

Testvorstellung der dgp Testbatterie K4

Amelie Kleinmanns & Anna-Lena Jobmann

Allgemeine Informationen

Der K4 ist eine Testbatterie zur Beurteilung der kognitiven Eignung für Fach- und Führungskräfte sowie den höheren Dienst, die seit September 2020 eingesetzt wird. Basierend auf einem Intelligenzstrukturmodell erfasst der Test verbale und numerische Verarbeitungskapazität mit jeweils vier Untertests. Für die Testbatterie kann ein Gesamtwert ausgegeben und interpretiert werden. Die Ergebnisse der Verarbeitungskapazitäten numerisch und verbal können ebenfalls ausgewiesen werden, aber man kann sie nur mit Vorsicht interpretieren. Der K4 kann als Einzel- oder Gruppentestung durchgeführt werden und dauert ca. 2 Stunden und 15 Minuten. Die einzelnen Aufgabengruppen sind zeitlich begrenzt.

Veränderungen zur Vorgänger Testbatterie K3

Die Vorgängerversion K3 wurde einer gründlichen psychometrischen Überprüfung auf Basis der Klassischen Testtheorie und der Item Response Theorie unterzogen. Teil der psychometrischen Überprüfung waren die Berechnung der Itemschwierigkeiten, die Reliabilitätsberechnung, Überprüfung der Eindimensionalität der einzelnen Untertests, sowie eine Fairnessanalyse der Items (Differential Item Functioning). Auch wenn der K3 schon sehr hohen psychometrischen Anforderungen gerecht wurde, ließen sich noch Ansätze zur Optimierung finden. So wurden einige Items

aufgrund ihrer zu geringen Schwierigkeit durch neue Items ersetzt. Außerdem wurden vereinzelt Aufgaben ausgetauscht, da sie Differential Item Functioning mit kleiner Effektstärke aufwiesen. Neben diesen psychometrischen Ansatzpunkten wurden Aufgaben auf ihre inhaltliche Aktualität hin überprüft und bei Bedarf angepasst oder ausgetauscht.

Mit einer konfirmatorischen Faktorenanalyse wurde die zweidimensionale Struktur (verbale und numerische Verarbeitungskapazität) des K3 überprüft. Der Modellfit konnte als zufriedenstellend mit $\chi^2 = 548.40$, $df = 34$, $p = .000$, $CFI = .940$, $RMSEA = .071$ [.066; .076] eingestuft werden. Neben diesem Ausgangsmodell wurden weitere Modelle aufgestellt, um das Modell mit dem besten Modellfit in Bezug auf die zweidimensionale Struktur des K3 zu finden. Der beste Fit konnte erreicht werden, nachdem die beiden Untertests Schlüsse und Tabellen & Statistiken entfernt wurden. Dieser kann als gut bezeichnet werden: $\chi^2 = 210.890$, $df = 19$, $p = .000$, $CFI = .970$, $RMSEA = .058$ [.051; .065]. Daher wurden beide Aufgabengruppen aus dem K4 entfernt.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Die theoretische Grundlage des K4 und somit Ausgangspunkt seiner Testkonstruktion ist das Berliner-Intelligenz-Strukturmodell (BIS-Modell). Im BIS-Modell sind

Fähigkeiten hierarchisch angeordnet, d.h. sie befinden sich auf unterschiedlichen Generalitätsebenen. Auf der obersten Ebene befindet sich die Allgemeine Intelligenz (AI). Sie spiegelt eine sehr allgemeine Sicht auf die kognitiven Fähigkeiten einer Person wider. Auf der zweiten Ebene wird eine differenzierte Sichtweise auf unterschiedliche kognitive Fähigkeiten einer Person möglich. Diese werden den Modalitäten operativ und inhaltlich zugeordnet. Operative Fähigkeiten sind Verarbeitungskapazität, Einfallsreichtum, Merkfähigkeit und Bearbeitungsgeschwindigkeit. Inhaltgebundene Fähigkeiten sind verbales (sprachgebundenes), numerisches (zahlelgebundenes) und figürliches (figural-bildhaftes) Denken (Jäger, Süß, & Beauducel, 1997). Der K4 kombiniert zwei inhaltsbezogene Fähigkeiten mit der operativen Fähigkeit der Verarbeitungskapazität. Hierbei geht es um die Verarbeitung komplexer Informationen zur Lösung von Aufgaben unter Zurhilfenahme von formallogischem Denken und korrekter Beurteilung von Informationen bezogen auf Sprache und Zahlen. Jäger, Süß und Beauducel (1997) sehen gerade die Verarbeitungskapazität als besonders wesentlich für die allgemeine Intelligenz einer Person an. Diese wiederum zeigt einen deutlichen Zusammenhang mit dem Berufserfolg (Kramer, 2009).

