



Stand 30.09.2019

---

Markus Scholz, Gerolf Renner, Manfred Mickley

**Zitationsempfehlung:**

Scholz, M., Renner, G. & Mickley, M. (2019). *Testinformation zur Kaufman Assessment Battery for Children – Second Edition (KABC-II)* (Dia-Inform Verfahrensinformationen 005-01). Ludwigsburg: Pädagogische Hochschule Ludwigsburg.

Manfred Mickley arbeitet als Diplom-Psychologe seit über 20 Jahren in der Sozialpädiatrie (SPZ Rostock; davor SPZ Berlin-Friedrichshain). Daneben ist er als Dozent und Supervisor in der Ausbildung von Kinder- und Jugendlichen-Psychotherapeut\*innen tätig und führt Diagnostik-Seminare im schulpsychologischen, sonderpädagogischen und kinder- und jugendpsychiatrischen Kontext durch.

**Dieser Text wird veröffentlicht unter der Creative-Commons Lizenz  
CC BY-NC-ND 4.0 DE**

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Diese Publikation darf dementsprechend nur unter Nennung der Urheber und ausschließlich zu nicht-kommerziellen Zwecken genutzt und weiterverbreitet werden, Modifikationen des Textes sind nicht zugelassen.

**Potenzieller Interessenkonflikt:**

Dia-Inform will unabhängige und nicht von kommerziellen Interessen beeinflusste Informationen zur Verfügung stellen. Wir legen deshalb Wert auf eine sorgfältige Dokumentation aller Tatsachen, die auf einen potenziellen Interessenkonflikt hinweisen könnten.

Manfred Mickley führt für die Pearson Deutschland GmbH Schulungs- und Supervisionsseminare zur KABC-II durch.

## Das Projekt Dia-Inform

---

### Projektleitung und Korrespondenz:

Prof. Dr. Gerolf Renner  
Pädagogische Hochschule Ludwigsburg  
Fakultät für Sonderpädagogik  
Reuteallee 46  
71634 Ludwigsburg  
renner@ph-ludwigsburg.de

Dr. Markus Scholz  
Pädagogische Hochschule Ludwigsburg  
Fakultät für Sonderpädagogik  
Reuteallee 46  
71634 Ludwigsburg  
markus.scholz@ph-ludwigsburg.de

### Projektinformation:

Ziel des Projektes Dia-Inform an der Pädagogischen Hochschule Ludwigsburg ist es, ein umfangreiches Angebot an Informationsmaterialien zu diagnostischen Verfahren bereitzustellen, die in der (sonder-)pädagogischen Diagnostik eingesetzt werden.

Informationen zu diagnostischen Verfahren sowie Informations-, Lehr- und Praxisvideos bilden ein abgestuftes System, das für unterschiedliche Informationsbedürfnisse von Lehrenden und Studierenden und verschiedene Einsatzzwecke (Veranstaltungen, Selbststudium, Vorbereitung auf diagnostische Gutachten) flexibel und passgenau verwendet werden kann und in dessen Weiterentwicklung Studierende aktiv mit eingebunden werden können.

Im Rahmen des Projektes soll ein umfassender, nicht von kommerziellen Interessen beeinflusster und anderweitig nicht verfügbarer Informationspool zu diagnostischen Verfahren entstehen.

### Dia-Inform Verfahrensinformationen:

Dia-Inform Verfahrensinformationen geben einen Überblick über theoretische Grundlagen, Aufbau und psychometrische Eigenschaften eines diagnostischen Instrumentes. Sie weisen eine einheitliche Struktur auf, die Leser/innen die Orientierung erleichtert. Die Verfahrensinformationen diskutieren Stärken und Schwächen der Verfahren und weisen auf sinnvolle Einsatzmöglichkeiten in der sonderpädagogischen und klinisch-psychologischen Diagnostik hin. Dabei werden insbesondere die spezifischen Anforderungen einer Diagnostik bei Kindern und Jugendlichen mit Behinderungen berücksichtigt.

Dia-Inform Verfahrensinformationen werden von mindestens zwei Autor\*innen gemeinsam verfasst, die unterschiedliche professionelle Perspektiven einbringen.

**Dia-Inform Verfahrensinformationen stehen Ihnen kostenfrei auf OPUS-PHLB, dem Hochschulschriftenserver der PH Ludwigsburg, unter folgendem Link zur Verfügung:**

<https://phbl-opus.phlb.de/solrsearch/index/search/searchtype/collection/id/16235>

## **KABC-II**

### **Kaufman Assessment Battery for Children – Second Edition**

#### **Autor\*innen:**

---

Alan S. Kaufman & Nadeen L. Kaufman

Bearbeiter der deutschen Fassung: Peter Melchers & Martin Melchers

#### **Erscheinungsjahr:**

---

2015

#### **Altersbereich**

---

3;0 – 18;11 Jahre

#### **Verlag**

---

Pearson, Frankfurt

#### **Preis des Testverfahrens und der Verbrauchsmaterialien laut Verlagsseite**

---

Gesamtsatz: €1.400,00 bzw. €1.508,00 inkl. Einzelplatzsoftware;  
PC-Auswertungsprogramm als Einzelplatzlizenz: €253,50; Netzwerkversion (3 Zugriffe gleichzeitig möglich): €464,00; Protokollbögen (25 Stück): €91,00 (Stand: 13.08.2019).  
Alle Angaben zuzüglich Mehrwertsteuer.

#### **Setting, Durchführungsdauer und Auswertungsmöglichkeiten**

---

Die KABC-II ist ein Einzeltest. Die durchschnittliche Testzeit ist abhängig von der gewählten Testzusammenstellung, von der Wahl der Ergänzungsuntertests und vom Alter der Testperson. Für die Kernuntertests wird im Handbuch eine Spanne zwischen 25 Minuten und 75 Minuten angegeben. Bei der Vorgabe aller Ergänzungsuntertests werden im Handbuch je nach Alter knapp 60 bis 100 Minuten veranschlagt. Eine Auswertung ist manuell oder mit Hilfe eines PC-Auswertungsprogramms möglich.

#### **Diagnostische Zielsetzung lt. Testmanual:**

---

Die KABC-II ist ein mehrdimensionaler Intelligenztest und dient der Erfassung kognitiver Fähigkeiten bei Kindern und Jugendlichen. Laut Handbuch wurde die KABC-II entwickelt, „um wichtige Beiträge zur psychologischen, klinischen, psychologisch-pädagogischen und neuropsychologischen Beurteilung auf allen Entwicklungsstufen zwischen Kindergartenaltern und später Adoleszenz zu erreichen“ (Handbuch, S. 25). Sie soll klinische Diagnosen und pädagogische Einschätzungen erleichtern. Als Einsatzbereiche werden Bildungsberatung, Behandlungsplanung und die „Zuordnung von Kindern und Jugendlichen zu bestimmten schulischen und außerschulischen Förder- und Unterbringungsangeboten“ (Handbuch, S. 39) genannt. Zudem soll die KABC-II eine kulturell faire Untersuchung und die Ermittlung individueller Stärken und Schwächen im Bereich kognitiver Leistungen ermöglichen.

#### **Theoretischer Hintergrund:**

---

Die Autoren der KABC-II sprechen von einer dualen theoretischen Grundlage. Das Verfahren orientiert sich nach ihren Aussagen sowohl an einem neuropsychologischen („Luria-Modell“) als auch an einem Intelligenztheoretischen (Cattell-Horn-Carroll-Theorie, CHC-Modell) Bezugsrahmen.

Die CHC-Theorie stellt im Sinne eines hierarchischen Intelligenzmodells einen Ordnungsrahmen für die Beschreibung kognitiver Fähigkeiten dar, bei dem die Allgemeinintelligenz (g-Faktor, Schicht III) in 16 „breite“ (Schicht II) und mehr als 80 „schmale“ Fähigkeitsbereiche (Schicht I) untergliedert wird. Die CHC-Theorie trifft also Aussagen über die Struktur kognitiver Leistungen. Nähere Informationen zur CHC-Theorie sind z. B. bei Schneider und McGrew (2018), Flanagan, Alfonso und Dixon (2013) und deutschsprachig bei Mickley & Renner (2010, 2019), Renner und Mickley (2015) sowie Süß und Beauducel (2011) zu finden. In der KABC-II wird der g-Faktor (Schicht III), der in den meisten Intelligenztests als Gesamt-Intelligenzquotient bezeichnet wird, im sogenannten *Fluid-Kristallin-Index* (FKI) abgebildet. Auf Schicht II werden fünf Faktoren berücksichtigt: *Fluide Fähigkeiten* (Gf), *Kristalline Fähigkeiten* (Gc), *Kurzzeitgedächtnis* (Gsm), *Visuelle Verarbeitung* (Gv) und *Langzeitspeicherung und -abruf* (Glr).

Das sogenannte Luria-Modell bezieht sich auf Theorien des einflussreichen sowjetischen Neuropsychologen Alexander R. Luria, der selbst einer standardisierten Erfassung neuropsychologischer Funktionen sehr skeptisch gegenüberstand (Luria & Majovski, 1977). Luria unterschied drei funktionale Systeme des Gehirns, die der Aufrechterhaltung von Wachheit (Block 1), der Kodierung und Speicherung von Informationen (Block 2) und der Planung und Organisation des Verhaltens (Block 3) dienen (Lurija, 1992). In der KABC-II soll die Skala *Lernen* mit allen drei Blöcken assoziiert sein, mit Schwerpunkt auf Block 1 und 2. *Sequentielle Verarbeitung* und *Simultane Verarbeitung* seien primär mit Block 2 assoziiert. Sequentielle Verarbeitung meint laut Handbuch die Verarbeitung von Informationen, bei denen die serielle Ordnung von Bedeutung ist, während simultane Verarbeitung „simultan-ganzheitlich integriert“ (Handbuch, S. 51) sei und häufig eine räumliche Komponente aufweise. *Planung* soll „Entscheidungs- und exekutive Prozesse auf hohem Niveau“ (Handbuch, S. 51) erfassen. Aus diesen vier Skalen kann als Gesamtwert der *Intellektuelle-Verarbeitungsindex* (IVI) gebildet werden. Die Orientierung am Luria-Modell stelle laut Handbuch im Gegensatz zum CHC-Modell vor allem die kognitive Verarbeitung in den Fokus. Allerdings gibt es auf Ebene der eingesetzten Untertests und der Skalenbildung nur einen einzigen Unterschied zwischen den beiden Modellen: *Kristalline Fähigkeiten* (Gc) werden nur im CHC-Modell erfasst und im Gesamtwert (FKI) berücksichtigt. Die 4 weiteren Skalen werden aus exakt den gleichen Untertests gebildet und jeweils mit Bezug auf die Terminologie beider Modelle bezeichnet: *Simultan/Gv*, *Lernen/Glr*, *Planung/Gf* und *Sequentiell/Gsm*. Eine besondere Bedeutung habe die Wahl des Modells, wenn die Einbeziehung sprachlicher (kristalliner) Fähigkeiten in den Gesamtwert die Validität der Untersuchung beeinträchtigen könnte. Dies wäre z. B. der Fall, wenn eine Intelligenzmessung bei Personen mit mangelnden Kenntnissen der deutschen Sprache erfolgen soll. Das Manual empfiehlt das Luria-Modell aber auch bei zweisprachig aufwachsenden Kindern und Jugendlichen, bei Testpersonen mit Migrationshintergrund, bei einem „schwachen kulturellen oder sozialen Hintergrund“ (Handbuch, S. 35), bei Testpersonen mit Sprachentwicklungsstörungen, autistischen Störungen oder Hörminderungen. Das CHC-Modell wird bei den meisten anderen Fragestellungen empfohlen, insbesondere bei Beeinträchtigungen des Lesens und Schreibens, allgemeinen Entwicklungsrückständen, emotionalen und Verhaltensstörungen und Aufmerksamkeitsdefizitstörungen.

