



# Psychometrische Eigenschaften der „Kaufman Assessment Battery for Children – II“ (KABC-II) bei 5- und 6-jährigen Kindern

Dieter Irblich, Anne Schroeder, Gerolf Renner

**Zusammenfassung:** Die KABC-II wurde bei 294 Kindern im Alter von 5 und 6 Jahren, die wegen Entwicklungs-, Verhaltens- oder emotionalen Störungen in drei Sozialpädiatrischen Zentren vorgestellt wurden, eingesetzt. Die Reliabilität der Gesamtskalen erwies sich mit Werten von .92 bis .96 als hoch. Die Reliabilitätswerte der Skalen können als mittel bis hoch bezeichnet werden ( $r_{tt} = .86 - .98$ ), auf Subtestebene wird mit Ausnahme von *Zahlen nachsprechen* eine mittlere bis hohe Zuverlässigkeit erreicht ( $r_{tt} = .74 - .97$ ). Für *Lernen/Langzeitspeicher und -erinnerung* fanden sich bei Kindern mit kognitiven Entwicklungsstörungen relativ höhere Werte als in anderen Skalen. Die Korrelation zwischen dem *Fluid-Kristallin-Index* und dem Gesamt-IQ des SON-R 2½–7 betrug in einer Teilstichprobe .71. Zur diskriminativen Validität für klinische Diagnosegruppen sowie zur divergenten Validität für Verhaltensvariablen (VBV 3–6) ergaben sich weitgehend den Erwartungen entsprechende Befunde, mit den niedrigsten Intelligenzwerten bei Kindern mit Intelligenzminderungen.

**Schlüsselwörter:** KABC-II, Intelligenzdiagnostik, Entwicklungsdiagnostik, Frühdiagnostik, Sozialpädiatrie

## Psychometric properties of the “Kaufman-Assessment Battery for Children – II” (KABC-II) in a clinical sample of 5- and 6-year-old children

**Summary:** Reliability and validity of KABC-II, German version, were examined in German Centers for Social Pediatrics (*Sozialpädiatrische Zentren*). 294 children aged 5 to 6 with developmental, behavioural, and emotional disorders were tested. Sociodemographic data and ICD-10 diagnoses were collected from clinical records. Reliability of the *Fluid-Crystallized Index* (FCI) and the *Mental Processing Index* (MPI) is high, ranging from .92 to .96. Reliability of the scales *Learning/Glr*, *Sequential/Gsm*, *Simultaneous/Gv* and *Knowledge/Gc* is average to high ( $r_{tt} = .86 - .98$ ), subtest reliabilities range from .74 to .97. All reliability coefficients are comparable to those of the normative sample. We found higher scores for *Learning/Glr* than for the other scales in children with developmental cognitive disorders. In a subsample of 77 children FCI and the global score of the SON-R 2½–7, a nonverbal intelligence test, showed convergent validity ( $r_{xy} = .71$ ). No significant correlations were found between KABC-II scores and parental ratings of behavioural problems (VBV 3–6), which supports the divergent validity of the test. The results also demonstrated discriminative validity for clinical groups.

**Keywords:** KABC-II, intelligence assessment, developmental assessment, preschool assessment

Eine der Kernaufgaben psychologischer Diagnostik in Sozialpädiatrie und Frühförderung ist die Abklärung kognitiver Entwicklungsstörungen im Hinblick auf Diagnosestellung, Interventionsplanung und Verlaufskontrolle. Die Ergebnisse solcher Untersuchungen können weitreichende Auswirkungen für die betroffenen Kinder und ihre Familien haben, woraus eine

hohe Verantwortung des Untersuchers bei der Testauswahl, Durchführung und Ergebnisinterpretation erwächst. Die Testgüte der zur Verfügung stehenden Messinstrumente wurde i. d. R. vorwiegend an repräsentativen Stichproben überprüft, und es kann nicht ohne Weiteres angenommen werden, dass bei klinischen Stichproben die gleiche Messgüte vorliegt.

Der Intelligenzdiagnostik kommt bei der Abklärung von Entwicklungsstörungen in Frühförderung und Sozialpädiatrie ein hoher Stellenwert zu. Bei manchen Diagnosen (z. B. Intelligenzminderung) ist sie unabdingbar. Eine besondere Bedeutung haben dabei Verfahren, die eine breite Abdeckung bedeutender Intelligenzfaktoren ermöglichen (Renner und Irblich 2009).

Die Kaufman Assessment Battery for Children – II (KABC-II; Melchers und Melchers 2015) ist dabei ein sehr gebräuchliches Verfahren. Es handelt sich um eine umfassende Testbatterie zur Erfassung kognitiver Fähigkeiten im Alter von drei bis 18 Jahren. Die KABC-II beruft sich, wie ihre in der klinischen und sonderpädagogischen Praxis weit verbreitete Vorläuferversion, die Kaufman-Assessment Battery for Children (K-ABC; Melchers und Preuß 1991), auf ein neuropsychologisches Modell der sequentiellen und simultanen Verarbeitung, unter Einbezug exekutiver und mnestischer Funktionen („Luria-Modell“). Zusätzlich bezieht sich die KABC-II auf die Cattell-Horn-Carroll-Intelligenztheorie (CHC-Theorie; vgl. Schneider und McGrew 2018). Die Interpretation des Testergebnisses kann daher nach Ansicht der Autoren aus zwei unterschiedlichen theoretischen Perspektiven vorgenommen werden.

Die KABC-II umfasst 18 Subtests, die je nach Alter des Kindes in unterschiedlichen Kombinationen vorgegeben werden. Auch die Teststruktur verändert sich in Abhängigkeit vom Alter des Kindes mehrfach. Gesamtskalen des Verfahrens sind der *Fluid-Kristallin-Index* (FKI) und der *Intellektuelle Verarbeitungsindex* (IVI). Im IVI werden Untertests zur Erfassung der kristallinen Intelligenz (sprachliche Leistungen) nicht berücksichtigt, ansonsten gibt es keinen Unterschied zwischen diesen beiden Kennwerten. In allen Altersgruppen kann aus Untertests, die keine expressivsprachlichen Leistungen des Kindes verlangen, auch ein sogenannter *Sprachfrei-Index* (SFI) bestimmt werden.

Bei den Fünf- und Sechsjährigen, auf die sich die vorliegende Studie bezieht, werden neben diesen Gesamtskalen vier weitere Skalenwerte ermittelt:

- **Wissen/Kristalline Fähigkeiten (Gc)** mit den Untertests *Wortschatz* und *Rätsel*, die vor allem lexikalisches Wissen ansprechen.
- **Sequentiell/Kurzzeitgedächtnis (Gsm)** mit den Untertests *Zahlen nachsprechen* und *Wortreihe* zur Erfassung der auditiven Gedächtnisspanne.
- **Simultane Verarbeitung/Visuelle Verarbeitung (Gv)** mit den Untertests *Konzeptbildung*, *Muster ergänzen*, *Dreiecke* und – nur bei den Sechsjährigen – *Rover*. Inhaltliche Schwerpunkte sind Visualisierung und das Erfassen räumlicher Zusammenhänge.
- **Lernen/Langzeitspeicher und -erinnerung (Glr)** mit den Untertests *Atlantis* und *Symbole*. Erfasst wird das assoziative Gedächtnis. Die in diesen Untertests erlernten Assoziationen können optional nach einem Intervall von jeweils 15 – 25 Minuten überprüft werden, was die Bestimmung des Skalenwertes *Abruf nach Intervall* (ANI) ermöglicht.

