1		
FPIR5 - ERREGBARKEIT	4,9	5	2,6	4,3	4	2,7		
FPIR6 - AGGRESSIVITAET	3,8	4	2,6	3,2	3	2,7		
FPIR7 - BEANSPRUCHUNG	4,1	4	3,3	4,8	5	3,4		
FPIR8 - KÖRPERLICHE BESCHWERDEN	4,3	4	3,0	3,5	3	2,7		
<b>FPIR9 - GESUNDHEITSSORGEN</b>	8,2	9	2,7	6,6	7	2,9	**	-0,57
FPIR10 - OFFENHEIT	4,9	5	2,8	5,3	6	2,8		
FPIRE - EXTRAVERSION	5,7	6	3,4	6,0	6	3,1		
FPIRN - EMOTIONALITAET	5,5	6	3,5	5,3	5	3,4		
N	131			143				
Durchschnittsalter nach Erhebung	<b>64</b>			<b>65</b>				

Cohen d: > 0,5 und < 0,7 mittlerer Effekt; > 0,8 großer Effekt

**Standardmessfehler und Vertrauensgrenzen der FPI-R Testwerte Erhebung 1999 (N = 3 740) :**

	Standard-Messfehler	Vertrauensgrenzen der Testwerte	Standardmessfehler Differenz	Kritische Differenz zweier Testwerte
<b>FPI-R-Skala</b>	$s_x \sqrt{1 - r_{tt}}$	CI	$s_x \sqrt{2(1 - r_{tt})}$	RCI
Lebenszufriedenheit	1.47	± 2.9	2.07	4.1
Soziale Orientierung	1.48	± 2.9	2.09	4.2
Leistungsorientierung	1.48	± 2.9	2.09	4.2
Gehemmtheit	1.46	± 2.9	2.06	4.1
Erregbarkeit	1.45	± 2.8	2.04	4.1
Aggressivität	1.46	± 2.9	2.06	4.1
Beanspruchung	1.43	± 2.8	2.02	4.0
Körperliche Beschwerden	1.36	± 2.7	1.92	3.8
Gesundheitssorgen	1.46	± 2.9	2.06	4.1
Offenheit	1.47	± 2.9	2.08	4.2
Extraversion	1.58	± 3.1	2.23	4.5
Emotionalität	1.56	± 3.1	2.21	4.4

Anmerkung:  $s_x$  = Standardabweichung,  $r_{tt}$  = Reliabilitätskoeffizient, CI = Konfidenzintervall, RCI = Reliable Change Index; Irrtumswahrscheinlichkeit von 5 %

**Standardmessfehler und Vertrauensgrenzen der FPI-R Testwerte Erhebung 2018 (N=3 450):**

	Standard-Messfehler	Vertrauensgrenzen der Testwerte	Standardmessfehler Differenz	Kritische Differenz zweier Testwerte
<b>FPI-R-Skala</b>	$s_x \sqrt{1 - r_{tt}}$	CI	$s_x \sqrt{2(1 - r_{tt})}$	RCI
Lebenszufriedenheit	1.43	± 2.8	2.02	4.0
Soziale Orientierung	1.48	± 2.9	2.10	4.1
Leistungsorientierung	1.48	± 2.9	2.10	4.1
Gehemmtheit	1.42	± 2.8	2.01	3.9
Erregbarkeit	1.41	± 2.8	1.99	3.9
Aggressivität	1.40	± 2.8	1.98	3.9
Beanspruchung	1.45	± 2.8	2.05	4.0
Körperliche Beschwerden	1.34	± 2.6	1.89	3.7
Gesundheitssorgen	1.49	± 2.9	2.10	4.1
Offenheit	1.48	± 2.9	2.10	4.1
Extraversion	1.57	± 3.1	2.21	4.3
Emotionalität	1.55	± 3.0	2.20	4.3

Anmerkung:  $s_x$  = Standardabweichung,  $r_{tt}$  = Reliabilitätskoeffizient, CI = Konfidenzintervall, RCI = Reliable Change Index; Irrtumswahrscheinlichkeit von 5 %

## V. Itemstatistiken

### Item Statistiken der Normstichprobe 2018 ( $N_{\min} = 3292$ , $N_{\max} = 3399$ )

Schwierigkeitsindex ( $p_i$ ) und Trennschärfeindex ( $r_{it}$ ) in der betreffenden Skala

<b>Skala 1 LEBENSZUFRIEDENHEIT, Cronbach's alpha = 0,77</b>	$p_i$	$r_{it}$
V88 Ich bin immer guter Laune	0,37	0,28
V29 Wenn ich noch einmal geboren würde, dann würde ich nicht anders leben wollen	0,54	0,31
V112 Ich grüble viel über mein bisheriges Leben nach (stimmt nicht)	0,58	0,43
V23 Ich lebe mit mir selbst in Frieden und ohne innere Konflikte	0,63	0,44
V131 Ich habe eine gute Partnerbeziehung (Ehe)	0,64	0,26
V100 Ich bin selten in bedrückter, unglücklicher Stimmung	0,65	0,33
V94 Oft habe ich alles gründlich satt (stimmt nicht)	0,66	0,51
V03 Ich habe (hatte) einen Beruf, der mich voll befriedigt	0,68	0,32
V58 In meinem bisherigen Leben habe ich kaum das verwirklichen können, was in mir steckt (stimmt nicht)	0,69	0,39
V119 Ich bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden (stimmt nicht)	0,69	0,57
V138 Meistens blicke ich voller Zuversicht in die Zukunft	0,75	0,50
V128 Alles in allem bin ich ausgesprochen zufrieden mit meinem bisherigen Leben	0,77	0,58

<b>Skala 2 SOZIALE ORIENTIERUNG, Cronbach's alpha = 0,72</b>	$p_i$	$r_{it}$
V36 Ich denke oft, dass ich meinen Konsum einschränken müsste, um dann an benachteiligte Menschen abzugeben	0,22	0,31
V87 Wenn mich ein Fremder um eine kleine Geldspende bittet, ist mir das ziemlich lästig (stimmt nicht)	0,39	0,28
V134 Ich habe schon unbezahlt in meiner Gemeinde oder in sozialen Einrichtungen wie dem Roten Kreuz geholfen	0,41	0,31
V80 Ich bekomme häufig ein schlechtes Gewissen, wenn ich sehe, wie schlecht es anderen Menschen geht	0,46	0,40
V22 Ich bin der Ansicht, die Menschen in den Entwicklungsländern sollten sich zuerst einmal selbst helfen (stimmt nicht)	0,50	0,32
V111 Ich gebe gelegentlich Geld u. Spenden für Katastrophenhilfe, Caritas, Brot für die Welt u. and. gemeinnützige Org.	0,51	0,37
V14 Ich würde kaum zögern, auch alte und schwerbehinderte Menschen zu pflegen	0,52	0,27
V34 Ich finde, jeder Mensch soll sehen, wie er zurechtkommt (stimmt nicht)	0,55	0,35
V121 Da der Staat schon für Sozialhilfe sorgt, brauche ich im Einzelnen nicht zu helfen (stimmt nicht)	0,55	0,48
V41 Wenn jemand weint, möchte ich ihn am liebsten umarmen und trösten	0,72	0,35
V26 Ich fühle mich auch über meine Familie hinaus für andere Menschen verantwortlich	0,73	0,40
V136 Ich nehme mir viel Zeit, anderen Menschen geduldig zuzuhören, wenn sie von ihren	0,79	0,33

<b>Skala 3 LEISTUNGSORIENTIERUNG, Cronbach's alpha = 0,77</b>	$p_i$	$r_{it}$
V61 Die beruflichen Aufgaben sind mir oft wichtiger als viel Freizeit oder interessante Hobbies	0,35	0,26
V48 Mit anderen zu wetteifern, macht mir Spaß	0,45	0,44
V129 Bei meiner Arbeit bin ich meist schneller als andere	0,46	0,43
V39 Ich habe gern mit Aufgaben zu tun, die schnelles Handeln verlangen	0,50	0,49
V13 Meine Bekannten halten mich für einen energischen Menschen	0,55	0,38
V54 Bei wichtigen Dingen bin ich bereit, mit anderen energisch zu konkurrieren	0,58	0,47
V21 Ich bin leicht beim Ehrgeiz zu packen	0,61	0,45
V05 Ich glaube, dass ich mir beim Arbeiten mehr Mühe gebe als die meisten anderen Menschen	0,61	0,31
V101 Ich ziehe das Handeln dem Planeschmieden vor	0,63	0,27
V74 Ich habe Spaß an schwierigen Aufgaben, die mich herausfordern	0,65	0,53
V44 Ich pflege schnell und sicher zu handeln	0,68	0,47
V43 Es gibt für mich noch eine Menge sinnvoller Aufgaben, die ich in der Zukunft anpacken werde	0,76	0,35

<b>Skala 4 GEHEMMTHEIT, Cronbach's alpha = 0,78</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V97 Ich werde leicht rot	0,23	0,36
V06 Ich scheue mich, allein in einen Raum zu gehen, in dem andere Leute bereits zusammensitzen und sich unterhalten	0,26	0,41
V08 Ich würde mich beim Kellner o. Geschäftsführer eines Restaurants beschweren, wenn schlechtes Essen serviert wird (stimmt nicht)	0,28	0,31
V85 Ich werde ziemlich leicht verlegen	0,28	0,50
V73 Ich bin im Grunde eher ein ängstlicher Mensch	0,32	0,44
V11 Ich bin ungern mit Menschen zusammen, die ich noch nicht kenne	0,32	0,47
V124 Es fällt mir schwer, den richtigen Gesprächsstoff zu finden, wenn ich jemanden kennenlernen will	0,37	0,52
V04 Ich habe fast immer eine schlagfertige Antwort bereit (stimmt nicht)	0,45	0,41
V63 Es fällt mir schwer, vor einer großen Gruppe von Menschen zu sprechen oder vorzutragen	0,48	0,46
V120 Beim Reisen schaue ich lieber auf die Landschaft als mich mit den Mitreisenden zu unterhalten	0,55	0,25
V81 Ich schließe nur langsam Freundschaften	0,56	0,38
V109 Bei Geselligkeiten und öffentlichen Veranstaltungen bleibe ich lieber im Hintergrund	0,60	0,45

<b>Skala 5 ERREGBARKEIT, Cronbach's alpha = 0,79</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V105 Ich kann oft meinen Ärger und meine Wut nicht beherrschen	0,19	0,47
V102 Im Allgemeinen bin ich ruhig und nicht leicht aufzuregen (stimmt nicht)	0,26	0,52
V60 Auch wenn mich etwas sehr aus der Fassung bringt, beruhige ich mich meistens wieder rasch (stimmt nicht)	0,26	0,40
V113 Ich neige oft zu Hast und Eile, auch wenn es überhaupt nicht notwendig ist	0,35	0,38
V108 Ich lasse mich durch eine Vielzahl von kleinen Störungen nicht aus der Ruhe bringen (stimmt nicht)	0,36	0,43
V93 Es gibt nur wenige Dinge, die mich leicht erregen oder ärgern (stimmt nicht)	0,38	0,44
V52 Auch wenn es eher viel zu tun gibt, lasse ich mich nicht hetzen (stimmt nicht)	0,39	0,37
V135 Ich bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde	0,45	0,53
V115 Oft rege ich mich zu rasch über jemanden auf	0,47	0,54
V86 Ich koche vor Wut, wenn man mich zum Narren hält	0,49	0,38
V30 Wenn mir einmal etwas schief geht, regt mich das nicht weiter auf (stimmt nicht)	0,60	0,36
V27 Ich neige dazu, bei Auseinandersetzungen lauter zu sprechen als sonst	0,70	0,31

<b>Skala 6 AGGRESSIVITÄT, Cronbach's alpha = 0,76</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V50 Wenn ich meine Rechte nur mit körperlicher Gewalt verteidigen kann, so tue ich es	0,14	0,43
V137 Es gab Leute, die mich so ärgerten, dass es zu einer tätlichen Auseinandersetzung kam	0,15	0,44
V40 Es macht mir Spaß, anderen Fehler nachzuweisen	0,20	0,35
V37 Als Kind habe ich manchmal ganz gerne anderen die Arme umgedreht, an Haaren gezogen, ein Bein gestellt usw.	0,21	0,36
V20 Sind wir in ausgelassener Runde, so überkommt mich oft eine große Lust zu groben Streichen	0,22	0,36
V90 Wenn ich wirklich wütend werde, bin ich in der Lage, jemandem eine runterzuhauen	0,24	0,53
V12 Wenn jemand meinem Freund etwas Böses tut, bin ich dabei, wenn es heimgezahlt wird	0,38	0,44
V132 Ich gehe lieber bis zum Äußersten, als dass ich feige bin	0,38	0,31
V98 Einem Menschen, der mich schlecht behandelt oder beleidigt hat, wünsche ich eine harte Strafe	0,43	0,42
V24 Ich male mir manchmal aus, wie übel es denen eigentlich ergehen müsste, die mir Unrecht tun	0,45	0,39
V56 Wenn mich jemand anschreit, schreie ich zurück	0,47	0,41
V15 Ich kann mich erinnern, mal so wütend gewesen zu sein, dass ich das nächstbeste Ding nahm und es zerriss oder zerschlug	0,49	0,36

<b>Skala 7 BEANSPRUCHUNG, Cronbach's alpha = 0,82</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V122 Die Anforderungen, die an mich gestellt werden, sind oft zu hoch	0,26	0,42
V103 Vor lauter Aufgaben und Zeitdruck bin ich manchmal ganz durcheinander	0,34	0,47
V64 Auch an Wochenenden bin ich stark eingespannt	0,39	0,34
V96 Nur selten kann ich richtig abschalten	0,41	0,48
V28 Ich bin oft nervös, weil zu viel auf mich einströmt	0,42	0,42
V125 Ich denke manchmal, dass ich zu viel arbeite	0,43	0,51
V130 Ich habe häufig das Gefühl, im Stress zu sein	0,43	0,65

V35 Die täglichen Belastungen sind so groß, dass ich davon oft müde und erschöpft bin	0,45	0,57
V79 Ich bin häufiger abgespannt, matt und erschöpft	0,48	0,50
V118 Ich arbeite oft unter Zeitdruck	0,50	0,48
V46 Ich hätte gern mehr Zeit für mich ohne so viele Verpflichtungen	0,56	0,47
V116 Ich denke manchmal, dass ich mich mehr schonen sollte	0,59	0,42

<b>Skala 8 KÖRPERLICHE BESCHWERDEN, Cronbach's alpha = 0,77</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V67 Ich habe häufiger Verstopfung	0,14	0,35
V99 Meine Hände sind häufiger zitterig, z.B. beim Schreiben mit einem Stift oder Halten einer Tasse	0,14	0,37
V133 Ich habe manchmal ein Gefühl erstickender Enge in der Brust	0,18	0,51
V16 Ich habe häufig Kopfschmerzen	0,22	0,36
V51 Ich habe manchmal Hitzewallungen und mir steigt das Blut in den Kopf	0,26	0,42
V72 Ich bemerke häufiger ein unwillkürliches Zucken, zB um meine Augen	0,28	0,36
V33 Ich habe selbst bei warmem Wetter häufiger kalte Hände und Füße	0,28	0,32
V47 Ich habe manchmal das Gefühl, einen Kloss im Hals zu haben	0,29	0,43
V57 Mein Herz beginnt manchmal zu jagen oder unregelmäßig zu schlagen	0,31	0,51
V92 Ich habe einen empfindlichen Magen	0,32	0,40
V123 Mein Körper reagiert deutlich auf Wetteränderung	0,38	0,44
V75 Ich habe Schwierigkeiten einzuschlafen oder durchzuschlafen	0,42	0,41

<b>Skala 9 GESUNDHEITSSORGEN, Cronbach's alpha = 0,75</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V127 Auch ohne ernste Beschwerden gehe ich regelmäßig zum Arzt, nur zur Vorsicht	0,26	0,35
V70 Ich hole sicherheitshalber ärztlichen Rat ein, wenn ich länger als zwei Tage erhöhte Temperatur (leichtes Fieber) habe	0,37	0,32
V31 Ich habe mich über die häufigsten Krankheiten und ihre ersten Anzeichen informiert	0,38	0,36
V89 Ich passe auf, dass ich nicht zu viel Autoabgase und Staub einatme	0,42	0,42
V65 Ich vermeide Zugluft, weil man sich zu leicht erkälten kann	0,47	0,47
V38 Um gesund zu bleiben, achte ich auf ein ruhiges Leben	0,48	0,41
V84 Weil man sich so leicht anstecken kann, wasche ich mir zu Hause gleich die Hände	0,53	0,47
V18 Ich achte aus gesundheitlichen Gründen auf regelmäßige Mahlzeiten und reichlichen Schlaf	0,56	0,43
V10 Im Krankheitsfall möchte ich Befund und Behandlung eigentlich von einem zweiten Arzt überprüfen lassen	0,62	0,23
V62 Ich vermeide es, ungewaschenes Obst zu essen	0,64	0,43
V117 Handtücher in viel benutzten Wächräumen sind mir wegen der Ansteckungsgefahr unangenehm	0,75	0,36
V68 Wenn jemand in meine Richtung hustet oder niest, versuche ich mich abzuwenden	0,75	0,39

<b>Skala 10 OFFENHEIT, Cronbach's alpha = 0,75</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V114 Ich spreche manchmal über Dinge, von denen ich nichts verstehe	0,26	0,37
V78 Ich spreche oft Drohungen aus, die ich gar nicht ernst meine	0,29	0,35
V95 Manchmal habe ich Gedanken, über die ich mich schämen muss	0,30	0,34
V83 Meine Tischmanieren sind zu Hause schlechter als im Restaurant	0,41	0,43
V71 Hin und wieder gebe ich ein bisschen an	0,51	0,43
V07 Manchmal bin ich zu spät zu einer Verabredung oder zur Schule gekommen	0,52	0,28
V104 Wenn ich irgendwo zu Gast bin, ist mein Benehmen meistens besser als zu Hause	0,53	0,43
V77 Manchmal bin ich beleidigt, wenn es nicht nach meinem Willen geht	0,57	0,40
V09 Ich habe manchmal hässliche Bemerkungen über andere Menschen gemacht	0,58	0,44
V107 Ab und zu erzähle ich auch mal eine Lüge	0,60	0,44
V69 Ich bin hin und wieder ein wenig schadenfroh	0,66	0,44
V66 Manchmal schiebe ich etwas auf, was ich sofort tun sollte	0,76	0,23



<b>Skala 11 EXTRAVERSION, Cronbach's alpha = 0,79</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V20_A Sind wir in ausgelassener Runde, so überkommt mich oft eine große Lust zu groben Streichen	0,22	0,31
V91 Ich spiele anderen Leuten gern einen harmlosen Streich	0,34	0,34
V17 Ich bin unternehmungslustiger als die meisten meiner Bekannten	0,39	0,39
V109_A Bei Geselligkeiten und öffentlichen Veranstaltungen bleibe ich lieber im Hintergrund (stimmt nicht)	0,40	0,45
V53 Ich kann in eine ziemlich langweilige Gesellschaft schnell Leben bringen	0,43	0,62
V32 Ich übernehme bei gemeinsamen Unternehmungen gern die Führung	0,43	0,46
V02 Ich gehe abends gerne aus	0,44	0,30
V81_A Ich schließe nur langsam Freundschaften (stimmt nicht)	0,44	0,34
V39_A Ich habe gern mit Aufgaben zu tun, die schnelles Handeln verlangen	0,50	0,37
V76 Ich bin ziemlich lebhaft	0,51	0,56
V04_A Ich habe fast immer eine schlagfertige Antwort bereit	0,55	0,47
V54_A Bei wichtigen Dingen bin ich bereit, mit anderen energisch zu konkurrieren	0,58	0,39
V59 Ich würde mich selbst als eher gesprächig bezeichnen	0,66	0,47
V25 In einer vergnügten Gesellschaft kann ich meistens ungezwungen und unbeschwert sein	0,88	0,30

<b>Skala 12 EMOTIONALITÄT, Cronbach's alpha = 0,82</b>	<b>p<sub>i</sub></b>	<b>r<sub>it</sub></b>
V42 Meine Familie und meine Bekannten können mich im Grunde kaum richtig verstehen	0,20	0,33
V45 Ich fühle mich oft wie ein Pulverfass kurz vor der Explosion	0,21	0,44
V126 Meine Laune wechselt ziemlich oft	0,27	0,48
V19 Ich habe manchmal ein Gefühl der Teilnahmslosigkeit und inneren Leere	0,30	0,48
V82 Manchmal habe ich ohne eigentlichen Grund ein Gefühl unbestimmter Gefahr oder Angst	0,34	0,50
V55 Ich mache mir oft Sorgen um meine Gesundheit	0,39	0,38
V49 Termindruck und Hektik lösen bei mir körperliche Beschwerden aus	0,41	0,44
V112_A Ich grübele viel über mein bisheriges Leben nach	0,42	0,49
V28_A Ich bin oft nervös, weil zu viel auf mich einströmt	0,42	0,52
V130_A Ich habe häufig das Gefühl, im Stress zu sein	0,43	0,42
V110 Ich träume tagsüber oft von Dingen, die doch nicht verwirklicht werden können	0,46	0,32
V106 Es gibt Zeiten, in denen ich ganz traurig und niedergedrückt bin	0,47	0,54
V115_A Oft rege ich mich zu rasch über jemanden auf	0,48	0,38
V79_A Ich bin häufiger abgespannt, matt und erschöpft	0,48	0,53

<b>Deskriptive Statistiken</b>	<b>M</b>	<b>s</b>
Skala 1 LEBENSZUFRIEDENHEIT	7,65	2,98
Skala 2 SOZIALE ORIENTIERUNG	6,37	2,80
Skala 3 LEISTUNGSORIENTIERUNG	6,81	3,07
Skala 4 GEHEMMTHEIT	4,69	3,03
Skala 5 ERREGBARKEIT	4,91	3,08
Skala 6 AGGRESSIVITÄT	3,76	2,81
Skala 7 BEANSPRUCHUNG	5,26	3,41
Skala 8 KÖRPERLICHE BESCHWERDEN	3,23	2,77
Skala 9 GESUNDHEITSSORGEN	6,23	2,97
Skala 10 OFFENHEIT	5,96	2,95
Skala E EXTRAVERSION	6,75	3,47
Skala N EMOTIONALITÄT	5,30	3,64

**FPI-R Items der Normstichproben 1982 und 2018 (nur West-Deutschland):**  
**Zustimmungshäufigkeiten nach Erhebungsjahr und Geschlecht (nennenswerte Differenzen)**

				1982 West D		2018 West D		Diff. 1982 minus 2018		
				M	F	M	F	M	F	
FPI - Iteminhalt (verkürzt)				Skala		stimmt %		Diff. stimmt %		
1	I12 Wenn meinem Freund Böses getan wird, zahle ich es heim	1	AGGR			32%	23%	43%	33%	-11% ↑ -9% ↑
2	I15 Schon so wütend gewesen, Dinge zerrissen oder zerschlagen	2	AGGR			44%	40%	53%	49%	-8% ↑ -10% ↑
3	I37 Habe als Kind gerne anderen die Arme umgedreht, an Haaren gezogen	3	AGGR			36%	26%	27%	15%	9% ↓ 11% ↓
4	I50 Verteidige meine Rechte auch mit körperlicher Gewalt	4	AGGR			34%	16%	20%	10%	14% ↓ 6% ↓
5	I56 Wenn mich jemand anschreit, schreie ich zurück	5	AGGR			63%	59%	51%	44%	12% ↓ 14% ↓
6	I132 Gehe lieber zum Äußersten, als dass ich feige bin	6	AGGR			53%	48%	41%	37%	12% ↓ 11% ↓
7	I137 Gab Leute, die mich so ärgerten, dass es zu tätlicher Auseinandersetzung kam	7	AGGR			32%	17%	21%	12%	11% ↓ 6% ↓
8	I105 Kann oft meinen Ärger und meine Wut nicht beherrschen	1	ERR			34%	36%	19%	19%	15% ↓ 17% ↓
9	I113 Neige oft zu Hast und Eile, auch wenn es nicht notwendig ist	2	ERR			44%	53%	31%	39%	13% ↓ 14% ↓
10	I135 Bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde	3	ERR			48%	60%	42%	48%	6% ↓ 12% ↓
11	I52 Auch wenn es viel zu tun gibt, lasse ich mich nicht hetzen	4	ERR-			59%	43%	64%	55%	-6% ↓ -12% ↓
12	I108 Lasse mich durch viele Störungen nicht aus der Ruhe bringen	5	ERR-			62%	49%	68%	61%	-6% ↓ -12% ↓
13	I78 Spreche oft Drohungen aus, die ich nicht ernst meine	1	OFF			41%	38%	32%	26%	9% ↓ 12% ↓
14	I130 Habe häufig das Gefühl, im Stress zu sein	1	BEAN			50%	53%	40%	43%	10% ↓ 10% ↓
15	I11 Bin ungern mit Menschen zusammen, die ich nicht kenne	1	GEH			40%	44%	31%	32%	9% ↓ 12% ↓
16	I63 Fällt mir schwer, vor großer Gruppe Menschen zu sprechen	2	GEH			51%	70%	43%	52%	8% ↓ 18% ↓
17	I124 Fällt schwer, richtigen Gesprächsstoff zu finden, wenn ich jmd. kennenlernen will	3	GEH			49%	45%	38%	35%	11% ↓ 10% ↓
18	I25 In vergnügter Gesellschaft kann ich ungezwungen sein	1	EXTR			81%	75%	88%	88%	-7% ↑ -13% ↑
19	I53 Kann in eine langweilige Gesellschaft schnell Leben bringen	2	EXTR			39%	29%	46%	42%	-6% ↑ -13% ↑
20	I10 Möchte im Krankheitsfall von zweitem Arzt überprüft werden	1	GES			33%	44%	61%	67%	-27% ↑ -23% ↑
21	I38 Um gesund zu bleiben, achte ich auf ruhiges Leben	2	GES			36%	36%	45%	51%	-9% ↑ -15% ↑
22	I84 Wasche mir zu Hause gleich die Hände, weil man sich leicht anstecken kann	4	GES			37%	46%	43%	60%	-6% ↑ -14% ↑
23	I70 Ärztlichen Rat, wenn ich zwei Tage und länger Fieber habe	3	GES			42%	52%	34%	40%	8% ↓ 11% ↓
24	I47 Manchmal das Gefühl, einen Kloß im Hals zu haben	1	KÖRP			28%	47%	24%	34%	4% ↓ 13% ↓
25	I57 Mein Herz beginnt manchmal zu jagen, unregelmäßig zu schlagen	2	KÖRP			39%	55%	25%	37%	13% ↓ 18% ↓
26	I67 Ich habe häufiger Verstopfung	3	KÖRP			14%	31%	11%	20%	3% ↓ 11% ↓
27	I99 Meine Hände sind häufiger zitterig	4	KÖRP			22%	26%	13%	15%	9% ↓ 11% ↓

Trend: Abnahme  
von aggressiver  
Erregbarkeit

Trend: Abnahme  
von Gehemmtheit

Trend: Zunahme von  
gesundheitsbewusstem  
Verhalten;  
Abnahme von  
körperlichen  
Beschwerden

Fortsetzung der Tabelle ...

FPI - Iteminhalt (verkürzt)				1982 West D		2018 West D		Diff. 1982 minus 2018				
				M	F	M	F	M	F			
				Skala		stimmt %		stimmt %			Diff. stimmt %	
28	I131 Ich habe eine gute Partnerbeziehung	2	LEB	81%	77%	65%	62%	15%	↓	14%	↓	Trend: <u>Zunahme</u> Lebenszufriedenheit, bis auf Partnerbeziehung (verschlechtert)
29	I23 Lebe mit mir in Frieden, keine inneren Konflikte	1	LEB	56%	50%	65%	60%	-8%	↑	-11%	↑	
30	I58 Konnte bisher kaum verwirklichen, was in mir steckt	3	LEB-	42%	46%	30%	32%	13%	↑	14%	↑	
31	I94 Oft habe ich alles gründlich satt	4	LEB-	48%	56%	32%	34%	16%	↑	22%	↑	
												kein klarer Trend
32	I05 Gebe mir bei Arbeit mehr Mühe als andere	1	LEI	49%	43%	63%	59%	-14%	↑	-16%	↑	
33	I44 Pflege schnell und sicher zu handeln	3	LEI	65%	54%	69%	65%	-4%	↑	-11%	↑	
34	I39 Habe gerne Aufgaben, die schnelles Handeln verlangen	2	LEI	63%	51%	53%	48%	10%	↓	3%	↓	
35	I48 Mit anderen zu wetteifern, macht mir Spaß	4	LEI	65%	45%	51%	36%	13%	↓	9%	↓	
36	I54 Bin bei wichtigen Dingen bereit, mit anderen zu konkurrieren	5	LEI	76%	62%	62%	54%	13%	↓	8%	↓	Trend: Abnahme soz. Orientierung
37	I111 Spende gelegentlich für Katastrophenhilfe, Caritas usw.	1	SOZ	57%	71%	48%	57%	9%	↓	14%	↓	
38	I34 Jeder Mensch soll sehen, wie er zurecht kommt	2	SOZ-	56%	50%	48%	39%	7%	↓	10%	↓	Trend: Abnahme Emotionalität
39	I19 Manchmal ein Gefühl der Teilnahmslosigkeit, innerer Leere	1	EMOT	35%	44%	28%	33%	7%	↓	11%	↓	
40	I82 Habe ohne Grund manchmal ein Gefühl von Gefahr oder Angst	2	EMOT	37%	56%	28%	40%	9%	↓	16%	↓	
41	I106 Gibt Zeiten, in denen ich ganz traurig, niedergedrückt bin	3	EMOT	52%	73%	39%	56%	13%	↓	18%	↓	

Anmerkung: 41 von 137 FPI-R-Items (30%) mit Zeitdifferenzen; ≥ 10% bei Männern und/oder Frauen

FPI - Itemprofile 1982 vs. 2018 (West-Deutschland); Veränderungen in den Zustimmungshäufigkeiten (Differenz %, Veränderungen ≥ 10%)

	Skala	LEB	SOZ	LEI	GEH	ERR	AGGR	BEAN	KÖRP	GES	OFF	EXTR	EMOT	total	total%
	n Items	12	12	12	12	12	12	12	12	12	12	8	9	137	100%
Männer	n Items verändert	3	0	4↓	1	2	5↓	0	1	1	0	0	1	18	13%
Frauen	n Items verändert	4↓	2	2	3	5↓	4↓	1	5↓	4↑	1	2	3	36	26%

↓ Abnahme    ↑ Zunahme

**FPI-R Items der Normstichproben 1999 und 2018: Zustimmungshäufigkeiten (%) nach Erhebungsjahr und Geschlecht**

	Erhebungsjahr		1999	2018	Diff.	1999	2018	Diff.
	Geschlecht		Männer			Frauen		
FPI-Iteminhalt (verkürzt)	Skala		stimmt%			stimmt%		
Habe Anleitung gelesen und bin bereit, jeden Satz offen zu beantworten			98%	99%	1%	98%	99%	1%
Bin immer guter Laune	1	LEB	36%	37%	1%	39%	37%	-2%
Wollte nicht anders leben, wenn ich noch einmal geboren würde	1	LEB	54%	54%	0%	53%	53%	0%
Grüble viel über mein bisheriges Leben nach (stimmt nicht)	1	LEB- EMOT	61%	59%	-2%	54%	56%	-2%
Konnte bisher kaum verwirklichen, was in mir steckt (stimmt nicht)	1	LEB-	65%	70%	5%	63%	68%	-5%
Oft habe ich alles gründlich satt (stimmt nicht)	1	LEB-	65%	66%	1%	57%	65%	-8%
Lebe mit mir in Frieden, keine inneren Konflikte	1	LEB	67%	65%	-2%	61%	60%	-1%
Bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden (stimmt nicht)	1	LEB-	67%	70%	3%	66%	68%	-2%
Bin selten bedrückt, unglücklich	1	LEB	67%	66%	-1%	63%	63%	-1%
Habe Beruf, der mich befriedigt	1	LEB	71%	70%	-1%	62%	67%	5%
Bin ausgesprochen zufrieden mit meinem bisherigen Leben	1	LEB	75%	76%	1%	74%	78%	4%
Ich habe eine gute Partnerbeziehung	1	LEB	77%	67%	-11%	74%	65%	-9%
Meistens blicke ich voller Zuversicht in die Zukunft	1	LEB	78%	74%	-3%	78%	76%	-2%
Lebenszufriedenheit	Mittelwert		65%	65%	-1%	62%	63%	-2%
Denke, ich sollte meinen Konsum einschränken, um Benachteiligten abzugeben	2	SOZ	18%	21%	3%	31%	25%	-5%
Mir ist es lästig, wenn mich ein Fremder um eine Geldspende bittet (stimmt nicht)	2	SOZ-	31%	37%	6%	42%	41%	1%
Habe schon unbezahlt in Gemeinde oder sozialer Einrichtung geholfen	2	SOZ	33%	38%	5%	38%	43%	5%
Häufig schlechtes Gewissen, wenn ich sehe, wie schlecht es anderen geht	2	SOZ	37%	41%	4%	53%	51%	-1%
Staat sorgt für Sozialhilfe, da brauche ich nicht zu helfen (stimmt nicht)	2	SOZ-	43%	48%	5%	58%	60%	-2%
Würde kaum zögern, alte und behinderte Menschen zu pflegen	2	SOZ	45%	44%	-1%	64%	59%	-5%
Jeder Mensch soll sehen, wie er zurechtkommt (stimmt nicht)	2	SOZ-	45%	50%	5%	56%	60%	-4%
Menschen in Entwicklungsländern sollten sich erst selbst helfen (stimmt nicht)	2	SOZ-	47%	44%	-3%	58%	53%	5%
Spende gelegentlich für Katastrophenhilfe, Caritas usw.	2	SOZ	50%	45%	-5%	62%	56%	-6%
Wenn jemand weint, möchte ich ihn umarmen, trösten	2	SOZ	57%	59%	2%	85%	84%	-1%
Fühle mich auch über Familie hinaus für andere verantwortlich	2	SOZ	67%	69%	2%	78%	75%	-3%
Nehme mir zu viel Zeit, anderen Menschen zuzuhören, wenn sie von Sorgen erzählen	2	SOZ	69%	70%	1%	84%	86%	1%
Soziale Orientierung	Mittelwert		45%	47%	2%	59%	58%	-1%

Fortsetzung nächste Seite

	Erhebungsjahr	1999	2018	Diff.	1999	2018	Diff.
	Geschlecht	Männer			Frauen		
Berufliche Aufgaben wichtiger als Freizeit und Hobbys	3 LEI	41%	37%	-3%	31%	33%	2%
Bin bei meiner Arbeit meist schneller als andere	3 LEI	53%	47%	-6%	46%	47%	1%
Bekannte halten mich für energisch	3 LEI	57%	52%	-4%	56%	56%	0%
Habe gerne Aufgaben, die schnelles Handeln verlangen	3 LEI EXTR	57%	52%	-5%	49%	48%	-2%
Mit anderen zu wetteifern, macht mir Spaß	3 LEI	58%	51%	-7%	42%	37%	-5%
Bin leicht beim Ehrgeiz zu packen	3 LEI	61%	63%	2%	56%	59%	2%
Gebe mir bei Arbeit mehr Mühe als andere	3 LEI	64%	62%	-2%	55%	61%	6%
Ziehe das Handeln dem Pläneschmieden vor	3 LEI	71%	65%	-6%	63%	62%	-1%
Bin bei wichtigen Dingen bereit, mit anderen zu konkurrieren	3 LEI EXTR	72%	63%	-9%	57%	53%	-3%
Habe Spaß an schwierigen Aufgaben	3 LEI	72%	69%	-3%	58%	60%	3%
Pflege schnell und sicher zu handeln	3 LEI	73%	68%	-5%	64%	66%	2%
Gibt noch viele sinnvolle Aufgaben, die ich anpacken werde	3 LEI	81%	79%	-2%	73%	72%	-1%
<i>Leistungsorientierung</i>	Mittelwert	<b>63%</b>	<b>59%</b>	<b>-4%</b>	<b>54%</b>	<b>55%</b>	<b>0%</b>
Ich werde leicht rot	4 GEH	17%	17%	0%	31%	31%	0%
Ich werde ziemlich leicht verlegen	4 GEH	21%	23%	3%	36%	33%	-3%
Würde mich im Restaurant über schlechtes Essen beschweren (stimmt nicht)	4 GEH-	22%	27%	5%	29%	29%	0%
Scheue mich, allein in Raum in dem andere sitzen zu gehen	4 GEH	22%	22%	0%	33%	31%	-2%
Bin im Grunde ein ängstlicher Mensch	4 GEH	23%	24%	1%	43%	41%	-2%
Bin ungern mit Menschen zusammen, die ich nicht kenne	4 GEH	34%	32%	-2%	38%	33%	-5%
Habe immer eine schlagfertige Antwort (stimmt nicht)	4 GEH- EXTR	39%	41%	2%	47%	49%	-2%
Fällt schwer, richtigen Gesprächsstoff zu finden, wenn ich jmd. kennenlernen will	4 GEH	40%	40%	0%	39%	36%	-3%
Fällt mir schwer, vor großer Gruppen Menschen zu sprechen	4 GEH	49%	47%	-3%	62%	53%	-9%
Schließe nur langsam Freundschaften	4 GEH EXTR+	57%	57%	0%	55%	56%	2%
Schaue bei Reisen lieber auf die Landschaft, als mich mit Mitreisenden zu unterhalten	4 GEH	58%	56%	-1%	57%	52%	-5%
Bei Geselligkeiten und Veranstaltungen bleibe ich lieber im Hintergrund	4 GEH EXTR+	58%	58%	0%	64%	63%	0%
<i>Gehemmtheit</i>	Mittelwert	<b>37%</b>	<b>37%</b>	<b>0%</b>	<b>45%</b>	<b>42%</b>	<b>-2%</b>
Kann oft meinen Ärger und meine Wut nicht beherrschen	5 ERR	24%	21%	-3%	25%	18%	-7%
Neige oft zu Hast und Eile, auch wenn es nicht notwendig ist	5 ERR	35%	32%	-3%	42%	39%	-3%
Wenn etwas schief geht, regt mich das nicht auf	5 ERR-	41%	39%	-2%	40%	39%	-1%
Rege mich oft zu rasch über jemanden auf	5 ERR EMOT	46%	48%	2%	50%	50%	-1%
Bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde	5 ERR	48%	44%	-4%	58%	49%	-8%
Koche vor Wut, wenn man mich zum Narren hält	5 ERR	54%	51%	-3%	57%	50%	-7%
Auch wenn es viel zu tun gibt, lasse ich mich nicht hetzen	5 ERR-	63%	65%	2%	50%	55%	4%
Nur wenige Dinge erregen oder ärgern mich leicht	5 ERR-	64%	64%	0%	56%	58%	2%
Lasse mich durch viele Störungen nicht aus der Ruhe bringen	5 ERR-	68%	68%	0%	61%	60%	0%
Bin im Allgemeinen ruhig und nicht leicht aufzuregen	5 ERR-	76%	75%	-1%	70%	72%	2%
Spreche bei Auseinandersetzungen lauter als sonst	5 ERR	77%	73%	-4%	73%	69%	-4%
Beruhige mich rasch wieder, wenn mich etwas aus der Fassung bringt	5 ERR-	79%	75%	-4%	69%	71%	2%
<i>Erregbarkeit</i>	Mittelwert	<b>56%</b>	<b>55%</b>	<b>-2%</b>	<b>54%</b>	<b>53%</b>	<b>-2%</b>