## Informationen für Anwender\*innen früherer Auflagen:

---

Die wichtigsten Änderungen von der K-ABC zur KABC-II auf einen Blick:

- Erweiterung des Altersbereichs von 12;5 Jahren auf bis zu 18;11 Jahre
- Anhebung des Testmindestalters von 2;6 auf 3;0 Jahre
- 10 neue Subtests: *Atlantis*, *Symbole*, *Atlantis - Abruf nach Intervall*, *Symbole - Abruf nach Intervall*, *Konzeptbildung*, *Bausteine zählen*, *Rover*, *Muster ergänzen*, *Geschichten ergänzen* sowie *Wort- und Sachwissen*
- 8 entfallene Subtests: *Zauberfenster*, *Räumliches Gedächtnis*, *Bildhaftes Ergänzen*, *Fotoserie*, *Gesichter & Orte*, *Rechnen*, *Lesen/Buchstabieren* und *Lesen/Verstehen*
- Modifikationen und Ergänzungen von Untertests (z. B. einfachere Aufgaben bei *Dreiecke*, Ersetzung von Aufgaben bei *Wiedererkennen von Gesichtern*, Verwendung von Bildern bei den ersten Aufgaben des Untertests *Rätsel*, Ergänzung schwierigerer Aufgaben bei verschiedenen Untertests)
- Veränderte Teststruktur (s. Testaufbau)

## Testaufbau:

---

Die KABC-II umfasst insgesamt 18 Untertests, die in Kern- und Ergänzungsuntertests unterschieden werden. Die Anzahl der Kern- und Ergänzungsuntertests sowie die Anzahl und Zusammensetzung der Skalen und die Skalenbildung variieren erheblich in Abhängigkeit vom Alter des untersuchten Kindes. Bei 3-Jährigen wird nur ein Gesamtwert bestimmt. Eine vollständige Erfassung der 5 CHC-Faktoren ist erst ab dem Alter von 7;0 Jahren möglich.

Neben FKI und IVI kann als weiterer Gesamtwert noch ein sogenannter *Sprachfrei-Index* (SFI) bestimmt werden, der auf Untertests basiert, die von den Testpersonen keine aktivsprachlichen Leistungen verlangen. Außerdem kann aus den Untertests *Symbole - Abruf nach Intervall* und *Atlantis - Abruf nach Intervall* der Gesamtwert *Abruf nach Intervall* (ANI) gebildet werden. Tabelle 1 zeigt den Aufbau der KABC-II, Tabelle 2 die Aufgabenstellung und die inhaltlichen Schwerpunkte der einzelnen Untertests sowie die Skalenzuordnung.

Es ist zulässig, einen Kernuntertest nach vorgegebenen Regeln durch einen Ergänzungsuntertest zu ersetzen. Können einzelne Untertests nicht durchgeführt oder nicht ausgeertet werden, kann eine Schätzung von FKI, IVI, SFI und *Simultan/Gv* erfolgen. Bei FKI und IVI hält das Handbuch (S. 91) dies für zulässig, wenn bis zu zwei Untertests nicht verwertbar sind, wobei für jede Skala mindestens ein Untertest vorliegen soll. Die Werte der fehlenden Untertests sollen durch den Mittelwert der weiteren Untertests ersetzt werden. Allerdings findet sich im Handbuch (S. 91) auch der Hinweis, dass auf der Basis von Schätzungen errechnete Standardwerte keine Interpretation der damit gewonnenen Ergebnisse erlauben. Das PC-Auswertungsprogramm berücksichtigt Substitutions- und Hochrechnungsregeln automatisch, mit klaren Warnhinweisen im Ergebnisprotokoll: „Der Skalenindex darf nicht interpretiert werden“ bei Hochrechnungen und „Interpretieren Sie den Skalenindex vorsichtig“ bei Ersetzungen.

**Tabelle 1**

*Aufbau und Untertests der KABC-II in Abhängigkeit vom Alter*

Skala / Untertest (Untertest-Nummer)	3;0-3;11	4;0-4;11	5;0-5;11	6;0-6;11	7;0-12,11	13,0-18;11
<b>Intellektueller Verarbeitungsindex (IVI)</b>						
<b>Fluid-Kristallin-Index (FKI)</b>						
<b>Sprachfrei-Index (SFI)</b>						
<b>Sequentiell/Kurzzeitgedächtnis (Gsm)</b>						
<i>Zahlen nachsprechen (5)</i>	Ergänzung	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI
<i>Wortreihe (14)</i>	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI
<i>Handbewegungen (16)</i>	SFI	SFI, Ergänzung	SFI, Ergänzung	SFI, Ergänzung	SFI, Ergänzung	SFI, Ergänzung
<b>Simultan/Visuelle Verarbeitung (Gv)</b>						
<i>Bausteine zählen (13)</i>			Ergänzung	Ergänzung	SFI, Ergänzung	IVI, FKI, SFI
<i>Konzeptbildung (2)</i>	IVI, FKI, SFI	IVI, FKI, SFI	IVI, FKI, SFI	IVI, FKI, SFI		
<i>Wiedererkennen von Gesichtern (3)</i>	IVI, FKI, SFI	IVI, FKI, SFI	SFI, Ergänzung			
<i>Muster ergänzen<sup>a</sup> (15)</i>			IVI, FKI, SFI	IVI, FKI, SFI		
<i>Rover (7)</i>				IVI, FKI	IVI, FKI	IVI, FKI
<i>Geschichten ergänzen<sup>a</sup> (4)</i>				Ergänzung, SFI		
<i>Dreiecke (12)</i>	IVI, FKI, SFI	IVI, FKI, SFI	IVI, FKI, SFI	IVI, FKI, SFI	IVI, FKI, SFI	SFI, Ergänzung
<i>Gestaltsschließen (6)</i>	Ergänzung	Ergänzung	Ergänzung	Ergänzung	Ergänzung	Ergänzung
<b>Planung/Fluide Fähigkeiten (Gf)</b>						
<i>Muster ergänzen<sup>a</sup> (15)</i>					IVI, FKI, SFI	IVI, FKI, SFI
<i>Geschichten ergänzen<sup>a</sup> (4)</i>					IVI, FKI, SFI	IVI, FKI, SFI
<b>Lernen/Langzeitspeicher (Glr)</b>						
<i>Atlantis (1)</i>	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI
<i>Symbole (11)</i>		IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI	IVI, FKI
<i>Atlantis – Abruf nach Intervall (8)</i>			Ergänzung	Ergänzung	Ergänzung	Ergänzung
<i>Symbole – Abruf nach Intervall (17)</i>			Ergänzung	Ergänzung	Ergänzung	Ergänzung
<b>Wissen/kristalline Fähigkeiten (Gc)</b>						
<i>Wortschatz (9)</i>	FKI	FKI	FKI	FKI	Ergänzung	Ergänzung
<i>Rätsel (18)</i>	FKI	FKI	FKI	FKI	FKI	FKI
<i>Wort- und Sachwissen (10)</i>	Ergänzung	Ergänzung	Ergänzung	Ergänzung	FKI	FKI

Anmerkungen. <sup>a</sup>Diese Untertests werden abhängig vom Alter des Kindes unterschiedlichen Skalen zugeordnet.

Bereiche mit Hintergrundschattierung kennzeichnen Altersgruppen, in denen Normen zur Verfügung stehen.

**Tabelle 2***Aufgabenstellung und inhaltliche Schwerpunkte der Untertests*

<b>Untertest</b>	<b>Aufgabenstellung</b>	<b>Erfasste Konstrukte</b>
<i>Atlantis</i>	Es sollen sinnfreie Namen von Figuren (Meerestiere und -pflanzen) gelernt und im Verlauf wiederholt in Abbildungen mit mehreren anderen Figuren gezeigt werden.	Auditives Kurzzeitgedächtnis, visuelles Kurzzeitgedächtnis, assoziatives Gedächtnis
<i>Konzeptbildung</i>	Die Testperson soll aus einer Auswahl aus vier bzw. fünf konkreten oder abstrakten Bildern erkennen, welches nicht zu den anderen Bildern passt.	Logisch schlussfolgerndes Denken
<i>Wiedererkennen von Gesichtern</i>	Der Testperson wird eine Vorlage mit ein oder zwei Gesichtern gezeigt. Diese Gesichter sollen anschließend auf einem Foto mit mehreren Personen identifiziert werden.	Visuelles Kurzzeitgedächtnis
<i>Geschichten ergänzen</i>	Die Testperson soll Lücken in einer Bildergeschichte durch Einfügen passender Karten an der richtigen Stelle vervollständigen.	Allgemeinwissen, logisch-schlussfolgerndes Denken, sequentielles Denken
<i>Zahlen nachsprechen</i>	Eine mündlich dargebotene Zahlenfolge soll von der Testperson unmittelbar in der gleichen Reihenfolge nachgesprochen werden.	Gedächtnisspanne, auditives Kurzzeitgedächtnis
<i>Gestaltschließen</i>	Die Testperson soll ein Objekt auf Basis einer unvollständigen Zeichnung erkennen und benennen.	Gestaltschließen
<i>Rover</i>	Eine Hundefigur soll auf einem mit Hindernissen und Erschwernissen versehenen Spielplan auf dem kürzesten Weg zu einem Knochen geführt werden.	Visuell-räumliche Fähigkeiten, sequentielles Denken
<i>Atlantis - Abruf nach Intervall</i>	Die Testperson soll die im Untertest <i>Atlantis</i> erlernten Figuren in einer Abbildung mit mehreren Ablenkern zeigen.	Assoziatives Gedächtnis (Paarassoziationslernen)
<i>Wortschatz</i>	Die Testperson soll abgebildete Objekte benennen.	Aktiver Wortschatz, lexikalisches Wissen

Untertest	Aufgabenstellung	Erfasste Konstrukte
<i>Wort- und Sachwissen</i>	Aus einer Auswahl aus 6 Abbildungen soll die Testperson diejenige zeigen, die zu einem vom Testleiter genannten Begriff passt.	Allgemeinwissen, lexikalisches Wissen
<i>Symbole</i>	Die Testperson soll die Bezeichnung bzw. Bedeutung von größtenteils abstrakten piktografischen Zeichen erlernen und Wortverbindungen und Sätze, die mit diesen Zeichen dargestellt werden, erlesen	Assoziatives Gedächtnis
<i>Dreiecke</i>	Zweidimensionale Muster sollen von der Testperson mit Hilfe von Formen nachgebaut werden.	Visuell-räumliche Zusammenhänge, visuell-räumliches Vorstellungsvermögen
<i>Bausteine zählen</i>	Der Testperson werden Abbildungen von gestapelten Bausteinen präsentiert. Die Anzahl aller Bausteine, einschließlich der zum Teil verdeckten Bausteine, soll genannt werden.	Visuell-räumliches Vorstellungsvermögen
<i>Wortreihe</i>	Der Testperson werden mehrere Worte vorgesprochen. Zu diesen soll die Testperson unmittelbar im Anschluss passende Bilder in der richtigen Reihenfolge zeigen. Im Verlauf wird eine Interferenzaufgabe dazwischengeschaltet.	Gedächtnisspanne, visuelles Kurzzeitgedächtnis
<i>Muster ergänzen</i>	Eine nach einem bestimmten Prinzip aufgebaute Reihe visueller Muster soll mit dem richtigen Muster ergänzt werden.	Logisch-schlussfolgerndes Denken
<i>Handbewegungen</i>	Die Testperson soll eine Abfolge von vorgemachten Handbewegungen unmittelbar wiederholen.	Visuell-räumliches Gedächtnis, Gedächtnisspanne
<i>Symbole – Abruf nach Intervall)</i>	Es sollen Wortverbindungen und Sätze erlesen werden, die mit den zuvor im Subtest <i>Symbole</i> erlernten piktografischen Zeichen dargestellt werden.	Assoziatives Gedächtnis, Lernfähigkeit



