



Torsten Joél

Das Dilemma der Intelligenzdiagnostik in der Sonderpädagogik – erläutert anhand der neuen KABC-II

Zusammenfassung

Die Anwendung eines Intelligenztests benötigt eine sorgfältige Vorbereitung, testdiagnostische Kenntnisse und vor allem Kenntnisse über die Intelligenztheorien, auf denen der Test aufbaut. Nur so ist gewährleistet, dass Fehlinterpretationen vermieden werden. Doch bereits bei der Testdurchführung könnte diesen Ansprüchen die Realität des sonderpädagogischen Schulalltags entgegenstehen. Schwierigkeiten im Umgang mit Intelligenztests im sonderpädagogischen Kontext werden am Beispiel der inzwischen verbreiteten KABC-II beschrieben, Lösungsmöglichkeiten vorgestellt sowie Kriterien vorgeschlagen, die über die Testgütekriterien hinaus die Nützlichkeit eines Intelligenztests bei Kindern mit sonderpädagogischem Unterstützungsbedarf belegen.

Neben Psychologen führen in Deutschland vielfach Sonderpädagogen Intelligenztests durch. Dies ist sinnvoll, denn bei der Erstellung von Förderplänen, insbesondere aber als Bestandteil bei der Erstellung eines sonderpädagogischen Gutachtens, liefern die Ergebnisse aus Intelligenztests wertvolle Hinweise auf intellektuelle Potenziale, Stärken und Schwächen der Kinder sowie Prognosen über zu erwartende Schwierigkeiten in der weiteren Schullaufbahn. Es ist auch sinnvoll, um so auf dem „kurzen Dienstweg“ ohne bürokratische Hürden zu den Ergebnissen der Testverfahren zu gelangen, denn würden in Deutschland lediglich Psychologen Intelligenztests durchführen, müsste zunächst eine Anmeldung in einer Facheinrichtung erfolgen, Wartezeiten bis zum Testtermin berücksichtigt werden und Absprachen und Gespräche abgewartet werden. Dies wäre ein deutlich erhöhter Aufwand im Vergleich zur direkten Anwendung von Intelligenztests durch Sonderpädagogen.

Weltweit werden Intelligenztests hauptsächlich durch Psychologen durchgeführt, in den USA z. B. in der Regel von zertifizierten Psychologen, die sich für die Anwendung der Intelligenztests besonders qualifizieren müssen. Der Pearson-Verlag erwartet – als einer der renommiertesten Anbieter von Intelligenztests – besondere Qualifikationen von den Psychologen, die in den USA Testmaterial bestellen („Level C“ Qualification), z. B. einen Dokortitel in Psychologie oder besondere Qualifikationsnachweise über das Psychologiestudium hinaus (Pearsonclinical, o. J.). Dies gilt z. B. für die Testverfahren WISC-IV und WPPSI-III, dessen Adaptionen auch in der Sonderpädagogik häufig Anwendung finden. Die hohen Anforderungen an die Nutzer von komplexen Intelligenztests durch Pearson in den USA sollen vor allem die korrekte Interpretation erzielter Testergebnisse sicher stellen. Es gehört zur Stellenbeschreibung von Sonderpädagogen, auch standardisierte normierte Testverfahren anzuwenden, entsprechend gehört zum universitären Curriculum im Rahmen der diagnostischen Ausbildung auch die Testdiagnostik mit Hilfe von normierten Verfahren. In Deutschland führen sowohl Psychologen als auch Sonderpädagogen Intelligenztests durch; beide Berufsgruppen haben die Anwendung von Testverfahren studiert und dementsprechend sind die in Deutschland verwendeten Intelligenztests auch für beide Berufsgruppen gleichermaßen geeignet.

KABC-II im Überblick:

Mit der KABC-II wird die veraltete K-ABC abgelöst. Nicht nur die theoretischen Grundlagen sind erweitert, auch die Komplexität hat sich deutlich erhöht. Während die für einen Intelligenztest typischen Regeln (Umkehrregel, Abbruchregel etc.) bei der K-ABC lediglich einmal gelernt werden mussten und dann auf jeden Subtest übertragbar waren, bestehen für die Subtests der KABC-II vielfältige und untereinander variierende Regeln.