Fortsetzung nächste Seite

	Erhebungsjahr	1999	2018	Diff.	1999	2018	Diff.
	Geschlecht	Männer			Frauen		
Macht mir Spaß, anderen Fehler nachzuweisen	6 AGGR	28%	25%	-3%	16%	15%	0%
Verteidige meine Rechte auch mit körperlicher Gewalt	6 AGGR	28%	21%	-7%	14%	10%	-4%
Gab Leute, die mich so ärgerten, dass es zu tätlicher Auseinandersetzung kam	6 AGGR	30%	22%	-8%	20%	11%	-9%
In ausgelassener Runde bekomme Lust zu groben Streichen	6 AGGR EXTR	33%	28%	-5%	25%	18%	-7%
Habe als Kind gerne anderen die Arme umgedreht, an Haaren gezogen	6 AGGR	37%	30%	-7%	25%	15%	-9%
Wütend bin ich in der Lage, jemandem eine runterzuhauen	6 AGGR	38%	29%	-9%	30%	22%	-8%
Wenn meinem Freund Böses getan wird, zahle ich es heim	6 AGGR	49%	44%	-4%	35%	33%	-2%
Wünsche dem, der mich schlecht behandelt oder beleidigt, eine harte Strafe	6 AGGR	49%	48%	-1%	41%	39%	-2%
Male mir aus, wie übel es denen gehen sollte, die mir Unrecht tun	6 AGGR	51%	48%	-3%	41%	43%	2%
Gehe lieber zum Äußersten, als dass ich feige bin	6 AGGR	51%	42%	-9%	43%	37%	-6%
Schon so wütend gewesen, Dinge zerrissen oder zerschlagen	6 AGGR	55%	53%	-2%	49%	48%	-1%
Wenn mich jemand anschreit, schreie ich zurück	6 AGGR	58%	52%	-6%	55%	46%	-9%
<b>Aggressivität</b>	Mittelwert	<b>42%</b>	<b>37%</b>	<b>-5%</b>	<b>33%</b>	<b>28%</b>	<b>-5%</b>
Anforderungen, die an mich gestellt werden, sind oft zu hoch	7 BEAN	22%	25%	3%	27%	29%	2%
Bin manchmal durch viele Aufgaben und Zeitdruck ganz durcheinander	7 BEAN	31%	31%	0%	35%	38%	3%
Bin oft nervös, weil zu viel auf mich einströmt	7 BEAN EMOT	40%	38%	-2%	51%	47%	-3%
Kann nur selten richtig abschalten	7 BEAN	41%	38%	-3%	44%	46%	2%
Bin häufiger abgespannt, matt und erschöpft	7 BEAN EMOT	46%	45%	-1%	57%	53%	-4%
Bin auch an Wochenenden stark eingespannt	7 BEAN	47%	39%	-7%	43%	41%	-2%
Habe häufig das Gefühl, im Stress zu sein	7 BEAN EMOT	47%	43%	-4%	49%	46%	-3%
Bin von täglichen Belastungen oft müde und erschöpft	7 BEAN	47%	43%	-5%	52%	51%	-2%
Ich denke manchmal, dass ich zu viel arbeite	7 BEAN	51%	45%	-6%	41%	44%	2%
Ich arbeite oft unter Zeitdruck	7 BEAN	57%	51%	-6%	48%	49%	1%
Denke manchmal, dass ich mich mehr schonen sollte	7 BEAN	58%	58%	0%	62%	62%	0%
Hätte gern mehr Zeit ohne viele Verpflichtungen	7 BEAN	58%	58%	0%	58%	55%	-3%
<b>Beanspruchung</b>	Mittelwert	<b>45%</b>	<b>43%</b>	<b>-3%</b>	<b>47%</b>	<b>47%</b>	<b>-1%</b>
Ich habe häufiger Verstopfung	8 KÖRP	9%	10%	1%	23%	20%	-3%
Meine Hände sind häufiger zitterig	8 KÖRP	13%	14%	0%	15%	15%	0%
Habe häufig Kopfschmerzen	8 KÖRP	14%	17%	3%	37%	28%	-9%
Habe manchmal Hitzewallungen, Blut steigt in den Kopf	8 KÖRP	15%	16%	2%	34%	37%	3%
Bei Wärme häufiger kalte Hände und Füße	8 KÖRP	15%	15%	0%	43%	41%	-2%
Habe manchmal das Gefühl erstickender Enge in der Brust	8 KÖRP	17%	15%	-2%	24%	22%	-3%
Manchmal das Gefühl, einen Kloß im Hals zu haben	8 KÖRP	24%	26%	2%	38%	35%	-3%
Mein Herz beginnt manchmal zu jagen, unregelmäßig zu schlagen	8 KÖRP	25%	25%	0%	40%	37%	-4%
Bemerke häufiger unwillkürliches Zucken, z.B. um die Augen	8 KÖRP	27%	27%	0%	35%	31%	-5%
Habe Schwierigkeiten, einzuschlafen oder durchzuschlafen	8 KÖRP	31%	34%	3%	44%	48%	4%
Habe einen empfindlichen Magen	8 KÖRP	31%	28%	-3%	40%	37%	-3%
Mein Körper reagiert deutlich auf Wetteränderung	8 KÖRP	33%	30%	-3%	54%	46%	-8%
<b>Körperliche Beschwerden</b>	Mittelwert	<b>21%</b>	<b>21%</b>	<b>0%</b>	<b>36%</b>	<b>33%</b>	<b>-3%</b>

Fortsetzung nächste Seite

	Erhebungsjahr		1999	2018	Diff.	1999	2018	Diff.
	Geschlecht		Männer			Frauen		
Gehe auch ohne Beschwerden regelmäßig zum Arzt	9	GES	25%	23%	-2%	34%	28%	-5%
Wasche mir zu Hause gleich die Hände, weil man sich leicht anstecken kann	9	GES	32%	43%	11%	43%	60%	16%
Habe mich über häufigste Krankheiten und Anzeichen informiert	9	GES	34%	30%	-4%	46%	44%	-2%
Ärztlichen Rat, wenn ich zwei Tage und länger Fieber habe	9	GES	37%	34%	-3%	50%	40%	-9%
Um gesund zu bleiben, achte ich auf ruhiges Leben	9	GES	41%	45%	4%	48%	50%	3%
Vermeide Zugluft, weil man sich leicht erkältet	9	GES	44%	40%	-4%	59%	54%	-5%
Achte aus gesundheitlichen Gründen auf regelmäßige Mahlzeiten, reichlichen Schlaf	9	GES	45%	50%	5%	60%	62%	1%
Passe auf, dass ich nicht zu viel Abgase und Staub einatme	9	GES	48%	38%	-10%	51%	46%	-4%
Vermeide es, ungewaschenes Obst zu essen	9	GES	56%	54%	-2%	73%	72%	0%
Möchte im Krankheitsfall von zweitem Arzt überprüft werden	9	GES	57%	58%	1%	61%	65%	4%
Handtücher in viel benutzten Waschräumen sind mir unangenehm	9	GES	65%	67%	2%	77%	81%	4%
Versuche mich abzuwenden, wenn jemand zu mir hustet oder niest	9	GES	70%	69%	-2%	73%	80%	7%
<i>Gesundheitssorgen</i>	Mittelwert		46%	46%	0%	56%	57%	1%
Spreche manchmal über Dinge, von denen ich nichts verstehe	10	OFF	25%	26%	1%	25%	27%	2%
Habe manchmal Gedanken, über die ich mich schämen muss	10	OFF	31%	32%	1%	32%	30%	-2%
Spreche oft Drohungen aus, die ich nicht ernst meine	10	OFF	42%	33%	-9%	36%	28%	-8%
Meine Tischmanieren sind zu Hause schlechter als im Restaurant	10	OFF	45%	50%	5%	29%	34%	6%
Wenn ich Gast bin, ist mein Benehmen besser als zu Hause	10	OFF	56%	62%	6%	41%	46%	6%
Hin und wieder gebe ich ein bisschen an	10	OFF	60%	60%	0%	46%	44%	-2%
Manchmal zu spät zu Verabredung, zur Schule	10	OFF	61%	57%	-4%	51%	48%	-3%
Habe schon schlechte Bemerkungen über andere gemacht	10	OFF	61%	63%	2%	60%	56%	-4%
Manchmal bin ich beleidigt, wenn es nicht nach meinem Kopf geht	10	OFF	62%	59%	-3%	64%	58%	-6%
Erzähle ab und zu auch mal eine Lüge	10	OFF	66%	66%	0%	55%	54%	0%
Bin hin und wieder ein wenig schadenfroh	10	OFF	74%	72%	-2%	69%	64%	-5%
Schiebe manchmal etwas auf, was ich sofort tun sollte	10	OFF	78%	78%	0%	80%	74%	-6%
<i>Offenheit</i>	Mittelwert		55%	55%	0%	49%	47%	-2%
In ausgelassener Runde bekomme ich Lust zu groben Streichen	6	AGGR EXTR	33%	28%	-5%	25%	18%	-7%
Bin unternehmungslustiger als die meisten meiner Bekannten	11	EXTR	40%	37%	-3%	39%	40%	2%
Kann in eine langweilige Gesellschaft schnell Leben bringen	11	EXTR	43%	45%	3%	36%	41%	4%
Schließe nur langsam Freundschaften (stimmt nicht)	4	GEH- EXTR	43%	43%	0%	45%	44%	2%
Spiele anderen gern einen harmlosen Streich	11	EXTR	50%	45%	-5%	33%	26%	-7%
Gehe abends gerne aus	11	EXTR	50%	47%	-4%	50%	43%	-7%
Habe immer eine schlagfertige Antwort (stimmt)	4	GEH- EXTR	51%	59%	8%	53%	51%	-2%
Übernehme bei gemeinsamen Unternehmungen gern die Führung	11	EXTR	51%	48%	-3%	35%	37%	2%
Bin ziemlich lebhaft	11	EXTR	52%	46%	-6%	58%	55%	-3%
Habe gerne Aufgaben, die schnelles Handeln verlangen	3	LEI EXTR	57%	52%	-5%	49%	48%	-2%
Bei Geselligkeiten und Veranstaltungen bleibe ich lieber im Hintergrund (stimmt nicht)	4	GEH- EXTR	58%	58%	0%	64%	63%	0%

Fortsetzung nächste Seite

	Erhebungsjahr		1999	2018	Diff.	1999	2018	Diff.
	Geschlecht		Männer			Frauen		
Würde mich als gesprächig bezeichnen	11	EXTR	64%	60%	-5%	69%	70%	1%
Bin bei wichtigen Dingen bereit, mit anderen zu konkurrieren	3	LEI EXTR	72%	63%	-9%	57%	53%	-3%
In vergnügter Gesellschaft kann ich ungezwungen sein	11	EXTR	85%	88%	3%	79%	88%	9%
<i>Extraversion</i>	Mittelwert		<b>54%</b>	<b>51%</b>	<b>-2%</b>	<b>49%</b>	<b>48%</b>	<b>-1%</b>
Familie, Bekannte können mich kaum richtig verstehen	12	EMOT	25%	21%	-4%	25%	20%	-5%
Fühle mich oft wie ein Pulverfass vor der Explosion	12	EMOT	26%	22%	-4%	30%	23%	-7%
Manchmal ein Gefühl der Teilnahmslosigkeit, innerer Leere	12	EMOT	26%	29%	3%	38%	32%	-6%
Meine Laune wechselt ziemlich oft	12	EMOT	29%	26%	-3%	32%	29%	-3%
Habe ohne Grund manchmal ein Gefühl von Gefahr oder Angst	12	EMOT	32%	28%	-4%	49%	40%	-8%
Termindruck, Hektik lösen körperliche Beschwerden aus	12	EMOT	33%	37%	3%	48%	48%	0%
Mache mir oft Sorgen um meine Gesundheit	12	EMOT	37%	37%	0%	49%	43%	-6%
Bin oft nervös, weil zu viel auf mich einströmt	7	BEAN EMOT	40%	38%	-2%	51%	47%	-3%
Gibt Zeiten, in denen ich ganz traurig, niedergedrückt bin	12	EMOT	45%	38%	-7%	64%	54%	-10%
Bin häufiger abgespannt, matt und erschöpft	7	BEAN EMOT	46%	45%	-1%	57%	53%	-4%
Rege mich oft zu rasch über jemanden auf	5	ERR EMOT	46%	48%	2%	50%	50%	-1%
Habe häufig das Gefühl, im Stress zu sein	7	BEAN EMOT	47%	43%	-4%	49%	46%	-3%
Träume tagsüber von Dingen, die nicht verwirklicht werden können	12	EMOT	49%	47%	-2%	49%	47%	-2%
Grüble viel über mein bisheriges Leben nach (stimmt nicht)	1	LEB- EMOT	61%	59%	-2%	54%	56%	-2%
<i>Emotionalität</i>	Mittelwert		<b>39%</b>	<b>37%</b>	<b>-2%</b>	<b>46%</b>	<b>42%</b>	<b>-4%</b>

Zusammenfassung der Tabelle oben

	Männer: Differenz 1999 2018		Frauen: Differenz 1999 2018		1999: Differenz Frauen   Männer	2018: Differenz Frauen   Männer
Skala	2018 minus 1999		2018 minus 1999		Frauen minus - Männer	Frauen minus - Männer
KÖRP	0%	-3%	0%	-3%	14%	12%
GES	0%	1%	0%	1%	10%	11%
SOZ	2%	-1%	2%	-1%	14%	11%
GEH	0%	-2%	0%	-2%	8%	5%
EMOT	-2%	-4%	-2%	-4%	7%	5%
BEAN	-3%	-1%	-3%	-1%	2%	4%
AGGR	-5%	-5%	-5%	-5%	-9%	-9%
LEI	-4%	0%	-4%	0%	-9%	-5%
OFF	0%	-2%	0%	-2%	-6%	-8%
EXTR	-2%	-1%	-2%	-1%	-4%	-3%
LEB	-1%	-2%	-1%	-2%	-3%	-2%
ERR	-2%	-2%	-2%	-2%	-2%	-2%



**FPI-R Items der Normstichproben 1982, 1999, 2018: Zustimmungshäufigkeiten (%) nach Altersgruppen: 16-25 und 60-70 Jahre**

Befragte im Alter von 16-25 Jahren (nur West-Deutschland)			1982	2018	Diff. ≥ 15%	
Item	Iteminhalt	Skala	stimmt%	stimmt%	stimmt%	
131	Ich habe eine gute Partnerbeziehung (Ehe)	LEB	71%	36%	-35%	↓
106	Es gibt Zeiten, in denen ich ganz traurig und niedergedrückt bin	EMOT	73%	47%	-26%	↓
82	Manchmal habe ich ohne eigentlichen Grund ein Gefühl unbestimmter Gefahr oder Angst	EMOT	46%	26%	-21%	↓
94	Oft habe ich alles gründlich satt	LEB-	62%	43%	-19%	↓
111	Ich gebe gelegentlich Geld u. Spenden für Katastrophenhilfe, Caritas, Brot für die Welt u. and. gemeinnützige Org.	SOZ	41%	22%	-19%	↓
57	Mein Herz beginnt manchmal zu jagen oder unregelmäßig zu schlagen	KÖRP	36%	18%	-18%	↓
54	Bei wichtigen Dingen bin ich bereit, mit anderen energisch zu konkurrieren	LEI	78%	61%	-17%	↓
30	Wenn mir einmal etwas schief geht, regt mich das nicht weiter auf	ERR-	49%	33%	-16%	↓
113	Ich neige oft zu Hast und Eile, auch wenn es überhaupt nicht notwendig ist	ERR	49%	34%	-15%	↓
108	Ich lasse mich durch eine Vielzahl von kleinen Störungen nicht aus der Ruhe bringen	ERR-	47%	62%	15%	↑
23	Ich lebe mit mir selbst in Frieden und ohne innere Konflikte	LEB	39%	53%	15%	↑
121	Da der Staat schon für Sozialhilfe sorgt, brauche ich im Einzelnen nicht zu helfen	SOZ-	33%	48%	15%	↑
5	Ich glaube, dass ich mir beim Arbeiten mehr Mühe gebe als die meisten anderen Menschen	LEI	32%	50%	18%	↑
86	Ich koche vor Wut, wenn man mich zum Narren hält	ERR	40%	59%	19%	↑
12	Wenn jemand meinem Freund etwas Böses tut, bin ich dabei, wenn es heimgezahlt wird	AGGR	36%	59%	23%	↑

**Kein klarer Trend erkennbar**

Befragte im Alter von 16-25 Jahren (Ost- und West-Deutschland)			1999	2018	Diff. ≥ 10%	
Item	Iteminhalt	Skala	stimmt%	stimmt%	stimmt%	
131	Ich habe eine gute Partnerbeziehung (Ehe)	LEB	62%	35%	-27%	↓
90	Wenn ich wirklich wütend werde, bin ich in der Lage, jemandem eine runterzuhauen	AGGR	52%	34%	-18%	↓
137	Es gab Leute, die mich so ärgerten, dass es zu einer tätlichen Auseinandersetzung kam	AGGR	39%	23%	-16%	↓
20	Sind wir in ausgelassener Runde, so überkommt mich oft eine große Lust zu groben Streichen	AGGR	48%	32%	-16%	↓
132	Ich gehe lieber bis zum Äußersten, als dass ich feige bin	AGGR	53%	39%	-14%	↓
37	Als Kind habe ich manchmal ganz gerne anderen die Arme umgedreht, an Haaren gezogen, ein Bein gestellt usw.	AGGR	39%	25%	-14%	↓
56	Wenn mich jemand anschreit, schreie ich zurück	AGGR	67%	54%	-13%	↓
24	Ich male mir manchmal aus, wie übel es denen eigentlich ergehen müsste, die mir Unrecht tun	AGGR	59%	48%	-11%	↓
78	Ich spreche oft Drohungen aus, die ich gar nicht ernst meine	OFF	49%	36%	-13%	↓
106	Es gibt Zeiten, in denen ich ganz traurig und niedergedrückt bin	EMOT	58%	45%	-13%	↓
126	Meine Laune wechselt ziemlich oft	EMOT	49%	36%	-13%	↓
30	Wenn mir einmal etwas schief geht, regt mich das nicht weiter auf	ERR-	41%	31%	-11%	↓
91	Ich spiele anderen Leuten gern einen harmlosen Streich	EXTR	66%	55%	-11%	↓
8	Ich würde mich beim Kellner o. Geschäftsführer eines Restaurants beschweren, wenn schlechtes Essen serviert wird	GEH	68%	56%	-12%	↓
94	Oft habe ich alles gründlich satt	LEB-	47%	35%	-12%	↓
84	Weil man sich so leicht anstecken kann, wasche ich mir zu Hause gleich die Hände	GES	27%	40%	13%	

**Trend: Abnahme von Aggressivität**

	Befragte im Alter von 60-70 Jahren (nur West-Deutschland)		1982	2018	Diff. ≥ 15%	
Item	Iteminhalt	Skala	stimmt%	stimmt%	stimmt%	
123	Mein Körper reagiert deutlich auf Wetteränderung	KÖRP	72%	46%	-25%	↓
57	Mein Herz beginnt manchmal zu jagen oder unregelmäßig zu schlagen	KÖRP	62%	39%	-22%	↓
99	Meine Hände sind häufiger zitterig, z.B. beim Schreiben mit einem Stift oder Halten einer Tasse	KÖRP	33%	16%	-17%	↓
16	Ich habe häufig Kopfschmerzen	KÖRP	38%	22%	-16%	↓
33	Ich habe selbst bei warmem Wetter häufiger kalte Hände und Füße	KÖRP	44%	30%	-14%	↓
92	Ich habe einen empfindlichen Magen	KÖRP	47%	33%	-14%	↓
133	Ich habe manchmal ein Gefühl erstickender Enge in der Brust	KÖRP	35%	22%	-13%	↓
75	Ich habe Schwierigkeiten einzuschlafen oder durchzuschlafen	KÖRP	62%	49%	-13%	↓
67	Ich habe häufiger Verstopfung	KÖRP	33%	21%	-12%	↓
72	Ich bemerke häufiger ein unwillkürliches Zucken, z.B. um meine Augen	KÖRP	42%	31%	-11%	↓
10	Im Krankheitsfall möchte ich Befund und Behandlung eigentlich von einem zweiten Arzt überprüfen lassen	GES	46%	70%	25%	↑
65	Ich vermeide Zugluft, weil man sich zu leicht erkälten kann	GES	74%	54%	-20%	↓
70	Ich hole sicherheitshalber ärztlichen Rat ein, wenn ich länger als zwei Tage erhöhte Temperatur (leichtes Fieber) habe	GES	57%	36%	-21%	↓
132	Ich gehe lieber bis zum Äußersten, als dass ich feige bin	AGGR	52%	36%	-16%	↓
86	Ich koche vor Wut, wenn man mich zum Narren hält	ERR	36%	52%	16%	↑
25	In einer vergnügten Gesellschaft kann ich meistens ungezwungen und unbeschwert sein	EXTR	71%	87%	16%	↑
48	Mit anderen zu wetteifern, macht mir Spaß	LEI	48%	32%	-16%	↓
9	Ich habe manchmal hässliche Bemerkungen über andere Menschen gemacht	OFF	39%	56%	17%	↑
80	Ich bekomme häufig ein schlechtes Gewissen, wenn ich sehe, wie schlecht es anderen Menschen geht	SOZ	59%	43%	-16%	↓
111	Ich gebe gelegentlich Geld u. Spenden für Katastrophenhilfe, Caritas, Brot für die Welt und andere gemeinnützige Org.	SOZ	75%	52%	-23%	↓
	N		327	431		

#### Trend für gesamten Zeitraum 1982-2018: Abnahme von körperlichen Beschwerden und Gesundheitssorgen

	Befragte im Alter von 60-70 Jahren (Ost- und West-Deutschland)		1999	2018	Diff. ≥ 10%	
Item	Iteminhalt	Skala	stimmt%	stimmt%	stimmt%	
127	Auch ohne ernste Beschwerden gehe ich regelmäßig zum Arzt, nur zur Vorsicht	GES	45%	29%	-16%	↓
70	Ich hole sicherheitshalber ärztlichen Rat ein, wenn ich länger als zwei Tage erhöhte Temperatur (leichtes Fieber) habe	GES	51%	35%	-16%	↓
65	Ich vermeide Zugluft, weil man sich zu leicht erkälten kann	GES	70%	55%	-15%	↓
89	Ich passe auf, dass ich nicht zu viel Autoabgase und Staub einatme	GES	61%	49%	-12%	↓
18	Ich achte aus gesundheitlichen Gründen auf regelmäßige Mahlzeiten und reichlichen Schlaf	GES	74%	62%	-12%	↓
62	Ich vermeide es, ungewaschenes Obst zu essen	GES	75%	64%	-11%	↓
38	Um gesund zu bleiben, achte ich auf ein ruhiges Leben	GES	67%	56%	-11%	↓
123	Mein Körper reagiert deutlich auf Wetteränderung	GES	57%	47%	-10%	↓
92	Ich habe einen empfindlichen Magen	KÖRP	44%	31%	-13%	↓
111	Ich gebe gelegentlich Geld u. Spenden für Katastrophenhilfe, Caritas, Brot für die Welt und andere gemeinnützige Org.	SOZ	73%	54%	-19%	↓
80	Ich bekomme häufig ein schlechtes Gewissen, wenn ich sehe, wie schlecht es anderen Menschen geht	SOZ	53%	43%	-10%	↓
23	Ich lebe mit mir selbst in Frieden und ohne innere Konflikte	LEB	77%	65%	-12%	↓
118	Ich arbeite oft unter Zeitdruck	BEAN	27%	37%	10%	↑
43	Es gibt für mich noch eine Menge sinnvoller Aufgaben, die ich in der Zukunft anpacken werde	LEI	61%	71%	10%	↑
25	In einer vergnügten Gesellschaft kann ich meistens ungezwungen und unbeschwert sein	EXTR	76%	88%	12%	↑
	N		305	546		

## V. Faktorenanalysen - Ladungsmuster

### Erhebung 1982 (Normstichprobe West-Deutschland)

Itemstatistiken von 138 Items der 10 Standardskalen und 2 Sekundärskalen des FPI-R (max. N=2.035). Item-Skalen-Zuordnung (Schlüssel), Schwierigkeitsindex  $p_i$  und Trennschärfeindex  $r_{it}$  in der betreffenden Skala, Faktorenladungen  $a_i \geq 0.20$  der 10-Faktoren- bzw. der 2-Faktoren-Hauptkomponenten-Lösungen, partielle Korrelationskoeffizienten der Items (stimmt=1, stimmt nicht=2) mit Geschlecht (männlich=1, weiblich=2)  $r_{G,AS}$ , Alter (3 Stufen)  $r_{A,GS}$  und Schulbildung (2 Stufen)  $r_{S,GA}$ .

Item	Skala	$p_i$	$r_{it}$	1	2	3	4	5	6	7	8	9	10	$r_{G,AS}$	$r_{A,GS}$	$r_{S,GA}$
3	1	64	37	44		30								9	-15	-05
23	1	53	48	51				-26					-23	8	-22	4
29	1	50	39	54										5	-08	-09
88	1	37	31		35				28				-22	0	-05	7
100	1	62	38	44										7	2	-03
128	1	74	58	70										3	-07	0
131	1	79	35	48										5	-08	2
138	1	73	44	51										8	3	-04
58	-1	56	39	-48					21					-03	5	7
94	-1	48	45	-39							23		20	-08	13	11
112	-1	53	39	-43						32				-10	-09	4
119	-1	65	57	-59										-05	11	9
14	2	55	36		51									-13	-06	1
26	2	66	36		48	20								0	-06	-12
36	2	31	39		54									-07	-02	-09
41	2	71	31		47									-23	0	-07
80	2	49	36		52									-12	-10	-08
111	2	64	35		40						20			-13	-23	-07
134	2	36	29		36									-02	-13	-10
136	2	75	33		44									-12	-06	-10
22	-2	52	33		-45				26					12	-11	11
34	-2	48	32		-45									7	-08	10
87	-2	40	28		-39									11	10	4
121	-2	61	44		-57				23					10	1	20
5	3	46	32			47								7	-10	-03
13	3	55	38			42	-35							4	-07	-14
21	3	59	46			58								8	4	-03
39	3	56	56			61								11	4	-08
43	3	75	32		21	31					-24			5	23	-14
44	3	60	43			44	-20							11	-02	-06
48	3	54	48			51								18	9	-06
54	3	68	45			45	-26						21	14	18	-11
61	3	39	26			40						-21		8	-10	-08
74	3	67	50			59								15	11	-16
101	3	66	28			33								7	-11	2
129	3	47	42			51								7	2	0
6	4	28	47				51							-15	6	12
11	4	42	43				54							-02	-06	14
63	4	61	43			-21	51							-18	1	20
73	4	36	42				43				28			-22	-03	6
81	4	54	30				42							4	-22	2
85	4	33	52				58							-15	6	10
97	4	27	37				44							-19	12	12
109	4	61	44				55							-05	-16	5
120	4	59	27				39							0	-13	2
124	4	47	52				65							5	2	14
4	-4	42	39			32	-43							13	5	-06
8	-4	31	31			23	-36							13	0	-06

Item	Skala	pi	rit	1	2	3	4	5	6	7	8	9	10	rG.AS	rA.GS	rS.GA
27	5	76	29					37					23	-01	8	1
86	5	44	34					38	30					-02	7	14
105	5	35	46					50	32				22	-02	14	9
113	5	49	36					40	31					-08	8	5
115	5	53	52					57	22					-09	7	6
135	5	54	50					55						-13	-01	9
30	-5	55	34					-48						7	1	0
52	-5	50	32					-42	25	-30				18	-10	1
60	-5	22	28					-34						14	-01	-05
93	-5	37	42					-51						11	-06	0
102	-5	29	41					-53						9	-06	-03
108	-5	45	39			24		-51						12	-08	-01
12	6	27	38						49					10	14	8
15	6	42	35					22	29				29	5	9	6
20	6	28	38						45				29	13	23	11
24	6	40	40						37				24	5	8	10
37	6	31	37						38				26	13	6	12
40	6	23	31		-23				23				29	14	8	9
50	6	24	41						54					20	12	9
56	6	61	36					35	38					5	12	13
90	6	32	43						47				20	9	20	12
98	6	37	40		-30				38					10	4	14
132	6	50	26			25			45					4	0	3
137	6	24	43						46				21	17	18	10
28	7	50	57					34		55				-15	7	7
35	7	42	60							62	21			-11	5	7
46	7	57	46							62				5	21	1
64	7	47	41							55				0	15	-02
79	7	52	48							43	43			-16	-08	5
96	7	49	39							34	25			-06	-02	3
103	7	37	56							58				-10	12	10
116	7	55	45							50	25	25		-04	-16	6
118	7	55	50							65				1	18	-08
122	7	26	37	-24						40				-09	1	4
125	7	45	56							67				2	2	3
130	7	52	66							73				-05	12	0
16	8	27	44								52			-17	-11	10
33	8	37	38								45			-27	-07	5
47	8	39	41								48			-19	1	10
51	8	27	47								58			-17	-24	12
57	8	47	45								55			-14	-17	7
67	8	23	38								45			-19	-17	6
72	8	33	44								52			-11	-10	10
75	8	41	44								54			-14	-30	6
92	8	36	35								43			-05	-18	2
99	8	24	35								46			-03	-10	5
123	8	45	45								54	23		-16	-36	5
133	8	26	53								61			-12	-18	8
10	9	39	33									41		-09	-11	-03
18	9	50	52									60		-04	-36	1
31	9	42	48								20	53		-10	-26	-04
38	9	36	45							-23		52		4	-33	7
62	9	63	49									60		-17	-20	-06
65	9	53	48									61		-08	-30	8
68	9	72	38									55		-02	-07	2
70	9	47	39									47		-08	-15	10
84	9	42	45									55		-06	-19	6
89	9	46	44									56		3	-19	0
117	9	79	33									47		-13	-11	0
127	9	25	35								27	38		-09	-19	2

Item	Skala	pi	rit	1	2	3	4	5	6	7	8	9	10	rG.AS	rA.GS	rS.GA
7	10	54	30										43	3	17	-03
9	10	52	46					23					54	6	19	1
66	10	77	29										41	2	12	-02
69	10	66	45										59	9	23	1
71	10	52	44										58	12	18	-05
77	10	54	36					33					43	-02	15	3
78	10	40	36					24	29				35	3	13	17
83	10	37	34										38	12	11	5
95	10	25	35										40	5	12	6
104	10	54	31				23		23				31	12	15	12
107	10	52	44										55	15	16	3
114	10	32	39										48	-02	15	2
2	E	51	34	44										-03	48	-02
4	E	42	46	56										13	5	-06
17	E	59	44	55										2	1	-01
20	E	72	34	43										13	23	11
25	E	22	37	46										7	12	-06
32	E	61	44	54										20	4	-14
39	E	44	37	47										11	4	-08
53	E	66	57	68										11	7	-01
54	E	32	40	50										14	18	-11
59	E	41	46	56										-08	3	-05
76	E	45	53	64										-06	13	0
91	E	60	39	48										17	28	5
81	-E	54	36	-46										4	-22	2
109	-E	61	48	-58										-05	-16	5
19	N	61	45		56									-09	-02	3
28	N	50	52		62									-15	7	7
42	N	76	35		46									1	10	5
45	N	76	43	20	55									-06	16	10
49	N	61	41		51									-16	-07	4
55	N	57	32		39									-06	-27	13
79	N	48	51		60									-16	-08	5
82	N	53	47		58									-19	-02	6
106	N	37	50		61									-21	7	3
110	N	49	33		44									-03	15	9
112	N	52	46		55									-10	-09	4
115	N	47	37		48									-09	7	6
126	N	67	41		53									-01	17	6
130	N	48	32		42									-05	12	0

# Erhebung 1999 (Normstichprobe, Ost-/West-Deutschland)

Faktorenanalyse von 120 Items der 10 Standardskalen des FPI-R (max. N=3.740). Item-Skalen-Zuordnung (Schlüssel);  
Faktorenloadungen  $\alpha \geq 0,20$  der 10-Faktoren-Hauptkomponenten-Lösung.

Varimax Rotierte Komponentenmatrix												
	Item	Skala	KÖRP	AGG	LEI	BEAN	GEH	GES	ERR	OFF	SOZ	LEB
1	57	8	0,61									
2	133	8	0,59									0,20
3	123	8	0,56					0,23				
4	51	8	0,56									
5	75	8	0,54									
6	92	8	0,48									
7	72	8	0,47									
8	67	8	0,47									
9	99	8	0,47									
10	16	8	0,44									
11	33	8	0,43									
12	47	8	0,43			0,22						0,20
1	90	6		0,61								
2	50	6		0,58								
3	137	6		0,56								
4	56	6		0,46					-0,28			
5	37	6		0,45								
6	98	6		0,43							-0,20	
7	12	6		0,42								
8	15	6		0,41						0,23		
9	24	6		0,40						0,19		0,23
10	132	6		0,34	0,34							
11	40	6		0,32	0,21						-0,19	
12	20	6		0,32						0,30		
1	39	3			0,60		-0,21					
2	21	3			0,57							
3	74	3			0,55		-0,25					
4	129	3			0,55							
5	48	3			0,52					0,25		
6	44	3			0,51		-0,25					
7	54	3			0,50		-0,24			0,22		
8	5	3			0,47							
9	13	3			0,47		-0,32					
10	61	3			0,46							
11	101	3			0,31							-0,19
12	43	3	-0,22		0,30	0,23				0,26		
1	130	7				0,70			-0,20			
2	125	7				0,67						
3	118	7			0,24	0,63						
4	46	7				0,62						
5	35	7	0,22			0,61						
6	122	7		0,20		0,49						0,21
7	116	7	0,25			0,48		0,25				
8	103	7	0,25			0,47						
9	79	7	0,41			0,47						
10	64	7			0,20	0,46						
11	96	7	0,25			0,43			-0,21			
12	28	7	0,29			0,40			-0,31			
1	124	4					0,63					
2	85	4					0,60			0,22		
3	109	4					0,52					
4	11	4					0,51					
5	63	4			-0,27		0,50					
6	97	4					0,50			0,21		
7	81	4					0,47					
8	73	4	0,23				0,46					
9	6	4					0,46					
10	120	4					0,39					
11	4	4			0,37		-0,38					
12	8	4			0,20		-0,36					

	Item	Skala	KÖRP	AGG	LEI	BEAN	GEH	GES	ERR	OFF	SOZ	LEB
1	68	9						0,60				
2	65	9	0,20					0,59				
3	89	9						0,58				
4	117	9						0,57				
5	62	9						0,57				
6	84	9						0,55				
7	70	9	0,19					0,47				
8	38	9						0,45	0,20	-0,31		
9	18	9						0,43		-0,31		
10	31	9						0,42			0,22	
11	127	9	0,27					0,40				
12	10	9						0,26				
1	102	5		-0,22					0,55			
2	108	5							0,53			
3	30	5							0,53			
4	93	5							0,51			
5	52	5				-0,20			0,50			
6	135	5		0,31			0,20		-0,47			
7	60	5							0,46			-0,20
8	115	5	0,19	0,37					-0,45			
	88	1							0,35			-0,21
9	86	5		0,32	0,20				-0,32			
10	27	5		0,29					-0,30	0,24		
11	105	5		0,50					-0,32			
12	113	5	0,21			0,30			-0,29			
1	69	10		0,21						0,52		
2	107	10		0,20						0,52		
3	71	10		0,20						0,50		
4	9	10		0,25						0,47		
5	83	10		0,19						0,39		
6	104	10		0,25						0,38		
7	66	10								0,38		
8	77	10		0,24					-0,29	0,37		
9	95	10		0,22						0,36		
10	7	10								0,35		
11	78	10		0,45					-0,20	0,22		
12	114	10	0,21	0,31						0,26		
1	121	2									-0,60	
2	80	2									0,59	
3	26	2			0,20						0,55	
4	36	2									0,52	
5	41	2						0,19			0,47	
6	111	2						0,20			0,46	-0,21
7	134	2									0,45	
8	22	2		0,23							-0,44	
9	34	2		0,25							-0,43	
10	14	2									0,42	
11	136	2		-0,26				0,23			0,38	
12	87	2								0,23	-0,36	
1	128	1										-0,67
2	119	1				0,22						0,62
3	138	1	-0,20									-0,52
4	58	1										0,51
5	112	1	0,28									0,46
6	131	1										-0,43
7	23	1							0,36	-0,22		-0,40
8	94	1	0,28	0,19		0,27			-0,20			0,40
9	100	1										-0,38
10	3	1			0,30							-0,36
11	29	1										-0,36

Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

# Erhebung 2018 (Repräsentative Bevölkerungstichprobe Juni 2018, Ost-/West-Deutschland, N=1.112)

**Faktorenanalyse** von 120 Items der 10 Standardskalen des FPI-R (max. N=3.450). Item-Skalen-Zuordnung (Schlüssel);  
Faktorenladungen  $\alpha \geq 0,20$  der 10-Faktoren-Hauptkomponenten-Lösung.

