

*Berufs-Interessen-
Test B-I-T.-II
(von Irle, M. & Allehoff, W.)*

Jörg-Tobias Kuhn

XXX

XXX

XXX

Alexander Zimmerhofer

XXX

XXX

XXX

XXX

<u>KURZE INFORMATION ZUR KONSTRUKTIONSGESCHICHTE UND INFORMATIONEN ZU DEN AUTOREN.....</u>	<u>3</u>
<u>TESTGRUNDLAGE</u>	<u>4</u>
<u>DIAGNOSTISCHE ZIELSETZUNG</u>	<u>4</u>
<u>THEORETISCHE GRUNDLAGE</u>	<u>6</u>
<u>(NACHVOLLZIEHBARKEIT DER) TESTKONSTRUKTION</u>	<u>7</u>
<u>TESTBESCHREIBUNG.....</u>	<u>8</u>
<u>VERHALTENS PROVOKATION.....</u>	<u>8</u>
<u>AUSWERTUNG.....</u>	<u>10</u>
<u>INTERPRETATION (INS. NORMIERUNG)</u>	<u>11</u>
<u>GÜTEKRITERIEN</u>	<u>12</u>
<u>OBJEKTIVITÄT.....</u>	<u>12</u>
<u>RELIABILITÄT</u>	<u>12</u>
<u>VALIDITÄT / EVALUATION.....</u>	<u>15</u>
<u>NEBENGÜTEKRITERIEN</u>	<u>19</u>
<u>KOSTEN</u>	<u>19</u>
<u>ANWENDUNGSEMPFEHLUNG</u>	<u>19</u>
<u>BEWERTUNG</u>	<u>20</u>
<u>ANHANG.....</u>	<u>23</u>
<u>AUSWERTUNGSHILFE</u>	<u>23</u>
<u>LITERATUR</u>	<u>24</u>

Kurze Information zur Konstruktionsgeschichte und Informationen zu den Autoren

Der Berufs-Interessen-Test (B-I-T.¹) in seiner ersten Version entstammt der Inauguraldissertation des Sozialpsychologen **PROFESSOR DR. MARTIN IRLE** (* 26.1.1927 in Witten/ Ruhr). Nach seinem Diplom der Psychologie im Jahre 1952 entwickelte er 1953 und 1954 das Testverfahren bei der Deutschen Gesellschaft für Personalwesen e.V., deren (psychologische) Abteilung er in Frankfurt am Main leitete. Ein Jahr darauf erhielt er von der Universität Göttingen die Doktorwürde. Bis zu seiner Habilitation im Jahre 1962 beschäftigte er sich besonders mit der Organisations- und der Gemeindeforschung. Erst 1964 wendete er sich der Sozialpsychologie zu, die er als Ordinarius der Universität Mannheim entscheidend mitprägen sollte. Unter den wichtigsten Veröffentlichungen zählen das „Lehrbuch der Sozialpsychologie“ (1975), „Das Instrument der Täuschung in der Verhaltens- und der sozialwissenschaftlichen Forschung“ (1979) und „Die Enzyklopädie der Psychologie: Marktpsychologie als Sozialwissenschaft“ (1983). Besonders im Bereich der Dissonanz- und Reaktanz-Forschung veröffentlichte er wegweisende Schriften. 1984 erschien in Zusammenarbeit mit Dr. Wolfgang Allehoff die revidierte Fassung des B-I-T. Heute ist Martin Irle emeritiert und Ehrenmitglied der Deutschen Gesellschaft für Psychologie.

DR. WOLFGANG ALLEHOFF (* 1949) studierte Psychologie, Medizin und Kybernetik in Heidelberg, Uppsala und Stockholm und beschäftigte sich im Studium besonders mit der empirischen Sozialforschung.

Der langjährige Prognos-Mitarbeiter mit Schwerpunkt psychischer und sozialer Perspektiven der Freiraumplanung und -gestaltung arbeitete – nach der Mithilfe am B-I-T.-II – am Zentralinstitut für seelische Gesundheit in Mannheim (*einer psychosomatischen Klinik*) und bei der Kodak AG. Heute ist er Mitarbeiter am Institut für Demoskopie, Allensbach.

Die erste frei veröffentlichte Version des B-I-T. beim Hogrefe-Verlag aus dem Jahr 1955 war eigentlich schon die dritte revidierte Version des Testverfahrens (*Irle, 1955 b*). Es bestand im Gegensatz zu dem heutigen B-I-T.-II nur aus forced-choice Bogen, wobei die als am positivsten bewerteten Tätigkeiten markiert werden sollten.

Bei der Entwicklung des B-I-T.-II (*veröffentlicht 1984*), gefördert von der Stiftung Volkswagenwerk, wurden im direkten Vergleich der Version aus dem Jahr 1955 folgendes modifiziert (*nach Irle & Allehoff, 1984*):

- Die Items wurden teilweise neu formuliert (so wurde z.B. aus „Stühle schreinern“ „In Schiffswerft Hochseejachten zimmern“).
- Pro Wahlentscheidung muss jetzt die höchste (markiert mit „+“) und die geringste Präferenz (markiert mit „-“) gekennzeichnet werden.
- Einführung der Free-choice-Technik mit je 81 Präferenz-Urteilen
- Die Zahl der Items der Forced-choice-Technik wurde halbiert (von 162 auf 81 vermindert).
- Für beide Techniken wurden Paralleltests eingeführt.
- Ferner wurde eine neue Normierungsstichprobe von insgesamt 4400 Personen erhoben.

Der B-I-T. gilt als eines der weitverbreitetsten Testverfahren in testdiagnostischen Untersuchungen und als wichtigster deutschsprachiger Interessentest (*Brambring, 1983*).

¹ Um präzise Aussagen treffen zu können, wird in diesem Text vom „B-I-T.“ (*damit ist das allgemeine Testverfahren gemeint*), vom „B-I-T.-I“ (*damit meinen die Autoren die erste veröffentlichte Fassung aus dem Jahr 1955; dieses Verfahren trägt eigentlich den Namen „B-I-T.“*) und vom „B-I-T.-II“ (*Testverfahren aus dem Jahr 1984*) gesprochen.

Testgrundlage

Diagnostische Zielsetzung

Interessentests sind - wie der B-I-T. als einer ihrer Vertreter - in vielerlei Hinsicht von den sogenannten Leistungstests verschieden. Die ersteren gehören nach Cronbach (1970), übrigens genauso wie die Persönlichkeitstests, zu den „typical performance tests“, während Testverfahren wie Intelligenztests zu den „maximum performance tests“ zählen. Bei „maximum performance tests“ wird versucht, die Leistungsgrenzen (*die maximale Leistung*) eines Testanden bzw. Bewerbers zu erkunden. Bei der Benutzung von „typical performance tests“ besteht eine ganz andere Absicht: Es werden bezeichnende, typische Leistungen bzw. Antworten der Personen verlangt, was nach Cronbach (1970) bedeutet, dass die Personen hinsichtlich der Interessentests die mit den Antworten verknüpften Gedanken häufig ausführen bzw. fühlen.

Irle (1955 a) beschreibt in seiner Inauguraldissertation, dass das Testverfahren „zur Beantwortung der Frage beitragen [soll], ob zur Ausübung eines Berufes nicht nur die erforderlichen Fähigkeiten [dies wäre mit einem „maximum performance test“ zu überprüfen; Anmerkung der Autoren], sondern auch genügend starke Motivationen und Antriebe vorhanden sind, um diese potentielle Eigenschaften im Beruf zu aktivieren.“

Nach Irle (1955 b) ist ein Zusammenspiel zwischen Leistungs- und Interessentest sinnvoll. Die reine Testung der Intelligenz scheint somit - seiner Meinung nach - kein guter alleiniger Prädiktor für den späteren Berufserfolg zu sein. Diese Formulierungen ähneln den Aussagen Schulers (1996), der eine Mehrgleisigkeit der Eignungsdiagnostik vorschlägt, in der neben den Fähigkeiten und dem Entwicklungspotential die Interessenkomponente nicht vernachlässigt werden sollte, um „Zufriedenheit und andere Aspekte des Wohlbefindens sowie das Verbleiben in der Organisation sicherzustellen“.

Auch wenn Irle (1955 a) in seiner Dissertationsschrift noch davon spricht, dass mit dem Testverfahren überprüft wird, ob die entsprechende Motivation für einen Berufserfolg vorhanden ist, sind die Aussage aus dem Jahre 1984 (Irle & Allehoff) schon vorsichtiger: Handlungsmotivationen können demnach nicht unbedingt mit dem Verfahren erkannt werden (Irle & Allehoff, 1984; Richter, 1997). Die herausgearbeiteten Interessen bedeuten nämlich nicht unbedingt, dass auch eine damit gleichgerichtete Motivation - später im Berufsalltag - sich entwickelt.

Schon dem Titel „Berufsinteressentest“ ist zu entnehmen, dass nicht die allgemeinen Interessen, z.B. Freizeit-Interessen oder auch Wissenspräferenzen, erfasst werden sollen. Vielmehr ist das Ziel des Tests, alleine die beruflichen Präferenzen offen zu legen (Irle & Allehoff, 1984). Mit dieser besonderen Betrachtung des beruflichen Aspektes distanzieren sich die Autoren klar von den Vorgehensweisen anderer Autoren, die ausgehend von z.B. der Lieblingsliteratur Rückschlüsse auf die allgemeine Interessenausprägungen vollziehen und dann damit auf berufliche Interessen schließen.

Die primäre Zielgruppe des Verfahrens wird von Irle & Allehoff (1984) bei Hauptschulabgängern und -abgängerinnen gesehen. Auch wenn Richter (1997) den Einsatz ab dem 13. Lebensjahr empfiehlt, fehlen bei Irle & Allehoff (1984) weitere Angaben zur Altersspanne. Durch eine Betrachtung der Normierungsstichprobe kann die obengenannte Aussage aber untermauert werden (*weitere Details zur Normierungsstichprobe siehe unten*).

Während in früheren Generationen die Wahl des Berufs häufig durch die Zugehörigkeit einer sozialen Schicht determiniert war, stehen heute Berufsanfänger vor einer unüberschaubaren Vielfalt von Ausbildungsgängen und Berufen. Somit ist nach Irle & Allehoff (1984) ein Einsatz des B-I-T. bei der Selektion und auch für die Klassifikation, besonders bei der Berufsberatung, angeraten.

Mit anderen Worten: Grundsätzlich gibt es zwei Möglichkeiten, den B-I-T. zu verwenden: entweder als Selektionsinstrument, wodurch also Bewerber für eine bestimmte Stelle ausgewählt werden können,

Berufs-Interessen-Test

die ein besonders hohes Interesse in einem bestimmten Bereich besitzen, andere dagegen werden abgelehnt. Oder aber als Klassifikationsinstrument, wobei mehrere Personen mehreren Arbeitsplätzen zugeordnet werden (Cronbach, 1970). Stehen z.B. zwei unterschiedliche Jobs – zusätzlich zur Ablehnung - zur Verfügung, so wäre eine Zuordnung aufgrund von zwei Testverfahren denkbar (*multivariat*). Abbildung 1 (ii) visualisiert dies. Bei der Selektion dagegen gibt es nur die Unterscheidung zwischen Ablehnen oder Annahme des Bewerbers. Dabei würde die Auswahl über die Ergebnisse in einem Testverfahren (*univariat*) erfolgen und gleichzeitig auch nur eine grundsätzliche Annahme ermöglichen, nicht aber die Zuordnung zu bestimmten Berufen.

