

NEUERE TESTVERFAHREN

Petermann, F. (Hrsg.) unter Mitarbeit von A. Toussaint (2010). **K-CAB. Kaufman – Computerized Assessment Battery**. Deutsche Adaptation des französischen Originalverfahrens K-Classic. Frankfurt: Pearson Assessment & Information. Anschaffungspreis 500,00 €, Preis pro Durchführung 2,00 €. ¹

Bei der K-CAB handelt es sich um ein computergestütztes Diagnoseinstrument zur Erfassung von Aufmerksamkeit und kognitiven Fähigkeiten bei Kindern im Alter von 6;5 bis 10;11 Jahren. Die Originalversion wurde in Frankreich entwickelt, normiert und unter dem Namen K-Classic (Kaufman & Kaufman, 2007) veröffentlicht.

Theoretischer Hintergrund

Der Aufbau des Verfahrens orientiert sich an der neuropsychologischen PASS-Theorie (Das, Naglieri, Kirby, 1994). Gestützt auf Lurias Theorie der zerebralen Spezialisierung postuliert die PASS-Theorie drei Basiskomponenten der Intelligenz. *Planungsfähigkeit* als erste Komponente ist beim problemlösenden Denken gefordert. Sie dient dazu, Lösungsmöglichkeiten zu finden und abzuschätzen, ob diese für die Problemlösung geeignet sind. Die zweite Komponente *Aufmerksamkeit* wird benötigt, um über die erforderliche Zeit hinweg auf relevante Reize zu achten, irrelevante Reize auszublenden und nicht darauf zu reagieren. Aufmerksamkeit wird nicht den kognitiven Leistungen im engeren Sinn zugerechnet, sondern kennzeichnet im Rahmen der PASS-Theorie die Tonusregulation des zentralen Kortex. *Simultane und sequentielle Reizverarbeitung* bilden gemeinsam die dritte Komponente. Simultane Verarbeitung meint dabei die sinnvolle Gruppierung gleichzeitig vorhandener Informationen, sequentielle Verarbeitung die Integration der Reize in eine bestimmte Abfolge. Das Zusammenspiel aller drei Basiskomponenten soll die erfolgreiche Bewältigung einer Reihe kognitiver Anforderungen ermöglichen. Obwohl die Faktoren keine sprachlichen oder bildungsabhängigen Inhalte aufweisen und daher als kulturfair angesehen werden, sollen sie dennoch geeignet sein, schulische Leistungen vorherzusagen. Die K-CAB soll alle drei Basiskomponenten der PASS-Theorie und zusätzlich kristalline Intelligenz erfassen. Somit werden vier Intelligenzfaktoren des Cattell-Horn-Carroll-Modells (CHC-Modell; McGrew, 2005) bei der Testung berücksichtigt (s. Tab. 1).

Mit Hilfe der K-CAB sollen Aufmerksamkeitsprobleme und kognitive Störungen im Sinne eines Screenings identifiziert und somit einer eingehenden Untersuchung zugeführt werden. Als weiteren Nutzen nennen die Autoren die ökonomische Testdurchführung. Die fünf Untertests weisen Ähnlichkeiten mit Aufgaben anderer

1 Der Autor dankt Jana Schultz für ihre Mithilfe bei der Erprobung des Verfahrens.

Testverfahren auf. Für die deutsche Ausgabe wurden die Items des französischen Originals beibehalten und mit deutschen Instruktionen versehen. Lediglich der Untertest *Wortschatz* erhielt neue Items.

Tabelle 1: Testaufbau und Reliabilitäten der K-CAB und theoretische Einordnung der Testaufgaben

Untertests der K-CAB	Faktoren der PASS-Theorie	CHC-Faktoren (gemäß Manual)	Skalen der K-CAB		Reliabilität ^b
			AKI ^a	ISSV	
1. Aufmerksamkeit ^c (AUFM)	Aufmerksamkeit	Kein CHC-Faktor			.76 (.78) ^d
2. Wortschatz (WOS)	(keine direkte Entsprechung)	G _c	X		.91 (.86-.94)
3. Sequentielle Verarbeitung (SQV)	Sequentielle Verarbeitung	G _{sm}	X	X	.81 (.77-.88)
4. Begriffsbildung (BEG)	Planungsfähigkeit	G _f	X		.90 (.87-.92)
5. Simultane Verarbeitung (SIV)	Simultane Verarbeitung	G _v	X	X	.89 (.87-.91)

Anmerkungen: ^a Alle Abkürzungen siehe Text; ^b Bis auf *Aufmerksamkeit* werden hier Fisher-z-transformierte Durchschnittswerte für die Altersgruppen angegeben, in Klammern steht der Altersgruppenrange der Konsistenzwerte; ^c Die Nummerierung der Untertests entspricht der Reihenfolge der Aufgaben bei der Testdurchführung; ^d Retest-Reliabilität (bezieht sich auf französische Daten), im Manual ohne Nennung der Altersgruppe, in Klammern ein korrigierter Reliabilitätskoeffizient unbekannter Herkunft.