Skala	Iteminhalt (verkürzt)	Komponente									
		1	2	3	4	5	6	7	8	9	10
KÖRP	V57 Mein Herz beginnt manchmal zu jagen oder unregelmäßig zu schlagen	53									
KÖRP	V123 Mein Körper reagiert deutlich auf Wetteränderung	49					25				
KÖRP	V133 Ich habe manchmal ein Gefühl erstickender Enge in der Brust	48						-28			
KÖRP	V92 Ich habe einen empfindlichen Magen	46									
KÖRP	V75 Ich habe Schwierigkeiten einzuschlafen oder durchzuschlafen	46									
KÖRP	V99 Meine Hände sind häufiger zitterig, zB beim Schreiben mit einem Stift oder Halten einer Tasse	45									
KÖRP	V47 Ich habe manchmal das Gefühl, einen Kloss im Hals zu haben	43	21					-27			
KÖRP	V51 Ich habe manchmal Hitzewallungen und mir steigt das Blut in den Kopf	43			25						
KÖRP	V33 Ich habe selbst bei warmem Wetter häufiger kalte Hände und Füße	39									
KÖRP	V67 Ich habe häufiger Verstopfung	39									
KÖRP	V16 Ich habe häufig Kopfschmerzen	37									
KÖRP	V72 Ich bemerke häufiger ein unwillkürliches Zucken, zB um meine Augen	37									
GEH	V124 Es fällt schwer, den richtigen Gesprächsstoff zu finden, wenn ich jmd. kennenlernen will		53								
GEH	V85 Ich werde ziemlich leicht verlegen		52								
GEH	V63 Es fällt mir schwer, vor einer großen Gruppe von Menschen zu sprechen oder vorzutragen		50								
GEH	V11 Ich bin ungern mit Menschen zusammen, die ich noch nicht kenne		49								
GEH	V109 Bei Geselligkeiten und öffentlichen Veranstaltungen bleibe ich lieber im Hintergrund		48								
GEH	V73 Ich bin im Grunde eher ein ängstlicher Mensch	22	48		22						
GEH	V04 Ich habe fast immer eine schlagfertige Antwort bereit (stimmt . nicht)		48							-20	-23
GEH	V06 Scheue mich, allein in Raum zu gehen, in dem and. Leute bereits zusammensitzen und sich unterhalten		41								
GEH	V81 Ich schließe nur langsam Freundschaften		41								
GEH	V97 Ich werde leicht rot		40								
GEH	V08 Ich würde mich im Restaurant beschweren, wenn schlechtes Essen serviert wird		39								
GEH	V120 Beim Reisen schaue ich lieber auf die Landschaft als mich mit Mitreisenden zu unterhalten		20								
OFF	V69 Ich bin hin und wieder ein wenig schadenfroh			57							
OFF	V107 Ab und zu erzähle ich auch mal eine Lüge			52							
OFF	V83 Meine Tischmanieren sind zu Hause schlechter als im Restaurant			51							
OFF	V104 Wenn ich irgendwo zu Gast bin, ist mein Benehmen meistens besser als zu Hause			50							
OFF	V71 Hin und wieder gebe ich ein bisschen an			48							
OFF	V09 Ich habe manchmal hässliche Bemerkungen über andere Menschen gemacht			46							
OFF	V114 Ich spreche manchmal über Dinge, von denen ich nichts verstehe			40							-20
OFF	V95 Manchmal habe ich Gedanken, über die ich mich schämen muss			40							
OFF	V77 Manchmal bin ich beleidigt, wenn es nicht nach meinem Willen geht			39	25						
OFF	V07 Manchmal bin ich zu spät zu einer Verabredung oder zur Schule gekommen			32							
OFF	V78 Ich spreche oft Drohungen aus, die ich gar nicht ernst meine			30	27					21	
OFF	V66 Manchmal schiebe ich etwas auf, was ich sofort tun sollte			21							
ERR	V52 Auch wenn es eher viel zu tun gibt, lasse ich mich nicht hetzen				56						
ERR	V135 Ich bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde			24	52						
ERR	V108 Ich lasse mich durch eine Vielzahl von kleinen Störungen nicht aus der Ruhe bringen				49						
ERR	V102 Im Allgemeinen bin ich ruhig und nicht leicht aufzuregen				46					25	
ERR	V115 Oft rege ich mich zu rasch über jemanden auf	25		37	44						
ERR	V93 Es gibt nur wenige Dinge, die mich leicht erregen oder ärgern				43						
ERR	V60 Auch wenn mich etwas aus der Fassung bringt, beruhige ich mich meistens wieder rasch				41						
ERR	V30 Wenn mir einmal etwas schief geht, regt mich das nicht weiter auf				40						
ERR	V113 Ich neige oft zu Hast und Eile, auch wenn es überhaupt nicht notwendig ist	21			37						
ERR	V105 Ich kann oft meinen Ärger und meine Wut nicht beherrschen			23	33			-20		21	
ERR	V27 Ich neige dazu, bei Auseinandersetzungen lauter zu sprechen als sonst			32	32						
ERR	V86 Ich koche vor Wut, wenn man mich zum Narren hält			27	30						21



Skala	Iteminhalt (verkürzt)	1	2	3	4	5	6	7	8	9	10
BEAN	V125 Ich denke manchmal, dass ich zu viel arbeite					67					
BEAN	V118 Ich arbeite oft unter Zeitdruck					61					
BEAN	V46 Ich hätte gern mehr Zeit für mich ohne so viele Verpflichtungen					56					
BEAN	V130 Ich habe häufig das Gefühl, im Stress zu sein				31	55		-21			
BEAN	V64 Auch an Wochenenden bin ich stark eingespannt					51					
BEAN	V35 Die täglichen Belastungen sind so groß, dass ich davon oft müde und erschöpft bin	32				43					
BEAN	V96 Nur selten kann ich richtig abschalten	25			24	38					
BEAN	V116 Ich denke manchmal, dass ich mich mehr schonen sollte	30				36					
BEAN	V103 Vor lauter Aufgaben und Zeitdruck bin ich manchmal ganz durcheinander	24			31	36					
BEAN	V122 Die Anforderungen, die an mich gestellt werden, sind oft zu hoch					34		-22			
BEAN	V79 Ich bin häufiger abgespannt, matt und erschöpft	42				30		-23			
BEAN	V28 Ich bin oft nervös, weil zu viel auf mich einströmt	30	21		30	27		-23			
GES	V62 Ich vermeide es, ungewaschenes Obst zu essen						56				
GES	V84 Weil man sich so leicht anstecken kann, wasche ich mir zu Hause gleich die Hände						53				
GES	V68 Wenn jemand in meine Richtung hustet oder niest, versuche ich mich abzuwenden						52				
GES	V65 Ich vermeide Zugluft, weil man sich zu leicht erkälten kann						47				
GES	V38 Um gesund zu bleiben, achte ich auf ein ruhiges Leben			-22	-21	-20	47				
GES	V117 Handtücher in viel benutzten Waschräumen sind mir wegen Ansteckungsgefahr unangenehm						47				
GES	V18 Ich achte aus gesundheitlichen Gründen auf regelmäßige Mahlzeiten und reichlichen Schlaf			-21			47				
GES	V89 Ich passe auf, dass ich nicht zu viel Autoabgase und Staub einatme						46				
GES	V127 Auch ohne ernste Beschwerden gehe ich regelmäßig zum Arzt, nur zur Vorsicht						41				
GES	V31 Ich habe mich über die häufigsten Krankheiten und ihre ersten Anzeichen informiert	25					37				
GES	V70 Hole sicherheitshalber ärztlichen Rat ein, wenn ich länger als zwei Tage erhöhte Temperatur habe						34				
GES	V10 Im Krankheitsfall möchte ich Befund u. Behandlung von einem zweiten Arzt überprüfen lassen						25				
LEB	V128 Alles in allem bin ich ausgesprochen zufrieden mit meinem bisherigen Leben				-21			60			
LEB	V119 Ich bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden	-23				-23		55			
LEB	V94 Oft habe ich alles gründlich satt	-35						46			
LEB	V138 Meistens blicke ich voller Zuversicht in die Zukunft	-30			-24			42	21		
LEB	V58 In meinem bisherigen Leben habe ich kaum das verwirklichen können, was in mir steckt		-28					39			
LEB	V23 Ich lebe mit mir selbst in Frieden und ohne innere Konflikte				-35			36			
LEB	V131 Ich habe eine gute Partnerbeziehung (Ehe)							32			
LEB	V112 Ich grübele viel über mein bisheriges Leben nach	-39						28			
LEB	V29 Wenn ich noch einmal geboren würde, dann würde ich nicht anders leben wollen							28			
LEB	V100 Ich bin selten in bedrückter, unglücklicher Stimmung	-26			-20			22			
LEB	V88 Ich bin immer guter Laune				-25			22			
LEB	V03 Ich habe (hatte) einen Beruf, der mich voll befriedigt							21			
SOZ	V80 Bekomme häufig schlechtes Gewissen, wenn ich sehe, wie schlecht es anderen Menschen geht								51		
SOZ	V121 Da der Staat schon für Sozialhilfe sorgt, brauche ich im Einzelnen nicht zu helfen								50		
SOZ	V22 Ich bin der Ansicht, Menschen in Entwicklungsländern sollten sich zuerst einmal selbst helfen								43		
SOZ	V26 Ich fühle mich auch über meine Familie hinaus für andere Menschen verantwortlich								42		
SOZ	V34 Ich finde, jeder Mensch soll sehen, wie er zurecht kommt								42		-24
SOZ	V36 Denke oft, dass ich Konsum einschränken müsste, um an benachteiligte Menschen abzugeben								42		
SOZ	V41 Wenn jemand weint, möchte ich ihn am liebsten umarmen und trösten								40		
SOZ	V136 Ich nehme mir viel Zeit, anderen Menschen geduldig zuzuhören, wenn sie von ihren								37		
SOZ	V87 Wenn mich ein Fremder um eine kleine Geldspende bittet, ist mir das ziemlich lästig								35		
SOZ	V111 Gebe gelegentlich Geld/Spenden für Katastrophenhilfe, Caritas, Brot für die Welt und andere gemeinnützige Organisationen								35		
SOZ	V134 Habe schon unbezahlt in Gemeinde o. soz. Einrichtungen wie Rotes Kreuz geholfen								30		
SOZ	V14 Ich würde kaum zögern, auch alte und schwerbehinderte Menschen zu pflegen								25		

Skala	Iteminhalt (verkürzt)	1	2	3	4	5	6	7	8	9	10
AGGR	V137 Es gab Leute, die mich so ärgerten, dass es zu einer tätlichen Auseinandersetzung kam									53	
AGGR	V50 Wenn ich meine Rechte nur mit körperlicher Gewalt verteidigen kann, so tue ich es									49	
AGGR	V90 Wenn ich wirklich wütend werde, bin ich in der Lage, jemandem eine runterzuhauen			26						47	
AGGR	V12 Wenn jemand meinem Freund etwas Böses tut, bin ich dabei, wenn es heimgezahlt wird									42	
AGGR	V132 Ich gehe lieber bis zum Äußersten, als dass ich feige bin		-25							35	34
AGGR	V40 Es macht mir Spaß, anderen Fehler nachzuweisen			25					-21	30	
AGGR	V24 Ich male mir manchmal aus, wie übel es denen eigentlich ergehen müsste, die mir Unrecht tun			22						27	
AGGR	V20 Sind wir in ausgelassener Runde, so überkommt mich oft eine große Lust zu groben Streichen			27						26	
AGGR	V98 Einem Menschen, der mich schlecht behandelt o. beleidigt hat, wünsche ich eine harte Strafe			29					-21	26	
AGGR	V37 Als Kind habe ich manchmal anderen die Arme umgedreht, an Haaren gezogen, ein Bein gestellt usw			32						23	
AGGR	V56 Wenn mich jemand anschreit, schreie ich zurück			33	30					22	
AGGR	V15 Ich kann mich erinnern, mal so wütend gewesen zu sein, dass ich das nächstbeste Ding nahm und es zerriss oder zerschlug			33						22	
LEI	V13 Meine Bekannten halten mich für einen energischen Menschen		-45								
LEI	V43 Es gibt für mich noch eine Menge sinnvoller Aufgaben, die ich in der Zukunft anpacken werde	-29				25			26		
LEI	V21 Ich bin leicht beim Ehrgeiz zu packen		-21								44
LEI	V44 Ich pflege schnell und sicher zu handeln		-31								41
LEI	V129 Bei meiner Arbeit bin ich meist schneller als andere		-21			22					38
LEI	V74 Ich habe Spaß an schwierigen Aufgaben, die mich herausfordern		-29		-20						36
LEI	V39 Ich habe gern mit Aufgaben zu tun, die schnelles Handeln verlangen		-36								35
LEI	V54 Bei wichtigen Dingen bin ich bereit, mit anderen energisch zu konkurrieren		-38	21							32
LEI	V05 Ich glaube, dass ich mir beim Arbeiten mehr Mühe gebe als die meisten anderen Menschen					22					31
LEI	V61 Die beruflichen Aufgaben sind mir oft wichtiger als viel Freizeit oder interessante Hobbies										31
LEI	V48 Mit anderen zu wetteifern, macht mir Spaß		-29	23						31	29
LEI	V101 Ich ziehe das Handeln dem Planeschmieden vor		-26								26

Extraktionsmethode: Hauptachsenfaktorenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung

# **Patientenstichprobe 2018 (N = 737)**

**Faktorenanalyse** von 120 Items der 10 Standardskalen des FPI-R (max. N = 737). Item-Skalen-Zuordnung (Schlüssel); Faktorenladungen  $a \geq 0,20$  der 10-Faktoren-Hauptkomponenten-Lösung.

Skala, Item-Nr., Iteminhalt (verkürzt)	Varimax-rotierte Komponenten									
	3	8	6	4	1	5	2	7	9	10
LEB	SOZ	LEI	GEH	ERR	AGG	BEAN	GES	OFF	?	
LEB V03 Habe Beruf, der mich befriedigt	32								-27	
LEB V23 Lebe mit mir in Frieden, keine inneren Konflikte	45				-35					
LEB V29 Wollte nicht anders leben, wenn ich noch einmal geboren würde	43									
LEB V58 Konnte bisher kaum verwirklichen, was in mir steckt	47					-21				22
LEB V88 Bin immer guter Laune	29				-26					-30
LEB V94 Oft habe ich alles gründlich satt	49				-25	-23	-32			
LEB V100 Bin selten bedrückt, unglücklich	51						-27			
LEB V112 Grüble viel über mein bisheriges Leben nach	52						-23			
LEB V119 Bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden	52									
LEB V128 Bin ausgesprochen zufrieden mit meinem bisherigen Leben	69									
LEB V131 Ich habe eine gute Partnerbeziehung	30									
LEB V138 Meistens blicke ich voller Zuversicht in die Zukunft	58		22		-26					
SOZ V14 Würde kaum zögern, alte und behinderte Menschen zu pflegen		43							-21	
SOZ V22 Menschen in Entwicklungsländern sollten sich erst selbst helfen		22				-35				24
SOZ V26 Fühle mich auch über Familie hinaus für andere verantwortlich		46	20							
SOZ V34 Jeder Mensch soll sehen, wie er zurecht kommt		24				-36				
SOZ V36 Denke, ich sollte meinen Konsum einschränken, um Benachteiligten		34								
SOZ V41 Wenn jemand weint, möchte ich ihn umarmen, trösten		49								
SOZ V80 Häufig schlechtes Gewissen, wenn ich sehe, wie schlecht es and. ergeht		56								
SOZ V87 Mir ist es lästig, wenn mich ein Fremder um eine Geldspende bitt		43		-23						20
SOZ V111 Spende gelegentlich für Katastrophenhilfe, Caritas usw.		33								40
SOZ V121 Staat sorgt für Sozialhilfe, da brauche ich nicht zu helfen		44								26
SOZ V134 Habe schon unbezahlt in Gemeinde oder sozialer Einrichtung geholfen		28								36
SOZ V136 Nehme mir viel Zeit, and. Menschen zuzuhören, wenn sie von Sorgen erzählen		35								
LEI V05 Gebe mir bei Arbeit mehr Mühe als andere			43				26			-25
LEI V13 Bekannte halten mich für energisch			39							
LEI V21 Bin leicht beim Ehrgeiz zu packen			55							
LEI V39 Habe gerne Aufgaben, die schnelles Handeln verlangen			51	-23						
LEI V43 Gibt noch viele sinnvolle Aufgaben, die ich anpacken werde	29		28							
LEI V44 Pflege schnell und sicher zu handeln	26		36	-22	-22					
LEI V48 Mit anderen zu wetteifern, macht mir Spaß			50						22	
LEI V54 Bin bei wichtigen Dingen bereit, mit anderen zu konkurrieren			47							
LEI V61 Berufliche Aufgaben wichtiger als Freizeit und Hobbys			27				25		-23	
LEI V74 Habe Spaß an schwierigen Aufgaben			50		-22					
LEI V101 Ziehe das Handeln dem Planeschmieden vor			25							
LEI V129 Bin bei meiner Arbeit meist schneller als andere			50							
GEH V04 Habe immer eine schlagfertige Antwort bereit			-34	36		-22				
GEH V06 Scheue mich, allein in Raum in dem andere sitzen zu gehen	-24			57						
GEH V08 Würde mich im Restaurant über schlechtes Essen beschweren				34		-21		-25		
GEH V11 Bin ungern mit Menschen zusammen, die ich nicht kenne	-26			59						
GEH V63 Fällt mir schwer, vor großer Gruppen Menschen zu sprechen			-22	53						
GEH V73 Bin im Grunde ein ängstlicher Mensch			-22	36			20			
GEH V81 Schließe nur langsam Freundschaften		-26		41						
GEH V85 Ich werde ziemlich leicht verlegen				59	21					
GEH V97 Ich werde leicht rot				52	24					
GEH V109 Bei Geselligkeiten und Veranstaltungen bleibe ich lieber im Hintergrund				55						
GEH V120 Schau bei Reisen lieber auf Landschaft, als mich mit Mitreisenden zu unterhalten				45						
GEH V124 Fällt schwer, richtigen Gesprächsstoff zu finden, wenn ich jmd. kennenlernen will			-21	55						
ERR V27 Spreche bei Auseinandersetzungen lauter als sonst					40	27				
ERR V30 Wenn etwas schief geht, regt mich das nicht auf	-25				48					
ERR V52 Auch wenn es viel zu tun gibt, lasse ich mich nicht hetzen					46	-28	26			
ERR V60 Beruhige mich rasch wieder, wenn mich etwas aus der Fassung bringt	-27				45					
ERR V86 Koche ERR Vor Wut, wenn man mich zum Narren hält					36	30				
ERR V93 Nur wenige Dinge erregen oder ärgern mich leicht	-26				53					
ERR V102 Bin im Allgemeinen ruhig und nicht leicht aufzuregen					59					
ERR V105 Kann oft meinen Ärger und meine Wut nicht beherrschen					38	46				
ERR V108 Lasse mich durch viele Störungen nicht aus der Ruhe bringen	-24		-24		48		21			
ERR V113 Neige oft zu Hast und Eile, auch wenn es nicht notwendig ist					46		37			
ERR V115 Rege mich oft zu rasch über jemanden auf					57	25	22			
ERR V135 Bin leicht aus der Ruhe zu bringen, wenn ich angegriffen werde					49					

Skala, Item-Nr., Iteminhalt (verkürzt)	3	8	6	4	1	5	2	7	9	10
AGGR V12 Wenn meinem Freund Böses getan wird, zahle ich es heim						45				
AGGR V15 Schon so wütend gewesen, Dinge zerrissen oder zerschlagen						39				
AGGR V20 In ausgelassener Runde, bekomme Lust zu groben Streichen			23			35				
AGGR V24 Male mir aus, wie übel es denen gehen sollte, die mir Unrecht t						37				
AGGR V37 Habe als Kind gerne anderen die Arme umgedreht, an Haaren gezogen usw						39				
AGGR V40 Macht mir Spaß, anderen Fehler nachzuweisen						26			29	
AGGR V50 Verteidige meine Rechte auch mit körperlicher Gewalt						52				
AGGR V56 Wenn mich jemand anschreit, schreie ich zurück					40	42				
AGGR V90 Wütend bin ich in der Lage, jemandem eine runterzuhauen						58				
AGGR V98 Wünsche dem, der mich schlecht behandelt oder beleidigt, eine h					25	37				
AGGR V132 Gehe lieber zum Äußersten, als dass ich feige bin			22			44				
AGGR V137 Gab Leute, die mich so ärgerten, dass es zu tätlicher Auseinandersetzung kam						57				
BEAN V28 Bin oft nervös, weil zu viel auf mich einströmt				34	31		40			
BEAN V35 Bin von täglichen Belastungen oft müde und erschöpft	-28						52			
BEAN V46 Hätte gern mehr Zeit ohne viele Verpflichtungen							42		31	
BEAN V64 Bin auch an Wochenenden stark eingespannt							46			
BEAN V79 Bin häufiger abgespannt, matt und erschöpft	-35				22		50			20
BEAN V96 Kann nur selten richtig abschalten	-41						44			
BEAN V103 Bin manchmal durch viele Aufgaben und Zeitdruck ganz durcheinander				20	23		44			
BEAN V116 Denke manchmal, dass ich mich mehr schonen sollte							54			
BEAN V118 Ich arbeite oft unter Zeitdruck			23				53			
BEAN V122 Anforderungen, die an mich gestellt werden, sind oft zu hoch			-30				41			
BEAN V125 Ich denke manchmal, dass ich zu viel arbeite			20				60			
BEAN V130 Habe häufig das Gefühl, im Stress zu sein					30		58			
KÖRP V16 Habe häufig Kopfschmerzen							25		-23	
KÖRP V33 Bei Wärme häufiger kalte Hände und Füße				22						
KÖRP V47 Manchmal das Gefühl, einen Kloß im Hals zu haben				36	23		25			
KÖRP V51 Habe manchmal Hitzewallungen, Blut steigt in den Kopf				25	28		22			31
KÖRP V57 Mein Herz beginnt manchmal zu jagen, unregelmäßig zu schlagen							43			29
KÖRP V67 Ich habe häufiger Verstopfung										
KÖRP V72 Bemerke häufiger unwillkürliches Zucken, z.B. um die Augen					33					
KÖRP V75 Habe Schwierigkeiten, einzuschlafen oder durchzuschlafen	-42						27			24
KÖRP V92 Habe einen empfindlichen Magen					28					
KÖRP V99 Meine Hände sind häufiger zittrig				27			24			
KÖRP V123 Mein Körper reagiert deutlich auf Wetteränderung		26				22				
KÖRP V133 Habe manchmal das Gefühl erstickender Enge in der Brust	-30			23			36			21
GES V10 Möchte im Krankheitsfall von zweitem Arzt überprüft werden								37		
GES V18 Achte aus gesundheitlichen Gründen auf regelmäßige Mahlzeiten, Schlaf	28							38		
GES V31 Habe mich über häufigste Krankheiten und Anzeichen informiert		27						37		
GES V38 Um gesund zu bleiben, achte ich auf ruhiges Leben	24							42		
GES V62 Vermeide es, ungewaschenes Obst zu essen								55		
GES V65 Vermeide Zugluft, weil man sich leicht erkältet								56		
GES V68 Versuche mich abzuwenden, wenn jemand zu mir hustet oder niest								58		
GES V70 Ärztlichen Rat, wenn ich zwei Tage und länger Fieber habe								36		-32
GES V84 Wasche mir zu Hause gleich die Hände, weil man sich leicht ansteckt								64		
GES V89 Passe auf, dass ich nicht zu viel Abgase und Staub einatme								60		
GES V117 Handtücher in viel benutzten Waschräumen sind mir unangenehm								57		
GES V127 Gehe auch ohne Beschwerden regelmäßig zum Arzt								34		-25
OFF V07 Manchmal zu spät zur Schule gekommen						23			26	26
OFF V09 Habe schon schlechte Bemerkungen über andere gemacht					31	25			21	
OFF V66 Schiebe manchmal etwas auf, was ich sofort tun sollte										28
OFF V69 Bin hin und wieder ein wenig schadenfroh			27		20			-23	28	
OFF V71 Hin und wieder gebe ich ein bisschen an			35						41	
OFF V77 Manchmal bin ich beleidigt, wenn es nicht nach meinem Kopf geht					44				25	
OFF V78 Spreche oft Drohungen aus, die ich nicht ernst meine					35	38				-22
OFF V83 Meine Tischmanieren sind zu Hause schlechter als im Restaurant									51	
OFF V95 Habe manchmal Gedanken, über die ich mich schämen muss		23		21					29	
OFF V104 Wenn ich Gast bin, ist mein Benehmen besser als zu Hause									47	
OFF V107 Erzähle ab und zu auch mal eine Lüge									37	
OFF V114 Spreche manchmal über Dinge, von denen ich nichts verstehe									41	
n Items mit höchster Ladung auf betreffendem Faktor/Skala	11	8	12	12	11	11	12	12	8	0

Zusammenfassung der FA: Die Ladungsmuster der 120 Items der Standardskalen sind über den Erhebungszeitraum 1982-2018 weitgehend stabil. Auffällig sind bei der Patientenstichprobe die komplexen Ladungen der Skala Körperliche Beschwerden. Dies ist vermutlich auf die „Heterogenität“ des KÖRP-Itempools zurückzuführen, die bei der Patientengruppe deutlicher hervortritt. Für die Skalenkonstruktion KÖRP hatte es wenig Sinn, sehr seltene Beschwerden einzubeziehen, weil die Verteilung in der Normalbevölkerung und die Faktorisierung deformiert würde.

## V. Gesundheits- und Belastungsindikatoren 1999 und 2018

**Tabelle 1: Gesundheitsindikatoren nach Erhebungsjahren**

Gesundheitsmerkmale	Kategorien	Antwort/Ausprägung	1999	2018		Diff.	3W
			N %	N %	N		
Hat man eine chronische Krankheit?	2	ja	13%	27%	3400	14%	nein
War man in Psychotherapie?	2	ja	7%	15%	3366	8%	nein
Rauchen Sie?	3	ja	31%	24%	3444	-7%	ja
		nicht mehr	27%	31%		4%	nein
Anzahl der Zigaretten	Anzahl	20 Zigaretten u. mehr	22%	12%	1110	-10%	
Schlaftabletten-Einnahme in letzten 12 Mon.	5	etwa 2-mal im Jahr	10%	9%	3291	-1%	nein
		nie	80%	81%		1%	
Beruhigungstabletten-Einnahme in letzten 12 Mon.	5	etwa 2-mal im Jahr	10%	8%	3271	-2%	nein
		nie	80%	82%		2%	
Schmerztabletten-Einnahme in letzten 12 Mon.	5	etwa 2-mal im Mon.	32%	32%	3381	0%	nein
		etwa 2-mal im Jahr	34%	37%		3%	
		nie	25%	19%		-6%	
Homöopathische Mittel-Einnahme in letzten 12 Mon.	5	etwa 2-mal im Jahr	14%	19%	3316	5%	nein
		nie	68%	60%		-8%	
Übergewichtig (subjektive Einschätzung)	3	ja	20%	21%	3439	1%	nein
		vielleicht	19%	18%		-1%	
		Nein	61%	62%		1%	
BMI - Kategorien	4	Untergewicht	3%	2%	3380	-1%	nein
		Normalgewicht	55%	48%		-7%	
		Übergewicht	34%	37%		3%	
		Adipositas	8%	13%		5%	
Body-Mass-Index	kg/m <sup>2</sup>	Mittelwert	24,7	25,6	3380		nein
		SD	3,8	4,8			
Arztbesuche (ohne Zahnarzt) während der letzten 12 Monate	Anzahl	Mittelwert	3,6	4,5	3220		ja
		SD	4,6	5,0			
		Median	2	3			
Krankenhausaufenthalte bisher	Anzahl	Mittelwert	2,1	3,4	2888		ja
		SD	2,8	3,7			
		Median	1	2			
Operationen bisher	Anzahl	Mittelwert	1,4	2,3	2941		ja
		SD	1,8	2,5			
		Median	1	2			
Kuraufenthalte bisher	Anzahl	Mittelwert	0,6	0,7	1049		nein
		SD	1,3	1,4			
		Median	0	0			

*Anmerkung:* Vergleich der beiden repräsentativen (gewichteten) Erhebungen von 1999 (N = 3360) und 2018 (N = 3450). Das N für 2018 schwankt zwischen minimal N=1049 (Kuraufenthalte und frühere Rente) und maximal N=3444. 3W = Prüfung auf Unterschiede zwischen den drei Erhebungswellen von Juni – November 2018.

Die Gesundheitsprofile der Jahre 1999 und 2018 sind in 7 der 13 erfassten Merkmale ähnlich. Auffällig ist aber 1. die Verdopplung von chronischen Krankheiten, 2. die Zunahme von Übergewicht und Fettleibigkeit um 8 Prozent, und 3. eine deutliche Zunahme von psychotherapeutischen Behandlungen, Arztbesuchen, Krankenhausaufenthalten und Operationen. Aus diesen Befunden muss gefolgert werden, dass sich der objektive Gesundheitszustand der Deutschen in fast zwei Dekaden nicht verbessert hat, wenngleich die subjektive Gesundheit von der überwiegenden Mehrheit der Befragten 1999 wie auch 2018 konstant als gut bis befriedigend bezeichnet wurde (1999 sehr gut/gut = 56%, 2018 = 53%). Der objektive Gesundheitsstatus hat sich im genannten Zeitraum auch unter Berücksichtigung des Alterseinflusses verschlechtert.

**Tabelle 2: Subjektive Gesundheit 1999 und 2018 nach Alter**

Subjektive Gesundheit (Selbsteinschätzung, 5 Stufen: von "sehr gut" bis "sehr schlecht")						
	sehr gut			schlecht/sehr schlecht		
Alter	2018	1999	Diff.	2018	1999	Diff.
16-19 J.	48%	38%	11%	5%	4%	1%
20-29 J.	26%	28%	-1%	5%	5%	1%
30-39 J.	19%	18%	1%	7%	7%	-1%
40-49 J.	13%	13%	0%	11%	9%	2%
50-59 J.	7%	5%	2%	14%	12%	1%
60-69 J.	6%	7%	-1%	15%	16%	-1%
70+ J.	4%	2%	2%	23%	25%	-2%
Gesamt	13%	14%	-1%	13%	11%	1%
N	456	520		436	407	

**Tabelle 3: Ausgewählte Gesundheitsindikatoren 1999 vs. 2018 nach Geschlecht und Alter**

	1999	2018	1999	2018	1999	2018
Männer	Arztbesuche		Krankenhausaufenthalte		Operationen	
Alter	Median	Median	Median	Median	Median	Median
16-19 J.	1	2	1	1	0	0
20-29 J.	2	2	1	1	0	1
30-39 J.	1	2	1	1	0	1
40-49 J.	2	2	1	2	1	1
50-59 J.	2	3	1	2	1	2
60-69 J.	3	4	2	3	1	2
70+ J.	4	4	3	4	2	2
Gesamt	2	3	1	2	1	2
Frauen	Arztbesuche		Krankenhausaufenthalte		Operationen	
Alter	Median	Median	Median	Median	Median	Median
16-19 J.	2	2	0	1	0	0
20-29 J.	2	3	1	1	0	1
30-39 J.	2	3	1	2	1	1
40-49 J.	2	3	2	2	1	1
50-59 J.	3	3	2	2	2	2
60-69 J.	4	3	3	3	2	2
70+ J.	4	5	3	4	2	3
Gesamt	3	3	2	2	1	2

Anmerkung: aus der Tabelle geht hervor, dass sich der Gesundheitsstatus mit zunehmendem Alter kontinuierlich verschlechtert, wie nicht anders zu erwarten.

**Tabelle 4: Durchschnittliche berufliche Belastung, Lebenszufriedenheit und subjektive Gesundheit (1999 und 2018)**

Jahr	Berufliche Belastung durch	Kat.*	N	Median	M	SD	ES
2018	Körperliche Arbeit	5	2124	2	2,5	1,4	.07
1999			2106	3	2,6	1,3	
2018	Tempo, Genauigkeit, Konzentration	5	2134	4	3,8	1,0	.00
1999			2166	4	3,8	0,9	
2018	Verantwortung für Menschen und Planung	5	2129	4	3,5	1,2	.08
1999			2109	4	3,4	1,2	
2018	Berufliche Belastung - Index (summativ)	3-15	2109	10	9,9	2,3	.08
1999			2149	10	9,7	2,4	
2018	Vorzeitige Rente, falls finanziell machbar (ja)	71%					
1999		79%					
Jahr	Lebenszufriedenheit mit	Kat.*	N	Median	M	SD	ES
2018	Beruflicher Situation	5	2675	2	2,4	1,1	.18
1999			2488	3	2,6	1,1	
2018	Familiärer Situation	5	3305	2	2,1	1,1	.18
1999			3148	2	2,3	1,1	
2018	Finanzieller Situation	5	3306	3	2,7	1,1	.09
1999			3147	3	2,8	1,1	
2018	Wohnsituation	5	1065	2	2,1	1,1	.00
1999			3147	2	2,1	1,0	
2018	Lebenszufriedenheit - Index (summativ, ohne Wohnsituation)	3-15	2665	7	7,4	2,7	.15
1999			2478	8	7,8	2,6	
2018	<b>Gesundheitszustand</b>	5	3340	2	2,5	0,9	.11
1999			3619	2	2,4	0,9	

Anmerkung: gewichtete repräsentative Stichproben für die Erhebungsjahre 1999 und 2018 (Ost-/Westdeutschland, Alter, Geschlecht)

\*Kategorien Belastung: 1 = gering – 5 = sehr stark; Zufriedenheit: 1 = sehr zufrieden – 5 = unzufrieden; Gesundheitszustand: 1 = sehr gut – 5 = schlecht

**Tabelle 5: Korrelationen der FPI-Skalen mit Merkmalen der Belastung, Zufriedenheit und Gesundheit**

Pearson Korrelation	Jahr	Polung	FPIR8 KÖRP	FPIR9 EMOT	FPIR1 LEB	FPIR3 LEI	FPIR7 BEAN	FPIR6 EXTR	FPIR4 GEH	FPIR5 ERR	FPIR9 GES	FPIR6 AGGR
Zufriedenheitsindex	2018	zufrieden	-0.26	-0.39	0.57	0.25	-0.25		-0.22	-0.23		
	1999		-0.21	-0.34	0.55	0.21	-0.24		-0.20	-0.26		
Belastungsindex	2018	belastet					0.26					
	1999					0.20						
Überstunden	2018	viele				0.22	0.24					
Gesundheitszustand	2018	gut	-0.43	-0.39	0.38	0.23		0.24				
	1999		-0.46	-0.39	0.32	0.22	0.22	0.28	0.23			
Tabletteneinnahme	2018	häufig	0.48	0.35	-0.29							
	1999		0.41	0.26	-0.23							
Psychotherapie	2018	ja	0.25	0.28	-0.22							
	1999			0.21	-0.20							
Chronische Krankheit	2018	ja	0.29	0.23	-0.23							
	1999		0.20									
Krankenhausaufenthalte	2018	mehrere	0.25									
	1999		0.22									
Operationen	2018	mehrere	0.22									
	1999		0.22									
Kuraufenthalte	2018	mehrere	0.23									
	1999		0.20								0.21	
Arztbesuche	2018	mehrere	0.29								0.22	
	1999		0.27								0.20	
Rauchen	2018	ja										0.23
	1999										0.22	0.25

Anmerkung: nur Korrelationen > 0.20; FPIR2 (SOZ) und FPIR10 (OFF) fehlen, da Korrelationen mit Kriterien < 0.20; grüne Schattierung: r > 0.30.



**Tabelle 6: FPI-R Skalen nach soziodemografischen Merkmalen und Belastungsindikatoren (Erhebung 1999 und 2018)**

Mittelwertunterschiede zwischen ausgewählten parallelisierten Gruppierungen (Kontrastgruppen)

SKALA		Beruf: Leitender Angestellter (M1) vs. Angestellter (M2) n = 329		Beruf: Leitender Angestellter (M1) vs. Facharbeiter (M2) n = 233		Berufstätig(M1) vs. Arbeitslos (M2) n = 293		Berufliche Belastung <8 >11 n = 208		Soziale Schicht obere untere n = 384		Chronische Krankheit ja nein n = 416 (1999) n = 915 (2018)		Arztbesuche im Jahr 0 > 10 n = 150 (1999) n = 104 (2018)		Psychotherapeutische Behandlung nein ja n = 219 (1999) n = 524 (2018)	
		M1	M2↑↓	M1	M2↑↓	M1	M2↑↓	M1	M2↑↓	M1	M2↑↓	M1↑↓	M2	M1	M2↑↓	M1	M2↑↓
LEB	1999					8.0	5.6↓			8.1	6.2↓					7.5	5.4↓
	2018															8.1	6.1↓
SOZ	1999			6.8	6.1↓					6.7	5.7↓	7.1↑	6.3			6.3	7.2↑
	2018											6.4↑	6.3			6.4	6.8↑
LEI	1999	8.3	6.8↓	8.02	6.7↓	7.3	6.6↓	6.5	8.4↑	7.4	6.0↓						
	2018																
GEH	1999	4.1	5.0↑	4.23	5.2↑	4.8	5.8↑	5.1	4.4↓	4.5	5.8↑						
	2018																
ERR	1999					5.1	5.8↑			4.9	5.6↑	5.7↑	5.1	5.1	6.1↑	4.9	6.7↑
	2018											5.4↑	5.2	4.0	5.8↑	4.6	6.3↑
AGGR	1999									4.1	5.0↑						
	2018																
BEAN	1999					6.4	4.9↓	5.4	7.0↑			6.0↑	5.1	4.9	6.5↑	5.6	6.8↑
	2018											5.6↑	5.3	4.8	6.2↑	5.1	6.5↑
KÖRP	1999					3.2	4.1↑			3.2	4.3↑	5.2↑	4.2	3.7	5.5↑	3.8	5.0↑
	2018											4.4↑	3.6	2.3	5.1↑	2.9	4.7↑
GES	1999													6.2	7.3↑		
	2018													5.3	7.6↑		
OFF	1999																
	2018																
EXTR	1999	7.8	6.4↓	7.8	6.6↓					7.4	6.2↓						
	2018																
EMOT	1999					5.8	7.1↑			5.3	6.7↑	7.1↑	5.9	5.7	7.8↑	5.8	8.6↑
	2018											6.6↑	5.9	4.2	7.4↑	4.7	7.7↑

Anmerkungen: Matching (statistische Zwillinge) von M1 und M2 (1999) nach Geschlecht, Alter, Schulabschluss, Einkommen und Ost-/Westdeutschland; M1 und M2 (2018) nach Geschlecht, Alter, Schulabschluss, Schicht. Bei den Gruppierungen (1999) nach Berufsgruppen und Berufstätigkeit wurden die Merkmale: Geschlecht, Altersdekade und Schulabschluss kontrolliert; bei beruflicher Belastung und Arztbesuchen zusätzlich die Einkommensgruppe; bei sozialer Schicht nur Geschlecht und Altersdekade; bei chronischer Krankheit und bei psychotherapeutischer Behandlung alle fünf Merkmale. Die Tabelle enthält nur Mittelwerte mit sehr signifikanten Mittelwertdifferenzen.



## V. VERGLEICH BEVÖLKERUNG 2018 UND PATIENTEN 2018

Aus dem Hogrefe Testsystem (HTS) liegen verschiedene Stichproben vor, die den FPI-Autoren freundlicherweise vom Hogrefe Verlag für Sekundäranalysen zur Verfügung gestellt wurden. Der HTS-Datensatz umfasst 2 985 Probanden aus den Jahren 2015–2018. Da es sich um einen anonymisierten Datensatz handelt, der außer den FPI-R-Items nur Angaben zu Alter und Geschlecht enthält, ist eine individuelle Zuordnung der Probanden zu bestimmten Populationen nicht möglich. Doch die HTS-Kundendaten, bezogen auf diesen Datensatz, weisen aus, dass 78 Prozent der Kunden, die das HTS nutzten, Kliniken (63 %) und therapeutische Praxen (15 %) waren. Deswegen kann für jenen Datensatz die Bezeichnung „Patientenstichprobe“ verwendet werden. Das Durchschnittsalter der weiblichen Probanden (52 %) beträgt 41 Jahre und das der männlichen Probanden (48 %) 37 Jahre (gesamter Altersbereich: 16–74 Jahre).

Für einen dezidierten Vergleich von Patienten mit der („normalgesunden“) Bevölkerung wurden zunächst die vorhandenen Stichproben Patienten und Bevölkerung per 'case-control-matching' in Hinsicht auf Geschlecht und Alter parallelisiert. Der Datensatz Patienten 2018 umfasst N = 737 männliche und weibliche Probanden (M: 49%, F: 51%) mit Durchschnittsalter 37 Jahre (Altersbereich 16-74). Nach Ausschluss von Probanden mit geringen Fallzahlen (Pbn im Alter von 16 und über 65) standen N = 698 Patienten für Analysen zur Verfügung (49% Männer und 51% Frauen; Altersbereich 17-65; Durchschnittsalter 37 Jahre). Für diese Stichprobe von Patienten wurde sodann eine für das Geschlecht und Alter angepasste Kontrollgruppe aus der Bevölkerungsstichprobe (Erhebung 2018) per matched pairs gesucht, sodass schließlich N = 698 Patienten und N = 698 „normalgesunde“ Probanden für gezielte Vergleiche zur Verfügung standen.

Die folgenden Tabellen und Abbildungen umfassen Profil-, Diskriminanz-, Cluster- und IRT-Analysen, welche insgesamt belegen, dass das FPI-R gut zwischen normalgesunden Probanden und Patienten differenziert, wobei zu berücksichtigen ist, dass die sogenannte Patientenstichprobe - wie oben erläutert - keine reine Patientenstichprobe und bei dieser Stichprobe keine Informationen zu ICD-Diagnosen vorhanden sind.

Im Vordergrund der explorativen Analysen standen 18 Items aus drei FPI-R-Skalen, welche die Normalbevölkerung und Patienten am besten trennen (siehe Itemliste und Zustimmungshäufigkeiten in Tabelle 3): (1) Lebenszufriedenheit [7 Items], (2) Emotionalität [7 Items] und (4) Körperliche Beschwerden [4 Items]). Aus den 18 Items wurde ein summativer Index - emotionale-körperliche Belastung 1 - gebildet mit Umpolung der LEB Items (= Lebensunzufriedenheit). Daneben wurde auch ein weiterer Index – emotionale-körperliche Belastung 2 – errechnet, der alle 37 Items der Skalen LEB-, KÖRP und EMOT umfasst.

**Tabelle 1: Vergleich der FPI-Skalenwerte von Bevölkerung 2018 und Patienten 2018**

	Bevölkerung 2018 N=698)			Patienten 2018 (N=698)			
Geschlecht	M + F			M + F			
Skala	M	Mdn	s	M	Mdn	s	ES*
1 LEB	7,4	8	3,0	4,1	4	3,0	-1,10
12 EMOT	5,7	5	3,6	9,0	10	3,6	0,92
8 KÖRP	2,9	2	2,6	5,4	5	3,0	0,89
4 GEH	4,7	4	3,0	7,0	7	3,2	0,74
11 EXTR	7,4	8	3,5	5,4	5	3,5	-0,57
5 ERR	5,4	5	3,1	6,8	7	3,3	0,44
7 BEAN	6,0	6	3,3	7,3	8	3,2	0,40
2 SOZ	6,2	6	2,9	7,2	7	2,5	0,37
4 LEI	7,1	7	3,0	6,0	6	2,9	-0,37
9 GES	5,5	5	2,8	4,6	4	2,8	-0,32
10 OFF	7,2	7	2,4	6,6	7	2,4	-0,25
6 AGGR	4,6	4	3,0	4,2	4	2,7	-0,14
Belastungsindex 1*	6,1	6	4,3	11,8	13	4,4	1,31

Alter und Geschlecht parallelisiert. Effektstärke\*  $|d| = 0,5$  mittlerer Effekt,  $|d| = 0,8$  großer Effekt.

\*Der summative Index „emotionale-körperliche“ Belastung 1 basiert auf 18 Items der Skalen LEB-, EMOT, KÖRP (Spannweite der Werte: 0 – 18).