Die KABC-II basiert auf zwei Intelligenzmodellen, dem Cattell-Horn-Carroll (CHC) und dem Luria-Modell, wobei letzteres bereits maßgeblich die theoretische Grundlage der K-ABC bestimmte. Der erweiterte Altersbereich beträgt 3;0 – 18;11 Jahre. Je nach Alter, Fragestellung und Intelligenzmodell kann aus 18 Subtests eine Auswahl vorgenommen werden, welche in drei bis fünf übergeordneten ‚Indices‘ dargestellt werden. Diese geben Hinweise auf Ressourcen und Schwächen innerhalb dieses mehrdimensionalen Tests. Alle Subtests und Indices münden in einem Gesamtwert, vergleichbar mit dem Generalfaktor der Intelligenz in hierarchischen Intelligenztheorien.

Die Subtests sind kindgerecht, das Bildmaterial ist im Gegensatz zur alten K-ABC ausgezeichnet und motivierend, die Konzeptionierung wird in Rezensionen positiv bewertet. Oft war die alte K-ABC in der Sonderpädagogik ein „Mittel der Wahl“, vor allem bei Kindern mit dem Förderbedarf Geistige Entwicklung und Lernen, da die Aufgaben einfach zu erläutern waren und viele einfache Items für kognitiv schwache Kinder zur Verfügung standen. Diese sonderpädagogische Relevanz bleibt auch bei der KABC-II erhalten und wird dadurch ergänzt, dass nun auch ältere und begabte Kinder getestet werden können.

Abbildung 1:
KABC-II im Überblick

Intelligenztests werden in den verschiedenen Regionen Deutschlands unterschiedlich von Sonderpädagogen genutzt. Dass Bildung Ländersache ist, soll hier nicht weiter ausgeführt werden. Doch selbst innerhalb der Bundesländer wird regional verschieden mit Intelligenztests gearbeitet, die Bandbreite reicht von der Ablehnung der „Statusdiagnostik“ bis zum Irrglauben daran, dass eine einzige Zahl („Intelligenzquotient“) die Intelligenz eines Kinds erfassen kann und sonderpädagogischen Unterstützungsbedarf nachweisen könnte. Im Folgenden sollen also typische Herangehensweisen skizziert werden, die sich auf Grund der großen Heterogenität im Bildungsbereich von selbst gemachten Erfahrungen unterscheiden können.

Die Anwendung von Intelligenztests im sonderpädagogischen Kontext ist im realen Arbeitsfeld nur ein kleiner Ausschnitt der vielfältigen Tätigkeiten. Häufig werden Intelligenztests angewendet, wenn ein sonderpädagogisches Gutachten angefertigt wird. Übliche Praxis ist die Verteilung aller anzufertigenden sonderpädagogischen Gutachten auf alle Kollegen eines Kollegiums, z. B. eines Förderzentrums. Daraus resultiert, dass ein Sonderpädagoge einige wenige Gutachten bearbeitet und entsprechend selten Intelligenztests durchführt. Nach dem Studium der Sonderpädagogik sollte jeder Sonderpädagoge in der Lage sein, jeden normierten Test durchführen und interpretieren zu können. Doch geht dies an der Realität vorbei. Gerade die komplexen Testverfahren WISC-IV (Petermann & Petermann, 2011), WPPSI-III (Petermann, 2009), IDS (Grob, Meyer & Hagmann von Arx, 2009) oder die inzwischen verbreitete KABC-II (Kaufman & Kaufman, 2015) benötigen sehr viel Vorbereitung, da eine Vielzahl von Regeln beachtet werden muss. Werden z. B. lediglich zwei, drei oder vier Tests jährlich durchgeführt, ist es unrealistisch zu erwarten, dass man sich an mehrere hundert Regeln bei einer Anwendung alle vier Monate im Jahr erinnert. Erschwerend kommt hinzu, dass in vielen Regionen die Gutachten innerhalb eines begrenzten Zeitraums fertiggestellt sein müssen, z. B. innerhalb von sechs Wochen. Daraus resultiert, dass die Pausen zwischen den Testanwendungen zehn Monate betragen können. Ohne die erneute Beschäftigung mit den Testverfahren ist es unvorstellbar, den Test einzig aus der Erinnerung durchzuführen. Ein Blick auf die Anzahl der Regeln der KABC-II soll dies belegen: Es sind ca. 580 Regeln und verpflichtende Hinweise, die bei der Durchführung beachtet werden müssen. In diese 580 Regeln sind nicht einbezogen „grundsätzliche Vorgehensweisen“ (Kaufman & Kaufman, S. 249 ff.) während der Testsituation sowie allgemeine Hinweise oder Wiederholungen. In diese Zählung der zu beachtenden Regeln sind Anweisungen an die Kinder

