

Kersting, M. (2000). Instrumente der Arbeits- und Organisationspsychologie: Rezension des "Intelligenz-Struktur-Test 2000". Zeitschrift für Arbeits- und Organisationspsychologie, 44, 96–101.

Instrumente der Arbeits- und Organisationspsychologie

Rezension des „Intelligenz-Struktur-Test 2000“
von R. Amthauer, B. Brocke, D. Liepmann und A. Beauducel

Martin Kersting

Testart

Intelligenztest

Testmaterial

Die Mappe für den I-S-T 2000 umfaßt die Handanweisung, Testhefte (pro Testform [„A“ und „B“] je ein Heft des Grundmoduls, des Merkfähigkeitstests und des Erweiterungsmoduls), einen Satz Auswertungsschablonen sowie je fünf Exemplare Antwortbogen, Rohwertprotokolle und Ergebnisprofilbogen. Alle Materialien können einzeln nachbestellt werden. Das Auswertungsprogramm lag zum Zeitpunkt der Rezensionserstellung noch nicht vor.

Testgliederung

Der I-S-T 2000 ist modular aufgebaut und ermöglicht drei unterschiedliche Anwendungsformen. (1) Zum Grundmodul gehören je drei Aufgabengruppen mit jeweils 20 Items zur verbalen, numerischen und figuralen Intelligenz, die auf einer höheren Generalitätsebene als Indikator für schlußfolgerndes Denken („Reasoning“) berücksichtigt werden können. (2) Zusätzlich zu diesen neun Aufgabengruppen können zwei, in einem separaten Testheft dargebotene Aufgabengruppen zur verbalen (10 Items) und figuralen (13 Items) Merkfähigkeit eingesetzt werden. (3) Das Erweiterungsmodul umfaßt 73 Wissensfragen. Dabei geht es um Kenntnisse aus fünf unterschiedlichen Wissensdomänen wie z.B. „Geographie/Geschichte“, „Naturwissenschaften“ und „Alltags-Sym-

bole“, wobei die Kenntnisfragen zu jeder Wissensdomäne mit verbalem, numerischem und figuralem Aufgabematerial dargeboten werden. Der Einsatz der Wissensfragen ermöglicht es, die Leistungen im Gesamtest über die Bestimmung von Faktorwerten nach fluider und kristallisierter Intelligenz zu differenzieren.

Bei den Aufgaben „Matrizen“, „Figurenpaare“ und „Rechenzeichen“ handelt es sich – ebenso wie bei den Wissensfragen – um Neukonstruktionen. Abweichend vom I-S-T 70 werden die „Rechenaufgaben“ im I-S-T 2000 ohne sprachliche Einkleidung vorgegeben. Die übrigen Aufgabengruppen wurden in überarbeiteter Form aus dem I-S-T 70 übernommen.

Grundkonzept / Theoretische Grundlagen

Der I-S-T 2000 stellt eine Weiterentwicklung des mit zahlreichen psychometrischen Mängeln behafteten I-S-T 70 dar. Zu nennen sind hier beispielsweise die niedrigen Reliabilitäten für die Aufgabengruppe „Wortauswahl“ und die Schwierigkeitsunterschiede zwischen den beiden Testformen bei dieser Aufgabe sowie bei der Aufgabe „Analogien“ (z. B. Schmidt-Atzert, Hommers und Heß, 1995). Die Autoren des I-S-T 2000 begnügen sich aber nicht mit einer Überwindung der handwerklichen Defizite des I-S-T 70, sondern ändern auch das Grundkonzept der Testkonstruktion. Während Amthauer bei der Konstruktion des I-S-T keinem expliziten theoretischen Ansatz folgte (siehe Brocke, Beauducel und Tasche, 1998), wurde die Konstruktion des I-S-T 2000 systematisch aus theoretischen Annahmen über Intelligenztestleistungen abgeleitet. Die theoretische Fundierung schafft die Vor-