**Tabelle 2: Vergleich FPI-Skalenwerte von Bevölkerung und Patienten nach Geschlecht (plus Index Belastung)**

	Bevölkerung 2018 (N=698)						Patienten 2018 (N=698)							Bev.		Pat.	
	Männer (A)			Frauen (B)			Männer (A)			Frauen (B)				M	F	M	F
Skala	M	Mdn	s	M	Mdn	s	M	Mdn	s	M	Mdn	s	t-Test	A	B	A	B
LEB	7,4	8	3,1	7,4	8	2,9	4,2	4	3,0	3,9	3	3,0	LEB				
SOZ	5,4	6	2,9	6,9	7	2,7	6,7	7	2,5	7,7	8	2,4	SOZ		A		A
LEI	7,3	8	3,1	7,0	7	2,8	6,3	7	3,1	5,8	6	2,6	LEI			B	
GEH	4,3	4	3,0	5,0	5	3,1	6,4	7	3,2	7,5	8	3,2	GEH		A		A
ERR	5,2	5	3,1	5,6	5	3,1	6,1	6	3,1	7,4	8	3,3	ERR				A
AGGR	5,2	5	3,2	4,0	4	2,7	4,5	4	2,8	3,9	4	2,6	AGGR	B		B	
BEAN	5,8	6	3,3	6,1	6	3,3	6,9	7	3,1	7,7	8	3,2	BEAN				A
KÖRP	2,2	1	2,3	3,6	3	2,7	4,6	4	2,9	6,2	6	2,9	KÖRP		A		A
GES	4,8	5	2,8	6,2	6	2,7	4,6	4	2,8	4,7	4	2,8	GES		A		
OFF	7,5	8	2,5	6,8	7	2,4	6,8	7	2,3	6,5	6	2,4	OFF	B			
EXTR	7,7	8	3,6	7,2	7	3,5	5,8	5	3,5	5,1	5	3,4	EXTR			B	
EMOT	5,3	5	3,7	6,0	6	3,5	8,5	9	3,5	9,6	11	3,7	EMOT				A
Belastung 1	5,7	5	4,4	6,5	6	4,3	11,2	12	4,3	12,4	14	4,4	Belastung 1				A
Alter	36,6	35	13,3	37,6	38	13,9	36,6	35	13,3	37,6	38	13,9	Alter				

Spaltenmittel (A) (B): Die Ergebnisse beruhen auf zweiseitigen Tests, wobei gleiche Varianzen angenommen werden. Für jedes signifikante Paar wird die Erläuterung der kleineren Kategorie in der Kategorie mit dem größeren Mittelwert angezeigt. Signifikanzniveau für Großbuchstaben (A, B): ,01

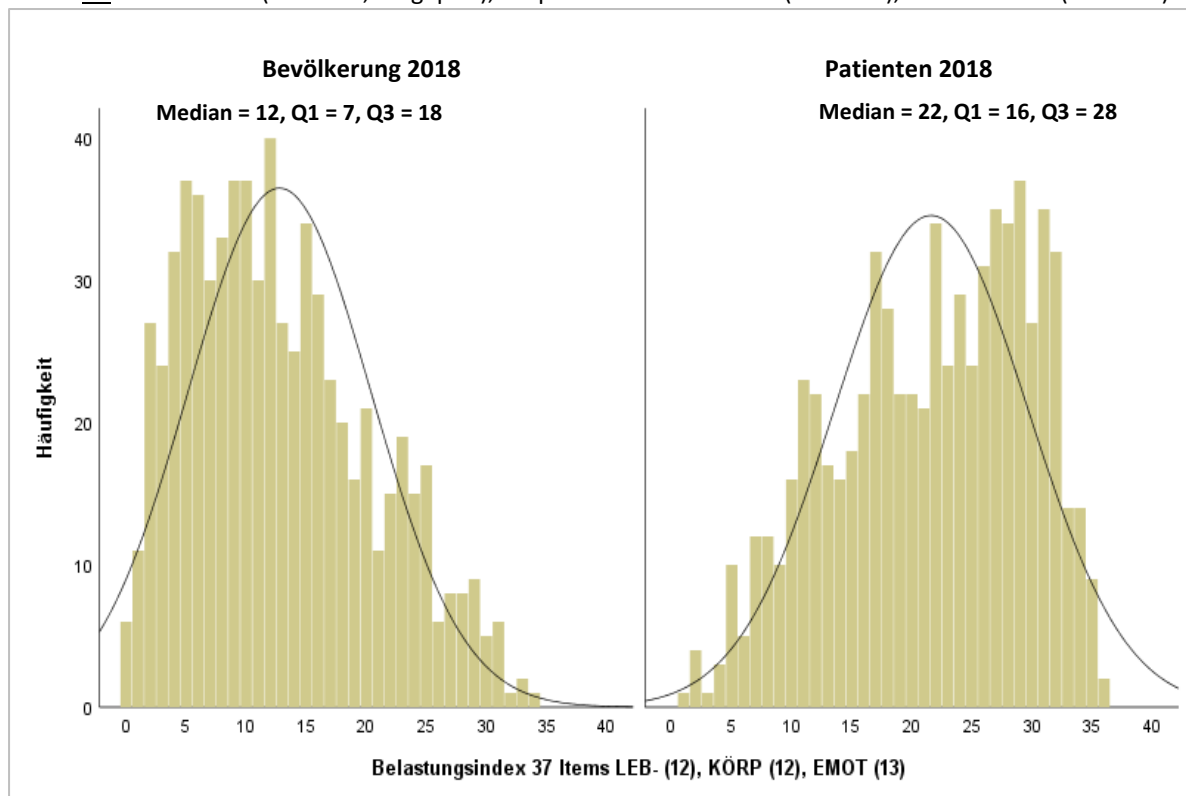
**Tabelle 3: Index emotionale-körperliche Belastung 1 (18 Items der Skalen LEB-, KÖRP, EMOT)**

	Item #	FPI-Items der Skalen LEB-, KÖRP, EMOT für summative Indexbildung	Bev. 2018	Pat. 2018	Diff.
			stimmt %		
1	N19	Ich habe manchmal ein Gefühl der Teilnahmslosigkeit und inneren Leere	31%	77%	46%
2	L128-	Alles in allem bin ich ausgesprochen <i>unzufrieden</i> mit meinem bisherigen Leben	24%	66%	42%
3	L119	Ich bin mit meinen gegenwärtigen Lebensbedingungen oft unzufrieden	32%	68%	36%
4	K75	Ich habe Schwierigkeiten einzuschlafen oder durchzuschlafen	35%	71%	36%
5	L112	Ich grüble viel über mein bisheriges Leben nach	43%	79%	36%
6	L23-	Ich lebe <i>nicht</i> mit mir selbst in Frieden und ohne innere Konflikte	45%	80%	35%
7	L100-	Ich bin <i>häufig</i> in bedrückter, unglücklicher Stimmung	32%	66%	34%
8	N106	Es gibt Zeiten, in denen ich ganz traurig, niedergedrückt bin	49%	83%	34%
9	L138-	<i>Selten</i> blicke ich voller Zuversicht in die Zukunft	23%	56%	33%
10	K57	Mein Herz beginnt manchmal zu jagen oder unregelmäßig zu schlagen	25%	55%	30%
11	N42	Meine Familie und Bekannte können mich kaum richtig verstehen	22%	49%	27%
12	N28	Ich bin oft nervös, weil zu viel auf mich einströmt	45%	72%	26%
13	K123	Mein Körper reagiert deutlich auf Wetteränderung	29%	54%	26%
14	K133	Ich habe manchmal ein Gefühl erstickender Enge in der Brust	16%	42%	26%
15	L29-	Wenn ich noch einmal geboren würde, dann würde ich anders leben wollen	48%	73%	26%
16	N126	Meine Laune wechselt ziemlich oft	32%	58%	26%
17	N79	Ich bin häufiger abgespannt, matt und erschöpft	50%	75%	25%
18	N55	Ich mache mir oft Sorgen um meine Gesundheit	34%	59%	25%
		Mittelwert	33%	61%	29%
		Interne Konsistenz (Cronbach $\alpha$ )	.84	.85	

Anmerkung: Items der Skala *Lebenszufriedenheit* umgepolt in *Lebensunzufriedenheit*. 18 aus 37 Items der Skalen LEB, KÖRP, EMOT ausgewählt, die mehr als 25% Differenz in der Zustimmungshäufigkeit zwischen Bevölkerung und Patienten aufweisen.

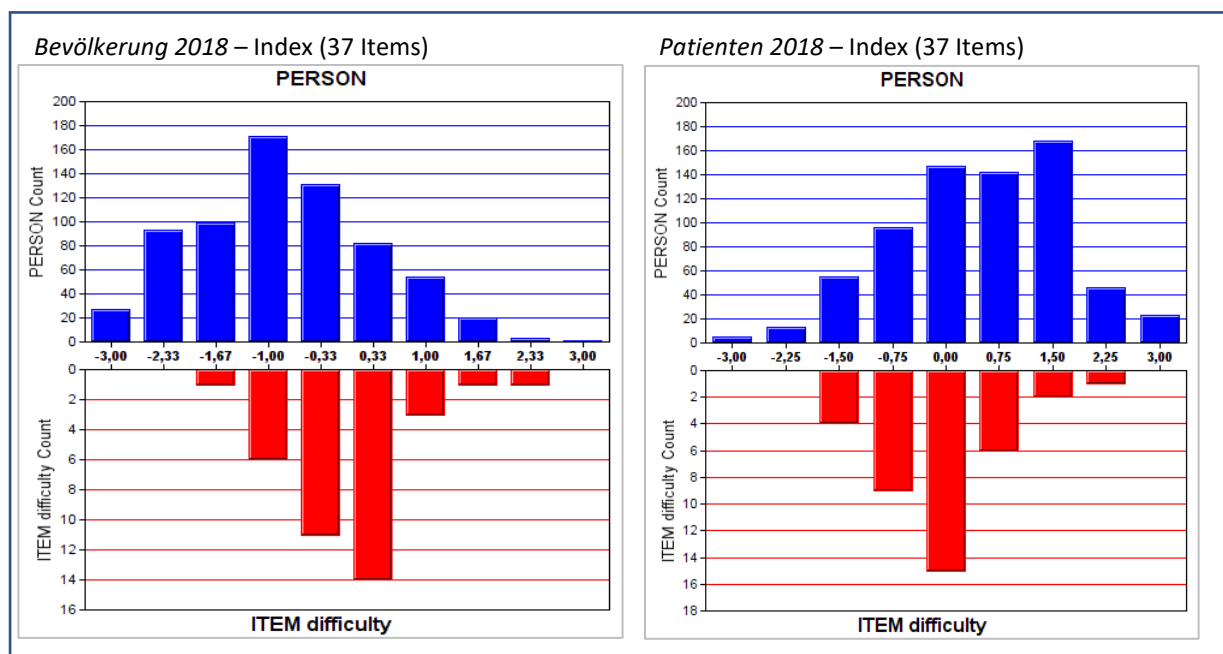
### Abbildung: Belastungsindex 2 (37 FPI-R Items) – Bevölkerung vs. Patienten 2018

Lebensunzufriedenheit (12 Items, umgepolt), Körperliche Beschwerden (12 Items), Emotionalität (13 Items)



Anmerkung: Skala Emotionalität umfasst 14 Items; ein Item aus Skala LEB doppelt verrechnet, deswegen n=14-1.

### Abbildung: Belastungsindex 2 – Personen und Itemschwierigkeiten auf der gemeinsamen Logit-Skala

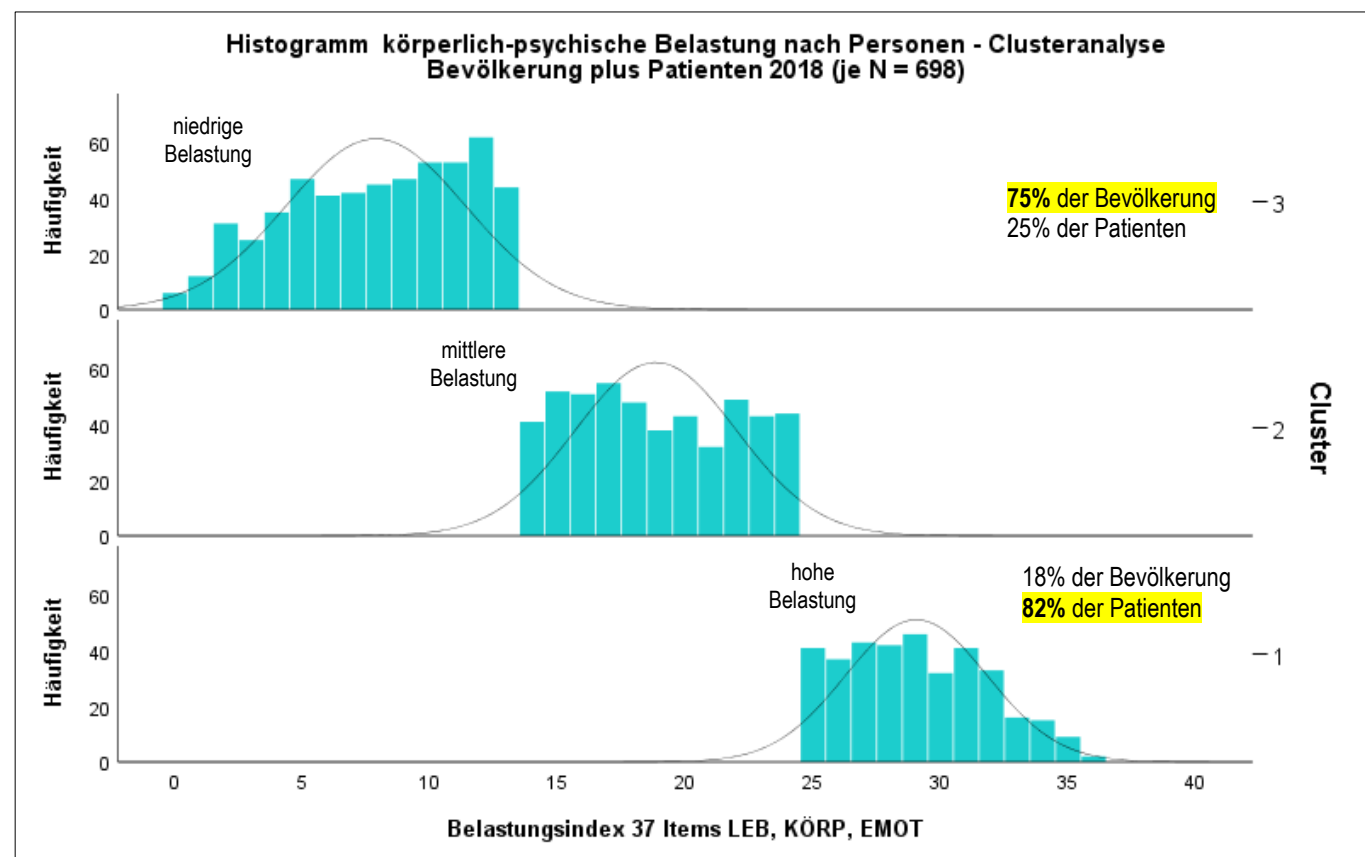


**Tabelle 4: Personen-Clusteranalyse( Two-Step, SPSS) von Bevölkerung und Patienten 2018 (je N = 698 parallelisiert)**  
 Klassifikation nach Belastungsniveau (niedrig, mittel, hoch) und Gruppe (Bevölkerung, Patienten)

Two-Step Clusteranalyse: Stetige Variable: <b>Belastungsindex 2</b> keine Clustervorgabe		Belastungsindex (37 Items LEB-, KÖRP, EMOT)											
		Bevölkerung 2018			Patienten 2018			Gesamt			Gesamt		
		Anzahl	Zeilen (%)	Spalten (%)	Anzahl	Zeilen (%)	Spalten (%)	Anzahl	Zeilen (%)	Spalten (%)	M	s	Mdn
Belastungs-niveau	niedrig	407	75%	58%	136	25%	19%	543	100%	39%	7,9	3,5	8
	mittel	228	46%	33%	268	54%	38%	496	100%	36%	18,8	3,2	19
	hoch	63	18%	9%	294	82%	42%	357	100%	26%	29,0	2,8	29
	Gesamt	698	50%	100%	698	50%	100%	1396	100%	100%	17,2	9,0	17

Anmerkung: Distanzmaß: Euklidisch, 3 Cluster automatisch ermittelt, Clusterkriterium BIC.

**Abbildung: Belastungsindex 2 – 37 FPI-R Items; 3 Belastungsniveaus (Resultat der Two-Step Personen-Clusteranalyse)**



**Tabelle 5: Vergleich von emotional-körperlich belasteten und unbelasteten Personen der Bevölkerung (N=698) nach soziodemografischen Merkmalen und Selbstauskunft zu Belastung und Gesundheit**

Bevölkerung 2018: Zwei Cluster (belastete und weniger belastete Personen) - 4 FPI-R-Skalenwerte: Lebensunzufriedenheit, Gehemmtheit, Körperliche Beschwerden, Emotionalität				
Sozio-demografische Merkmale		mehr belastet	weniger belastet	Diff. >10%
Haushaltseinkommen in Quartilen	<= 1.999 €	38%	21%	17%
Soziale Schicht - Interviewer-Eindruck	Unterschicht	18%	5%	13%
Schulabschluss	Haupt-/Volksschule	23%	12%	12%
Familienstand	verheiratet	30%	41%	-12%
Berufstätigkeit	Ja, berufstätig	62%	74%	-12%
	arbeitslos	10%	1%	9%
Personen im HH mit Einkommen	2 Personen	46%	56%	-10%
Selbstauskünfte (Fragebogen)		mehr belastet	weniger belastet	Diff. >10%
War man in Psychotherapie?	Ja	33%*	11%*	22%
Hat man eine chronische Krankheit?	Ja	28%	16%	12%
Eigener Gesundheitszustand während der letzten Zeit	(sehr) gut	35%	71%	-36%
Grad der Zufriedenheit mit der familiären Situation	(sehr) zufrieden	45%	73%	-28%
Grad der Zufriedenheit mit der finanziellen Situation	(sehr) zufrieden	23%	45%	-22%
Grad der beruflichen Belastung durch körperliche Arbeit	(sehr) gering	31%	44%	-13%
N = 698		145	553	

\* Spaltenprozente pro Merkmal; z.B. Psychotherapie: Ein Drittel aller Befragten mit Psychotherapie ist emotional-körperlich belastet vs. 11% bei den Unbelasteten.

**Tabellen 6.1 – 6.3: Ergebnisse von Diskriminanzanalysen Bevölkerung und Patienten 2018**

**Tabelle 6.1: 12 FPI-Skalen**

12 FPI-Skalen als unabhängige Variablen: Klassifizierungsergebnisse <sup>a</sup>					
		Gruppe	Vorhergesagte Gruppenzugehörigkeit		Gesamt
			Patienten 2018	Bevölkerung 2018	
Original	Anzahl	Patienten 2018	520	178	698
		Bevölkerung 2018	146	552	698
	%	Patienten 2018	74,5	25,5	100
		Bevölkerung 2018	20,9	79,1	100
a. 76,8% der ursprünglich gruppierten Fälle wurden korrekt klassifiziert.					

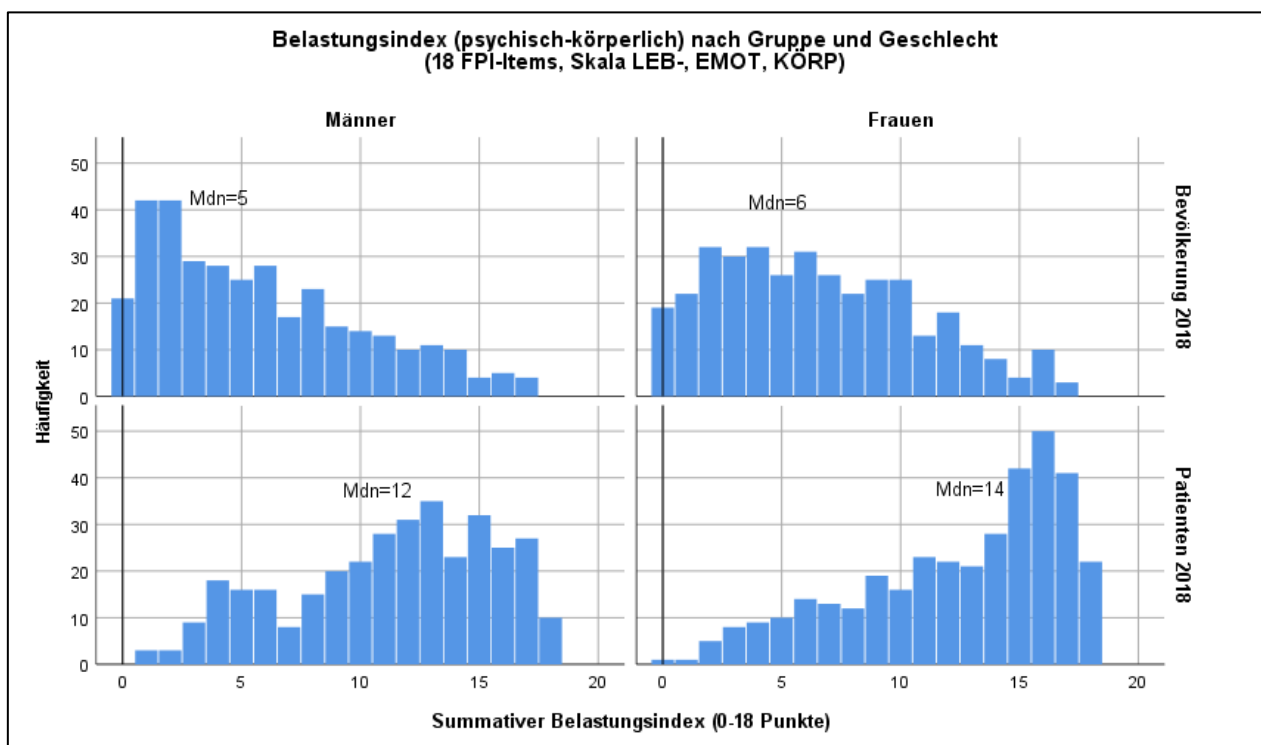
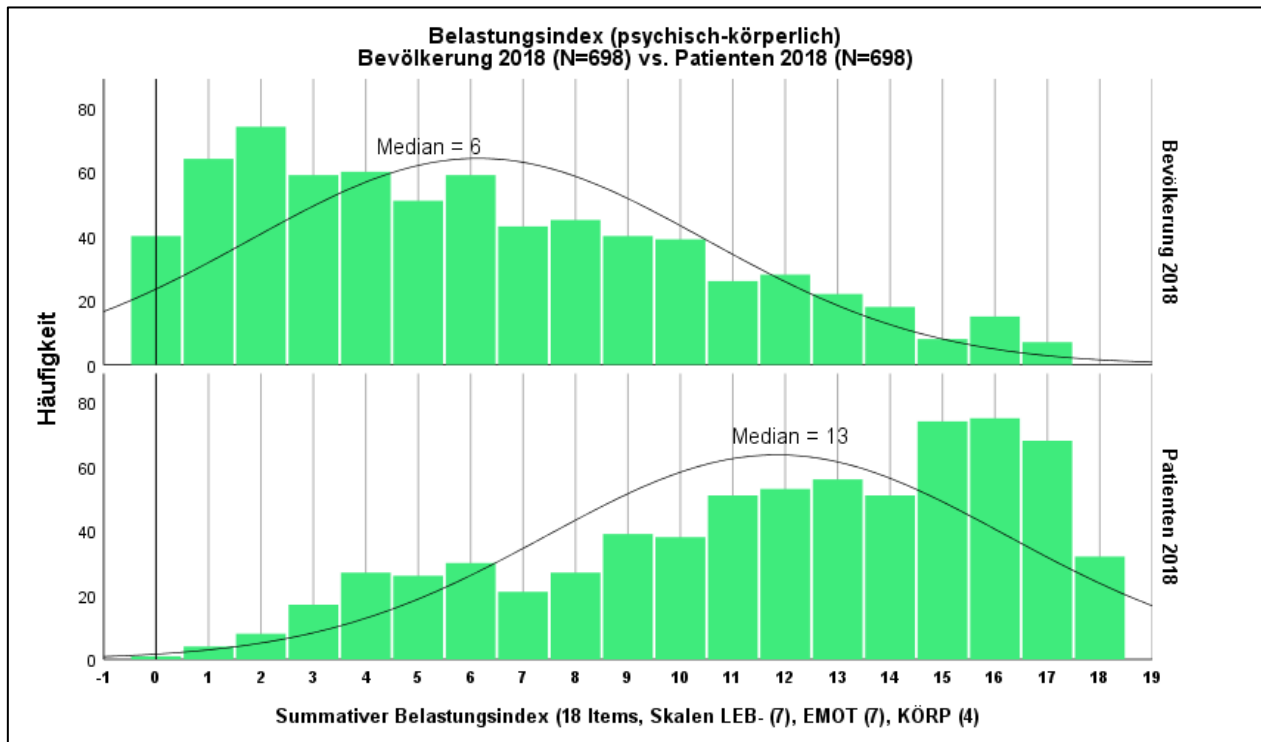
**Tabelle 6.2: 3 FPI-Skalen (37 Items)**

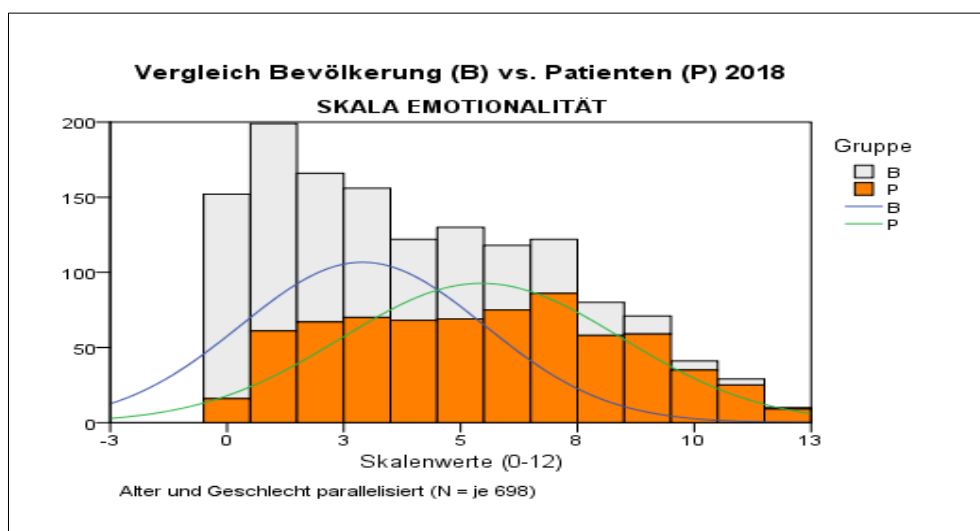
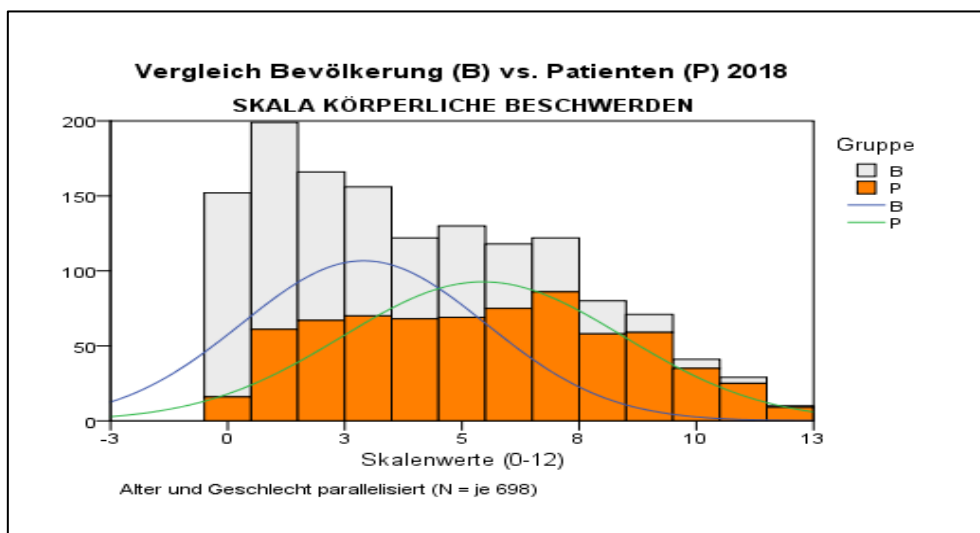
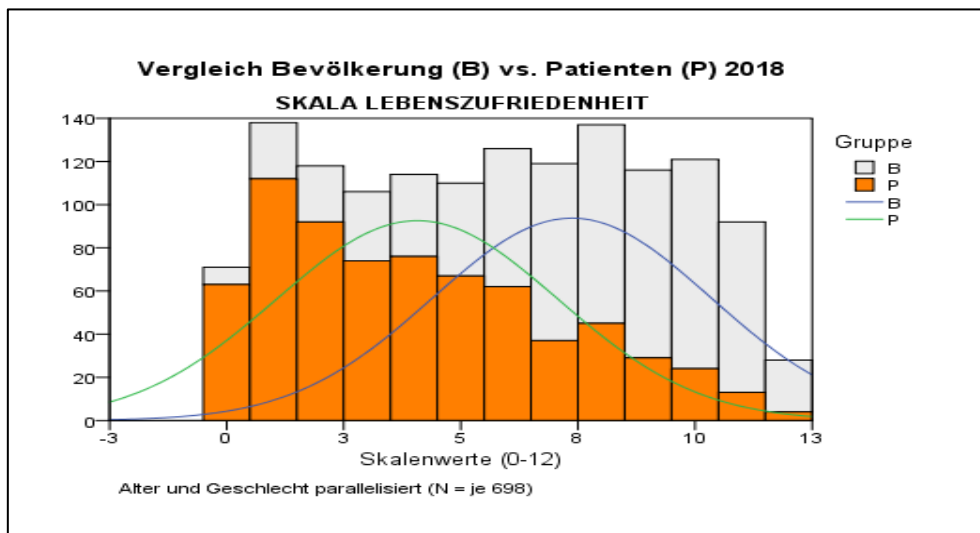
Skala LEB, KÖRP, EMOT als unabhängige Variablen: Klassifizierungsergebnisse <sup>a</sup>					
		Gruppe	Vorhergesagte Gruppenzugehörigkeit		Gesamt
			Patienten 2018	Bevölkerung 2018	
Original	Anzahl	Patienten 2018	487	211	698
		Bevölkerung 2018	187	511	698
	%	Patienten 2018	69,8	30,2	100
		Bevölkerung 2018	26,8	73,2	100
a. 71,5% der ursprünglich gruppierten Fälle wurden korrekt klassifiziert.					

**Tabelle 6.3: Belastungsindex 1 (18 Items)**

Belastungsindex 1 (18 Items) als unabhängige Variable: Klassifizierungsergebnisse <sup>a</sup>					
		Gruppe	Vorhergesagte Gruppenzugehörigkeit		Gesamt
			Patienten 2018	Bevölkerung 2018	
Original	Anzahl	Patienten 2018	537	161	698
		Bevölkerung 2018	203	495	698
	%	Patienten 2018	76,9	23,1	100
		Bevölkerung 2018	29,1	70,9	100
a. 73.9% der ursprünglich gruppierten Fälle wurden korrekt klassifiziert.					

Abbildungen: Belastungsindex 1 für Bevölkerung und Patienten 2018 (18 Items, siehe Tabelle 6.3)

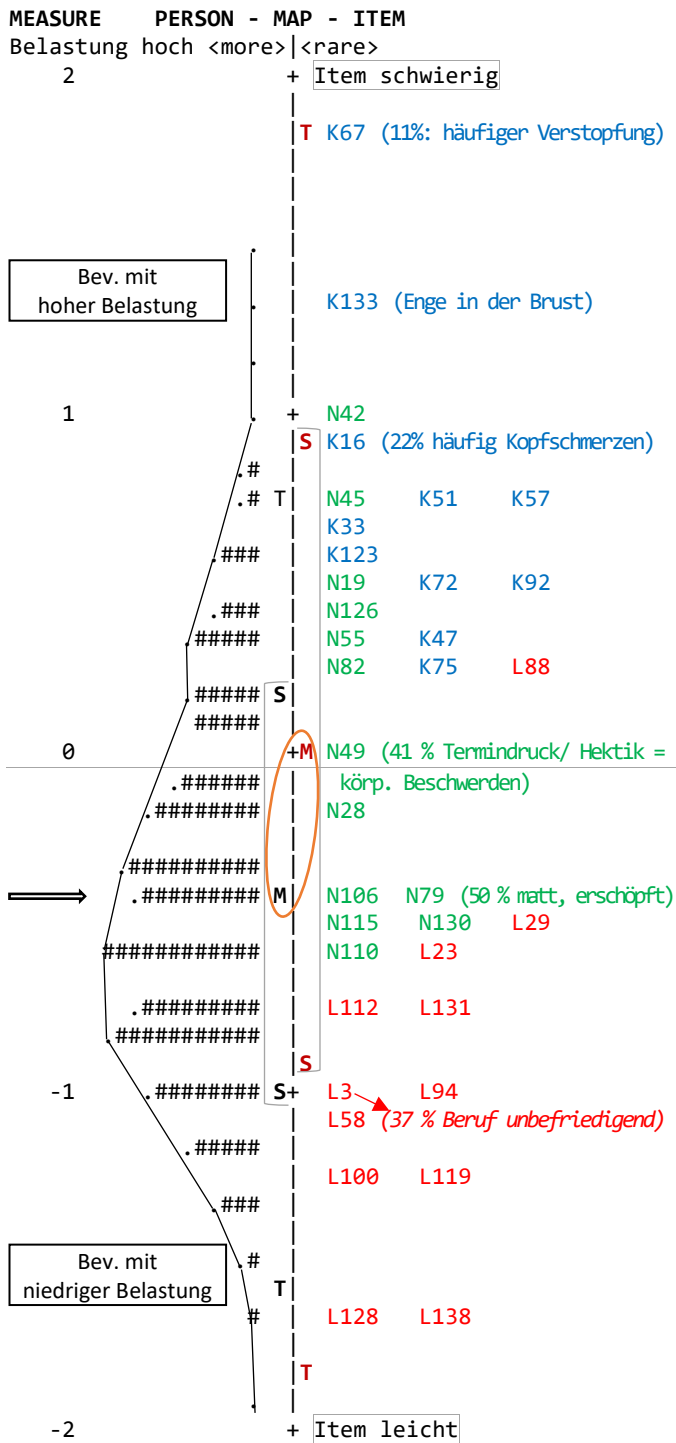






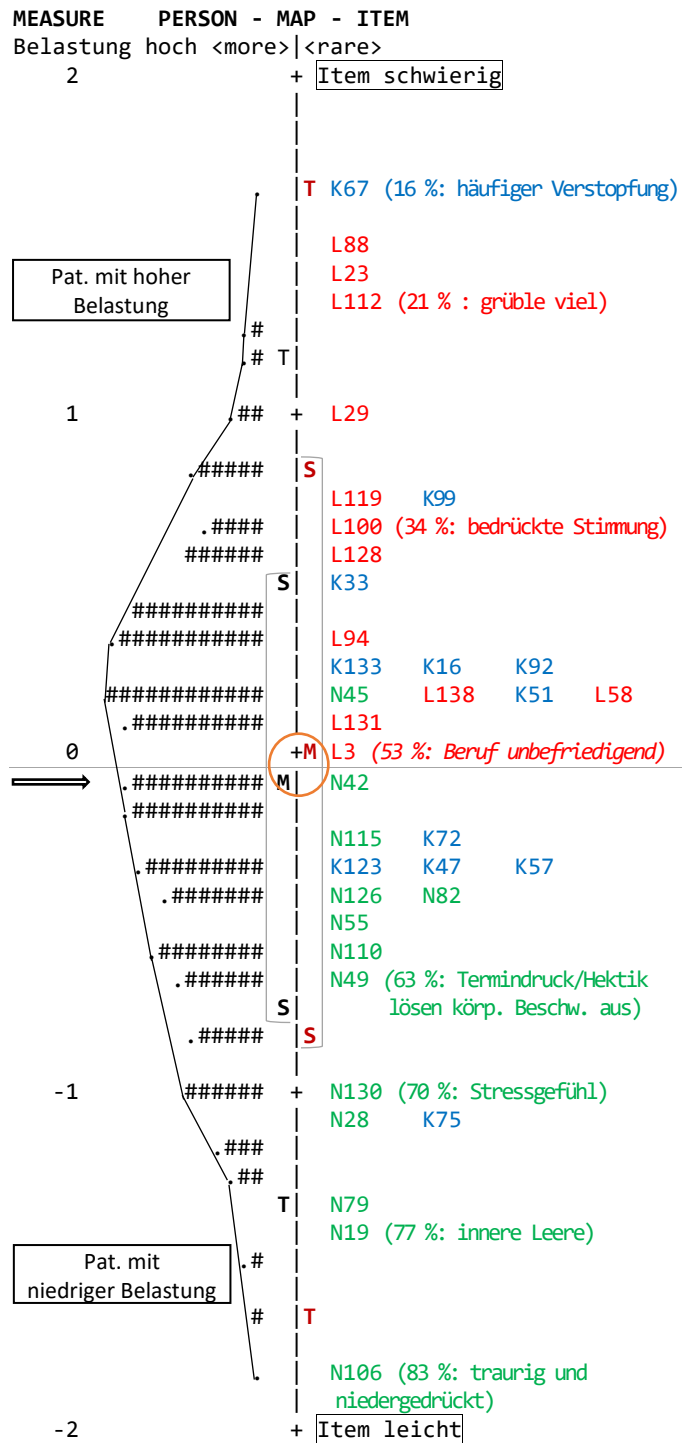
## Abbildungen: Wright-Maps (Gegenüberstellung von Probanden und Items)

**BEVÖLKERUNG 2018: 698 Personen, 37 Items**  
3 Skalen: LEB-(L), KÖRP (K), EMOT (N)



Belastung 0 <less>|<freq>  
EACH "#" IS 6 persons: EACH "." IS 1 TO 5  
M=mean, S=1 st.dev., T=2 st.dev.

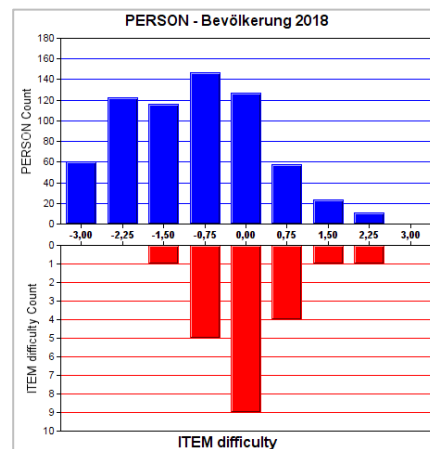
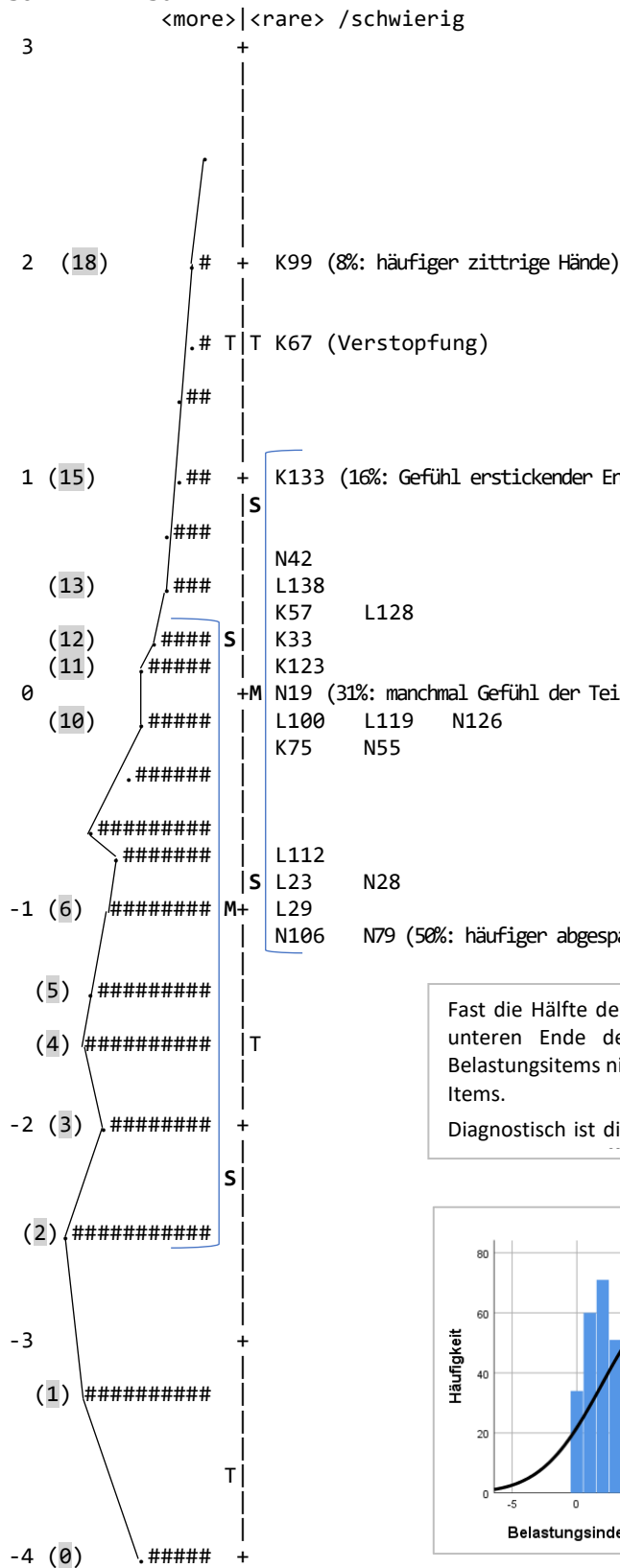
**Patienten 2018: 698 Personen, 37 Items**  
3 Skalen: LEB-(L), KÖRP (K), EMOT (N)



Belastung 0 <less>|<freq>  
EACH "#" IS 5 persons: EACH "." IS 1 TO 4  
M=mean, S=1 st.dev., T=2 st.dev.