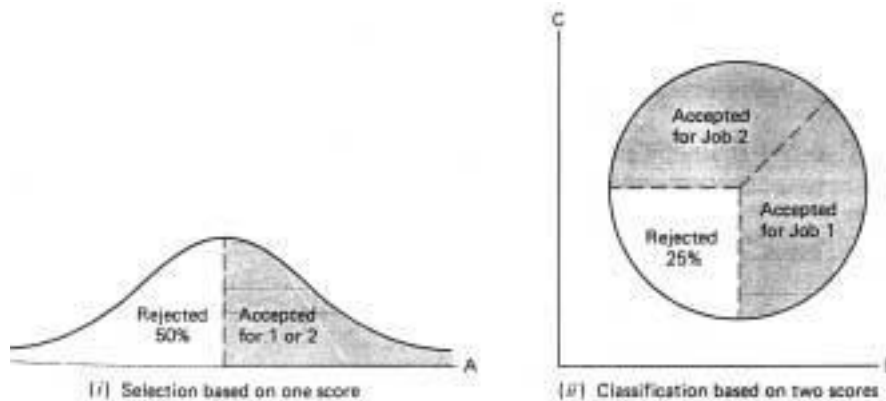


Abb.: 1 Die Unterscheidung von Selektion und Klassifikation (Cronbach, 1970)

Nach Amelang & Zielinski (1997) ist die oben genannte Klassifikation eine Zuweisung von Testanden bzw. Bewerbern entsprechend einer ermittelten Merkmalskonfiguration (*multivariater Ansatz*). Notwendige Abgrenzungs- bzw. Zuordnungskriterien und deren multiple Verknüpfungen werden aber von Irle und Allehoff nicht mitgeteilt (Richter, 1997). Klare Regeln und Kriterien werden für die Klassifikation den Testleitern also nicht an die Hand gegeben.

Teilt man das Testverfahren weiter ein, so gehört der B-I-T. nach den Zuordnungskriterien von Pawlik (1982) zur Statusdiagnostik. Prägnante Interessenrichtungen der Gegenwart werden herausgearbeitet und nach Irle (1955 b) und Irle & Allehoff (1984) mit einem prognostischem Zweck gesehen. Die Benutzung als Instrument zur Prozessdiagnostik wird von den Testautoren nicht erwähnt (Irle, 1955 a; Irle, 1955 b; Irle & Allehoff, 1984). Nach den Ausführungen von Kettel & Simmat (1968) könnte aber eine solche Benutzung möglich sein. Sie konnten nämlich nachweisen, dass das Alter, aber auch die Bildung, einen Einfluss auf die Interessenausprägungen hat. Somit wäre, auch wenn in diesem Bereich zu wenige Studien vorliegen, eine Prozessdiagnostik anzudenken, die die Veränderungsmessung als Ziel hat.

Besonders aus den Erklärungen von Irle (1955 a) kann der Schluss gezogen werden, dass das Testverfahren zur normorientierten Diagnostik entwickelt wurde. Dabei ist das Ziel, das einzelne Ergebnis im Hinblick auf statistische Normwerte, also relativ zur Verteilung der Testergebnisse in einer Bezugsgruppe auszudrücken und zu interpretieren (Pawlik, 1982).

Wenn nun die diagnostische Zielsetzung geklärt ist, liegt die Frage nicht fern, warum die bloße Exploration der Berufsinteressen kaum reichen kann. Besonders als Ersatz eines Gesprächs bzw. Interviews sei der Einsatz des B-I-T. angeraten, da aufgrund von mangelndem Bewusstsein der eigenen Interessen seitens des Bewerbers, Formulierungsschwierigkeiten, das Fehlen von Anhaltspunkten zum Vergleich verschiedener Interessen einer Person und gleicher Interessen verschiedener Personen, mit dem B-I-T. ein objektiveres Instrument eingesetzt werden könne (Irle, 1955 a).

Auch wenn es – verfolgt man den Gedankenweg – naheliegend wäre, ein Profil der Interessen bzw. eine Struktur der Präferenzen zu erhalten und dann auszuwerten: Grundsätzlich ist das diagnostische Ziel des B-I-T. nicht, ein komplettes Berufsinteressenprofil oder gar eine -struktur zu erstellen. Vielmehr sollen nur die hohen/höchsten und geringen/geringsten Ausprägungen herausgearbeitet werden. Hier kann dann die Arbeit des Berufsberaters beginnen (*Irle & Allehoff, 1984; Richter, 1997*).

Theoretische Grundlage

Die hier zu erfolgende Erklärung der theoretischen Grundlage des Testverfahrens muss chronologisch und zwar zeitlich von der Gegenwart in die Vergangenheit vollzogen werden, da auch hier - genauso wie bei der Unterscheidung zwischen Handlungsmotivation und Interesse - ein Wandel vom B-I-T.-I zum B-I-T.-II zu verzeichnen ist.

Liest man im Manual des B-I-T.-II (*Irle & Allehoff, 1984*), dann erscheint den Testautoren eine Theorie des Berufsinteresses abwegig, sie existiert ihrer Meinung nach nicht. Hypothesen über die Entstehung, Entwicklung und Variabilität beruflicher Interessen und deren Zusammenhang mit Berufswahlentscheidungen werden allemal als Konglomerate angesehen von Annahmen, die aus verschiedenen Theorien hergeleitet werden und/oder ad hoc erfunden wurden (*Irle & Allehoff, 1984*).

Irle & Allehoff (1984) verstehen den B-I-T. nicht als Operationalisierung eines Explanans, also einer Theorie des beruflichen Interesses. Auch misst er ihrer Meinung nach nicht - im Rahmen einer Theorie - konkrete Anteile einer Anfangsbedingung, einer Randbedingung oder einer Endbedingung. Der zu klärende Sachverhalt (*Explanandum*), nämlich der ausgeübte oder der chancenreiche Beruf (*mit bestimmten Berufserfolg*), wird nach *Irle & Allehoff (1984)* nur selten durch eine dahinterstehende allgemeine Theorie (*Explanans*) des beruflichen Interesses bestimmt.

Der B-I-T. wird also als „unvollkommenes Produkt“ angesehen, das nicht vorschreibt, „welche Theorie-Perspektive vorzuziehen sei“ (*Irle & Allehoff, 1984*). Der B-I-T. soll ein Einschätzungs-Instrument sein, das „unter ganz unterschiedlichen wissenschaftlichen und technologischen Perspektiven genutzt werden kann“ (*Irle & Allehoff, 1984*) und nur versucht, ein hypothetisches Konstrukt in Korrespondenz zu empirischen Fakten (Präferenz-Urteilen) zu setzen.

Während im B-I-T.-II somit keine Angaben zur Definition des Interesses, des beruflichen Interesses im speziellen, noch zur Theorie gemacht wird, die zu den neun Interessenrichtungen geführt hat, wird in der ersten Version des B-I-T. zumindest noch eine Interessendefinition genannt: Berufliche Interessen sind nach *Irle (1955 b)* komplexe Eigenschaftsorganisationen, die sowohl perzeptive und kognitive, als auch emotionale und motivationale Züge tragen. Sie sollen - anders als geistige Begabung - weniger starr sein und sich im Zusammenspiel von personalen und umweltbedingten Faktoren herausbilden. Auch spricht er von Interessen als „latente Dispositionen, welche die aktuellen Handlungen in einer konkreten Lebenssituation des Individuums allgemein antreiben, ordnen und steuern“ (*Irle, 1955 b*).

Noch weiter geht *Irle* in seiner Inauguraldissertation (*1955 a*). Hier sieht er „entscheidende Gesichtspunkte [...] zur Entwicklung und Strukturierung der Interessen“ und somit die Basis seines Testverfahrens durch die Interessentheorie Lewins erklärt. Die mit dieser Theorie verbundenen Grundidee ist, dass sich das Interesse - zuerst allgemein, unspezifisch - und anschliessend als Teilorganisation der Persönlichkeit - aus einem allgemeingegebenen Interesse an der Umwelt heraus - spezifiziert. Neben einer allgemeinen Darstellung der Theorie wird die Bedeutung Lewins für sein Testverfahren nicht erläutert.

Weiter wird dargelegt, dass er die Interessen-Synthese von *Super (1949)* als theoretischen Ausgangspunkt seiner Entwicklung der neun Interessenrichtungen (*-skalen*) sieht. Die von *Super* erarbei-

tete Zusammenstellung der Autoren Thurstone, Allport, Vernon, Lurie, Strong und Kuder beschreibt acht Interessenrichtungen, die die Grundlage zur Testkonstruktion Irls waren: scientific (*wissenschaftliches Interesse*), social welfare (*soziales Interesse*), material (*handwerkliches Interesse*), system (*Interesse hinsichtlich der Verwaltung*), contact (*kaufmännisches Interesse*), artistic (*künstlerisches I.*) und musical (*musikalisches Interesse*). Über die Modifikation und Erweiterung dieser Interessenskalen werden keine weiteren Angaben gegeben.

(Nachvollziehbarkeit der) Testkonstruktion

Genauso wie an anderen Stellen zuvor werden im B-I-T-II auch zur Testkonstruktion nur dürftige Informationen gegeben. Allerdings bietet besonders die Dissertation Irls eine genaue Beschreibung des Vorganges.

So werden im neusten Manual (*Irle & Allehoff, 1984*) keine Angabe darüber gemacht, wie überhaupt neun Interessenrichtungen zustande kamen. In der Dissertationsschrift allerdings wird erklärt, dass die Interessenrichtungen introspektiv bzw. durch Beobachtung des Berufslebens festgelegt wurden. Die Anzahl der Interessenausprägungen (*-skalen*) scheint willkürlich. Genaue Angaben dazu lassen sich nicht finden. Zusätzlich kam es zu der erwähnten Einarbeitung des Synthese-Ansatzes von Super (1949).

Zu den neun Interessenrichtungen wurden je 50 Items konstruiert. 126 Items dieses Pools wurden aufgrund direkter Berufsbezeichnungen (*anstelle von Tätigkeitsumschreibungen*) entnommen. Die noch übrig gebliebenen 324 Items wurden jeweils mit ihrer zugehörigen Interessenrichtung korreliert. Die 162 Items mit der höchsten Trennschärfe bildeten den B-I-T.-I. Nach leichten Itemänderungen und einer weiteren Auswahl bilden heute Items mit einer Trennschärfe von $r \geq .70$ (*untere Grenze; tetrachorisch*) gegenüber dem Gesamtpunktwert derselben Interessenrichtung den B-I-T.-II (*Amelang & Zielinski, 1997*).

Durch die Itemselektion gelang es, dass mindestens 50 % der Varianz einer Tätigkeit (*Item*) durch die Varianz aller übrigen Tätigkeiten derselben Interessenrichtung beschrieben werden kann (*Irle & Allehoff, 1984*).