Untertest	Aufgabenstellung	Erfasste Konstrukte
<i>Rätsel</i>	Die Testperson soll einen verbal beschriebenen Begriff erraten. Bei den ersten Items geschieht dies mit Hilfe von Abbildungen.	Allgemeinwissen, passiver Wortschatz

### Ergebnis(-werte)

#### Zentrale Ergebniswerte:

- Standardwerte für die Gesamtskalen ( $M = 100$ ,  $SD = 15$ ) mit 90%- und 95%-Konfidenzintervallen, bei Nutzung der Computerauswertung auch exakte Prozentränge
  - Luria-Modell: Intellektueller Verarbeitungsindex (IVI)
  - CHC-Modell: Fluid-Kristallin-Index (FKI)
  - Sprachfrei-Index (SFI)
- Standardwerte ( $M = 100$ ,  $SD = 15$ ) der Skalenindices mit 90%- und 95%-Konfidenzintervallen, bei Nutzung der Computerauswertung auch exakte Prozentränge (je nach Alter 3 bis 4 Skalenindices bei Verwendung des Luria-Modells, 4 bis 5 bei Verwendung des CHC-Modells). Bei 3-Jährigen werden nur IVI oder FKI berechnet.
- Skalenwerte ( $M = 10$ ,  $SD = 3$ ) für alle Untertests, bei Nutzung der Computerauswertung auch exakte Prozentränge. Bei *Geschichten ergänzen*, *Dreiecke* und *Muster ergänzen* werden ab dem Alter von 7;0 Jahren zusätzlich Bonuspunkte für schnelle Aufgabenlösungen vergeben, wobei eine alternative Auswertung ohne Berücksichtigung des Zeitbonus möglich ist.

#### Weitere Ergebniswerte:

- Altersäquivalente (Entwicklungsalter) für die Untertest-Rohwertsummen

#### Daten zur Profilinterpretation

- Spannweite der Standardwerte der Untertests (Differenz zwischen dem höchsten und niedrigsten Untertestwert innerhalb der Skalen des Luria- und CHC-Modells). Das Handbuch rät, bei Spannweiten, die in weniger als 10% der Normierungstichprobe auftraten, auf die Interpretation des Skalenwertes zu verzichten. Bei der PC-Auswertung werden entsprechende Ergebnisse als „nicht interpretierbar“ gekennzeichnet.
- Angabe von individuellen Stärken und individuellen Schwächen mit einem Signifikanzniveau von  $p < .05$ . Basis hierfür bilden die Abweichungen der einzelnen Skalen vom individuellen Mittelwert aller Skalen. Das Handbuch gibt zusätzlich Hinweise auf die Häufigkeit der jeweiligen intraindividuellen Stärke oder Schwäche in der Normstichprobe (für Häufigkeiten  $< 10\%$ ,  $< 5\%$  und  $< 1\%$  der Untersuchungsfälle).

### Objektivität:

*Durchführungsobjektivität:* Die ausführlichen allgemeinen und spezifischen Instruktionen garantieren weitgehend die Durchführungsobjektivität. Unter Umständen könnten sich in einzelnen Untertests (*Dreiecke*, *Muster ergänzen*, *Geschichten ergänzen*) bei der sekundengenauen Zeitmessung auswertungsrelevante Fehler ergeben. Die an sich

gesicherte Durchführungsobjektivität gilt jedoch nicht für den *Sprachfrei-Index*, da für eine sprachfreie Testung keine standardisierten Instruktionen vorliegen.

*Auswertungsobjektivität:* Bei fast allen Untertests ist die Bewertung der Antworten eindeutig möglich. Bei den Untertests *Gestaltschließen*, *Wortschatz* und *Rätsel* kann auch nach Einschätzung der Testautoren eine eigenständige Entscheidung der Testleiter\*innen über noch zulässige und nicht mehr zulässige Antworten erforderlich werden.

*Interpretationsobjektivität:* Die Interpretationsobjektivität ist durch den Bezug auf Normwerte gesichert.

### **Reliabilität:**

Die Reliabilitätskoeffizienten der KABC-II wurden mittels Split-half Methode berechnet.

Auf Ebene der Untertests liegen die über die Altersjahrgänge gemittelten Reliabilitätswerte bei den 3- bis 6-Jährigen bei .70 (*Handbewegungen*) oder höher. Bei den 7- bis 18-Jährigen bei .79 (*Geschichten ergänzen*, *Gestaltschließen*) oder höher. Betrachtet man die einzelnen Altersjahrgänge, ergibt sich die geringste Reliabilität mit .55 bei 6 Jahren im Untertest *Handbewegungen*. Alle anderen Werte liegen höher als .71 (*Wiedererkennen von Gesichtern* bei 4 Jahren).

Tabelle 3 zeigt die Reliabilitäten für die Gesamtskalen und Skalenindices aus dem Manual.

### **Tabelle 3**

*Reliabilitäten von Gesamt IQ und Indices*

Gesamtskala / Skalenindex	Reliabilität
<b>Fluid-Kristallin-Index (FKI)</b>	<b>.96-.98</b>
<b>Individueller Verarbeitungsindex (IVI)</b>	<b>.94-.98</b>
<b>Sprachfrei-Index (SFI)</b>	<b>.90-.96</b>
Sequentiell/Gsm	.85-.93
Simulant/Gv	.88-.97
Lernen/Glr	.96-.97
Planung Gf	.84-.92
Wissen/Gc	.92-.96
Abruf nach Intervall (ANI)	.90-.95

*Anmerkungen.* Angegeben ist die Spannweite der Reliabilitätskoeffizienten in den Altersjahrgängen.

Zur Retest-Reliabilität der deutschsprachigen KABC-II wurde keine eigene Studie durchgeführt. Bei der US-amerikanischen KABC-II fanden sich über einen Zeitraum von durchschnittlich 4 Wochen bei 205 Testpersonen für FKI und IVI hohe Stabilitäten zwischen .87 und .92. Deutliche Übungeffekte traten in dieser Untersuchung bei *Lernen/Glr* (z. B. 14.6 IQ-Punkte in der Altersgruppe 7-12), *Simultan/Gv* (9.2 IQ-Punkte bei den 7- bis 12-Jährigen) und *Planung/Gf* (10.8 IQ-Punkte bei den 13- bis 18-Jährigen auf (Kaufman & Kaufman, 2004). Eine Testwiederholung sei deswegen vor Ablauf eines Jahres nicht sinnvoll (Handbuch, S. 186). Daten zur längsschnittlichen Stabilität von Testprofilen liegen nicht vor.

**Validität:**

*Inhaltliche Validität:* Die KABC-II enthält größtenteils bewährte Aufgabenarten. Die Teststruktur ist explizit an aktuelle Intelligenzvorstellungen angelehnt (Cattell-Horn-Carroll Intelligenztheorie), was für die inhaltliche Validität des Tests spricht.

Problematisch erscheint die Skalenbezeichnung *Planung* im Luria-Modell. Planung bedeutet, mental einige Handlungsschritte zu antizipieren, um Fehler zu vermeiden (Drechsler, 2007). Dies wird bei den entsprechenden Untertests kaum gefordert.

Im Hinblick auf die inhaltliche Validität des Untertests *Wort- und Sachwissen* ist zu beachten, dass verschiedene Items und Abbildungen deutlich von der anglo-amerikanischen Kultur geprägt sind. Auch einzelne Abbildungen aus dem Ergänzungsuntertest *Gestaltschließen* sind für Kinder und Jugendliche aus anderen Kulturkreisen evtl. schwerer zu deuten.

*Lernen/Langzeitgedächtnis* enthält nur Aufgaben zum Paarassoziationslernen, deckt also das Konstrukt Langzeitgedächtnis nur begrenzt ab. Die Bezeichnung „Lernen“ im Luria-Modell kann bei Personen, die mit dem theoretischen Hintergrund nicht vertraut sind, zu der Einschätzung führen, dass hier Lernen (einschließlich des schulischen Lernens) als globales Konstrukt erfasst wird. Das ist definitiv nicht der Fall.

*Faktorielle Validität:* Zur Faktorenstruktur berichtet das Handbuch summarisch die Ergebnisse von konfirmatorischen Faktorenanalysen für 4 Altersgruppen (4 Jahre, 5-6 Jahre, 7-12 Jahre, 13-18 Jahre). Die Zusammenfassung der Altersgruppen 5 und 6 Jahre ist allerdings nicht gerechtfertigt, da der Untertest Rover erst ab 6 Jahren durchgeführt werden kann. Die Stichprobengröße ( $N = 102$ ) entspricht genau der Anzahl der 6-Jährigen in der Normstichprobe, so dass es sich hier um eine Falschangabe im Manual handeln könnte.

Überprüft wurde ein hierarchisches Modell mit Allgemeinintelligenz (g-Faktor) als Faktor höherer Ordnung und den fünf Faktoren erster Ordnung des CHC-Modells. Dabei wurden einmal alle Untertests, einmal nur die Kernuntertests berücksichtigt. Die Stichprobengrößen für die 4- bis 6-Jährigen sind für eine konfirmatorische Faktorenanalyse recht gering. Die Modellpassung kann nach gängigen Kriterien (z. B. Moosbrugger & Schermelleh-Engel, 2012) für die auf den Kernuntertests beruhenden Untertests als akzeptabel bis gut beurteilt werden, was nur bedingt für die Modelle gilt, die alle Untertests einbeziehen. In den Altersgruppen 7-12 und 13-18 zeigt *Planung/Gf* die höchsten Ladungen auf dem g-Faktor (.99), die niedrigsten Ladungen finden sich bei *Sequentiell/Gsm* (.69 bzw. .65). Bei den 4- und bei den 6-Jährigen lädt *Simultan/Gv* mit jeweils .93 am höchsten auf dem g-Faktor, die niedrigsten Ladungen finden sich mit .59 bei *Lernen/Glr* (4-Jährige) und *Sequentiell/Gsm* mit .73 (6-Jährige).

Bei eigenen Nachberechnungen auf Basis der im Manual angegebenen Korrelationsmatrizen ließen sich die Ergebnisse bei den 4-Jährigen recht gut replizieren. Der Fit des CHC-Modells ist bei den 5-Jährigen, nicht jedoch bei den 6-Jährigen, akzeptabel. Bei den Altersgruppen 7-12 und 13-18 stimmten die Ergebnisse der Nachberechnungen nicht mit den Angaben im Manual überein. Für das CHC Modell fanden sich unzulässige Lösungen (sog. Heywood-Fälle, bei denen die Modellschätzung negative Varianzen und standardisierte Faktorladungen  $> 1.0$  ergab). Auch im Manual der amerikanischen KABC-II findet sich zumindest eine unzulässige Lösung mit einer Faktorladung größer 1.0.

