Hamburg-Wechsler-Intelligenztest für Kinder - IV

F. Petermann und U. Petermann

EINSATZBEREICH:

Der HAWIK-IV ist ein Intelligenztest zur Erfassung allgemeiner und spezifischer intellektueller Fähigkeiten
bei Kindern von 6;0 bis 16;11 Jahren

DAS VERFAHREN:

Neben einer Aktualisierung des Testmaterials und einer Neunormierung steht eine Änderung der Gesamtstruktur im Vordergrund der Testneuentwicklung.
Der HAWIK-IV enthält insgesamt 15 Untertests.
Die Erfassung der fünf Intelligenzwerte
· Sprachverständnis
· Wahrnehmungsgebundenes Logisches Denken
· Arbeitsgedächtnis
· Verarbeitungsgeschwindigkeit
· Gesamt-IQ-Wertes ermöglichen es dem Anwender, ein fundiertes und differenziertes Bild des kognitiven Entwicklungsstandes zu erstellen.
Zusätzlich können auf der Untertestebene weitere Analysen vorgenommen werden.
So ermöglicht die Profilanalyse sehr differenzierte Aussagen über Stärken und Schwächen eines Kindes. Zusätzlich liefern Prozessanalysen wertvolle Hinweise für eine gezielte Therapieplanung. Der HAWIK-IV bietet außerdem:
· eine repräsentative und umfassende Neunormierung
· attraktive und kindgerechte Testmaterialien
· anwenderfreundliches Manual sowie Protokollbogen
· ökonomische Testdurchführung
· wissenschaftliche Fundierung

ZUVERLÄSSIGKEIT:

Die Reliabilität der Untertests variiert zwischen r = .76 und r = .91
Indexebene zwischen r = .87 und r = .94.
Gesamttest r = .97.

GÜLTIGKEIT:

Es liegen Studien zur faktoriellen und zur kriteriumsbezogenen Validität vor.

NORMEN:

1.650 Kinder aus der Bundesrepublik Deutschland, Österreich und der deutschsprachigen Schweiz.

BEARBEITUNGSDAUER:

zwischen 60 und 90 Minuten.

In Anwendung seit 2007, in 2., ergänzter Auflage seit 2008.

TAT

Thematic Apperception Test

3., überarbeitete Auflage

Von H. A. Murray

AREA OF APPLICATION:

Kinder ab 4 Jahre und Erwachsene.

Zu beantwortende Fragen:
Was passiert auf der Tafel?
Was hat dazu geführt?
Gefühle und Emotionen
Ausgang der Szene

Der TAT hat sich auch in Europa neben dem Rorschach-Test eine zentrale Stellung unter den persönlichkeitsdiagnostischen Verfahren erobert. Er dient zur Persönlichkeits- und Neurosendiagnostik von erwachsenen Personen, aber auch von Kindern ab vier Jahren.

TEST DURATION:

ca. 100 Minuten.

<u>Bewertung Projektiver Testverfahren</u>

• es kommt sehr darauf an, dass wir Verfahren weiterentwickeln und Auswertungen
weiterentwickeln, so dass klar wird, welche Intentionen hinter welcher Aussage
stehen
• das Problem ist, dass man nicht denken sollte, hier wird das innerste der
Person/die Wahrheit klar
→ naive Interpretation
• es gibt auch Entwicklungen, die wieder sehr nah an die projektiven Testverfahren
herankommen: Implizite Tests (Reaktionszeitmessung): nicht bewusste Annahmen
sollen damit gemessen werden (in der Sozialpsychologie z.B. Einstellungsmessung, aber auch Verfahren zur Messung von Selbsteinstellung oder
Ängstlichkeit)
• aufgrund der schwachen theoretischen Begründung besteht hier noch

großer
Forschungsbedarf, allerdings wird auch in der neueren Zeit viel daran geforscht

BIT II

Berufs-Interessen-Test II

2. Auflage

von M. Irle und W. Allehoff

EINSATZBEREICH:

Jugendliche ab 14 Jahren; junge Erwachsene. Einzel- und Gruppenverfahren. Verwendung in Berufsberatung und Eignungsdiagnostik.

DAS VERFAHREN:

Der BIT II ist eine umfassende Weiterentwicklung des Berufs-Interessen-Tests (BIT). Er erlaubt, berufliche Interessen einzuschätzen, berücksichtigt dabei jedoch nicht musische, Bildungs- oder Freizeit-Interessen. Der BIT II liegt in zwei Parallelformen des forced choice-Verfahrens und zusätzlich in zwei Parallelformen des free choice-Verfahrens vor. Der Test diagnostiziert 9 Interessenrichtungen. Pro Interessenrichtung finden sich zwei mal 9 Items (konkrete berufliche Tätigkeiten), die in beiden Verfahren jeweils identisch sind. Die Probanden äußern ihre Präferenzen zu jeder Tätigkeit im Vergleich zu Tätigkeiten aus anderen Interessengebieten (forced choice) bzw. unabhängig von solchen (free choice) auf einer Rating-Skala.

ZUVERLÄSSIGKEIT:

Im forced choice-Verfahren werden Interkorrelationen zwischen den Paralleltests von r = .71 bis r = .89 erreicht, im free choice-Verfahren von r = .60 bis r = .76.

GÜLTIGKEIT:

Es liegen Muster der faktoriellen Gültigkeit vor. Die Interkorrelationen zwischen den Interessen streuen zwischen r = -.44 und r = .56. Die Konstrukt-Gültigkeit wird in einer gesonderten Studie nachgewiesen.

NORMEN:

Geschlechtsspezifische Normen für alle 4 Paralleltests (jeweils N = 500).

BEARBEITUNGSDAUER:

Etwa 20 bis 40 Minuten.

BIP: Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung

Ziel des BIP ist die standardisierte Erfassung des Selbstbildes eines Testkandidaten in Hinblick auf relevante Beschreibungsdimensionen aus dem Berufsleben.“ (Hossiep & Paschen, 1998) Die aktuelle Forschungsversion des BIP umfasst 251 Fragen, die als Aussagen formuliert sind und bei denen der Teilnehmer jeweils angibt, inwieweit sie für ihn zutreffen.