Installation und Betrieb

Der Test umfasst Manual (181 Seiten), Installations-CD und Schutzstecker. Für den Start des Testprogramms benötigt man den Schutzstecker, der an einen USB-Port angeschlossen wird und beim Kauf für 25 Testdurchführungen aufgeladen ist. Benötigt werden außerdem Lautsprecher oder Kopfhörer und Drucker. Um weitere Testdurchführungen zu erhalten, muss man den Verlag anrufen und eine 32-stellige Auftragsnummer durchgeben. Daraufhin bekommt man einen anderen, ebenso langen Code mitgeteilt, der wiederum manuell in den PC einzugeben ist, wodurch weitere Anwendungen aktiviert werden. Die Rechnung kommt konventionell auf dem Postweg. Weitere Schutzstecker können beim Verlag für 95,00 Euro bezogen werden.

Testaufbau, Material und Durchführung

(1) Der erste Untertest dient der Erfassung der *Aufmerksamkeit* und besteht aus zwei Teilen. Im ersten Teil werden innerhalb von drei Minuten auf dem Bildschirm 90 lächelnde und traurige Gesichter in unregelmäßiger Folge kurzzeitig gezeigt. Das Kind soll immer dann die Leertaste des PC's drücken, wenn das lächelnde Gesicht auftaucht, was bei 42 % der Reize der Fall ist. Damit sollen selektive *Aufmerksam-*

keit und Daueraufmerksamkeit gemessen werden. Im zweiten Teil werden in wechselnder Abfolge ebenfalls drei Minuten lang sieben verschiedene Gesichtsausdrücke gezeigt, und es soll jetzt bei allen Gesichtern außer dem lächelnden die Leertaste gedrückt werden. Dadurch werden höhere Anforderungen an die Reaktionshemmung gestellt als im ersten Teil, da jetzt in der Mehrzahl der Fälle die zuvor angebahnte Reaktion gehemmt werden muss. Gemessen werden in beiden Teilen die Zahl der falsch positiven und falsch negativen Antworten, Reaktionszeit und Reaktionszeitstreuung. Ausgegeben wird für Aufmerksamkeit aber nur ein Gesamtwert, der die vier Messkomponenten über beide Testteile hinweg additiv verknüpft. Es gibt keine Angaben darüber, ob hier Roh- oder Standardwerte zusammenfasst wurden. Der Untertest *Aufmerksamkeit* ist keinem Intelligenzfaktor des CHC-Modells zugeordnet und fließt auch nicht in die Indices AKI und ISSV ein.

(2) Im Untertest *Wortschatz* nennt eine Computerstimme Alltagsobjekte oder Begriffe der Allgemeinbildung. Das Kind muss aus einer Auswahl von jeweils sechs Abbildungen die richtige mit der PC-Maus anklicken. Bei fehlerhafter Antwort wird die Aufgabe unter Weglassen der falschen Alternative wiederholt. Die Lösung im ersten Durchgang ergibt zwei Punkte, die richtige Antwort im zweiten Durchgang bringt noch einen Punkt. Der Untertest soll erworbenes Wissen messen, ohne Anforderungen an den Aktivwortschatz zu stellen. Der Wissenserwerb ist von den jeweiligen Bildungsbedingungen abhängig. Der Untertest gilt als Indikator für kristalline Intelligenz G_c .

(3) Die *Sequentielle Verarbeitung* ähnelt „Wortreihe“ aus der K-ABC. Dem Kind wird eine Reihe von Gegenständen genannt. Anschließend soll es die entsprechenden Objekte in der richtigen Reihenfolge auf dem Bildschirm mit der PC-Maus anklicken. Der Untertest wird mit dem CHC-Faktor Kurzzeitgedächtnis G_{sm} in Verbindung gebracht.

(4) Bei der *Begriffsbildung* erscheinen auf dem Bildschirm nach und nach mehrere Objekte. Bei den ersten bekommt der Proband zusätzlich verbal die Information, ob das Objekt „dazu gehört“ oder „nicht dazu gehört“. Auf diese Weise soll er selbst das Klassifizierungsmerkmal herausfinden. Bei den nachfolgenden Objekten soll das Kind dann selbst über die Klassenzugehörigkeit entscheiden und mit der PC-Maus „Ja“ oder „Nein“ anklicken. Klassifikationskriterien sind Merkmale oder Merkmalskombinationen. Der Test soll Planungsfähigkeit messen und gemäß PASS-Theorie Zusammenhänge mit exekutiven Funktionen, Arbeitsgedächtnis und dem Sprachzentrum aufweisen und repräsentiere die fluide Intelligenz G_f im CHC-Modell. Bei der Profilverinterpretation gilt er als Kennwert für nicht sprachliche Intelligenz und wird dem Untertest *Wortschatz* gegenüber gestellt.

(5) Zur Überprüfung der *Simultanen Verarbeitung* werden Aufgaben vorgegeben, die dem Untertest „Räumliches Gedächtnis“ der K-ABC entsprechen. Das Kind soll sich die räumliche Anordnung von Objekten merken, die für fünf Sekunden auf dem Bildschirm sichtbar sind, und deren Lokalisation anschließend auf der nunmehr mit diagonalen Gitterlinien unterteilten Bildschirmfläche anklicken. Der Untertest

misst die Wahrnehmung, Speicherung und Manipulation visueller Informationen (CHC-Faktor G_v).