aussetzung für die Konstruktvalidität des neuen Tests. Insbesondere hinsichtlich dieses Gütekriteriums, der Konstruktvalidität, erweisen sich die „Vorgänger“ I-S-T und I-S-T 70 als defizitär. So ließ sich die von Amthauer postulierte Intelligenzstruktur anhand der mit dem I-S-T 70 erhobenen Daten zumeist nicht aufzeigen. Entsprechende Analysen deuteten eher in Richtung einer Zwei- bis Dreifaktorenstruktur des I-S-T 70 (z. B. Brocke et al., 1998; Heyde, 1996; Greif, 1972; Schmidt-Atzert et al., 1995). Außerdem wird der für den I-S-T 70 berechnete Gesamtwert durch ein theoretisch unbegründetes Übergewicht sprachlicher Fähigkeitsskalen dominiert und kommt somit nur bedingt als Indikator der Allgemeinen Intelligenz in Frage.

Diese „Geburtsfehler“ des I-S-T konnten nur durch ein modifiziertes Grundkonzept und dessen konsequente Umsetzung überwunden werden. Als theoretischer Ausgangspunkt der Konstruktion des I-S-T 2000 wird ein „*hierarchisches Rahmen- bzw. Protomodell der Intelligenzstrukturforschung (HPI)*“ formuliert. Dieses Modell wird von den Autoren aus den Konvergenzen verschiedener Intelligenztheorien abgeleitet. Angeführt werden hier sowohl strukturelle Konvergenzen wie die Annahme unterschiedlicher Hierarchie- oder Generalitätsebenen als auch inhaltliche Konvergenzen auf der Ebene der Primärfaktoren. Die fünf Intelligenzkomponenten verbale, numerische und figurale Intelligenz sowie schlußfolgerndes Denken („Reasoning“) und Merkfähigkeit werden von den Autoren als modellübergreifend bedeutsame Intelligenzfaktoren gewertet und mit dem I-S-T 2000 erfaßt. Neben dieser inhaltlichen Festlegung folgt die Testkonstruktion darüber hinaus der Strukturannahme der *multi-trait-Determination* von Intelligenztestleistungen, derzufolge sich in den Leistungen bei der Bearbeitung einzelner Aufgaben verschiedene Fähigkeiten zugleich manifestieren können. Eine Fähigkeit, die bei der Manifestation verschiedener anderer Fähigkeiten Einfluß nimmt, kann durch die Berücksichtigung einer höheren Generalitätsebene der Auswertung derselben Aufgaben diagnostiziert werden. Dieser Aspekt findet im I-S-T 2000 zweifach Berücksichtigung. Das schlußfolgernde Denken („Reasoning“) wird auf einer übergeordneten Generalitätsebene als Summenscore der verbalen, numerischen und figuralen Intelligenz ausgewertet. Im Erweiterungsmodul werden auf übergeordneter Generalitätsebene die Faktoren fluide und kristallisierte Intelligenz gebildet. Diese auf Cattell und Horn zurückgehende Differenzierung stellt nach Ansicht der Autoren eine spezielle Version des „HPI“ dar. Beim Einsatz des Erweiterungsmoduls kann das schlußfolgernde Denken somit zum einen in einer stärker kulturgebundenen Form („Reasoning“) und zum anderen in einer weniger kulturgeprägten Form („fluide Intelligenz“) diagnostiziert werden. Als Markiervariable für kristallisierte Intelligenz dient der als „Erweiterungsmodul“ bezeichnete neue Wissenstest.

Durchführung

Die paper-pencil Version des I-S-T 2000 kann einzeln oder in Gruppen bearbeitet werden. Angaben zur maximalen Gruppengröße fehlen in der Handanweisung ebenso wie auf das Alter bezogene Empfehlungen zum Testeinsatz. Der Test ist für Personen im Alter zwischen 15–25 Jahren normiert. Angeboten werden zwei pseudo-parallele Formen, d.h., die Aufgabensubstanz ist in beiden Formen identisch, lediglich die Darbietungsfolge der Items und ggf. der Distraktoren wurde variiert. Die Voraussetzungen für die Durchführungsobjektivität des Verfahrens sind durch die vorgedruckten Erläuterungen der Aufgabengruppen, durch Beispielaufgaben und durch die Vorgaben der Bearbeitungszeit gegeben.