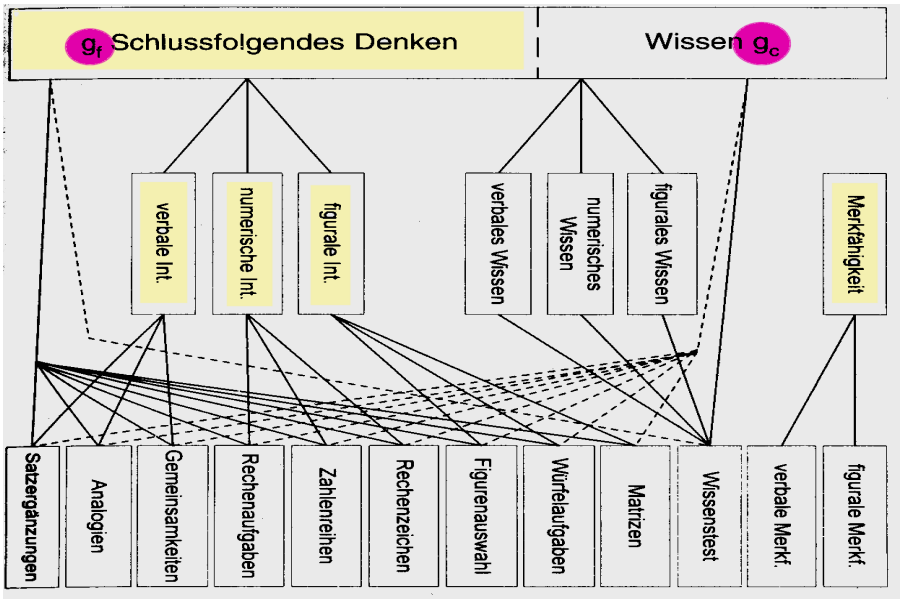
Das BIP erfasst 17 Persönlichkeitseigenschaften, die sich neben der fachlichen Qualifikation im Berufsleben als bedeutsam erwiesen haben. Sie werden vier Persönlichkeitsbereichen zugeordnet. Hierbei wird keineswegs unterstellt, dass für eine erfolgreiche Berufstätigkeit alle Dimensionen auf überdurchschnittlich hohem Niveau ausgeprägt sein müssen. Es hängt wesentlich vom Anforderungsprofil der Position ab, welche der aufgeführten Eigenschaften für die erfolgreiche Ausübung einer bestimmten Tätigkeit mehr oder weniger von Bedeutung sind.

Leistungsmotivation Gestaltungsmotivation Führungsmotivation Wettbewerbsorientierung	BERUFLICHE ORIENTIERUNG		ARBEITS-VERHALTEN	Gewissenhaftigkeit Flexibilität Handlungsorientierung Analyseorientierung
	ÜBERFACHLICHE KOMPETENZEN			
Sensitivität Kontaktfähigkeit Soziabilität Teamorientierung Durchsetzungsstärke Begeisterungsfähigkeit	SOZIALE KOMPETENZEN		PSYCHISCHE KONSTITUTION	Emotionale Stabilität Belastbarkeit Selbstbewusstsein

Objektivität:
Das Verfahren ist objektiv und somit in der Durchführung und Auswertung vom Durchführenden unabhängig. Genaue Hinweise und Instruktionen zu den o.g. Aspekten des Fragebogens finden sich in der Handanweisung des BIP, ebenso zur Interpretation.
Reliabilität:
internen Konsistenz (Cronbachs α zwischen .74-.91),
Split Half (zwischen r=.72-.91)
Retest Reliabilität nach einem Zeitraum von 8-10 Wochen und nochmals nach 5-36 Monaten (r >70). Die Werte bewegen sich auf einem Niveau, das für Persönlichkeitsverfahren angezeigt ist, teilweise auch höher, und können somit als zufriedenstellend betrachtet werden.
Validität:
Die Kriteriumsvalidität wurde anhand von objektiven und subjektiven Maßen des beruflichen Erfolges wie
- berufliches Entgelt (bereinigt um Alterseffekte, r= .40),
- hierarchische Position (r= .41),
- Einschätzung der eigenen Arbeitszufriedenheit (r= .38) und
- des eigenen Berufserfolges (r= .42) berechnet
- für die Zielgruppe der Hochschulabsolventen an Maßen des Studienverlaufes wie Gesamtdauer von außeruniversitären Praktika (r= .48) und gesellschaftlichem Engagement (r= .56).

Zusammenhänge des BIP mit anderen Persönlichkeitsverfahren wie mit dem Eysenck-Persönlichkeitsinventar (EPI), dem Neo-Fünf-Faktoren-Inventar (NEO-FFI) und dem 16-Persönlichkeits-Faktoren-Test (16 PF-R) konnten sehr zufriedenstellende Korrelationen für verwandte Konstrukte zwischen r= .54-.84 aufzeigen und somit substantielle Zusammenhänge nachgewiesen werden.

IST-2000-R



Intelligenz-Struktur-Test 2000 R

2., erweiterte und überarbeitete Auflage

D. Liepmann, A. Beauducel, B. Brocke und R. Amthauer

EINSATZBEREICH:

Jugendliche ab 15 Jahren und Erwachsene. Einzel- und Gruppentest.

DAS VERFAHREN:

Der I-S-T 2000 R ist ein vielseitig einsetzbarer, ökonomischer Intelligenztest, der in der 2. Auflage durch eine echte Parallelform erweitert und durch aktuelle Normdaten und Untersuchungen zu den Gütekriterien ergänzt wurde. Die Kennwerte für die Parallelform (Form C) entsprechen nahezu vollständig den Statistiken der Pseudoparallelformen A und B. Weiterhin besteht die Möglichkeit einen separat auswertbaren Wissenstest einzusetzen. Gegenüber dem ursprünglichen I-S-T 70 bleiben weiterhin sechs der ursprünglich neun Aufgabengruppen in überarbeiteter Form erhalten

Satzergänzung

Analogien

Gemeinsamkeiten

Zahlenreihen
Figurenauswahl
und Würfelaufgaben)
durch drei Aufgabengruppen (Rechenaufgaben ohne verbalen Anteil, Vorzeichenaufgaben und Matrizenaufgaben) sowie einen Test zum Allgemeinwissen ergänzt
Der I-S-T 2000 R ist modular aufgebaut, d.h. bestimmte Komponenten können in Abhängigkeit von inhaltlichen und ökonomischen Anforderungen hinzugenommen oder weggelassen werden. Im Rahmen des neuen, theoretisch fundierten und empirisch begründeten Strukturkonzepts können mit dem I-S-T 2000 R folgende 11 Fähigkeiten erfasst werden:
verbale Intelligenz
figural-räumliche Intelligenz
rechnerische Intelligenz
figurale Merkfähigkeit
schlussfolgerndes Denken
verbales Wissen
figural-bildhaftes Wissen
numerisches Wissen und Wissen (Gesamt) sowie fluide und kristallisierte Intelligenz.