Vier der fünf Untertests können zu zwei Indices zusammengefasst werden (s. Tab. 1):

(1) In den *Allgemeinen Kognitiven Index* (AKI) gehen alle Untertests mit Ausnahme von *Aufmerksamkeit* ein. AKI wird als Kennwert für die Allgemeine Intelligenz (g) angesehen.

(2) Außerdem wird ein *Index der sequentiellen und simultanen Verarbeitung* (ISSV) gebildet, der aus den Untertests 3 und 5 besteht. Er gilt als neuropsychologisches Maß für die Kodierung und Speicherung von Informationen.

Der Testaufbau soll insbesondere Vergleiche zwischen verschiedenen theoretisch fundierten Konzepten ermöglichen, nämlich zwischen Aufmerksamkeit und kognitiven Leistungen (*Aufmerksamkeit* vs. AKI), fluider und kristalliner Intelligenz bzw. nonverbaler und verbaler Intelligenz (*Begriffsbildung* vs. *Wortschatz*), einzelheitlicher und ganzheitlicher Verarbeitung (*Sequentielle Verarbeitung* vs. *Simultane Verarbeitung*). Alle Skalen und Untertests sind normalverteilt mit einem Mittelwert von 100 und einer Streuung von 15. Als durchschnittlich werden Werte von 85 bis 115 bewertet.

Bei allen Untertests gibt es Einführungsaufgaben, die sicherstellen sollen, dass das Kind die Aufgaben verstanden hat und den PC sachgerecht bedienen kann. Dabei wird vom Programm automatisch erfasst, wie viele Fehler das Kind bei den Übungsaufgaben macht, was in den abschließenden Untersuchungsbericht mit einfließt. Die Durchführung der Untertests 2 bis 5 erfolgt adaptiv. Das Lösen bzw. Nichtlösen von Aufgaben entscheidet nach bestimmten Regeln darüber, ob anschließend einfachere oder schwierigere Aufgaben präsentiert werden. Dadurch werden Ober- und Untergrenzen des individuellen Leistungsvermögens ermittelt. Dieses Vorgehen ermöglicht im günstigen Fall eine ökonomische Testdurchführung durch Minimierung der zu bearbeitenden Aufgaben. Außerdem soll dadurch die Motivation der getesteten Person aufrechterhalten werden.

Nach der Testdurchführung, die bei Administration aller Aufgaben ca. eine Stunde dauert, gibt der Untersucher einige Angaben zum Testverhalten des Kindes in einen elektronischen Fragebogen ein. Die Testergebnisse werden dann vom Computerprogramm ausgerechnet. Aus allen verfügbaren Informationen wird automatisch ein ausführlicher Bericht generiert, der die Ergebnisse darstellt und zugleich psychologische und pädagogische Empfehlungen zum weiteren Vorgehen enthält. Für die Skalen- und Untertestdifferenzen werden signifikante Unterschiede angegeben, allerdings ohne Angaben zur Häufigkeit solcher Differenzen in der Normstichprobe. Für die Vergleiche zwischen SQV und SIV sowie zwischen BEG und WOS sind die Differenzen nach Altersgruppen differenziert, ansonsten werden pauschal Unterschiede von 1½ Standardabweichungen als signifikant angenommen. Es können nur bestimmte Untertestkombinationen ausgewählt werden. Bei auffälligen Befunden wird empfohlen, die jeweiligen Merkmale mit bestimmten Aufgaben, insbesondere aus der K-ABC und dem HAWIK-IV eingehender zu überprüfen.

Normierung

Für die Normierung wurden 899 Grundschüler an verschiedenen Orten Deutschlands rekrutiert. Auf Repräsentativität wurde hinsichtlich Geschlecht und regionaler Verteilung geachtet. Angaben zu Sozialstatus oder ethnischer Zugehörigkeit blieben unberücksichtigt. Kinder mit sprachlichen, motorischen oder sensorischen Defiziten wurden ausgeschlossen. Es wurden neun Altersgruppen gebildet, die jeweils sechs Lebensmonate umfassen. Mit Ausnahme der jüngsten und der beiden ältesten Gruppen bestehen die Normgruppen aus jeweils mehr als 100 Kindern. Angaben zur Auswahl und Qualifikation der Testleiter und der Rahmenbedingungen, unter denen die Testung stattfand, fehlen.

Die Normen weisen nur wenige Boden- und Deckeneffekte auf, lediglich der Untertest *Wortschatz* hat einen ausgeprägten Deckeneffekt bei Kindern ab 7;6 Jahren (Maximalwert weniger als zwei Standardabweichungen über dem Mittelwert; s. Tab. 2). Die anderen aufgeführten leichten Boden und Deckeneffekte beziehen sich auf den Untertest *Begriffsbildung* und schränken die Brauchbarkeit des Verfahrens lediglich bei der Testung von Kindern in den Extrembereichen ein. Die Normen sind im Manual abgedruckt und weisen für die Untertests eine gute Stetigkeit aus. Der Leistungsanstieg zwischen benachbarten Altersgruppen beträgt bei den Untertests 2, 3, 4 und 5 nahezu über den gesamten Normbereich hinweg konstant 3-4 Standardwertpunkte. Wie diese erstaunlich gleichmäßigen Normwertanstiege zustande kommen, wird nicht berichtet. Es fehlen auch Angaben, wie die Normalverteilung der Untertestnormen überprüft wurde und ob die adaptive Testdurchführung Auswirkungen auf die Bildung der Testnormen hatte.