Die standardmäßig vorzugebenden Untertests werden als Kerntests bezeichnet. Sogenannte Ergänzungsuntertests können zur breiteren Erfassung der kognitiven Leistungsbereiche dienen oder nach bestimmten Regeln Kerntests ersetzen.

Das Manual der KABC-II bietet bereits umfangreiche Daten zur Reliabilität sowie zur faktoriellen, konvergenten und divergenten Validität. Befunde zu klinischen Diagnosegruppen werden dagegen nicht berichtet. Zur deutschsprachigen KABC-II finden sich bisher nur wenige Folgestudien, z. B. fand sie bei Kindern aus sozial-benachteiligten Familien Einsatz als Referenzmaß zur IQ-Bestimmung (Eckstein-Madry und Ahnert 2016). Untersuchungen an größeren klinischen Stichproben und gezielte Über-

prüfungen der psychometrischen Qualität des Verfahrens liegen nicht vor (zur amerikanischen Version s. z. B. Kaufman et al. 2005, Whipple Drozdick et al. 2018). Dieser Mangel war Anlass, im Sinne einer klinisch orientierten Praxisforschung die psychometrischen Eigenschaften der KABC-II in einer sozialpädiatrischen Stichprobe zu überprüfen<sup>1</sup>. Dabei stehen folgende Aspekte im Mittelpunkt:

- 1) Deskriptive Statistiken zu Skalen- und Subtestwerten sowie zur Heterogenität von Testprofilen,
- 2) die Reliabilität der Skalen und Subtests,
- 3) die konvergente Validität (durch Vergleiche mit dem Nonverbalen Intelligenztest SON-R 2½ – 7; Tellegen et al. 2007),
- 4) die divergente Validität gegenüber Verhaltensvariablen (durch Vergleiche mit der Elternfassung des Verhaltensbeurteilungsbogens für Vorschulkinder VBV-EL 3 – 6; Döpfner et al. 1993),
- 5) die diskriminative Validität für bedeutsame diagnostische Gruppen,
- 6) der Einfluss migrationsbezogener Variablen auf die Testwerte.

## Methoden

### Stichprobe

Die Gesamtstichprobe umfasst 294 Kinder, die zwischen 2015 und 2018 in drei Sozialpädiatrischen Zentren (SPZ) in Hamburg (n = 70), Rheinland-Pfalz (n = 196) und Mecklenburg-Vorpommern<sup>2</sup> (n = 28) vorgestellt wurden<sup>3</sup>. Das mittlere Alter der Kinder betrug 70.9 Monate (SD = 6.7 Monate). Weitere Stichprobenmerkmale sind in Tabelle 1 zusammengefasst.

Tab. 1: Stichprobenbeschreibung

Alter (n)	
5 Jahre	162
6 Jahre	132
Geschlecht (n, %)	
Mädchen	88 (29.9%)
Jungen	206 (70.1%)
Schwangerschaftsdauer (n, %)	
≤ 27. SSW	7 (2.4%)
28. – 33. SSW	15 (5.1%)
34. – 36. SSW	17 (5.8%)
≥ 37. SSW	238 (81.0%)
Geburtsgewicht (n, %)	
< 1000 g	8 (2.7%)
1000 – 1499 g	5 (1.7%)
1500 – 2499 g	27 (9.2%)
≥ 2500 g	232 (78.9%)
Bildungseinrichtung (n, %)	
Kindertagesstätte	222 (75.5%)
Förder-/Inklusionskindergarten	33 (11.2%)
Eingangs-/Diagnoseklasse	9 (3.1%)
Grundschule	18 (6.1%)
Förder-/Sonderschule	5 (1.7%)
Migrationshintergrund (n, %)	
Migration Kind	15 (5.1%)
Migration ein Elternteil	27 (9.2%)
Migration beide Eltern	31 (10.5%)
Migration nur Großeltern	3 (1.0%)
Kein Migrationshintergrund	216 (73.5%)

Anmerkung: SSW = Schwangerschaftswoche. Abweichungen zu 100% aufgrund fehlender Angaben oder durch Rundungseffekte.

<sup>1</sup> Die Datenerhebung zu dieser Untersuchung wurde durch ein Forschungsstipendium der Deutschen Gesellschaft für Sozialpädiatrie und Jugendmedizin (DGSPJ) unterstützt.

<sup>2</sup> Wir danken Dipl.-Psych. Manfred Mickley, SPZ Rostock, für die Bereitstellung dieser Daten.

<sup>3</sup> Wir danken den ärztlichen Leitungen aller beteiligten SPZ für die Unterstützung der Datenerhebung.

In der klinisch-psychologischen Diagnostik fand sich bei 99.7 % aller Kinder wenigstens eine Diagnose aus dem Kapitel V (Psychische und Verhaltensstörungen) der ICD-10 (Dilling et al. 2015), im Durchschnitt wurden 1.8 Diagnosen gestellt. Am häufigsten vertreten waren:

- 34.1 % Kombinierte Entwicklungsstörungen (F83)
- 21.8 % Störungen des Sozialverhaltens (F91.x)
- 19.0 % Hyperkinetische Störungen (F90.x)
- 16.0 % Emotionale Störungen des Kindesalters (F98.x)
- 15.6 % Rezeptive Sprachentwicklungsstörungen (F80.2x)
- 8.5 % Expressive Sprachentwicklungsstörungen (F80.1)
- 7.5 % Intelligenzminderungen (F7x.x)

Somatische Diagnosen gemäß ICD-10 fanden sich am häufigsten in den Diagnosegruppen Q00–Q99 (angeborene Missbildungen, Deformitäten und Chromosomenanomalien) mit 13.6 % und E00–E90 (endokrine, Ernährungs- und Stoffwechselkrankheiten) mit 7.5 %.

## Erhebungsinstrumente

Bei allen Kindern wurde die KABC-II durchgeführt. Dabei kamen zumindest die Subtests zum Einsatz, die zur Bestimmung des *Fluid-Kristallin-Index* erforderlich waren, bzw. Untertests, die im Rahmen der zulässigen Substitutionsregeln durchgeführt wurden. Bei 228 Kindern wurde zusätzlich die Skala *Abruf nach Intervall* berechnet. Ergänzend wurde die Kooperation der Kinder in der Testsituation mit einem Item aus der „Verhaltensbeobachtung während der Untersuchung“ (VEWU; Döpfner und Petermann 2008) eingeschätzt. Für Teilstichproben lagen weitere Daten vor: Der SON-R 2½–7 wurde bei 77 Kindern durchgeführt, davon bei den meisten Kindern (81.8 %) zeitlich vor der KABC-II. Der mittlere Testabstand zur

Durchführung der KABC-II betrug 434 Tage ( $SD = 228.2$ ; Spannweite 20–784 Tage). Bei 49.4% der Fälle lag mehr als ein Jahr zwischen den beiden Testungen. Der VBV-EL 3–6 wurde bei 59 Kindern zeitnah (Abstand max. 30 Tage) zur Testung mit der KABC-II eingesetzt. Alle diagnostischen Instrumente wurden entsprechend der Vorgaben der Testmanuale durchgeführt und ausgewertet. Auf einer dreistufigen Schätzskala wurde beurteilt, ob Beeinträchtigungen in den Zugangsfertigkeiten Hören, Sehen, Motorik, Sprachverständnis/mangelnde Deutschkenntnisse und Expressivsprache die Testdurchführung nicht, leicht oder stark beeinträchtigten.