ZUVERLÄSSIGKEIT:

Die Reliabilitätsschätzungen für die einzelnen Skalen liegen zwischen $\alpha = .87$ und $\alpha = .97$ (Cronbachs Alpha) bzw. $r = .88$ und $r = .96$ (Split-Half-Reliabilität).

GÜLTIGKEIT:

Validitätsbelege ergeben sich aus der Überprüfung der Struktur des Instruments in multivariaten Analysen ($N = 2.208$ bzw. $N = 2.363$). Korrelationen mit verschiedenen Testverfahren (Test d2, HAWIE-R, MWT-B, CFT 20, Raven und FRT), Aspekten des Fremdsprachenerwerbs (START-E), Dimensionen der BIG-FIVE (NEO-FFI), arbeitspsychologischen Konzepten (Commitment, Arbeitszufriedenheit) und Schulnoten weisen auf konvergente und diskriminante Gültigkeit hin.

NORMEN:

Basis von inzwischen mehr als 5.800 Personen im Alter zwischen 15 und 60 Jahren vor (Gymnasiasten, Nicht-Gymnasiasten). Für eine nach Schulbildung gewichtete Gesamtstichprobe wurde eine Zufallsauswahl von 2.020 Probanden vorgenommen.

BEARBEITUNGSDAUER:

In Abhängigkeit der verwendeten Module zwischen ca. 77 (Grundmodul) und 130 Minuten (einschließlich Wissenstest).

Explorix: Werkzeug zur Berufsauswahl und Laufbahnberatung

Der EXPLORIX wurde hauptsächlich aus zwei Gründen entwickelt:
a) als Werkzeug für Berufs- und LaufbahnberaterInnen
b) als Selbstexplorationsinstrument zur Berufswahl und Laufbahnplanung für Personen, die keinen Zugang zu Beratungsstellen haben oder keine Beratung wünschen. Er kann somit von Fachpersonen im Rahmen einer Beratung (Variante Beratungsinstrument) oder als Selbst-Test von Rat suchenden Personen völlig eigenständig durchgeführt werden (Variante Selbst-Test).
Als Anwendungsmöglichkeiten zeichnen sich die folgenden drei Hauptfunktionen ab:
- Unterstützung der Berufswahl und Laufbahnplanung
- Strukturierung von Berufs- und Persönlichkeitsdaten
- Persönlichkeitsforschung.

Die Anwendung eignet sich auch für Gruppen oder Schulklassen, etwa im Rahmen des Berufswahlunterrichts.

Zielgruppe (Testpersonen)
EXPLORIX wurde für Jugendliche und Erwachsene mit guten Deutschkenntnissen und erweiterten oder höheren Schulansprüchen konzipiert. Das Mindestalter für die Durchführung beträgt 15 Jahre.
Theoretische Grundlagen
→Self-directed Search kann mit "Selbstdurchführbare Suche" übersetzt werden: Ein Selbsterkundungsverfahren zur Berufswahl und Laufbahnplanung, das die Persönlichkeit in verschiedenen Facetten erfragt und diese in ein Gesamtprofil integriert.
→wissenschaftlich fundierter Test zur Berufswahl und Laufbahnplanung. Er ist selbsterklärend, erfasst die berufsbezogene Persönlichkeit (Interessen und Fähigkeiten) und liefert Berufsideen.
Die sechs Dimensionen beruhen auf den sogenannten RIASEC-Persönlichkeits-Umwelttypen (hexagonales Modell).
Nachfolgend werden die Typen übersetzt und umschrieben:
R (Realistic): realistisch; handwerklich-technisch, land-/forstwirtschaftlich
I (Investigative): intellektuell, untersuchend-forschend
A (Artistic): kreativ, künstlerisch, sprachlich, gestalterisch
S (Social): sozial, erziehend-pflegend
E (Enterprising): unternehmerisch, führend-organisierend-verkaufend
C (Conventional): konventionell, ordnend-verwaltend

Reliabilität:
-interne Konsistenz der Gesamtskalen zwischen .86 und .91
-Subskalen zwischen .70 und .84 (N = 415 bis 1020)

-Retest-Reliabilität einer Kurzform liegt bei einem Zeitabstand von 15 bis 18 Monaten zwischen .36 und .91 (für Subskalen und Teilstichproben) bzw. zwischen .70 und .87 (für Summenskalen).
Validität:
Die Skalen-Interkorrelationen können das Hexagon zum Teil bestätigen, insgesamt resultiert jedoch eher die Form von drei Clustern (S-A, E-C, R-I). Mittelwertsvergleiche zeigen deutliche Geschlechtsunterschiede. Faktoranalysen auf Subskalen- und Itemebene bestätigen die 6 Faktoren und die Struktur.
Kriteriumsbezogene Validierungsstudien
zeigen die erwarteten Zusammenhänge mit anderen Tests (AIST: konvergente Validität = .68, divergente Validität = .19; NEO-FFI: z.B. Enterprising & Extraversion: .44, Artistic & Offenheit für Erfahrung: .47, Social & Verträglichkeit: .36). Ausführliche Darstellungen zur Frage nach der Wirkung und der Akzeptanz zeigen positive Ergebnisse (z.B. 76 % passende Berufe gefunden, 76 % Bestätigung für bisherige Pläne, 68 % neue Ideen/mehr Möglichkeiten, hohe Übereinstimmung von Testergebnis und Berufswunsch. Das Ausfüllen des Testheftes war für die Mehrzahl interessant, klar verständlich und einfach).
Normen
Die Auswertung von EXPLORIX erfolgt auf der Basis von Rohwerten und verzichtet auf Normen von Eichstichproben.
Begründet wird dies mit der identischen Skalierung, der relativ ähnlichen Attraktivität der Skalen, der Anwendung von Berufscodes und der Problematik der angemessenen Referenzgruppe bei Persönlichkeitstests.
Die Konstruktionsstichprobe beruht auf insgesamt 1815 Personen der Länder Deutschland, Österreich und Schweiz.