Für den Einsatz der neun Aufgaben des Grundmoduls sind insgesamt 92 Minuten zu veranschlagen. Werden zusätzlich die Merkfähigkeitsaufgaben appliziert, kommt man – inklusive einer dann notwendigen Pause – auf eine Gesamtzeit von 114 Minuten. Die vollständige Durchführung aller Komponenten des I-S-T 2000 (Grundmodul inkl. Merkfähigkeit und Erweiterungsmodul) nimmt einschließlich zweier Pausen insgesamt 156 Minuten in Anspruch.

Auswertung

Bei der Auswertung des Grundmoduls und der Merkfähigkeitsaufgaben kommen Antwortschablonen zum Einsatz, mit denen der Rohwert pro Aufgabengruppe bestimmt wird. Die Auswertung des Erweiterungsmoduls ist vergleichsweise aufwendiger, da hier keine Antwortschablonen, sondern lediglich sogenannte „Rohwertprotokolle“ angeboten werden. Über die Spalten dieser Rohwertprotokolle wird die Zuordnung der Rohwertpunkte pro Item zu den Subskalen des Wissenstests realisiert. Anschließend werden die Rohwertsummen anhand von Tabellen einmal in Punktwerte für fluide und einmal in Punktwerte für kristallisierte Intelligenz umgelesen. Diese Punktwerte stellen das Ergebnis der Multiplikation der Rohwerte mit den regressionsanalytisch ermittelten B-Gewichten zur Berechnung der Faktorwerte dar. Dieses Vorgehen zeigt, daß auch bei „Handauswertungen“ das Arbeiten mit Faktorwerten möglich ist. Gleichwohl wird das angekündigte Auswertungsprogramm eine willkommene Erleichterung bei der Faktorwertbestimmung darstellen. Faktorwerte erlauben eine diskretere bzw. den intendierten Konstrukten angemessenere Abbildung von Fähigkeiten. Alle ermittelten Werte können auf dem Vordruck „Ergebnisprofil“ eingetragen werden.

Gütekriterien

Reliabilität

Zur Reliabilität des Tests werden Maße der inneren Konsistenz und Halbierungskoeffizienten berichtet. Die so ermittelten Reliabilitäten der Einzelskalen und der Gesamtskala „Reasoning“ können mit Werten über .80 bis .97 als gut bis sehr gut bezeichnet werden. Diese Reliabilitätsschätzungen basieren auf dem Pool der Items der jeweiligen *Skala*, in der mehrere Aufgabengruppen Berücksichtigung finden. Die Reliabilitätsschätzungen für die einzelnen Aufgabengruppen fallen teilweise deutlich geringer aus, insbesondere die Items der Aufgabengruppen „Satzergänzungen“ und „Matrizen“ sind wenig homogen. Man sollte daher unbedingt den Vorgaben der Autoren folgen und auf einen separaten Einsatz und/oder eine separate Interpretation der Aufgabengruppen oder Subskalen verzichten.

Zur Bestimmung der Reliabilität der Faktorwerte für fluide und kristallisierte Intelligenz schlagen die Autoren ein methodisch interessantes neues Verfahren vor. Grundidee dieser neuen Bestimmungsmethode ist es, die Faktorwerte zweimal – aufgrund jeweils der Hälfte der Items der Aufgabengruppen – zu schätzen und dann die so gebildeten Faktorwerte miteinander zu korrelieren. Für zwei unterschiedliche Datensätze werden entsprechende Spearman-Brown korrigierte „split-half-Reliabilitäten“ in Höhe von $r = .86$ bis $r = .93$ für die Faktorwerte fluide und kristallisierte Intelligenz berichtet.