Das Dilemma

Achillessehne Durchführungs- objektivität

ebenfalls nicht aufgeführt und auch nicht alle Regeln, die im Kapitel 4 die Protokollierung, Bewertung und das Ausfüllen des Testbogens beschreiben. Die ca. 580 gezählten Regeln setzen sich zusammen aus den konkreten Anweisungen der Subtests, die während der Testdurchführung den vier Stellordnern entnommen werden müssen und den korrekten Ablauf gewährleisten sollen und den konkreten Regeln des Kapitels Durchführung (ebd., S. 73 ff.). Durchschnittlich ca. 26 Regeln werden in den Stellordnern für jeden Subtest beschrieben, die wenigsten bei dem Subtest ‚Konzeptbildung‘ (8 gezählte Regeln), die meisten bei den Subtests ‚Rover‘ (mehr als 65 Regeln) und ‚Dreiecke‘ (mehr als 40 Regeln, zusätzlich ca. 30 spezielle Bewertungsregeln). Bei einer seltenen Anwendung eines dermaßen komplexen Intelligenztests wie der KABC-II kann ohne „eigene Hochbegabung“ und ein fotografisches Gedächtnis ausgeschlossen werden, dass ohne Vorbereitung (wozu auch Probetestungen gehören) alle Regeln erinnert werden. Doch was bedeutet dies konkret:

- Ein Arbeitgeber bzw. Auftraggeber, z. B. die Schulleitung, stellt mehrere Stunden Arbeitszeit für die Vorbereitung zur Verfügung, damit der Test korrekt durchgeführt werden kann.
- Der Arbeitgeber stellt die Arbeitszeit nicht zur Verfügung und es findet eine mehrstündige Vorbereitung in der Freizeit statt.
- Der Test wird ohne angemessene Vorbereitung durchgeführt.
- Der Test wird nicht durchgeführt.

Jede dieser Ausgänge ist unwahrscheinlich bzw. unbefriedigend, ein Umdenken im Umgang mit den immer komplexeren Intelligenztests in der Sonderpädagogik scheint notwendig.

Wird ein nach Deutschland geflüchtetes syrisches Kind getestet, wird jedes Item, bei dem die deutsche Sprache eine wichtige Rolle spielt, genau so auf Deutsch vorgetragen, wie es im Handbuch steht und es wird kein Wort erläutert, auch wenn das Kind kein Deutsch spricht. Kann ein Kind auf Grund einer Spastik die Hände nicht gut verwenden und es soll nach einer Vorlage unter Zeitdruck in 90 Sekunden ein Gebilde mit dreieckigen Schaumstoffteilen nachbauen, wird die Zeit nicht aus Gründen der Fairness verlängert oder ganz weggelassen im Sinne eines „Nachteilsausgleichs“. Obwohl diese Beispiele absurd klingen, sind sie unter dem Aspekt der Durchführungsobjektivität inhaltlich richtig.

Selbstverständlich werden Intelligenztests fachgerecht durchgeführt, doch soll dieser Allgemeinplatz bezüglich der Durchführungsobjektivität einmal genauer unter die Lupe genommen werden.