D2 test: Aufemrksamkeitsbelastungstest

Der Test d2 stellt eine standardisierte Weiterentwicklung der sogenannten Durchstreichtests dar. Er mißt Tempo und Sorgfalt des Arbeitsverhaltens bei der Unterscheidung ähnlicher visueller Reize (Detail-Diskrimination) und ermöglicht damit die Beurteilung individueller Aufmerksamkeits- und Konzentrationsleistungen.
Mit Aufmerksamkeit und Konzentration meint der Autor das Vermögen eines Individuums, sich bestimmten (aufgaben-)relevanten internen oder externen Reizen selektiv, d.h. unter Abschirmung gegenüber irrelevanten Stimuli, ununterbrochen zuzuwenden und diese schnell und korrekt zu analysieren. Die Herausforderung des Tests basiert auf den Zeitdruck und der monotonen Wiederholung der Aufgaben. Das Testblatt zeigt 14 Zeilen zu je 47 Zeichen. Es gilt die Grundzeichen d und p sowie deren Markierungen (ein bis vier senkrechte Striche) zu unterscheiden. Durchzustreichen sind alle d mit zwei Strichen (targets). Sie sind eingestreut unter d mit mehr oder weniger als zwei Strichen sowie unter p mit ein bis vier Strichen, die alle unmarkiert zu belassen

sind. Jede Zeile enthält in unregelmässiger Abfolge 21 oder 22 Targets. Jede der 14 Zeilen wird als Minitest betrachtet, was eine Beobachtung des Leistungsverlaufs ermöglicht. Das Verfahren zählt einerseits die Menge der bearbeiteten Zeichen und andererseits die Qualität dieser Bearbeitung (Anzahl Fehler). Es existieren verschiedene Vorschläge, wie diese zwei Indikatoren in einem Messwert kombiniert werden sollen.Der qualitative Leistungsaspekt, die Leistungsgüte, ergibt sich aus dem Fehlerrohwert (F). Der besteht aus den Auslassungsfehlern (F1) und den Verwechslungsfehlern (F2), die addiert und prozentual auf die Leistungsmenge bezogen werden (Messwert F%).
Der seit 1994 neu eingeführte Konzentrationsleistungswert (KL) besteht aus der Summe aller zutreffend durchgestrichenen Zeichen, abzüglich der Verwechslungsfehler (F2). KL entspricht besser dem Sinn und Geist des Tests und seiner Anweisungen als der früher übliche Messwert (GZ - 2F).
Die vielfältige Absicherung der Testgütekriterien (Objektivität, Zuverlässigkeit und Gültigkeit), die einfache Anwendung, der geringe Aufwand an Zeit und Material sind Ursache dafür, daß der Test d2 zu den am häufigsten verwendeten psychodiagnostischen Verfahren gehört und auch im Ausland weite Verbreitung gefunden hat. Inzwischen liegt das Verfahren in vielen Sprachen vor.
Einsatz:
9 bis 59 Jahre, Einzel- und Gruppentest. Verwendung in Berufsberatung und Eignungsdiagnostik, in der Arbeits-, Betriebs- und Verkehrspsychologie, in der Klinischen Psychologie, Schulpsychologie und Erziehungsberatung sowie als Forschungsinstrument (z.B. in der Pharmakopsychologie).
Zuverlässigkeit
Die Werte GZ-F (einfach korrelierte Mengenleistung) und GZ-2F (doppelt korrigierte Mengenleistung) und KL (Konzentrationsleistung) korrelieren sehr hoch (Mittelwert r > .93). Erwartungsgemäß sehr hoch ist auch der Zusammenhang zwischen absolutem und relativem Fehlerwert (F, F%). Hinsichtlich der internen Konsistenz erweist sich der d2 in zahlreichen Untersuchungen als hoch reliabel (r > .90); auch die Stabilitätskoeffizienten sind sehr zufriedenstellend.
Gültigkeit
Konstruktvalidität, faktorielle Validität wie auch empirische Validität sind durch eine große Zahl von Untersuchungen belegt. Wer Zweifel an der Konzentrationsfähigkeit eines Probanden hat, wer vor schwer erklärbaren Schul- oder Testleistungen steht, wird nützliche Hinweise (Schnelligkeit, Arbeitsqualität, Ausdauer, Lernkurve) aus einem d2 erfahren. Für viele Lebens- und Lernsituationen ist eine normale Konzentration, wie im d2 erforderlich, Vorbedingung zum Erfolg. Fast die Hälfte der Handanweisung sind der Validität des Verfahrens gewidmet und bilden somit die Basis für eine sachgerechte Interpretation. Bei normalen Probanden sind die Korrelationen der gemessenen Konzentrationswerte mit Intelligenztestleistungen eher positiv aber gering, bei beeinträchtigten Testpersonen viel deutlicher, wodurch das Konstrukt einer "für höhere Leistungen erforderliche Grundhaltung" unterstützt wird.

Normen
Geschlechts-, altersgruppen- und schulspezifische Normen für den Altersbereich von 9 bis 20 Jahren (N = 3132) sowie Normen für Erwachsene von 19 bis 59 Jahren (N = 3000).
Bearbeitungsdauer maximal 15 Minuten, die reine Testzeit liegt bei 4 Minuten und 40 Sekunden.
Der Proband soll unter Zeitdruck aus einer Vielzahl sich mehr oder weniger ähnlicher Reize die per Instruktion als relevant definierten (=Zielitems) anstreichen, die irrelevanten dagegen nicht. In einigen Verfahren werden Indices für den Leistungsverlauf bestimmt; bei der Bearbeitung solcher Tests werden die Probanden daher aufgefordert, jeweils nach Ablauf von Bruchteilen der Gesamtbearbeitungszeit mit der Bearbeitung neurer Testabschnitte zu beginnen.

Problem:
Zwischen den Zielitems verbleiben unmarkierte Abschnitte, die keinen eindeutigen Rückschluss auf ein konzentriertes Arbeitsverhalten erlauben.

Abhilfe:
Zur Überwindung dieses Problems wurde von Moosbrugger das vollständige Markierungsprinzip vorgeschlagen, bei dem sowohl die Zielitems als auch die Nicht – Zielitems zu markieren sind, jedoch auf unterschiedliche Weise.

Fakt II

Frankfurter Adaptiver Konzentrationsleistungs-Test

2., überarbeitete und neu normierte Auflage des FAKT von Moosbrugger und Heyden (1997)

H. Moosbrugger und F. Goldhammer

EINSATZBEREICH:

Der FAKT-II dient der adaptiven Ermittlung der individuellen Konzentrationsfähigkeit. Die Anwendungsbereiche des FAKT-II liegen in der Eignungsdiagnostik, der Arbeits- und Betriebspsychologie, Klinischen Psychologie, Schulpsychologie, Pädagogischen Psychologie, Verkehrspsychologie, Sportpsychologie, im neuro- und pharmakopsychologischen Bereich, in der Psychiatrie sowie in der Grundlagenforschung.