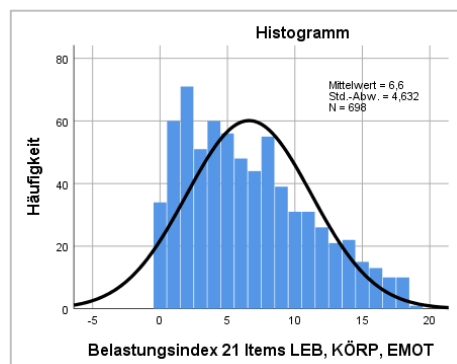
Bevölkerung 2018 - 21 Belastungsitems der Skalen LEB-(7), KÖRP(7), EMOT(7)  
 INPUT: 698 PERSON 21 ITEM REPORTED: 698 PERSON 21 ITEM 2 CATS WINSTEPS 4.5.0

MEASURE PERSON - MAP - ITEM



Fast die Hälfte der normalgesunden Bevölkerung 2018 wird am unteren Ende der hier definierten Belastungsskala mit 21 Belastungsitems nicht adäquat erfasst, weil zu wenige leichte FPI-Items.

Diagnostisch ist dies nicht weiter entscheidend, weil das FPI für

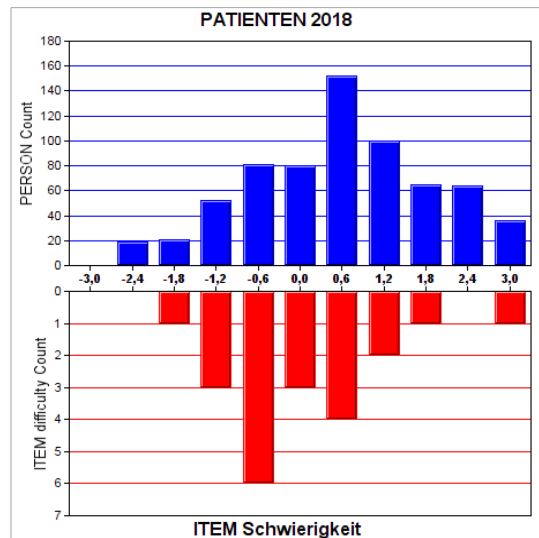
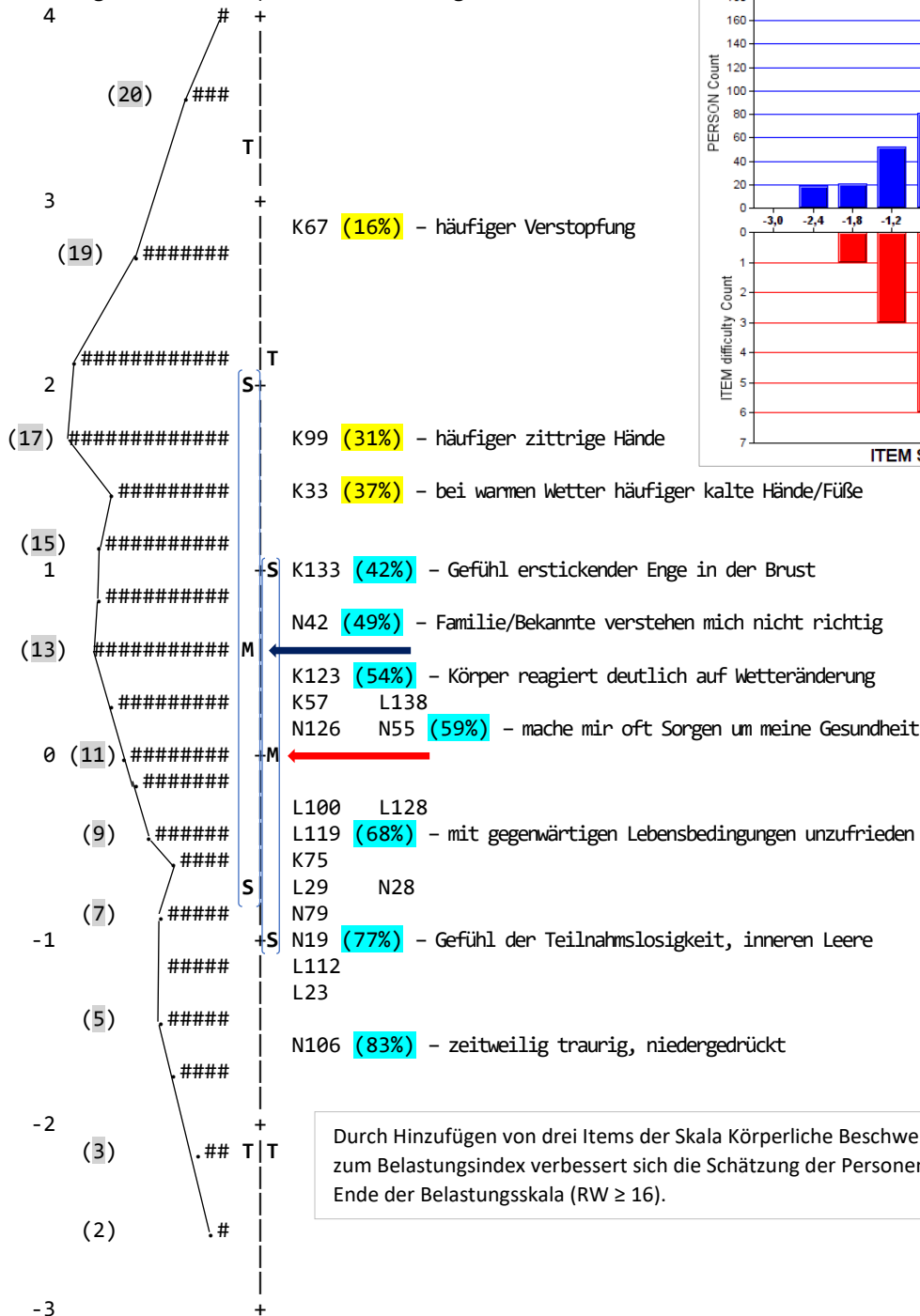


<less>|<freq> /leicht  
 EACH "#" IS 6: EACH "." IS 1 TO 5  
 M = mean, S = 1 Std.Dev, T = 2 Std.Dev.  
 In Klammern Indexwerte 0 - 21.

Patienten 2018 - 21 Belastungsitems der Skalen LEB-(7), KÖRP(7), EMOT(7)  
 INPUT: 698 PERSON 21 ITEM REPORTED: 698 PERSON 21 ITEM 2 CATS WINSTEPS 4.4.8

# MEASURE PERSON - MAP - ITEM

Belastung hoch <more>|<rare>/schwierig



Durch Hinzufügen von drei Items der Skala Körperliche Beschwerden (K67, K99, K33) zum Belastungsindex verbessert sich die Schätzung der Personenparameter am oberen Ende der Belastungsskala (RW ≥ 16).

Bel. niedrig/<less>|<freq>/leicht  
 EACH "#" IS 5: EACH "." IS 1 TO 4TABLE 12.2  
 M = mean, S = 1 Std.Dev, T = 2 Std.Dev.  
 In Klammern Indexwerte 2 - 21.

**Tabelle 7: Patienten 2018 – Item- und Personenparameter Raschanalyse**

INPUT: 698 PERSON 21 ITEMS BELASTUNGSINDEX 2 CATS WINSTEPS 4,5,0

SUMMARY OF 692 MEASURED (NON-EXTREME) PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	12.6	21.0	.59	.58	.99	.04	1.00	.04
SEM	.2	.0	.05	.00	.01	.03	.02	.04
P.SD	4.7	.0	1.37	.13	.24	.91	.49	.93
S.SD	4.7	.0	1.37	.13	.24	.91	.49	.93
MAX.	20.0	21.0	3.57	1.08	1.84	3.86	4.71	3.77
MIN.	1.0	21.0	-3.39	.48	.55	-2.91	.14	-2.22
REAL RMSE	.62	TRUE SD	1.22	SEPARATION	1.97	PERSON RELIABILITY		.79
MODEL RMSE	.59	TRUE SD	1.23	SEPARATION	2.08	PERSON RELIABILITY		.81
S.E. OF PERSON MEAN = .05								

SEPARATIONINDEX =  $(4G+1)/3 = (4*1.97+1)/3 = 2.96 = 3$  groups

MAXIMUM EXTREME SCORE: 5 PERSON .7%

MINIMUM EXTREME SCORE: 1 PERSON .1%

SUMMARY OF 698 MEASURED (EXTREME AND NON-EXTREME) PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	12.7	21.0	.61	.59				
SEM	.2	.0	.05	.01				
P.SD	4.8	.0	1.42	.17				
S.SD	4.8	.0	1.42	.17				
MAX.	21.0	21.0	4.89	1.87				
MIN.	.0	21.0	-4.65	.48				
REAL RMSE	.64	TRUE SD	1.27	SEPARATION	1.98	PERSON RELIABILITY		.80
MODEL RMSE	.61	TRUE SD	1.28	SEPARATION	2.09	PERSON RELIABILITY		.81
S.E. OF PERSON MEAN = .05								

SEPARATIONINDEX =  $(4G+1)/3 = (4*1.98+1)/3 = 2.97 = 3$  groups

PERSON RAW SCORE-TO-MEASURE CORRELATION = .98

CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .85 SEM = 1.87

SUMMARY OF 21 MEASURED (NON-EXTREME) ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	421.6	698.0	.00	.10	.99	-.14	1.00	-.02
SEM	26.8	.0	.24	.00	.03	.75	.06	.69
P.SD	119.8	.0	1.06	.01	.15	3.36	.28	3.06
S.SD	122.8	.0	1.08	.01	.16	3.44	.28	3.14
MAX.	582.0	698.0	2.80	.12	1.30	6.83	1.67	6.79
MIN.	114.0	698.0	-1.55	.09	.79	-4.40	.60	-4.24
REAL RMSE	.10	TRUE SD	1.05	SEPARATION	10.53	ITEM	RELIABILITY	.99
MODEL RMSE	.10	TRUE SD	1.05	SEPARATION	10.81	ITEM	RELIABILITY	.99
S.E. OF ITEM MEAN = .24								

Zusammenfassung: ausreichende Personen-Reliabilität (Cronbach- $\alpha$ : 0,85) und gute Item-Reliabilität (0,99) bei Verwendung von 21 FPI-R Items. Der Personen-Separationsindex weist auf 3 Belastungsniveaus hin. Der Mittelwert für das Personenmaß von .59 logits und die zugehörige Standardabweichung von 1.37 zeigen an, dass der „Belastungstest“ für diese Patientenstichprobe zu leicht ist. Die Item-Schätzungen sind verlässlicher als die Personen-Schätzungen.

**Tabelle 8: Patienten 2018 (N=698) – Item-Parameter (Raschanalyse mit 21 Items)**

PERSON: REAL SEP.: 1.98 REL.: .80 ... ITEM: REAL SEP.: 10.53 REL.: .99

**ITEM STATISTICS: MISFIT ORDER**

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-AL CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM
20	257	698	1.36	.09	1.30	6.70	1.67	6.79	A .29	.50	66.3	74.1	K33
11	413	698	.11	.09	1.30	6.83	1.48	6.11	B .32	.51	64.0	73.5	N55
2	512	698	-.76	.10	1.22	4.14	1.40	3.41	C .34	.48	74.9	79.0	L29
17	380	698	.38	.09	1.24	5.68	1.38	5.29	D .37	.51	63.6	72.6	K123
19	218	698	1.70	.09	1.06	1.40	1.15	1.45	E .44	.48	75.1	75.9	K99
10	340	698	.69	.09	1.04	1.15	1.11	1.67	F .48	.51	70.7	72.3	N42
21	114	698	2.80	.12	1.06	.82	1.09	.56	G .37	.41	84.2	85.2	K67
16	494	698	-.59	.10	.97	-.51	1.02	.29	H .50	.49	78.9	77.8	K75
15	384	698	.35	.09	1.00	-.05	.96	-.63	I .52	.51	72.8	72.7	K57
5	475	698	-.42	.09	.99	-.30	.90	-1.19	J .51	.50	75.6	76.5	L119
9	500	698	-.65	.10	.96	-.71	.88	-1.24	K .51	.48	78.0	78.1	N28
3	463	698	-.31	.09	.92	-1.85	.95	-.54	j .54	.50	79.0	75.8	L100
1	561	698	-1.29	.11	.94	-.91	.77	-1.73	i .49	.44	82.8	83.0	L23
6	461	698	-.29	.09	.90	-2.21	.86	-1.77	h .56	.50	78.2	75.7	L128
4	550	698	-1.16	.11	.88	-2.17	.79	-1.74	g .53	.45	83.8	81.9	L112
7	389	698	.31	.09	.88	-3.24	.86	-2.26	f .58	.51	77.3	72.7	L138
13	582	698	-1.55	.12	.88	-1.83	.82	-1.18	e .49	.42	86.6	85.0	N106
18	295	698	1.05	.09	.86	-3.68	.86	-1.96	d .58	.51	78.8	72.9	K133
14	404	698	.18	.09	.83	-4.40	.74	-4.24	c .62	.51	78.9	73.2	N126
12	525	698	-.89	.10	.81	-3.88	.65	-3.51	b .59	.47	84.1	79.9	N79
8	536	698	-1.01	.10	.79	-4.03	.60	-3.93	a .59	.46	84.0	80.8	N19
MEAN	421.6	698.0	.00	.10	.99	-.1	1.00	.0			77.0	77.1	
P.SD	119.8	.0	1.06	.01	.15	3.4	.28	3.1			6.4	4.1	

Anmerkung: Infit MNSQ zwischen .70 und 1.3 gilt als akzeptabel. Infit ZSTD > 2.0 deutet auf Modellunver-träglichkeit (misfit) hin (8 von 21 Items). ZSTD (t-Statistik) ist stichprobenabhängig, wie die folg. Tabelle zeigt.

**Tabelle 9: Patienten 2018 - Zufallsstichprobe (N=100 aus N=698)**

PERSON: REAL SEP.: 1.95 REL.: .79 ... ITEM: REAL SEP.: 4.32 REL.: .95

**ITEM STATISTICS: MISFIT ORDER**

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-AL CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM
8	34	100	1.74	.25	1.45	3.59	2.81	4.74	A .15	.49	62.6	75.1	K33
18	66	100	-.10	.25	1.29	2.25	1.78	2.99	B .32	.51	69.7	76.4	N55
13	53	100	.66	.24	1.28	2.49	1.51	2.51	C .35	.52	63.6	72.9	K123
2	76	100	-.78	.27	1.14	.99	1.25	.86	D .39	.48	77.8	80.8	L29
17	56	100	.49	.24	1.05	.48	1.16	.92	E .48	.53	73.7	73.5	N42
4	83	100	-1.35	.30	.83	-.98	1.12	.41	F .52	.44	86.9	84.8	L112
7	62	100	.14	.24	.87	-1.13	1.05	.31	G .58	.52	79.8	75.1	L138
3	70	100	-.36	.26	.93	-.52	1.02	.16	H .54	.50	80.8	78.0	L100
16	69	100	-.29	.25	1.01	.16	.95	-.14	I .50	.51	77.8	77.6	N28
10	21	100	2.62	.28	.98	-.07	.81	-.40	J .45	.44	82.8	81.5	K67
9	53	100	.66	.24	.95	-.42	.88	-.62	K .56	.52	71.7	72.9	K57
5	70	100	-.36	.26	.94	-.42	.94	-.19	j .53	.50	78.8	78.0	L119
12	25	100	2.32	.26	.94	-.41	.80	-.49	i .50	.46	78.8	78.7	K99
6	71	100	-.43	.26	.91	-.68	.79	-.79	h .56	.50	79.8	78.4	L128
1	83	100	-1.35	.30	.88	-.65	.60	-1.00	g .52	.44	84.8	84.8	L23
19	77	100	-.85	.28	.80	-1.41	.88	-.28	f .58	.48	84.8	81.3	N79
21	57	100	.43	.24	.88	-1.10	.76	-1.34	e .60	.52	74.7	73.7	N126
11	69	100	-.29	.25	.87	-.99	.78	-.95	d .58	.51	79.8	77.6	K75
14	45	100	1.10	.24	.84	-1.61	.83	-.85	c .60	.52	83.8	72.9	K133
20	93	100	-2.61	.43	.84	-.44	.63	-.30	b .40	.33	93.9	93.4	N106
15	83	100	-1.35	.30	.80	-1.19	.47	-1.47	a .58	.44	86.9	84.8	N19
MEAN	62.7	100.0	.00	.27	.98	-.1	1.04	.2			78.7	78.7	
P.SD	18.7	.0	1.23	.04	.17	1.3	.49	1.5			7.4	5.0	

Abbildung: Item - Fit-Statistiken aus Tabelle 9 (N=100 Patienten, n=21 Belastungsitems)

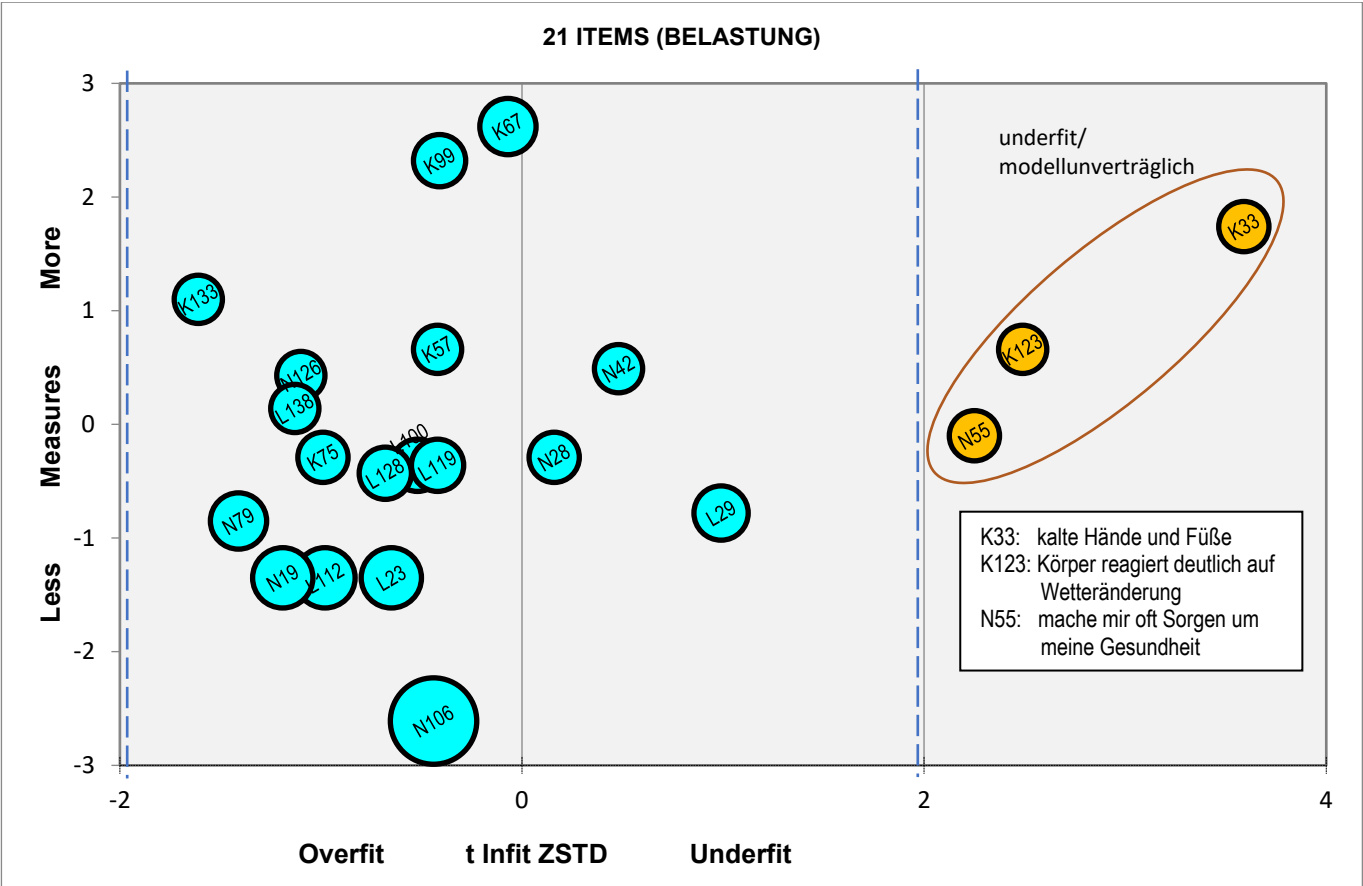


Abbildung: Personen - Fit-Statistiken (N=100 Patienten, n=21 Belastungsitems)

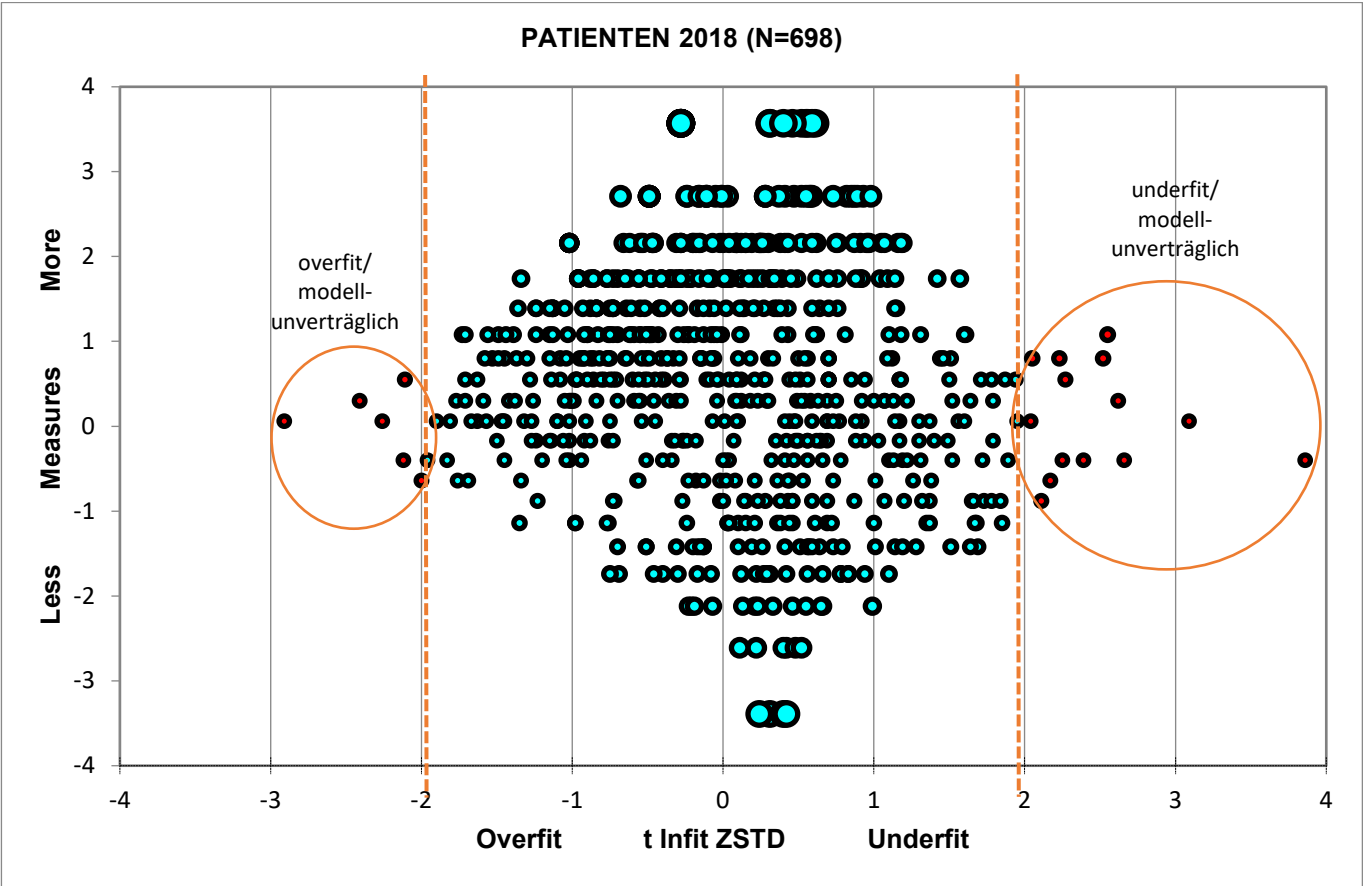


Abbildung: ICC für 3 ausgewählte Items (K67, N42, L112; Patienten 2018)

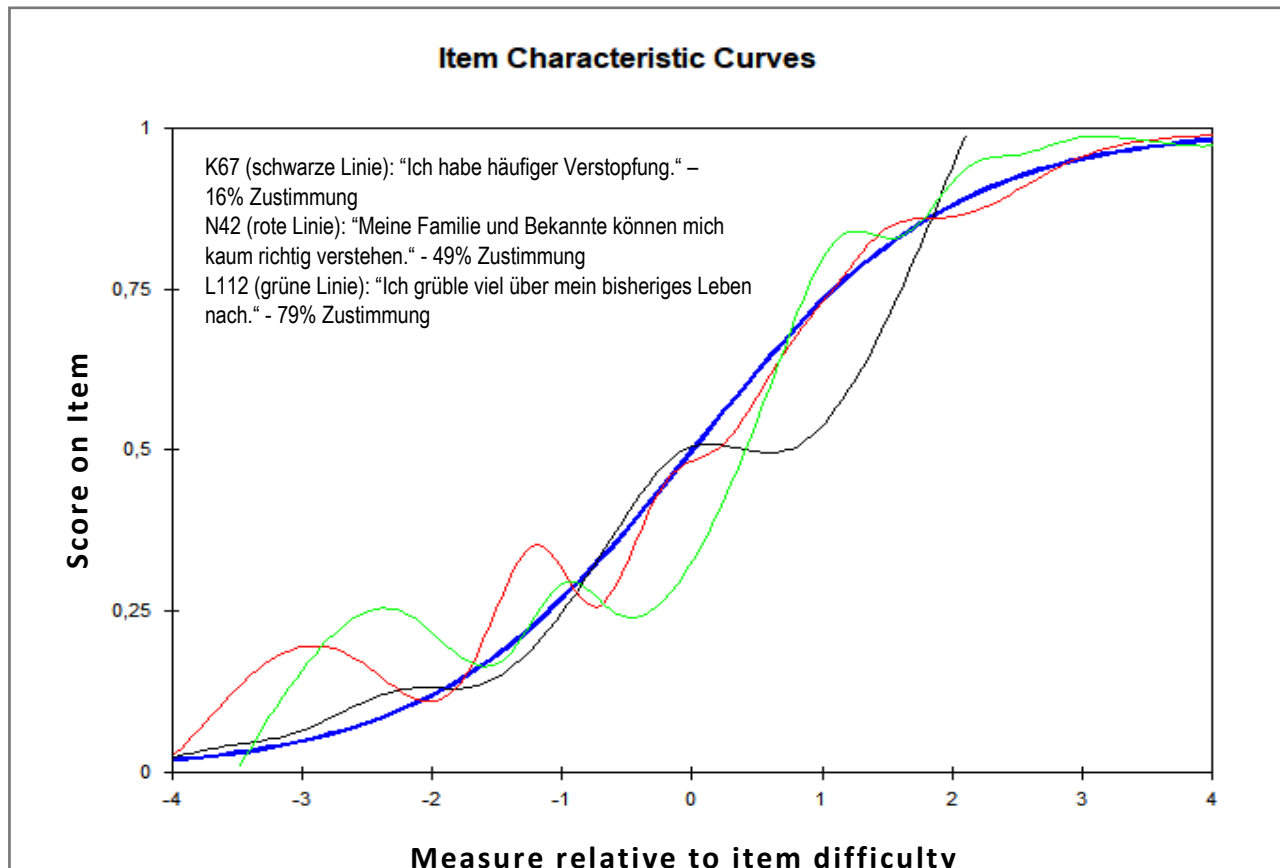


Abbildung: DIF (Differential Item Functioning) nach Gruppe (Bevölkerung vs. Patienten)

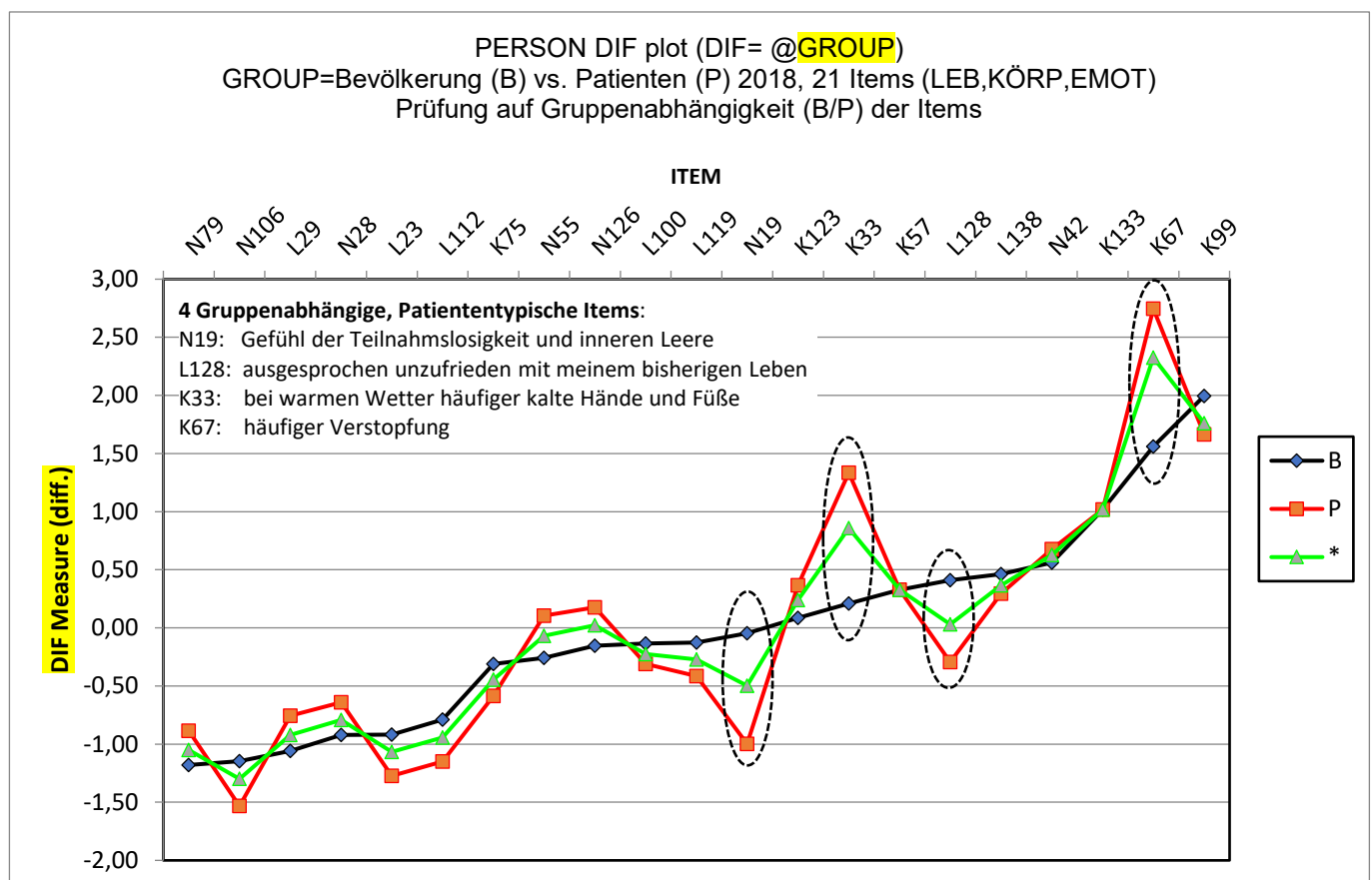


Abbildung: DIF nach Geschlecht (Patienten)

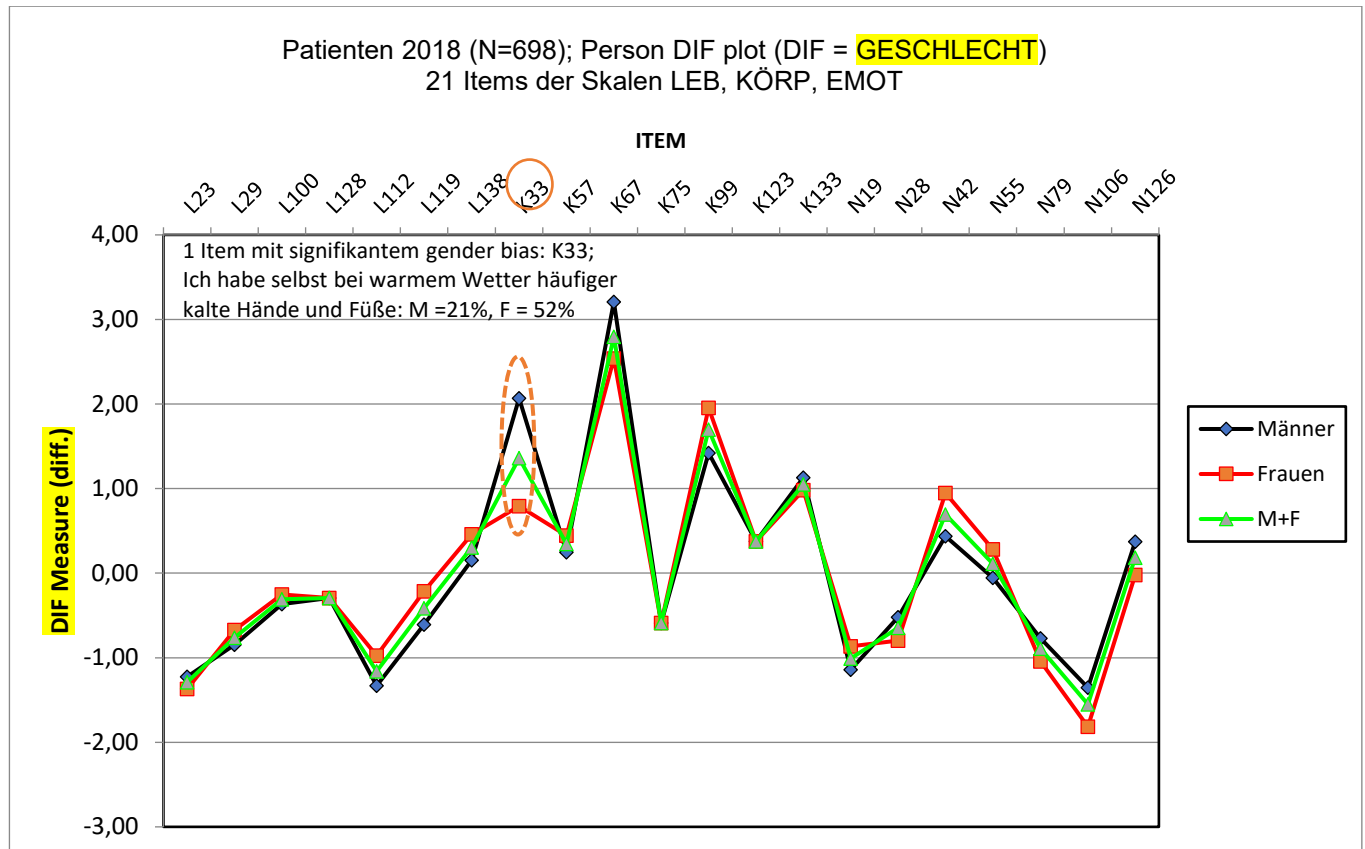
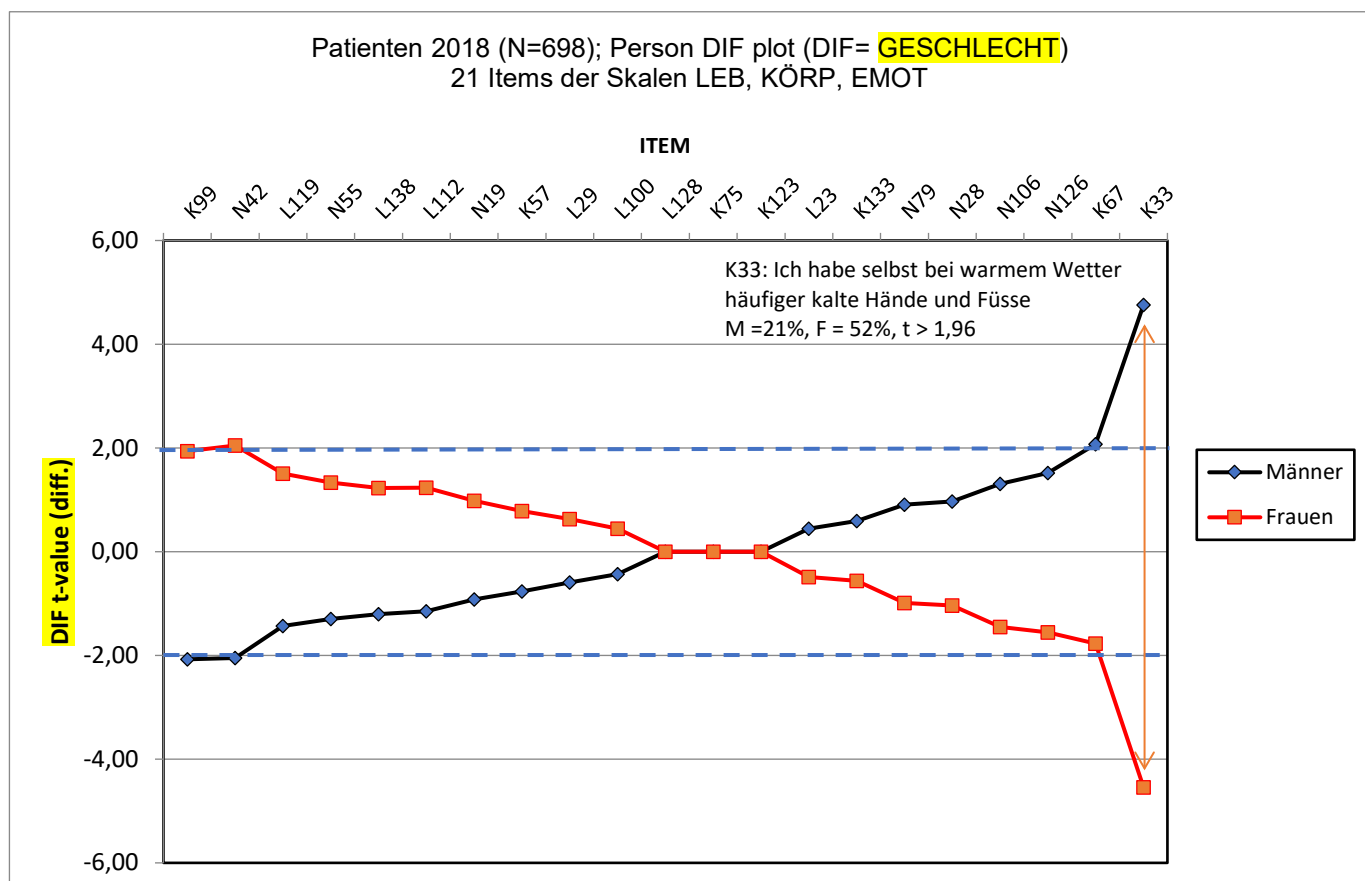


Abbildung: DIF T-Test für Geschlecht (Patienten)





### Zusammenfassung der Ergebnisse des Vergleichs von Bevölkerung und Patienten 2018:

Die Ergebnisse sind ein Beleg für die diskriminante- und Konstrukt-Validität des FPI-R.

