

Psychologische Rundschau

Herausgeber

Karl Christoph Klauer
Jens Bölte
Oliver Dickhäuser
Barbara Krahe
Klaus Moser

Offizielles Organ der Deutschen Gesellschaft für Psychologie (DGPs)

Themenheft

Gewissheit und Skepsis:
Wissenschaftskommunikation als
Forschungsthema der Psychologie

Gastherausgeber

Rainer Bromme und Dorothe Kienhues

TBS-TK Rezension

Adaptives Intelligenz Diagnostikum 3 (AID 3)

Matthias Ziegler, Humboldt-Universität zu Berlin
Andreas Reichert, Diakonisches Werk Würzburg e.V.

Allgemeine Informationen

Im Manual geben die Autoren Testzweck, Testziel und Zielpopulation an; die Mindestanforderungen an ein Manual sind somit erfüllt. Der AID 3 diene der Erfassung komplexer und basaler Kognitionen bei Kindern und Jugendlichen (6;0 bis 15;11). Die angegebenen Einsatzmöglichkeiten beinhalten Schulpsychologie, Berufs- und Bildungsberatung und Klinische Psychologie. Es wird eine Profilinterpretation auf Untertestebene angeraten. Von der Bestimmung eines Gesamt-IQ raten die Testautoren ab. Als Globalmaß wird dagegen zum Beispiel die Intelligenzquantität (kognitive Mindestfähigkeit) angeboten. Die im

Manual angeführten Untersuchungen erlauben eine Einschätzung verschiedener Reliabilitäts- und Validitätsaspekte. Klassische Nachweise der Kriteriumsvalidität finden sich nicht. Weiterhin ist anzumerken, dass einige der angegebenen Studien Teil von unveröffentlichten Abschlussarbeiten sind, die somit nicht frei zugänglich sind.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Die Autoren definieren Intelligenz als die „Gesamtheit aller kognitiven Voraussetzungen, die notwendig sind, um Wissen zu erwerben und Handlungskompetenz zu ent-

wickeln“ (S. 26). Weiterhin geben sie an, die Testaufgaben seien inhaltlich an Wechslers Intelligenzmodell orientiert (S. 6) und dienten vor allem förderungsorientierten diagnostischen Prozessen. Schließlich weisen die Autoren darauf hin, dass die Annahme eines g-Faktors im AID 3 nicht gerechtfertigt sei. Die Autoren grenzen ihr Testmodell zudem von anderen Intelligenzstrukturtheorien ab. Aus unserer Sicht erschwert dies die Vergleichbarkeit mit den Ergebnissen anderer Tests. Zudem ist es so auch nur schwer möglich, aktuelle Befunde der Intelligenzforschung basierend auf international anerkannten Modellen zu berücksichtigen. Eine Stärke des AID 3 ist die adaptive Vorgabe von Items, was durch die Konstruktion der meisten Untertests (UT) auf Basis des Rasch-Modells ermöglicht wurde.

Objektivität

Es existieren detaillierte Hinweise zur Testdurchführung, teilweise wörtliche Instruktionvorgaben und Angaben zur Gestaltung der Testsituation, sodass die Durchführungsobjektivität für die Standardversion des AID 3 gegeben ist. Außer bei den UT 6, 9, 11 und 10a, bei denen der Testleiter einen gewissen Ermessensspielraum bei der Antwortbewertung besitzt, ist die Auswertungsobjektivität durch Verrechnungsvorschriften ebenfalls gegeben. Die Interpretationsobjektivität wird durch Eichung der UT und Zusatztests (ZT) erhöht. Die Interpretationsobjektivität könnte noch gesteigert werden, wenn auch in den Fallbeispielen die Messgenauigkeit berücksichtigt würde und Hinweise auf die Berechnung individueller Konfidenzintervalle vorhanden wären. Hierzu gibt es Tabellen mit Standardschätzfehlern. Ausführliche Schilderungen zum Protokollbogen, Kommentare zur Benutzung der Eichentabelle sind ebenfalls im Hinblick auf die Objektivität hilfreich.