Validität

Zur Konstruktvalidität des Grundmoduls werden Dimensionsanalysen der faktoriellen Struktur angeführt. Die Analysen mit den neun Aufgabengruppen zur verbalen, numerischen und figuralen Intelligenz liefern eine überzeugende Bestätigung der diesbezüglich dreifaktoriellen Struktur. Erste Hinweise auf die konvergente und diskriminante Validität des Grundmoduls können aus den Korrelationen der Skalen „verbal“, „numerisch“, „figural“ sowie der Gesamtskala „Reasoning“ einerseits und vier verschiedenen Intelligenzmaßen andererseits abgeleitet werden. Die anhand von Untersuchungen mit 170 bis 180 Schülern¹ errechneten Daten zeigen verhältnismäßig geringe Zusammenhänge mit dem als Indikator für Speed eingesetzten „d2“-Test sowie substantielle Zusammenhänge mit dem „CFT-20“ und mit dem Wissensanteil des „HAWTE-R“. Als einziges externes Kriterium stehen Schulnoten in fünf Fächern von 151 bis 202 Schülern zur Verfügung. Der Gesamtwert („Rea-

soning“) kovariiert insbesondere ($r = -.45$) mit der Schulleistung im Fach „Mathematik“ sowie – in nominell geringerem Ausmaß – mit den naturwissenschaftlichen Fächern. Die Schulleistungen in den sprachlichen Fächern korrelieren mit den sprachlichen intellektuellen Fähigkeiten ($r = -.20$ bis $r = -.25$).

Aus nicht genannten und nicht nachvollziehbaren Gründen bleiben die Merkfähigkeitsaufgaben bei den Analysen zur Validität vollständig ausgespart. Obwohl auf Seite 15 der Handanweisung explizit angekündigt wird, daß der I-S-T 2000 die Erfassung von vier Intelligenzfaktoren ermöglicht, werden keine entsprechenden Analysen berichtet. Auch bei dem Bericht des Zusammenhangs der Skalen mit anderen Tests und mit Schulnoten bleibt die Merkfähigkeit außer Betracht, so daß über die Validität dieser Skala kaum etwas ausgesagt werden kann.

Die mit dem Wissenstest des Erweiterungsmoduls gewonnenen Daten wurden mit Hilfe der Multidimensionalen Ähnlichkeitsstruktur-Analyse umfassend untersucht. Die Ergebnisse der Dimensionsanalysen rechtfertigen die Bildung von mehreren Wissensskalen, deren gemeinsame Analysen mit den neun Skalen des Grundmoduls (ausgenommen der Merkfähigkeit) den Ausgangspunkt für die Dimensionsanalysen zur fluiden und kristallisierten Intelligenz bilden. Die Daten stimmen mit der intendierten zweifaktoriellen Struktur gut überein.

Anhand der Daten von 173 nicht näher beschriebenen Schülern konnte die konvergente Validität des Maßes für kristallisierte Intelligenz anhand der substantiellen Korrelationen zu dem Wissensanteil des „HAWTE-R“ sowie zum „Mehrfachwahl-Wortschatz“ nachgewiesen werden. Als Indikator der konvergenten Validität der fluiden Intelligenz gilt den Autoren der substantielle Zusammenhang mit den Matrizenaufgaben des CFT-20. Auch der Nachweis der diskriminanten Validität konnte anhand der Korrelationen der Faktorwerte mit dem „d2“ sowie mit den bereits genannten Intelligenzmaßen geführt werden. Schließlich werden Zusammenhänge zwischen den Indikatoren der fluiden und kristallisierten Intelligenz einerseits mit ausgewählten Schulnoten andererseits berichtet. Das berichtete Korrelationsmuster unterstreicht nach Ansicht der Autoren die Differenzierung in fluide und kristallisierte Intelligenz.