## Statistische Auswertung

Die Datenauswertung erfolgte mit der Statistiksoftware SPSS 25. Vergleiche mit den Daten der Normstichprobe erfolgten mittels Einstichproben t-Tests. Reliabilitätsschätzungen erfolgten nach der Split-half-Methode. Die Zuverlässigkeit der Skalen und Gesamtskalen wurde nach Formel 14.4 bei Lienert und Raatz (1994) berechnet. Gruppenvergleiche wurden auf Ebene der Gesamtskalen (FKI, IVI) mit Varianzanalysen vorgenommen. Für die Skalen wurden Gruppenvergleiche mit Varianzanalysen (mixed ANOVA) mit den Skalenwerten als Innersubjektfaktor und der jeweiligen Gruppenvariable (z. B. Diagnosegruppe) als Zwischensubjektfaktor eingesetzt, ergänzt durch univariate Varianzanalysen für jede Skala. Als Maß der Effektstärke wurde das partielle Eta-Quadrat ( $\eta^2$ ) genutzt. Effektstärken werden als klein ( $\eta^2 > .01$ ), mittel ( $\eta^2 > .06$ ) und groß ( $\eta^2 > .14$ ) klassifiziert (Maher et al. 2013). Als Indikatoren heterogener Leistungsprofile wurde für die Skalen und Gesamtskalen jeweils die Spannweite (absolute Differenz zwischen dem höchsten und niedrigsten Standardwert innerhalb der *individuellen* Testprofile) berechnet. Zur Kontrolle des globalen Fehlerniveaus wurde ergänzend die

Bonferroni-Holm-Korrektur verwendet, wenn eine statistische Analyse mehrfach für die Skalen und Gesamtskalen beziehungsweise für die Untertests durchgeführt wurde. Dies hatte durchgehend keine Auswirkung auf die Signifikanzbeurteilung.

Im Manual der KABC-II werden in Tabelle D.7 a Grenzwerte angegeben, mit denen beurteilt werden kann, ob es sich bei der Differenz zwischen einer einzelnen Skala und dem Mittelwert aller Skalen um seltene Ereignisse handelt. Unter Anlegung der Grenzwerte für eine Frequenz von 10% wurde mittels Binomialtests überprüft, ob die Häufigkeit seltener Ereignisse in unserer Stichprobe von den Normierungsdaten abweicht.

## Ergebnisse

### Testdurchführung

Daten zur Anzahl an benötigten Terminen für die Durchführung der KABC-II lagen für 221 Kinder vor. Davon wurden 75.2% an einem Termin untersucht. Durchgeführt wurden durchschnittlich 12.5 Untertests ( $SD = 1.4$ ). Die Testung dauerte bei den Fünfjährigen im Mittel 52.1 Minuten ( $SD = 15.5$ ; Spannweite 30 – 130 Minuten), bei den Sechsjährigen 61.0 Minuten ( $SD = 18.9$ ; Spannweite 35 – 130). Die Testdauer korrelierte signifikant mit dem FKI ( $r = .35$ ,  $p < .001$ ).

Der Mittelwert der von 1 bis 5 reichenden Schätzskala zur Beurteilung der Kooperation in der Testsituation lag bei 1.6 ( $SD = 1.0$ ). Dabei erwies sich die Mitarbeit in 51.7% der Fälle als unproblematisch („befolgt gut und leicht die Anweisungen ...“). Bei 29.9% wurden die Anweisungen nach einfachen Hinweisen eingehalten, bei 11.9% waren „deutliche“ und bei 2.4% „erhebliche“ Interventionen erforderlich. FKI, IVI und die Skalen zeigten mit Ausnahme

von *Wissen/Gc* signifikante Korrelationen mit der Kooperation, die mit Werten zwischen  $-.13$  (*Lernen/Glr*) und  $-.19$  (*Sequentiell/Gsm*) niedrig ausfielen.

Erschwernisse in der Testdurchführung aufgrund eingeschränkter Zugangsfertigkeiten fanden sich in folgenden Bereichen: Hören (leichte Beeinträchtigung 0.7%/starke Beeinträchtigung 0.0%), Sehen (3.1%/0.0%), Motorik (4.4%/0.3%), Sprachverständnis (5.1%/2.7%), Expressivsprache (5.1%/2.7%).

### Deskriptive Skalenstatistiken und Profilheterogenität

Die Mittelwerte der Gesamtskalen und Skalen lagen durchgehend signifikant unter den Werten der Normstichprobe. Bei *Lernen/Glr* und *Abruf nach Intervall* handelt es sich um kleine, sonst um mittlere bis große Effekte (Tab. 2). Auf Subtestebene fanden sich mit Ausnahme von *Atlantis – Abruf nach Intervall* ebenfalls durchgehend signifikant niedrigere Werte, mit dem stärksten Effekt bei *Zahlen nachsprechen*.

Die Varianzanalyse ergab signifikante Unterschiede zwischen den Skalen ( $F(3, 876) = 25.5$ ;  $p < .001$ ;  $\eta^2 = 0.08$ ). Paarweise Vergleiche waren mit Ausnahme der Differenz zwischen *Wissen/Gc* und *Simultan/Gv* signifikant. Auf Subtestebene bestanden für die Kerntests ebenfalls signifikante Unterschiede bei den 5-Jährigen ( $F(8, 1256) = 11.2$ ;  $p < .001$ ;  $\eta^2 = 0.07$ ) und bei den 6-Jährigen ( $F(9, 1125) = 11.6$ ;  $p < .001$ ;  $\eta^2 = 0.09$ ). In beiden Altersgruppen fanden sich die höchsten Werte für *Atlantis*, das sich in paarweisen Vergleichen signifikant von allen anderen Subtests unterschied.

Die Interkorrelationen der Gesamtskalen und Skalen waren durchgehend signifikant (Tab. 3.). FKI und IVI zeigten mit  $r = .95$  einen sehr engen Zusammenhang.