DAS VERFAHREN:

Der Frankfurter Adaptive Konzentrationsleistungs-Test in der 2. Auflage ist die grundlegend neu bearbeitete und neu normierte Realisierung eines seit 1997 bewährten computerbasierten Konzepts zur adaptiven Messung der individuellen Konzentrationsfähigkeit. Zur Bestimmung der Konzentrationsfähigkeit werden die Aspekte

Konzentrations-Leistung (KL)
Konzentrations-Genauigkeit (KG)
und Konzentrations-Homogenität (KH) erfasst.
Es existieren drei Testformen (FAKT-E, FAKT-S und FAKT-SR), welche die Variation des situativen Belastungsgrades erlauben und sich in der Darbietungsweise der Items unterscheiden. FAKT-E bietet nur jeweils ein Item auf dem Bildschirm dar (hohe situative Belastung), FAKT-S gleichzeitig 10 Items (mittlere situative Belastung). Die Testform FAKT-SR erhebt Reaktionszeiten für die Itembearbeitung (niedrige situative Belastung). Zu jeder Testform kann zusätzlich zwischen den zwei Durchführungsarten »adaptive Testzeit« oder »wählbare Testzeit« gewählt werden, wobei 6 Minuten als Standardtestzeit voreingestellt sind. Die Auswertung des Tests erfolgt automatisch und wird entweder auf dem Bildschirm oder auf dem Drucker ausgegeben. Die Testergebnisse können für den wissenschaftlichen Einsatz exportiert und mit Statistik-Programmen weiterverarbeitet werden.

ZUVERLÄSSIGKEIT:

Für die Konzentrations-Leistung wurden in einer Reihe von Untersuchungen interne Konsistenzen (Cronbachs Alpha) zwischen $\alpha = .91$ und $\alpha = .97$ gefunden, die Retest-Reliabilitäten (bei verschiedenen zeitlichen Abständen) liegen zwischen $rtt = .70$ und $rtt = .96$. Auch die Konzentrations-Genauigkeit weist relativ hohe Cronbachs Alpha-Werte zwischen $\alpha = .68$ und $\alpha = .91$ auf. Für die Konzentrations-Homogenität wurden interne Konsistenzen zwischen $\alpha = .77$ und $\alpha = .93$ ermittelt.

GÜLTIGKEIT:

Die Konstruktvalidität des FAKT-II kann einerseits durch das einheitliche Itemkonstruktionsprinzip und andererseits durch Modellkonformität mit dem polytomen Rasch-Modell als gegeben betrachtet werden. Sowohl die diskriminante als auch die konvergente Validität zu anderen Testverfahren wurden geprüft und können als zufriedenstellend bezeichnet werden. Durch den Einsatz der adaptiven Testformen wird die verschiedentlich beobachtbare Konfundierung von Konzentrationsleistung und Aktivierung, Ärger, Prüfungsangst, negativem Affekt und Lärmeinfluss beseitigt. Zudem fallen situative Einflüsse auf die Messungen nach zwei vorausgehenden Übungsdurchläufen nur noch gering aus. Die Messungen gelten somit als transsituativ konsistent und können als Merkmalsmessung betrachtet werden.

NORMEN:

Die aktualisierten Normen stammen aus den Jahren 1996 bis 2003 und basieren auf einer Normierungsstichprobe von N = 859 Probanden beiderlei Geschlechts zwischen 16 und 55 Jahren. Die Anwendbarkeit des FAKT-II ab einem Alter von 10 Jahren sowie bei Senioren ist erprobt. Es liegen differenzierte Testnormen für die drei Testformen, die adaptive Testzeit, die Standardtestzeit,

die Anzahl vorangegangener Testungen, verschiedene Testlängen und verschiedene Längen der Übungsphase vor.

BEARBEITUNGSDAUER:

etwa 10 Minuten, reine Testdauer 6 Minuten;Zur Beurteilung des Leistungsverlaufes kann die Testlänge in 6-Minuten-Schritten auf 30 Minuten ausgedehnt werden.

NEO-PI-R NEO-PI-R
NEO-Persönlichkeitsinventar nach Costa & McCrae
deutsch: Ostendorf und Angleitner, 2004
• Theoretische Grundlagen
 Basierend auf den lexikalischen „Big Five“ sowie auf Clusteranalysen des 16PF
(Cattell)
→ Clusteranalysen sind der eigentliche Ursprung des NEO-PI-R
 5 Domains: „vielgestaltige Ansammlungen von spezifischen Kognitionen,
Affekten und von Verhaltenstendenzen, die unterschiedlich gruppierbar sind“
(Costa & McCrae, 1995).
 Je Domain 6 Facetten. Rationale Entscheidung. Konsistent mit der Literatur und
empirisch aufweisbar.
→ die 6 Facetten entstammen nicht empirischen Analysen (sonst hätte man
nicht gleich viele Facetten pro Domain)
 Fünfstufiges Antwortformat
→ es gibt eine Mittelkategorie (hat sowohl Vor- als auch Nachteile, es gibt keine
empirische Begründung dafür, welches besser ist)
• Persönlichkeitsbereiche und Facetten des NEO-PI-R
→ Neurotizismus: geringes Selbstwertgefühl fehlt, ist aber indiziert
→ Offenheit für Erfahrungen nicht mit Offenheitsskalen verwechseln
(Offenheitsskalen sollen Lügen herausfinden, Offenheit im NEO-PI-R soll das
Persönlichkeitsmerkmal erfassen, z.B. Offenheit für Ideen, Offenheit des Werteund
Normensystems → negative Korrelation der Offenheit des Werte- und Normensystems mit Konservatismus)

NEO-Persönlichkeitsinventar nach Costa und McCrae, Revidierte Fassung

1. Auflage

von F. Ostendorf und A. Angleitner

EINSATZBEREICH:

Ab 16 Jahre; Einsatz in vielfältigen Bereichen der Forschung und Praxis, wie z.B. in der Klinischen Psychologie, der Psychiatrie, der Arbeits- und Organisationspsychologie, der Berufsberatung oder der Pädagogischen Forschung. Selbst- und Fremdbeurteilung.

DAS VERFAHREN:

Dieser weltweit in der Forschung und klinischen Praxis wohl am häufigsten eingesetzte Fragebogen zur Messung des Fünf-Faktoren-Modells der Persönlichkeit basiert auf den Ergebnissen jahrzehntelanger faktorenanalytischer Forschung mit umfangreichen Bevölkerungsstichproben und klinischen Probandengruppen. Das revidierte NEO-Persönlichkeitsinventar (NEO-PI-R) erfasst mit 240 Items die Hauptbereiche interindividueller Persönlichkeitsunterschiede. Darüber hinaus besteht die Möglichkeit einer differenzierten Messung der Hauptskalen
Neurotizismus
Extraversion
Offenheit für Erfahrung
Verträglichkeit
und Gewissenhaftigkeit
durch insgesamt 30 Fassetten. Im Vergleich zum NEO-Fünf-Faktoren-Inventar (NEO-FFI) ermöglicht das NEO-PI-R somit eine vollständige Erfassung des Fünf-Faktoren-Modells. Der Vorteil des NEO PI-R liegt in seiner großen Bandbreite. Das Inventar ermöglicht somit eine umfassende und zugleich detaillierte Persönlichkeitsbeschreibung. Auch für die Erhebung von Fremdbeurteilungen, z.B. durch Familienangehörige oder Klinikpersonal, steht eine entsprechende Form (Form F) zur Verfügung.