Wird ein Kind getestet, werden die daraus resultierenden Testergebnisse mit einem repräsentativen Durchschnitt altersgleicher Kinder abgeglichen. Hilfreich ist hierzu die Annahme, dass ein Konstrukt wie Intelligenz innerhalb einer Population, z. B. innerhalb der deutschen Kinder, normalverteilt ist. Diese Normalverteilung wird mittels der berühmten Glockenkurve nach Gauß dargestellt. Die meisten Menschen befinden sich im mittleren durchschnittlichen Normbereich (ca. 68%), Kinder mit Lernschwächen (ca. 14%) oder Intelligenzminderungen (ca. 2%) im unter-, bzw. weit unterdurchschnittlichen Bereich. Der repräsentative Durchschnitt altersgleicher Kinder wird ermittelt über die Normstichprobe: So wurden z. B. 1745 Kinder für die Normierung mit der KABC-II getestet (Kaufman & Kaufman, 2015, S. 162 ff.). Aus dem Abgleich der Testergebnisse eines getesteten Kindes mit der Normstichprobe resultiert zwingend, dass das getestete Kind exakt so getestet wurde, wie die Kinder aus der Normstichprobe getestet worden sind. Dies bedeutet, dass nicht eine der ca. 580 Regeln außer Acht gelassen werden darf, dass nicht an einer einzigen Stelle die Zeiten geändert werden, die dem Kind zur Bearbeitung zur Verfügung steht, dass kein Punkt gegeben wird, wenn das Kind auch nur eine Sekunde die zur Verfügung stehende Zeit für die Bearbeitung eines Items überschreitet und dann auch, dass ein Kind mit einer Spastik nicht zusätzliche Zeit erhielt. Dies wäre ansonsten Willkür – man hätte aus dem normierten Test einen informellen Test gemacht.

Das eingangs erwähnte nicht Deutsch sprechende syrische Kind würde natürlich mit einem non-verbalen Test getestet werden (z. B. SON-R 6-40, Tellegen, Laros & Petermann, 2012; WNV,

Petermann, 2014) und das Kind mit einer Spastik würde mit Subtests getestet werden, bei denen die Stoppuhr keine Rolle spielt. Es würde also ein passender Test für die Kinder gesucht, nicht der Test passend gemacht. Eigenmächtige Modifikationen der Durchführungsregeln müssten als Kunstfehler bezeichnet werden.

Sowohl Forschungsbelege als auch die Betrachtung des Umgangs mit den immer komplexer werdenden Intelligenztests in der sonderpädagogischen Praxis stimmen nicht optimistisch bezüglich der Wahrung der Durchführungsobjektivität:

- Praxisalltag ist die seltene Anwendung von Intelligenztests, deren Komplexität aber eine häufige Anwendung verlangt. Bei einer seltenen Anwendung wäre man während der Testung vor allem damit beschäftigt, den Test korrekt mit allen seinen hunderten Regeln durchzuführen, unter Einbezug der Stoppuhr, der Umkehr-, Abbruch- und Anfangsregeln, der richtigen Handhabung des Formulars, der Materialien usw. Die Gefahr bestünde, dass man mehr mit dem Test beschäftigt wäre als mit dem getesteten Kind.
- Wenige Studien haben die Auswirkungen von fehlerhaft durchgeführten Intelligenztests und die Anzahl gemachter Fehler untersucht. Alfonso, Johnson, Patinella und Rader (1998) analysierten 60 Testprotokolle. Auch nach intensiver Einarbeitung von Studierenden in den damals aktuellen WISC-III betrug die Fehlerquote 100 %, es wurden insgesamt 468 Fehler gefunden. In einem Vergleich mehrerer Studien stellten Lipsius, Petermann und Daseking (2008) gar fest, dass erfahrene Testleiter anfälliger für Fehler sind. Bei der Anwendung des HAWIK-IV (jetzt WISC-IV) erkannten Lipsius und Kollegen (2008) in einer eigenen Untersuchung zur Durchführungsobjektivität vielfältige Fehler, z. B. beim Abbruchkriterium, beim Weglassen von Nachfragen, wo diese vorgeschrieben sind, bei der Umkehrregel (beginnt ein älteres Kind nicht mit Aufgabe 1 eines Subtests, sondern mit einer späteren und altersgerechteren Aufgabe, muss es zu den einfacheren Aufgaben „umkehren“, sofern die altersgerechte Anfangsaufgabe nicht oder falsch gelöst wird) und auch beim Umgang mit den Zeitvorgaben. Lipsius und Kollegen (2008) beschreiben Beispiele von falsch berechneten Gesamtwerten bis zu einer Standardabweichung. Huber (2000) befragte 313 Lehrer aus 14 Förderschulen mit dem Schwerpunkt Körperlich-motorische Entwicklung nach dem Umgang mit Intelligenztests und stellte fest, dass die Durchführungsobjektivität nicht nur verletzt, sondern sogar mit Absicht verletzt worden ist. Lediglich 2,4 % der Lehrkräfte gaben die Einhaltung der standardisierten Vorgaben an. Sogar leichte Hinweise zum Lösungsweg (ebd., S. 412) wurden gegeben, Verletzungen der Durchführungsobjektivität schienen eher die Regel als die Ausnahme (ebd., S. 412) zu sein. Huber (2000, S. 413) resümiert, dass „zumindest ein Teil der quantitativen Testergebnisse [...] durch einen fragwürdigen Umgang mit dem Testmaterial aussagelos ist“.