Objektivität

Um die Durchführungsobjektivität zu gewährleisten sind die Testinstruktionen standardisiert. Die Instruktionen sind schriftlich in den Testheften vorhanden und werden durch die/den Testanweiser*in vorgelesen. Neben den Instruktionen wird jede Aufgabengruppe durch mindestens ein Beispiel erläutert. Des Weiteren wird der K4 nur durch Diplom-Psycholog*innen, Psycholog*innen mit Master of Science

Psychologie und psychologisch-technische Assistent*innen angewiesen. Die Auswertung erfolgt automatisiert und computer-gestützt mit entsprechend hoher Auswertungsobjektivität. Als Interpretationshilfe werden Bewerber*innen Empfehlungsgerade zugeordnet. Diese Empfehlungsgerade mit den Werten 1-5 umfassen einen a priori festgelegten Bereich an Standardwerten. Diese werden mit Hilfe einer Tabelle auf den Profilkarten (der Ergebnisdarstellung) visualisiert. Eine Eingruppierung in die Gruppen: Leistung entspricht „...den Anforderungen“, „...weitgehend den Anforderungen“, „nur teilweise den Anforderungen“, „weitgehend nicht den Anforderungen“ und „nicht den Anforderungen“ bietet eine weitere Hilfe bei der Interpretation der Testergebnisse.

Normierung

Für den K4 liegt derzeit eine Gesamtnorm in Form von Standardwerten mit einem Mittelwert von $M = 100$ und einer Standardabweichung von $SD = 10$ vor. Die Normstichprobe umfasst 673 Bewerber*innen für den höheren Öffentlichen Dienst in Deutschland. 54,83% der Personen sind männlich, 45,17% weiblich. Die Norm wurde im September 2020 erhoben.

Zuverlässigkeit

Die Berechnung der internen Konsistenzen mithilfe von Cronbach's Alpha ergeben für Verarbeitungskapazität numerisch $\alpha = .83$ und verbal $\alpha = .80$. Die Interne Konsistenz der gesamten Testbatterie liegt bei $\alpha = .87$. Die Split-Half Reliabilitäten liegen für die Skalen numerisch bei $rr = .80$ und verbal $rr = .66$. Die gesamte Batterie hat eine Split-Half Reliabilität von $rr = .80$.

Gültigkeit

Konstruktvalidität

Mit Hilfe von konfirmatorischen Faktorenanalysen wurde an Hand der Normstichprobe überprüft, ob die durch die theoretischen Annahmen postulierte Struktur der Daten mit der tatsächlichen Struktur der Daten

übereinstimmt (Kersting, 2009). Es zeigt sich eine gute Modellpassung ($ML = 482.13$, $df = 315$, $p = .000$, $CFI = .952$, $RMSEA = .028$, $90\% CI [.023; .033]$). Abbildung 1 zeigt die postulierte Faktorstruktur des K4.