ZUVERLÄSSIGKEIT:

Selbstbeurteilungsform (Form S):
Interne Konsistenz der fünf Hauptskalen .87 - .92
der Fassetten .53 - .85.
Retestreliabilität der Skalen nach einem Monat für die fünf Hauptskalen im Bereich von .88 - .91
für die 30 Fassetten .48 - .91 (Median: .83).
Nach fünf Jahren: .74 - .78 (Hauptskalen) bzw. .53 - .78 (Fassetten).

Fremdbeurteilungsform (Form F):
Interne Konsistenz der Hauptskalen .89 - .95
der Fassetten .49 - .84 (Mittelwert .75).

GÜLTIGKEIT:

Die Kongruenz der deutschen und amerikanischen varimax-rotierten Faktoren

der NEO-PI-R-Form S liegt im Bereich von .96 - .98, die der varimax-rotierten Faktoren der deutschen Formen S und F zwischen .98 und 1. Zahlreiche Untersuchungsergebnisse zur konvergenten und diskriminanten Validität liegen vor. Das Manual enthält u.a. Angaben über die Beziehungen der NEO-PI-R-Skalen zu verschiedenen Persönlichkeitstests (GT, TPF, FPI-R, BIP, LMI, MBTI, FRBS, 16 PF-R, EPQ-RK, PRF), Maßen kognitiver Kontrolle und habitueller Affektlagen (FKK, DC-D, PANAS), Intelligenzmaßen (BOTE, WMT, WST, LPS), Interessen (AIST) und berufsbezogenen Merkmalen (Testsystem pro facts).

NORMEN:

T-, Stanine- und Prozentrangnormen für die Hauptskalen und Fassetten der NEO-PI-R-Formen S und F (Gesamtstichprobe: N = 11.724). Bevölkerungsrepräsentative sowie Alters- (16 bis über 50 Jahre), geschlechts- und z.T. bildungsspezifische Normierung.

BEARBEITUNGSDAUER:

ca. 35 Minuten. Etwa 90% der Probanden bearbeiten die Items des Inventars innerhalb von 45 Minuten.

Leistungsmotivationsinventar LMI
Schuler, H. & Prochaska, M (2001). Gottingen: Hogrefe

Die Autoren verfolgten bei der Entwicklung des LMI das Ziel, ein Verfahren zur breiten Messung der berufsbezogenen Leistungsmotivation bereit zu stellen. Sie kamen bei ihren Vorarbeiten zu der Erkenntnis, dass es sich bei der Leistungsmotivation um ein breites Konzept ohne scharfe Grenzen zu handeln scheint. Diesen Zustand charakterisieren sie durch ihr Zwiebelmodell: Einige Merkmale sind zentral fur die Leistungsmotivation und bilden die Kernfacette. Dazu gehoren etwa Beharrlichkeit und Erfolgshoffnung. Andere wie z. B. Selbststandigkeit und Statusorientierung liegen weiter in der Peripherie (Randfacetten). Noch weiter in der Peripherie liegen Merkmale wie etwa Selbstvertrauen und Kontrolluberzeugung, die der Leistungsmotivation zumindest theoretisch verbunden sind. Im auseren Randbereich schlieslich sind Merkmale wie Gewissenhaftigkeit und Neurotizismus angesiedelt, die als Hintergrundmerkmale einen Einfluss auf die Leistungsmotivation ausuben. Vor dem Hintergrund dieser Konzeption ist es nicht verwunderlich, dass die Autoren insgesamt 728 Items in die Vorauswahl nahmen. Diese Zahl reduzierte sich allerdings in mehreren Auswahlschritten.
Die Autoren beschreiben ihr Vorgehen als Wechsel von phanomenologischrationaler und empirischer Strategie (S. 12). Am Ende

sahen sie 17 Dimensionen als angemessen zur Beschreibung der beruflichen Leistungsmotivation an.

Gliederung. Das LMI umfasst 17 Skalen mit je zehn Items (. Tab. 3.20). Die Items
konnen jedoch auch zu einem Gesamtwert verrechnet werden. Anwender, die nicht an einem differenzierten Personlichkeitsbild ihrer Probanden interessiert sind, sondern nur □die□ Leistungsmotivation messen wollen, konnen die Kurzskala mit 30 Items einsetzen. Die Items wurden nach ihren Trennscharfen fur den Gesamtwert aus der Langform herausgefiltert.

Durchfuhrung und Auswertung. Das LMI kann einzeln oder in Gruppensitzungen
durchgefuhrt werden. Die Bearbeitung der 170 Items, die auf einer Skala von 1 (trifft
gar nicht zu) bis 7 (trifft vollstandig zu) zu beantworten sind, nimmt etwa 30 – 40 min
in Anspruch; fur die Kurzform genugen etwa 10 min. Die Auswertung der Langform
mit Schablonen ist auserordentlich muhsam, weil zehn Seiten des Testhefts durchzusehen, bei einigen Items Invertierungen vorzunehmen, die 170 Zahlenwerte auf einem Auswertungsbogen einzutragen und zu addieren sind.

Reliabilitat. Die interne Konsistenz (α) der Skalen liegt uberwiegend im Bereich von
.80. Solche Werte sind fur eng umschriebene Merkmale bei 10 Items pro Skala vollig
angemessen. Fur die Kurzversion mit 30 Items betragt α =. 94. Die Retestreliabitat
nach einem Intervall von etwa drei Monaten wird mit .66 (Flow) bis .82 (Furchtlosigkeit und Statusorientierung) (Kurzform: .78) angegeben. Offenbar erfasst das LMI stabile Merkmale.