Tab. 2: Deskriptive Kennwerte der KABC-II-Skalen und Subtests und Vergleich mit der Normstichprobe

Skala/Subtest	Mittelwert	Standardabweichung	Spannweite	t <sup>a</sup>	d
FKI	86.9	16.3	43 – 135	-13.8**	-0.80
IVI	86.3	17.4	42 – 138	-13.6**	-0.79
<b>Sequentiell/Gsm</b>	85.9	16.3	45 – 128	-14.8**	-0.86
■ Zahlen nachsprechen	7.0	3.1	1 – 17	-16.3**	-0.95
■ Wortreihe	8.1	3.2	1 – 18	-10.4**	-0.61
<b>Simultan/Gv</b>	88.3	16.6	41 – 136	-12.0**	-0.70
■ Konzeptbildung	8.1	3.0	1 – 16	-10.8**	-0.63
■ Muster ergänzen	8.2	2.7	2 – 17	-11.4**	-0.67
■ Dreiecke	8.3	3.4	1 – 19	-8.9**	-0.52
■ Rover <sup>a</sup>	8.1	2.6	1 – 16	-8.2**	-0.70
<b>Lernen/Glr</b>	93.7	14.1	55 – 148	-7.6**	-0.44
■ Atlantis	9.5	2.9	2 – 19	-3.1**	-0.18
■ Symbole	8.1	3.0	1 – 17	-11.1**	-0.65
<b>Wissen/Gc</b>	90.2	17.8	40 – 146	-9.4**	-0.55
■ Wortschatz	8.7	3.0	1 – 18	-7.5**	-0.44
■ Rätsel	8.0	3.4	1 – 17	-9.9**	-0.58
<b>Abruf nach Intervall</b>	97.4	12.0	65 – 140	-3.2**	-0.22
■ Atlantis	10.1	2.9	2 – 19	0.6	
■ Symbole	9.1	2.6	4 – 18	-5.2**	-0.34

Anmerkungen: FKI=Fluid-Kristallin-Index. IVI=Intellektueller Verarbeitungsindex. d=Effektstärke. <sup>a</sup> Einstichproben t-Tests bei einem Vergleichswert von 100 (Skalen) bzw. 10 (Subtests).

Tab. 3: Interkorrelationen der Skalen und Gesamtskalen der KABC-II

Skala	IVI	Gsm	Gv	Glr	Gc	ANI
FKI	.95**	.79**	.85**	.74**	.74**	.54**
IVI		.80**	.88**	.74**	.59**	.55**
Gsm			.58**	.48**	.51**	.35**
Gv				.54**	.47**	.40**
Glr					.48**	.73**
Gc						.29**

Anmerkungen: FKI=Fluid-Kristallin-Index; IVI=Intellektueller Verarbeitungsindex; Gsm=Sequentiell/Kurzzeitgedächtnis; Gv=Simultan/visuelle Verarbeitung; Glr=Lernen/Langzeitspeicher und -erinnerung; Gc=Wissen/kristalline Fähigkeiten; ANI=Abruf nach Intervall.

\*  $p < .05$ ; \*\*  $p < .01$ .

Testrohwerte von Null als Indiz einer deutlichen Überforderung der untersuchten Kinder fanden sich in geringer Anzahl bei den Untertests *Konzeptbildung* (0.7% der Fälle), *Zahlen nachsprechen* (0.3%), *Rover* (0.3%),

*Atlantis – Abruf nach Intervall* (1.4%), *Symbole* (0.7%), *Muster ergänzen* (3.1%), *Symbole – Abruf nach Intervall* (2.0%) und *Rätsel* (0.3%). Deckeneffekte traten in keinem der Untertests auf.

Die mittlere Spannweite auf Skalenebene betrug 22.1 (SD=9.6) Standardwerte bei den 5-Jährigen und 26.6 (SD=11.5) Standardwerte bei den 6-Jährigen. Heterogene Profile mit einer Spannweite von mindestens 2 Standardabweichungen sind sowohl auf Skalen- als auch auf Untertestebene häufig zu finden und stellen eher den Regelfall dar. Auf Skalenebene betrifft dies 19.1 % der 5-Jährigen und 37.1 % der 6-Jährigen. Auf Untertestebene weisen in beiden Altersgruppen mehr als 70 % der Fälle solche großen Differenzen auf (Tab. 4). Auf Basis der im Manual angegebenen kritischen Werte für die Beurteilung signifikanter Differenzen einer Skala vom Mittelwert aller Skalen fanden sich folgende Anteile statistisch signifikanter Abweichungen, angegeben jeweils für die 5- und 6-Jährigen: *Sequentiell/Gsm* 36.4%/46.2%, *Simultan/Gv* 49.4%/47.7%, *Lernen/Glr* 45.7%/67.4%, *Wissen/Gc* 53.1%/62.1%. Auf Untertestebene (Kerntests) fand sich eine intraindividuelle Spannweite von 6.8 (SD=2.0) bei 5-Jährigen und von 7.3 (SD=2.3) bei 6-Jährigen. Bezogen auf den Grenzwert für eine Häufigkeit von 10 % ergaben Binomialtests bei den 5-Jährigen keine signifikanten Abweichungen von der Verteilung in der Normstichprobe. Bei den 6-Jährigen fanden sich signifikante Unterschiede bei *Simultan/Gv* ( $p < .001$ ), *Lernen/Glr* ( $p = .03$ ) und *Wissen/Gc* ( $p < .001$ ). Dabei traten die als selten klassifizierten Abweichungen häufiger auf als in der Normstichprobe.

Tab. 4: Prozentualer Anteil der Kinder mit bestimmten Ausprägungen der Spannweite der Skalen und der Untertestwerte

	5-Jährige	6-Jährige
<b>Skalen</b>		
■ $\geq 15$ IQ-Punkte	78.4%	84.8%
■ $\geq 30$ IQ-Punkte	19.1%	37.1%
■ $\geq 45$ IQ-Punkte	3.1%	6.8%
<b>Untertests</b>		
■ $\geq 3$ SW	99.4%	99.2%
■ $\geq 6$ SW	73.5%	76.5%
■ $\geq 9$ SW	18.5%	30.3%

Anmerkungen: SW = Skalenwerte ( $M = 10$ ,  $SD = 3$ )

## Reliabilität

Die Reliabilitäten der Kerntests lagen mit Ausnahme von *Zahlen nachsprechen* bei den 5-Jährigen über .80. Werte von .90 und höher erreichten bei den 5-Jährigen vier und bei den 6-Jährigen sechs Kerntests sowie für beide Altersgruppen *Atlantis/Abruf nach Intervall* und *Symbole/Abruf nach Intervall*. Die Zuverlässigkeit der Gesamtskalen betrug zwischen .92 und .96, die Werte der Skalen lagen bei den 5-Jährigen zwischen .86 und .97, bei den 6-Jährigen zwischen .92 und .98 (Tab. 5). Die mittlere Abweichung zu den im Manual angegebenen Werten lag für die Kerntests bei -.01 (5-Jährige) und .02 (6-Jährige), identische Werte ergaben sich bei Betrachtung aller Skalen und Gesamtskalen. Die größte Abweichung betrug bei den Skalen .07 (*Sequentiell/Gsm*) und bei den Kerntests .09 (*Wortreihe, Zahlen nachsprechen*).