Normen

Umfang und Differenzierung der für den I-S-T 2000 angebotenen Normen variieren in Abhängigkeit von den jeweils getesteten Fähigkeitskomponenten. Die Skalen zur verbalen, numerischen und figuralen Intelligenz sowie der diesbezügliche Summenwert wurden anhand der Daten von insgesamt 1285 Personen im Alter zwischen 15–25 Jahren normiert. Diese Normgruppe

¹ Personen und Funktionsbezeichnungen gelten jeweils in weiblicher und männlicher Form.

wird nach Schulbildung (Gymnasiasten versus Nicht-Gymnasiasten) untergliedert. Innerhalb der 925 Personen umfassenden Gymnasialgruppe werden vier Altersgruppen, innerhalb der anderen Gruppe werden zwei Altersgruppen unterschieden. Für den Gesamtwert („Reasoning“) werden zusätzlich zu den genannten sechs Gruppen spezifischer Alters- und Bildungskombinationen drei weitere Normgruppen angeboten. Möglich ist zusätzlich eine Eichung aufgrund der Gruppe (1) aller Gymnasiasten und (2) aller Nicht-Gymnasiasten (jeweils ohne Altersdifferenzierung) sowie (3) aufgrund einer stratifizierten Gesamtgruppe ($N = 625$). Bei dieser zuletzt genannten Gruppe wurde der Anteil der Gymnasiasten auf ca. 40% beschränkt. Während alle bisher angeführten Normen die Z-Skala ($M=100$; $s=10$) nutzen, wird für die zuletzt genannte Vergleichsbasis des „Reasoning“-Wertes (Gesamtgruppe) zusätzlich auch die IQ-Skala ($M=100$; $s=15$) angeboten.

Den für die Merkfähigkeitsaufgaben gebildeten Normen liegt mit 435 Gymnasiasten und 331 Nicht-Gymnasiasten eine weniger umfangreiche Normbasis zugrunde. Zusätzlich wird auch hier eine stratifizierte Gesamtgruppe (40% Gymnasiasten, $N = 534$) als Referenz angeboten. Auf die Bildung von Altersgruppen wurde hier verzichtet, da sich keine altersspezifischen Ausprägungen der Merkfähigkeit zeigten.

Die Faktorwerte zur fluiden und kristallisierten Intelligenz wurden anhand der Daten von 248 Schülern (15 bis 25 Jahre) normiert, eine Differenzierung nach Alter und Schulbildung erfolgt nicht.

Kritik

Der I-S-T 2000 weist zahlreiche Vorzüge gegenüber dem I-S-T 70 auf, die Neuentwicklung geht über die seit langer Zeit notwendige, aber nicht hinreichende bloße technische Revision des I-S-T 70 hinaus. Durch die konsequente Ausrichtung der Testkonstruktion auf eine explizite theoretische Grundlage konnte nach Ausweis der Strukturanalysen ein konstruktvalides Instrument geschaffen werden. Die theoriegeleitete Konstruktion macht sich vor allem auf der Ebene der aggregierten Werte positiv bemerkbar, wobei die Aggregationen der Annahme der *multi-trait-Determination* von Intelligenztestleistungen folgen. Die über verschiedene Inhaltskaskalen (verbal, numerisch, figural) hinweg erfolgende Summierung der Leistungen zu einer Gesamtskala des schlußfolgernden Denkens erinnert stark an die im Kontext des Berliner Intelligenzstrukturmodells (Jäger, 1982) entwickelte und für den „BIS-Test“ (Jäger, Süß und Beauducel, 1997) genutzte Aggregationstechnik. Diesbezüglich wurde es versäumt, eine explizite Bezugnahme herzustellen. Die durch die theoriegeleitete Aggregation gewährleistete Ausmittlung der Inhaltskomponenten

führt dazu, daß der I-S-T 2000 eine stringent operationalisierte Erfassung des schlußfolgernden Denkens ermöglicht. Damit hebt sich der Test positiv von seinem „Vorgänger“ ab. Die bedauerndwert häufig vorkommende Interpretation des Gesamtwerts des I-S-T 70 als globales Intelligenzmaß ist insofern unberechtigt, da aufgrund der pro Dimension *unterschiedlichen* Anzahl von Aufgabengruppen diese Dimensionen in (theoretisch unbegründeten) *unterschiedlichem* Ausmaß in den Gesamtwert des I-S-T 70 eingehen. Diesem Mangel wurde mit dem Konzept des Grundmoduls des I-S-T 2000 abgeholfen.

