

Modularer Kurzintelligenztest (M-KIT)

KARIN FUNSCH, Philipps-Universität Marburg
LENA ROEMER, GESIS – Leibniz-Institut für Sozialwissenschaften

Beschreibung des Tests und seiner diagnostischen Zielsetzung

Der M-KIT soll der Erfassung der fluiden Intelligenz bei deutschsprachigen Personen ab 15 Jahren dienen. Er wird als ökonomisches, genderfaires und flexibel anwendbares Verfahren beschrieben. Eingesetzt werden kann der Test laut Autor*innen zur Eignungsdiagnostik, in der Bildungsberatung und im Rahmen wissenschaftlicher Untersuchungen.

Der Test besteht aus sechs Aufgabengruppen, die sich in fünf Module gruppieren lassen. Modul A (ausbalanciert) setzt sich aus drei Aufgabengruppen zusammen, welche neben abstrakt logischem Denken auch verbale, figural-bildhafte und numerische Aspekte erfordern. Modul K (kernfokussiert) besteht ebenfalls aus je einer verbalen, figural-bildhaften und numerischen Aufgabengruppe. Im Unterschied zu Modul A sind die Aufgaben in Modul K stärker sequenziell-deduktiv ausgerichtet, was die fluid-schlussfolgernde Komponente der Intelligenz stärker betonen soll. Die weiteren drei Module setzen sich aus je zwei Aufgabengruppen mit der gleichen Materialprägung zusammen (Modul verbal, figural-bildhaft, numerisch). Da alle fünf Module fluide Intelligenz messen sollen, soll jedes Modul separat durchgeführt werden können. Der Gesamttest dauert mit Instruktionen ca. 85 Minuten, für die fünf Module variiert die Dauer zwischen ca. 25 bis 50 Minuten. Der Test kann als Einzel- oder Gruppentestung, papierbasiert oder computergestützt durchgeführt werden.

Bewertung des Informationsgehalts der Verfahrenshinweise

Informationen zu den theoretischen Grundlagen, der praktischen Durchführung und dem Konstruktionsprozess werden detailliert beschrieben. Zur testtheoretischen Überprüfung werden im Manual mehrere empirische Untersuchungen aufgeführt, doch fehlen einige Informationen; etwa zu den Rahmenbedingungen bei der Datenerhebung, zum Stichprobenplan und teils zu durchgeführten statistischen Analysen. Die Anzahl der Proband*innen lässt in der Regel statistisch gesicherte Aussagen über die jeweilige Fragestellung zu. Eine tabellarische Übersicht zu den nach Diagnostik- und Testkuratorium (DTK) notwendigen Informationen ist nicht enthalten.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Als theoretische Grundlage des M-KIT wird ein integratives »Operationalisierungsmodell der fluiden Intelligenz« beschrieben. Fluide Intelligenz (Gf) ist darin

definiert als die »Fähigkeit, Problemlösungen ohne Zuhilfenahme von Vorwissen zu erschließen« (S. 13). Sie wird mit schlussfolgerndem Denken gleichgesetzt und als Kern der allgemeinen Intelligenz angesehen. Laut dem Modell ist die Messung fluider Intelligenz immer – mehr oder weniger stark – auch von anderen Fähigkeiten, Eigenschaften oder Kompetenzen beeinflusst (S. 16), was zu konstruktfernen Prägungen führen soll.

Visualisiert wird das Modell mit einer faktorenbasierten Radex-Darstellung. In deren Zentrum steht der fluide Intelligenzfaktor. Darum gruppiert werden auf entsprechenden Achsen die verbalen, figural-bildhaften und numerischen Aufgabengruppen dargestellt. Je stärker die Materialprägung der Aufgabengruppen ist, desto größer ist ihre Entfernung vom Zentrum. Im Manual werden Parallelen des Operationalisierungsmodells mit insgesamt neun Modellen zu (fluider) Intelligenz skizziert. Als Mehrwert gegenüber bestehenden Tests werden genderfaire Konstruktionsprinzipien sowie die ökonomische, modularisierbare Erfassung angeführt. Für die Modulwahl werden im Manual Entscheidungshilfen präsentiert. Diese fußen vor allem auf kontextuellen Randbedingungen; inhaltliche Kontrastierungen der Module fallen knapp aus.