In dem B-I-T.-II kommen 81 Tätigkeiten der insgesamt 162 in den Formen AA (*Forced-choice*) und BA (*Free-choice*), weitere 81 in den Formen AB (*Forced-choice*) und BB (*Free-choice*) zum Einsatz. Das Itemmaterial beider Varianten ist, anders als Brickenkamp (1997) und Richter (1997) darlegen, identisch. Weitere Informationen zu der free-choice-Version liegen nicht vor.

Die Items sind neun Skalen (*Interessenrichtungen*) zugeordnet:

- Ø **Technisches Handwerk (TH-1)**
 - z.B. Maschinen in neuen Fabriken montieren
- Ø **Gestaltendes Handwerk (GH-2)**
 - z.B. Innenausstattung für ein neues Hotel aussuchen
- Ø **Technische und Naturwissenschaftliche Berufe (TN-3)**
 - z.B. Verwendbarkeit von Kunststoffen prüfen
- Ø **Ernährungshandwerk (EH-4)**
 - z.B. Erfrischende Obstsalate aus Südfrüchte zubereiten
- Ø **Land- und Forstwissenschaftliche Berufe (LF-5)**
 - z.B. Tannenwälder aufforsten
- Ø **Kaufmännische Berufe (KB-6)**
 - z.B. Urlaubswünsche von Kunden für ein Reisebüro erheben
- Ø **Verwaltende Berufe (VB-7)**
 - z.B. Formularwesen einer Verwaltung vereinfachen

- Ø **Literarische und Geisteswissenschaftliche Berufe (LG-8)**
 - z.B. Im Fernsehen über neue Bücher berichten
- Ø **Sozialpflege und Erziehung (SE-9)**
 - z.B. Selbstmordgefährdete beraten

Testbeschreibung

Verhaltensprovokation

Beschreibung der Items

Dieser Abschnitt bezieht sich auf den B-I-T.-II aus dem Jahr 1984. Dabei sind zwei Versionen mit unterschiedlichem Testmaterial zu unterscheiden. Auf der einen Seite wurde die alte (*allerdings modifizierte*) „forced-choice-Variante“ des B-I-T.-I übernommen. Bei dieser Technik werden gleichzeitig vier Aussagen vorgelegt, die Tätigkeiten aus verschiedenen Tätigkeitsbereichen benennen und von denen jeweils die am stärksten bevorzugte (*mit einem „+“*) und die am meisten abgelehnte Aufgabe (*„-“*) gekennzeichnet werden muss. Eine Gesamtzahl von 81 Vergleichen soll ein aussagekräftiges Bild bezüglich der neun Interessenskalen ermöglichen. Die Anordnung der Items gewährleistet darüber hinaus, dass maximal +36 und minimal -36 Punkte auf eine Interessenrichtung sowie ± 0 im Durchschnitt entfallen. Diesen Umstand verdankt der Test dem ipsativen Messansatz, wodurch die Werte bei gleichen Mittelwerten aller Versuchspersonen in der Population um den Mittelwert einer Person streuen. Somit können die Testwerte nicht von Individuum zu Individuum, sondern nur innerhalb eines Testanden verglichen werden (*Häcker & Stapf, 1998; siehe dazu die rechte Seite der Abbildung 2*).

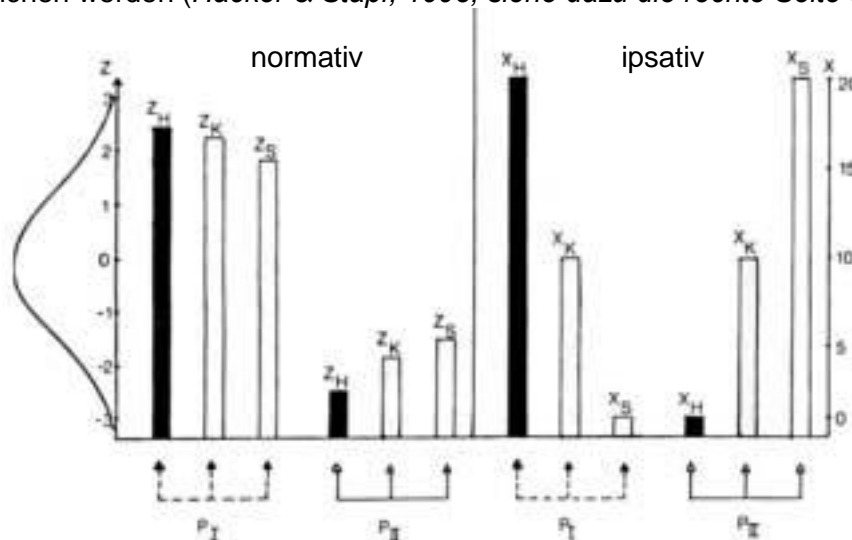


Abbildung 2: Vergleich der Ergebnisse eines normativen (links) und eines ipsativen Messansatzes (rechts) (Jäger & Petermann, 1992)

Die Grundidee, die hinter diesem verwendeten Messansatz steckt, ist die Möglichkeit der intraindividuellen Kontrastverstärkung (*Jäger & Petermann, 1992*).

Bei der im B-I-T.-II neu eingeführten **free-choice**-Variante (*normativer Messansatz*) besteht die Aufgabe nicht darin, Vergleiche zwischen unterschiedliche Tätigkeiten (*wie bei der forced-choice-Technik*) anzustellen. Vielmehr wird hier ein absolutes Rating anhand einer fünfstufigen Skala erwartet (*sehr gern [1], gern [2], weder gern noch ungern [3], ungern [4], sehr ungern [5]*). Die Items selbst sind der forced-choice-Version identisch.

Dabei streuen die Messungen aufgrund des normativen Ansatzes um den Mittelwert der Population (*Amelang & Zielinski, 1997*). Der Messwert sagt etwas über den Ausprägungsgrad einer Person relativ

Berufs-Interessen-Test

zu anderen Personen aus (Rost, 1996). Interindividuelle Vergleiche sind somit - anders als beim ipsativen Messansatz - uneingeschränkt möglich (siehe Abbildung 2 linke Seite).

Die Items sind in beiden Versionen nicht direkt berufsbenennend (denkbar wäre z.B. „Arzt für Innere Medizin“), sondern umschreibend gehalten (z.B. „Wohnungen für Obdachlose beschaffen“).

Die Anordnung der Test-Items stammt bei der forced-choice-Version in abgewandelter Form aus dem „Motivation Indicator“ (Baldwin, 1947) ab, einem Test für die Messung von Unterrichtsinteressen in der Schule. Jede Tätigkeit ist beim B- I-T. aus einer Richtung von acht Tätigkeiten umgeben. Diese acht Tätigkeiten teilen sich gleichmäßig auf die unterschiedlichen Interessenrichtungen auf, d.h. jede Tätigkeiten einer Interessenrichtung wird mit acht anderen Tätigkeiten aus acht unterschiedlichen Interessenrichtungen verglichen.

Dies gilt auch für die Tätigkeiten an den Ecken des Bogens: Die Tätigkeit der Interessenrichtung „TH“ (1) an der linken oberen Ecke des Feldes wird zum Beispiel an den übrigen drei Ecken des Feldes wiederholt. Streicht man nun die doppelten Tätigkeiten, so wird auch hier die obengenannte Regel nicht verletzt.

1	4	7	3	6	8	2	5	9	1
2	6	9	2	4	9	3	4	8	2
3	5	8	1	5	7	1	6	7	3
1	4	7	3	6	8	2	5	9	1
2	6	9	2	4	9	3	4	8	2
3	5	8	1	5	7	1	6	7	3
1	4	7	3	6	8	2	5	9	1
2	6	9	2	4	9	3	4	8	2
3	5	8	1	5	7	1	6	7	3
1	4	7	3	6	8	2	5	9	1

Abb. 3: Anordnung der Interessenrichtungen der forced-choice-Versionen

Betrachtet man den Aufbau zeilenweise, so fällt ein besonderes Muster auf, wobei sich drei Zeilen jeweils drei Mal finden lassen. Zum Abschluss wird die erste Zeile nochmals dargeboten, was ebenfalls mit der obenbeschrieben Regel zu tun hat. Denn nur so ist auch wirklich gewährleistet, dass jede Tätigkeit mit acht anderen aus unterschiedlichen Interessenrichtungen verglichen wird.

Durchführungsmodalität

Die Autoren empfehlen, die Hinweise zur Bearbeitung des Tests langsam und betont vorzulesen, die Testandengruppe auf 20 Personen zu beschränken und die Bewerber auch auf Fehler zu kontrollieren. Mehrmalige Kontroll-Rundgänge sind angebracht, da fehlerhaft ausgefüllte Bogen nicht ausgewertet werden können („unterstützende Rolle des Durchführenden“; Irle & Allehoff, 1984).

Kettel & Simmat (1967) haben - bezüglich des B-I-T.-I - darauf hingewiesen, dass es häufig bei der Durchführungen der forced-choice-Variante nicht allen Testanden auf Anhieb gelingt, die Instruktionen zu verstehen, was wahrscheinlich aus dem recht ungewöhnlichen Testaufbau resultiert. Besonders in einer Gruppe von unterdurchschnittlichen Intelligenzwerten (gemessen durch den I-S-T) mussten Nachinstruktionen gegeben werden. Im Vergleich zu einer Gruppe mit überdurchschnittlich intelligenten Testanden zeigte sich in der Differenz ein signifikantes Ergebnis (n=20; überprüft mit Chi-Quadrat). Die Instruktionen des B-I-T.-II zeigen im Vergleich zum B- I-T.-I - trotz diverser Anregungen von Kettel & Simmat - keine Veränderungen. Der Kritik begegnen die Testautoren (Irle & Allehoff, 1984), indem die Versuchsleiter Psycholog(inn)en oder Psychologisch-Technische Assistent(inn)en sein sollen, da „der B-I-T.-II [...] ein unvollkommenes diagnostisches Instrument mit Stärken und Schwächen“ sei. Die Instruktion, aber natürlich besonders die Interpretation, sei von diesen Berufszweigen zu leisten.

Zeitbedarf

Die Bearbeitungszeit ist bei beiden Varianten nicht limitiert, sollte aber normalerweise nicht über 80 (*forced-choice*) bzw. 40 Minuten (*free-choice*) liegen (Irle & Allehoff, 1984). Die Schwankungsbreite zwischen den Testanden ist aber enorm: In der *forced-choice*-Variante wird ein Spielraum von 20-80 Minuten, sonst von 10-40 Minuten angegeben (Brickenkamp, 1997). Daraus resultiert auch ein Hauptproblem in der Durchführungsphase: Auf der einen Seite ist man angehalten, den Bewerbern so viel Zeit wie nötig zu geben, auf der anderen Seite besteht aber die Gefahr, dass sich aufgrund der hohen Streuung in den Bearbeitungszeiten eine Unruhe unter den Testteilnehmern breit macht. Irle & Allehoff (1984) raten dazu, den B-I-T. als letzten Test vor einer Pause zu verwenden. So wäre es den Bewerbern, die frühzeitig mit dem Ausfüllen fertig geworden sind, ermöglicht, eine ausgedehntere Pause zu nehmen, ohne die anderen Bewerber zu stören.