*Konvergente Validität:* Im Handbuch werden Korrelationen mit verschiedenen anderen Intelligenztests angegeben, die hier nur auszugsweise dargestellt werden können. Die

Korrelationen der Gesamtskalen der KABC-II und dem IQ-Wert der IDS (Grob, Meyer & Hagmann-von Arx, 2009) liegen mit .52 (IVI) und .54 (FKI) im mittleren Bereich, mit der Wechsler Intelligence Scale for Children – IV (WISC-IV; Petermann & Petermann, 2011), die einige ähnliche Aufgabenstellungen aufweist, mit .69 (IVI) und .79 (FKI), im hohen Bereich. Bei Kindern im Vorschulalter lagen die Korrelationen mit dem Gesamt-IQ der Wechsler Primary and Preschool Scale of Intelligence - III (WPPSI-III; Petermann & Lipsius, 2009) bei .59 (IVI) und .75 (FKI). Neben der Betrachtung der Korrelationen sind auch Niveauunterschiede zwischen verschiedenen Testverfahren bedeutsam. Diese fielen meist gering aus. Im Vergleich zum FKI lagen der Gesamt-IQ der WISC-IV 2.3 IQ-Punkte höher, der IQ der IDS 1.3 IQ-Punkte niedriger und der Gesamt-IQ der WPPSI-III 3.6 IQ-Punkte höher. Zum Zusammenhang mit Schulleistungen bzw. Schulleistungstests finden sich keine Angaben.

*Divergente Validität:* Mit der Kaseler Konzentrationsaufgabe (KKA; Krampen, 2007), einem Maß der selektiven Aufmerksamkeit, korrelieren die Skalenindices und Gesamtwerte der KABC-II fast durchgehend nur sehr niedrig.

*Diskriminative Validität:* Angaben zur diskriminativen Validität für klinische Diagnosegruppen liegen nicht vor. Die entsprechenden Daten der amerikanischen Version werden im Manual nicht dargestellt, da die Diagnosekriterien in den USA und Deutschland nicht vergleichbar seien.

*Prognostische Validität:* Zur prognostischen Validität liegen keine Daten vor.

Zu der im Handbuch empfohlenen Auswertung von Testprofilen werden keine Validitätsdaten berichtet.

### **Normierung:**

Die deutschsprachige Normierungsstichprobe umfasste 1745 Probanden (102 bis 123 in den Altersjahrgängen). Diese wurde von April 2013 bis Februar 2014 in Verantwortung eines Marktforschungsinstituts (IFM Mannheim) durchgeführt. Zur Orientierung bei der Stichprobenwahl wurden bevölkerungsstatistische Ergebnisse herangezogen. Zur Gewährleistung der Repräsentativität wurden Geschlecht, elterlicher Bildungsgrad, Anteil von Probanden mit Migrationshintergrund, besuchte Schulform und regionale Verteilung in der Stichprobe berücksichtigt. In Bezug auf den deutschen Sprachraum liegt eine leichte Überrepräsentation von Österreich und der Schweiz vor. Die Autoren sprechen von einer leichten Unterrepräsentation in Bezug auf Eltern mit fehlenden Schulabschlüssen oder Eltern mit Hauptschulabschluss, wobei sich die Bildungsabschlüsse in den Altersjahrgängen durchaus unterschiedlich verteilten. In der Normstichprobe enthalten waren auch 117 Kinder und Jugendliche aus stationär-klinischen Kontexten oder Förderschulen.

Nur die Korrelationstabellen im Manual lassen erkennen, dass die Stichprobengrößen der Normierung bei den Ergänzungsuntertests durchgehend kleiner sind als bei den Kernuntertests. Offensichtlich wurden nicht mit allen Testpersonen alle Untertests durchgeführt. Im Manual finden sich keine Hinweise, nach welchen Kriterien der Einsatz von Ergänzungsuntertests erfolgte.

Nach der Normierung wurden noch verschiedene Anpassungen vorgenommen (Eliminierung einzelner Items, Änderungen der Aufgabenreihung, Testeinstiege). Die in der Normierung eingesetzte Fassung entsprach daher nicht ganz der Endfassung.

Die Normen seien von Psychometrikern des Testverlags auf Basis einer nicht näher beschriebenen regressionsanalytischen Glättungsprozedur berechnet worden, ergänzend wurden „per Hand“ (Manual, S. 178) Glättungen vorgenommen.

Die Normgruppen umfassen bei den 3- bis 5-Jährigen jeweils drei Monate, bei den 6- bis 14-Jährigen vier Monate und bei den 15- bis 18-Jährigen sechs Monate.

Bodeneffekte gibt es in der KABC-II hauptsächlich im niedrigsten Altersbereich 3;0 bis 3;2 Jahre. Beim Kernuntertest *Konzeptbildung* ist mit einem Rohwertpunkt bereits ein Skalenwert von 5 und bei *Wiedererkennen von Gesichtern* einer von 6 erreicht. Ähnlich sieht es im Bereich von 5;0 bis 5;2 Jahren bei den Untertests *Muster ergänzen* und *Rätsel* aus. Auch hier führt ein Rohwertpunkt bereits zu 6 Skalenwertpunkten. Ab einem Alter von acht Jahren ist die KABC-II in den Kernuntertests frei von Bodeneffekten. Deckeneffekte tauchen in den meisten Kernuntertests erst in den höchsten Altersstufen auf. Schon vorher und am stärksten zeigt sich dies beim Untertest *Geschichten ergänzen* mit und ohne Zeitvorgabe. Ohne Zeitvorgabe treten schon bei 12-Jährigen größere Deckeneffekte auf. Differenzierungen im hohen Leistungsbereich bei älteren Probanden werden daher nur durch die Berücksichtigung der Bearbeitungsgeschwindigkeit erreicht.

Der tabellierte Wertebereich umfasst bei den Untertests  $\pm 3$  Standardabweichungen (1 bis 19 Wertpunkte), bei den Gesamtwerten und Skalen  $\pm 4$  Standardabweichungen (IQ 40 bis 160).

Problematische Itemgradienten finden sich selten. Im Extremfall führt ein zusätzlicher Rohwertpunkt zu einem Sprung von einer Standardabweichung bei den Normwerten (*Handbewegungen*, Altersgruppe 3;0 bis 3;2 Jahre), Sprünge um 0.66 Standardabweichungen kommen bei den jüngeren Altersgruppen mehrfach vor und sind bei *Zahlen nachsprechen* über den gesamten Altersbereich zu beobachten.

#### **Weitere Gütekriterien:**

---

*Ökonomie:* Von der KABC-II wird ein breites Spektrum kognitiver Leistungen abgedeckt. Wenn das Ziel der Intelligenzdiagnostik eine umfassende und differenzierte Erfassung kognitiver Leistungen ist, kann die Durchführungsdauer daher als akzeptabel bewertet werden. Da keine Kurzform vorliegt, ist der Einsatz weniger ökonomisch, wenn nur eine grobe Abschätzung des Intelligenzniveaus angestrebt wird.

#### **Zielgruppengerechte Gestaltung:**

---

Bisherige Erfahrungen von Studierenden bei der Anwendung des Verfahrens in sonderpädagogischen Übungsgutachten im Rahmen des Lehramtstudiums Sonderpädagogik haben gezeigt, dass die KABC-II von Kindern und Jugendlichen gut angenommen wird.

Auch in der sozialpädiatrischen Praxis ist das Verfahren insbesondere bei jungen Kindern mit Entwicklungsbeeinträchtigungen, Verdacht auf Intelligenzminderungen, schulischen Leistungsproblemen gut einsetzbar.

#### **Zugangsfertigkeiten / Einsatz bei behinderten Kindern / Testfairness:**

---

Die Untertests der KABC-II setzen teilweise Zugangsfertigkeiten voraus, über die Kinder mit körperlichen und motorischen Beeinträchtigungen und/oder Sinnesbeeinträchtigungen nicht oder nur eingeschränkt verfügen. Dadurch kann die Testdurchführung erschwert und die Testfairness gefährdet werden:

- Die aktivsprachlichen Anforderungen der KABC-II sind insgesamt im Vergleich zu anderen Testbatterien eher als niedrig einzuschätzen, da verbale Untertests auch mit 1-Wortäußerungen oder mit dem Zeigen auf ein Bild erfolgreich bewältigt werden können. Sprechstörungen (z. B. Dysarthrie/Anarthrie bei Kindern mit cerebralen Bewegungsstörungen) werden sich vor allem auf *Zahlen nachsprechen*, *Symbole* und *Wortreihe* (bei den Interferenzaufgaben) auswirken und – bei starker Beeinträchtigung – auch die Durchführung aller weiterer Untertests, die verbale Reaktionen verlangen, erschweren.
- Motorische Anforderungen in der KABC-II beschränken sich in den meisten Untertests auf Zeigen sowie Aufnehmen und Ablegen von Testmaterial. Grafomotorische Leistungen werden nicht verlangt. Bei deutlichen – insbesondere bilateralen – handmotorischen Beeinträchtigungen ist eine erschwerte Bearbeitung von *Dreiecke*, *Handbewegungen*, *Geschichten ergänzen* und *Rover* zu erwarten. Motorische Beeinträchtigungen dürften sich generell am stärksten auf die Untertests auswirken, die handmotorische Leistungen bei begrenzter Bearbeitungszeit verlangen.
- In allen Untertests werden die Instruktionen standardmäßig verbal gegeben. Da eine konsequent sprachfreie Testdurchführung nicht möglich ist, können sich Hörbeeinträchtigungen insbesondere negativ auf Untertests auswirken, die ohne visuelles Material durchgeführt werden, bzw. deren Aufgabendarbietung nicht selbsterklärend ist. Längere verbale Erklärungen finden sich auch beim sog. *Sprachfrei-Index*, wenn bei falsch gelösten Anfangsaufgaben Hilfen gegeben werden, z. B. bei *Handbewegungen*: „Siehst Du, wie meine Hand aussieht? Sie liegt flach auf dem Tisch. Mache, dass Deine Hand genau so aussieht wie meine. Jetzt berühre den Tisch zwei Mal damit. Mache das zusammen mit mir.“ Die angeblich sprachfreien Untertests *Dreiecke*, *Geschichten ergänzen*, *Bausteine zählen* und *Muster ergänzen* enthalten einen Hinweis zur Bedeutung des Arbeitstempos, der kaum sprachfrei gegeben werden kann: „Ich werde ab jetzt schauen, wie lange wir hierfür brauchen. Du sollst Dich aber nicht hetzen! Das Wichtigste ist, die richtige Antwort zu geben“. Es muss zudem bedacht werden, dass Kinder mit leichteren Hörbeeinträchtigungen unter Umständen zwar Aufgaben verstehen, aber aufgrund der damit verbundenen vermehrten Anstrengung schneller ermüden.
- Fast alle Untertests der KABC-II erfordern die Verarbeitung von visuellen Reizen. Das visuelle Material der KABC-II ist in den meisten Untertests optisch ansprechend und klar gestaltet. Zum Teil (z. B. bei *Muster ergänzen*) wird die Lösung von Aufgaben nur gelingen, wenn visuelle Details sicher erkannt werden können. Im Untertest *Wort- und Sachwissen* ist die Qualität der Abbildungen schlechter. Dies zeigt sich vor allem bei einigen kontrastarmen Fotos. Sehstörungen oder Wahrnehmungseinschränkungen können die Bearbeitung von Aufgaben erschweren und/oder zu einer schnelleren Erschöpfung führen. Bei Sehstörungen ist nicht nur an Visusminderungen zu denken. Auch cerebrale Sehfunktionsstörungen (z. B. Gesichtsfeldeinschränkungen, okulomotorische Störungen, reduzierte Fixation, gestörte Kontrastsensitivität oder Beeinträchtigungen des Farbsehens) beeinflussen unter Umständen die Durchführung.

Im Handbuch gibt es nur sehr vereinzelt Empfehlungen, wie mit Beeinträchtigungen bei Zugangsfertigkeiten umzugehen ist. So wird empfohlen, bei Untertests mit Zeitbonus bei Kindern mit motorischen Beeinträchtigungen eine Auswertung ohne Vergabe von Zusatzpunkten vorzunehmen. Im oberen Leistungsbereich kommt es dadurch allerdings zu Deckeneffekten, worauf im Handbuch aber auch hingewiesen wird.