1. Patienten und normalgesunde Probanden lassen sich deutlich voneinander trennen. Vor allem die Items der Skalen Lebenszufriedenheit, Emotionalität und Körperliche Beschwerden sind effektstarke Indikatoren psychisch-körperlicher Belastung. Aufgrund der Diskriminanzanalysen kann gefolgert werden, dass 3 von 4 Probanden der Bevölkerung sowie psychotherapeutische Patienten mit dem FPI-R adäquat diagnostiziert werden können.
2. Personen-Clusteranalysen (Bevölkerung plus Patienten 2018) auf Basis der Skalenwerte von Lebenszufriedenheit, Emotionalität und Körperliche Beschwerden führten zu drei Clustern von niedrig-, mittel- und hoch-belasteten Probanden. Dem Cluster der hoch-belasteten Probanden wurden 18% aus der Bevölkerungs- und 82% aus der Patientenstichprobe zugeordnet. In der Bevölkerungsstichprobe können 9% als hoch-belastet und 58% als niedrig-belastet beschrieben werden. Für die Patientenstichprobe lauten die vergleichbaren Prozentwerte: 42% hoch-belastet und 19% niedrig-belastet.
3. Die Rasch-Analysen für die Gruppe der Patienten ergaben, dass mit den Items der Skalen Lebenszufriedenheit, Emotionalität und Körperliche Beschwerden verschiedene Belastungsniveaus in der Mehrzahl der Fälle annähernd treffend erfasst werden. Die Item-Reliabilität ist hoch (0,99) und die Personen-Reliabilität befriedigend (0,85). Es gibt zwar einige Gruppenabhängige Items (DIF) wie z.B. „Ich habe manchmal ein Gefühl der Teilnahmslosigkeit und inneren Leere“, doch die Zahl der invarianten Items für Patienten und Bevölkerung (DIF) beläuft sich auf ungefähr 80 Prozent. Auch das Geschlecht hat keinen Item-Bias, bis auf das Item „Ich habe selbst bei warmen Wetter häufiger kalte Hände und Füße“ (Männer 21% vs. Frauen 52%).

## IX. EIGENSCHAFTSPROFILE AUFGRUND VON CLUSTERANALYSEN (Bevölkerung 2018)

Im FPI-Manual wurden Clusteranalysen der *Items* im Hinblick auf die *Skalenkonstruktion* ausführlich dargestellt. Auch Cluster von *Personen* aufgrund von Skalenwerten können mit dieser statistischen Methodik erkundet werden.

In der älteren psychologischen Testdiagnostik wurde, insbesondere bei den Untertests von Intelligenztests, auch nach der relativen Ausprägung der Skalenwerte gefragt, intraindividuell (ipsativ) und interindividuell vergleichend. Die Begriffe *Intelligenzprofil*, *Begabungsprofil* oder *Einstellungsprofil* erinnern an diese Sichtweise. Methodisch schlossen sich hier metrische Typenanalysen an die ältere, vorwiegend deskriptive Typenlehre an. Der Übergang von der typologischen zur dimensional Beschreibung von Persönlichkeitseigenschaften war ein wichtiger Schritte in der frühen Aus-einander-setzung Eysencks mit der älteren Charakterkunde. Solche Typisierungen sind jedoch erhalten geblieben, alltagspsychologisch, aber auch fachpsychologisch: „die Extravertierte“ oder der „Emotional Labile“.

Hinsichtlich der individuellen Profile aufgrund der 12 FPI-Testwerte ist zu fragen: Wie lautet die Zufallserwartung und welche Arbeitshypothesen gibt es?

Die Zufallserwartung ist, dass es eine diffuse „Normal“-Verteilung von *Profilgestalten* gibt, extrem dicht im mittleren Bereich und ausdünnend zu den Extrema. Dieser Ansatz muss jedoch noch spezifiziert werden: Die beiden Testwerte E und N sind mit Testwerten anderer Skalen konfundiert. Für explorative Analysen ist es zweckmäßig, nur die Skalen 1 bis 9, ohne die Skala Offenheit, aufzunehmen. Auch zwischen den Testwerten der verbleibenden 9 Standardskalen bestehen Korrelationen, von denen einige eine mittlere Effektstärke erreichen. Die modifizierte Arbeitshypothese lautet demnach: es sind relative Maxima von Profilgestalten aufzuweisen, die zunächst hinsichtlich der bekannten Skalen-Korrelationen zu interpretieren sind, aber darüber hinaus eventuell auf heuristisch interpretierbare Profilgestalten hinweisen.

Einen approximativen Eindruck könnte bereits eine Zufallsstichprobe der Gesamtstichprobe geben. Psychologische Hilfs-hypothesen könnten lauten: hohe Profilstreuung bedeutet pauschal eine extreme Abweichung der Selbstbeschreibung von dem Mittelwert in der Normierungsstichprobe, niedrige Profilstreuung eine durchschnittlich undifferenzierte Selbstbeschreibung auf der Ebene der FPI-Testwert-Profile. Beide Extrema legen eine Interpretation (1) hinsichtlich der Zuverlässigkeit der Antworten (Antwortstil, Missverständnisse, Ankreuzverhalten) und (2) hinsichtlich der persönlichkeitspsychologischen Bedeutung nahe: kontrastreiche, höchst individuelle, eventuell widersprüchliche gegenüber uniform-unauffälligen „durchschnittlichen“ Selbstbeschreibungen. Explorativ könnten anhand der FPI-Rohwerte dieser Personen die Beziehungen zu Geschlecht, Alter, Schulbildung und anderen Daten der Erhebung geprüft werden.

Prägnantere Antworten als diese Betrachtung von Profilgestalt und Profilstreuung könnte vielleicht die Suche nach Personen-Clustern ergeben. Hier wird die Individuen x Variablen-Matrix hinsichtlich der statistischen Gruppierung von Individuen auf der Basis der Skalenwerte vorgenommen (nicht auf der Ebene der nur dichotomen Itemantworten wegen der nur geringen Variation).

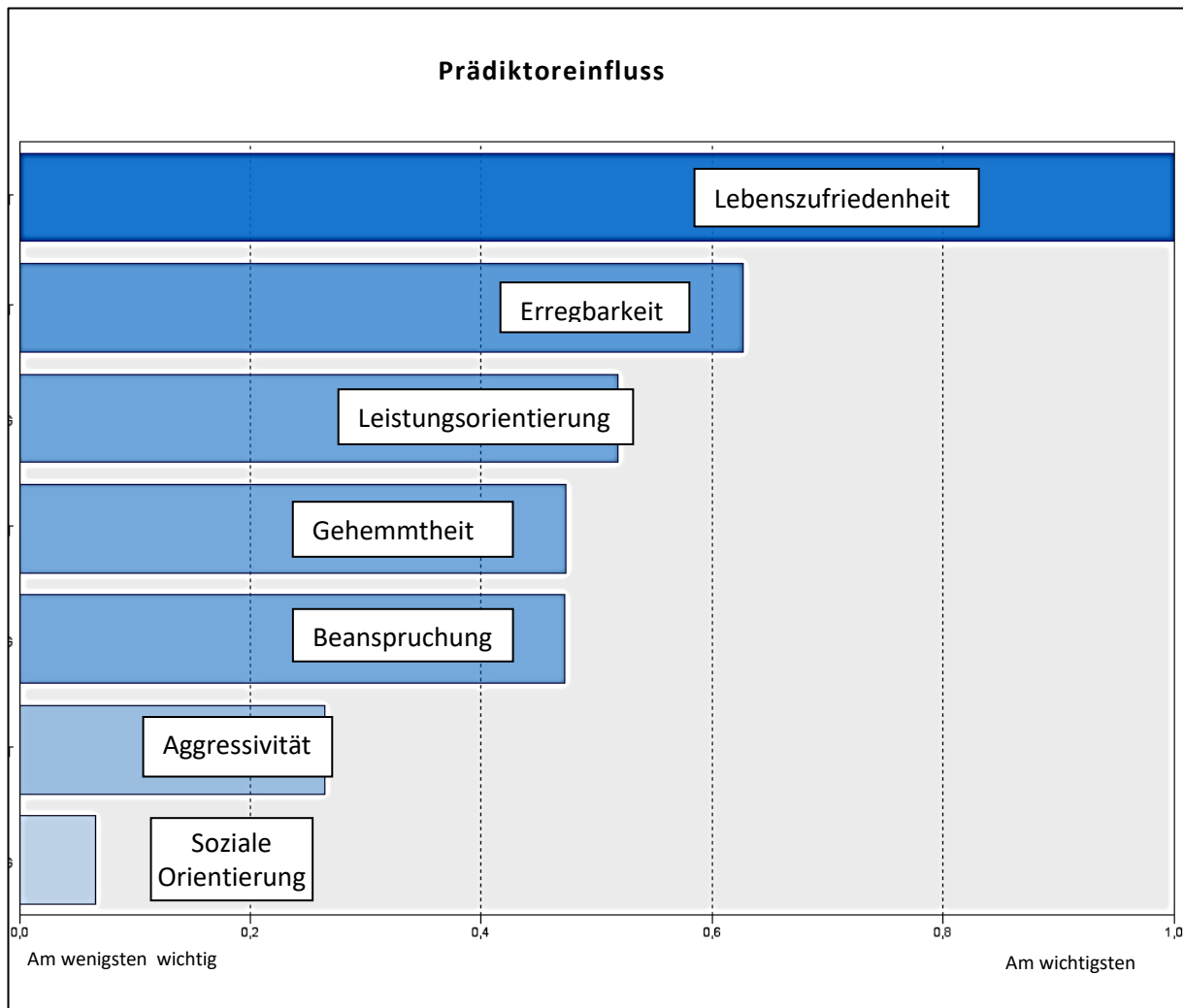
Mit dem SPSS-Programm *Two-Step-Clusteranalyse* (BIRCH-Algorithmus: Balances Iterative Reducing and Clustering using Hierachies) wurde explorativ eine Reihe von Personen-Clusteranalysen durchgeführt. Die hier mitgeteilten Ergebnisse beschränken sich nur auf die männliche Altersgruppe der 50-65-Jährigen, weil, so die Vermutung, in dieser Altersgruppe relativ prägnant ausgebildete und relativ stabile Persönlichkeitsdispositionen häufiger zu finden sind. Zur Cluster-gruppierung wurden nur 7 der 10 FPI-R-Standardskalen verwendet, ohne die Skala Offenheit und den Gesundheitskomplex (Körperliche Beschwerden und Gesundheitssorgen). Der Gesundheitskomplex wurde ausgeklammert, weil bekannt ist, dass sich (1) Männer und Frauen jener Altersgruppe in Hinsicht auf Körperliche Beschwerden und Gesundheitssorgen von vornherein signifikant unterscheiden (Frauen haben im Durchschnitt höhere Skalenwerte und der Prädiktoreinfluss von Körperlichen Beschwerden und Gesundheitssorgen auf die Clusterbildung ist erheblich).

Da keine Vorinformation zur möglichen Clusterzahl vorliegt, musste im SPSS-Programm per Voreinstellung die automatische Clusterermittlung gewählt werden (log-likelihood Distanzmaß). Die automatische Clusterlösung ergab 3 bzw. 4 Gruppen: (1) selbstsichere/kontaktbereite Personen (35%, GEH sehr niedrig); (2) kontrollierte, zurückhaltende Personen (31%, AGGR niedrig); (3) Lebensunzufriedene (18%, LEB niedrig); (4) Erregbare, Beanspruchte (16%, ERR und BEAN hoch). Im Folgenden werden nur die Ergebnisse der 3-Clusterlösung im Einzelnen dargestellt.

## Two-Step-Personen-Clusteranalyse

50-65-jährige Männer (N = 447, Bevölkerung 2018)

Prädiktoreinfluss von 7 ausgewählten FPI-R Standardskalen, 3 Clusterlösung



**Cluster 1:** Emotional Labile – Erregbarkeit, Beanspruchung, Gehemmtheit hoch;  
Lebenszufriedenheit, Leistungsorientierung niedrig (43% von Gesamt)

**Cluster 2:** Zufriedene und Ausgeglichene – Lebenszufriedenheit hoch; Erregbarkeit,  
Beanspruchung, Aggressivität niedrig (22% von Gesamt)

**Cluster 3:** Selbstsichere, Leistungsorientierte, Zufriedene – Gehemmtheit niedrig,  
Leistungsorientierung hoch, Lebenszufriedenheit hoch (35% von Gesamt)

### 3 Clusterlösung: 50-65-jährige Männer (Bevölkerung 2018)

#### Clustervergleich

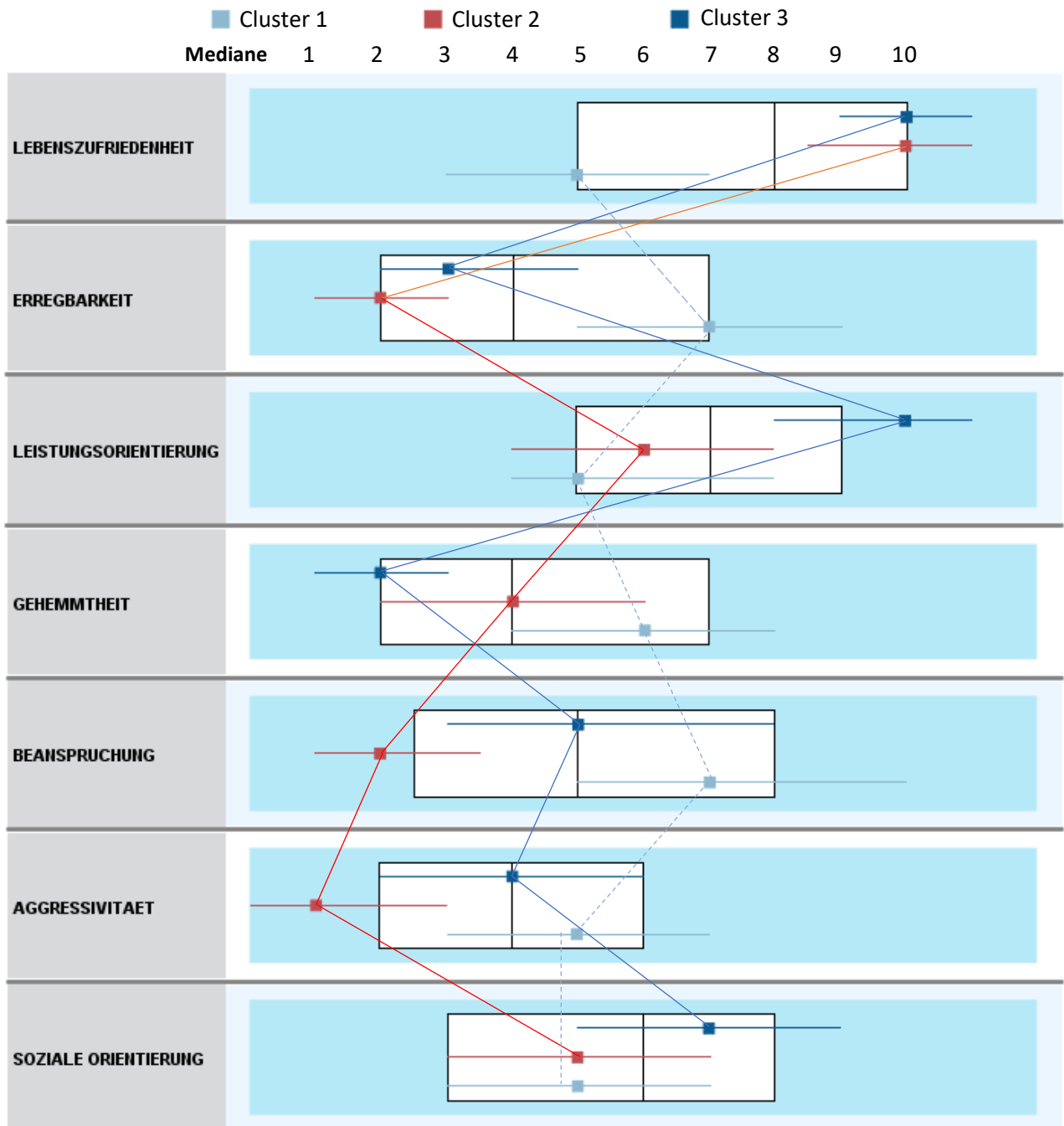
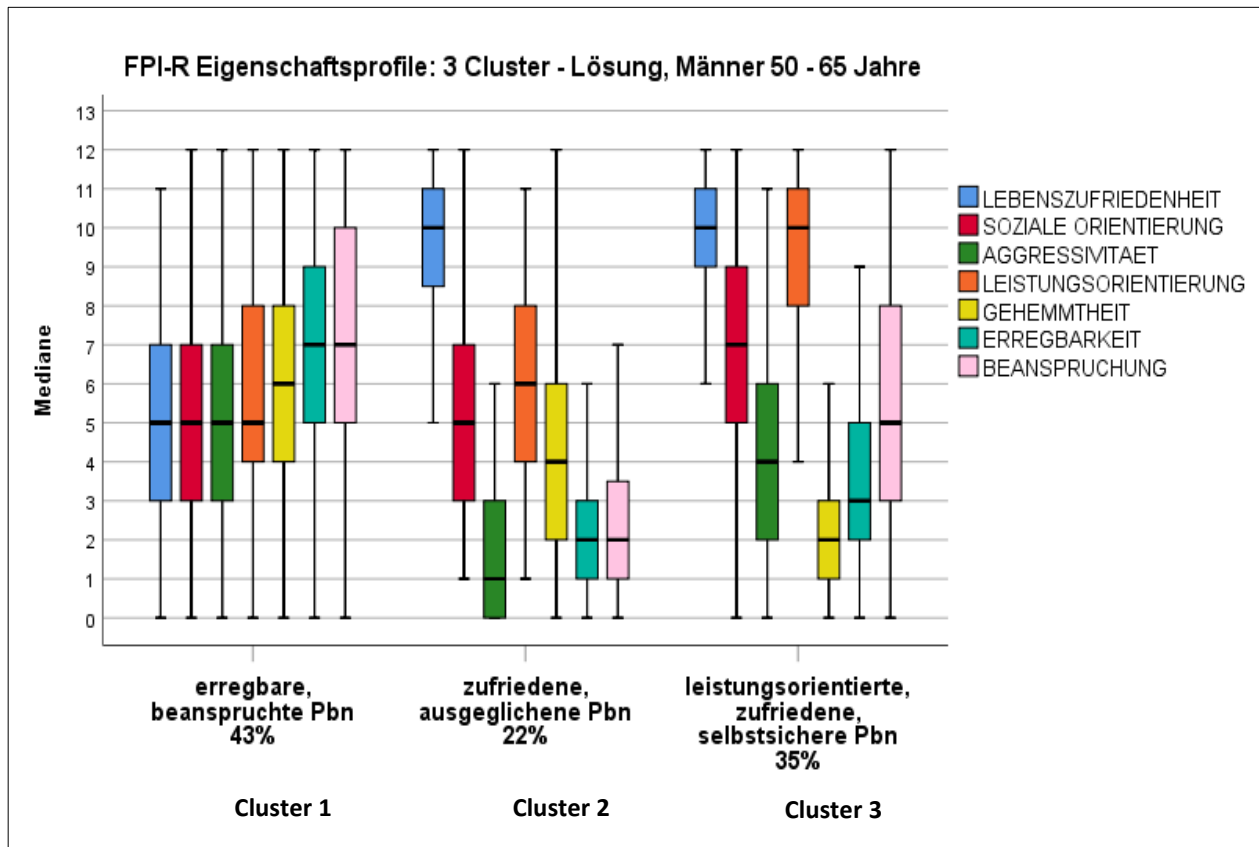


Abbildung: Boxplot Darstellung der Personen-Clusterlösung (vgl. Abbildung vorherige Seite)



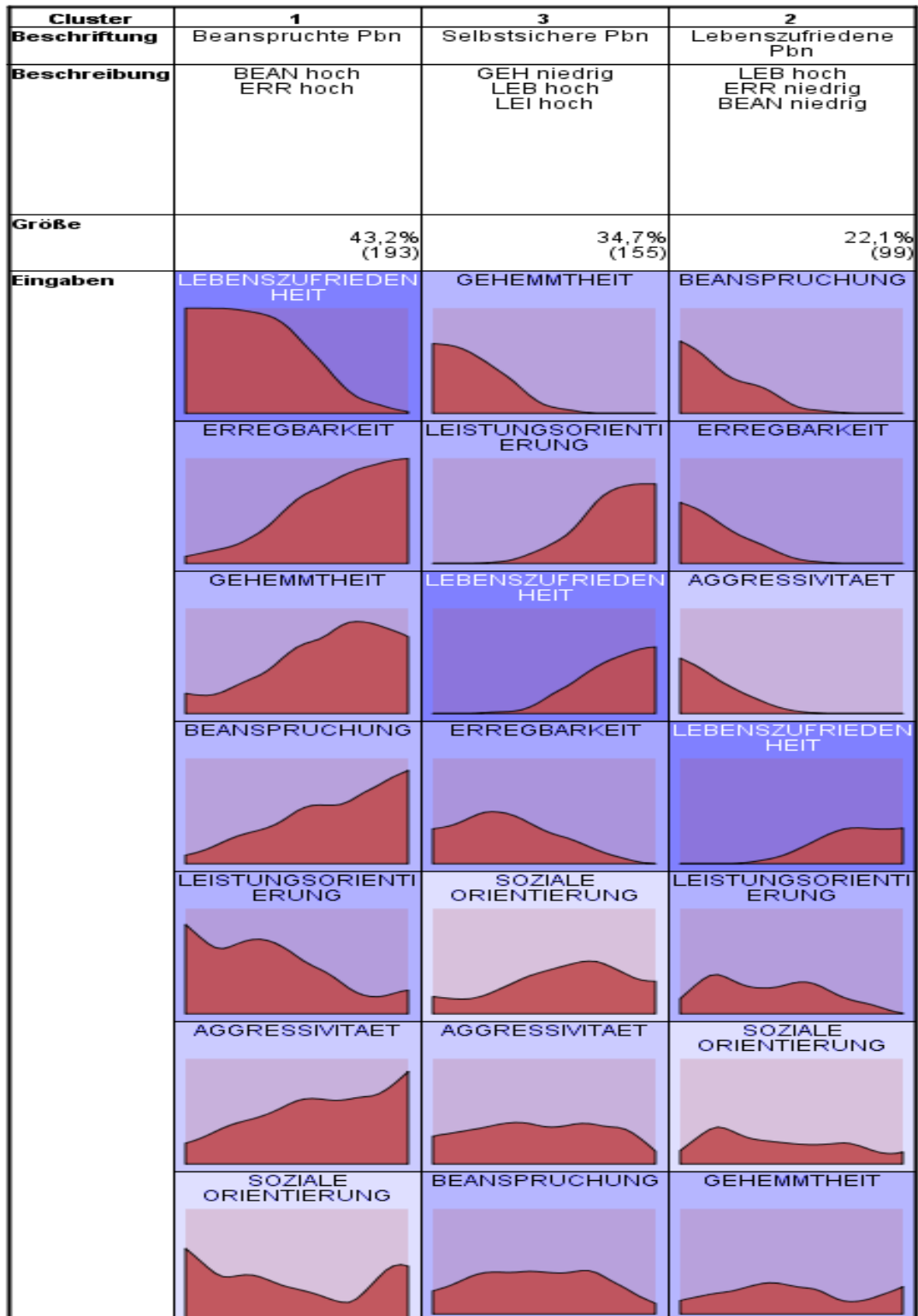
Mittelwerte und Mediane nach Clusterbildung

	Cluster 1			Cluster 2			Cluster 3		
	M	Median	N	M	Median	N	M	Median	N
LEBENSZUFRIEDENHEIT	4.9	5	193	9.6	10↑	99	9.9	10↑	155
SOZIALE ORIENTIERUNG	5.3	5	193	5.2	5	99	6.7	7	155
LEISTUNGSORIENTIERUNG	5.7	5	193	5.6	6	99	9.5	10↑	155
GEHEMMTHEIT	6.1	6↑	193	4.5	4	99	2.2	2↓	155
ERREGBARKEIT	7.0	7↑	193	2.2	2↓	99	3.8	3↓	155
AGGRESSIVITÄT	5.0	5	193	1.7	1↓	99	4.2	4	155
BEANSPRUCHUNG	7.3	7↑	193	2.2	2↓	99	5.4	5	155
<i>in Clusteranalyse ausgeklammert</i>									
KÖRPERLICHE BESCHWERDEN	4.2	4	193	1.3	1↓	99	1.7	1↓	155
GESUNDHEITSSORGEN	5.9	6	193	5.8	6	99	6.0	6	155
OFFENHEIT	7.2	8↑	193	4.8	5	99	5.5	5	155
EXTRAVERSION	5.5	5	193	5.1	4	99	9.1	9↑	155
EMOTIONALITÄT	8.0	8↑	193	2.2	2	99	3.2	3	155

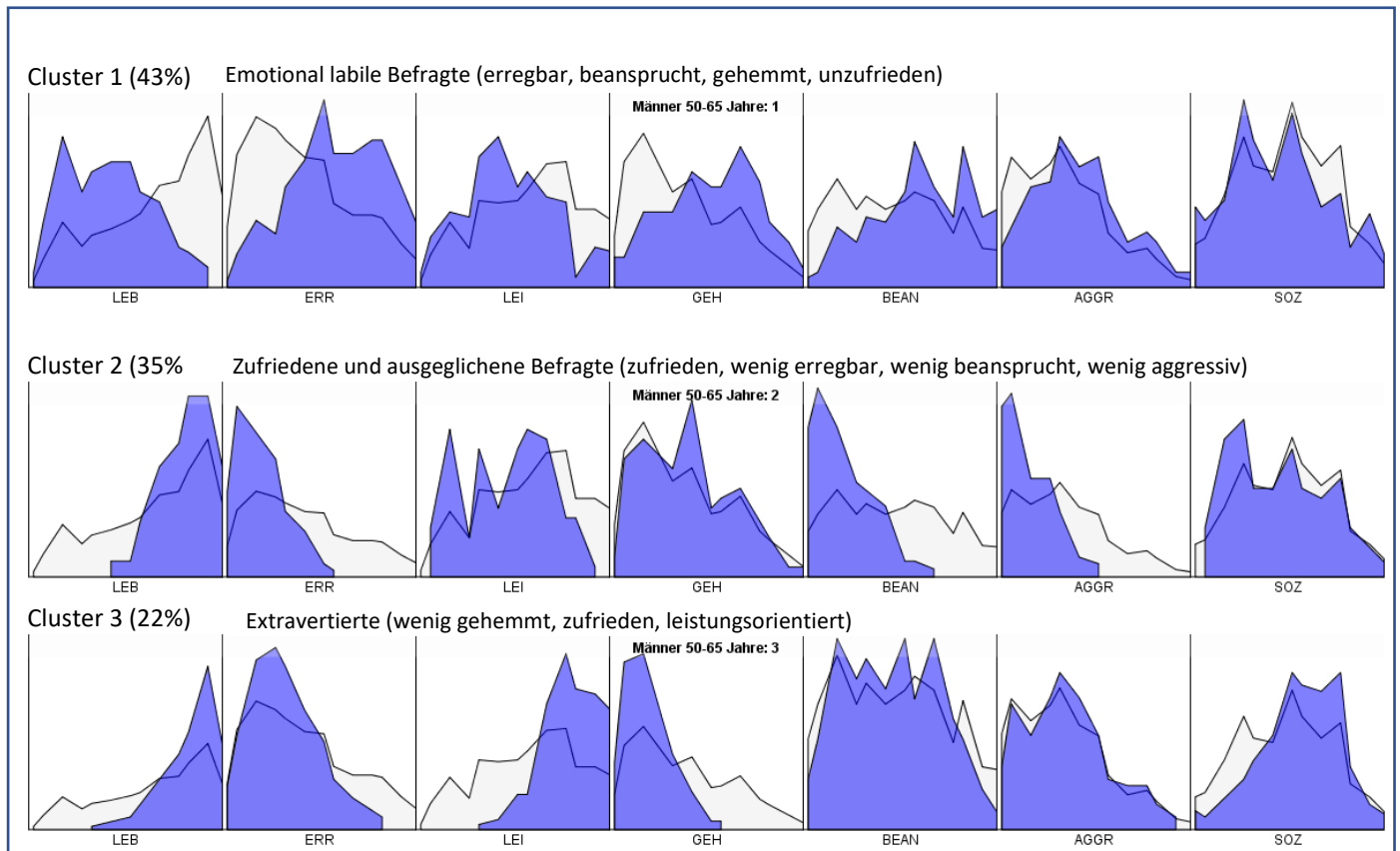
**3 Clusterlösung: Skalenwert-Verteilung pro Prädiktor**  
(sortiert nach Prädiktoreinfluss innerhalb des Cluster)

**Cluster**

Wichtigkeit der Eingabe (Prädiktor)  
 1,0 0,8 0,6 0,4 0,2 0,0



## FPI-R Skalenwertverteilung pro Prädiktor (ohne Sortierung nach Prädiktoreinfluss)



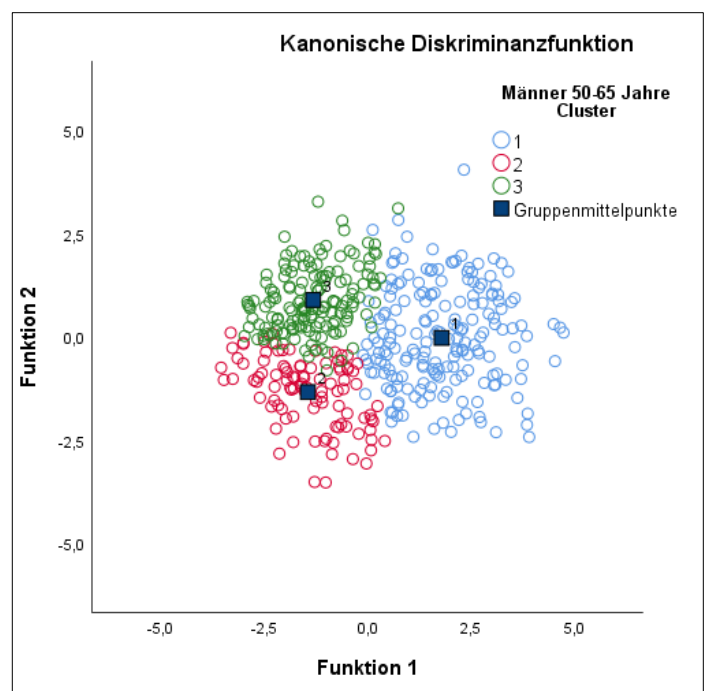
## Clusterüberprüfung mittels Diskriminanzanalyse (7 Prädiktoren, 2 kanonische Diskriminanzfunktionen)

Funktion	Eigenwert	Eigenwerte		Kanonische Korrelation
		% der Varianz	Kumulierte %	
1	2.463 <sup>a</sup>	78.5	78.5	.843
2	.675 <sup>a</sup>	21.5	100.0	.635

a. Die ersten 2 kanonischen Diskriminanzfunktionen werden verwendet.

	Struktur-Matrix	
	Funktion	
	1	2
LEBENSZUFRIEDENHEIT	-.723*	.108
ERREGBARKEIT	.509*	.259
LEISTUNGSORIENTIERUNG	-.275	.722*
BEANSPRUCHUNG	.362	.483*
GEHEMMTHEIT	.381	-.431*
AGGRESSIVITÄT	.227	.420*
SOZIALE ORIENTIERUNG	-.088	.234*

\* Größte absolute Korrelation zwischen jeder Variablen und einer Diskriminanzfunktion



Diskriminanzanalyse auf Clusterbasis - Klassifizierungsergebnisse <sup>a</sup>					
		Vorhergesagte Gruppenzugehörigkeit			Gesamt
		Männer 50-65 Jahre	1	2	
Original	Anzahl	Cluster 1	187	1	193
		Cluster 2	0	85	99
		Cluster 3	0	3	155
	Prozent	Cluster 1	97	0	100
		Cluster 2	0	86	100
		Cluster 3	0	2	100

a. 95% der ursprünglich gruppierten Fälle wurden korrekt klassifiziert.

Männer 50-65 Jahre	Sozio-demografische Merkmale, Interviewer-Urteile, Selbstauskünfte	Emotional Labile	Zufriedene/Ausgeglichene	Extravertierte/Leistungs-orientierte	
		Cluster 1	Cluster 2	Cluster 3	Gesamt
		Spalten (%)			
Schulabschluss	Haupt-/Volksschule	25% ↑	14%	11%	18%
	ABI/Studium	31%	54%	47%	42%
Berufstätigkeit	Ja, berufstätig	82%	95%	96%	89%
	Arbeitslose	18% ↑	5%	2%	10%
Berufsstellung	Arbeiter	39% ↑	19%	21%	21%
	Nichtleitender Angestellter	33%	40%	27%	33%
Familienstand	Verheiratet (zusammenlebend)	59%	71%	79%	68%
	Ledig	23% ↑	13%	8%	15%
Haushaltseinkommen	<= 1.999 €	32% ↑	20%	13%	22%
Soziale Schicht	OS/Obere Mittelschicht	20%	40%	40%	31%
	UMS/Unterschicht	80% ↑	60%	60%	69%
Parteisympathie	AfD	25% ↑	8%	7%	15%
Interviewer-Beurteilung	sehr selbstsicher, von sich überzeugt	27% ↓	34% ↔	48%	36%
	sehr ungezwungen, kontaktbereit	26% ↓	41%	38%	33%
	sehr durchsetzungsbereit	12% ↓	14% ↔	27%	18%
	eher erregbar, empfindlich	20% ↑	4%	11%	13%
Psychotherapie	Ja	25% ↑	8%	7%	15%
Überstunden	deutlich mehr als vorgesehen	23% ↓	25% ↔	36%	29%
Kontakt mit Freunden in der Freizeit	Täglich, mehrmals täglich	25% ↓	24% ↔	40%	30%

Personen des Cluster 1 (Emotional Labile) sind im Vergleich zu Personen in den Clustern 2 und 3 weniger gebildet, weniger berufstätig bzw. häufiger arbeitslos; eher Arbeiter, häufiger ledig; haben ein geringeres Einkommen; sind häufiger Angehörige der unteren Mittelschicht oder Unterschicht; häufiger AfD-Parteisympathisanten; sind/waren häufiger in Psychotherapie und wurden von den Interviewern als eher erregbar, weniger durchsetzungsbereit, eher gehemmt beschrieben.

Das soziodemografische Profil der Personen in Cluster 2 und 3 ist ähnlich. Auffällige Unterschiede zwischen den Personen im Cluster 2 und 3 gibt es aber in Hinsicht auf die Fremdurteile, Überstunden und Freundeskontakte. Cluster 3 Personen (Extravertierte/Leistungsorientierte) wurden im Gegensatz zu den Cluster 2 Personen von den Interviewern häufiger als selbstsicher und durchsetzungsbereit eingeschätzt; sie leisten mehr Überstunden als im Arbeitsvertrag vorgesehen und haben häufiger Kontakt zu Freunden in der Freizeit.



### Zusammenfassung:

In diesem Kapitel sind zusätzliche Untersuchungsergebnisse hinsichtlich der personen-orientierten Analyse im Unterschied zu item- und skalen-orientierten Auswertung der Datensatzes berichtet. Diese wurden exemplarisch durchgeführt an der männlichen Altersgruppe der 50-65-Jährigen. Die wichtigsten Ergebnisse:

Es konnten ohne Vorinformation über jene Männer aus der Normalbevölkerung mindestens drei Personentypen mittels Clusteranalyse entdeckt werden, deren Eigenschaftsprofile sich deutlich voneinander unterscheiden - auf der Basis von nur sieben der 12 FPI-R-Skalen:

1. Emotional Labile mit erhöhten Skalenwerten in Erregbarkeit, Beanspruchung und Gehemmtheit. Diese Gruppe zeichnet sich auch durch erhöhte Werte in der Skala Emotionalität aus, wobei diese Sekundärskala nicht zur Clusterbildung verwendet wurde. Emotional Labile sind soziodemografisch charakterisiert durch einen eher geringeren Schulabschluss, geringeres Einkommen, Sympathie für die AfD-Partei, seltenere Kontakte mit Freunden und relativ häufige psychotherapeutische Behandlung.
2. Ausgeglichenheit und hohe Lebenszufriedenheit sind typisch für die Personen des zweiten Clusters. Es handelt sich um eine kleinere Gruppe von psychisch und körperlich gesunden Männern, die eher eine höhere Bildung und höheres Einkommen haben.
3. Bei der Gruppe der extravertierten älteren Männer stehen hohe Lebenszufriedenheit und hohe Leistungsorientierung im Vordergrund. Diese Personengruppe arbeitet deutlich mehr als vorgesehen und hat auch häufigere Kontakte mit Freunden in der Freizeit. Von den Interviewern wurden Selbstsicherheit und Durchsetzungsbereitschaft im Fremdurteil notiert.

Bevor nicht weitere Analysen dieser Art unternommen sind, wäre es zu früh, die Ergebnisse in persönlichkeits-theoretischer Sicht zu diskutieren. Zu erinnern ist jedoch, dass die heutigen mehrdimensionalen Persönlichkeitsinventare nur Testwerte in einem dimensional Beschreibungssystem liefern, zwar jedes Individuum einordnen lassen, jedoch ohne psychologische Aussagen über das funktionale und dynamische Zusammenwirken dieser abstrahierten Persönlichkeitseigenschaften. Vielleicht kann die genauere Beschreibung von Personen-Clustern mithelfen, typische Verhaltensmuster in typischen Lebenssituationen zu unterscheiden.

## IX. Multitrait-Multimethod Hinweise

Multitrait und Multimethod bedeuten: es gibt viele theoretische Konstrukte (Eigenschaftskonzepte) und viele empirische Indikatoren (Methoden verschiedener Art). Ein bestimmtes theoretisches Konstrukt sollte durch eine Kombination solcher Indikatoren *konvergent*, d.h. statistisch signifikant und effektstark operationalisiert werden, wobei diese Indikatoren mit anderen theoretischen Konstrukten nur geringfügig zusammenhängen und deshalb *diskriminieren* können. Ein typischer Untersuchungsplan würde drei Konstrukte, etwa mathematische, musikalische und sportliche Begabung aufgrund von mindestens drei Indikatoren aus den Bereichen objektiver Schulleistungstests, spezieller Lehrerbeurteilungen und Selbstbeurteilungen, hinsichtlich konvergenter und diskriminanter Validität der Indikatoren prüfen. Die Ergebnisse könnten zu einer Revision der theoretischen Konstrukte, beispielsweise durch die Definition von Sub-Konstrukten, oder zu einer prägnanteren Auswahl der Indikatoren führen. (Dieser Prozedur entspricht auch die empirische Itemselektion im Hinblick auf die Bildung von Skalen.)

Im Manual wurden bereits die Absicht und das methodische Vorgehen einer MTMM-Analyse geschildert (S. 89-92, 117, 150-153). Untersuchungspläne dieser Art sind in der Persönlichkeitsforschung organisatorisch und methodisch sehr viel schwieriger als jene der Begabungs- und Schulleistungs-Forschung. Eine methodisch befriedigende Datenerfassung auf mehreren Datenebenen ist bei einer größeren Personenzahl unrealistisch bzw. unmöglich. Zumindest können jedoch explorative Analysen nach der Leitidee der MTMM-Analysen unternommen werden. Der Vorzug des FPI-Datensatzes liegt einerseits in der Stichprobengröße, andererseits im Umfang der zusätzlich erhobenen Daten, d.h. der ergänzenden Selbstbeschreibungen, zusätzlichen Selbstberichte und der speziell für einige Skalen entwickelten „verhaltensnahen“ Items. Problematisch bleibt die Ebene der Verhaltensbeurteilungen, doch bestand die seltene Gelegenheit, zumindest einige aktuelle Verhaltensbeurteilungen durch die Interviewer zu gewinnen. Wegen der speziellen und einzigartigen Beobachtungssituation und der zahlreichen und untrainierten Beurteiler sind diese Daten zweifellos sehr problematisch, aber dennoch nützlich und – wie die Analysen zeigen – wertvoll. Es mangelt natürlich an Verhaltensdaten, Selbstbeschreibungen und Fremdbeurteilungen unter Alltagsbedingungen. Auf die Forschungschancen des Ambulanten Assessments wurde bereits anhand eigener Projekte eingegangen (Manual S. 97-99, 158-161).