Die Qualität eines Intelligenztests wird mit den Hauptgütekriterien belegt. Über diese Kriterien hinaus sind in der praktischen Anwendung eines Intelligenztests weitere Aspekte von Bedeutung, die sowohl die Arbeitsbedingungen der Sonderpädagogen, aber vor allem die Besonderheiten von Kindern mit sonderpädagogischem Unterstützungsbedarf berücksichtigen.

Im Folgenden sollen diese Aspekte, die nicht ausreichend mit den bekannten Haupt- und Nebengütekriterien beschrieben sind, mit einem Fokus auf die KABC-II skizziert werden.

Aspekt ‚Itemdichte‘

Eine große Anzahl an Items in einem Subtest erlaubt differenziertere Aussagen bei Kindern mit sonderpädagogischem Unterstützungsbedarf, da ansonsten das Anforderungsniveau zu schnell steigt. Bei einem Subtest, der z. B. lediglich aus sieben Items besteht, ist die „Treppe zu steil“; kognitiv schwache Kinder fallen schnell heraus und differenzierte Aussagen sind kaum möglich. Eine sehr geringe Itemdichte bietet z. B. die IDS (Grob et al., 2009), weswegen dieser Test für den Bereich Geistige Entwicklung nicht geeignet ist.

KABC-II: geeignet oder ungeeignet für die Sonderpädagogik?

Im Durchschnitt stellt die KABC-II in den 18 Subtests etwas mehr als 31 Items zur Verfügung. Eine große Anzahl an Items in einem Subtest stellt in der Regel auch sicher, dass nicht wenige

Veränderungen im Rohwert zu rasanten Sprüngen in den Normwerten führen, was zu Interpretationsverzerrungen führen könnte, da entsprechende Sprünge in den Testergebnissen dann eher auf eine schwache Testkonstruktion als auf Stärken oder Schwächen der Kinder zurückgeführt werden könnten.

Aspekt ‚große Altersbandbreite‘

Für kognitiv schwache Kinder sollten Intelligenztests Anwendung finden, mit denen auch deutlich jüngere Kinder getestet werden könnten. So sollte ein kognitiv schwaches siebenjähriges Kind mit einem auch für sehr viel jüngere Kinder geeignetem Verfahren überprüft werden. Damit wäre gewährleistet, dass für die kognitiv schwachen Kinder die einfachen Items der jüngeren Kinder zur Verfügung stehen, z. B. nach Anwendung der Umkehrregel. Ansonsten bestünde die Gefahr, dass die Items zu schnell zu schwer werden und differenzierte Aussagen kaum möglich sind. Um Boden- und Deckeneffekte zu vermeiden, sollten in der Sonderpädagogik Intelligenztests nicht auf wenige Altersgruppen beschränkt sein, tendenziell sollten sie vor allem aber sehr junge Kinder testen können. Die KABC-II hat eine Altersbandbreite von 3;0–18;11 Jahren, somit können sowohl kognitiv schwache als starke Kinder und Jugendliche angemessen getestet werden. Als Nebeneffekt ist es für das schmale Budget vieler Schulen von Vorteil, wenn ein Test alle Altersbandbreiten abdeckt, da nicht viele teure Tests angeschafft werden müssen.

Aspekt ‚große Normstichprobe‘

Eine repräsentative und umfangreiche Normstichprobe vermindert die Gefahr von Verzerrungen beim Abgleich der vom Kind erzielten Rohwerte mit den Normdaten. Die absolute Größe der Normstichprobe ist dabei nicht ausschlaggebend, denn sie muss im Zusammenhang mit der Anzahl der Normtabellen gesehen werden, die meist in Vierteljahresschritten unterteilt sind. Positiv hervorgehoben werden kann der CFT 1-R (Weiss & Osterland, 2012), dessen 4700 Kinder der Normstichprobe sich auf 16 Altersnormen verteilen, durchschnittlich also 294 Kinder je Altersnorm. Dies ist von Vorteil, da Kinder mit Unterstützungsbedarf oft am „Rand“ der Gauß'schen Glockenkurve getestet werden.