Tab. 5: Split-half Reliabilität der KABC-II-Skalen und Subtests im Vergleich zur Normierungsstichprobe

Skala/Subtest	5-Jährige	6-Jährige
FKI	.95 (.96)	.96 (.96)
IVI	.92 (.97)	.95 (.97)
<b>Sequentiell/Gsm</b>	.86 (.89)	.92 (.85)
■ Zahlen nachsprechen	.74 (.83)	.86 (.79)
■ Wortreihe	.81 (.83)	.87 (.78)
<b>Simultan/Gv</b>	.95 (.94)	.97 (.95)
■ Konzeptbildung	.88 (.83)	.86 (.82)
■ Muster ergänzen	.90 (k. A.)	.94 (k. A.)
■ Dreiecke	.91 (.93)	.94 (.93)
■ Rover	–	.89 (.92)
<b>Lernen/Glr</b>	.97 (.96)	.98 (.97)
■ Atlantis	.96 (.96)	.96 (.97)
■ Symbole	.96 (.95)	.97 (.96)
<b>Wissen/Gc</b>	.94 (.95)	.96 (.94)
■ Wortschatz	.86 (.90)	.91 (.89)
■ Rätsel	.92 (.92)	.95 (.93)
<b>Abruf nach Intervall</b>	.94 (.91)	.95 (.92)
■ Atlantis	.91 (k. A.)	.89 (k. A.)
■ Symbole	.93 (k. A.)	.96 (k. A.)

Anmerkungen: In Klammern Reliabilitätskennwerte der Normstichprobe (Melchers und Melchers 2015, 182). FKI = Fluid-Kristallin-Index; IVI = Intellektueller Verarbeitungsindex. k. A. = keine Angabe.

Tab. 6: Korrelationen der KABC-II-Skalen und Gesamtskalen mit Intelligenz- und Verhaltensvariablen

Skala	FKI	IVI	Gsm	Gv	Glr	Gc	ANI
<b>SON-R 2 ½ – 7</b>							
■ Gesamt-IQ	.71**	.73**	.42**	.74**	.49**	.41**	.30*
■ Handlungsskala	.67**	.70**	.50**	.69**	.41**	.34**	.22
■ Denkskala	.71**	.71**	.36**	.73**	.55**	.43**	.44**
<b>VBV 3 – 6</b>							
■ Emot. Kompetenzen	-.05	.01	-.19	.09	.09	-.16	.11
■ Oppos.-aggr. Verhalten	.08	.03	.17	-.12	.04	.20	-.05
■ Hyperaktivität	-.04	-.12	.07	-.21	-.09	.13	-.16
■ Emot. Auffälligkeiten	-.11	-.09	.03	-.17	-.06	-.11	-.03

Anmerkungen: FKI = Fluid-Kristallin-Index; IVI = Intellektueller Verarbeitungsindex; ANI = Abruf nach Intervall. \*  $p < .05$ ; \*\*  $p < .01$ . Gsm = *Sequentiell/Kurzzeitgedächtnis*; Gv = *Simultan/visuelle Verarbeitung*; Glr = *Lernen/Langzeitspeicher und -erinnerung*; Gc = *Wissen/kristalline Fähigkeiten*; ANI = *Abruf nach Intervall*.

## Konvergente & divergente Validität

Die Überprüfung der konvergenten Validität ergab, dass die Gesamtwerte des SON-R 2 ½ – 7 substantiell mit FKI und IVI (Tab. 6) korrelierten, dieser Zusammenhang wurde nicht durch den Zeitabstand moderiert. Der mittlere IQ des SON-R 2 ½ – 7 fiel 4.4 IQ-Punkte niedriger als der FKI ( $t = 2.8$ ,  $p < .001$ ; mittlerer Effekt) und 3.9 IQ-Punkte niedriger als der IVI ( $t = 2.4$ ,  $p < .02$ ; kleiner Effekt). Zwischen der KABC-II und dem VBV 3 – 6 als konstruktfernes Verfahren (divergente Validität) fanden sich keine signifikanten Zusammenhänge (Tab. 6).

## Diskriminative Validität für klinische Gruppen

Für die Analysen zur diskriminativen Validität (Tab. 7) wurden nur Hauptdiagnosen berücksichtigt, die mindestens 20 Fälle umfassten.

- In Bezug auf die Gesamtskalen (FKI, IVI) ergaben sich jeweils große Effekte der Diagnosegruppen, mit den niedrigsten Werten für Kinder mit Intelligenzminderungen, gefolgt von Kindern mit kombinierten Entwicklungsstörungen und rezeptiven Sprachstörungen,

und Werten im Durchschnittsbereich für Kinder mit Störungen des Sozialverhaltens (Tab. 8). Die Interaktion zwischen Gesamtskala und Diagnosegruppe war signifikant ( $F(4, 153) = 7.56$ ,  $p < .001$ ,  $\eta^2 = .17$ ), was auf die Differenzen zwischen FKI und IVI bei Kindern mit Intelligenzminderung (FKI > IVI) und Kindern mit Rezeptiven Sprachentwicklungsstörungen (FKI < IVI) zurückzuführen ist.

- Auf Skalenebene fand sich ein signifikanter Haupteffekt für den Innersubjektfaktor ( $F(3, 459) = 14.08$ ,  $p < .001$ ,  $\eta^2 = .08$ ). Außerdem war die Interaktion zwischen Skala und Diagnosegruppe signifikant ( $F(12, 459) = 6.69$ ,  $p < .001$ ,  $\eta^2 = .15$ ), es zeigten sich also unterschiedliche Ergebnismuster in den einzelnen Gruppen. So erreichten z. B. Kinder mit rezeptiven Sprachentwicklungsstörungen in *Wissen/Gc* nur weit unterdurchschnittliche Ergebnisse, während ihre Werte für *Simultan/Gv* und *Lernen/Glr* noch im Durchschnittsbereich lagen. *Lernen/Glr* fiel vor allem bei Kindern mit Intelligenzminderungen, kombinierten Entwicklungsstörungen und rezeptiven Sprachstörungen deutlich höher aus als das allgemeine Intelligenzniveau (zu Details s. Tab. 8).
- Für die Indikatoren der Profilheterogenität ergab sich in keiner Analyse ein signifikanter Effekt der Diagnosegruppe.



Tab. 7: Deskriptive Statistik der Skalen und Gesamtskalen für ausgewählte Diagnosegruppen und Ergebnisse der univariaten Varianzanalysen

