

Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). 2., erweiterte und überarbeitete Auflage

L. SCHMIDT-ATZERT, Fachbereich Psychologie
 der Philipps-Universität Marburg

W. RAUCH, Institut für Psychologie der
 Johann Wolfgang Goethe-Universität Frankfurt am Main

Allgemeine Informationen über den Test, Beschreibung des Tests und seiner diagnostischen Zielsetzung

Der I-S-T 2000 R (erw.) soll mit dem »Grundmodul« die fluide und mit dem »Erweiterungsmodul« die kristallisierte Intelligenz (Wissen) möglichst breit erfassen. Der Test ist ab 15 Jahren anwendbar und dauert mit Instruktionen ca. 1 Std. 30 Min. (Grundmodul; Merkaufgaben mit Pause weitere 20 Min.), der Wissenstest erfordert weitere ca. 40 Min. Das Grundmodul liefert einen Gesamtwert für schlussfolgerndes Denken sowie Kennwerte für die verbale, numerische und figurale Intelligenz (basierend auf entsprechenden Subtests). Zwei nicht obligatorische Subtests erfassen die Merkfähigkeit. Das »Erweiterungsmodul« dient der Erfassung von verbalem, numerischem und figuralem Wissen. Bei Durchführung des Erweiterungsmoduls können »reine« Kennwerte für das »schlussfolgernde Denken/gf« (um Wissen bereinigt) bzw. »Wissen/gc« (um schlussfolgerndes Denken bereinigt) bestimmt werden. Form A und B unterscheiden sich nur in der Reihung der Aufgaben. Eine echte Parallelform C existiert für das Grundmodul.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Der I-S-T 2000 R (erw.) basiert auf dem »alten« I-S-T 70. Der Überarbeitung wurde das sogenannte »Hierarchische Rahmen- bzw. Protomodell der Intelligenzstrukturforschung (HPI)« zugrunde gelegt. Diese Formulierung kennzeichnet, dass beim I-S-T 2000 R (erw.) ein Kompromiss zwischen einer Kontinuität beim Übergang von der »alten« auf die neue Testversion auf der einen Seite und der Berücksichtigung der Ergebnisse der neueren Intelligenzforschung auf der anderen Seite hergestellt wurde. Inhaltlich wird das Modell durch die Primärfaktoren charakterisiert, die übereinstimmend in den Intelligenzmodellen von Thurstone, Vernon und Cattell vorkommen, formal durch eine hierarchische Struktur mit den übergeordneten Faktoren gf und gc auf der höchsten Ebene. Messtechnisch bedeutsam ist, dass verbale, numerische und figurale Aufgabengruppen gleich stark vertreten sind (bei der Merkfähigkeit jedoch keine numerischen Aufgaben).

Objektivität

Die Durchführungs- und Auswertungsobjektivität wird durch präzise Anweisungen und standardisiertes Testmaterial bzw. Auswertungsmaterialien (Schablonen) ge-

währleistet. Der Interpretationsobjektivität dienen die nach Alter und Bildung aufgeschlüsselten Normtabellen. Welche Fähigkeiten mit den einzelnen Kennwerten erfasst werden sollen, geht aus den Interpretationshinweisen hervor. Allerdings fallen die Beschreibungen der drei Subfacetten des Wissens (verbal, numerisch und figural) sehr knapp aus. Vorschläge zur sprachlichen Vermittlung unterschiedlich ausgeprägter Ergebnisse werden nicht gemacht; einzig die Fallbeispiele auf den Seiten 83 bis 88 geben Anhaltspunkte.