Vor diesem Hintergrund wurden für drei Persönlichkeitseigenschaften – FPI-R Extraversion, FPI-R Emotionalität, FPI-R Soziale Orientierung – Indikatoren zusammengestellt, die aus drei Datenquellen stammen: (1) objektivierbare Selbstauskünfte, (2) Selbstbeurteilungen aufgrund von Zusatzitems, (3) Fremdbeurteilungen und Beobachtungen durch Interviewer.

# Korrelationen von FPI-R Skalen mit ausgewählten Außenkriterien der Erhebung 2018 (N = 3450)

GAS: partielle Korrelation (Geschlecht, Alter und Schulabschluss)

INHALT (verkürzt)	KORRELATIONEN		
FPI-R Skala Soziale Orientierung	r	r (GAS)	R
Index Empathie/Gerechtigkeitsempfinden (hoch = 7, niedrig = 1)	.53	<b>.49</b>	<b>.68</b>
Spendenbereitschaft (4 = hoch, 0 = spende nicht)	.42	<b>.40</b>	
Index Autoritarismus (niedrig = 9, hoch = 1)	.42	<b>.39</b>	
Index Interesse an Sinnfragen (4 = interessiert; 1 = nicht interessiert)	.31	.28	
Item - Kontakt mit Freunden (4 = täglich, 1 = seltener)	.13	.12	
Fremdbeurteilung - freundlich	.21	.18	
FPI-R Skala Lebenszufriedenheit	r	r (GAS)	R
Index Zufriedenheit (hoch = 15, gering = 3)	.58	<b>.57</b>	<b>.62</b>
Item - Psychotherapie (nein = 2, ja = 1)	.21	.23	
Index Multiple Ängste (wenige = 7, viele = 1)	.21	.19	
Fremdbeurteilung - selbstsicher	.24	.22	
Fremdbeurteilung - durchsetzungsbereit	.20	.18	
FPI-R Skala Leistungsorientierung	r	r (GAS)	R
Item - Überstunden (3 = deutlich mehr, 1 = wie vorgesehen)	.23	.23	<b>.52</b>
Item - Leistungssport (3 = leistungsorientiert, 1 = gar kein Sport)	.21	.15	
Index Berufliche Belastung (15 = hoch, 3 = gering)	.20	.14	
Fremdbeurteilung - durchsetzungsbereit	.37	<b>.35</b>	
Fremdbeurteilung - selbstsicher	.36	<b>.32</b>	
Fremdbeurteilung - energisch	.32	<b>.30</b>	
Fremdbeurteilung - ausdrucksstark	.25	.23	
FPI-R Skala Extraversion	r	r (GAS)	R
Item - Kontakt mit Freunden (4 = täglich, 1 = seltener)	.30	.26	<b>.50</b>
Fremdbeurteilung - durchsetzungsbereit	.34	<b>.33</b>	
Fremdbeurteilung - ausdrucksstark	.33	<b>.32</b>	
Fremdbeurteilung - energisch	.31	<b>.32</b>	
Fremdbeurteilung - selbstsicher	.33	<b>.31</b>	
Fremdbeurteilung - ungezwungen	.26	.25	
Fremdbeurteilung - auffälliger persönlicher Stil	.20	.18	
FPI-R Skala Aggressivität	r	r (GAS)	R
Item - Opfer körperlicher Gewalt (4 = häufig, 1 = nie)	.39	<b>.36</b>	<b>.48</b>
AfD- vs. CDU/CSU-Sympathie (1 = AfD, 0 = CDU/CSU)	.32	.27	
Index Autoritarismus (1 = niedrig, 9 = hoch)	.21	.18	
Fremdbeurteilung - Sichtbare Tätowierung (2= ja, 1 = nein)	.23	.16	
FPI-R Skala Emotionalität	r	r (GAS)	R
Index Tablettenkonsum (15 = hoch, 3 = niedrig)	.35	<b>.34</b>	<b>.48</b>
Item - Psychotherapie (ja = 2, nein = 1)	.28	.28	
Index Multiple Ängste (viele = 7, wenige = 1)	.27	.25	
Fremdbeurteilung - unsicher	.21	.17	
FPI-R Skala Beanspruchung	r	r (GAS)	R
Item - Überstunden (3 = deutlich mehr, 1 = wie vorgesehen)	.24	.28	<b>.41</b>
Index Berufliche Belastung (15 = hoch, 3 = gering)	.24	.21	
Item - Zeit für Hobbies (4 = wenig Zeit, 1 = viel Zeit)	.17	.16	

Anmerkung: Cohen (1988) gibt für r die folgenden Intervalle an: .1 bis .3: kleiner Effekt; .3 bis .5: mittlerer Effekt; .5 und höher: starker Effekt; gelbe Schattierung: r > 0.30. SB: Selbstbeschreibung FPI-R; FB: Fremdbeurteilung durch Interviewer.

Interkorrelationen von ausgewählten FPI-R Skalen und Fremdbeurteilung (Eigenschafts-Ratings durch IfD – Interviewer)

FPI und Fremdbeurteilung									Interkorrelationen - Fremdbeurteilung						
		LEB	SOZ	LEI	AGGR	BEAN	EXTR	EMOT	selbst-sicher	energisch	freundlich	unge-zwungen	durchset-zungsbereit	auffällig	ausdrucks-stark
Selbstbeschreibung	LEB														
	SOZ	.16		▽			▽								
	LEI	.36	.12												
	AGGR	-.23	-.26	.24											
	BEAN	-.36	.01	.07	.23										
	EXTR	.28	.08	.63	.36	.00									
	EMOT	-.69	-.06	-.20	.30	.64	-.18								
Fremdbeurteilung	selbstsicher	.25	.04	.36	.06	-.05	.32	-.21							
	energisch	.15	.02	.32	.14	-.02	.31	-.10	.51						
	freundlich	.14	.22	.04	-.14	-.04	.03	-.09	.16	.04					
	ungezwungen	.23	.12	.23	-.01	-.06	.27	-.17	.47	.30	.41				
	durchsetzungsbereit	.21	.02	.38	.09	-.02	.34	-.15	.55	.51	.12	.41			
	auffälliger pers. Stil	.03	.02	.14	.08	.06	.20	.02	.16	.17	.02	.15	.16		
	ausdrucksstark	.15	.12	.26	.05	.02	.34	-.09	.40	.36	.22	.42	.44	.25	

Konvergente und diskriminante Validität MTMM - Korrelationen			Methode 1			Methode 2			Methode 3		
			FPI Selbstbeschreibung			verhaltensnahe Auskünfte			Verhaltensbeobachtung		
			E 1	N 1	S 1	E 2	N 2	S 2	E 3	N 3	S 3
			EXTR	EMOT	SOZ	Kontakt Freunde	Tabletten-einnahme	Spenden in €	ausdrucksstark	unsicher	freundlich
Methode 1 Selbstbeschreibung FPI	E 1	Extraversion (FPI)									
	N 1	Emotionalität (FPI)									
	S 1	Soziale Orientierung (FPI)									
Methode 2 verhaltensnahe Auskünfte (objektivierbar)	E 2	Kontakt Freunde seltener (1) fast täglich (4)									
	N 2	Tabletteneinnahme gering (3) hoch (9)									
	S 2	Spenden in € (0 - 1000+)									
Methode 3 Verhaltensbeobachtung	E 3	ausdruckslos (1) - lebhaft, ausdrucksstark (4)									
	N 3	sehr selbstsicher, von sich überzeugt (1) - ziemlich unsicher (4)									
	S 3	eher unfreundlich (1) - sehr freundlich (4)									

Hauptkomponenten-Analyse - Varimax-Rotierte Komponentenmatrix			
	1	2	3
FPI-R EXTRAVERSION	.72		.12
Kontakt mit Freunden in der Freizeit	.46		
Ausdrucksverhalten des Befragten - ausdrucksstark	.76	.12	
FPI-R EMOTIONALITÄT	-.11	-.12	-.81
Tabletteneinnahme			-.81
Befragter während des Interviews – ziemlich unsicher	-.66		.22
FPI-R SOZIALE ORIENTIERUNG		.84	
Spenden in Euro		.79	
Befragter - freundlich	.29	.38	

- Korrelations-Koeffizienten in Rot: Monotrait-Heteromethod-Korrelation = konvergente Validität (ein Merkmal, z.B. Extraversion, erfasst mit 3 Methoden).
- grüne Schattierung: Heterotrait-Heteromethod-Korrelationen = diskriminante Validität (z.B. EMOT mit Methode 1 und EXTR mit Methode 2).
- gelbe Schattierung: Heterotrait-Monomethod-Korrelation = diskriminante Validität (Merkmale EXTR, EMOT, SOZ mit einer Methode erfasst).

### Interpretation der MTMM-Matrix (EXTR, EMOT, SOZ):

Nachweis der *konvergenten* Validität:

Die *konvergente* Validität, gemessen mit den drei Methoden (Selbstbeschreibung mittels FPI-R, verhaltensnahe Selbstauskunft durch Befragten und Verhaltensbeobachtung durch einen lfd-Interviewer), erreicht für die FPI-Konstrukte Extraversion, Emotionalität und Soziale Orientierung angenähert mittlere Effektstärken zwischen  $r \geq .20$  und  $r \leq .35$ . Acht der 9 Validitätskoeffizienten (rot) sind statistisch signifikant und von null verschieden, aber numerisch gering (z.B.  $S2/S3 = .06$ ).

Nachweis der *diskriminanten* Validität:

Die *diskriminante* Validität ist insgesamt noch geringer:

(1) Die Korrelationen der Heterotrait-Heteromethod-Blöcke sind lediglich bei 28 von allen 36 Vergleichen (12 pro Block) niedriger als die Validitätskoeffizienten. Das erste Kriterium nach Campbell und Fiske (1959) ist damit nicht voll erfüllt.

(2) Die Heterotrait-Monomethod-Korrelationen (gelbe Dreiecke) sind zwar geringer als die Monotrait-Heteromethod-Korrelationen (rot), doch nur für die Methoden 1 und 2 (FPI und verhaltensnahe Auskünfte). Die E-, N-, S-Koeffizienten der Methode 3 ( $r=-.40$ ,  $r=-.22$ ,  $r=-.16$ ) sind höher als die 9 Validitätskoeffizienten und erfüllen damit nur teilweise das zweite Kriterium der diskriminanten Validität.

(3) Die Muster der 3 Merkmalskorrelationen E, N, S sind innerhalb und zwischen den Methoden verschieden, insbesondere bei Methode 3 (Verhaltensbeobachtung). Damit ist auch das dritte Kriterium der diskriminanten Validität nur teilweise erfüllt.

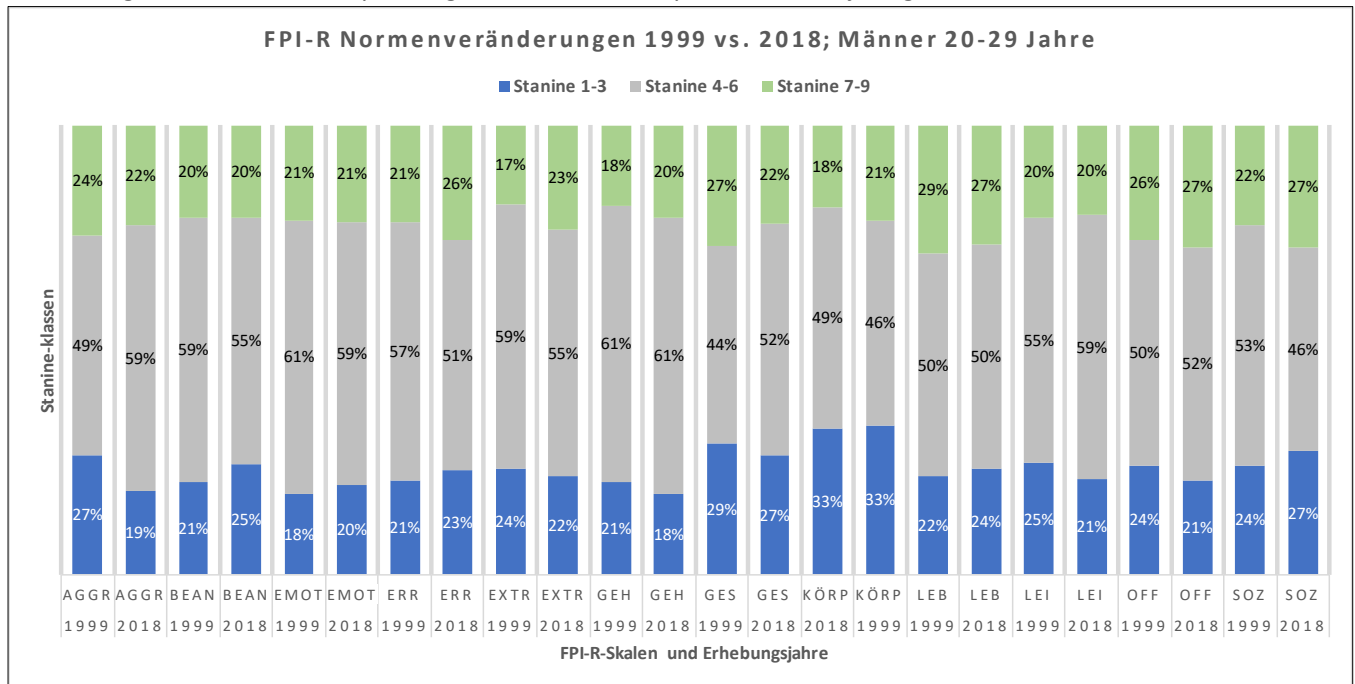
Während die Selbstbeurteilung mittels FPI und die Selbstauskünfte über objektivierbare Indikatoren relativ gut übereinstimmen, ist die Verhaltensbeobachtung lediglich in Hinsicht auf das Konstrukt Extraversion mehr oder weniger konvergent. Die diskriminante Validität des FPI-R, hier nur hinsichtlich der Eigenschaften EXTR, EMOT und SOZ untersucht, erscheint hinsichtlich Methode 1 und 2 ausreichend, fehlt aber hinsichtlich Methode 3 (Verhaltensbeobachtung).

Als mögliche Gründe dieser unbefriedigenden explorativen Ergebnisse sind u.a. die methodische Unsicherheit der Verhaltensbeurteilungen und die hochaggregierten (facettenreichen) Konstrukte *Emotionale Labilität* und *Soziale Orientierung* im Vergleich zu manifesten Zeichen der Extraversion zu nennen.

Im Zusammenhang mit der Kontrolle der Skalenkonstruktion und der Normierung des FPI wurde wenigstens diese explorative Analyse unternommen und in die Dokumentation aufgenommen. Der Forschungsbedarf und dessen Schwierigkeiten sind offenkundig. Auch in der Persönlichkeitsforschung sind MTMM-Analysen unverzichtbar: Werden die interessierenden Eigenschaften auf den wichtigen Datenebenen (Selbstbeurteilung, Fremdbeurteilung, Verhaltensbeobachtung) so prägnant, konvergent und diskriminant durch einen Persönlichkeitsfragebogen erfasst, dass ein spezieller Index (beispielsweise FPI-Testwert der Skala Extraversion) allein als Indikator des theoretischen Konstrukts E gelten kann? Andernfalls sollte der Index E nur mit dem Suffix E (FPI), E (MPI), E (HEXACO) usw. verwendet werden. Zwar ist es üblich, dass Untersucher die verwendeten Tests nennen, aber in Diskussion und Anwendung eventueller Befunde erfolgt regelmäßig die – empirisch nicht belegte – Verallgemeinerung auf „die Extraversion“ schlechthin.

## IX. Normenveränderung seit der Erhebung 1999

Notwendigkeit der Normenanpassung des FPI-R am Beispiel der 20-29-jährigen Männer und Frauen



### FPI-R - Normenveränderung 1999/2018

Männer 20-29 Jahre; N = 215		
Skala	korrekt klassifiziert VP+VN	Falsch klassifiziert FP+FN
Gehemmtheit	100%	0%
Beanspruchung	100%	0%
Körperliche Beschwerden	100%	0%
Gesundheitssorgen	100%	0%
Emotionalität	96%	4%
Lebenszufriedenheit	94%	6%
Soziale Orientierung	90%	10%
Erregbarkeit	90%	10%
Offenheit	90%	10%
Extraversion	81%	19%
Leistungsorientierung	79%	21%
Aggressivität	60%	40%
	VP = Valide Positive VN = Valide Negative	FP = Falsche Positive FN = Falsche Negative

Frauen 20-29 Jahre; N = 195		
Skala	korrekt klassifiziert VP+VN	falsch klassifiziert FP+FN
Gehemmtheit	100%	0%
Körperliche Beschwerden	100%	0%
Leistungsorientierung	100%	0%
Extraversion	96%	4%
Beanspruchung	93%	7%
Erregbarkeit	92%	8%
Offenheit	92%	8%
Gesundheitssorgen	90%	10%
Emotionalität	88%	12%
Soziale Orientierung	87%	13%
Lebenszufriedenheit	78%	22%
Aggressivität	74%	26%
	VP = Valide Positive VN = Valide Negative	FP = Falsche Positive FN = Falsche Negative

VP = Valide Positive: Auffällige = ST 1-3 plus ST 7-9 – Übereinstimmung 2018 und 1999

VN = Valide Negative: Normbereich ST 4-6 – Übereinstimmung 2018 und 1999

FP = Falsche Positive: ST 4-6 2018, Abweichung zu 1999

FN = Falsche Negative: ST 4-6 1999, Abweichung zu 2018

## XII. Zur Qualitätssicherung von Persönlichkeitsfragebogen

### Vorbemerkung:

Der folgende wissenschaftliche Text wurde nach Abschluss der 9. Auflage des Freiburger Persönlichkeitsinventars (2020) bei einer deutschen Fachzeitschrift für Psychologie zur Veröffentlichung eingereicht, aber abgelehnt. Die FPI-Autoren haben den Text mit folgender Intention verfasst:

- (1) Zur wichtigen Frage der Qualitätskontrolle von Tests gibt es eine pauschale Forderung im Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologenvereinigungen, die ursprünglich aus einer Daumenregel stammt, aber bisher, zumindest für deutschsprachige Persönlichkeitsfragebogen, keine *empirische* Basis hatte: Neue Normierung nach 8 bis 10 Jahren!
- (2) Nur für das Freiburger Persönlichkeitsinventar FPI existieren zwei bzw. jetzt (2020) drei Normierungsstudien auf „bevölkerungsrepräsentativer“ Basis, jeweils verbunden mit replizierter Skalenkonstruktion und Aktualisierung der Validitätsbelege.
- (3) Das FPI ist auch nach der dritten Umfrage bei Mitgliedern des Berufsverbandes Deutscher Psychologinnen und Psychologen (Roth, Schmitt & Herzberg, 2010) der in der Praxis am weitesten verbreitete Persönlichkeitsfragebogen.

**Zusammenfassung:** Für psychologische Tests und deren Anwendung wurden Richtlinien der Qualitätssicherung entwickelt, insbesondere in dem *Testbeurteilungssystem TBS-DTK des Diagnostik- und Testkuratorium* (2018) und in den *Guidelines der International Test Commission* (2005, 2013, 2015, 2017). Im weiteren Sinn umfasst Qualitätssicherung den gesamten Prozess eines kontinuierlichen Qualitätsmanagements, an dem die Testautoren und der Verlag sowie die Anwender und die Rezensenten Anteil haben. Jede neue Auflage eines Tests bietet die Gelegenheit, qualitätsverbessernde Befunde und Überlegungen aufzunehmen.

Einige dieser Aspekte werden am Beispiel der dritten Normierungsstudie zum *Freiburger Persönlichkeitsinventar FPI* diskutiert. Nach den bevölkerungsrepräsentativen Normierungen in den Jahren 1982 (nur Westdeutschland) und 1999, die keine erheblichen Abweichungen ergaben, war es zur aktuellen Qualitätskontrolle angebracht, die Normen und die Skalenkonstruktion erneut zu überprüfen. Tatsächlich sind die Normwerte der Altersgruppen 16-19 und 20-29 anzupassen. Die Unterschiede sind hauptsächlich in den Bereichen Leistungsorientierung, Aggressivität, Extraversion und Emotionalität zu erkennen. Weitere Aspekte der Qualitätssicherung werden nur kurz diskutiert. Über die testkonstruktiven Analysen und die Integration von neueren Validitätshinweisen wird im Testmanual der 9. Auflage (Fahrenberg, Hampel & Selg, 2020, im Druck) ausführlich berichtet – auch in der Absicht, die geforderte kritisch-methodenbewusste Anwendung von Persönlichkeitsfragebogen zu unterstützen. Der zeitliche und finanzielle Aufwand für Skalenkonstruktion, Normierung und Validierung hinsichtlich externer Kriterien ist so hoch, dass kooperative Projekte zur Qualitätskontrolle – und künftig bereits zur Entwicklung – mehrdimensionaler Persönlichkeitsfragebogen zu erwägen sind.

### 1. Aspekte der Qualitätssicherung

„Qualitätssicherung verfolgt das Ziel durch kontinuierliche Beschäftigung mit Aspekten der Struktur- und Prozessqualität die Erreichung von Qualitätsanforderungen (Qualität) sicherzustellen. ... Charakteristisch ist eine Orientierung am PDCA-Zyklus [Planen-Ausführen-Prüfen-Handeln-Zyklus], der die systematische Erfassung von Qualitätsindikatoren verlangt, und nach Identifikation von Defiziten die Einleitung potenziell qualitätsverbessernder Maßnahmen einfordert, deren Effektivität möglichst unmittelbar kritisch im Hinblick auf die Zielsetzungen geprüft werden muss.“ (Wirtz, 2019, S. 1465). Die Aufgabe der Qualitätssicherung stellt sich in sehr vielen Bereichen, und ein Seitenblick beispielsweise auf die Entwicklung von Software-systemen zeigt, welcher Aufwand erforderlich ist. Die gesamten Kosten verteilen sich nach einer Schätzung zu 30 Prozent auf die Entstehungs- und Entwicklungsphase, zu 40 Prozent auf die Evolutionsphase (Korrekturen, Änderungen, Optimierungen, Erweiterungen), zu 25 Prozent auf die Erhaltungsphase und zu 5 Prozent auf die Ablösungsphase bzw. Ausmusterung; die Erhaltungsphase beansprucht etwa den doppelten Zeitaufwand der Evolutionsphase (Sneed, Baumgartner & Seidl, 2011, S. 199).



Im Hinblick auf einen psychologischen Test interessiert die systematische Verbesserung der Qualität aufgrund der gewonnenen Erfahrungen. Für dieses Qualitätsmanagement sind außer den Testautoren und dem Verlag auch die Anwender und die Test-Rezensenten wichtig, auch wenn sie keinen organisierten Qualitätszirkel bilden. – Gibt es von einem eingeführten Test überhaupt eine ergänzte, überarbeitete oder sogar revidierte Neuauflage und nicht bloß einen unveränderten Nachdruck?

Hauptsächliche Themen wären:

- Überprüfung der Normen;
- Replizierbarkeit der Skalen und andere testkonstruktive Aspekte;
- weiterführende Diskussion der zugrundeliegenden psychologischen Konstrukte;
- Integration von Validitätshinweisen aufgrund eigener Untersuchungen der Testautoren und aufgrund der Fachliteratur;
- Testrezensionen und Reviews ggf. mit einem kritischen Vergleich ähnlicher Tests;
- Stellungnahme der Testautoren zu Testrezensionen und zu ähnlichen Test-Publikationen;
- Engagement des Testverlags hinsichtlich Neuauflagen und Unterstützung bevölkerungsrepräsentativer Normierungen;
- Verfügbarkeit der Datensätze für unabhängige Reanalysen;
- Umfang und Zugänglichkeit der Dokumentation.

Eine Umfrage unter Mitgliedern des Berufsverbandes BDP hinsichtlich psychologischer Diagnostik in der Praxis (Roth, Schmitt & Herzberg, 2010) ergab außer einer Liste der am häufigsten verwendeten Testverfahren eine Reihe von Verbesserungsvorschlägen für das Studium. An erster Stelle wurde gefordert, dass „in der universitären Diagnostikausbildung eine kritische Diagnostik gelehrt wird. So soll auf die Grenzen der Testverfahren hingewiesen werden, und es soll für Fehlerquellen beim Testen eine Sensibilität erzeugt werden.“ (S. 126). Diese kritisch-methodenbewusste Einstellung kann durch die Weiterentwicklung eines Testmanuals unterstützt werden. Gerade bei den so einfach erscheinenden Persönlichkeitsfragebogen sind die Prinzipien der Testkonstruktion und der empirischen Validierung wichtig, um kritisch auswählen und interpretieren zu können. Persönlichkeitsfragebogen erfassen Selbstbeurteilungen und Selbstbeobachtungen und sie sollen facettenreiche Persönlichkeitseigenschaften repräsentieren. Folglich können die psychometrischen Postulate (Messmodelle) und die Konstruktionsweisen von objektiven Intelligenz- und Leistungstests nicht einfach auf Persönlichkeitsfragebogen übertragen werden, und die Nachweise externer Kriterienvalidität sind hier noch wichtiger.

Das Freiburger Persönlichkeitsinventar FPI wurde im Jahr 1970 publiziert, in der 4. Auflage (Fahrenberg, Hampel & Selg, 1984) durch weitere Skalen ergänzt und wiederholt hinsichtlich der Normierung und Replizierbarkeit der Skalenkonstruktion überprüft (siehe auch Bengel & Wittmann, 1983; Fahrenberg, Selg & Hampel, 1983). Nach 50 Jahren, anlässlich der 9. Auflage, wird hier aufgrund der aktuellen dritten Repräsentativerhebung über die Gesichtspunkte und Erfahrungen dieser „Testpflege“ berichtet.

## Richtlinien

Vorauszuschicken ist die Frage, inwieweit das *Testbeurteilungssystem TBS-DTK des Diagnostik- und Testkuratorium* (2018) oder die *Guidelines der International Test Commission* (ITC) bereits Grundsätze oder Beurteilungshinweise für diesen *Prozess der Qualitätsverbesserung* enthalten. Das TBS-DTK Testbeurteilungssystem fasst zahlreiche Bewertungskriterien zum Zweck von Testrezensionen zusammen. Das zugehörigen *DIN Screen* ist ein Bewertungsbogen, in dem diese Bewertungskriterien formalisiert sind. Sie wurden ursprünglich für berufsbezogene Eignungsbeurteilungen nach DIN 33430 entwickelt (Kersting, 2018a,b) und werden in Deutschland oft für andere Gebiete der psychologischen Diagnostik und Beurteilung übernommen. – Für Persönlichkeitsfragebogen und Klinische Skalen, so ist hier einzuschränken, können diese primär an Berufseignung und Berufsanforderungsanalysen orientierten Beurteilungen kaum ausreichen. So werden das für Persönlichkeitsfragebogen wesentliche Qualitätsmerkmal einer *bevölkerungsrepräsentativen* Normierung und die typischen Probleme der Skalenkonstruktion nur allgemein erwähnt.

Im TBS-TK (2018, S. 18-19) werden zahlreiche „Qualitätssichernde und qualitätsoptimierende Maßnahmen“ und Aspekte der „Qualitätsforderungen“ aufgezählt. Die differenzierte Bewertung hinsichtlich dieser qualitätsverbessernden Maßnahmen fehlt jedoch im TBS-TK weitgehend. Beispielsweise gilt die generelle Forderung, dass für die praktische Anwendung eines psychologischen Tests dessen Normierung eine zentrale Voraussetzung ist. Im *DIN Screen* (Kersting, 2018b) ist zu beurteilen „Die Angemessenheit der Normwerte wurde in den letzten acht Jahren überprüft“ (S. 239). Es handelt sich um eine Soll- und nicht um eine Muss-Forderung, und es folgt der Hinweis: „Es geht nur um eine Überprüfung der Angemessenheit

der Normwerte. Ob eine Neunormierung durchgeführt werden muss, ergibt sich in Abhängigkeit von den Ergebnissen der Überprüfung. In der DIN 33430 wird nicht gefordert, dass spätestens alle acht Jahre neu normiert werden muss.“ Das folgende Beurteilungskriterium im *DIN Screen* lautet jedoch als Muss-Bestimmung: „In den Verfahrenshinweisen wird begründet, warum und unter welchen Umständen das Verfahren für einen Anwendungsfall ausgewählt werden kann, obwohl die Angemessenheit der Normwerte nicht in den letzten acht Jahren überprüft wurde.“

Demgegenüber hat die *International Test Commission* ausführliche *Guidelines for Practitioner Use of Test Revision, Obsolete Test, and Test Disposal* (2015) entwickelt, um die Beurteilung von revidierten Tests, aber auch die Gründe für den Verzicht auf überholte Tests zu vereinheitlichen: „Test Developers and Users Have a Reciprocal Relationship.“ (p. 8). „Test revisions may be driven by knowledge that the assessed behaviours are subject to substantial change over time, by significant demographic changes, from research that leads to improvements in theories and concepts that should impact test use, from changes in diagnostic criteria, or in response to test consumer’s demands for improved versions. The presence of these qualities warrants consideration of whether a revised test may be more suitable than an older version.“ (p. 9). „Revisions of tests that provide norm-referenced interpretations typically focus on updating norms. Thus, decisions regarding the purchase and use of a revised test may be more favourable when the norms for the revised test were obtained recently, are sufficient, and better represent the population.“ (p. 11). Diese Guidelines enthalten keine Daumenregel für den zweckmäßigen Abstand von Testrevisionen. – „Practitioners should have a broad knowledge of quality control practices, as these are critical to the accurate use of tests.“ (ITC Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores, 2013, p. 6).

Die mehr oder minder differenzierten Richtlinien dieser Kommissionen werden die fachliche Diskussion wesentlich anregen können. Jedoch fehlen, beispielsweise zu den Anforderungen an bevölkerungsrepräsentative Normierungen, Hinweise auf empirische Untersuchungsergebnisse oder Vorbilder, so dass keine Maßstäbe oder Bedingungen solcher Maßstäbe vermittelt werden. An der Grundforderung nach regelmäßiger Überprüfung der Normen lässt sich zeigen, wie schwierig es bleibt, empirisch fundierte Anforderungen zu formulieren. Aus einer verbreiteten Daumenregel „8-10 Jahre“ wird bei Kersting (2018b) die Forderung nach einer Überprüfung im Intervall von 8 Jahren. Geeignete empirische Untersuchungen, um diese Forderung zu begründen, werden nicht zitiert. Regeln in dieser Allgemeinheit, ohne Rücksicht auf die untersuchten psychologischen Merkmale und ohne Konvention zur Evaluation von Effektstärken, kann es auch nicht geben.

Bemerkenswert sind die *ITC Guidelines for Translating and Adapting Tests* (2017), in denen die Methodenprobleme, u.a. der wesentliche Unterschied zwischen der sprachlichen und der psychologischen Äquivalenz der Übersetzung von Items, gründlich dargelegt sind. Das TBS-TK geht auf solche Fragen nicht näher ein, obwohl es sich bei vielen der in Deutschland publizierten Tests um einfache Übersetzungen angloamerikanischer Vorbilder handelt. (In den meisten der ITC-Kommissionen scheint kein Experte aus den deutschsprachigen Ländern mitgearbeitet zu haben.)

## **Zur Normierung von Persönlichkeitsfragebogen**

Der vielfach beschriebene gesellschaftliche Wandel in Deutschland während der letzten Dekaden, eventuell zurückzuführen auf demografische Veränderungen, technologischen Fortschritt, Massenmedien und Globalisierung, könnte sich auch in veränderten Selbstbildern, Verhaltensweisen und Lebensstilen zeigen. In der Sozialforschung wurde u.a. auf eine Auflösung geschlechtsspezifischer Rollenbilder und Karrierepläne hingewiesen und erläutert, dass ältere Menschen heute aktiver seien als vor 20 Jahren, häufiger erwerbstätig, häufiger ehrenamtlich engagiert, mehr Sport treiben und mehrheitlich sozial gut eingebunden sind. Im Unterschied zu Intelligenz- und Leistungstests sind Persönlichkeitsfragebogen, klinische Skalen und Einstellungsskalen stärker von Sprachgewohnheiten abhängig, so dass einzelne Wörter und ganze Items, auch die Rechtschreibung, bereits nach zwei oder drei Jahrzehnten „altmodisch“ wirken können. Auch hinsichtlich der Offenheit der Antworten oder der sozialen Erwünschtheit könnte es solche Trends geben.

Bei einem typischen Persönlichkeitsfragebogen ist eine bevölkerungsrepräsentative Normierung und bei einer klinischen Skala die Normierung für eine bestimmte Patientenpopulation zu erwarten. Ein Vergleich der häufiger verwendeten (deutschen) Persönlichkeitsinventare zeigt, dass „repräsentativ“ unterschiedlich verstanden wird. Die bevölkerungsrepräsentative Normierung im engeren Sinn sollte aufgrund einer einheitlichen Erhebung durch ein auf Umfragen spezialisiertes Institut vorgenommen werden. Dieses Verfahren ist an dem Quotenplan, bezogen auf die offizielle Bevölkerungsstatistik, und an anderen Informationen zu erkennen. Je nachdem wie detailliert bestimmte Quotierungsmerkmale berücksichtigt sind und welche Zellenbesetzungen bzw. Schätzfehler vorgesehen sind, steigen der erforderliche Stichprobenumfang und die Kosten an. Für die Normierungen des FPI wurde die Größenordnung von 3000 Personen als hinreichend angesehen, um die Normen nach Geschlecht

und sieben Altersgruppen differenzieren zu können. Der Datensatz wurde in bewährter Zusammenarbeit mit dem Institut für Demoskopie IfD, Allensbach, in drei Wellen erhoben. Jeweils waren ca. 300 routinierte freie Mitarbeiter des *Instituts für Demoskopie Allensbach* beteiligt, um den Fragebogen direkt an die nach Quotenplan zu suchenden Teilnehmer auszugeben und das ergänzende mündliche Interview zu führen.

Daneben gibt es weit verbreitete Persönlichkeitsfragebogen, deren Normierungsbasis auf andere Weise gewonnen wurde. Es werden diverse Datensätze aus verschiedenen Erhebungen und laufenden Projekten als sogenannte „Eichstichprobe“ zusammengesetzt, um eine möglichst breite und eventuell „repräsentative“ Grundlage anzunähern. Es sind „zuhandene“ Daten bzw. „Gelegenheitsstichproben“, eventuell sogar durch telefonische Umfragen oder im Internet gewonnen. Über den methodischen Bias und dessen Effektstärken hinsichtlich der wichtigsten Personenmerkmale kann bei diesem Verfahren ohne eine tatsächliche Repräsentativerhebung nichts ausgesagt werden.

Auch die direkten Erhebungen durch routinierte Interviewer haben noch einige typische Methodenprobleme, doch sind sie zur Qualitätssicherung zweifellos angemessener als die Kombination diverser zuhandener Datensätze. Dass die Konstruktion und die Normierung von Persönlichkeitsfragebogen bevölkerungsrepräsentativ anzulegen sind, muss gerade angesichts der zunehmenden Publikation von Internet-basierten und deshalb hoch-selektiven Erhebungen betont werden (vgl. ITC-Guidelines 2005). Auch für die Publikation einzelner Forschungsvorhaben aufgrund solcher Daten muss auf den völlig unbekannten und unkontrollierbaren groben Selektions-Bias hinsichtlich der Kontaktaufnahme, der Bereitschaft und der technischen Voraussetzungen zur Teilnahme hingewiesen werden. Wenn außerdem in solchen Untersuchungen nur die Signifikanzen statt der Effektstärken von Gruppenunterschieden interpretiert werden (siehe u.a. Obschonka, Wyrwich, Fritsch, Gosling, Rentfrow & Potter, 2019) entstehen höchst fragwürdige Publikationen.

## **2. Zur Qualitätssicherung des Freiburger Persönlichkeitsinventars**

### **2.1 Normierung**

Trotz der nur geringen Differenzen der mittleren Testwerte zwischen den Erhebungen von 1982 und 1999, damals nur im Vergleich der westdeutschen Bevölkerung möglich, war für die 9. Auflage nach Auffassung der Autoren wie auch des Verlags eine Überprüfung erforderlich, um eventuell die Normen anzupassen. Für die aktuelle Normierungsstudie wurde außerdem, auch auf Anregung des IfD und des Verlags, die Formulierung von neun der 138 Items sprachlich leicht verändert bzw. modernisiert. Beispiele sind: Nr. 15: Ich kann mich erinnern, mal so wütend (statt: so zornig) gewesen zu sein, dass ich das nächstbeste Ding nahm und es zerriss oder zerschlug; Nr. 97: Ich werde leicht rot (statt: Ich erröte leicht). Ein möglicher differenzieller Effekt dieser Anpassung lässt sich ohne Kontrollgruppe nicht angeben. Die Antworthäufigkeiten bei diesen modifizierten Items differieren zwischen den Erhebungen 1999 und 2018 zwischen 3 und maximal 8 Prozent.

Die Daten der neuen bevölkerungsrepräsentativen Stichprobe würden die Aktualisierung der Normen-Tabellen des FPI-R für die heutige Bevölkerung ermöglichen. Diesem für die praktische Anwendung wichtigen Schritt gehen statistische Analysen voraus, ob die Veränderungen so deutlich sind, dass überhaupt eine Anpassung notwendig ist. Kriterien sind die Effektstärke ES und in praktischer Hinsicht ein anderer Normwert (Skalenstufe) für einen bestimmten Test-Rohwert. Es genügt jedoch nicht, die erhobenen Daten direkt zu vergleichen, denn der Bevölkerungsaufbau hat sich hinsichtlich Lebensalter und Bildungsabschluss kontinuierlich verändert. Im Prinzip könnten nur *repräsentative* Kohorten- bzw. Langzeitstudien eine zuverlässige Grundlage liefern.

Die Kohortenstudie des *Sozio-oekonomische Panel* (SOEP) konnte in den Jahren 2003, 2009 und 2013 mit der Kurzform eines Big Five Inventory verknüpft werden (Wagner, Lüdtke & Robitzsch, 2019). Die spezielle Auswahl und der spezielle Antwortmodus der Items lassen jedoch keine Verallgemeinerung auf andere Versionen dieser Fünf-Faktoren-Inventare zu. Der Effekt sozioökonomischer Merkmale wurde nicht untersucht, außerdem sind die Intervalle der Erhebung gering, so dass Schlussfolgerungen für die in der Praxis verwendeten Persönlichkeitsfragebogen kaum angebracht sind.

Zur *Schätzung* der Effekte mittelfristiger Veränderungen werden hier zwei statistische Verfahren verwendet (siehe Tabelle 1 und 2):

(1) die *Bildung statistischer Zwillinge* (matched pairs). Dabei standen für das Jahr 1999  $N = 3\,660$  und für 2018  $N = 3\,450$  Personen zur Verfügung. Die Suche nach statistischen Zwillingen ergab  $N = 2\,304$  Paare, die in Hinsicht auf Geschlecht, Alter in Jahren (7 Kategorien), Schulabschluss (3 Kategorien) und Bundesland (Ost/West) identisch sind. Dieser soziographisch kontrollierte Mittelwertvergleich ist zur Schätzung mittelfristiger Trends zweckmäßig, wenngleich der matched pairs-Datensatz nicht mehr bevölkerungsrepräsentativ ist.

(2) die *Konstruktion eines Längsschnitts (Pseudo-Kohorte)*. Da auch die Veränderungen im FPI-Profil über den längeren Zeitraum von 1982, 1999 und 2018 interessierten, aber dieselben Personen nicht erneut befragt werden konnten, wurden ersatzweise Analysen mit konstruierten Alterskohorten durchgeführt. Beispielsweise wurde für den 20-jährigen männlichen Jugendlichen mit Hauptschulabschluss aus der Erhebung 1982 ein 37-jähriger Mann mit Hauptschulabschluss aus der Erhebung 1999 als statistischer Zwilling gesucht. Es verblieb schließlich eine Analyse-Stichprobe von N = 966 Paaren. Diese Stichprobe wurde dann in vier Altersgruppen unterteilt. Dieselbe Auswahlprozedur von Personen erfolgte für die Datensätze 1999 und 2018, wobei aber Ost- und Westdeutsche gemäß ihrem Anteil an der Gesamtbevölkerung berücksichtigt werden konnten. Diese Analysestichprobe besteht aus N = 1 113 Paaren.