Auswertung**Beschreibung**

Die insgesamt neun Interessenrichtungen werden bei der ***forced-choice-Version*** durch das Auflegen von sechs DIN-A4-Schablonen ausgewertet, von denen je zwei - zu einer DIN-A3-Schablone zusammengelegt - drei Richtungen berechnen lassen.

Dabei muss streng darauf geachtet werden, dass...

- ∅ immer die beiden richtigen Schablonen zu einer DIN-A3-Schablone kombiniert werden.
- ∅ dass die Schablonen immer korrekt zusammengefügt werden und nicht z.B. seitenverkehrt.
- ∅ dass die Summe der einzelnen Spalten mathematisch korrekt berechnet wird, wobei besonders darauf geachtet werden muss, dass auch die passenden Spalten zu einer Interessenrichtung verrechnet werden. Die Angabe der Richtungen sind schlecht erkennbar und lassen sich einmal oben am Beginn der Spalte finden, manchmal auch unten am Fuß.
- ∅ dass beim Eintragen auf ein Extrablatt die richtige Reihenfolge eingehalten wird. Die Reihenfolge der Schablonenspalten und der Übersichtstabelle auf der letzten Seite des Testbogens stimmen nämlich nicht überein.

Aufgrund dieser Unwägbarkeiten müssen unbedingt folgende Kontrollmechanismen benutzt werden:

Keine Interessenrichtungen darf eine Ausprägung unterhalb von -36 bzw. oberhalb von +36 einnehmen. Die Summe der einzelnen Interessenrichtungen muss normalerweise Null ergeben.

Insgesamt müssen 81 „+“- und 81 „-“-Bewertungen vorgenommen werden. Nach Irle (1955 b) und Irle & Allehoff (1984) dürfen Bogen nicht ausgewertet werden, die eine Abweichung dieser Zahlen von über 5% beinhalten.

Die Auswertung der ***Free-choice-Variante*** erfolgt mit einer DIN-A4-Schablone, auf der pro Interessenrichtung drei Fenster ausgestanzt sind. Diese Schablone muss drei Mal aufgelegt und pro Richtung müssen Rohpunkt-Werte von neun Tätigkeiten summiert werden. Bei mehr als 5% fehlenden Ankreuzungen darf der Bogen nicht ausgewertet werden.

Das Ziel der Auswertung ist nicht, ein Profil der beruflichen Interessen zu erstellen, sondern es sollen nach Irle & Allehoff (1984) nur die prägnanten Interessenrichtungen herausgearbeitet werden (s.o.). Dementsprechend werden Diagnostiker angehalten, keine Grafik im Sinne einer Profillinie zu entwerfen. Vielmehr sollten nur die besonders hohen bzw. niedrigen Ausprägungen in der Auswertung ermittelt werden, genauer gesagt die Richtungen sollen interpretiert werden, die den Prozentrang von 85 (*hohe Ausprägung*), 95 (*sehr hohe Ausprägung*) überschreiten bzw. 5 (*sehr geringe Ausprägung*) und 15 (*geringe Ausprägung*) unterschreiten.

Zeitbedarf

Der Zeitbedarf der ***forced-choice-Variante*** beträgt ohne technische Hilfen etwa 15 Minuten (*inkl. Eintragen in den Bogen*). Dies sind eigene Erfahrungswerte. Angaben dazu lassen sich nicht im Testhandbuch von Irle bzw. Irle & Allehoff finden.

Mit einer technischen Unterstützung in Form eines Excel-Sheets (*siehe Anhang*) reduziert sich die Auswertungszeit um durchschnittlich 5 Minuten.

Ein großes Problem ist und bleibt auch in der neueren Fassung die Benutzung der sechs Schablonen, die Kalkulation der einzelnen Richtungen und das Nachschauen in den Normierungstabellen. Aufgrund der Fehleranfälligkeit sollte geübtes Personal die Auswertung durchführen. Irle & Allehoff (1984) verweisen - genauso wie bei der Durchführung - wieder auf die Verwendung von Psycholog(inn)en und Psychologisch-Technische Assistent(inn)en.

Die **Free-choice-Fassung** gelingt schneller und ist normalerweise nach fünf Minuten pro Bewerber beendet.

Interpretation (ins. Normierung)

Nach Irle & Allehoff (1984) sollen nur die sehr hohen (*Prozentrang* ≥ 95), hohen (*Prozentrang* ≥ 85), geringen (*Prozentrang* < 15), sehr geringen (*Prozentrang* < 5) Ausprägungen interpretiert werden. Eine Profilerstellung und eine grafische Darstellung der einzelnen Interessenrichtungen soll nicht vollzogen werden. Der Zusammenhang von den registrierten Phänomenen und den dahinter liegenden dynamischen Gegebenheiten ist nicht ausreichend geklärt, um eine Struktur bzw. ein Profil zu zeichnen (Irle, 1955 a).

Die ipsativen Werte, die bei der Auswertung des forced-choice-Bogens aus dem B-I-T.-I sowie aus dem B-I-T.-II anfallen, werden zu Prozenträngen umgeformt. Die Prozentrangnorm wird statt Standardwerten benutzt, da nur wenige Interessenrichtungen Normalverteilung erreichen (Irle & Allehoff, 1984).

Am Beispiel des B-I-T.-I soll dies veranschaulicht werden: Der methodische Aufbau des Interessentests, bei dem zwischen Tätigkeiten verschiedener Richtungen eine Wahl getroffen wird (*also forced-choice*), führt zu Verteilung der Ergebnisse, die der Gaußschen Kurve insofern nicht entsprechen, als die Gipfel der Verteilung sich zum unteren Ende der Skala verschieben (Irle, 1955a). Wenn z.B. ein Bewerber in einer Richtung 72 mal ankreuzt, verbleiben für die restlichen 8 Richtungen nur noch $162 - 72 = 90$ Wahlen, pro Richtung also durchschnittliche etwa 11 Wahlen. Da von den Personen durchweg bestimmte Richtungen bevorzugt werden, verschiebt sich die durchschnittliche Zahl der Wahlen in jeder Richtung von $72:8 = 9$ nach 0 hin. Eine Schiefgipfeligkeit entsteht. Im B-I-T.-II wurde dies durch eine leichte Modifikation abgedämpft: Die durchschnittliche Wahl wurden von 9 auf 0 verschoben (Irle & Allehoff, 1984).

Zur Normierung wurden 4410 Personen mit dem B-I-T.-II getestet. Allerdings wurde keine Zufallsauswahl der Testanden durchgeführt. Vielmehr wurde zuerst eine bestimmte Region in Deutschland ausgesucht und dann - dort angereist - eine bestimmte Schulklasse (Irle & Allehoff, 1984). Es wurde darauf geachtet, dass alle Bundesländer (*bis auf Bayern, dessen Kultusminister die Teilnahme verweigert hat*) repräsentativ erfasst wurden.

Die Normierung anhand von Hauptschulen (9. und 10. Klassen), Realschulen (10. Klassen), Gymnasien (13. Klassen) und Gesamtschulen (9. Klassen & Abschlussklassen) lässt schon die diagnostische Zielsetzung erkennen (s. dort). 2187 männliche und 1747 weibliche Testanden haben an der Normierung teilgenommen, im Durchschnitt etwa 1100 pro Testform.

Gütekriterien

Objektivität

Durchführung

Die Angaben im Manual zur Durchführung des B-I-T.-II sind detailliert und präzise. Es werden beispielsweise genaue Angaben zur maximalen Gruppengröße oder zum Zeitbedarf gemacht. Die Durchführungsobjektivität scheint somit gesichert.

Im Manual wird jedoch auch darauf hingewiesen, dass der Testleiter individuelle Fragen der Testanden ausführlich beantworten soll (*besonders bezüglich der nicht unbedingt intuitiv verständlichen Forced-choice-Version*). Da eine solche Interaktion nicht durch Vorgaben des Manuals eingeschränkt wird, verringert sie die ansonsten gegebene Standardisierung des Tests und somit auch die Durchführungsobjektivität.

Auswertung

Die objektive Testauswertung wird durch ein standardisiertes Auswertungsverfahren, das im Manual klar und ausführlich beschrieben wird, gesichert. Der Test verwendet zur Auswertung beispielsweise Schablonen, gibt ein konkretes Auswertungsbeispiel und liefert Hinweise zum Umgang mit fehlenden oder fehlerhaften Daten.

Interpretation

Die Angaben zur Interpretation sind sehr spärlich. Neben Normtabellen geben die Autoren im Manual noch an, ab welchem Prozentrang ein Ergebnis (*abhängig von der betrachteten Dimension und dem Geschlecht*) als „sehr hoch“, „hoch“, „gering“ oder „sehr gering“ zu beurteilen ist. Von der Bildung von Berufsinteressenprofilen und von der Verwendung des B-I-T.-II als alleiniges Klassifikationsinstrument wird außerdem abgeraten. Mehr ist zur Interpretation der Testergebnisse nicht zu finden. Diese Informationen reichen zu einer objektiven Interpretation aber nicht aus, da es so letztendlich in der Hand des interpretierenden Psychologen liegt, was er aus der gewonnenen Information herausliest bzw. was er beispielsweise einem Testanden rät, der ein sehr hohes Interesse im Bereich technisches Handwerk aufweist. Hierzu wären Klassifikationsregeln erforderlich, sinnvoll wären auch (*für die Free-choice-Version*) typische Ausprägungsgrade verschiedener Berufe (*wie sie im B-I-T.-I noch vorhanden waren*). Der Test kann also nicht als objektiv interpretierbar gelten.

Reliabilität

Im Manual des B-I-T.-II lassen sich keine quantitativen Informationen zur Reliabilität des Testverfahrens finden, weder zur Retest- oder Paralleltestreliabilität noch zur internen Konsistenz (*obwohl im Manual des B-I-T.-I noch sehr wohl zumindest Angaben zur internen Konsistenz der Subskalen zu finden waren; s.u.*). Die einzige Aussage der Autoren zu diesem Testgütekriterium ist qualitativer Natur, nämlich dass die Ergebnisse von über die Interitemkorrelationen gerechneten Faktorenanalysen „[...] strukturell über verschiedene Stichproben hinweg („split-sample“-Analysen) sehr stabil [sind]“ (*Irle & Allehoff, 1984*). Diese qualitative Aussage erlaubt jedoch keinerlei Beurteilung (*Richter, 1997*).

Mangels Angaben im Testmanual soll nun auf andere Untersuchungen verwiesen werden, die die Reliabilitätsformen dieses Testverfahrens analysierten.