Tabelle 2: Boden- und Deckeneffekte der Untertests der K-CAB

	6;6-6;11	7;0-7;5	7;6-7;11	8;0-8;5	8;6-8;11	9;0-9;5	9;6-9;11	10;0-10;5	10;6-10;11
AUFM									
WOS	D	D	D	D	D	D	D	D	D
SQV									
BEG	B	B	B			D	D	D	D
SIV									

Anmerkungen: leichter Bodeneffekt B: Rohwert (RW) 1 entspricht einem Standardwert (STW) von mehr als 55 (-3 SD); schwerer Bodeneffekt B: RW entspricht einem STW von mehr als 70 (-2SD); leichter Deckeneffekt D: maximale RW-Zahl entspricht einem STW von weniger als 145 (3 SD); schwerer Deckeneffekt D: maximale RW-Zahl entspricht einem STW von weniger als 130 (2 SD).

Die Altersnormgruppen sind bei den meisten Untertests in Halbjahresschritte unterteilt. Die Inspektion der Normtabellen zeigt aber, dass es bei *Aufmerksamkeit* nur zwei Altersgruppen 6;0 bis 8;11 sowie 9;0 bis 10;11 gibt, was zu einem deutlichen Leistungssprung zwischen den beiden Gruppen zu führen scheint. Allerdings wird nicht erklärt, wie aus den verschiedenen Messwerten in diesem Untertest jener Wert gebildet wird, aus dem sich der Standardwert ergibt.

Objektivität

Es fehlen genauere Angaben zu den äußeren Testbedingungen (z. B. Bildschirmgröße, Bildqualität, Raumhelligkeit, Lautstärke, Mauseinstellungen, Größe der Leertaste etc.). Die computergestützte Darbietung verspricht bei reibungslosem Ablauf eine objektive Durchführung und Auswertung. Regeln für den Umgang mit auftretenden Schwierigkeiten gibt es aber nicht. In die Befunde gehen auch die standardisierten Verhaltensbeschreibungen ein, die der Untersucher nach der Testung anhand grober Einschätzungskategorien vorzunehmen hat. Hierbei kommt es zur Konfundierung von Ausmaß des Problemverhaltens und Gültigkeit der Testergebnisse (z. B. „Er hatte große Schwierigkeiten, die Aufmerksamkeit über den Zeitraum der Testung aufrecht zu erhalten. Dies hat vermutlich Auswirkungen auf die Zuverlässigkeit der Ergebnisse.“). Außerdem gibt es keine Möglichkeit, Veränderungen des Testverhaltens über den Testverlauf hin zu protokollieren, z. B. nachlassende Aufmerksamkeit. Da es keine Angaben zur Beurteilerübereinstimmung für den auszufüllenden Fragebogen gibt, ist anzunehmen, dass diese nicht überprüft wurde.

Reliabilität

Die interne Konsistenz der Untertests 2 bis 5 wurde getrennt für die verschiedenen Altersgruppen berechnet. Die Angaben liegen zwischen .77 und .94. Im Mittel liegt die z-transformierte Reliabilität bei .80 (*Sequentielle Verarbeitung*) bis .92 (*Wortschatz*) und kann somit als überwiegend zufriedenstellend bis gut bezeichnet werden. Schwächen in einzelnen Altersgruppen zeigt lediglich der Untertest *Sequentielle Verarbeitung* (s. Tab. 1). Die Reliabilität der Intelligenzskala AKI beträgt .92 (für die einzelnen Altersgruppen .91 bis .95) und erreicht somit gute bis sehr gute Werte. Die Skalenreliabilität des ISSV fällt mit .84 (.79 bis .92) geringer aus und ist nicht in allen Altersgruppen zufriedenstellend.

Für den Untertest *Aufmerksamkeit* wird von einer Retest-Studie mit 45 Kindern unbekanntes Alters berichtet, die mit der französischen Originalversion durchgeführt wurde. Das Zeitintervall zwischen den beiden Messpunkten lag bei zwei bis vier Wochen. Es wurde eine Retest-Reliabilität von .76 ermittelt, die somit niedriger liegt als die Konsistenzmaße der übrigen Untertests. Im Manual wird hierzu zusätzlich ein „korrigierter“ Wert von .78 angegeben, es bleibt aber unklar, um welche Art von Korrektur es sich hier gehandelt hat.

Validität

Die Testautoren sind der Überzeugung, dass die *Inhaltsvalidität* der Testaufgaben keiner eigenen Überprüfung bedürfe. Dieser Automatismus ist jedoch zu bezweifeln. *Begriffsbildung* ist möglicherweise konfundiert mit der Kapazität des Kurzzeitgedächtnisses und ungeschicktes Hantieren mit der PC-Maus erschwert vor allem

bei den Aufgaben des *ISSV* die Antworteingabe am PC, sodass konstruktirrelevante Einflüsse nicht von vornherein ausgeschlossen werden können. Da in der K-CAB jeder Untertest für sich allein relativ komplexe kognitive Leistungsbereiche zu repräsentieren hat, sollte unbedingt geprüft werden, inwieweit die gewählten Aufgaben dieser Funktion tatsächlich auch gerecht werden.