Objektivität

Zur Gewährleistung der Objektivität werden Anweisungen zur Durchführung, Auswertung und Interpretation des Verfahrens dargestellt. Präzise beschrieben werden Rahmenbedingungen zur Testvorbereitung und -durchführung. Auch sind Anweisungen für Gruppentestungen enthalten, wobei Angaben zur maximalen Gruppengröße fehlen. Die Auswertung der Papier-Bleistift-Version ist genau erläutert. Weniger präzise sind Anweisungen zur Interpretation der Ergebnisse. Die Transformation der Rohwerte in Standardwerte sowie deren Einordnung werden noch ausreichend erläutert, und es bestehen Normtabellen zweier Referenzgruppen. Es wird darauf hingewiesen, dass in beiden Referenzgruppen überdurchschnittlich hohe Gf-Werte vorliegen, doch es bleibt unklar, wie dies in der Interpretation konkret zu berücksichtigen ist. Auch wann welche Referenzgruppe heranzuziehen ist, wird nicht erläutert. Darüber hinaus wären genauere Ausführungen zur Berücksichtigung der Messgenauigkeit bei der Ergebnisinterpretation angebracht. Beschrieben werden Vergleiche von Testergebnissen verschiedener Personen miteinander oder mit Cut-off-Werten; exakte Konfidenzintervalle sind nur indirekt ermittelbar. Zuletzt wären für Anwender*innen auch inhaltliche Interpretationshilfen und Fallbeispiele hilfreich.

Testbeurteilungssystem – Diagnostik- und Testkuratorium der Föderation Deutscher Psychologinnenvereinigungen

Modularer Kurzintelligenztest (M-KIT)	Die TBS-DTK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Informationsgehalt der Verfahrensweise		●		
Objektivität			●	
Zuverlässigkeit		●		
Validität		●		
			ja	nein
In den Verfahrenshinweisen ist verzeichnet, wo die nach dem DTK-Testinformationsstandard notwendigen Informationen zu finden sind.				●

Normierung

Die Daten zur Normierung und empirischen Prüfung wurden von Februar 2013 bis Mai 2014 erhoben. Normiert wurde der M-KIT an 608 Gymnasiast*innen der 11. und 12. Klasse im Alter von durchschnittlich ca. 17 Jahren. Zudem kann eine Studierendenstichprobe ($n = 360$; $M \approx 22$ Jahre) als Referenzgruppe herangezogen werden. Beide Gruppen sind nicht repräsentativ für die Zielpopulation der deutschsprachigen Allgemeinbevölkerung. Das Fehlen von Normwerten für die gesamte Zielgruppe ist ein erheblicher Mangel des Verfahrens. Dem gender fairen Konstruktionsprinzip des M-KIT entsprechend (s. u.) liegen keine geschlechtsspezifischen Normen vor. Die Rohwerte werden nicht in IQ-Testwerte verrechnet, sondern in durch Flächentransformation ermittelte Standardwerte (SW und PR).

Zuverlässigkeit

Entsprechend den status- und prognosediagnostischen Zielen des M-KIT werden Angaben zur internen Konsistenz und Retest-Reliabilität (rund zwei Monate) gemacht. Die Reliabilitätsschätzungen erfolgten – wie die gesamte empirische Überprüfung inklusive Normierung – an (Teil-)Stichproben, die nicht repräsentativ für die Allgemeinbevölkerung sind ($n = 608$ Gymnasiast*innen, $n = 105$ Studierende, $n = 52$ Auszubildende; sowie eine Gesamtstichprobe mit $N = 1.054$).

Die Werte liegen zwischen .78 und .94 für die interne Konsistenz und .75 und .90 für die Retest-Reliabilität. Die höheren Schätzwerte bestehen für den Gesamttest (α und $r_{tt} \geq .90$) und Modul K (kernfokussiert; $\alpha \geq .90$ und $r_{tt} = .86$).