Normierung (Eichung)

Für die Normierung des Grundmoduls wurden insgesamt 3 484 Probanden mit der Form A/B und 2 363 mit der Form C untersucht. Die Probanden stammen aus sieben deutschen Bundesländern. Normentabellen liegen für Gymnasiasten, Nicht-Gymnasiasten und eine kombinierte Gesamtgruppe, jeweils mit unterschiedlichen Altersstaffelungen, vor. Mit acht Altersgruppen sind die Grundmodulnormen für Gymnasiasten am stärksten unterteilt (von 15-16 bis 51 und mehr Jahre). Zur Merkfähigkeit liegen nur drei breite Altersnormen vor. Mit den Wissenstests wurden insgesamt 1 107 Probanden untersucht. Bei den fünf Altersgruppen wird nicht nach Bildung differenziert. Dafür liegen aber zwei Gesamtnormen für Gymnasiasten und Nicht-Gymnasiasten vor. Die Größe der Teilstichproben variiert in Abhängigkeit von Alter, Testmodul und Testform. Die Stichprobengröße reicht beim Grundmodul von 69 (Form C, Gymnasiasten, Alter 51+) bis 578 (Form A/B, Nicht-Gymnasiasten, 31-40 Jahre).

Zuverlässigkeit (Reliabilität, Messgenauigkeit)

Es werden ausschließlich Angaben zur internen Konsistenz gemacht (Cronbachs Alpha und – mit ähnlichen Ergebnissen – Split-Half). Cronbachs Alpha wird für das schlussfolgernde Denken mit 0,96 (Form C: 0,94) angegeben, für die Merkfähigkeit und das Wissen jeweils mit 0,93. Auf der Ebene verbaler, numerischer und figuraler Intelligenz liegen die Reliabilitätsschätzungen zwischen 0,87 (Form C: 0,82) und 0,95 (0,92), bei den drei Wissensbereichen zwischen 0,82 und 0,84. Die Reliabilitäten (Cronbachs Alpha) der Subtests des Grundmoduls variieren zwischen 0,69 und 0,86; niedrige Werte sind aber unerheblich, weil die Autoren davon abraten, einzelne Subtestergebnisse zu interpretieren.

	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weit- gehend	teil- weise	nicht
Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). 2., erweiterte und überarbeitete Auflage				
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung	●			
Objektivität	●			
Zuverlässigkeit		●		
Validität		●		

Gültigkeit (Validität)

Die Struktur des Tests wurde mittels konfirmatorischer Faktorenanalysen untersucht. Die Ergebnisse sprechen dafür, dass die Subtests des erweiterten Grundmoduls zu Kennwerten für verbale, numerische und figurale Intelligenz sowie Merkfähigkeit verrechnet werden dürfen. Für die Konstruktvalidität sprechen Korrelationen von 0,63, 0,69 und 0,49 zwischen schlussfolgerndem Denken und drei konvergenten Tests (CFT 20, Raven, FRT). Die diskriminante Validität wird durch niedrige Korrelationen mit einem Konzentrationstest (Test d2, $r = 0,22$) und einem Wortschatztest (MWT-B, $r = 0,31$) belegt. Strukturanalysen zum Wissensmodul zufolge fügen sich die Aufgabengruppen sowohl in ein Modell mit verbalem, numerischen und figuralen Wissen ein als auch in eines nach inhaltlichen Themenbereichen. Weitere Validitätsbelege liegen nur für Faktorwerte vor: Der Globalwert g_c korreliert anders als g_f relativ hoch mit Wissenstests, das Korrelationsmuster der Faktorwerte mit Schulnoten folgt dem erwarteten Muster.

Weitere Gütekriterien (Störanfälligkeit, Unverfälschbarkeit und Skalierung)

Nach oben kann ein Leistungstest wie der I-S-T 2000 R (erw.) praktisch nur dann verfälscht werden, wenn ein Proband sich vorher gründlich mit dem Test vertraut macht oder indem er in einer Gruppenuntersuchung abschreibt. Da der Test vom Verlag nicht an jedermann abgegeben wird und ein Abschreiben sich durch Einsatz der Parallelförmchen weitgehend verhindern lässt, kann die Verfälschung nach oben bei sachgemäßem Einsatz gut verhindert werden. Eine Simulation von niedriger Leistungsfähigkeit ist alleine am Testverhalten nicht zu erkennen. Bei einem gründlichen diagnostischen Vorgehen sollten aber Diskrepanzen zwischen Testergebnis und anderen Intelligenzindikatoren (u.a. schulische und berufliche Leistungen) bemerkt und abgeklärt werden. Zur Frage der Rasch-Skalierbarkeit des Grundmoduls (Form A) verweisen die Autoren auf eine Untersuchung von Bühner, Ziegler, Krumm und Schmidt-Atzert (2006), der zufolge die meisten Untertests dem eindimensionalen Rasch-Modell genügen.