Normierung

Die Normen ($N = 2165$ aus Deutschland und Österreich, T-Werte) stammen aus den vorgesehenen Zielpopulationen. Die Datenerhebung erfolgte anhand eines Klumpenerhebungsverfahrens. Die Normstichproben eignen sich für altersbezogene Auswertungen im Jahresabstand. Zwischen benachbarten Altersgruppen kommt es aufgrund der breiten Altersbereiche (1 Jahr) teilweise allerdings zu deutlichen Differenzen bzgl. der Normwerte. Das im Manual empfohlene Interpolieren kann aufgrund anzunehmender unterschiedlicher Vorgehensweisen die Interpretationsobjektivität einschränken. Hier wäre ein Beispiel ratsam. Geschlechtsspezifische Normierungen werden nur für zwei UT (Alltagswissen und Angewandtes Rechnen) vorgenommen. Die übrigen empirischen Unterschiede

zwischen den Geschlechtern werden als Zufallsergebnisse oder Auswirkungen mangelnder Repräsentativität als nicht bedeutsam verworfen. Unter Berücksichtigung des enormen Erhebungsaufwands kann die Normierung insgesamt als gelungen angesehen werden.

Zuverlässigkeit

Es finden sich Reliabilitätsschätzungen, die sowohl Konsistenz als auch Stabilität der Testwertinterpretation nahelegen. Die Reliabilität wird über die Standardschätzfehler der Fähigkeitsparameter angegeben. Diese werden tabellarisch für die einzelnen UT dargestellt. Reliabilitätsschätzungen im Sinne der klassischen Testtheorie liegen lediglich für AID und AID 2 vor. Entsprechende Ergebnisse werden im Manual berichtet. Dies ist vor allem für die Retestreliabilität von Bedeutung. Aufgrund der vorgenommenen Verbesserungen im AID 3 kann jedoch eine mindestens vergleichbare Stabilität vermutet werden. Die Bestimmung der Messgenauigkeit für einzelne Fähigkeitsbereiche aufgrund der eingesetzten Item-Response-Modellierung ist wieder eine Stärke des Verfahrens. Die Güte dieser Schätzungen lässt sich aufgrund der Breite der Konfidenzintervalle der einzelnen Subtests beurteilen. Diese können manuell oder mittels des separaten Auswerteprogramms berechnet werden.

Gültigkeit

Auch zur Validität finden sich oft Angaben, die sich auf Vorgängerversionen beziehen. Bezüglich der Konstruktvalidität werden aus Studien selektiv Korrelationskoeffizienten (z.B. mit Persönlichkeit, Leistungsmotivation, Konzentration und räumlichem Vorstellungsvermögen) aufgeführt, so dass eine Beurteilung dieses Aspekts für alle UT ohne Sekundärliteratur nicht möglich ist. Angaben zur Kriteriumsvalidität, Sensitivität oder Spezifität fehlen. Die Analysen zur faktoriellen Validität legen nahe, dass sich im AID 2 anerkannte andere Intelligenzstrukturtheorien nicht widerspiegeln. Zudem scheinen die hohen Korrelationen zwischen den drei vorgeschlagenen Bereichen Wahrnehmen, Merken und Verarbeiten eher die Existenz eines g-Faktors zu stützen und die in der PCA angenommene Unkorreliertheit (Varimax-Rotation) zu widerlegen.

Weitere Gütekriterien

Hinsichtlich der Störanfälligkeit wird im Handbuch angemessen auf zu berücksichtigende Zustände der Klienten und situative Faktoren eingegangen. Es werden zum Beispiel Hinweise für die Testvorbereitung und die Gestaltung der Testsituation gegeben. Insgesamt erscheint daher die Störanfälligkeit als gering.

Bezüglich des Gütekriteriums der Unverfälschbarkeit wäre ein „faking good“ nur bei vorheriger Kenntnis der Testaufgaben denkbar; ansonsten ist dieser Faktor bei

Intelligenztests generell ausgeschlossen. „Faking bad“ kann eher als unwahrscheinlich betrachtet werden. Eine Übereinstimmung der Zahlenrelationen der Testwerte mit den Relationen der beobachtbaren Verhaltensweisen (innerhalb und zwischen Pb) ist gegeben. Die meisten Subtests sind nach dem Rasch-Modell skaliert. Das Kriterium der Skalierung ist so als erfüllt anzusehen.