Gültigkeit

Die Autor*innen begründen die Inhaltsvalidität des M-KIT theoretisch u. a. durch die Kompatibilität mit anderen Intelligenzmodellen. Die theoretischen Rationale, die Item-Universum und -Auswahl zugrunde liegen, werden summarisch beschrieben, jedoch ohne empiri-

sche Angaben. Im Hinblick auf die faktorielle Validität legt die Passung eines Einfaktormodells nahe, dass die Aufgabengruppen zu einem Gesamtscore verrechenbar sind. Latente Korrelationen nahe 1 weisen darauf hin, dass sich die fünf Modulscores kaum voneinander differenzieren lassen.

Mit Blick auf die konvergente Validität sprechen hohe Korrelationen mit dem Grundmodul des »I-S-T 2000R« (Lipmann, Beauducel, Brocke & Amthauer, 2007) dafür, dass der M-KIT und dessen fünf Module »Schlussfolgerndes Denken« erfassen. Wegen fehlender Nachweise der diskriminanten Validität wird jedoch die empirische Abgrenzung der Modulscores nicht ganz klar. Etwa wären Angaben zu Korrelationen mit den Subfaktoren des »I-S-T 2000R« (kristallisierte Intelligenz) oder mit anderen kognitiven Verfahren (z. B. Konzentration) hilfreich, um die spezifischen empirischen Bedeutungen der fünf Modulscores zu verdeutlichen.

Zur Kriteriumsvalidität werden zufriedenstellend hohe Korrelationen mit diversen Leistungsindikatoren (Schulnote, Ausbildungsnote, Leistungsbeurteilung) angegeben ($.13 \leq r \leq .51$).

Angaben zur inkrementellen und auch prognostischen Validität fehlen.

Weitere Gütekriterien

Die Genderfairness wird im Manual anhand nur kleiner Mittelwertsunterschiede zwischen Männern und Frauen festgemacht (z. B. auf Modulebene $|d| \leq .11$); Messinvarianz auf Aufgaben-/Modulebene oder differenzielle Validität für Geschlechter werden jedoch nicht geprüft. Die per Fragebogen erfasste Akzeptanz (Kersting, 2008) des M-KIT war im Mittel eher hoch. Zur Störanfälligkeit wird im Manual auf situative Rahmenbedingungen oder den Umgang mit Störungen eingegangen, sodass diese gering erscheint. Zur Skalierbarkeit wird faktorenanalytisch dargestellt, dass sich die sechs Aufgabengruppen zu einem Gesamtscore verrechnen lassen; doch fehlen Überprüfungen auf Item-Ebene. Zur Verfälschbarkeit werden keine Untersuchungen aufgeführt. Es ist davon auszugehen, dass zumindest »faking bad« möglich ist.

Abschlussbewertung

Beim M-KIT handelt es sich um ein modulares Verfahren zur Messung der fluiden Intelligenz. Stärke des Tests ist insbesondere, dass die fünf alternativ einsetzbaren Module eine recht zeitökonomische Erfassung des Schlussfolgernden Denkens (Gf) ermöglichen. Gleichwohl geht der Einsatz eines einzelnen Moduls mit verminderter Reliabilität und so mit einzelfalldiagnostischen Einschränkungen einher.

Auch die Integration des Operationalisierungsmodells in bestehende Intelligenztheorien ist positiv anzusehen. Jedoch würden hier eine detaillierte konzeptuelle Ausarbeitung und weitere empirische Überprüfungen die Validität der Testwerte belegen und ggf. weitere Hilfen

für die Modulwahl und den Einsatz des aufwendigen Gesamttests liefern.

Der Konstruktionsansatz, Mittelwertsdifferenzen zwischen Geschlechtern zu reduzieren, ist eine Besonderheit des Tests. Um die Genderfairness weiter beurteilen zu können, wären Angaben zur differenziellen Validität wünschenswert.

Einschränkend ist insbesondere die Nicht-Repräsentativität der Normierungsstichproben. Da Testwerte nur mit jungen Erwachsenen mit überdurchschnittlich hohen Gf-Werten verglichen werden können, sind einzelfalldiagnostische Anwendungen nicht in allen Gruppen der Allgemeinbevölkerung möglich.