Als Nachweis der *Konstruktvalidität* werden die Interkorrelationen der Untertests und Subskalen referiert (s. Tab. 3). Für die Untertests des AKI ergeben sich in der Gesamtstichprobe durchweg niedrige Korrelationen ($r = .19$ bis $.40$; Durchschnitt $r = .30$), was für eine heterogene Aufgabenzusammenstellung spricht. Die Korrelationen der vier kognitiven Untertests mit der Gesamtskala AKI fallen mit $.66$ bis $.75$ dagegen deutlich höher aus. Bezogen auf die Gesamtstichprobe sind die meisten Interkorrelationen konform mit den theoretischen Annahmen. So korrelieren SIV und SQV höher miteinander als mit allen übrigen Untertests. Allerdings variieren in diesem Fall die Korrelationskoeffizienten stark zwischen den einzelnen Altersgruppen (von $-.04$ bis $.55$) und nicht selten korreliert AUFM mit den Untertests des AKI höher als diese untereinander. Der Aussagewert der Befunde hinsichtlich der Konstruktvalidität ist jedoch mangels geeigneter Faktorenanalysen stark eingeschränkt.

Tabelle 3: Interkorrelationen der Untertests und Indices der K-CAB für die Gesamtstichprobe ^a

	BEG	WOS	SIV	SQV	AUFM	AKI
WOS	.35					
SIV	.40	.19				
SQV	.32	.31	.24			
AUFM	.31	.30	.32	.30		
AKI	.75	.66	.67	.67	.44	
ISSV	.45	.31	.80	.77	.40	.85

Anmerkungen: ^a Abkürzungen der Untertest- und Skalenbezeichnungen siehe Tabelle 1

Weitere Hinweise auf die Konstruktvalidität stammen aus Untersuchungen mit dem französischen K-CLASSIC. Für 40 Kinder wurden Korrelationen zwischen K-Classic und K-ABC und für 41 Kinder zwischen K-Classic und WISC-IV (HAWIK-IV) ermittelt. Angaben über Zusammensetzung, Alter und Werteverteilung der beiden Stichproben fehlen. Es finden sich in diesen Studien Hinweise sowohl auf konkurrente als auch diskriminante Validität in meist konstruktcompatibler Weise. Die Aussagen stützen sich aber vornehmlich auf korrigierte Korrelationskoeffizienten, die von den unkorrigierten Werten z. T. erheblich abweichen, wobei nicht mitgeteilt wird, welche Korrektur hier vorgenommen wurde. Somit ist der Aussagegehalt der Ergebnisse unklar. Erwartungswidrig korreliert aber *Wortschatz* höher mit der Skala Intellektueller Fähigkeiten als mit der Fertigkeitenskala der K-ABC und *Aufmerksamkeit* bildet höhere Zusammenhänge mit sprachlichem Verständnis als mit allen anderen Skalen der WISC-IV. Außerdem korreliert *Begriffsbildung* höher mit SV als mit WLD im WISC-IV, sodass die Annahme

der Sprachfreiheit dieses Untertests nicht haltbar zu sein scheint und der Vergleich zwischen *Begriffsbildung* und Wortschatz keine Aussage über sprachliche vs. nichtsprachliche Intelligenz, wie in der K-CAB postuliert, zulässt. Eine Replikation der Ergebnisse an deutschsprachigen Kindern wäre ebenso erforderlich wie nachvollziehbare Angaben über die im Manual referierten statistischen Prozeduren.

Zur *klinischen Validität* der K-CAB weist das Manual auf eine Studie mit 28 ADHS-Kindern und einer ebenso großen Kontrollgruppe hin, die der deutschen Normierungsstichprobe entnommen wurde. ADHS-Kinder schneiden im Untertest *Aufmerksamkeit* signifikant schlechter ab als unauffällige Kinder, in allen anderen Untertests erzielen sie nur geringfügig schlechtere Ergebnisse. Standardwerte unter 80 im Untertest AUFM werden laut Auswertungsprogramm als auffällig gewertet. Da die ADHS-Kinder in dieser Untersuchung im Mittel einen Standardwert von 87 erzielten, spricht dies jedoch nicht für eine Eignung dieses Kennwerts für die Feststellung von Aufmerksamkeitsstörungen. Zur klinischen Validität der kognitiven Indices AKI und ISSV scheinen noch keine Befunde vorzuliegen.