Um Boden- und Deckeneffekte zu vermeiden, sollten in der Sonderpädagogik Intelligenztests nicht auf wenige Altersgruppen beschränkt sein, tendenziell sollten sie vor allem aber sehr junge Kinder testen können.

Für die KABC-II wurden 1745 Kinder für die Normierungsphase getestet. Aufgeteilt in 47 Altersgruppen ergibt dies eine durchschnittliche Anzahl von 37 Kindern je Alterstabelle. Entsprechend der Verteilung nach Gauß müssten – statistisch betrachtet – fünf von diesen 37 Kindern eine Lernschwäche und hochgerechnet ein Kind eine Intelligenzminderung haben. Auch ohne Darstellung der komplexen statistischen

Abläufe, die zur Erstellung einer Normtabelle führen, kann die Normstichprobe der KABC-II für den sonderpädagogischen Einsatz als schwaches Merkmal bezeichnet werden. Erschwerend kommt hinzu, dass für einige Ergänzungsuntertests lediglich eine reduzierte Normstichprobe vorliegt (Kaufman & Kaufman, 2015, S. 32).

Aspekt ‚aktuelle Normstichprobe‘

Vor allem auf Grund der kontinuierlichen Zunahme der Intelligenz („Flynn-Effekt“) entsteht die Notwendigkeit, Intelligenztests mit aktuellen Normdaten zu verwenden. Aus der postulierten Zunahme der Intelligenz um etwa 3 IQ-Punkte pro Jahrzehnt (Flynn, 1987) resultiert die gedankliche Berücksichtigung von Gesamtergebnissen, die mit veralteten Tests erzielt worden sind. Der in der Praxis noch angewendete SON-R 5 ½–17 (Tellegen & Larros, 2005) wurde vor über 30 Jahren normiert, durch den Abgleich von Testergebnissen mit einer veralteten Normstichprobe resultieren unter Berücksichtigung des Flynn-Effekts deutlich erhöhte und somit verzerrende Testergebnisse.

Auch wenn der Flynn-Effekt zumindest in der Computerauswertung des SON-R 5 ½–17 berücksichtigt wird, bleibt die Nutzung veralteter Normen fragwürdig. Die Verwendung veralteter Intel-

Intelligenztests ist verbreitet und hat auch mit den hohen Anschaffungskosten zu tun. Huber (2000, S. 413) stellt fest, dass nicht nur in Ausnahmen über vierzig Jahre alte Intelligenztests in Förderschulen angewendet worden sind.

Die Normierung der KABC-II fand 2013/2014 statt, die Berücksichtigung demografischer Faktoren ist belegt. Eine aktuellere Normstichprobe ist derzeit nicht denkbar.

Aspekt ‚Praktikabilität‘

Im sonderpädagogischen Kontext ist es üblich, Intelligenztests eher selten als häufig anzuwenden. Deshalb wäre es nützlich, wenn die Verfahren zügig gelernt und dann zügig wiedererlernt werden können. Ein gutes Beispiel für eine kluge und einfache Konstruktion bietet der SON R 6–40. Die Regeln ähneln sich in jedem Subtest und das adaptive Durchführungssystem verhindert komplizierte und damit fehleranfällige Regeln wie die Umkehrregel.

Es muss festgestellt werden, dass die KABC-II aus einer Regelflut besteht. Jeder Subtest ist anders gestaltet und selbst bei häufigerer Anwendung ist die Klage über verwirrend empfundene Erläuterungen bei den Anwendern groß. Bereits weiter oben wird beschrieben, wie (über-)komplex die KABC-II gestaltet ist. Ursprünglich erschien dieser, von dem Ehepaar Kaufman konzipierte Test im Jahr 2004 in den USA und gilt dort als „Level C“-Test – vorgesehen also für besonders befähigte bzw. zertifizierte Psychologen. Die relativ nah am Original adaptierte deutsche Version der für häufig testende Psychologen konzipierten KABC-II berücksichtigt allerdings ungenügend die Rahmenbedingungen, unter denen Sonderpädagogen normierte Intelligenztests anwenden.