Diagnosegruppe (ICD-Codes)	n	FKI M (SD)	IVI M (SD)	Gsm M (SD)	Gv M (SD)	Glr M (SD)	Gc M (SD)
Intelligenzminderungen (F70.x)	20	64.2 <sub>a</sub> (6.5)	60.4 <sub>a</sub> (6.5)	64.8 <sub>a</sub> (9.7)	65.1 <sub>a</sub> (9.0)	77.1 <sub>a</sub> (9.3)	69.6 <sub>a</sub> (15.9)
Kombinierte Entwicklungsstörungen (F83)	61	77.7 <sub>b</sub> (7.8)	77.2 <sub>b</sub> (10.3)	79.1 <sub>b</sub> (10.6)	78.9 <sub>b</sub> (10.0)	89.7 <sub>b</sub> (12.5)	83.6 <sub>b</sub> (12.7)
Rezeptive Sprachstörungen (F80.2x)	20	79.7 <sub>b</sub> (9.9)	84.1 <sub>b</sub> (11.2)	80.3 <sub>b</sub> (6.6)	91.6 <sub>b</sub> (12.8)	90.5 <sub>b</sub> (10.5)	69.6 <sub>a</sub> (16.2)
Hyperkinetische Störungen (F90.x)	32	88.3 <sub>c</sub> (11.8)	88.6 <sub>c</sub> (12.7)	89.9 <sub>c</sub> (14.2)	88.3 <sub>c</sub> (12.1)	94.6 <sub>b,c</sub> (11.1)	94.2 <sub>b</sub> (13.2)
Störungen des Sozialverhaltens (F91.x)	25	101.7 <sub>d</sub> (16.9)	101.0 <sub>d</sub> (17.9)	103.0 <sub>d</sub> (16.9)	99.4 <sub>d</sub> (14.4)	100.6 <sub>c</sub> (11.0)	101.5 <sub>c</sub> (15.0)
F (df)		40.77 (4, 153)	36.54 (4, 153)	32.70 (4, 153)	30.96 (4, 153)	13.01 (4, 153)	24.12 (4, 153)
p		<.001	<.001	<.001	<.001	<.001	<.001
Effektstärke $\eta^2$		.52	.49	.46	.45	.25	.39

Anmerkung: Innerhalb jeder Spalte unterscheiden sich Mittelwerte mit unterschiedlichem Subskript statistisch bedeutsam voneinander (Duncan-Test,  $p < .05$ ). FKI = Fluid-Kristallin-Index; IVI = Intellektueller Verarbeitungsindex; Gsm = Sequentiell/Kurzzeitgedächtnis; Gv = Simultan/visuelle Verarbeitung; Glr = Lernen/Langzeitspeicher und -erinnerung; Gc = Wissen/kristalline Fähigkeiten.

## Migrationsbezogene Variablen

Die Varianzanalysen (Tab. 8) ergaben folgende Ergebnisse:

- Beim Migrationsstatus fanden sich für FKI und IVI keine signifikanten Unterschiede. Auf Skalenebene ergab die Varianzanalyse einen signifikanten Haupteffekt für den Innersubjektfaktor ( $F(3, 855) = 18.90, p < .001, \eta^2 = .06$ ), außerdem fand sich eine signifikante Interaktion zwischen Gruppe und Skalenwert ( $F(9, 840) = 8.87, p < .001, \eta^2 = .09$ ). In der univariaten Analyse zeigte sich nur bei *Wissen/Gc* ein signifikanter Gruppenunterschied (großer Effekt), mit den niedrigsten Werten bei eigenen Migrationserfahrungen der Kinder und Migrationshintergrund bei beiden Elternteilen.
- Der Vergleich mono- und bilingualer Kinder zeigte für den FKI, nicht jedoch für den IVI, signifikant höhere Werte (kleiner Effekt) für monolingual aufwachsende Kinder. Auf Skalenebene waren der Haupteffekt ( $F(3, 840) = 20.96, p < .001, \eta^2 = .07$ ) und die Interaktion zwischen Gruppe und Index signifikant ( $F(3, 840) = 20.96, p < .001, \eta^2 = .08$ ). Univariat fand sich nur bei *Wissen/Gc* ein deutlicher Gruppenunterschied.

## Diskussion

Die Reliabilitäten der KABC-II zeigten sich in den an der Studie beteiligten SPZ bei 5- und 6-jährigen Kindern vergleichbar mit den Werten aus der Normierungsstichprobe. Die Gesamtskalen weisen nach den Kriterien von Flanagan et al. (2006) eine hohe Reliabilität auf. Auch die Skalen erreichen fast durchgehend hohe Werte, mit Ausnahme von *Sequentiell/Gsm* bei den 5-Jährigen (mittlere Reliabilität). Auf Ebene der Untertests wurde die für klinische Interpretationen wünschenswerte Reliabilität von  $r_{tt} \geq .80$  (Yousfi und Steyer 2006) nur bei *Zahlen nachsprechen* in der Gruppe der 5-Jährigen nicht

Tab. 8: Deskriptive Statistik der Skalen und Gesamtskalen für migrationsbezogene Variablen und Ergebnisse der univariaten Varianzanalysen

	n	FKI M (SD)	IVI M (SD)	Gsm M (SD)	Gv M (SD)	Glr M (SD)	Gc M (SD)
<b>Migration</b>							
■ Keine Migration	216	87.9 (16.8)	86.6 (18.0)	86.0 (16.4)	88.3 (16.8)	93.8 (14.9)	94.0 <sub>a</sub> (16.1)
■ Migration ein Elternteil	27	85.8 (12.5)	87.4 (13.8)	86.2 (13.3)	90.4 (15.2)	94.4 (10.2)	85.3 <sub>b</sub> (15.6)
■ Migration beide Eltern	31	81.6 (12.7)	82.7 (13.6)	84.4 (15.5)	85.5 (13.8)	94.3 (12.8)	74.9 <sub>c</sub> (15.9)
■ Migration Kind	15	84.3 (17.6)	84.9 (17.9)	85.1 (20.7)	89.4 (16.4)	90.3 (10.9)	76.7 <sub>b,c</sub> (23.0)
F (df)		1.56 (3, 285)	0.55 (3, 285)	0.10 (3, 285)	0.47 (3, 285)	0.34 (3, 285)	16.84 (3, 285)
p		.20	.67	.96	.70	.80	<.001
Effektstärke $\eta^2$							.15
<b>Bilingualität des Kindes</b>							
■ Nur Deutsch	236	87.8 (16.7)	86.6 (18.0)	85.9 (16.5)	88.3 (16.7)	93.8 (14.8)	93.8 (15.9)
■ Deutsch & Fremdsprache	46	81.7 (13.3)	83.9 (13.8)	83.5 (14.1)	87.6 (16.2)	92.5 (9.8)	75.7 (17.6)
F (df)		5.31 (1, 280)	0.91 (1, 280)	0.86 (1, 280)	0.07 (1, 280)	0.29 (1, 280)	48.33 (1, 280)
p		.02	.34	.35	.79	.59	<.001
Effektstärke $\eta^2$		.02					.15

Anmerkung: Mittelwerte mit unterschiedlichem Subskript unterscheiden sich statistisch bedeutsam voneinander (Duncan-Test,  $p < .05$ ). FKI = Fluid-Kristallin-Index; IVI = Intellektueller Verarbeitungsindex; Gsm = Sequentiell/Kurzzeitgedächtnis; Gv = Simultan/visuelle Verarbeitung; Glr = Lernen/Langzeitspeicher und -erinnerung; Gc = Wissen/kristalline Fähigkeiten.

erreicht. Dies stellt gegenüber sozialpädiatrischen Daten zur K-ABC (vgl. Renner et al. 2012) eine Verbesserung dar.