Validitat. Bei einem Personlichkeitsfragebogen mit 17 Skalen stellt sich die Frage nach der faktoriellen Struktur. Die Autoren berichten Korrelationen zu den funf Skalen des NEO-FFI. Inzwischen liegt eine Untersuchung an 121 Sportstudenten vor, in der die Skalen des LMI zusammen mit denen des NEO-PI-R (7 Abschn. 3.2.3) faktorisiert wurden (Ostendorf & Angleitner, 2004). Die Ergebnisse sind in . Tab. 3.20 aufgefuhrt. Sie belegen, dass sich die LMI-Skalen sehr gut in das □Big-Five□-Modell einordnen lassen, also offenbar verschiedene Facetten von Neurotizismus, Extraversion etc. erfassen. Vor diesem Hintergrund interessiert die Korrelation der Gesamtskala, die ein Konglomerat der 17 Skalen darstellt, mit den grosen Personlichkeitsfaktoren.

Schuler und Prochaska (2001) berichten signifikante Korrelationen von .57 mit Gewissenhaftigkeit, –.40 mit Neurotizismus und .23 mit Extraversion.
Die Befunde zur Kriteriumsvaliditat sind relativ unubersichtlich, weil die Autoren
zahlreiche Korrelationen mit unterschiedlichen Leistungsindikatoren wie Abiturnoten,
Note des Ausbildungsabschlusses und Jahresgehalt berichten. Viele Korrelationen
sind insignifikant und nur wenige liegen uber .30. Einzelne herausragend hohe Korrelationen erscheinen nachtraglich plausibel, so die zwischen Dominanz und Stellung in der Hierarchie (r = .43) oder die zwischen Lernbereitschaft und Bildungsniveau (r = .35).

Normen. Es liegen (zum Teil geschlechtsspezifische) Normen fur Wirtschaftsstudenten (N = 259), Berufsschuler in kaufmannischen Ausbildungsberufen (N = 1008), Schuler eines Wirtschaftsgymnasiums (N = 160), Berufstatige in Finanzdienstleistungsunternehmen (N = 166) und Hochleistungssportler (N = 78) vor. Daraus konstruieren die
Autoren zusatzlich eine nicht reprasentative Gesamtnorm (auch getrennt fur Manner und Frauen).

Bewertung. Das LMI ist ein objektives und hinreichend reliables Verfahren zur Messung der Leistungsmotivation. Die Kriteriumsvaliditat in den vorgesehenen Anwendungsbereichen Personalauswahl und -entwicklung, Schul-, Studien- und Berufsberatung sowie Sportpsychologie ist noch umfassender zu belegen. Die starke Differenzierung in 17 Dimensionen erweist sich zugleich als Chance und Problem (Schmidt-Atzert, 2001, S. 144). Die vielen Skalen eroffnen die Chance, in bestimmten Anwendungsfeldern gute Einzelpradiktoren beispielsweise fur berufliche Leistungen zu finden. Fur eine Profilauswertung ist die grose Zahl von zum Teil erheblich korrelierten Skalen hinderlich.

WIT II

Wilde-Intelligenz-Test 2

1. Auflage

M. Kersting, K. Althoff und A.O. Jäger

Hrsg. von , Deutschen Gesellschaft für Personalwesen e.V.

EINSATZBEREICH:

Der WIT-2 dient der differenzierten Erfassung klar unterscheidbarer und theoretisch fundierter kognitiver Fähigkeiten von Jugendlichen und Erwachsenen.
Seine Differenzierungsfähigkeit ist für den Altersbereich von 14 bis 42 Jahren

nachgewiesen. Der WIT-2 eignet sich für die Intelligenzdiagnostik im Allgemeinen sowie für die Forschung und Eignungsdiagnostik im Besonderen. Als Anwendungsbereiche innerhalb der Eignungsdiagnostik sind die Personalauswahl und Personalentwicklung ebenso zu nennen wie die Berufs- / Studienwahl und die Berufs- / Studienlaufbahnplanung. Der WIT-2 erlaubt eine treffsichere Vorhersage des Erfolgs in Schule, Ausbildung, Studium und Beruf.

DAS VERFAHREN:

Als theoretischer Ausgangspunkt der Konstruktion des »neuen« Wilde Tests galt das modifizierte Modell der Primary Mental Abilities (MMPMA). Der neue WIT-2 repräsentiert mit elf Subtests (darunter eine Arbeitsprobe)
8 Dimensionen:
(1) Schlussfolgerndes Denken,
(2) Sprachliches Denken,
(3) Rechnerisches Denken,
(4) Räumliches Denken,
(5) Merkfähigkeit,
(6) Arbeitseffizienz,
(7) Wissen Wirtschaft
(8) Wissen Informationstechnologie
Die Denkaufgaben des Tests fokussieren nicht so sehr das abstrakte Denkvermögen, sondern berufliche Schlüsselqualifikationen und Grundfertigkeiten. Im Gegensatz zu einigen anderen Tests erfasst der WIT-2 die Merkfähigkeit und nicht die unmittelbare Behaltensleistung. Die acht WIT-2 Module können einzeln oder in beliebiger Kombination miteinander eingesetzt werden, so dass sich ein Höchstmaß an Flexibilität und Zeitökonomie ergibt. Der WIT-2 ist ein Test aus der Praxis für die Praxis. Die fehlerfreie und komfortable Instruktion wird beispielsweise durch ein separates Instruktionsheft gewährleistet. Andere Beispiele für die Anwendungsorientierung des WIT-2 sind die Informationsbroschüre für die Testteilnehmer, die anforderungsanalytisch gewichtete Gesamtwertbildung sowie die umfassenden Darstellungen zur Auswertung und Interpretation. Die Verfahrenshinweise zum WIT-2 orientieren sich explizit an den Anforderungen der DIN 33430. Der WIT-2 erfreut sich bei den Testteilnehmern nachweislich einer hohen Akzeptanz.

ZUVERLÄSSIGKEIT:

Die internen Konsistenzen liegen in einem sehr zufrieden stellenden Bereich von $\alpha = .77$ bis $\alpha = .98$.

GÜLTIGKEIT:

Zu kaum einem anderen Test dürften bei der Erstpublikation derart umfassende Validitätsinformationen vorliegen. Dies betrifft beispielsweise die Einbettung in das nomologische Netz anderer Fähigkeiten (Konstruktvalidierung). Zu nennen

sind hier zahlreiche Analysen zur konvergenten und diskriminanten Validität durch Bestimmung der Zusammenhänge zu 17 anderen Testverfahren. Hinsichtlich der Kriteriumsvalidität wurden zu jedem der sieben Subtests, die aus dem »alten« WIT in modifizierter Form übernommen wurden, die bisherigen Ergebnisse zur Kriteriumsvalidität metaanalytisch zusammengefasst (pro Aufgabe 11 bis 24 Studien im Umfang von 698 bis 2.039 Personen). Der Nachweis der Kriteriumsvalidität für die Aufgaben des WIT-2 wurde in mehreren Studien anhand von Daten aus der Praxis erbracht.