Das adaptive Testen führt zu einem ökonomischen Vorgehen. Bei Anschaffung des verfügbaren Auswerteprogramms gestaltet sich auch die Auswertung als recht bequem.

Abschlussbewertung/Empfehlung

Die adaptive Vorgabe im Sinne eines „branched testings“ ist eine Stärke des Verfahrens. Die Überarbeitung des Testmaterials kann als gelungen bezeichnet werden. Sehr begrüßenswert ist die Aufnahme eines Untertests, der das Konstrukt „Reasoning“ berücksichtigt.

Kritisch wird die theoretische Abgrenzung der untersuchten Kognitionen von bestehenden Intelligenzstrukturtheorien gesehen. Die Autoren schildern hier ein pragmatisches Vorgehen und legen die inhaltlichen Gestaltungsprinzipien der Untertests ausführlich dar. Eine empirische Absicherung, welche Untertests oder spezifischen Maße bei welchen diagnostischen Fragestellungen speziell zu berücksichtigen sind oder eine besonders gute prognostische Validität aufweisen, fehlt jedoch teilweise. Fehlende Angaben zur Sensitivität, Spezifität und Kriteriumsvalidität sind weitere wichtige Limitationen. Wünschenswert wären Nachweise mit Kriterien (z.B. Berufserfolg), die sich auf die einzelnen intendierten Testeinsatzzwecke (z.B. Berufsberatung) beziehen (Ziegler, 2014). Problematisch ist die nicht näher spezifizierte Verwendung der unterschiedlichen Testkennwerte, bezogen auf spezifische diagnostische Fragestellungen, auch wenn einige der angegebenen Kennwerte von den Testautoren eigentlich gar nicht erst zur Verwendung empfohlen werden. Vor allem die fehlenden Nachweise klinischen Nutzens erschweren die fragebezogene Auswahl unter diesen Testkennwerten. Abschließend ist festzustellen, dass vor allem theoretische Fundierung und fehlende Belege für die Kriteriumsvalidität den Gesamteindruck trüben.

AID-3	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Allgemeine Informationen, Beschreibung und diagnostische Zielsetzung		X		
Objektivität		X		
Zuverlässigkeit		X		
Validität			X	

Diese Testrezension wurde im Auftrag des Testkuratoriums der Föderation deutscher Psychologinnenvereinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2009, 2010) erstellt.

Testkuratorium. (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie*, 34, 470–478.

Testkuratorium. (2010). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Psychologische Rundschau*, 61, 52–56.

Literatur

Ziegler, M. (2014). Stop and state your intentions! Let's not forget the ABC of test construction. *European Journal of Psychological Assessment*, 30, 239–242. <https://doi.org/10.1027/1015-5759/a000228>

Testinformationen

Kubinger, K. D. & Holoher-Ertl, S. (2014). *Adaptives Intelligenz Diagnostikum 3*. Göttingen: Beltz Test Gesellschaft.
 Bezugsquelle: Testzentrale; Herbert-Quandt-Str. 4; 37081 Göttingen
 Test komplett bestehend aus: Manual, Testanweisung, Spiralheft zu Untertest 2, Spiralheft zu Untertest 3, Bildkarten zu Untertest 4, Karte zu Untertest 7, Puzzle zu Untertest 8, Würfel zu Untertest 10, Spiralheft zu Untertest 10, Spiralheft zu Untertest 12, Holzplättchen zu Untertest 12, Bildtafel zu Untertest 5a, Bildkarten zu Untertest 5c, Spiralheft zu Untertest 5c, 10 Protokoll- und Arbeitsblätter, Auswertungsschabloneinsatz und Koffer: 1.248 Euro

Bitte zitieren Sie diesen Artikel wie folgt: Ziegler, M. & Reichert, A. (2017). TBS-TK Rezension: „Adaptives Intelligenz Diagnostikum 3 (AID 3)“. *Psychologische Rundschau*, 68, 237–239.

DOI: 10.1026/0033-3042/a000368