Das Leistungsniveau in unserer sozialpädiatrischen Stichprobe lag insgesamt knapp eine Standardabweichung unter der Norm, der Untertest *Zahlen nachsprechen* wies dabei den niedrigsten Wert auf. In den Gesamtskalen fanden sich bei Kindern mit Intelligenzminderungen und kombinierten Entwicklungsstörungen die niedrigsten Testleistungen. Rezeptive Sprachentwicklungsstörungen waren durch weit unterdurchschnittliche Leistungen bei *Wissen/Gc* gekennzeichnet. Eine nähere Untersuchung der Sprachentwicklung erscheint daher bei Vorliegen entsprechender Befunde sinnvoll. Kinder mit Störungen des Sozialverhaltens zeigten dagegen ein homogenes Profil mit Leistungen im Durchschnittsbereich. Dies kann insgesamt als Hinweis auf die diskriminative Validität der KABC-II interpretiert werden, wobei berücksichtigt werden muss, dass diese Diagnosen in SPZ zwar interdisziplinär und unter Berücksichtigung verschiedener Informationsquellen gestellt werden, aber auch von den Ergebnissen der Intelligenzdiagnostik beeinflusst sind. Die Diagnosen wurden daher nicht unabhängig von den Testwerten der KABC-II ermittelt.

Heterogene Testprofile kamen in unserer Stichprobe sehr häufig vor (vgl. die Daten zur K-ABC in Renner 2013). Dies gilt besonders für das Profil der Untertests, bei denen die mittlere intra-individuelle Spannweite 6.8 (5-Jährige) bzw. 7.3 (6-Jährige) Skalenwerte umfasst. Die Analyse individueller Stärken und Schwächen auf Skalenebene hätte in unserer Stichprobe je nach Altersgruppe und Skala bei 36.4 % bis 67.4 % der Kinder zu einem signifikanten Ergebnis geführt. Ein direkter Vergleich dieser Ergebnisse mit der Normstichprobe war nicht möglich, da entsprechende Daten im Manual der KABC-II nicht berichtet werden. Die ersatzweise betrachtete Häufigkeit von Abweichungen der einzelnen Skalen vom Mittelwert aller Skalen zeigte in den

Altersgruppen keine einheitliche Tendenz. Bei den 5-Jährigen bestanden keine Unterschiede zur Normstichprobe, bei den 6-Jährigen fanden sich in drei von vier Skalen eine höhere Frequenz von Abweichungen, die nach den Vorgaben des Manuals als selten eingeordnet werden können. Ebenfalls bei 5- und 6-Jährigen hatte Renner (2013) bei der K-ABC eine etwas höhere Variabilität in zwei sozialpädiatrischen Stichproben im Vergleich zu einer klinisch unauffälligen Stichprobe berichtet, allerdings bei kleiner Effektstärke. Die im Manual der KABC-II vorgeschlagene Absicherung signifikanter Profildifferenzen durch die Betrachtung der Grundraten sollte in der klinischen Praxis unbedingt genutzt werden. Eine erhöhte Profilvariabilität kann ohne weitere Evidenz nur wenig zur Diagnosestellung beitragen (vgl. Hurks et al. 2013). Ein diagnostisches Vorgehen, bei dem gezielt aus der Fragestellung abgeleitete Hypothesen über individuelle Stärken und Schwächen überprüft werden, sollte daher als Basis einer Profilinterpretation bevorzugt werden.

Die divergente Validität wird durch sehr geringe Zusammenhänge mit Verhaltensvariablen (VBV-EL 3–6) bestätigt. Für die konvergente Validität sprechen die substanziellen Korrelationen mit dem SON-R 2½–7, die sich selbst bei einer zeitverzögerten Testung zeigten. Der sprachfrei erfasste Gesamt-IQ des SON-R 2½–7 korreliert dabei in nahezu identischer Höhe mit dem FKI und dem IVI. Auch die hohe Interkorrelation von FKI und IVI deutet an, dass die im Manual der KABC-II ausführlich dargelegte duale theoretische Perspektive auf Ebene der quantitativen Testergebnisse eher nicht zu bedeutsamen Erkenntnissen führt. Eine intensive Beschäftigung mit der Frage, für welches Modell man sich entscheidet, lohnt vielleicht in Bezug auf den Gesamtwert den Aufwand nicht. In diesem Zusammenhang sei darauf hingewiesen, dass der FKI vermutlich eine bessere Prognose von Schulleistungen bei Kindern mit diversen kulturellen und sprachlichen Hintergründen erlaubt als IVI und SFI (Scheiber und Kaufman 2015).

Die Überprüfung von assoziativen Gedächtnisleistungen kann im Einzelfall eine hohe Relevanz haben. Im Hinblick auf die Interpretation der KABC-II ist jedoch der Befund bemerkenswert, dass sich in unserer klinischen Stichprobe für die Skala *Lernen/Glr* und den Kennwert *Abruf nach Intervall* im Gegensatz zu den übrigen Skalen Werte im Durchschnittsbereich ergaben. Besonders *Atlantis* und *Atlantis Abruf nach Intervall* stellten für die untersuchten Kindern eine relativ geringere Herausforderung dar. In der kleinen Gruppe von Kindern mit Intelligenzminderungen lag *Lernen/Glr* 12.9 IQ-Punkte über dem FKI und erwies sich damit im Gruppenvergleich als weniger sensitiv für globale kognitive Entwicklungsstörungen.

Eine ähnliche Tendenz fand sich in einer Studie zur amerikanischen KABC-II (vgl. Kaufman et al. 2005), bei der die Differenz zwischen *Lernen/Glr* und dem FKI bei Kindern mit Intelligenzminderungen mit 7.9 IQ-Punkten allerdings weniger deutlich ausfiel. Eine Absicherung unseres Befunds an einer weiteren klinischen Stichprobe ist im Hinblick auf die Interpretation der KABC-II sehr zu wünschen, zumal die Skalenbezeichnung *Lernen/Glr* bei Personen, die den theoretischen Hintergrund der KABC-II nicht kennen, einen besonderen Zusammenhang zwischen diesem Intelligenzfaktor und dem Schulerfolg suggerieren könnte.

Kinder mit Migrationshintergrund zeigten bei großer Effektstärke ausschließlich bei *Wissen/Gc* niedrigere Testwerte. Eine Benachteiligung in den weiteren Skalen ist nicht erkennbar, da in diesen keine Unterschiede zu Kindern ohne Migrationshintergrund gefunden wurden. In den beteiligten SPZ wurde die KABC-II allerdings kaum bei Kindern eingesetzt, die erst geringen Kontakt mit der deutschen Sprache hatten, sodass zur Testfairness der KABC-II in dieser Gruppe keine Aussage getroffen werden kann.

Daten zur Kooperation der untersuchten Kinder, die unseres Wissens aus klinischen Anwendungskontexten bisher noch für keinen deutschsprachigen Intelligenztest berichtet wurden, weisen auf eine überwiegend gute Mitarbeit. Allerdings wurden in unseren Analysen nur vollständige Testdurchführungen einbezogen, so dass vorzeitige Testabbrüche nicht berücksichtigt werden konnten.

Bei der Bewertung dieser Ergebnisse müssen einige Limitationen berücksichtigt werden. Die Stichprobe ist stark selektiert. Es wurden nur Kinder berücksichtigt, die an ein SPZ überwiesen wurden und bei denen eine intelligenzdiagnostische Untersuchung als notwendig erachtet wurde. Die Entscheidung für den Einsatz der KABC-II trafen die jeweiligen Untersucher, wobei neben den Besonderheiten des jeweiligen Falles und der diagnostischen Fragestellung auch institutionelle Gepflogenheiten und persönliche Präferenzen eine Rolle gespielt haben können. Die Generalisierbarkeit der Ergebnisse ist daher zurückhaltend zu bewerten.