Abschlussbewertung/Empfehlung

Beim I-S-T 2000 R (erw.) handelt es sich um einen gründlich konstruierten Intelligenztest mit einer langen Tradition. Der Test könnte aber im Detail noch optimiert werden.

Die Herleitung des HPI-Modells ist nachvollziehbar, entbehrt aber mit der Auswahl der drei Ausgangsmodelle nicht einer gewissen Beliebigkeit. Die Unterscheidung von fluider Intelligenz und Wissen wird durch die Intelligenzforschung unterstützt. Das Modell stellt wohl einen Kompromiss zwischen wissenschaftlichem Anspruch und praktischen Konstruktionsanforderungen dar. Die Konsistenzmaße für die globalen Kennwerte liegen über 0,90 und damit in einem sehr hohen Bereich. Leider werden keine Retestreliabilitäten berichtet, auch keine Reliabilitäten für die einzelnen Normgruppen (zur Berechnung von Konfidenzintervallen).

Dass das Grundmodul schlussfolgerndes Denken erfasst, wird durch erwartungsgemäß hohe Korrelationen mit konvergenten Tests sowie mit der Mathematiknote belegt. Für die verbale, numerische und figurale Intelligenz fehlen hingegen noch überzeugende Validitätsbelege; die drei Facetten korrelieren in ähnlicher Höhe mit figuralen Intelligenztests und mit Schulnoten (z. B. beträgt die Korrelation zwischen Mathematiknote und numerischer, verbaler und figuraler Facette $-0,40$, $-0,29$ und $-0,36$). Für die Validität der verbalen Intelligenz spricht jedoch, dass sie höher als die numerische und figurale Intelligenz mit Aspekten des Spracherwerbs korreliert. Für die Wissenstests fehlen auf der Ebene der verbalen, numerischen und figuralen Facetten noch Validitätsbelege in Form von Korrelationen mit anderen Tests oder Kriterien. Die Unterscheidung der drei Facetten ist konsequent; ob sie im Wissensbereich auch praktisch relevant ist, erscheint fraglich. Für viele Altersgruppen sind die Normierungstichproben hinreichend groß, für einige aber mit $N < 100$ zu klein. Bei den Wissenstests sollten Angaben zu den Stichprobengrößen berichtet werden. Wünschenswert wären genaue Angaben zu Anwendungszweck und Zielgruppen.

Diese Testrezension wurde im Auftrag des Testkuratoriums der Föderation deutscher Psychologengvereinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2006, 2007) erstellt.

Testkuratorium (2006). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologengvereinigungen. Report Psychologie, 31, 492-499.

Testkuratorium (2007). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologengvereinigungen. Psychologische Rundschau, 58, 25-30.

TESTINFORMATIONEN

Liepmann, D., Beauducel, A., Brocke, B. & Amthauer, R. (2007). Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R). 2., erweiterte und überarbeitete Auflage. Göttingen: Hogrefe.

Bezugsquelle: Testzentrale Göttingen, Robert-Bosch-Breite 25, D-37079 Göttingen (E-Mail: testzentrale@hogrefe.de; URL: <http://www.testzentrale.de>)

Test komplett bestehend aus: Manual, je 1 Grundmodul Form A, B + C, je 1 Grundmodul Merkaufgaben Form A + B, je 1 Erweiterungsmodul Wissenstest Form A + B, 15 Antwortbogen Grundmodul A, B + C, 15 Antwortbogen Wissenstest A + B, 15 Ergebnisprotokolle Erweiterungsmodul, 15 Ergebnisprofile, je 1 Schablone Wissenstest Form A + B, je 2 Schablonen Grundmodul Form A, B + C und Mappe: 188 Euro