Paralleltestreliabilität

In seiner Monographie kommt Allehoff (1984) bei der Analyse der Paralleltestreliabilität des B-I-T.-II zu folgenden Ergebnissen:

Berufs-Interessen-Test**Tabelle 1: Paralleltestreliabilitäten zum B-I-T.-II, unterschieden nach Interessenskalen und Messansätzen (Allehoff, 1984)**

Skala	Forced-choice (N = 139)	Free-choice (N = 214)
	corr A mit B	corr A mit B
Technisches Handwerk	.83	.61
Gestaltendes Handwerk	.81	.63
Techn.-naturw. Berufe	.89	.63
Ernährungshandwerk	.74	.61
Land- und Fortswirtschaft	.77	.76
Kaufmänn. Berufe	.71	.62
Verwaltungsberufe	.73	.68
Lit.-geisteswiss. Berufe	.76	.65
Sozialpflege und Erzieh.	.89	.65

Korreliert wurden hier einmal die Paralleltestformen der Forced-choice-Version ($N = 139$) und der Free-choice-Version ($N = 214$). Wie zu erkennen ist, korrelieren die Paralleltestformen der Forced-choice-Version trotz kleinerer Stichprobe höher als die der Free-choice-Version; dies hängt damit zusammen, dass die Items in der Forced-choice-Version aufgrund der hier vorgenommenen Ipsativierung nicht unabhängig voneinander sind und die Korrelationen somit künstlich erhöht werden. Andererseits ist die Stichprobe der Forced-choice-Version eben auch kleiner, weshalb man trotz Ipsativierung zumindest eine ähnliche Effektgrößenausprägung wie in der Free-choice-Version vermuten könnte (*genauere Angaben sind kaum möglich, da es schwierig ist, das Ausmaß der Verzerrung durch die Ipsativierung zu quantifizieren*).

Allerdings stellt sich hier die Frage, inwieweit die hier genannten Paralleltestformen die psychometrischen Anforderungen an Paralleltests (*gleiche Mittelwerte und Varianzen, hohe Interkorrelationen der Testformen*) erfüllen. Allehoff (1984) macht hierzu keinerlei Angaben, und auch an anderer Stelle ist zu dieser Frage nichts zu finden. Schuler (1996) sieht zwei Testformen erst ab einer Korrelation von $r \geq .80$ als parallel an, was hier nur bei einigen Subskalen der Forced-choice-Version gegeben ist. Die Parallelität der Testformen ist daher zumindest zweifelhaft.

Retestreliabilität

Allehoff (1984) ermittelte folgende Test-Retest-Korrelationen (*Halbjahresabstand*) der Free-choice-Version des B-I-T.-II an kurz vor der Berufswahl stehenden Schülern:

Tabelle 2: Re-Test-Reliabilität (Allehoff, 1984)

Skala	Konsistente Berufswähler (N = 83)	Inkonsistente Berufswähler (N = 337)
	r	r
Technisches Handwerk	.33	.51
Gestaltendes Handwerk	.49	.51
Techn.-naturw. Berufe	.59	.46
Ernährungshandwerk	.43	.32
Land- und Fortswirtschaft	.57	.45
Kaufmänn. Berufe	.44	.47
Verwaltungsberufe	.57	.42
Lit.-geisteswiss. Berufe	.56	.47
Sozialpflege und Erzieh.	.69	.54

Berufs-Interessen-Test

Allehoff operationalisierte hier die Stabilität des Berufsinteresses durch die Konsistenz des Berufswunsches. Konsistente Berufswähler ($N = 83$) sind jene, die nach einem halben Jahr noch immer denselben Berufswunsch wie vorher haben, während inkonsistente Berufswähler ($N = 337$) einen anderen Berufswunsch angeben. Allehoff nahm an, dass die Korrelationen bei konsistenten Berufswählern höher oder zumindest annähernd gleich hoch sein müssten, was auch (*bis auf das technische Handwerk*) zutraf. Zu kritisieren wäre hier die unterschiedliche Stichprobengröße, zumal eine Betrachtung der reinen Effektgrößen Allehoffs Hypothese (*Konsistenz des Berufswunsches moderiert die Test-Retest-Korrelationen*) deutlicher untermauern dürfte.

Dennoch sind die Test-Retest- Korrelationen unabhängig von der Konsistenz des Berufswunsches nicht befriedigend. Es stellt sich die Frage, warum sie trotz eines konsistenten Berufswunsches nur im mittleren Bereich liegen. Dies spricht entweder dafür, dass die Operationalisierung des Berufsinteresses durch die Konsistenz des Berufswunsches inadäquat ist, was bedeutet, dass ein konsistenter Berufswunsch keineswegs mit einem stabilen Berufsinteresse gleichzusetzen ist; diese Frage wird im Abschnitt Kriteriumsvalidität näher behandelt.

Es könnte aber auch dafür sprechen, dass die Stabilität des gemessenen Konstruktes Berufsinteresse (*trotz gegenteiliger Angabe der Testanden*) gering ist; dies würde dann die zu interpretierende Korrelation $r_{12} = \rho_{12} \sqrt{rel_1 rel_2}$ stark senken (r_{12} = Test-Retest-Korrelation zwischen Zeitpunkten 1 und 2, ρ_{12} = Stabilität des Berufsinteresses zwischen Zeitpunkten 1 und 2, rel_x = Reliabilität des Tests zum Zeitpunkt x ; Rudinger, 1995).

In der Tat ist das Berufsinteresse in der Pubertät offenbar sehr instabil (*Super, 1949*). Dies spricht gegen die Interpretation der Test-Retest-Korrelationen als Retestreliabilitäten, denn „[...] wenn die beobachtete Korrelation als Schätzung der Test-Retest-Reliabilität angesehen wird, dann impliziert diese Interpretation eine perfekte Stabilität, eine Annahme, die von der Differentiellen Psychologie und der Diagnostik häufig stillschweigend vorausgesetzt wird“ (*Rudinger, 1995*).

Eine weitere Möglichkeit wäre natürlich, dass die mediokren Korrelationen einfach auf eine unbefriedigende Reliabilität des Verfahrens zurückzuführen sind. Mangels Untersuchungen zu diesem Thema ist es allerdings nicht möglich, die Konfundierung von Stabilität und Reliabilität in diesen Test-Retest-Korrelationen aufzuheben und gesicherte Aussagen zur Ursache ihrer unbefriedigenden Ausprägung zu machen (*hierzu wäre beispielsweise eine präzise Erfassung der Stabilität des Berufsinteresses nötig*).

Interne Konsistenz

Im Gegensatz zum Manual der (*jüngeren!*) Testversion B-I-T.-II weist das Manual des B-I-T.-I Daten zur internen Konsistenz auf. Diese konnten durch eine Untersuchung von Deimann & Kastner-Koller (1994) recht gut bestätigt werden:

Tabelle 3: Gegenüberstellung der Internen Konsistenz (KR 20) des Handbuchs (B-I-T.-I) und einer Analyse von Deimann & Kastner-Koller (1994)

	Rel. It. Deimann & Kastner-Koller N = 308	Rel. It. B-I-T Handbuch N = 310
Skala		
Technisches Handwerk	.92	.94
Gestaltendes Handwerk	.86	.87
Techn.-naturw. Berufe	.89	.83
Ernährungshandwerk	.74	.81
Land- und Fortswirtschaft	.87	.95
Kaufmänn. Berufe	.72	.75
Verwaltungsberufe	14 .92	.90
Lit.-geisteswiss. Berufe	.88	.86
Sozialpflege und Erzieh.	.92	.90

Berufs-Interessen-Test

Sowohl Deimann & Kastner-Koller als auch Irle (1955A) verwendeten als Konsistenzmaß die KR 20, da eine sorgfältige Itemschwierigkeits- und Trennschärfeanalyse des B-I-T.-I bisher noch aussteht und die KR 20 diese Daten nicht benötigt (*andere Konsistenzmaße hingegen sehr wohl, z.B. KR 8; Lienert & Raatz, 1998*). Die interne Konsistenzbestimmung wurde hier wohl deshalb nicht mittels Cronbachs α durchgeführt, da dieser Koeffizient die Testhomogenität voraussetzt, die hier schon von der inhaltlichen Seite bewusst nicht gegeben ist.

Die guten internen Konsistenzen sprechen dafür, dass der Messfehler des B-I-T.-I gering ist und das Berufsinteresse recht präzise gemessen wird; nicht vergessen werden darf hier aber die Tatsache, dass der B-I-T.-I ein reines Forced-choice-Verfahren und somit eine artifizielle Erhöhung der internen Konsistenz zu erwarten ist.

Validität / Evaluation

Inhaltsvalidität

Die Items des B-I-T.-II sprechen zumindest für eine hohe Augenscheinvalidität, da sie sehr anschaulich sind (z.B.: „*Maschinen in neuen Fabriken montieren*“, „*Fremdsprachige Romane übersetzen*“ etc.). Es werden jedoch im Manual keinerlei explizite Angaben zur Itemkonstruktion oder –auswahl gemacht, die erstens zur Nachvollziehbarkeit der Testkonstruktion und zweitens zur Kontrolle der Inhaltsvalidität erforderlich wären. Eine der möglichen Ursachen für das Fehlen derartiger Information ist sicher die Tatsache, dass es keine zugrundeliegende Theorie der Berufsinteressen gibt, aus der Items zur Messung derselben abgeleitet werden könnten. Auf eine mögliche Quantifizierung der Inhaltsvalidität (z.B. *Übereinstimmungs-Koeffizient nach Fricke, 1974*) wurde gänzlich verzichtet.

Die Inhaltsvalidität muss also als nicht gegeben angesehen werden, was auch bedeutet, dass die Genauigkeit des Repräsentationsschluss vom im B-I-T.-II gezeigten Interesse auf das Interesse außerhalb der Testsituation als gering angesehen werden muss.

Kriteriumsvalidität

Weder das Manual des B-I-T.-II noch das Manual des B-I-T.-I enthalten Aussagen zur Kriteriumsvalidität.

Die Untersuchung von Deimann & Kastner-Koller (1994) validierte den B-I-T.-I am Kriterium „beruflicher Erfolg“. Als beruflich erfolgreich galten Testanden, die einen Meisterabschluss erlangt hatten oder kurz vor der Meisterprüfung standen. Als beruflich nicht erfolgreich galten Testanden, die länger als ein halbes Jahr arbeitslos waren. Problematisch an der Bestimmung der (*hier der Logik nach eigentlich als prognostisch anzusehenden*) Kriteriumsvalidität war allerdings, dass das Berufsinteresse retrospektiv ermittelt wurde, d.h. die Testanden sollten (*mittels B-I-T.-I*) angeben, welchen Berufswunsch sie früher hatten. Außerdem wurden noch weitere Daten (*Geschlecht, Alter, Familienstand, Intelligenz*) erhoben.

Die Tatsache, dass beruflich Erfolgreiche signifikant häufiger ihren Wunschberuf ausübten als beruflich nicht Erfolgreiche, kann somit nicht nur als Bestätigung der Wichtigkeit von Berufsinteressen, sondern könnte stattdessen auch als Nachweis zur kognitiven Dissonanzreduktion oder eines mere-exposure-Effektes interpretiert werden (*was die folgenden Ergebnisse zumindest fragwürdig erscheinen ließe*).