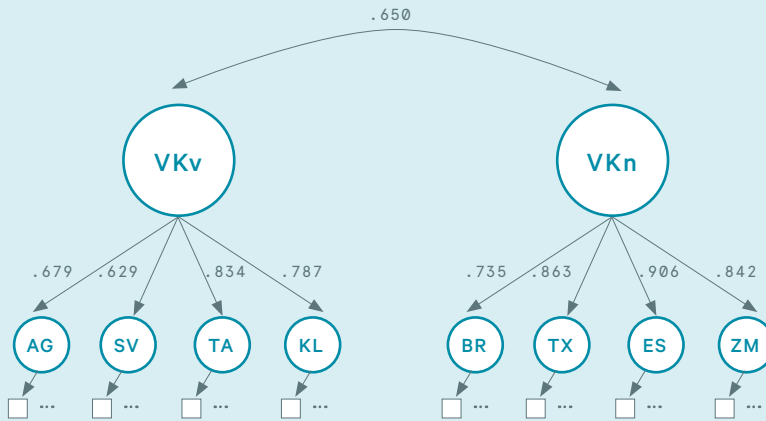


Abbildung 1. Faktorenmodell der Testbatterie K4

Weitere Gütekriterien

Ökonomie

Mit einer Gesamtzeit von ca. 2 Stunden 15 Minuten ist der K4 eine zeitökonomische Testbatterie, um die numerische und verbale Verarbeitungskapazität mit jeweils vier unterschiedlichen Aufgabengruppen abzudecken. Des Weiteren kann der K4 als Gruppentestung durchgeführt werden. Je nach zur Verfügung stehender Raumkapazität können somit viele Personen gleichzeitig getestet werden. Da in die Testhefte selbst nicht geschrieben werden darf, sind diese wiederverwendbar, nur die Antwortbögen und das Notizpapier können pro Person nur einmal genutzt werden. Der Materialaufwand ist daher als eher gering einzustufen.

Testsicherheit

Der K4 ist eine unveröffentlichte Testbatterie. Die Aufgaben sind den zu testenden Personen nur vor Ort zugänglich, während Aufsichtspersonen verhindern, dass Testhefte entwendet oder Bildmaterial von den Aufgaben gemacht wird. Das Abschreiben bei einem Tischnachbarn wird durch die Benutzung zweier pseudoparalleler Testformen unterbunden. Verfälschungen nach unten sind wie bei allen Leistungstests möglich. Um die Verfälschbarkeit nach oben durch Trainingseffekte zu vermeiden, sind Personen nach Bearbeitung der Testbatterie für sechs Monate von der erneuten Teilnahme am Test ausgeschlossen.

Fazit

Die Testbatterie K4 erlaubt die faire, ökonomische und unverfälschte Erfassung kognitiver Fähigkeiten von Fach- und Führungskräften sowie dem höheren Dienst. Die Objektivität in Durchführung, Auswertung und Interpretation ist gesichert. Die vorliegenden Normen sind aktuell und wurden auf Basis einer Stichprobe von Bewerber*innen aus den relevanten Berufen berechnet. Die Reliabilität der Testbatterie kann als gut bewertet werden, sodass eine Interpretation des Gesamtergebnisses empfohlen werden kann. Die Evidenz für faktorielle Validität des Tests liegt vor und spricht für ein gutes Auswahlinstrument.

Die Qualität der Testbatterie wird in zukünftigen Untersuchungen weiter gesichert: Detaillierte und aktualisierte Analysen zur Fairness werden durchgeführt, eine Überprüfung der Normen des K4 erfolgt im kommenden Jahr und Bewährungskontrollen werden durchgeführt, sobald hierfür Daten vorliegen.

Literatur:

- Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). BIS-TEST Berliner Intelligenzstruktur-Test – Form 4. Göttingen: Hogrefe.
- Kersting, M. (2009). Die DGP Testverfahren - Ein kurzer Rückblick und eine aktuelle Studie zur Konstrukt und Kriteriumsvalidität des BIS-r-DGP Tests. *dgp Informationen*, S. 22-37.
- Kramer, J. (2009). Allgemeine Intelligenz und beruflicher Erfolg in Deutschland. *Psychologische Rundschau*, 60(2), S. 82–98.

Kontakt:

M. Sc. Psychologie Amelie Kleinmanns
kleinmanns@dgp.de

Deutsche Gesellschaft für Personalwesen e. V.
Hohenzollernstr. 11-13, 40211 Düsseldorf

Dipl.-Psych. Dr. Anna-Lena Jobmann
jobmann@dgp.de

Deutsche Gesellschaft für Personalwesen e. V.
Kantstr. 153, 10623 Berlin