NORMEN:

einzelne Aufgaben des WIT-2 bei annähernd 42.000 Personen eingesetzt. der überwiegende Teil der Daten für den WIT-2 wurde im Kontext des Ernstfalls von beruflichen Bewerbungssituationen erhoben

BEARBEITUNGSDAUER:

variiert in Abhängigkeit vom Umfang der eingesetzten Module zwischen 6 Minuten (Wissenstest Wirtschaft) und 143 Minuten (alle acht Module inkl. allgemeiner Einführung und einer Pause). Für die Messung des schlussfolgernden Denkens (Modul 4) werden 50 Minuten benötigt (inkl. Einführung und Instruktion).

Minnesota-Multiphasic-Personality-Inventory-MMPI-2

Hathaway, S. R., McKinley, J. C., Engel, R. R. (2000). MMPI-2. Bern: Huber

Konstruktionsansatz. Am Anfang aller Entwicklungsarbeiten stand die Anlegung einer Liste von 1000 Items, die sich auf allgemeine Gesundheit, familiare und eheliche Beziehungen, sexuelle und religiose Einstellungen sowie emotionale Zustande bezogen und letztlich psychopathologische Symptome erfassen sollten. Spater kamen Items zu Geschlechtsrollencharakteristika und abwehrender Selbstdarstellung dazu. Gruppen von klinisch auffalligen Personen, die von Psychiatern als Schizophrene, Hysteriker, Hypochonder usw. diagnostiziert worden waren, bearbeiteten die Items ebenso wie □unauffallig-normale□ Kontrollpersonen (Einwohner von Minnesota, Bewerber um einen Studienplatz, Besucher des Krankenhauses). Jene 566 Items wurden schlieslich in Skalen zusammengestellt, die die Patienten von den Kontrollpersonen signifikant diskriminierten. Von daher eignet sich das Verfahren vorrangig zur Entscheidungsfindung uber die Zuordnung von Probanden zu entweder psychiatrischen Kategorien oder zur unauffalligen □Normal□-Population.

Gliederung. Das MMPI wie auch die revidierte Form MMPI-2 umfasst vier Validitatsskalen
(?, L, F, K) und 13 Basisskalen. Daruber hinaus konnen aus den Items zahlreiche

Zusatzskalen gebildet werden (z. B. Soziale Verantwortlichkeit, Posttraumatische Belastungsstorung, Suchtgefahrdung).

Durchfuhrung. Das MMPI-2 kann in Einzel- und Gruppensituation durchgefuhrt werden. Die Bearbeitung dauert gut 1 h, bei Patienten etwas langer. Der Proband kreuzt auf einem separaten Auswertungsblatt fur jedes Item Richtig oder Falsch an.

Auswertung. Die Rohwerte werden skalenweise mit Schablonen ermittelt. Die individuellen Rohwerte werden direkt in ein Profilblatt fur Frauen oder Manner eingetragen. Darin sind die Rohwerte bei jeder Skala grafisch so angeordnet, dass praktisch eine Transformation in T-Werte erfolgt. Bei einigen Skalen sind zuvor die Rohwerte um eine bestimmte Punktzahl zu erhohen. Bei dieser so genannten K-Korrektur wird der Rohwert einer Basisskala fur mangelnde Offenheit des Probanden, die sich aus der K-Skala ablesen lasst, nach oben korrigiert. Dazu wird der Punktwert der K-Skala mit dem angegebenen Faktor, z. B. 0.4, multipliziert. Der resultierende Korrekturwert wird zum Rohwert der Basisskala addiert. Die muhsame manuelle Auswertung lasst sich bei Nutzung des kostenpflichtigen Fax-Services umgehen. Das Antwortblatt wird an die angegebene Adresse gefaxt und nach wenigen Minuten kommt die komplette Auswertung zuruck. Als nicht wirklich attraktive Alternative bietet sich ein Auswertungsprogramm an, bei dem die Items eingegeben werden mussen. Sehr komfortabel ist dagegen die
Computerversion des MMPI-2, bei der die Auswertung per Mausklick zu haben ist.

Interpretation. Zuerst wird anhand der Validitatsskalen gepruft, ob das Protokoll gultig ist. Die Basisskalen konnen einzeln interpretiert werden. Dazu stehen im Manual Interpretationshinweise nach dem in . Tabelle 3.13 gezeigten Schema zur Verfugung.

Normierung. Die Eichstichprobe (N = 958) zur Normierung des deutschen MMPI-2
ist bezuglich Alter, Geschlecht und geographischer Herkunft reprasentativ fur die 18–
70jahrige deutsche Bevolkerung. Bei der Erbhebung entstand ein gewisser Schwund
dadurch, dass immerhin 192 Personen die Bearbeitung des MMPI ablehnten und Protokolle wegen extrem hoher F-Werte oder zu vielen unbeantworteten Items eliminiert werden mussten.

Reliabilitat. Die Retestreliabilitat nach 10 Tagen der Skalen wird im Manual mit .66

(Ma) bis .90 (D) bei Mannern (N = 49) und .71 (Ma) bis .92 (Sc) bei Frauen (N = 56)
angegeben.

Validitat. Im Testmanual werden lediglich Faktorenanalysen der Skalen berichtet. Es
fanden sich ahnliche Strukturen fur Manner und Frauen sowie gute Ubereinstimmung
mit US-Ergebnissen. Die vier Faktoren sind:
4 F1: Psychotische Gedankeninhalte (Sc, Pp, Pa, F)
4 F2: Neurotische Verhaltensweisen (Hy, L, K)
4 F3: Introversion (Si, D)
4 F4: Geschlechtsrollenidentifikation (Mf)
Wegen der Itemuberlappung sind die Ergebnisse der Faktorenanalysen mit Vorsicht zu bewerten.

Bewertung. Es ist sehr zu begrusen, dass ein so bewahrtes und gut erforschtes Verfahren wie das MMPI nun auch in der aktualisierten Form (MMPI-2) in deutscher Sprache vorliegt. Positiv zu werten ist die extrem grose Informationsausbeute durch die vielen klinischen Skalen, Validitatsskalen und Zusatzskalen. Zur fehleranfalligen und zeitraubenden Auswertung mit Schablonen stehen attraktive Alternativen zur Verfugung. Am Manual ist zu bemangeln, dass Angaben zur Profilreliabiltat fehlen und die zur Validitat durftig sind. Ein grundsatzliches Problem des MMPI ist, dass das Verfahren auf veralteten diagnostischen Kriterien basiert und eine klinische Diagnostik nach ICD-10 oder DSM-IV nicht unterstutzt.