Bewertung des Verfahrens

Theoretisches Konzept: Die K-CAB basiert auf neuropsychologischen Modellvorstellungen und dient der Aufmerksamkeit- und Intelligenzdiagnostik. Die kognitiven Untertests bilden nur eine Minderheit der zehn CHC-Intelligenzfaktoren ab und es ist auch fraglich, ob damit die PASS-Theorie ausreichend repräsentiert werden kann, so wird z. B. sequentielle Informationsverarbeitung ausschließlich auditiv und nicht visuell geprüft. Die Untertests ähneln weitgehend Aufgaben, die auch in anderen Testbatterien Verwendung finden. *Aufmerksamkeit* erfasst selektive Aufmerksamkeit und Reaktionshemmung unter verschiedenen Bedingungen, lässt aber keine differenzierte Betrachtung der verschiedenen Aufmerksamkeitskomponenten zu. Wie die Testwerte zustande kommen, ist anhand des Manuals nicht nachvollziehbar, was eine Bewertung der Ergebnisse unmöglich macht. Man fragt sich z. B., ob eine geringe Aufmerksamkeit eher durch kürzere oder durch längere Reaktionszeiten gekennzeichnet ist. Im Untertest *Wortschatz* besteht eine hohe Wahrscheinlichkeit (36 %), auch bei zufälliger Antwortwahl einen oder zwei Punkte pro Aufgabe zu bekommen. Die Bezeichnung *Begriffsbildung* passt nach Ansicht des Rezensenten nicht zur Aufgabenstellung dieses Untertests, geht es hier doch vielmehr um das Finden von Klassifizierungsmerkmalen. Die Art der Darbietung lässt eine erhebliche Konfundierung mit Merkfähigkeit und bestimmten Aufmerksamkeitskomponenten vermuten. Der Profilvergleich wird laut Manual eine große Bedeutung bei der Auswertung zugemessen. Die Signifikanzangaben kritischer Differenzen ermangeln aber der klinisch-empirischen Bestätigung, ein Vergleich sprachgebundener mit nicht sprachgebundenen Leistungen ist aufgrund von Konfundierung unbrauchbar.

Im Großen und Ganzen ist das Manual hinsichtlich Theorie und Testaufbau informativ und enthält einen Abdruck der Normtabellen, so dass sich der Testanwender

ein eigenes Bild von Boden- und Deckeneffekten sowie von Itemgradienten machen kann. An einigen Stellen weist es aber sinnentstellende Fehler auf.

Testgüte: Der Test verfügt über eine deutsche Normierung an einer hinreichend großen Stichprobe, deren Repräsentativität jedoch nicht umfassend genug nachgewiesen ist. Die Normwerte der Untertests und Skalen weisen bemerkenswerte Stetigkeit auf und haben nur wenige Boden- und Deckeneffekte, allerdings fehlen Angaben zu Verteilungseigenschaften und zur Methodik der Normbildung. Im AKI sorgen Halbjahresschritte für eine ausreichende Altersabstufung, die Abstufungen bei *Aufmerksamkeit* sind nicht zu beurteilen. Die Reliabilitäten sind in allen Altersgruppen bis auf die Subskala ISSV zufriedenstellend bis gut. Die Validität der deutschen Version sollte gründlicher überprüft werden. Durchführungs- und Auswertungsobjektivität für die Testung selbst sind gegeben, die Verbalbeschreibungen der numerischen Testergebnisse sind eindeutig. Es fehlen allerdings Hinweise zur Gestaltung der Rahmenbedingungen der Untersuchung und zum Umgang mit Störsituationen (z. B. wenn das Kind beliebige Antwortreaktionen gibt, die Mitarbeit verweigert oder den PC unsachgemäß bedient). Nicht belegt ist die Objektivität des Fragebogens zur Verhaltensbeschreibung, dem bei der Formulierung der Empfehlungen oft die gleiche Gewichtung gegeben wird wie den Testergebnissen selbst.

Handhabung: Das Programm ist für die Benutzung unter den gängigen Windows-Betriebssystemen vorgesehen. Herstellerangaben zu den System-Mindestvoraussetzungen findet man an versteckter Stelle jedoch erst, nachdem man das Programm installiert hat (1 Ghz-Prozessor, 256 MB RAM Arbeitsspeicher, Bildschirmauflösung 800x600, 110 MB verfügbarer Speicherplatz). Rechner, die auf 64bit Basis arbeiten, erkennen den Schutzstecker nicht, ein passender Treiber kann bei Pearson oder SafeNet, dem Hersteller des Schutzsteckers, kostenlos bezogen werden. Die Kosten pro Testdurchführung in Höhe von 2,- € sind akzeptabel, allerdings unterscheidet die Zählung nicht zwischen der Durchführung des gesamten Verfahrens und der Nutzung von Teilbereichen. Die Nachbestellung weiterer Testdurchführungen ist aufwändiger als bei konventionellen Verbrauchsmaterialien. Das Programm lief in der Erprobung weitgehend störungsfrei. Zu Programmabstürzen kam es lediglich im Untertest AUFM bei sehr hohen Fehlerzahlen.

Durchführung: Die Testdurchführung erfolgt programmgesteuert, wobei sensorische oder motorische Schwierigkeiten sich stärker als bei herkömmlichen Tests negativ auswirken können. Die Abbildungen des Untertests WOS sind teilweise nicht gut zu erkennen und die Handhabung der Maus bei SIV und SQV erfordert mehr feinmotorisches Geschick als das Zeigen auf die Abbildungen bei „Räumliches Gedächtnis“ in der K-ABC. Der gesamte Test dauert ca. eine Stunde, vorausgesetzt, die Einführungsaufgaben bereiten keine erheblichen Schwierigkeiten und die Aufgaben werden nicht besonders langsam bearbeitet. Laut Manual sind die Untertests spielerisch gestaltet, ein Urteil, das nach eigener Erfahrung von Versuchspersonen eher nicht geteilt wird. Die K-CAB wurde im Rahmen der Erprobung von den Kindern überwiegend als anstrengender und eintöniger bewertet als andere Leistungsdiagnostiken.