Deimann & Kastner-Koller führten anschließend eine Diskriminanzanalyse durch und ermittelten den prozentualen Anteil der Variablen zur Diskriminanzfunktion. Sie erhielten folgende Ergebnisse:

Berufs-Interessen-Test**Tabelle 4: Ergebnis der Diskriminanzanalyse von Deimann & Kastner-Koller (1994) mit den einzelnen Beiträgen der Variablen zur Erklärung des Berufserfolgs (zur Operationalisierung siehe Text)**

	Prozentuale Gewichte
Variablen	In %
<i>IQ im IST-70</i>	26.3
<i>Berufsadäquate Ausbildung</i>	4.1
<i>Übereinstimmung mit Wunschberuf</i>	3.5
<i>BIT-VB</i>	3.5
<i>BIT-TN</i>	3.2
<i>Familienstand "geschieden"</i>	3.1
<i>BIT-TH</i>	2.2
<i>BIT-LF</i>	1.9
<i>Familienstand "ledig"</i>	1.6

Es ist klar zu erkennen, dass die anhand des B-I-T-I erfassten Berufsinteressen kaum einen Beitrag zur Erklärung des Berufserfolgs leisten, sondern dass der gewichtigste Prädiktor die erfasste Intelligenz ist. Dies kann zwei Gründe haben. Zum einen könnte die retrospektive Erfassung der Berufsinteressen das Ausmaß geteilter Varianz zwischen Berufsinteresse und Berufserfolg senken, zumal beruflich nicht Erfolgreiche noch sehr wohl Interessen äußern könnten, die ihrem ehemaligen Beruf entsprechen; tun dies die beruflich Erfolgreiche auch, ist es klar, dass das Berufsinteresse somit keine Rolle für die Unterscheidung zwischen den Gruppen spielt.

Dies führt auf den zweiten Grund, der für ein so geringes Erklärungsgewicht des Berufsinteresses sprechen würde: Inwieweit ist das Kriterium angemessen? Stillschweigend und ohne Erklärung setzen Deimann & Kastner-Koller voraus, dass ein hohes Berufsinteresse automatisch mit Berufserfolg einhergeht (s. auch oben bei Allehoff: *Er setzt voraus, dass das Interessenprofil konsistenter Berufswähler eine höhere Stabilität aufweist*). Diese Annahme ist allerdings nicht ohne weiteres zu rechtfertigen. Das zugrundeliegende Problem ist hier folgendes: „Eine Theorie der Berufs-Interessen erscheint den Autoren abwegig; sie existiert auch nicht“ (Irle & Allehoff, 1984). Da es keine Theorie und auch keine präzise Definition des Konstruktes Berufs-Interessen gibt, kann es schwerlich ein adäquates Kriterium geben. Ghiselli fand heraus, dass der Validitätskoeffizient zwischen Berufserfolg und Interessentests meist kleiner als $r = 0.20$ ist, was Zweifel an der Angemessenheit des Kriteriums Berufserfolg aufkommen lässt (Brambring, 1983). Nötig wären in diesem Zusammenhang nicht nur follow-up-Studien, die die „wahre“ (und nicht retrospektive) prognostische Validität des B-I-T validieren, sondern zuvor auch die theoretische Ausarbeitung einer Berufsinteressentheorie und ein daraus abgeleitetes adäquates Kriterium.

Der B-I-T-II ist allerdings explizit kein Instrument, das allein die Berufswahl oder gar den beruflichen Erfolg aufklären kann und soll: „Mit bescheidenem Anspruch wird ein Instrument vorgelegt, das ein relativ isoliertes, hypothetisches Konstrukt in Korrespondenz zu empirischen Fakten, hier Präferenzurteilen, zu bringen sucht“ (Irle & Allehoff, 1984).

Konstruktvalidität

Im Manual des B-I-T-II finden sich zur Konstruktvalidität einzig Interkorrelationstabellen der neun Skalen innerhalb der Normstichproben (*unterteilt nach Testform und Geschlecht*), worüber jeweils Faktorenanalysen berechnet wurden. Die Interskalenkorrelationen sind bei der Forced-choice-Version größtenteils (*schwach*) negativ (*d.h. besitzt ein Proband beispielsweise auf der Skala „Technisches Handwerk“ hohe ipsative Interessenausprägungen, muss er zumindest auf einigen anderen Skalen eine geringe Ausprägung besitzen* \rightarrow negative Korrelation).

Berufs-Interessen-Test

Nach Rost (1996) sind ipsative Meßwerte untereinander im Mittel stets negativ korreliert, was durch die stochastische Abhängigkeit der Skalen bedingt wird. Die Interkorrelationen der Subskalen der Free-choice-Version sind nahezu ausschließlich positiv und besitzen schwache bis mittlere Ausprägungen. Einige Korrelationen weisen allerdings eine beachtliche Höhe auf, was für die Durchführung einer Faktorenanalyse spricht. Angaben zu den Medianen der Korrelationskoeffizienten fehlen allerdings.

Die Ergebnisse der durchgeführten Faktorenanalysen sprechen dafür, dass der B-I-T.-II faktoriell nicht valide ist, da er laut Manual nicht (*die inhaltlich postulierten*) 9, sondern nur 2 Faktoren aufweist (*gilt für beide Testversionen*). Irle & Allehoff (1984) erhielten folgende Faktorenmuster:

Tabelle 5: Faktorenmuster nach Irle & Allehoff (1984)

		männlich	weiblich
Forced-choice-Version	<i>Dimension 1</i>	TH + TN + VB versus GH + SE + LG	KB + VB versus LF + LG
	<i>Dimension 2</i>	KB + VB versus LF	TH + TN + VB versus GH + KB
Free-choice-Version	<i>Faktor 1</i>	LG + SE + GH + EH + KB + (VB + TN)	TH + TN + KB + VB
	<i>Faktor 2</i>	TH + TN	LG + SE + (EH + GH + TN)

(Abkürzungen: **TH** = Technisches Handwerk, **GH** = Gestaltendes Handwerk, **TN** = Technische und Naturwissenschaftliche Berufe, **EH** = Ernährungs-Handwerk, **LF** = Land- und Forstwirtschaftliche Berufe, **KB** = Kaufmännische Berufe, **VB** = Verwaltende Berufe, **LG** = Literarische und Geisteswissenschaftliche Berufe, **SE** = Sozialpflege und Erziehung)

In Klammern gesetzte Skalen sind über verschiedene Stichproben hinweg „nicht stabil“ (Irle & Allehoff, 1984). Die Interpretierbarkeit der Faktoren der Forced-choice-Version im klassischen (*faktorenanalytischen*) Sinne ist nicht gegeben, da die zugrundeliegenden Items nicht unabhängig voneinander und somit die Interkorrelationen verschiedener Skalen hauptsächlich negativ sind. Aufgrund der nicht gegebenen Unabhängigkeit der Items der Forced-choice-Version werden die extrahierten Faktoren in bipolare Dimensionen abgebildet, die sich „interpretativ gegenseitig ergänzen“ (Irle & Allehoff, 1984). Weitere Angaben zur Interpretation der bipolaren Dimensionen werden nicht gemacht.

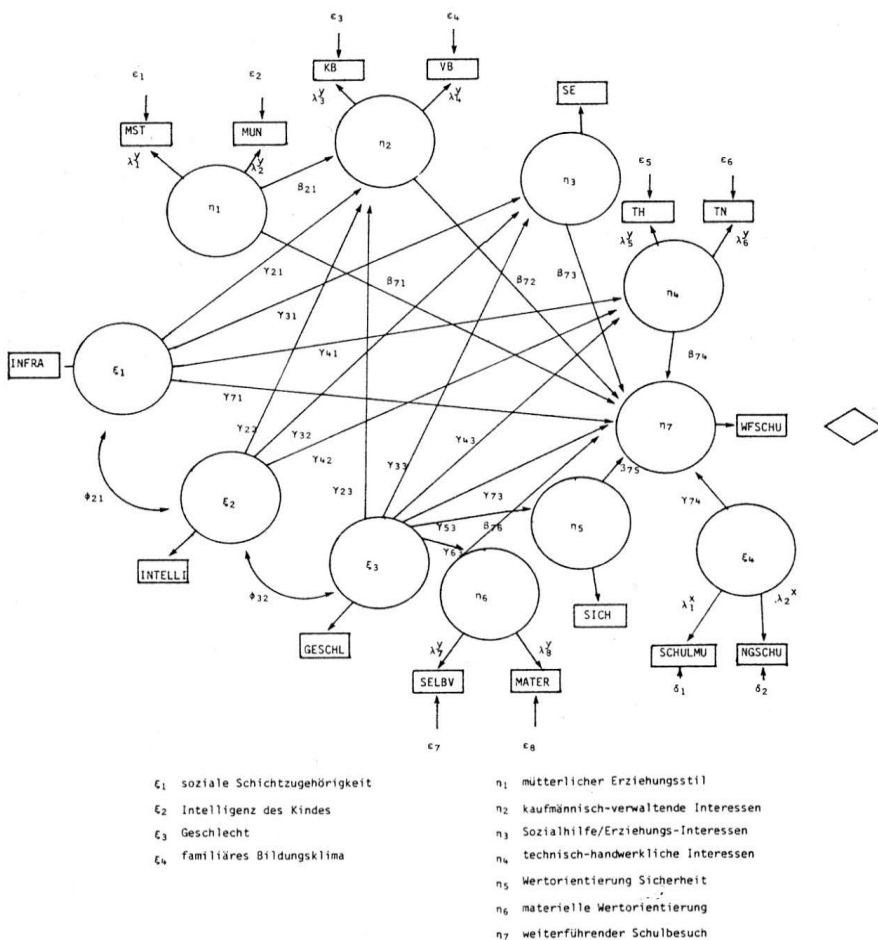
Die Faktoren der Free-choice-Version wären hingegen im klassischen Sinne faktorenanalytisch interpretierbar, Angaben hierzu fehlen hingegen völlig.

Eine Begründung, warum gerade zwei Faktoren extrahiert wurden (*Screeplot? Kaiser-Kriterium?*), fehlt an dieser Stelle, ebenso wie Angaben zur Kommunalität. Die angegebenen Eigenwerte der Faktoren sind in der Forced-choice-Version trotz VARIMAX-Rotation nicht besonders hoch, in der Free-choice-Version zufriedenstellend.

Allehoff (1984) führte eine umfangreiche LISREL-Analyse im Hinblick auf berufliche Interessen und Berufswahl durch (*N = 6629*).

Berufs-Interessen-Test

Abbildung 4: Ergebnis der LISREL-Analyse im Hinblick auf berufliche Interessen und Berufswahl (Allehoff, 1984)



Er verwendet neben den Berufsinteressen noch 10 weitere latente Variablen (z.B. soziale Schichtzugehörigkeit, Intelligenz; diese Variablen wiesen untereinander ebenfalls strukturelle und korrelative Zusammenhänge auf), wobei die Berufsinteressen als drei nach inhaltlichen Kriterien getrennte Faktoren (kaufmännisch-verwaltende Interessen, Sozialhilfe/ Erziehungsinteressen, technisch-handwerkliche Interessen) berücksichtigt wurden.

Die entscheidende endogene latente Variable war „Besuch einer weiterführenden Schule“, da nach Allehoffs umfangreicher theoretischer Fundierung des Strukturgleichungsmodells diese latente Variable diejenige ist, welche den späteren Verlauf des Berufslebens (aus Sicht der hier untersuchten Schüler) am stärksten determiniert bzw. zwischen späteren Berufen am besten diskriminiert.