Daß es mit dem Erweiterungsmodul des I-S-T 2000 im deutschen Sprachraum erstmalig möglich ist, Diagnosen der fluiden und kristallisierten Intelligenz (junger) Erwachsener zu erstellen und somit die Investmenttheorie von Cattell und Horn deutlich besser zu nutzen, stellt einen weiteren Vorzug des I-S-T 2000 dar. Dabei wurde die häufig praktizierte, aber theoretisch unbegründete reduktionistische Operationalisierung fluiden Intelligenz über figurale Reasoningaufgaben (z.B. Advanced Progressive Matrices) und kristallisierten Intelligenz über verbale Aufgaben (z.B. Mehrfachwahl-Wortschatztest) überwunden. Im I-S-T 2000 wird die fluide und kristallisierte Intelligenz sensu Cattell vollständig mit der Inhaltsfacette (verbal, numerisch, figural) kreuzklassifiziert, wodurch eine Kontamination dieser Fähigkeiten mit Inhaltskomponenten vermieden wird. Die Diagnose der kristallisierten Intelligenz könnte insbesondere für die Berufseignungsdiagnostik eine hohe Bedeutung erlangen. Fachkenntnisse gehören zu den validesten Prädiktoren beruflichen Erfolgs (siehe z.B. Schuler, 1996, S. 138 f.). Die kristallisierte Intelligenz könnte ein Indikator dafür sein, wie gut es Personen gelingt, ihre Intelligenz erfolgreich zum Erwerb von Kenntnissen zu nutzen. Möglicherweise erlaubt die Diagnose allgemeinbildender Kenntnisse somit Rückschlüsse auf den Erfolg beim späteren Erwerb von Fachkenntnissen. Diesbezüglich dringend benötigte empirische Erkenntnisse stehen noch aus, der I-S-T 2000 schafft aber die Voraussetzung zur Erhebung entsprechender Daten. Im Kontext der Berufseignungsdiagnostik muß allerdings die Gefahr der Verbreitung von Items beachtet werden. Wissensfragen „verbrennen“ vergleichsweise schneller als herkömmliche Intelligenzitems. Die Anwendung des Erweiterungsmoduls sollte daher langfristig auf potentielle Effekte der inoffiziellen Verbreitung der Kenntnisfragen hin überprüft werden.

Andere Aspekte des I-S-T 2000 geben Anlaß zur Kritik und indizieren Ergänzungs- und Entwicklungsbedarf. Hinsichtlich der Validität muß zunächst das „theoretische Modell“ der Testkonstruktion hinterfragt werden. Die mit dem „hierarchischen Rahmen- bzw. Protomodell der Intelligenzstrukturforschung“ (HPI) formulierte Idee einer integrativen Betrachtung konkurrieren-