**CHC-theoretische Einordnung:**

Die Zuordnung von Untertests zu CHC-Faktoren erfolgt in der Regel auf empirischer Basis unter Einbeziehung von Expertenurteilen. Dabei werden in der Fachliteratur die Untertests der KABC-II nicht durchgängig so eingeordnet, wie es der Testaufbau vorsieht.

**Tabelle 4**

*Zuordnung der Untertests zu den CHC Faktoren*

CHC Schicht-II	Untertest [CHC Schicht-I]
Fluide Intelligenz ( <i>Gf</i> )	Muster ergänzen [Induktives Denken] (ab 7;0 Jahre)
Kristalline Intelligenz ( <i>Gc</i> )	Wortschatz [Lexikalisches Wissen] Rätsel [Lexikalisches Wissen] Wort- und Sachwissen [Lexikalisches Wissen]
Kurzzeitgedächtnis ( <i>Gsm</i> )	Zahlen nachsprechen [Gedächtnisspanne] Wortreihe [Gedächtnisspanne]
Visuelle Verarbeitung ( <i>Gv</i> )	Dreiecke; Bausteine zählen [Visualisierung] Muster ergänzen [Visualisierung] (5;0 – 6;11 Jahre) Gestaltschließen [Geschwindigkeit der Gestalterkennung] Wiedererkennen von Gesichtern [Visuelles Gedächtnis]
Langzeitspeicherung- und abruf ( <i>Glr</i> )	Atlantis [Assoziatives Gedächtnis] Symbole [Assoziatives Gedächtnis] Atlantis – Abruf nach Intervall [Assoziatives Gedächtnis] Symbole – Abruf nach Intervall [Assoziatives Gedächtnis]
Sonstige & unklare Zuordnungen	Geschichten ergänzen: <i>Gf, Gv, Gc</i> Rover: <i>Gf, Gv</i> Konzeptbildung: <i>Gv, Gf</i> Handbewegungen: <i>Gsm, Gv</i>
Nicht berücksichtigte CHC-Faktoren (Auswahl): Auditive Verarbeitung ( <i>Ga</i> ) Verarbeitungsgeschwindigkeit ( <i>Gs</i> ) Quantitatives Wissen ( <i>Gq</i> ) Lesen & Schreiben ( <i>Grw</i> )	

Tabelle 4 orientiert sich an den umfassenden Klassifikationen von Testverfahren durch Flanagan, Ortiz & Alfonso (2013, 2017; s. a. Mickley & Renner, 2019). Sie gibt daher nicht die CHC-orientierte Skalenbildung der KABC-II wieder, sondern macht deutlich, welche Untertests auf Schicht-II des CHC-Modells inhaltlich nicht nur einem breiten Intelligenzfaktor zugeordnet werden können. Diese Unklarheiten werden auch im Handbuch der KABC-II (Seite 127ff.) deutlich, wenn die Autoren auf Schicht I im CHC-Modell bei verschiedenen Untertests (z. B. *Handbewegungen, Konzeptbildung, Rover*)

Einzelfähigkeiten aufführen, die unterschiedlichen Schicht-II-Faktoren zugeordnet werden. Beispiel: Der Untertest *Rover* wird in der KABC-II der Skala *Visuelle Verarbeitung (Gv)* zugeordnet. Als Einzelfähigkeiten nennen die Autoren neben „Räumliche Erfassung“ (Schicht-II-Faktor Visuelle Verarbeitung) auch „Generelles sequentielles Denken“ (Schicht-II-Faktor Fluide Intelligenz) und „Mathematische Leistung“ (Schicht-II-Faktor Quantitatives Wissen). Diese Inkonsistenzen werden im Handbuch nicht näher thematisiert. Flanagan und Ortiz et al. (2013) halten es für sinnvoll, dass jeder Schicht-II-Faktor durch mindestens zwei unterschiedliche Schicht-I-Faktoren erfasst werden sollte. Dies soll zum einen verhindern, dass ein Testergebnis auf Schicht II ausschließlich durch einen sehr spezifischen Leistungsbereich bestimmt wird, zum anderen können nur so intraindividuelle Leistungsunterschiede zwischen verschiedenen Schicht-I-Faktoren zumindest ansatzweise abgebildet werden. In der KABC-II wird diese Forderung nur für *Simultan/Gv* erfüllt.

### **Stärken:**

---

- Abdeckung eines breiten Altersbereichs
- Aktuelles Intelligenzmodell (CHC-Modell)
- Breite Abdeckung des Konstruktes Intelligenz durch die Berücksichtigung von fünf Faktoren des CHC-Modells bei Nutzung des Fluid-Kristallin-Index (FKI)
- Abwechslungsreiche Aufgabenstellungen gewährleisten die Motivation auf Seiten der Testperson
- Testmaterial ist für Testpersonen ohne motorische und Sinnesbeeinträchtigungen gut handhabbar
- Einführungs- und Lernaufgaben zur Sicherung des Instruktionsverständnisses
- Aufstellordner mit klaren Kennzeichnungen für Übungsaufgaben erleichtern die Durchführung
- Bei der sprachlichen Durchführung klar vorgegebene und eindeutige Instruktionen
- Möglichkeit zur Überprüfung des Langzeitgedächtnisses (Abruf nach Intervall)
- Expressivsprachliche Anforderungen sind eher gering, was Kindern mit expressiven Sprachstörungen oder Sprechstörungen entgegenkommt.
- Sehr hohe Reliabilitäten der Skalen und der Gesamtwerte
- Umfangreiche statistische Daten zu Signifikanz und Häufigkeit von Profildifferenzen, die eine systematische Beurteilung individueller Stärken und Schwächen ermöglichen
- Auch wenn die Entscheidungsregeln über die Interpretierbarkeit von Skalen kritisch diskutiert werden könnten (vgl. Schneider, 2011), regen sie doch an, bei Diskrepanzen zwischen ähnlichen Aufgabenstellungen nach möglichen Ursachen (z. B. inhaltliche Aspekte, Motivation, Aufmerksamkeit) zu fragen und ggf. ergänzende Überprüfungen vorzunehmen.
- Auf Ebene der Gesamtwerte und Skalenindices können – abgesehen von den jüngsten Altersgruppen – auch extrem unterdurchschnittliche kognitive Leistungen gemessen werden (der Testboden bei einem IQ von 40 entspricht einem Prozentrang von 0.003, die Testdecke von 160 einem Prozentrang von 99.997). Die KABC-II ist daher über einen sehr breiten Leistungsbereich einsetzbar.
- Aktuelle Normierung
- Kinder mit psychischen und Entwicklungsstörungen wurden nicht aus der Normierung ausgeschlossen.



- PC-Auswertungsprogramm ermöglicht schnelle und zuverlässige Ermittlung einer Vielzahl von Testwerten und Daten zur Profilinterpretation. Auch die vergleichende grafische Darstellung mehrerer Testuntersuchungen bei einer Person ist möglich. Bei Hochrechnungen, Substitutionen und Testrohwerten von Null erfolgen Warnungen, die auf die eingeschränkte Interpretierbarkeit hinweisen.

### **Einschränkungen / Schwächen:**

---

- Teilweise intransparente und nicht nachvollziehbare Darstellungen im Manual, z. B. im Abschnitt zur faktoriellen Validität, fehlende Signifikanzprüfungen bei einzelnen statistischen Daten
- Aufgrund der teilweise nicht nachvollziehbaren und nicht replizierbaren Daten zur faktoriellen Validität kann diese nicht als empirisch gesichert angenommen werden.
- Es fehlen Daten zur Validität der Profilinterpretationen
- Es fehlen Daten zur längsschnittlichen Stabilität von Testprofilen
- Die im Handbuch (S. 25) zu findende Aussage, dass die KABC-II Unterschiede in den Testleistungen „zwischen verschiedenen ethnischen und kulturellen Gruppen reduziert“, wird durch die entsprechenden Daten nicht hinreichend belegt.
- Für den sog. *Sprachfrei-Index* steht keine standardisierte sprachfreie Instruktion zur Verfügung
- Die Normierungsstudie wurde durch ein Meinungsforschungsinstitut durchgeführt. Zur Qualifikation der Testleiter\*innen und zur Qualitätskontrolle werden keine Angaben gemacht.
- Die Stichprobengrößen der Normierung sind für die Ergänzungsuntertests durchgehend kleiner als für die Kernuntertests. Im Manual finden sich keine Hinweise, nach welchen Kriterien Ergänzungsuntertests durchgeführt wurden oder nicht.
- Die Stichprobengrößen des sog. *Sprachfrei-Index* sind zum Teil sehr klein ( $N = 35$  für 6-Jährige).
- KABC-II deckt im Prinzip einen breiten Altersbereich ab; die wechselnde Testzusammenstellung erschwert jedoch vor allem einen längsschnittlichen Vergleich von Testwerten, wenn die Ersttestung vor dem siebten Geburtstag erfolgte.
- Die Auswertungsregeln lassen zu, dass IVI und FKI noch gebildet werden können, wenn zwei Untertests einen Rohwert von Null aufweisen. Da Rohwerte von Null nicht eindeutig interpretierbar sind, ist es fraglich, ob in diesem Fall tatsächlich ein Gesamtwert berechnet werden sollte. Kritisch ist dies vor allem bei den 3-Jährigen, bei denen der IVI nur aus 5 Untertests gebildet wird und bei denen zudem Bodeneffekte vorliegen.
- Der Untertest *Wort- und Sachwissen* ist angloamerikanisch (Edison, San Francisco, Wilder Westen, Florence Nightingale etc.) geprägt und prüft daher zum Teil spezifisches Kulturwissen, was in Bezug auf Allgemeinwissen bei Kindern aus anderen Ländern oder anderen Kulturkreisen zu Fehlinterpretationen führen kann.
- Beim Untertest *Wortreihe* ist die Folgeinstruktion nach der ersten Aufgabe möglicherweise missverständlich. Hier heißt es nur noch „Berühre die Bilder, die ich gesagt habe“. Die für die Bewertung ausschlaggebende Information, dass dies in derselben Reihenfolge gemacht werden muss, wird nur dann wiederholt, wenn das Item falsch gelöst wurde.

- Bei Aufgaben mit der Vergabe von Bonuspunkten für eine schnelle Lösung wird den Probanden nicht vermittelt, dass Geschwindigkeit ein wichtiger Faktor bei der Bewertung ist. Explizit heißt es im Handbuch sogar, dass dieser Aspekt nicht hervorgehoben werden sollte. Der Vorschlag zu sagen „Ich schau dann mal, wie lange wir hierfür brauchen“ im Handbuch (S. 67) lässt die Testperson über den Zweck der Zeitmessung im Unklaren.
- Durch die Vergabe von Bonuspunkten in den Untertests *Dreiecke*, *Muster ergänzen* und *Geschichten ergänzen* entsteht ein hoher Einfluss der Bearbeitungsgeschwindigkeit auf das Testergebnis. Dies betrifft insbesondere die höheren Altersstufen. Beispielsweise beträgt bei einem Jugendlichen mit 16;6 Jahren der Unterschied zwischen den richtigen Lösungen in der Maximalzeit und den richtigen Lösungen in der Minimalzeit im Untertest *Geschichten ergänzen* 11, im Untertest *Dreiecke* sogar 13 Skalenwertpunkte. Bei den 7-Jährigen können in diesen Untertests mit jeweils 14 Skalenwertpunkten auch ohne Bonuspunkte leicht überdurchschnittliche Ergebnisse erreicht werden.
- Die Ringordner sind sehr anfällig für Beschädigungen, wenig haltbar und wiesen in mindestens einer uns vorliegenden Version des Verfahrens Fehler in der Zusammenstellung auf. Diese Fehler sind vermutlich in neueren Versionen behoben, dennoch empfiehlt sich die Kontrolle unbekannter Testkoffer.
- Bildmaterial nicht durchgehend von hoher Qualität
- Der Untertest *Symbole* verwendet auch konkret-anschauliche Symbole, die u. U. nicht erst in der Testsituation gelernt werden müssen.
- In den Untertests *Atlantis* und *Symbole* erfolgt in Abhängigkeit von der Anzahl der vorgegebenen Items eine Umrechnung der erzielten Punkte in einen Testrohwert. Eine empirische Grundlage dieser Berechnung wird im Handbuch nicht angegeben.
- Im Untertest *Atlantis* werden auch Punkte vergeben, wenn ein Kind eine Figur zeigt, die der gleichen Kategorie (Fisch, Pflanze, Muschel) wie der Zielreiz angehört. Die Kategorie kann jedoch auch anhand der Silbenzahl der Namen erkannt werden. Daher ist es denkbar (aber nicht wahrscheinlich), dass auch ohne Gedächtnisabruf Punkte erzielt werden.