Betrachtet man die ermittelten Pfadkoeffizienten der Berufsinteressen auf die latente endogene Variable „Besuch einer weiterführenden Schule“, findet man teilweise erhebliche Gewichtungen (je nach „fitting“: technisch-handwerkliche Berufe .44, kaufmännisch-verwaltende Berufe .86, Sozialhilfe- und Erziehungsberufe -.04). Offenbar hat das berufliche Interesse abhängig von der Interessenrichtung einen Einfluss auf das hier verwendete Kriterium „Besuch einer weiterführenden Schule“.

Deimann & Kastner-Koller (1994) führten in ihrer Studie eine Faktorenanalyse über die Interitemkorrelationen des B-I-T.-I durch (N = 308) und konnten eine recht gute faktorielle Validität nachweisen: 7 von 9 Faktoren wurden sehr klar bestätigt (und korrelierten auch hoch mit jeweils einer Skala), 2 weitere Faktoren korrelierten mit 2 oder mehr Skalen. Dies Ergebnis ist erstaunlich, da der B-I-T.-I die Berufsinteressen (wie die Forced-choice-Version des B-I-T.-II) ausschließlich ipsativ misst und somit eher ein diffuses Bild wie beim B-I-T.-II zu erwarten wäre. Mögliche Erklärung für diese Diskrepanz wäre vielleicht, dass Irle & Allehoff (1984) orthogonal rotierten (VARIMAX), während Deimann & Kastner-Koller (1994) eine oblique Rotation durchgeführt haben könnten (detaillierte Angaben fehlen hier).

Eine MTMM-Analyse zum Vergleich des B-I-T. mit anderen Testverfahren und Kriterien wurde bisher nicht durchgeführt, wäre aber von methodischer und inhaltlicher Seite sehr hilfreich (z.B. zur Einbettung des Konstruktes „Berufsinteresse“ in ein nomologisches Netzwerk anhand der Bestimmung der diskriminanten und konvergenten Validität).

Nebengütekriterien

Transparenz

Der B-I -T. II ist für die Testanden transparent, da sie wissen, welche Verhaltenserwartungen an sie gestellt werden, welchem Zweck er dient und die Items anschaulich sind. In einer Testanwendung des B-I-T.-I an 52 Bewerbern (*Tachler, 1983*) gaben 57% der Befragten an, die im B-I-T.-I untersuchten Fähigkeiten seien klar, allerdings gaben auch 21% an, eine Eignungsbestimmung anhand des B-I-T.-I sei nicht vorstellbar.

Verfälschbarkeit

Der B-I-T. II ist ähnlich wie herkömmliche Persönlichkeitstests verfälschbar (*Richter, 1997*). Nach Irle (1963) spricht dies aber in keiner Weise gegen den Test, sondern eher für ihn, da das zu messende Konstrukt den Testanden bewusst und somit wesentlich präziser erfassbar ist. Tendenzen zu „sozial erwünschten“ oder anderweitig nicht interessengesteuert angestrebten Berufen i.S. v. „impression management“ sind nach Irle nicht zu befürchten: „Die Fälschung von Interessen ist nicht in dem Maße möglich, dass berufsspezifische Konstellationen von verschiedenen Interessenrichtungen vorgetäuscht werden können“ (*Irle, 1963*).

Zumutbarkeit

Richter (1997) hinterfragt zumindest die Zumutbarkeit der Forced-choice- Version, da hier eben ein unter Umständen zeitaufwändiger Entscheidungszwang ausgeübt wird. Die Free-choice-Version dürfte jedoch ohne weiteres zumutbar sein.

Akzeptanz

Da das Verfahren keine intimen oder moralisch bedenklichen Fragen stellt, dürfte der B-I-T.-II eine hohe Akzeptanz aufweisen. Außerdem handelt es sich beim B-I -T.-II um einen Interessentest, was den Testanden durch die Instruktion auch klar vermittelt wird; die Bewertungsangst ist somit geringer als bei psychologischen Leistungstests.

Kosten

Die komplette Mappe (*Handanweisung, je 5 Testbögen der 4 Testformen, 6 Schablonen*) des B-I -T.-II kostet 138 DM. Nur die Handanweisung kostet 49,80 DM, jeweils 25 Testbögen kosten 26,25 DM. Der Preis für den Schablonensatz der Testformen A (*Forced-choice- Version*) liegt bei 44 DM, der für den Schablonensatz der Testformen B (*Free-choice-Version*) liegt bei 16,50 DM.

Preis pro Person bei

10 Testanden:	16,43DM	8,40 €
100 Testanden:	2,43DM	1,24 €
1.000 Testanden:	1,19DM	0,61 €

Der B-I -T. II erweist sich als relativ günstiges Verfahren, schon bei 100 Testanden liegt der Preis pro Person bei 2,43 DM und ist somit gerade für berufsberatende Institute (z.B. *Arbeitsamt*) von der finanziellen Seite betrachtet empfehlenswert.

Anwendungsempfehlung

Allgemein

Der B-I-T.-II hat als Zielgruppe Schüler der 9.-13. Klasse, die unmittelbar vor dem Übergang ins Berufsleben stehen (z.B. *vor Beginn einer Lehre*). Die Hauptzielgruppe wird von Irle & Allehoff festgesetzt: „Er [der B- I-T. II] soll auch (und gerade) zur Einschätzung beruflicher Interessen von Hauptschul-Abgängern und -Abgängerinnen diagnostisch brauchbar sein“ (*Irle & Allehoff, 1984*). Laut Irle &

Allehoff (1984) reicht der B-I-T. II allerdings als alleiniges Selektions- oder Klassifikationsinstrument nie aus; er sollte darum immer in einem Ensemble psychologischer Tests angewendet werden.

Konkret

Ein Psychologe, der beim Arbeitsamt als Berufsberater tätig ist, möchte einem unentschlossenen Hauptschüler zu einer seiner Berufsinteressen entsprechenden Lehre raten und führt deshalb den B-I-T.-II (*neben anderen Testverfahren*) zur Klassifikation durch.

Bewertung**Positiv****Allgemeine Vorteile**

Der B-I-T. stellt - besonders für den Einsatz während der Berufsberatung Jugendlicher vor der Lehre - ein nützliches und originelles Werkzeug dar, das eine Transformation von generellen Interessen auf berufliche Interessen umgeht; stattdessen wird sehr anschauungsnah versucht, die Berufsinteressen direkt zu ermitteln.

Vorteile des ipsativen Messansatzes

Im Gegensatz zum D-I-T. (Toldt, 1971) u.ä. muss sich der Proband bzw. Bewerber beim B-I-T. aufgrund der forced-choice-Konstruktion stets für eine von vier dargebotenen Tätigkeiten entscheiden. Dadurch wird es unmöglich, dass sich der Proband in allen Dimensionen als hochinteressant darstellt, und der Einfluss der sozialen Erwünschtheit ist beim B-I-T. dementsprechend verkleinert.

Gerade durch die Art der intraindividuellen Kontrastverstärkung der einzelnen Interessenskalen ist stets die Aussage möglich, in welchem Interessenbereich der Proband die stärkste Ausprägung besitzt, so dass eine klare intraindividuelle Rangfolge erkennbar wird. Genau das ist gerade für die Berufsberatung von großem Nutzen.

Die Testautoren schreiben den Benutzern des Verfahrens aber nicht den Messansatz vor, sondern bieten seit 1984 auch eine free-choice-Version an, wodurch der Testanwender flexibler ist und auch andere Fragestellungen angehen kann.

Normstichprobe

Besonders die sehr große ($N=4400$ Personen) und die geographisch breit verteilte Normstichprobe ist ein großer Vorteil des Verfahrens.

Gütekriterien

Die Durchführungs- und die Auswertungsprozesse sind sehr detailliert beschrieben. Auch besitzt zumindest der B-I-T. zufriedenstellende interne Konsistenz und faktorielle Validität.

Negativ**Theoretische Grundlage**

Die theoretische Grundlage des B-I-T. ist schwer nachvollziehbar. Eine gewisse Willkürlichkeit des Testverfahrens, gar eine „Theorielosigkeit“ wird von den Autoren nicht verschwiegen. Dies dadurch zu rechtfertigen, dass man keinem Benutzer die Theorie-Perspektive vorschreiben möchte, scheint allerdings unangebracht. Zumindest die Definition des Interesses müsste gegeben sein, immerhin wird der Begriff im Manual mehrmals verwendet. Die Testkonstruktion ist nur in der Dissertation Irls nachlesbar, in den Manuals selbst - die den meisten Benutzern nur vorliegen werden - sind sie nur ansatzweise vorhanden.

Allgemeine Probleme

Es kann nicht davon ausgegangen werden, dass es den Bewerbern gelingt, die Tätigkeiten ganz ohne den Einfluss von Gehalt bzw. gesellschaftlichem Ansehen zu bewerten. Eine Konfundierung wird immer einen verfälschenden Einfluss haben.

Auch ist eine Verzerrung durch die Wunschkomponente denkbar (*Amelang & Zielinski, 1997*), das heißt, man könnte eine Tätigkeit hoch einschätzen, weil man gerade dafür - aus externen Gründen - Interesse empfinden sollte. Diese Konfundierung ist aber wahrscheinlich bei allgemeinen Interessentests (z.B. *DIT*) stärker. Mit anderen Worten: Wie bei allen Persönlichkeitstests können auch beim B-I-T. im Sinne der Erwünschtheit Antworten vorgetäuscht werden. Natürlich stellt sich die Frage, ob man Berufsinteresse als völlig unabhängig von Gehalts- oder Ansehensvorstellungen auffassen kann; auch hier würde eine präzise Definition von „Berufsinteresse“ eine große Hilfe sein.

Es stellt sich allgemein natürlich die Frage, inwieweit standardisierte Testverfahren zur Erfassung der Berufsinteressen überhaupt nötig sind. *Dolliver (1969)* konnte nachweisen, dass die Selbsteinschätzung hinsichtlich der Berufsinteressen der Erfassung durch standardisierte Testverfahren ebenbürtig sind. Der erhöhte Aufwand einer Testung und einem späteren Gespräch, z.B. mit dem Berufsberater, könnte somit um den ersten Schritt, dem Testverfahren zum Berufsinteresse, gekürzt werden.

Durchführung & Auswertungsökonomie

Durchführung und Auswertung sind zeitlich aufwendig und sehr fehleranfällig (*der mitgelieferte Schablonensatz von 6 Schablonen hätte z.B. ohne weiteres auf 2 Schablonen reduziert werden können*). Die fehlenden Angaben zur Korrektur einer falschen Tätigkeitsbewertung seitens der Testanden sind zudem problematisch.

Gütekriterien

Im Manual sind weder interpretierbare Reliabilitäts- noch Validitätsdaten zu finden, deren Angabe eigentlich eine wissenschaftliche Forderung an heutige psychologisch-diagnostische Verfahren darstellt. Auch zur Interpretation werden nur sehr dürftige Angaben gegeben, wodurch falsche diagnostische Entscheidungen getroffen werden können. Die Befunde zur Retest- und Paralleltestreliabilität, zur Inhaltsvalidität und zur Kriteriumsvalidität (*wenn „Berufserfolg“ als valides Kriterium angesehen wird*) sprechen gegen das Testverfahren.