der Intelligenzmodelle bleibt insgesamt Kryptographie. So ist es offen, welche Intelligenzmodelle letztlich konkret auf inhaltliche und formale Konvergenzen geprüft wurden. Während im Text zahlreiche Modelle angerissen werden, sind in Abbildung 2.1 auf Seite 7 der Handanweisung lediglich drei Ansätze (warum diese drei?) herausgegriffen. Die Kriterien für die Annahme einer Konvergenz werden nicht spezifiziert, die Regeln, nach denen der I-S-T 2000 aus dem HPI-Modell abgeleitet wurde, sind unvollständig. Die „wahrnehmungsbezogenen Fähigkeiten“ werden beispielsweise zwar zu den modellübergreifend bedeutsamen Intelligenzfaktoren gezählt, bleiben im I-S-T 2000 gleichwohl aber – ohne Begründung – unberücksichtigt. Empirische Modellüberprüfungen des „HPI“ werden nicht angestellt, Fragen der Konstanz, Invarianz, Stabilität, Generalität und Universalität der Faktoren des „HPI“ nicht thematisiert. Es ist einzuräumen, daß eine Handanweisung möglicherweise nicht der Ort ist, ein Intelligenzmodell hinreichend darzustellen – dieser Einwand bleibt aber fruchtlos, solange das „HPI“ auch an keiner anderen Stelle zufriedenstellend erläutert ist. Angesichts der in der I-S-T 2000 Handanweisung lediglich skizzenhaft bleibenden Überlegungen sowie angesichts der ausstehenden empirischen Prüfungen erscheint somit der Modellbegriff strapaziert – zumindest wenn man den Begriff „Modell“ so versteht, wie er in der Intelligenzstrukturforschung üblicherweise verwendet wird. Die Überlegungen stellen gleichwohl einen interessanten und diagnostisch tragfähigen, wenngleich unvollständigen theoretischen Bezugsrahmen dar, der auf Gemeinsamkeiten existierender Intelligenzmodelle verweist. Die Grundidee einer weitergehenden Einbettung von Intelligenztests in den breiten Kontext von Strukturmodellen, wie sie mit dem I-S-T 2000 ansatzweise verfolgt wurde, ist grundsätzlich begrüßenswert. Vor dem Hintergrund dieser Idee wären Strukturvergleiche zwischen den mit dem I-S-T 2000 erhobenen Daten und solchen Daten, die mit empirisch sehr gut ausgeloteten Strukturtests (z. B. dem BIS-4 Test von Jäger et al., 1997) erhoben wurden, eine interessante und für die Interpretation äußerst hilfreiche Ergänzung. Dies gilt insbesondere für die fluide und kristallisierte Intelligenz, deren Bezug zu anderen intellektuellen Fähigkeiten bislang nur unzureichend geklärt ist. Unbedingt nachzutragen sind Strukturanalysen, die auch die Merkfähigkeitsaufgaben einbeziehen.

Daß bei einem neuen Test zunächst ein Bedarf an Kriteriumsvalidierungen besteht, ist nicht ungewöhnlich. Bei dem hier gegebenen Sonderfall der Überarbeitung und Erweiterung eines der im deutschen Sprachraum am häufigsten eingesetzten Tests ist es aber überraschend, daß die Autoren im Kontext der Validitätsdiskussion nicht auf den Zusammenhang zwischen dem I-S-T 70 und dem I-S-T 2000 eingehen. Da zahlreiche Aufgabengruppen aus dem I-S-T 70 fast unverändert übernommen wurden, könnten – mit zu explizierenden Ausnahmen und

Einschränkungen – einige der mit dem I-S-T 70 gewonnenen Validitätsbelege auch bei Überlegungen zum Einsatz und zur Interpretation des I-S-T 2000 hilfreich sein.