### **Sinnvolle Ergänzungen:**

Für die Erfassung von in der KABC-II nicht abgedeckten bedeutsamen Intelligenzfaktoren und Leistungsbereiche kommen u. a. in Frage:

- Zur Erfassung der auditiven Verarbeitung (*Ga*) Basiskompetenzen für Lese-Rechtschreibleistungen (BAKO 1-4; Stock, Marx & Schneider, 2003), bei diesem Verfahren wird auch das auditive Arbeitsgedächtnis stark gefordert. Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (TEPHOBE; Mayer, 2013). Ergänzende pädaudiologische Untersuchung
- Zur Erfassung der Lese- und Schreibfähigkeit Lese- und Rechtschreibtests
- Zur Erfassung mathematischer Leistungen Rechentests
- Zur Erfassung des Sprachverständnisses ohne expressivsprachliche Komponente der Leseverständnistest für Erst- bis Siebtklässler - Version II (ELFE II; Lenhard, Lenhard & Schneider, 2018; sichere Lesekenntnisse werden vorausgesetzt), *Handlungssequenzen* aus dem Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren (SET 5-10; Petermann, 2015), der Test zur Überprüfung des Grammatikverständnisses (TROG-D; Fox-Boyer & Bäumer, 2016).

- Zur Erfassung der Verarbeitungsgeschwindigkeit z. B. entsprechende Untertests aus WPPSI-IV und WISC-V
- Ergänzende Überprüfungen zum Langzeitgedächtnis mit Aufgabenstellungen, die nicht nur Paarassoziationslernen verlangen, z. B. der Verbale Lern- und Merkfähigkeitstest (VLMT; Helmstaedter, Lendt M. & Lux S., 2001), die Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-Jährige (BASIC-MLT; Lepach & Petermann, 2008).

### **Anforderungen an die Testleiterinnen & Testleiter:**

---

Wie jeder psychometrische Test sollte die KABC-II nur von Fachkräften eingesetzt werden, die über fundierte Kenntnisse in Testtheorie und Teststatistik verfügen. Ein tieferes Verständnis für den theoretischen Hintergrund der KABC-II ist auch erforderlich, weil nicht alle Empfehlungen des Manuals unbedacht übernommen werden sollten. Testleiter\*innen sollten auch über ein solides Grundwissen im Bereich der kognitiven Entwicklung verfügen und in der Lage sein, eine kindzentrierte und leistungsförderliche Arbeitsatmosphäre herzustellen.

Bei der KABC-II handelt es sich um ein komplexes Verfahren, bei dem vielfältige Anwendungs- und Auswertungsregeln zu beachten sind. Die Testanweisungen sind weitgehend ausführlich, klar und nachvollziehbar beschrieben, verlangen aber vor einer ersten Durchführung eine intensive Beschäftigung mit dem Handbuch (vgl. Joél, 2017). Es ist definitiv mit einem hohen Aufwand bei der Einarbeitung zu rechnen. Ohne regelmäßige Textpraxis wird eine sichere Handhabung des Tests, wie sie für klinische und sonderpädagogische Diagnostik unbedingt erforderlich ist, nicht zu erreichen sein. Mehrere Probestellungen sind unabdingbar, die Kontrolle der Testdurchführung durch kollegiales Feedback oder Videoaufnahmen ist nachdrücklich anzuraten. Testanwender\*innen müssen auch darauf vorbereitet sein, dass in unterschiedlichen Altersgruppen jeweils andere Testzusammenstellungen erforderlich sind.

### **Zusammenfassende Bewertung von Markus Scholz:**

---

In der praktischen Anwendung innerhalb sonderpädagogischer diagnostischer Fragestellungen kenne ich nur die Vorgängerversion Kaufman Assessment Battery for Children (K-ABC; Melchers & Preuß, 1991). Die Erfahrungen mit der KABC-II entstammen der Betreuung von studentischen Übungsgutachten, bei denen das Verfahren zum Einsatz kam, sowie der theoretischen Analyse im Rahmen von Seminarveranstaltungen im Hochschulkontext.

Was die praktische Anwendung anbelangt, ist mir bei der KABC-II im Gegensatz zur Vorgängerversion sofort der stärker motivierende Testeinstieg aufgefallen. Auch die weiteren Kernuntertests sind so gestaltet, dass sie Kinder und Jugendliche zur Mitarbeit anregen. Die visuelle Gestaltung des Materials ist weitgehend gelungen. In dem Kontext muss jedoch kritisch die Bildqualität im Untertest *Wort- und Sachwissen* erwähnt werden. Die Ringordner helfen dem Testleiter in der Durchführung mit präzisen und für das jeweilige Item passenden Hinweisen. Dies erleichtert die praktische Anwendung, da auch immer Informationen zu Anpassungen und Abbruch gegeben werden. Bei häufigen Anwendungen zeigten die Ordner aber schnell Auflösungserscheinungen.

Vereinzelt unpräzise oder die tatsächliche Bewertung mit Bonuspunkten verschleiende Instruktionen (s. Abschnitt Schwächen) haben in den Lehrveranstaltungen bei den Studierenden zum Teil für Verwirrung gesorgt. Die Anweisungen sollten gegenüber der Testperson fair und transparent sein.

Aus theoretischer Perspektive ist die umfängliche und breite Fundierung des Verfahrens zu nennen. Zwar ist mehr als fraglich, inwieweit man tatsächlich von einer dualen theoretischen Perspektive sprechen kann, da für beide Theorien mit Ausnahme der Wissensbereiche dieselben Untertests eingesetzt werden. Dennoch ist der explizite Bezug zum CHC-Modell positiv. Dem Testleiter erleichtert dies die Interpretation sowie die Auswahl möglicher Ergänzungen für diagnostische Fragestellungen erheblich.

Bei einzelnen Items – insbesondere im Untertest *Wort- und Sachwissen* – muss eine US-amerikanische Prägung angemerkt werden. Die Items erfassen aus meiner Sicht entgegen der Behauptung im Handbuch (S. 150) kulturspezifisches Wissen der „westlichen Welt“, welches durchaus von ethnischer oder kulturelle Gruppenzugehörigkeit abhängig ist. Die Daten zu Unterschieden aufgrund ethnischer Zugehörigkeit im amerikanischen Handbuch stützen diese Vermutung. Große Unterschiede zwischen der Skala *Wissen/Kristalline Fähigkeiten (Gc)* und den anderen Fähigkeitsbereichen könnten möglicherweise hier eine Ursache haben. Testleiter\*innen sollten daher genau überlegen, inwieweit sie diese Skala in einem derartigen Fall zur Berechnung des Gesamtwertes heranziehen wollen, da dadurch unter Umständen das tatsächliche Leistungsvermögen der Testperson unterschätzt wird.

Leider macht das Handbuch der KABC-II keinerlei Angaben zu Modifikationen für Kinder mit Beeinträchtigungen der Zugangsfertigkeiten. Zwar sind die verbalsprachlichen und motorischen Anforderungen an die Testperson gering, jedoch kann insbesondere bei schwerer motorisch eingeschränkten Kindern und Jugendlichen die klare Vorgabe von Zeitgrenzen in einzelnen Untertests zu einer Benachteiligung führen. Instruktionen für die Untertests werden in der KABC-II ausschließlich verbal gegeben. Eine faire Testung für Kinder mit Hörbeeinträchtigung ist daher insbesondere bei den Aufgaben ohne visuelle Stimuli kaum möglich.

### **Zusammenfassende Bewertung von Gerolf Renner:**

---

Die Vorläuferversion der KABC-II, die K-ABC, war während meiner Tätigkeit in der Sozialpädiatrie eines der am häufigsten von mir eingesetzten intelligenzdiagnostischen Verfahren im Kontext der Einschulung und im Grundschulalter. Die KABC-II hätte heute vermutlich einen ähnlichen Stellenwert. Sie ermöglicht eine umfassende und sehr reliable Erfassung kognitiver Leistungen. Die hohe Reliabilität des Verfahrens bestätigte sich auch in einer ersten unabhängigen Studie zur klinischen Anwendung (Irblich, Schroeder & Renner, in Vorb.).

Die eher geringen expressivsprachlichen und motorischen Anforderungen kommen einer Diagnostik bei Kindern mit Kommunikationsstörungen und leichten handmotorischen Einschränkungen entgegen. Die Aufgabenstellungen sind abwechslungsreich. Eine gewisse Flexibilität bei der Testzusammenstellung ist gegeben, so dass in begrenztem Rahmen auf fehlende Zugangsfertigkeiten bei den untersuchten Kindern eingegangen werden kann. Während bei vielen Intelligenztests IQ-Werte unter 55 oder 50 prinzipiell nicht abgebildet werden können, liegt der allgemeine Testboden bei Gesamtwerten von 40. Unter der Annahme, dass in jedem Untertest zumindest ein Rohwertpunkt erzielt wurde, können ab dem Alter von vier Jahren Intelligenzwerte unter 50 und ab dem Alter von acht Jahren Werte von 40 erreicht werden.

Einer rundum positiven Bewertung der KABC-II stehen allerdings einige Mängel und Unklarheiten gegenüber, die im Folgenden ergänzend zu den anderen Bewertungen näher ausgeführt werden sollen.

Hinsichtlich des theoretischen Hintergrunds kann ich die Aussagen im Manual zur Bedeutung der Wahl des Modells (CHC vs. Luria) nicht teilen. Die Doppelbenennung von Skalen ändert überhaupt nichts an den erfassten Fähigkeiten. Der einzige Unterschied in der Testzusammenstellung und -auswertung liegt darin, dass im IVI die kristalline Intelligenz nicht berücksichtigt wird. Dieser Schritt sollte gut überlegt sein, denn *Gc* ist neben *Gf* der Intelligenzfaktor mit dem engsten Zusammenhang zur Allgemeinintelligenz und hat eine hohe Bedeutung für schulisches Lernen und Alltagsbewältigung. Individuelle Stärken und Schwächen in *Gc* sollten daher in einer Intelligenzdiagnostik nicht übersehen werden – auch dann nicht, wenn es sich um Testpersonen handelt, für die laut Handbuch (S. 35) das Luria-Modell bevorzugt werden sollte.