### Fazit für die Praxis

- Die KABC-II konnte in der klinischen Diagnostik bei Kindern mit Entwicklungsstörungen meist problemlos eingesetzt werden. Die untersuchten Kinder zeigten überwiegend eine gute Mitarbeit.
- Die Testung dauerte im Durchschnitt eine knappe Stunde, wobei sich eine beträchtliche Variabilität zeigte. Kinder mit niedrigeren Gesamtwerten wiesen eine kürzere Testdauer auf als Kinder mit höheren Gesamtwerten.
- Auch in der klinischen Anwendung zeigten die Gesamtskalen und Skalen eine hohe Reliabilität.
- Die KABC-II erwies sich als robust gegenüber von den Eltern beschriebenen Verhaltensmerkmalen wie Aggressivität, Unruhe oder Ängstlichkeit.

- Die Gesamtwerte FKI (CHC-Modell) und IVI (Luria-Modell) korrelierten sehr hoch miteinander und zeigten keinen bedeutsamen Mittelwertsunterschied. Auch hinsichtlich der konvergenten und divergenten Validität zeigen sich sehr ähnliche Ergebnismuster. Möglicherweise ist die Unterscheidung der beiden Gesamtwerte in der Praxis weniger bedeutsam als im Manual dargestellt.
- Die Skala *Lernen/Glr* ergab bei Kindern mit Intelligenzminderungen deutlich höhere Werte als die anderen Skalen. Dies könnte darauf hinweisen, dass *Lernen/Glr* weniger sensitiv für kognitive Entwicklungsstörungen ist.

Insgesamt erwies sich die KABC-II in der sozialpädiatrischen Praxis als gut handhabbares Verfahren, für das wesentliche Gütekriterien bestätigt werden konnten.

#### Dieter Irblich

Sozialpädiatrisches Zentrum  
kreuznacher diakonie  
Auf dem Schmiedel 22  
55469 Simmern

#### Anne Schroeder

Werner Otto Institut Hamburg gGmbH  
Bodelschwinghstr. 23  
22337 Hamburg

#### Gerolf Renner

Pädagogische Hochschule Ludwigsburg  
Fakultät für Sonderpädagogik  
Reuteallee 46  
71634 Ludwigsburg

**Döpfner, M., Berner, W., Fleischmann, T., Schmidt, M. H. (1993):** Verhaltensbeurteilungsbogen für Vorschulkinder (VBV 3 – 6). Beltz Test, Weinheim

**Eckstein-Madry, T., Ahnert, L. (2016):** Kinder aus sozial benachteiligten Familien. Wie Bindungsdefizite und Verhaltensauffälligkeiten durch KiTa-Betreuung beeinflusst werden. *Familiendynamik* 41 (4), 304 – 311, <http://dx.doi.org/10.21706/FD-41-4-304>

**Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., Mascolo, J. T. (2006):** The Achievement Test Desk Reference. 2. Aufl. Wiley, Hoboken/NJ

**Hurks, P. P., Hendriksen, J. G., Dek, J. E., Kooij, A. P. (2013):** Normal variability of children's scaled scores on subtests of the Dutch Wechsler Preschool and Primary scale of Intelligence – third edition. *The Clinical Neuropsychologist* 27 (6), 988 – 1003, <http://dx.doi.org/10.1080/13854046.2013.797502>

**Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., Kaufman, N. L. (2005):** Essentials of KABC-II assessment. Essentials of psychological assessment series. John Wiley & Sons, New York

**Lienert, G. A., Raatz, U. (1994):** Testaufbau und Testanalyse. 5. Aufl. Beltz PVU, Weinheim

**Maher, J. M., Markey, J. C., Ebert-May, D. (2013):** The other half of the story: Effect size analysis in quantitative research. *CBE – Life Sciences Education* 12 (3), 345 – 351, <http://dx.doi.org/10.1187/cbe.13-04-0082>

**Melchers, P., Melchers, M. (2015):** KABC-II. Kaufman Assessment Battery for Children – II von Alan S. Kaufman & Nadeen L. Kaufman. Deutschsprachige Fassung. Pearson, Frankfurt a. M.

**Melchers, P., Preuß, U. (1991):** Kaufman-Assessment Battery for Children – Deutschsprachige Fassung (K-ABC). Swets & Zeitlinger, Amsterdam

**Renner, G. (2013):** Heterogenität von Intelligenztestprofilen im Vorschulalter am Beispiel der „Kaufman-Assessment Battery for Children“ (K-ABC) und des „Nonverbalen Intelligenztests“ (SON-R 2 ½ – 7): Diskrepanzen sind die Regel, nicht die Ausnahme. *Frühförderung interdisziplinär* 32 (3), 169 – 174

**Renner, G., Irblich, D. (2009):** Intelligenzdiagnostik. In: Irblich, D., Renner, G. (Hrsg.): Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre. Göttingen, Hogrefe, 136 – 151

**Renner, G., Schmid, S., Irblich, D., Krampen, G. (2012):** Psychometrische Eigenschaften der „Kaufman-Assessment Battery for Children“ (K-ABC) bei 5- und 6-jährigen Kindern I: Reliabilität und Validität in einer klinischen Stichprobe. *Frühförderung interdisziplinär* 31 (4), 197 – 206, <http://dx.doi.org/10.1080/13854046.2013.797502>

## Literatur

**Dilling, H., Mombour, W., Schmidt, M. H. (Hrsg.) (2015):** Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F) – Klinisch-diagnostische Leitlinien. 10. Aufl. Bern, Verlag Hans Huber

**Döpfner, M., Petermann, F. (2008):** Diagnostik psychischer Störungen im Kindes- und Jugendalter. 2. Aufl. Hogrefe, Göttingen



**Scheiber, C., Kaufman, A. S. (2015):** Which of the three KABC-II global scores is the least biased? *Journal of Pediatric Neuropsychology* 1 (1), 21 – 35, <http://dx.doi.org/10.1007/s40817-015-0004-6>

**Schneider, W. J., McGrew, K. S. (2018):** The Cattell-Horn-Carroll theory of cognitive abilities. In: Flanagan, D. P., McDonough, E. M. (Hrsg.): *Contemporary intellectual assessment. Theories, tests, and issues*. 4. Aufl. The Guilford Press, New York, 73 – 163

**Tellegen, P. J., Laros, J. A., Petermann, F. (2007):** SON-R 2½ – 7. Non-verbaler Intelligenztest. Hogrefe, Göttingen

**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: Flanagan, D. P., McDonough, E. M. (Hrsg.): *Contemporary intellectual assessment. Theories, tests, and issues*. 4. Aufl. New York, The Guilford Press, 333 – 359

**Yousfi, S., Steyer, R. (2006):** Klassische Testtheorie. In: Petermann, F., Eid, M. (Hrsg.): *Handbuch der Psychologischen Diagnostik Göttingen*, Hogrefe, 288 – 303