Bezüglich der Normen ist zu kritisieren, daß für die unterschiedlichen Anwendungsformen des I-S-T 2000 (ausschließlich) unterschiedliche Normgruppen angeboten werden. Bei einer vollständigen Nutzung aller Testmodule bedeutet dies, daß die auf dem Ergebnisprofil verzeichneten Werte auf einem Vergleich mit einer jeweils anderen Referenzgruppe beruhen. Die Normbasis ist teilweise knapp, teilweise (für die Faktorwerte zur fluiden und kristallisierten Intelligenz) sehr knapp. Äußerst unbefriedigend ist die Tatsache, daß ausgerechnet für die stärker kulturgebundene kristallisierte Intelligenz eine Differenzierung der Normen nach Alter und Schulbildung unterbleibt. Nachzutragen sind Angaben zu geschlechtsspezifischen Leistungsausprägungen sowie vor allem eine weitergehende Differenzierung der Kategorie „Nicht-Gymnasiasten“. Ein solcher Sammelbegriff ist für den Anwendungsalltag nur bedingt tauglich. Durch zwei Nachlässigkeiten auf der Detailebene ergibt sich bei der Bestimmung der Standardwerte für die Merkfähigkeitsaufgaben ein durch das Handbuch nicht vollständig gedeckter Regelungsbedarf. Zum einen sind diesbezüglich zwei unterschiedliche Tabellenvarianten vorgesehen, und zum anderen bestehen innerhalb einer Tabelle teilweise Freiheitsgrade in der Zuordnung der Rohwerte zu den Standardwerten. Um Einbußen in der Auswertungsobjektivität zu vermindern, müssen Anwender diesbezüglich eine eindeutige Festlegung treffen (für eine ausführliche Darstellung dieser Problematik siehe Kersting, 1999, S. 102 ff.).

Empfehlung

Mit dem I-S-T 2000 wurde jetzt endlich eine Version des I-S-T vorgestellt, die vorhandene Vorzüge des I-S-T 70 – wie z. B. das attraktive und zahlreich erprobte Aufgabenmaterial – beibehält, psychometrische und theoretische Defizite des I-S-T 70 ausräumt und darüber hinaus innovative diagnostische Möglichkeiten wie die Diagnostik der fluiden und kristallisierten Intelligenz eröffnet. Angesichts der teilweise massiven Probleme des I-S-T 70, die bei Funktionsmängeln einzelner Items anfangen, über unterschiedliche Schwierigkeiten der beiden Testformen gehen und bis zu empirisch nicht zu rechtfertigenden Strukturannahmen reichen, muß der wider besseren Wissens erfolgende weitere Einsatz des veralteten I-S-T 70 eindeutig als fachlich begründungsbedürftig gekennzeichnet werden. Verlag, Autoren und Anwender des I-S-T 2000 sollten gemeinsam dazu beitragen, daß die nun anstehende Aufgabe der Sammlung umfangreicherer Normdaten sowie der Gewinnung

weitergehender Erkenntnisse zur Validität rasch in Angriff genommen wird.

Literatur

- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (1999). *Intelligenz-Struktur-Test 2000*. Göttingen: Hogrefe.
- Brocke, B., Beauducel, A. & Tasche, K. (1998). Der Intelligenz-Struktur-Test: Analysen zur theoretischen Grundlage und technischen Güte. *Diagnostica*, 44, 84–99
- Greif, S. (1972). *Gruppenintelligenztests. Untersuchungen am WIT, IST, LPS und AIT*. Frankfurt/ M.: Lang.
- Heyde, G. (1996). Weitere Daten zum I-S-T 70. *Report Psychologie*, 21, 138.
- Jäger, A.O. (1982). Mehrmodale Klassifikation von Intelligenzleistungen. Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells. *Diagnostica*, 28, 195–226.
- Jäger, A.O., Süß, H. M. & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. Form 4. Handanweisung*. Göttingen: Hogrefe.
- Kersting, M. (1999). Rezension des Intelligenz-Struktur-Test 2000. In E. Fay (Hrsg.), *Tests unter der Lupe. Band 2* (S. 86–113). Lengerich: Pabst.
- Schmidt-Atzert, L., Hommers, W. & Heß, M. (1995). Der I-S-T 70. Eine Analyse und Neubewertung. *Diagnostica*, 41, 108–130.
- Schuler, H. (1996). *Psychologische Personalauswahl. Einführung in die Berufseignungsdiagnostik*. Göttingen: Verlag für Angewandte Psychologie.
- Dipl.-Psych. Dr. Martin Kersting
 Deutsche Gesellschaft für Personalwesen e.V., Grassistr. 12, D-04107 Leipzig