Die Kinder haben bei der Bearbeitung des Tests entweder die Leertaste oder die PC-Maus zu bedienen. Ablenkung durch die übrige PC-Ausstattung, die das Kind nicht betätigen darf, und unsachgemäße Handhabung sind dabei nicht ausgeschlossen. Für Laptop-Benutzer empfiehlt es sich, das Touchpad zu deaktivieren oder eine externe Tastatur zu benutzen.

Insbesondere impulsiven Kindern fällt es schwer, beim Untertest *Aufmerksamkeit* in beiden Einführungsphasen, die jeweils 27 Reize umfassen, fehlerfrei zu arbeiten, sodass sie stereotyp immer wieder dieselben Belehrungen durch die Computerstimme erhalten, was erhebliche Frustration bis hin zur Verweigerung auslösen kann. Ob es zulässig ist, in dieser Phase den Ton abzuschalten, wie von manchen Kindern gefordert, wird nicht mitgeteilt. Nach eigener Erfahrung ist der Aufmerksamkeitsanteil der K-CAB für Kinder mit erheblichen Aufmerksamkeitsproblemen nicht zu empfehlen.

Die Einführung in den Untertest *Begriffsbildung* ist umständlich und kann nicht auf den individuellen Erklärungsbedarf des Kindes abgestimmt werden. Bei *Wortschatz* wird eine Eingabe verlangt, auch wenn das Kind die Antwort nicht weiß. Wählt es dann falsch, teilt ihm die Computerstimme jedes Mal mit, dass dies nicht die richtige Antwort sei, was ebenfalls Aversionen auslösen kann. Während der kognitiven Untertests kann die Durchführung unterbrochen oder vorzeitig beendet werden, bei Abbruch wird aber kein Bericht und folglich auch kein Ergebnis ausgegeben.

Bei schwankendem Leistungsvermögen erweist sich die interaktive Testdurchführung als zeitaufwändig und trotz oder gerade wegen des beständigen Wechsels von Aufgabenschwierigkeiten auch nicht als motivationsfördernd. Dieses Phänomen ist auch von anderen Tests bekannt und wird mit der niedrigeren subjektiven Einschätzung der Erfolgswahrscheinlichkeit in Verbindung gebracht (vgl. Frey, Hartig, Moosbrugger, 2009). Die Durchführung der K-CAB ist wesentlich unflexibler als andere Testverfahren, bei denen der Testleiter die Aufgaben selbst stellt, was insbesondere bei jungen, kognitiv beeinträchtigten oder schwierig zu untersuchenden Kindern ein gravierendes Problem darstellen kann. Durch die weitgehende Reduktion der sozialen Interaktion zwischen Untersucher und Kind entfallen auch Beobachtungsmöglichkeiten, die für erfahrene Untersucher in der Einzeltestung eine wichtige Informationsquelle hinsichtlich Motivation, Arbeitsweise, Selbstsicherheit und Frustrationstoleranz darstellen. Der Testleiter nimmt stattdessen eine passive Beobachterrolle ein, was von einigen Probanden als unangenehm erlebt wird. Manche Untersucher werden sich der Versuchung ausgesetzt sehen, währenddessen anderen Tätigkeiten nachzugehen bzw. die Durchführung zu delegieren.

Auswertung und Testbericht: Der Fragebogen, der vom Untersucher nach der Testdurchführung auszufüllen ist, weist Mängel auf. Bei der Beurteilung des Testverhaltens sind die Wahlmöglichkeiten stark eingeschränkt und es kommt zur Konfundierung der Verhaltensbeschreibung mit der Gültigkeit der Befunde. Verhaltensveränderungen während der Testung, z. B. nachlassende Konzentration, können nicht dokumentiert werden.

Der Testbericht ist ausführlich und enthält die erzielten Ergebnisse. Wünschenswert wäre noch eine Tabelle mit den zur Interpretation herangezogenen Testwertdifferenzen.

Bei der Beurteilung der Aufmerksamkeit wird das jeweilige Testniveau nicht berücksichtigt. So wird auch bei einem durchschnittlich aufmerksamen Kind immer dann eine Aufmerksamkeitsstörung nahe gelegt, wenn es im AKI signifikant besser als in der *Aufmerksamkeit* abgeschnitten hat. Das hier verwendete Diskrepanzkriterium entspricht nicht den diagnostischen Leitlinien für Aufmerksamkeitsstörungen. Unauffällige Befunde werden lediglich als Ergebnisse aufgeführt, es findet sich aber keine Formulierung, die darauf hinweist, dass damit Störungen ausgeschlossen werden konnten oder dass das Ergebnis keinen Hinweis auf das Vorliegen einer möglichen Störung geliefert hat, was eine wesentliche Aufgabe einer psychologischen Untersuchung mit der K-CAB wäre. Die Auswertung der AKI-Untertests berücksichtigt auch nicht, wenn der Test für ein Kind eindeutig zu schwierig war, z. B. wenn keine richtige Antwort abgegeben wurde.