## Ergebnisse

Die Tabelle 1 zeigt, dass die *Unterschiede zwischen den FPI-Profilen 1999 und 2018* insgesamt eher gering sind, wenn Effektstärken statt t-Test-Signifikanzen zugrunde gelegt werden. Die höchste Effektstärke ergibt sich für die Skala *Aggressivität*: der Mittelwert hat bei 16-29-Jährigen abgenommen ( $d = .34$ ), etwa in der Größenordnung einer Skalenstufe. In diesem Kontext ist daran zu erinnern, dass es sich nur um mittelfristige Veränderungen der Selbstbeurteilungen handelt und nicht um manifestes Verhalten. Deshalb ist kein direkter Bezug zur Gewaltkriminalität herzustellen. Die Studien von Pfeifer, Baier und Kliem (2018, S. 5) haben ergeben, dass – entgegen landläufiger Meinung und medialer Berichterstattung über schwerste Gewalttaten – die Gewaltkriminalität nicht zugenommen, sondern abgenommen hat. – Auch die Testwerte für *Körperliche Beschwerden* und *Gesundheitssorgen* bei Älteren sind niedriger als im Jahr 1999 ( $d = .25$  und  $d = .29$ ). Dieser Trend ist in Anbetracht einer verbesserten Gesundheitsvorsorge und veränderter Lebensweise nachvollziehbar (Gesundheitsbericht des Robert Koch Instituts, 2015; S. 31). Die große Mehrzahl der berechneten Effektstärken ist unbedeutend.

Tabelle 1: Effektstärken in den FPI-Skalenwerten 1999 vs. 2018 nach Alter, Schulabschluss und West-/Ost-Deutschland

Einflussfaktoren	1999 vs. 2018			1999 vs. 2018			1999 vs. 2018	
	Alter 16-29	Alter 30-64	Alter 65+	HS	RS	ABI+	West	Ost
Effektstärke	ES	ES	ES	ES	ES	ES	ES	ES
Lebenszufriedenheit	-.01	.04	.03	.12	-.01	.00	.06	-.04
Soziale Orientierung	-.03	.12	.20	.10	.09	.11	.08	.15
Leistungsorientierung	.15	.11	-.03	-.03	.12	.14	.07	.13
Gehemmtheit	-.08	-.02	.20	.16	.04	-.13	.00	.05
Erregbarkeit	.08	.02	-.01	-.12	.12	.03	.00	.09
Aggressivität	.34	.05	.10	.00	.18	.15	.10	.18
Beanspruchung	.03	-.10	-.02	-.10	-.01	-.10	-.02	-.15
Körperliche Beschwerden	-.04	.02	.25	.12	.09	-.04	.01	.16
Gesundheitssorgen	-.07	.02	.29	.19	.06	-.05	.04	.11
Offenheit	.10	-.06	-.06	-.19	.00	.06	.00	-.10
Extraversion	.21	.08	-.18	-.12	.11	.11	.08	-.01
Emotionalität	.14	.01	.06	-.03	.13	.01	.05	.04
Stichprobengröße (N)	471	1.366	467	601	945	758	1.606	698

Anmerkungen: N = 2.304 statistische Zwillinge (matched pairs). HS = Hauptschule, RS = Realschule, ABI+ =, Abitur/Studium. Polung der ES positiv = Mittelwert im Jahr 1999 höher. ES ist die aus der Mittelwertdifferenz und  $s_M$  berechnete Effektstärke. Das Vorzeichen bei ES gibt die Richtung des Effekts an. Ein Minus-Zeichen bedeutet: höherer Mittelwert 2018 im Vergleich zu 1999.

Die Tabelle 2 enthält Ergebnisse des konstruierten Längsschnitts, der wegen dieser Konstruktion und mit nur zwei Zeitpunkten eine begrenzte Aussagekraft hat. – Die Mittelwertvergleiche der FPI-Skalen (1982/1999 und 1999/2018) ergaben zwar zahlreiche signifikante Differenzen, doch nur die Alterskohorte der 50-59-Jährigen wies im Vergleich mit den um 17 und 19 Jahre älteren Personen nennenswerte Veränderungen in drei FPI-Skalen auf: eine geringer ausgeprägte Leistungsorientierung, Beanspruchung und Gesundheitssorgen (Effektstärken mit Beträgen von  $|.32|$  bis  $|.73|$ ). Personen in der zweiten Lebenshälfte beschreiben sich mit wachsendem Alter - nicht überraschend - als weniger beansprucht und nicht mehr „im Stress“; sie sind weniger leistungsorientiert und ihre Gesundheitssorgen nehmen zu.

Tabelle 2: Mittelwertvergleiche und Effektstärken für Alterskohorte 50-59 aus 1982 (1999) und 1999 (2018)

Bundesländer	nur Westdeutschland					Deutschland gesamt				
Jahr	1982		1999			1999		2018		
Alterskohorte	50-59 Jahre		67-76 Jahre			50-59 Jahre		69-78 Jahre		
Altersdurchschnitt	55		72			55		74		
Skala	M	s	M	s	ES	M	s	M	s	ES
Lebenszufriedenheit	7.1	3.2	7.7	2.9	-.19	7.8	3.1	7.7	2.8	.06
Soziale Orientierung	6.8	3.0	6.6	3.0	.07	6.6	3.1	6.6	2.7	-.01
Leistungsorientierung	6.8	3.2	5.3	3.3	<b>.49</b>	7.2	3.2	6.0	3.2	<b>.40</b>
Gehemmtheit	5.6	3.0	5.7	3.2	-.03	4.9	3.1	4.8	2.8	.02
Erregbarkeit	5.9	2.9	4.8	2.9	.38	5.4	3.0	4.6	3.0	.25
Aggressivität	3.9	2.6	3.4	2.9	.18	4.3	2.8	3.3	2.5	.35
Beanspruchung	6.4	3.6	3.8	3.1	<b>.73</b>	5.9	3.5	4.0	3.0	<b>.54</b>
Körperliche Beschwerden	4.8	3.2	4.8	3.1	.02	3.8	3.0	4.1	3.2	-.11
Gesundheitssorgen	6.9	3.0	8.2	2.8	<b>-.41</b>	6.7	2.8	7.6	2.7	<b>-.32</b>
Offenheit	5.4	2.8	4.8	2.9	.23	5.9	2.8	5.2	2.8	.26
Extraversion	6.1	3.6	4.9	3.4	.34	6.6	3.5	6.0	3.2	.20
Emotionalität	6.5	3.7	5.5	3.5	.28	6.0	3.7	5.5	3.6	.13
Stichprobengröße (N)	192		192			258		258		

Anmerkung: Vergleich konstruierter Kohorten nur Westdeutschland 1982 zu 1999 (N = 192) und Deutschland insgesamt 1999 zu 2018 (N = 258), siehe auch Tabelle 1.

Zusammenfassend ergibt sich aus der neuen Repräsentativerhebung: Die beobachteten Veränderungen sind hinsichtlich ihrer Effektstärke gering und würden sich nicht oder nur geringfügig auf die Normen auswirken. Für die Testpraxis hatten die Autoren die Stanine-Skala ausgewählt, d.h. relativ grobe Standardwerte, welche der Spannweite der Testwerte angemessen sind. – Ausnahmen bestehen jedoch hinsichtlich der Skalen *Leistungsorientierung*, *Aggressivität*, *Extraversion* und *Emotionalität*. Hier bestehen in den Mittelwerten der Jahre 1999 und 2018 Trends, die teils etwas mehr als einen Rohwertpunkt betragen, am deutlichsten in den zwei Altersklassen der 16-19 und der 20-29-Jährigen. Folglich sind die Normen für diese Altersgruppen zu ändern. Zugleich werden die Normen-Tabellen insgesamt aktualisiert, d.h. auch die geringen Effekte berücksichtigt. – Darüber hinaus wird eine zusätzliche Tabelle mit Normen aufgenommen, die nach vier Bildungskategorien, aber nur hinsichtlich der drei auffälligen Skalen *Lebenszufriedenheit*, *Gehemmtheit* und *Emotionalität* differenziert. Generell bleibt die Gliederung der Normen nach Männern und Frauen sowie nach sieben Altersklassen.

Da sich die Mittelwerte der FPI-Skalen seit der ersten Normierung 1982, d.h. in einem Zeitraum von 36 Jahren, nur geringfügig änderten, ist jene, in den zitierten TBS-TK Richtlinien pauschal erwartete Neu-Normierung im Abstand von 8 bis 10 Jahren, zu relativieren. Erst aufgrund weiterer Kontrollen bei ähnlichen Persönlichkeitsfragebogen könnten empirisch fundierte Empfehlungen gegeben werden.

## 2.2 Skalenkonstruktion

Die Konstruktion der 12 FPI-Skalen wurde durch Itemanalysen und Reliabilitätsanalysen, Faktorenanalysen (orthogonal, Varimax-Rotation) und durch Clusteranalysen (Ward-Algorithmus) überprüft. Explorative Versuche mit IR-Modellierungen wurden kritisch kommentiert. Hier wird nur eine zusammenfassende Tabelle wiedergegeben. Aufgrund der Faktorenanalysen sind 92 Prozent der FPI-Items nach ihren höchsten Ladungen den betreffenden Skalen zuzuordnen. In der Clusteranalyse sind es 85 Prozent der Items, wobei insbesondere die Skala Gesundheitssorgen durch Tendenz zu zwei Subclustern auffällt (Tabelle 3).

Tabelle 3: Reliabilitätskoeffizienten, Faktorenanalysen der Daten aus drei Befragungswellen hinsichtlich der 10 Standardskalen und separat für E und N sowie eine Clusteranalyse (FPI-R Daten 2018, N = 3 450).

FPI-R-Skala	Anzahl Items	Bevölkerungsrepräsentativer Datensatz 2018					
		# Items mit höchster Ladung auf betreffendem Faktor			Reliabilität Cronbach $\alpha$	Clusterlösung (Anzahl Items)	
		FA Welle 1	FA Welle 2	FA Welle 3		korrekt zugeordnet	anderes Cluster
1 Lebenszufriedenheit	12	11	9	10	.77	11	1
2 Soziale Orientierung	12	12	12	12	.72	9	3
3 Leistungsorientierung	12	8	10	12	.77	11	1
4 Gehemmtheit	12	11	11	9	.77	12	0
5 Erregbarkeit	12	12	10	11	.78	12	0
6 Aggressivität	12	9	12	11	.77	10	2
7 Beanspruchung	12	11	11	11	.82	12	0
8 Körperliche Beschwerden	12	12	12	12	.77	12	0
9 Gesundheitssorgen	12	12	12	11	.75	5	7
10 Offenheit	12	11	11	12	.75	8	4
Itemsumme Skalen 1–10	120	109	110	111		102	18
E Extraversion	14	14	14	14	.80		
N Emotionalität	14	14	14	14	.82		

Anmerkungen: (1) Der Datensatz besteht aus drei repräsentativen Stichproben der deutschen Bevölkerung. Die Daten wurden in drei Erhebungswellen gesammelt (Juni und September-November 2018; N Welle 1 = 1093, N Welle 2 = 1137, N Welle 3 = 1163). (2) 12 zufällig ausgewählte Items aus dem gesamten Itempool der Standardskalen (n = 120) ergaben eine untere Grenze der Reliabilität von  $\alpha = .27$ .

Die multistrategische Kontrolle der Skalenkonstruktion ergab keine perfekte, doch eine weitgehende Übereinstimmung der Faktoren- und Clusteranalysen untereinander und eine Kontinuität hinsichtlich der früheren Analysen. Deshalb sehen die Testautoren keinen hinreichenden Anlass, einzelne FPI-Skalen zu revidieren. Die Skalen entsprechen weiterhin den erklärten Absichten der Testentwicklung. Die Forschungsvorhaben der Testautoren gaben den theoretischen Rahmen: psychophysiologische Persönlichkeitsforschung im Sinne Eysencks, d.h. hinsichtlich Emotionalität und Extraversion, außerdem Aggressivitätsforschung, Soziale Orientierung (prosoziale Einstellung), Leistungsorientierung, Beanspruchung und Körperliche Beschwerden. In diesem theoretischen Rahmen ist die Konstruktionsstrategie hypothetisch-deduktiv und empirisch-induktiv. Die persönlichkeits-theoretische Diskussion findet also hinsichtlich der *einzelnen Eigenschaftskonstrukte* statt, um die Iteminhalte so zu formulieren, dass die als wichtig angesehenen Facetten jedes Konstrukts repräsentiert sind – ohne Anspruch auf eine bestimmte, umfassende Theorie der Persönlichkeit und ohne Zuversicht in eine angeblich die gesamte Persönlichkeitsdomäne repräsentierende, lexikalisch-psychologische Reduktionsmethode (aufgrund englischer Wörterbücher?) mit wahlweise 15, 6, 5, 4 oder noch weniger Generalfaktoren – selbst wenn die Konstruktion und die Normierung bevölkerungsrepräsentativ durchgeführt wären (zur Diskussion, siehe Testmanual).

Einzelne FPI-Skalen weisen untereinander Korrelationen bis zu einer Größenordnung von etwa  $r = .45$  (gemeinsame Varianz ca. 20%) auf, wie es für einige der Eigenschaften persönlichkeitspsychologisch erwartet werden kann. Es verbleibt jedoch ein großer Anteil eigenständiger und psychologisch-deskriptiv nützlicher, wahrer Varianz, wie aus den Reliabilitätskoeffizienten ersichtlich ist. Die Thematik des FPI-R stimmt mit einigen anderen Fragebogen mehr oder minder überein. Ein gründlicher Vergleich, der über bloß statistische Aussagen hinsichtlich der Kovarianzen hinausginge, erfordert jedoch ein, offensichtlich noch nie organisiertes, übergeordnetes Bezugssystem, das wesentliche Kriterieninformation (multimodal, aggregierend und als MTMM-Analyse) aus den Anwendungsbereichen einschließen müsste, um pragmatisch zu informieren.

Zum Dilemma von Kontinuität und Vergleichbarkeit versus Verbesserung durch Revision sind als die bedeutendsten Vorbilder zu nennen: das *Minnesotas Multiphasic Personality Inventory* MMPI und die Entwicklung des *Maudsley Medical Questionnaire* MMQ zum *Maudsley Personality Inventory* MPI, *Eysenck Personality Inventory* EPI, *Eysenck Personality Questionnaire* EPQ (vgl. die Übersicht in Schmidt-Atzert & Amelang, 2012). Aus der ersten Etappe der Entwicklung des FPI bleibt die Erinnerung, dass bei Revisionen auch ein möglicher Schaden zu bedenken ist. Für die herausragende *Zürich Studie*, eine von

der Psychiatrischen Universitätsklinik organisierte Kohortenstudie, wählten Jules Angst und Mitarbeiter das FPI in der ersten Version aus. Die Zürich-Studie, ein vielseitiges Projekt mit verschiedenen Patientengruppen sowie repräsentativ erhobenen Vergleichsgruppen, führte zu einer Serie wichtiger Publikationen (u.a. Angst, Dobler-Mikola & Binder, 1984; Hengartner, Tyrer, Ajdacic-Gross, Angst & Rössler, 2018) – hinsichtlich des FPI jedoch nur teilweise mit dessen heutigen Skalen vergleichbar. Insofern könnte es zur Qualitätssicherung gehören, dass die Testautoren bei einer eingreifenden Revision wenigstens eine Kombination von Marker-Items entwickeln, um den Anschluss an solche seltenen Forschungsvorhaben zu erleichtern.

### 2.3 Integration von Validitätshinweisen

Von den Autoren eines eingeführten Tests ist zu erwarten, dass bei jeder Neuauflage mindestens eine Fortschreibung der Validitätshinweise erfolgt und in adäquaten Abständen eine neue Normierung, die mit der Kontrolle der Testkonstruktion verbunden wird. Diese „Pflege“ ist umso mehr erforderlich, wenn ein Test relativ weit verbreitet ist.

Die Chancen der dritten bevölkerungsrepräsentativen Normierung des FPI wurden erneut genutzt, um durch schriftliche Zusatzfragen und durch Mitwirkung der Interviewer weitere Daten und Validierungshinweise zu gewinnen:

- 24 soziodemografische Merkmale, einschließlich Bildungsabschluss, Schichtzugehörigkeit, Kirchenbindung, aufgrund des standardmäßigen mündlichen Interviews (IfD);
- Angabe der Parteipräferenz (Selbstbericht);
- Berufliche Belastung, u.a. Anforderungen und Dauer, und Gesundheitszustand, u.a. Arztbesuche, chronische Krankheit, Psychotherapie, Medikamente (Selbstberichte);
- Politisch-soziale Einstellungen und religiös-weltanschauliche Einstellungen mit abgeleiteten Miniskalen bzw. Aggregaten, u.a. autoritär-konservative Einstellung, ideelle Einstellung hinsichtlich Gottesglaubens, Religiosität und Sinnfragen (Selbstbeurteilungen);
- Gegenwärtige Sorgen und Ängste (Selbstbeurteilungen);
- verhaltensnahe Frage hinsichtlich Aggressivität, Extraversion, Leistungsorientierung, Beanspruchung und Soziale Orientierung, z.B. erfahrene körperliche Angriffe seit dem Alter von 16 Jahren (Anzahl), Kontakt mit Freunden in der Freizeit (Häufigkeit), leistungsorientierter Sport, tatsächlich gegebene Spenden (Selbstberichte).
- Eindruck vom Verhalten der Befragten auf acht 4-stufigen Skalen, z. B. wie selbstsicher, freundlich, lebhaft jemand wirkte (Verhaltensbeurteilungen durch die Interviewer);
- Einstufungen der Wohnung und Wohnumgebung sowie Beobachtungen hinsichtlich des Körperschmucks, d.h. der Anzahl sichtbarer Tätowierungen, Piercings und Ringe (Beobachtungen der Interviewer).

Es sind Daten unterschiedlicher Quellen bzw. Datenebenen, wobei zu unterscheiden ist zwischen Selbstbeurteilungen der Befragten und Selbstberichten, die im Prinzip objektiviert werden könnten, beispielsweise die Antworten hinsichtlich beruflicher Belastung oder Psychotherapie. Die Beurteilungen und Beobachtungen durch die Interviewer bilden eine besondere Datenklasse. Den Interviewern wurden keine Hinweise auf typische Methodenprobleme solcher Beobachtungen gegeben. Zweifellos werden diese Einstufungen durch sprachlich-semantische Schwierigkeiten und durch Beurteiler-Tendenzen beeinflusst sein, doch ist angesichts der sehr großen Anzahl von ca. 300 erfahrenen Interviewern anzunehmen, dass sich solche Tendenzen teilweise „ausmitteln“. Jedenfalls können auf diese Weise alltagsnahe psychologische Informationen zusätzlich gewonnen werden, wie sie sonst bei der Konstruktion und Normierung psychologischer Tests nicht möglich sind.

Die hypothesengeleiteten und die explorativen Analysen lieferten psychologisch wichtige Beiträge zur Validierung, die auch in dieser Kombination von Datenquellen neu sind. Die Ergebnisse sind im Testmanual tabellarisch und graphisch dargestellt, teils auch durch die *Aggregation* ausgewählter Kriterieninformation weitergeführt. Beispiele sind:

- Gesundheitszustand, d.h. ein Aggregat aufgrund der angegebenen Häufigkeit von Arztbesuchen, Psychotherapie, Medikamenten, mit den FPI-Testwerten *Körperliche Beschwerden* Effektstärke  $ES = +0.82$ ; Emotionalität  $+0.65$ ; Lebenszufriedenheit  $-.52$ .
- Verhaltensaussage „bereits Opfer körperlicher Angriffe gewesen“ mit dem FPI-Testwert *Aggressivität*  $ES = +1.82$ .
- Autoritarismus und AfD-Parteipräferenz mit dem FPI-Testwert *Soziale Orientierung*  $ES = -0.80$ .
- Ideelle Einstellung, d.h. ein Aggregat aufgrund der Aussagen über Gottesglauben, Religiosität, Sinnfragen, mit dem FPI-Testwert *Soziale Orientierung*  $ES = +0.58$ .

- Verhaltensbeurteilung „gehemmt“ durch den Interviewer mit den FPI-Testwerten *Gehemmtheit* ES = +0.87; Leistungsorientierung = +0.85; Extraversion –0.80.

Die Sekundärliteratur mit empirischen und methodischen Beiträgen zum FPI ist heute nicht mehr überschaubar. Die FPI-Bibliographien wurden für insgesamt 655 Arbeiten bis zum Jahr 1994 geführt (Potreck, 1983; Jehle & Fahrenberg, 1994). Diese Dokumentation wird heute durch die Literaturbanken PubPsych, PsycINFO, MEDLINE u.a. geleistet, jedoch nur unvollständig, denn die unveröffentlichten Diplom-, Master- und Doktorarbeiten fehlen. Im Manual der 9. Auflage werden neben früheren, aber herausragenden Beiträgen insbesondere jene referiert, die zur Persönlichkeitsforschung und zu zentralen Anwendungsgebieten, insbesondere Klinische Psychologie, Psychotherapieforschung einschließlich Sozialtherapie, wichtig erscheinen. Hervorgehoben wurden auch Studien, die eine Verbindung zwischen Persönlichkeitsskalen und Ergebnissen des Ambulanten Assessment herstellen oder ungewöhnliche Befunde enthalten, wie die FPI-Daten von auffälligen Personen mit extremistischer Orientierung.

In zahlreichen Untersuchungen wurden Unterschiede in den FPI-Testwerten bestimmter Personengruppen beschrieben. Diese Gruppierungen erfolgten je nach Fragestellung der Untersucher aufgrund von bestimmten Statusmerkmalen, psychosozialen oder klinisch-psychologischen Merkmalen. Demgegenüber sind Publikationen über die Vorhersage praktisch relevanter Kriterien, über die Evaluation individueller Begutachtungen oder Evaluation von Klassifikations- und Selektionsentscheidungen in bestimmten Anwendungsfeldern selten, und anspruchsvolle persönlichkeitspsychologische Grundlagenforschung zu den theoretischen Konstrukten (Persönlichkeitsdispositionen), ihrer funktionellen Dynamik und zur biographischen Entwicklung fehlen weitgehend. Oft handelt es sich um kleine Personengruppen, oft um eine Gelegenheitsauswahl, d.h. mit unklarer Präselektion, oder es mangelt an einer äquivalenten Kontrollgruppe. Falls wichtige Befunde mitgeteilt werden, fehlen in der Regel Replikationen. Die erwähnte Züricher Kohortenstudie ist auch durch ihren Bezug auf ein sehr breites Spektrum klinisch-psychologisch relevanter Befunde eine weit herausragende Forschung.

Insgesamt mangelt es jedoch an methodisch anspruchsvolleren und zugleich praxisnahen Untersuchungen der externen Validität im Hinblick auf wichtige, diagnostisch und prognostisch wichtige Kriterien bzw. den tatsächlichen „Entscheidungsnutzen“. Eine *vertiefte Qualitätskontrolle* und Evaluation dieser Art übersteigt die Möglichkeiten der meisten Testautoren, so dass Kooperationen notwendig sind.

## 2.4 Künftige Evaluationen

Die Frage nach geeigneten Forschungsprojekten zur Kriterienvvalidierung von FPI-Skalen, insbesondere im klinisch-psychologischen, rehabilitationspsychologischen und sozialtherapeutischen Bereich liegt nahe. Die Testautoren waren durchaus motiviert, sich weiterhin dieser Aufgabe zu stellen. Zwei relativ große Projekte scheiterten, trotz der Kooperationsbereitschaft der Ärzte und Psychologen in den betreffenden Kliniken und der bereits gesicherten Finanzierung, an der Universitätsbürokratie und an einem Datenschutzbeauftragten und dessen zusätzlicher Forderung, die Projektleiter dürften nicht mit den eingestellten Projektmitarbeitern sprechen, da andernfalls die Anonymität der von diesen Mitarbeitern untersuchten Patienten nicht sicher gewährleistet sei. Projekte dieser Art können heute wahrscheinlich nur noch aus den großen Institutionen der Praxisfelder (Kliniken, soziale Einrichtungen, Betriebe, Verwaltungen u.a.) heraus initiiert und organisiert werden, denn nur dort sind die wichtigsten Indikatoren und Prädiktoren sowie die wesentlichen Informationen über den Entscheidungsnutzen zu gewinnen und die Prinzipien der modernen Assessmenttheorie adäquat anzuwenden.

Das Drängen auf gründlichere externe Kriterienvvalidierung kann im Zusammenhang mit der erwähnten Umfrage unter Mitgliedern des Berufsverbandes BDP (Roth u.a., 2010) gesehen werden und der vorrangigen Forderung, dass in der universitären Diagnostikausbildung eine kritische Diagnostik gelehrt wird. Gerade die so häufig verwendeten Persönlichkeitsfragebogen verlangen eine methodenbewusst-kritische Anwendung und möglichst auch Strategien der multimodalen Diagnostik, d.h. Absicherungen durch andere Methoden und vorsichtige Interpretation. Wegen der äußerlich leichten Handhabung können vielleicht die Schwierigkeiten der Testkonstruktion übersehen werden.

Die kritische Diskussion der Vorzüge und der Probleme von Persönlichkeitsfragebogen gehören nach der Auffassung der Autoren auch in das Test-Manual. Wie in den früheren Auflagen wird ausführlich über die Absichten der Konstruktion, über Validitätshinweise, über die Anwendung und auch über die Kritik an Persönlichkeitsfragebogen berichtet. Auch eine Einführung in wichtige Prinzipien der modernen Assessmenttheorie ist notwendig – in Erinnerung an Cattells (1973) *Personality and Mood by*



*Questionnaires*. Wichtige Leitkonzepte der modernen Assessmenttheorie werden an dieser Stelle nur genannt: Multitrait-Multimethod, Multimodale Strategie, Generalisierbarkeitstheorie, Aggregation und Brunswiks Linsenmodell, Wittmanns Multivariate Reliabilitätstheorie. Hinzu kommt die Frage nach der ökologischen (externen) Validität von Testergebnissen hinsichtlich der tatsächlich im Alltag registrierten Indikatoren des individuellen Verhaltens und Befindens mit der Methodik des *Ambulanten Assessment* (Ecological Momentary Assessment, Experience Sampling Method), technisch unterstützt durch miniaturisierte, ggf. auch interaktive Systeme. Diese innovative Methodik, die in wichtigen Ansätzen in den deutschsprachigen Ländern entwickelt wurde, ist auf einigen Gebieten bereits die Methode der Wahl gegenüber den natürlich viel einfacheren Persönlichkeitsfragebogen. Die Methodik und Anwendung des *ambulanten Assessment* haben sich international mit jährlich mehr als 500 Publikationen zu einer breiten und höchst aktiven Forschungsrichtung entwickelt, die jedoch in Lehrbüchern der Differentiellen Psychologie und der psychologischen Diagnostik auch nach 40 Jahren noch kaum rezipiert ist.

Diese Informationen als Grundlage kritisch-methodenbewusster Diagnostik sind heute noch wichtiger, da im Vergleich zum früheren Diplom-Studiengang wahrscheinlich nur ein Teil der Absolventen eines Psychologie-Studiums (mit dem speziellem Vertiefungsgebiet) über eine hinreichende diagnostische Ausbildung und Gutachtenpraxis unter Supervision, einschließlich Testkonstruktion und Einführung in die wichtigsten Prinzipien der modernen Assessmenttheorie verfügen wird.

## 2.5 Rezensionen und Kommentare, Datensätze open access

Zum FPI wurden im Laufe der Jahre 15 Testrezensionen, außerdem eine Anzahl von Reviews in Lehrbüchern und in Handbuchbeiträgen verfasst. Rezensionen tragen zur Qualitätssicherung bei, insbesondere wenn sie explizit methodenkritisch und nicht pauschal formuliert sind. Möglichst prägnante Testrezensionen, die bestehende Defizite hervorheben, könnten tatsächlich eine motivierende Wirkung auf Testautoren und Testverlag haben. – Andererseits gab es in einigen der FPI-Rezensionen und Reviews sachliche Fehler und Missverständnisse. So wurde die Skalenkonstruktion unter anderem als „faktorenanalytisch“, „blind faktorenanalytisch“ oder „willkürlich“ beurteilt statt das hypothetisch-deduktive und empirisch-induktive Verfahren nachzuvollziehen, dem bereits Eysenck in der Entwicklung der Dimensionen Neuroticism und Extraversion-Introversion hinsichtlich klinischer Kriterien folgte (Eysenck, 1950). Die beiden Dimensionen bilden noch immer das Rückgrat der meisten heutigen Persönlichkeitsinventare. Die erwähnten Rezensionen motivierten dazu, in der Neuauflage des Testmanuals vier verschiedene Konstruktionsstrategien zu beschreiben: die hauptsächlich kriterienbezogene Strategie, die auf eine spezielle Persönlichkeitstheorie bezogene deduktive Strategie, die lexikalisch-induktive Strategie und die kombinierte, hypothetisch-deduktive und empirisch-induktive Strategie.

Eine bessere fachliche Kooperation zwischen den Testrezensenten und den Testautoren ist wünschenswert. Selbst wenn den Testautoren von der Redaktion intern die Möglichkeit einer Stellungnahme gegeben wird, könnte eventuell ein gedruckter Kommentar der wechselseitigen Qualitätskontrolle dienlich sein.

Zum FPI wurden seit der 1. Auflage Datensätze zu Vergleichszwecken zur Verfügung gestellt. Sie stammten aus eigenen Untersuchungen oder konnten von anderen Autoren anonymisiert zur Verfügung gestellt werden. Für solche Vergleichszwecke und für kritische Reanalysen oder Unterrichtszwecke sind die Datensätze der Repräsentativerhebungen zur 4. und 7. Auflage des FPI-R in ZPID PsychData archiviert und können dort aufgrund eines Datennnehmer-Antrags mit dem ZPID angefordert werden. Diese Datensätze enthalten eine beträchtliche Anzahl von zusätzlichen psychologischen und sozioökonomischen Daten. Auch für den aktuellen Datensatz der 9. Auflage ist diese open access Regelung vorgesehen. Rückblickend ist jedoch deutlich, dass die einzelnen Datensätze nur sehr selten angefordert wurden; eine gezielte Umfrage bei den Datennnehmern hinsichtlich der Normierungs-Datensätze ergab eine nur minimale Nutzung, am ehesten noch für Unterrichtszwecke.

## 3. Ausblick

Ist „Testpflege“ ein zusätzliches Gütekriterium? Die Normierung von Persönlichkeitsfragebogen sollte regelmäßig überprüft werden wie es in der produktiven Kooperation mit dem Testverlag und einem auf Umfragen spezialisierten Institut zu leisten ist. Eine Daumenregel hinsichtlich des adäquaten Zeitintervalls kann nicht vorgeschlagen werden, denn es mangelt noch an weiteren Erfahrungsberichten über die Effektstärke solcher Trends. In der „Testpflege“ erfordern bevölkerungsrepräsentative Normierungen den größten finanziellen und zeitlichen Aufwand. Deshalb ist es zweckmäßig, in einer Repräsentativerhebung mehrere Fragebogen, je nach Länge und Inhalt, zu kombinieren. An eine fachliche Kooperation ist bereits beim kompetenten Entwurf eines mehrdimensionalen Persönlichkeitsfragebogen zu denken, um die speziellen Erfahrungen aus den verschiedenen Bereichen der Persönlichkeitsforschung zu verbinden. Methodisch überzeugende Validitätsstudien und

Evaluationen des Entscheidungsnutzens erfordern gemeinsame Projekte von Testautoren und von Psychologen, die in größeren Organisationen tätig sind und die im Interesse dieser Organisationen anonymisierte Kriterieninformationen verwenden dürfen. Zunehmend wichtig sind auch die mit der Methodik des ambulanten Assessment unter Alltagsbedingungen durchgeführten Studien. Auch solche Projekte sind ohne Kooperation kaum möglich.

Künftige Neukonstruktionen von mehrdimensionalen *Persönlichkeitsfragebogen* sollten sich grundsätzlich nicht mit einer Datenreduktion auf wenige Sekundärfaktoren oder mit einem angeblich universellen System mit minimaler Anzahl von Skalen begnügen. Die eigenschaftstheoretisch, im deduktiv-induktiven Verfahren oder primär kriterienorientiert konstruierten Skalen haben den Vorzug, dass sie eventuell sogar in Modulen – näherungsweise wie bei einem Intelligenztest – zusammengesetzt werden könnten, um dem Anwender eine aufgabenrelevante Auswahl zu ermöglichen. Wegen der erforderlichen Forschungskompetenz in vielen Bereichen und wegen des hohen Arbeits- und Kostenaufwands wären solche Vorhaben nur auf eine neue Weise kooperativ möglich – wie es sich in der Labormethodik der Medizin und Naturwissenschaften bewährt hat.

## Literatur

- Angst, J., Dobler-Mikola, A. & Binder, J. (1984). The Zurich Study – A prospective epidemiological study of depressive, neurotic and psychosomatic syndromes. I. Problem, Methodology. *European Archives of Psychiatry and Neurological Sciences*, 234, 13-20.
- Bengel, J. & Wittmann, W. W. (1983). Urteilsfehler aufgrund zeitlicher Veränderungen von Testnormen. *Diagnostica*, 29, 101–117.
- Diagnostik- und Testkuratorium (2018). TBS-DTK. Testbeurteilungssystem des Diagnostik- und Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 3. Januar 2018. *Psychologische Rundschau*, 69 (2), 109–116. <https://econtent.hogrefe.com/doi/abs/10.1026/0033-3042/a000401>
- Cattell, R. B. (1973). *Personality and mood by questionnaire*. San Francisco: Jossey-Bass.
- Eysenck, H. J. (1950). Criterion analysis: An Application of the hypothetico-deductive Method as Factor Analysis. *Psychological Review*, 57, 38–53.
- Fahrenberg, J., Selg, H. & Hampel, R. (1983). Die bevölkerungsrepräsentative Normierung des Freiburger Persönlichkeitsinventars FPI-A. *Diagnostica*, 29, 336–343.
- Fahrenberg, J., Hampel, R. & Selg, H. (1984–2001). *Das Freiburger Persönlichkeitsinventar (revidierte Fassung FPI-R und teilweise geänderte Fassung FPI-A1)*, 4. Aufl. 1984; ergänzte 5. Aufl. 1989, ergänzte 6. Aufl. 1994, revidierte 7. Auflage 2001; 1. bis 3. Aufl. siehe Fahrenberg & Selg). Göttingen: Hogrefe.
- Fahrenberg, J., Hampel, R. & Selg, H. (2020, im Druck). *Freiburger Persönlichkeitsinventar FPI-R. Neue Normierung und Validitätshinweise, Prinzipien der Testkonstruktion und modernen Assessmenttheorie*. (9. erweiterte Aufl.). Göttingen: Hogrefe.
- Fahrenberg, J. & Selg, H. (1970). *Das Freiburger Persönlichkeitsinventar FPI* (1. Auflage 1970, 2. erweiterte Aufl. 1973, ergänzte 3. Aufl. 1978). Göttingen: Hogrefe.
- Hengartner, M. P., Tyrer, P., Ajdacic-Gross, V., Angst, J. & Rössler, W. (2018). Articulation and testing of a personality-centered model of psychopathology: evidence from a longitudinal community study over 30 years. *European Archives Psychiatry & Clinical Neuroscience*, 268 (5), 443–454.
- International Test Commission (2005). *ITC Guidelines on Computer-Based and Internet Delivered Testing*. [https://www.intestcom.org/files/guideline\\_computer\\_based\\_testing.pdf](https://www.intestcom.org/files/guideline_computer_based_testing.pdf)
- International Test Commission (2013). *ITC Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores*. [https://www.intestcom.org/files/guideline\\_quality\\_control.pdf](https://www.intestcom.org/files/guideline_quality_control.pdf)
- International Test Commission (2015). *The ITC Guidelines on Practitioner Use of Test Revisions, Obsolete Tests, and Test Disposal*. [https://www.intestcom.org/files/guideline\\_test\\_disposal.pdf](https://www.intestcom.org/files/guideline_test_disposal.pdf)
- International Test Commission (2017). *ITC Guidelines for Translating and Adapting Tests* (Sec. Edition, Version 2.4). [https://www.intestcom.org/files/guideline\\_test\\_adaptation\\_2ed.pdf](https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf)

- Jehle, M. & Fahrenberg, J. (1994). *Literaturverzeichnis zum Freiburger Persönlichkeitsinventar FPI 1983–1993*. (Forschungsberichte des Psychologischen Instituts der Albert-Ludwigs-Universität Nr. 105). Freiburg i. Br.: Psychologisches Institut. [http://www.jochen-fahrenberg.de/uploads/media/FPI\\_Bibliographie\\_1983-1993\\_01.pdf](http://www.jochen-fahrenberg.de/uploads/media/FPI_Bibliographie_1983-1993_01.pdf)
- Kersting, M. (2018a). Qualitätssicherung und -optimierung in der Eignungsdiagnostik. In: Diagnostik- und Testkuratorium (Hrsg.) *Personalauswahl kompetent gestalten: Grundlagen und Praxis der Eignungsdiagnostik nach DIN 33430* (S. 2-20). Berlin: Springer. (DOI 10.1007/978-3-662-53772-5)
- Kersting, M. (2018b). Zur Information über und Dokumentation von Instrumenten zur Erfassung menschlichen Erlebens und Verhaltens – Die DIN SCREEN Checkliste 1, Version 3. In: Diagnostik- und Testkuratorium (Hrsg.) *Personalauswahl kompetent gestalten: Grundlagen und Praxis der Eignungsdiagnostik nach DIN 33430* (S. 223-244). Berlin: Springer. [http://kersting-internet.de/pdf/Kersting\\_2017\\_in\\_DTK\\_DIN-SCREEN\\_Checkliste\\_1.pdf](http://kersting-internet.de/pdf/Kersting_2017_in_DTK_DIN-SCREEN_Checkliste_1.pdf)
- Obschonka, M., Wyrwich, M., Fritsch, M., Gosling, S. D., Rentfrow, P. J. & Potter, J. (2019). Von unterkühlten Norddeutschen, gemütlichen Süddeutschen und aufgeschlossenen Großstädtern: Regionale Persönlichkeitsunterschiede in Deutschland. *Psychologische Rundschau*, 70(3), 173-194.
- Pfeifer, C., Baier, D. & Kliem, S. (2018). *Zur Entwicklung der Gewalt in Deutschland. Schwerpunkte: Jugendliche und Flüchtlinge. Zentrale Befunde eines Gutachtens im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend*. <https://www.zhaw.ch/storage/shared/sozialarbeit/News/gutachten-entwicklung-gewalt-deutschland.pdf>
- Potreck, F. (1983). *Kommentiertes Literaturverzeichnis zum Freiburger Persönlichkeitsinventar FPI. Veröffentlichungen seit 1978*. Forschungsberichte des Psychologischen Instituts der Universität Freiburg, Nr. 11. [http://www.jochen-fahrenberg.de/uploads/media/FPI\\_Bibliographie\\_seit\\_1978\\_01.pdf](http://www.jochen-fahrenberg.de/uploads/media/FPI_Bibliographie_seit_1978_01.pdf)
- Roth, M., Schmitt, V. & Herzberg, P.Y. (2010). Psychologische Diagnostik in der Praxis. Ergebnisse einer Befragung unter BDP-Mitgliedern. *Report Psychologie*, 35, 118-128.
- Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik* (5. Aufl.). Berlin: Springer.
- Sneed, H.M. & Baumgartner, M. & Seidl, R. (2011). Testpflege und -fortschreibung. In: R. Richard (Hrsg.). *Der Systemtest. Von den Anforderungen zum Qualitätsnachweis* (S. 199-219). München: Carl Hanser, S. 199
- Wagner, J., Lüdtke, O. & Robitzsch, A. (2019). Does personality become more stable with Age? Disentangling state and trait effects for the Big Five across the life span using local structural equation modelling. *Journal of Personality and Social Psychology*, 116(4), 666-680.
- Wirtz, M.A. (2019). Qualitätssicherung. In: M.A. Wirtz (Hrsg.). *Dorsch. Lexikon der Psychologie* (S. 1465). (19. überarbeitete Aufl.). Göttingen: Hogrefe.