Einen völligen Verzicht auf die Erfassung von *Gc* kann ich mir nur vorstellen, wenn die Fragestellung dies definitiv nicht erfordert oder wenn das untersuchte Kind noch keine oder nur begrenzte Gelegenheit hatte, die deutsche Sprache zu erwerben. In diesem Fall wäre allerdings primär der Einsatz eines nonverbalen Verfahrens sinnvoll, denn auch die Skalen des Luria-Modells sind weder sprachfrei noch kulturunabhängig. Der sog. *Sprachfrei-Index* der KABC-II kann eine konsequent sprachfreie Testung nicht ersetzen (s. Abschnitt Schwächen). Bei Kindern mit Migrationshintergrund generell auf *Gc* zu verzichten, hat zum einen den Nachteil, dass ein wichtiger prognostischer Faktor nicht berücksichtigt wird, zum anderen sind zumindest Einzelfälle dokumentiert, bei denen schon nach wenigen Jahren Aufenthalt in Deutschland kristalline Intelligenz eine Leistungsstärke darstellt (Irblich & Renner, 2016). In einer Studie von Scheiber und Kaufman (2015) zeigte sich bei der US-amerikanischen KABC-II, dass der FKI bei ethnischen Minoritäten eine bessere Prognose von Schulleistungen ermöglichte als IVI und SFI.

Es spricht aus meiner Sicht überhaupt nichts dagegen, in Zweifelsfällen sowohl IVI und FKI zu berechnen und mögliche Diskrepanzen in der Testinterpretation zu berücksichtigen. IVI und FKI sind mit Werten von .95 und .98 sehr hoch korreliert, und nur deutliche Schwächen bei *Gc* werden zu einem nennenswerten Unterschied zwischen den beiden Gesamtwerten führen. Ich sehe keine überzeugende Begründung, warum man sich, wie im Manual gefordert, *vor* der Testung für das eine oder andere Modell entscheiden sollte.

Im Luria-Modell halte ich die Skalenbezeichnungen *Lernen* und *Planung* für wenig glücklich gewählt. Wie schon ausgeführt, erfasst *Lernen* nur das assoziative Gedächtnis. Daraus kann wenig über die allgemeinen Lernfähigkeiten eines Kindes abgeleitet werden. Bei durchschnittlichen Werten in *Lernen/Glr* sind daher verallgemeinerte Interpretationen wie „Lernen und Langzeitgedächtnis sind nicht beeinträchtigt“ zu vermeiden. Daten aus dem amerikanischen Manual der KABC-II und erste klinische Daten aus Deutschland (Irblich et al., in Vorb.) zeigen auch, dass *Lernen/Glr* im Leistungsprofil von Kindern mit Intelligenzminderungen eine intraindividuelle Stärke darstellt. Diese Skala scheint daher weniger sensitiv für kognitive Entwicklungsstörungen zu sein.

Empirische Daten zur Validität liegen bisher nur für die Faktorenstruktur und – sehr umfangreich – für die konvergente Validität mit anderen Intelligenztests vor. Daten zur prognostischen Validität, zum Zusammenhang mit Schulleistungen und zu klinischen Diagnosegruppen fehlen für die deutschsprachige KABC-II noch ganz. Angesichts der Betonung der neuropsychologischen Grundlagen des Luria-Modells ist es unverständlich, dass im Handbuch keinerlei Daten zum Einsatz der KABC-II bei Kindern und Jugendlichen mit neurologischen Erkrankungen berichtet werden. Es besteht auch ein deutliches Ungleichgewicht zwischen dem Stellenwert, den die Profilanalyse bei der

Interpretation der Testergebnisse haben soll, und empirischen Belegen für deren Nützlichkeit.

In verschiedener Hinsicht hätte ich im Handbuch eine transparentere Darstellung und zusätzliche Informationen erwartet. Dies betrifft u. a. die konfirmatorischen Faktorenanalysen und fehlende Reliabilitätskennwerte für den Abruf nach Intervall bei *Atlantis* und *Symbole*.

Im Handbuch finden sich keine Hinweise auf systematische Verzerrungen bei der Normstichprobe, und das mittlere Intelligenzniveau zeigt nur geringe Unterschiede zu relativ aktuellen Intelligenztests. Dennoch bleibt angesichts der fehlenden Transparenz hinsichtlich der Durchführung der Normierung ein gewisses Unbehagen.

In der klinischen und sonderpädagogischen Anwendung ist insbesondere bei älteren Testpersonen zu beachten, dass bei den Untertests *Geschichten ergänzen*, *Dreiecke* und *Muster ergänzen* Beeinträchtigungen des Arbeitstempos, z. B. bei motorischen Störungen, neurologischen Erkrankungen oder einer geruhsamen Arbeitsweise, zu einer deutlichen Gefährdung der Testfairness führen können.

Bei der Testinterpretation sind bei Wiederholungsmessungen die deutlichen Retest-Effekte in einigen Skalen zu beachten. Es sollte vor einer Testung stets nachgefragt werden, ob das Verfahren bereits durchgeführt wurde.

Intelligenztestungen in der Sonderpädagogik finden vor allem im Rahmen von Feststellungsverfahren zur Ermittlung des sonderpädagogischen Förderbedarfs bzw. des sonderpädagogischen Bildungsanspruches statt. Sonderpädagog\*innen werden angesichts dieses umschriebenen Einsatzgebietes je nach Organisationsform der Diagnostik oft nur wenige Untersuchungen im Laufe eines Jahres durchführen. Die komplexe Struktur der KABC-II, die vielen bei der Durchführung zu beachtenden Regeln (vgl. Joél, 2017) und die Notwendigkeit, sich intensiv und kritisch mit dem theoretischen Hintergrund auseinanderzusetzen, erfordern jedoch einen hohen Aufwand für die Einarbeitung, der vor ersten Anwendungen auf keinen Fall unterschätzt werden sollte. Feststellungsverfahren sind keine Spielwiese für ungeübte Testleiter\*innen. Auch nach mehrmonatigen Phasen ohne Testpraxis wird eine korrekte Durchführung in der Regel eine gründliche Auffrischung der eigenen Kenntnisse erfordern.

Empirische Befunde zum Einsatz der KABC-II bei sonderpädagogischen Fragestellungen liegen – wie bei fast allen Intelligenztests im deutschsprachigen Bereich – noch nicht vor. Aufgrund der allgemeinen Testeigenschaften und der erforderlichen Zugangsfertigkeiten können – bei aller Variabilität der individuellen Voraussetzungen der Schüler\*innen – meines Erachtens folgende Empfehlungen gerechtfertigt werden: Im *Förderschwerpunkt (FSP) Lernen* werden Testungen mit der KABC-II bereits im Kontext der Einschulung bei den meisten Kindern möglich sein. Im *FSP geistige Entwicklung* wird die Untersuchung von Kindern und Jugendlichen mit leichten und teilweise mit mittelgradigen Intelligenzminderungen ohne zusätzliche sensorische oder motorische Beeinträchtigungen überwiegend gelingen. Vor der Einschulung und in den ersten Schuljahren können sich im Einzelfall durchaus einige Kernuntertests als zu schwierig erweisen. Für die Diagnostik bei schweren und schwersten Intelligenzminderungen ist die KABC-II nicht geeignet. Im *FSP körperliche und motorische Entwicklung* können deutliche handmotorische und sprechmotorische Beeinträchtigungen eine Testung erschweren oder unmöglich machen. Im *FSP Hören* muss vor einer Testung beurteilt werden, ob das vollständige Verständnis der Instruktionen gesichert ist. Dies gilt auch im *FSP Sprache* für Kinder mit rezeptiven Sprachentwicklungsstörungen. Der Einsatz im *FSP Sehen* ist erschwert, da die meisten

Untertests visuelles Material verwenden. In allen FSP sollten bei Testpersonen mit (psycho-)motorischen Verlangsamungen die Auswertungsvarianten ohne Zeitbonus bevorzugt werden.

### **Zusammenfassende Bewertung von Manfred Mickley:**

---

Seit dem Erscheinen der KABC-II in Deutschland ist dieses Testverfahren in meiner Arbeit in Sozialpädiatrischen Zentren das Verfahren, das ich am meisten einsetze.

Die KABC-II ist der erste deutschsprachige Intelligenztest, der explizit das CHC-Modell benutzt, um Testergebnisse der untersuchten Kinder zu interpretieren. Für mich ist dieses Modell zum einen ein Rahmen, um Testergebnisse gegenüber Eltern auch mit deren Erfahrungen im Lebensalltag des Kindes zu verknüpfen. Zum anderen erlaubt mir dieses Modell, auch Testergebnisse des Kindes, die bei früheren Intelligenzdiagnostischen Untersuchungen des Kindes erhoben wurden, mit den gleichen begrifflichen Kategorien darzustellen und miteinander zu vergleichen. Nach meiner Erfahrung hilft dies Eltern, erfasste intellektuelle Besonderheiten und Herausforderungen ihres Kindes besser zu verstehen.

Mit nur sehr wenigen Ausnahmen tragen die beiden ersten Untertests (*Atlantis* und *Geschichten ergänzen* für die älteren Kinder, *Atlantis* und *Konzeptbildung* für die jüngeren Kinder) dazu bei, die Kinder in eine gute Arbeitshaltung zu ‚locken‘. Mir gefällt an der KABC-II die hier deutlich werdende Einstellung der Testautor\*innen Alan und Nadeen Kaufman, sicher zu stellen, dass das zu untersuchende Kind erfasst hat, was von ihm mit den Testaufgaben verlangt wird.

Diese Haltung realisiert die KABC-II auch mit ihren Einführungs- und Lernaufgaben, die Feedback- und Korrektur-Möglichkeiten für das Kind ermöglichen. Habe ich den Eindruck, dass das Kind das Anforderungsprinzip der Aufgabe auch nach Einführungs- und Lernaufgaben noch nicht erfasst hat, kann ich weitere nicht standardisierte Erläuterungen geben. In meiner Arbeit mit der KABC-II werde ich daher immer wieder insbesondere bei kognitiv schwachen Kindern herausgefordert, eine Balance zwischen inhaltlicher Validität und Durchführungsobjektivität herzustellen. Mir gefällt es, dass dabei nicht nur auf die Durchführungsobjektivität geachtet wird.

Die Reihenfolge der Aufgaben schafft eine schöne Variabilität und Abwechslung für die Kinder. Oft genug hatte ich den Eindruck, dass die Kinder regelrecht neugierig waren, welche nächste Aufgabe jetzt kommen könnte.

Es gibt ja nicht „die“ KABC-II: hinter dem Test-Kürzel KABC-II verbergen sich mindestens 2 Testverfahren mit unterschiedlichen Aufgabenzusammenstellungen für Kinder ab 7 Jahren und Kinder unter 7 Jahren. Dies hat für mich als Untersucher den Vorteil, dass ich Aufgabenitems zur Verfügung habe, die sehr leicht sind und daher geeignet sind, kognitiv auch sehr schwachen Kindern die Erfahrung zu vermitteln, dass sie etwas können. Ich muss dafür nicht die Testkoffer wechseln. Die Wahrscheinlichkeit, dass mein Kind Null-Lösungen produziert, reduziert sich nach meiner Erfahrung.

Die KABC-II stellt sehr hohe Anforderungen an mich als Testleiter, die erst erfüllt werden können, wenn ich die KABC-II ca. 10-mal durchgeführt habe und dann auch regelmäßig anwende. Dies kann man als Nachteil interpretieren. Ich interpretiere es als Vorteil. Wende ich die KABC-II an, wird immer wieder deutlich, welche hohen Anforderungen auch ich als Testleiter genügen muss, um ein klinisch aussagekräftiges Intelligenzdiagnostisches Ergebnis zu erzielen. Intelligenzdiagnostik aus der Perspektive des

Diagnostikers ist mehr, als nur Kreuze im Testbogen zu machen, die ich dann ausgezählt dem Auswertungsprogramm übergebe.