Die automatisch erstellten Vorschläge zu psychologischen und pädagogischen Maßnahmen sind in der Mehrzahl der geprüften Fälle fachlich unzutreffend. So wird z. B. nach der Untersuchung mit der K-CAB eine weitere Abklärung der Aufmerksamkeit empfohlen, selbst wenn das Kind den entsprechenden Untertest fehlerfrei bewältigt hat und auch die Verhaltensbeobachtung keine entsprechenden Anhaltspunkte ergab. Grund allein für diese Empfehlung ist dabei die Angabe im Fragebogen, das Kind werde hinsichtlich seiner Aufmerksamkeit untersucht. Bei signifikantem Unterschied zwischen SQV und SIV werden je nach Anmeldegrund Empfehlungen zu Lese-, Rechtschreib- oder Rechenförderung gegeben, ohne die Durchführung entsprechender Schulleistungstests zu empfehlen. Ein Wert kleiner 80 im Untertest *Wortschatz* führt zur Empfehlung, das Kind vom Logopäden auf eine Sprachstörung untersuchen zu lassen, anstatt zunächst auf die Notwendigkeit einer genaueren Abklärung des Verbal-IQ hinzuweisen. Auch die pädagogischen Empfehlungen werden in vielen Fällen der Komplexität der Störungen nicht gerecht. So beschränken sich beispielsweise die Tipps bei Aufmerksamkeitsproblemen auf das Einhalten kurzer Arbeitsphasen und Erholungspausen, als ob es keine anderen pädagogischen Vorgehensweisen gäbe. Die Empfehlungen bei Verdacht auf Rechenstörungen beschränken sich auf verschiedene Möglichkeiten der Zahlenrepräsentation, ohne dass überhaupt bekannt ist, ob das Kind Probleme mit dem arabischen Zahlensystem hat.

Solche fachlich unhaltbaren Empfehlungen können in den Händen unkritischer Testanwender Schaden anrichten und Kosten durch unnötige weitere Untersuchungen verursachen. Da es nicht möglich ist, im Bericht abwegige Schlussfolgerungen zu unterdrücken, ist dringend zu empfehlen, die durch das Programm erstellten Befundberichte nicht in Umlauf zu setzen. Auch eine Einschränkung des potentiellen Benutzerkreises auf ausreichend qualifizierte Anwender, die in der Lage sind, die Plausibilität der Empfehlungen zu überprüfen, wäre eine daraus abzuleitende Konsequenz. Da dies in Deutschland leider nicht durchzusetzen ist, sollten die Empfehlungsfunktion seitens des Verlags dringend aus dem Testerstellungsprogramm entfernt werden.

Anwendungsbereich: Die K-CAB wird vom Herausgeber als Screening-Instrument bezeichnet, das nicht in der Lage sei, Intelligenz und Aufmerksamkeit abschließend zu

beurteilen. Da Diagnostik in den meisten Fällen mit dem Ziel einer Diagnosestellung oder eines sicheren Ausschlusses von Störungen betrieben wird, ergibt sich aus der Anwendung der K-CAB zwangsläufig die Notwendigkeit einer weitergehenden Diagnostik, womit der Zeitvorteil der K-CAB gegenüber herkömmlichen Verfahren rasch mehr als ausgeglichen wird. Dies sollte vor allen Dingen beachtet werden, wenn die Dauer einer Testdurchführung als wesentliches Argument für den Einsatz dieses Verfahrens angesehen wird. Die Verwendung der K-CAB können Untersucher erwägen, die sich mit übersichtlichen Informationen begnügen wollen oder müssen. Allerdings wäre auch hier zu fragen, ob die zumeist vorläufigen Ergebnisse den höheren finanziellen, zeitlichen und apparativen Aufwand im Vergleich zu globalen Testmaßen wie CFT, CPM oder d2 rechtfertigen. Zu den herkömmlichen mehrdimensionalen Leistungstests stellt K-CAB keine gleichwertige Alternative dar. Eine Einsatzmöglichkeit könnte darin bestehen, Kinder zu testen, die in ihrem Leistungsverhalten durch die Interaktion mit dem Untersucher beeinträchtigt werden. *Wortschatz* bietet sich evtl. bei Kindern an, die trotz guten Sprachverständnisses geringe aktivsprachliche Kompetenzen aufweisen.

Es ist zu erwarten, dass auch in der kinderpsychologischen Diagnostik in Zukunft computergestützte Verfahren vermehrt eingesetzt werden. Dazu müssen sie aber in der Durchführung deutlich mehr Flexibilität ermöglichen als die K-CAB in ihrer gegenwärtigen Form. Außerdem benötigen sie eine bessere testtheoretische Absicherung. Solange automatisch generierte Behandlungsempfehlungen nicht in der Lage sind, die klinischen Erfordernisse angemessen abzubilden, sollte darauf verzichtet werden. Bis auf einige spezielle Anwendungsbereiche stellt die K-CAB in ihrer gegenwärtigen Form keine empfehlenswerte Alternative zu herkömmlichen Aufmerksamkeits- oder Intelligenztests dar.

Literatur

- Das, J. P., Naglieri, J. A., Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Ally & Bacon.
- Frey, A., Hartig, J., Moosbrugger, H. (2009). Effekte des adaptiven Testens auf die Motivation zur Testbearbeitung am Beispiel des Frankfurter Adaptiven Konzentrationsleistungstests. *Diagnostica*, 55, 20-28.
- Kaufman, A. S., Kaufman, N. L. (2007). *K-CLASSIC. Evaluation informatisée des capacités cognitives et attentionnelles*. Paris: ECPA.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan, P. L. Harrison (Hrsg.), *Contemporary intellectual assessment: Theories, tests, and issues* (2. Aufl., S. 136-181). New York: Guilford.

Dieter Irblich, Auel