Literatur

Kersting, M. (2008). Zur Akzeptanz von Intelligenz- und Leistungstests. *Report Psychologie*, 33(9), 420–433.

Liepmann, D., Beauducel, A., Brocke, B. & Amthauer, A. (2007). *Intelligenz-Struktur-Test 2000 R. Manual* (2., erweiterte und überarbeitete Aufl.). Göttingen: Hogrefe.

Testinformation

Dantlgraber, M. (2015). *Modularer Kurzintelligenztest (M-KIT)*. Göttingen: Hogrefe.

Bezugsquelle

Testzentrale Göttingen, Robert-Bosch-Breite 25, 37079 Göttingen. Test komplett 295,00 €, 5 Antwortbogen (für 1 der 6 Teile) 8,70 €, 5 Auswertebogen 9,50 € (Netto-Preise).

Diese Testrezension wurde im Auftrag des Diagnostik- und Testkuratoriums der Föderation Deutscher Psychologengruppen (DGPs und BDP) gemäß TBS-DTK (Diagnostik- und Testkuratorium, 2018) erstellt.

Diagnostik- und Testkuratorium (2018). TBS-DTK. Testbeurteilungssystem des Diagnostik- und Testkuratoriums der Föderation Deutscher Psychologengruppen. Revidierte Fassung vom 3. Januar 2018. *Report Psychologie*, 43(3), 106–113.

Diagnostik- und Testkuratorium (2018). TBS-DTK. Testbeurteilungssystem des Diagnostik- und Testkuratoriums der Föderation Deutscher Psychologengruppen. Revidierte Fassung vom 3. Januar 2018. *Psychologische Rundschau*, 69(2), 109–116.

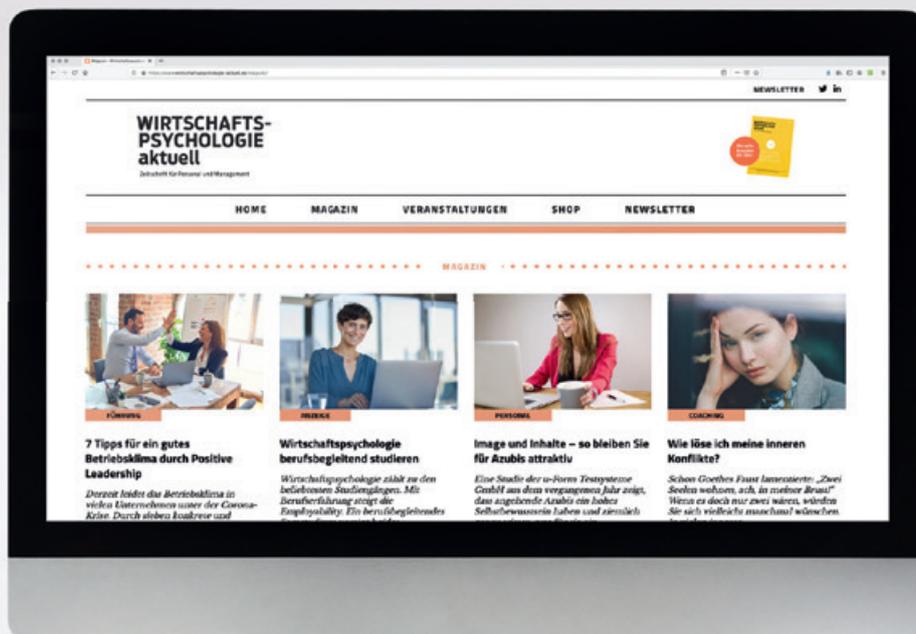
Bitte zitieren Sie diesen Artikel wie folgt:

Funsch, K. & Roemer, L. (2024). TBS-DTK-Rezension: »Modularer Kurzintelligenztest (M-KIT)« *Report Psychologie*, 49(5), 23–25.

**WIRTSCHAFTS-
PSYCHOLOGIE
aktuell**

Jetzt neu im Online-Magazin:
Wöchentlich aktuelle Artikel mit praktischen
Tipps – kurz und prägnant aufbereitet.

❖ www.wirtschaftspsychologie-aktuell.de



110903 - Foto: Freepik