Die KABC-II ist nicht das Verfahren erster Wahl, um Kinder im oberen Leistungsspektrum (Verdacht auf Hochbegabung, so genannte Teilleistungsdiagnostik ab ca. Schulklasse 4) oder im oberen Altersbereich (ab ca. 12 bis 13 Jahren) zu untersuchen. Wie bereits in dieser Rezension beschrieben, sind hier die Testwerte der KABC-II schnell mit unklar präsentierten kognitiven Verarbeitungsgeschwindigkeitsleistungen konfundiert.

Die KABC-II erfasst den u. U. relevanten CHC-Bereich kognitive Verarbeitungsgeschwindigkeit nicht explizit; sie kann hierfür keinen Skalenwert liefern. Ebenso ist nach meiner Einschätzung der Bereich der Arbeitsgedächtnisleistungen nicht ausreichend repräsentiert. Ein sehr großer Teil der Kinder, mit denen ich es im SPZ zu tun habe, erreicht z.B. im Untertest *Wortreihe* nicht die Items, bei deren Bewältigung über die Gedächtnisspanne hinaus mentale Operationen notwendig sind (Bsp. Zahlen nicht nur wiederzugeben, sondern rückwärts wiederzugeben, oder wie in der KABC-II zwischendurch Interferenzaufgaben zu bewältigen). Und selbst wenn diese Aufgaben erreicht würden, erlaubt die KABC-II dann nicht, hierfür einen separaten Kennwert zu berechnen. Dementsprechend ist es nötig, bei entsprechenden diagnostischen Fragestellungen auf Untertests anderer Verfahren zurückzugreifen.

### **Aktuelle Entwicklungen:**

---

Da die deutschsprachige KABC-II erst 2015 veröffentlicht wurde, ist eine baldige Überarbeitung nicht zu erwarten. In den USA erschien 2018 eine KABC-II mit aktualisierten Normen (KABC-II NU; Kaufman & Kaufman, 2018), daher dürfte in nächster Zeit auch keine umfassende Revision des amerikanischen Originals erfolgen.

### **Testrezensionen:**

---

#### *Rezensionen der deutschen Bearbeitung (ohne Anspruch auf Vollständigkeit):*

- Irblich, D. (2015). Neuere Testverfahren. KABC-II. Kaufman Assessment Battery for Children - II. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 64(8), 635-648.
- Kastner-Koller, U. & Deimann, P. (2016). KABC-II. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 48(3), 165-169.
- Kuschel, A., Kamp-Becker, I., & Ständer, D. (2017). TBS-TK Rezension: Kaufman Assessment Battery for Children-2 (KABC-II). *Psychologische Rundschau*, 68(4), 321-323.
- Renner, G. (2015). Neue diagnostische Verfahren für die Sonderpädagogik. Melchers, P. / Melchers, M. (2015): KABC-II. Kaufman Assessment Battery for Children – II von Alan S. Kaufman / Nadeen L. Kaufman. Deutschsprachige Fassung. *Sonderpädagogische Förderung heute*, 60(4), 439-444.

#### *Rezensionen der amerikanischen Version (ohne Anspruch auf Vollständigkeit):*

- Bain, S. K. & Gray, R. (2008). Test Reviews: Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman Assessment Battery for Children, Second edition. *Journal of Psychoeducational Assessment*, 26(1), 92-101.
- Braden, J. P., Ouzts & Sandye M. (2005). Review of the Kaufman Assessment Battery for Children, Second Edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 517-520). Lincoln, Nebraska: Buros Institute of Mental Measurements.



Thorndike, R. M. (2005). Review of the Kaufman Assessment Battery for Children, Second Edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 520-522). Lincoln, Nebraska: Buros Institute of Mental Measurements.

### Zitierte Literatur:

- Drechsler, R. (2007). Exekutive Funktionen. *Kindheit und Entwicklung*, 18(3), 233–248. <https://doi.org/10.1026/0942-5403.15.1.27>
- Flanagan, D. P., Alfonso, V. C. & Dixon, S. G. (2013). Cross-battery approach to the assessment of executive functions. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning*. New York: Springer Science+Business Media.
- Flanagan, D. P., Ortiz, S. O. & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3<sup>rd</sup> ed.). Hoboken, NJ: Wiley.
- Flanagan, D. P., Ortiz, S. O. & Alfonso, V. C. (2017). *Cross-Battery Assessment Software System 2.1 (X-BASS 2.1)*. Hoboken, NJ: Wiley.
- Fox-Boyer, A. V. & Bäumer, T. (2016). *TROG-D. Test zur Überprüfung des Grammatikverständnisses* (Materialien zur Sprachtherapie, 7. Aufl.). Idstein: Schulz-Kirchner-Verl.
- Grob, A., Meyer, C. S. & Hagmann-von Arx, P. (2009). *Intelligence and Development Scales (IDS). Intelligenz- und Entwicklungsskalen für Kinder von 5-10 Jahren*. Bern: Hans Huber.
- Helmstaedter, C., Lendt M. & Lux S. (2001). *Verbaler Lern- und Merkfähigkeitstest (VLMT)*. Göttingen: Hogrefe.
- Irblich, D. & Renner, G. (2016, April). *KABC-II: Kritische Analyse und Bewertung*, Wels.
- Irblich, D., Schroeder, A. & Renner, G. (in Vorb.). *Psychometrische Eigenschaften der „Kaufman Assessment Battery for Children - II“ (KABC-II) bei 5- und 6-jährigen Kindern*.
- Joél, T. (2017). Das Dilemma der Intelligenzdiagnostik in der Sonderpädagogik – erläutert anhand der neuen KABC-II. *Zeitschrift für Heilpädagogik*, 68(1), 12–21.
- Kaufman, A. S. & Kaufman, N. L. (2018). *Kaufman Assessment Battery for Children, second edition, normative update*. Bloomington: Pearson.
- Krampen, G. (2007). *Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige. KKA ; Manual*. Göttingen: Hogrefe.
- Lenhard, W., Lenhard, A. & Schneider, W. (2018). *ELFE II. Ein Leseverständnistest für Erst- bis Siebtklässler - Version II* (3., korr. Aufl.). Göttingen: Hogrefe.
- Lepach, A. C. & Petermann, F. (2008). *BASIC-MLT Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-Jährige*. Bern: Huber.
- Luria, A. R. & Majovski, L. V. (1977). Basis approaches used in american and soviet clinical neuropsychology. *American Psychologist*, 32(11), 959–968.
- Lurija, A. R. (1992). *Das Gehirn in Aktion*. Reinbek: Rowohlt.
- Mayer, A. (2013). *Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (TEPHOBE)*. (2., akt. Aufl.). München: Ernst Reinhardt.
- Melchers, P. & Preuß, U. (1991). *Kaufman-Assessment Battery for Children - Deutschsprachige Fassung (K-ABC)*. Amsterdam: Swets & Zeitlinger.
- Mickley, M. & Renner, G. (2010). Intelligenztheorie für die Praxis: Auswahl, Anwendung und Interpretation deutschsprachiger Testverfahren für Kinder und Jugendliche auf Grundlage der CHC-Theorie. *Klinische Diagnostik und Evaluation*, 3(4), 447–466.
- Mickley, M. & Renner, G. (2019). Auswahl, Anwendung und Interpretation Auswahl, Anwendung und Interpretation deutschsprachiger Intelligenztests für Kinder und

- Jugendliche auf Grundlage der CHC-Theorie: Update, Erweiterung und kritische Bewertung. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 68(4), 323–343.  
<https://doi.org/10.13109/prkk.2019.68.4.323>
- Moosbrugger, H. & Schermelleh-Engel, K. (2012). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Aufl., S. 325–343). Berlin: Springer.
- Petermann, F. (2015). *SET 5-10. Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren* (3., überarb. & teilw. neu normierte Aufl). Göttingen: Hogrefe.
- Petermann, F. & Lipsius, M. (2009). *Wechsler Preschool and Primary Scale of Intelligence-III - Deutsche Version (WPPSI-III)* (3<sup>rd</sup> ed.). Frankfurt am Main: Pearson.
- Petermann, F. & Petermann, U. (Hrsg.). (2011). *WISC-IV. Wechsler Intelligence Scale for Children - Fourth Edition*. Frankfurt/M.: Pearson.
- Renner, G. & Mickley, M. (2015). Berücksichtigen deutschsprachige Intelligenztests die besonderen Anforderungen von Kindern mit Behinderungen? *Praxis Kinderpsychologie und Kinderpsychiatrie*, 64(2), 88–103.  
<https://doi.org/10.13109/prkk.2015.64.2.88>
- Scheiber, C. & Kaufman, A. S. (2015). Which of the three KABC-II global scores is the least biased? *Journal of Pediatric Neuropsychology*, 1(1-4), 21–35.  
<https://doi.org/10.1007/s40817-015-0004-6>
- Schneider, W. J. (2011). *Do large subtest score differences invalidate composite scores?* Zugriff am 29.08.2019. Verfügbar unter  
<https://assessingpsyche.wordpress.com/2011/07/15/do-large-subtest-score-differences-invalidate-composite-scores/>
- Schneider, W. J. & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment. Theories, tests, and issues* (4<sup>th</sup> ed., pp. 73–163). New York: The Guilford Press.
- Stock, C., Marx, P. & Schneider, W. (2003). *BAKO 1–4. Basiskompetenzen für Leserechtschreibleistungen. Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr*. Göttingen: Beltz Test.
- Süß, H.-M. & Beauducel, A. (2011). Intelligenztests und ihre Bezüge zu Intelligenztheorien. In L. Hornke, M. Amelang & M. Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (Enzyklopädie der Psychologie, Bd. 3, S. 97–234). Göttingen: Hogrefe.

#### **Ausgewählte weiterführende Literatur:**

---

- McGill, R. J. (2015). Interpretation of KABC-II scores: An evaluation of the incremental validity of Cattell-Horn-Carroll (CHC) factor scores in predicting achievement. *Psychological Assessment*, 27(4), 1417–1426.
- McGill, R. J. & Dombrowski, S. C. (2018). Factor structure of the CHC model for the KABC-II: Exploratory factor analyses with the 16 core and supplementary subtests. *Contemporary School Psychology*, 22(3), 279-293.
- McGill, R. J. & Spurgin, A. R. (2017). Exploratory higher order analysis of the Luria interpretive model on the Kaufman Assessment Battery for Children-Second Edition (KABC-II) School-Age Battery. *Assessment*, 24(4), 540–552.
- Morgan, K. E., Rothlisberg, B. A., McIntosh, D. E. & Hunt, M. S. (2009). Confirmatory factor analysis of the KABC-II in preschool children. *Psychology in the Schools*, 46(6), 515-525.

- Potvin, D. C. H., Keith, T. Z., Caemmerer, J. M. & Trundt, K. M. (2015). Confirmatory factor structure of the Kaufman Assessment Battery for Children-Second Edition with preschool children. Too young for differentiation? *Journal of Psychoeducational Assessment*, 33(6), 522-533.
- Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E. & Low, J. A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children--Second Edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly*, 22(4), 511-539
- Scheiber, C. (2016). Do the Kaufman tests of cognitive ability and academic achievement display construct bias across a representative sample of Black, Hispanic, and Caucasian school-age children in grades 1 through 12? *Psychological Assessment*, 28(8), 942–952.
- Whipple Drozdick, L., Kaufman Singer, J., Lichtenberger, E. O., Kaufman, J. C., Kaufman, A. S., & Kaufman, N. L. (2018). The Kaufman Assessment Battery for Children - second edition and KABC-II normative update. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 333–359). New York: The Guilford Press.