Probleme ipsativer Messung

Die Wahl der einen Möglichkeit schließt die Wahl der anderen aus, auch wenn nur ein minimaler Unterschied der Interessen besteht, d.h. die ipsative Messung kann keinerlei Aussagen über die Abstände zwischen den Interessenausprägungen machen. Der erzwungene Wahlakt offenbart nur die Rangfolge, aber nicht, ob die einzelnen Tätigkeiten ähnlich oder sehr unterschiedlich bewertet werden (*lediglich Ordinalskalenniveau ≠ Informationsdefizit gegenüber der Free-choice-Version*).

Des weiteren kann eine mittlere Ausprägung in allen Interessen für zwei Dinge sprechen: nach *Meili & Steingrüber (1978)* könnte sie sowohl auf ein Desinteresse an allen diesen Bereichen deuten oder aber auch auf ein breites Interesse an sehr vielen Tätigkeiten.

Ein weiteres Problem der ipsativen Messung ist die nach Meinung verschiedener Autoren nicht erlaubte interindividuelle Vergleichbarkeit ipsativer Meßwerte. „Der aus der Beantwortung solcher Items resultierende Punktwert lässt keinen Schluss zu auf die Intensität oder die Generalität der Interessen der Testanden im Vergleich zu einer bestimmten Bezugsgruppe“ (*Todt, 1978*). Laut *Jäger & Petermann (1992)* sind interindividuelle Vergleiche und auf Personenstichproben zugeschnittene statistische Operationen selbst nach weiteren Skalentransformationen nicht zulässig. Die von *Irle* in der Dissertation (*1955 a*) vorgenommene Erstellung von Berufsprofilen anhand von Mittelwertsbestimmungen der ipsativen Messwerte ist somit fragwürdig. In späteren Publikationen zum B-I-T. wurde diese Berechnung bekanntermaßen auch nicht mehr durchgeführt.

Andere Autoren (z.B. *Rost, 1996*) sehen zumindest den Vergleich von relativen Variablenausprägungen zwischen den Personen als erlaubt an.

Die Problematik der ipsativen Messung wird im Manual nicht erwähnt und somit wird dem fehlerhaften Gebrauch in der Praxis nicht vorgebeugt.

Interpretation

Der Anwendungsbereich des B-I-T. ist eingeschränkt, da aufgrund der Auswahl der Tätigkeiten eine Fokussierung auf Lehr-Berufe vollzogen wurde. Allgemein muss die Berufsfeldstruktur angepasst werden, da die Veränderung der Berufswelt seit den fünfziger Jahren zu wenig berücksichtigt wurde. Es lässt sich eine klare Übergewichtung handwerklicher und land- und forstwirtschaftlicher Tätigkeiten feststellen: 45 % der 81 Testitems gehören zu diesen Bereichen, obwohl die genannten Tätigkeiten nur noch 10% der heute ausgeübten Berufe darstellen (*Richter, 1997*).

Die Prozentrangnormen werden nur nach dem Geschlecht differenziert. Die Einflüsse der Schulbildung beziehungsweise des Alters werden nicht berücksichtigt, zeigten aber einen Einfluss auf das berufliche Interesse (*Kettel & Simmat, 1968*). Auch ist eine Erstellung der naheliegenden Interessenstruktur nach *Irle & Allehoff (1984)* nicht erlaubt.

Irle & Allehoff (1984) haben das Testverfahren primär zur Klassifikation konstruiert, geben aber keine klaren Informationen zur Zuordnung, die im B-I-T. noch vorhanden waren, an (*Richter, 1997*).

Selbst die Testautoren kommen zum Fazit, dass „der B-I-T.-II nie ausreichen wird, um Selektions- oder Klassifikationsentscheidungen, um Auslesen oder Beratung allein von diesen Informationen abhängig zu machen“ (*Irle & Allehoff, 1984*).

Die Instabilität von Berufsinteressen in der Pubertät ist ein weiteres Problem (*Super, 1949*). Eine Klassifikation von Testanden wäre demnach zeitlich wenig stabil und könnte sogar hinderlich sein.

Fazit

Der B-I-T.-II ist ein interessantes, nützliches und weitverbreitetes Werkzeug zur beruflichen Interessenbestimmung. Aus methodischer Sicht (*Reliabilität, Validität, Itemkonstruktion, präzise Begriffsdefinition „Berufsinteresse“*) benötigt er allerdings eine Revision.

Wichtig wäre es, die Benutzern präzise über die grundsätzlich unterschiedlichen Messansätze der Testformen in Kenntnis zu setzen und auch Tipps für die Nutzung (*wann und wo z.B. der forced-choice-Bogen sinnvoll wäre*) zu geben.

Anhang

Auswertungshilfe

Berufs-Interessen-Test (BIT-II)

Form AA

Auswertungshilfe von Alexander Zimmerhofer

Schablone 1	TH1	EH1	VB1	TH2	EH2	VB2	TH3	EH3	VB3	TH4	
		-5	3	-4	-11	3	2	-10	6	-2	-6
Schablone 2	GH1	LF1	LG1	GH2	LF2	LG2	GH3	LF3	LG3	GH4	
		2	-10	8	-3	-7	3	1	-7	5	1
Schablone 3	TN1	KB1	SE1	TN2	KB2	SE2	TN3	KB3	SE3	TN4	
		1	5	5	1	12	-3	3	6	-2	3

Technisches Handwerk	Gestaltendes Handwerk	Technische und Naturwissenschaftliche Berufe	Ernährungs-Handwerk	Land- und Forstwirtschaftliche Berufe	Kaufmännische Berufe	Verwaltende Berufe	Literarische und Geisteswissenschaftliche Berufe	Sozialpflege und Erziehung
-32	1	8	12	-24	23	-4	16	0

m	sehr gering			sehr hoch	sehr gering	sehr hoch	hoch	
w	sehr gering	gering	hoch	hoch	sehr gering	sehr hoch	hoch	gering

m - 95	17	18	32	12	24	14	14	22	27
m - 85	8	9	26	7	17	7	6	11	15
m - 15	-21	-11	-6	-10	-7	-11	-18	-14	-12
m - 5	-27	-16	-16	-15	-13	-17	-26	-20	-21
w - 95	-4	25	16	16	19	10	14	19	33
w - 85	-10	21	7	10	12	6	5	9	27
w - 15	-26	1	-15	-6	-9	-9	-19	-13	2
w - 5	-31	-5	-20	-11	-17	-13	-26	-19	-5

Literatur

- ALLEHOFF**, Wolfgang (1984). Berufswahl und berufliche Interessen. Göttingen: Hogrefe.
- AMELANG & ZIELINSKI** (1997). Psychologische Diagnostik und Intervention. Berlin [u.a.]: Springer.
- BALDWIN**, G.B. (1946). Motivation Indicator. Minneapolis: Educational Testing Service.
- BRAMBRING**, M. (1983). Spezielle Eignungsdiagnostik. In: K.-J. Groffmann und L. Michel [Hrsg.] Intelligenz- und Leistungsdiagnostik. Enzyklopädie der Psychologie. BII/2, Göttingen: Hogrefe, Seite 414-481.
- BRICKENKAMP**, Rolf (1997). Handbuch psychologischer Test. Göttingen: Hogrefe-Verlag.
- CRONBACH**, Lee J. (1970). Essentials of psychological testing. New York [u.a.]: Harper and Row.
- DOLLIVER**, Robert H. (1969). Strong Vocational Interest Blank versus expressed voc. interests: A Blank. Psychological Bulletin. 72 (2), Seite 95 -107.
- DEIMANN**, Pia & **KASTNER-KOLLER**, Ursula (1994). Wozu Berufsinteressentests? In: Psychologie in Erziehung und Unterricht 41, S. 49-58.
- FRICKE**, R. (1974). Kriteriumsorientierte Leistungsmessung. Stuttgart: Kohlhammer.
- HÄCKER**, Hartmut & **STAPF**, Kurt H. (1998) [Hrsg.]. Dorsch Psychologisches Wörterbuch. Bern: Verlag Hans Huber.
- IRLE**, Martin (1955 A). Eine Methode zur Bestimmung beruflicher Interessen und erste Ergebnisse ihrer Anwendung. Dissertation Uni Göttingen.
- IRLE**, Martin (1955 B). Der Berufs-Interessen-Test (B-I-T.). Weinheim: Hogrefe.
- IRLE**, Martin (1963). Die Vortäuschung von Interessen im Berufs-Interessen-Test (B-I-T.). In: Diagnostika, S. 49 - 55.
- IRLE**, Martin & **ALLEHOFF**, Wolfgang (1984). Der Berufs-Interessen-Test (B-I-T). Weinheim: Hogrefe.
- JÄGER**, R.S., **PETERMANN**, F. (Hrsg.) (1992). Psychologische Diagnostik. Weinheim: Beltz – Psychologische VerlagsUnion.
- KETTEL**, Karl Josef & **SIMMAT**, William E. (1967). Über B-I-T.-Instruktion und Kurzform. Diagnostika, S. 30 - 37.
- KETTEL**, Karl Josef & **SIMMAT**, William E. (1968). Geschlecht, Alter und Bildung als Bedingungen der Interessenausprägung. In: Diagnostica, Band XIV, 156-174.
- LIENERT**, Gustav A. & **RAATZ**, U. (1998). Testaufbau und Testanalyse. 6. Auflage. Weinheim: Beltz – Psychologische VerlagsUnion.
- PAWLIK**, K. (1982). Modell- und Praxisdimensionen psychologischer Diagnostik. In K. Pawlik [Hrsg.] (1982). Diagnose der Diagnostik. (S. 13-42). Stuttgart: Klett.
- RICHTER**, Peter (1997). Testrezension zu Berufs-Interessen-Test (B- I-T.-II). In: Zeitschrift für Differentielle und Diagnostische Psychologie, 18. Jahrgang, Heft 1/2, S. 100 - 101. **ROST**, Jürgen (1996). Testtheorie Testkonstruktion. Bern: Verlag Hans Huber.
- RUDINGER**, Georg (1995). Strukturgleichungsmodelle in der Entwicklungspsychologie. In: Oerter & Montada, Entwicklungspsychologie. 6. Auflage. Weinheim: Beltz – Psychologische VerlagsUnion.
- SCHULER**, Heinz (1996). Psychologische Personalauswahl - Einführung in die Berufseignungsdiagnostik. Göttingen: Verlag für Angewandte Psychologie.
- SUPER**, D. E. (1949). Appraising Vocational Fitness. New York: Harper & Brothers.
- TACHLER**, E. (1983). Empirischer Beitrag zur Erforschung der sozialen Validität, zit. nach **SCHULER**, H. (1990). Personalauswahl aus der Sicht der Bewerber: zum Erleben eignungsdiagnostischer Situationen. Zeitschrift für Arbeits- und Organisationspsychologie 34 (N.F. 8), 4.
- TODT**, Eberhard (1978). Das Interesse. Bern [u.a.]: Huber.
- TODT**, Eberhard (1971). Differentieller Interessentest DIT. Bern: Huber