Aspekt ‚Computerauswertung‘

Die Computerauswertung der KABC-II ist sehr übersichtlich und klug programmiert, die Ergebnisse werden zunächst entsprechend des Layouts des Formular-Deckblatts dargestellt. Neben statistisch signifikanten Stärken- und Schwächen-Analysen werden die Testergebnisse zudem grafisch erstellt. Die Betonung des Konfidenzintervalls durch farblich markierte Bänder wirkt der oft zu beobachtenden Simplifizierung der Gesamtergebnisse durch einen „Gesamt IQ“ entgegen. Am Ende gibt ein Text Interpretationshinweise. Hilfreich sind eindeutig rot markierte Hinweise, die vor Überinterpretationen warnen, z. B. auf Grund fehlender Werte.

Aspekt ‚Freies und intuitives Erklären‘

Kognitiv schwache Kinder benötigen eine eindeutige und schnell zu verstehende Subtesteinweisung. Günstig sind Subtests, die vom Kind intuitiv verstanden werden können. Bei Kindern mit Unterstützungsbedarf sind bei der Subtesteinweisung zwei Aspekte sinnvoll: erstens die Möglichkeit, die Anforderungen mit eigenen Worten dem Kind „frei“ erläutern zu können und zweitens nicht bewertete Einführungsaufgaben. Bewertete Einführungsaufgaben würden wenig Spielraum lassen bei der Erläuterung, da sich ein bewertetes (Einführungs-)Item zur Wahrung der Durchführungsobjektivität streng an die Vorgaben im Handbuch halten muss.

Der CFT 20-R (Weiss, 2006) ist ein schwaches Beispiel bezüglich dieser Aspekte: Das Manual schreibt eine wortgetreue Wiedergabe der im Handbuch abzulesenden Anweisungen vor und im Subtest ‚Topologien‘ (die Position eines kleinen schwarzen Punkts soll einer Auswahl unter bestimmten Vorgaben zugeordnet werden) fällt es kognitiv schwachen Kindern schwer, die komplizierten Anweisungen zu verstehen.

Die K-ABC (Melchers & Preuss, 1991) war einer der wenigen Tests, der vor allem auf Grund seiner intuitiven und sprachlich wenig einschränkenden Vermittlung oft komplett mit intelligenzgeminderten Kindern durchgeführt werden konnte. Die KABC-II hat die positive Möglichkeit der freien Erläuterung beibehalten, die Subtests bleiben auch überwiegend intuitiv nachvollziehbar für die Kinder. Leider sind in einigen Subtests die Lernaufgaben bereits bewertet und der Subtest ‚Symbole‘ ist gelegentlich schwer zu vermitteln.

*Abb. 2:
Bewertung der KABC-II nach
praktischen Erwägungen
(von sehr positiv ++
bis sehr negativ --)*

Itemdichte	++
Altersbandbreite	++
große Normstichprobe	–
aktuelle Normstichprobe	++
Praktikabilität	--
Computerauswertung	++
Freies und intuitives Erklären	+
motivierende Stimuli	+
Ableitungsmöglichkeiten	++

Alle Urheberrechte liegen beim Verband Sonderpädagogik e. V. – Veröffentlichung und Wiedergabe sind nur mit Genehmigung des Rechteinhabers gestattet.

Aspekt ‚motivierende Stimuli‘

Kinder mit sonderpädagogischem Unterstützungsbedarf haben oft unlustvolle Erfahrungen in Leistungssituationen erlebt, da sie – subjektiv empfunden und objektiv nachgewiesen – oft weniger können als gleichaltrige Kinder. Kognitiv schwache Kinder wachsen in einer leistungsorientierten, kapitalistischen Gesellschaft auf und entwickeln ein Gefühl dafür, dass sie eigentlich den Anforderungen der Gesellschaft nicht gerecht werden. Die Testsituation ist eine „privilegierte“ Situation: Es besteht eine Eins-zu-eins-Zuwendung, was Kinder meist als angenehm empfinden, es besteht eine wohlwollende und zugewandte Haltung des Testleiters dem Kind gegenüber und es wird das grundsätzliche Bedürfnis von Menschen nach kognitiver Stimulation

Gerade im pädagogischen Bereich wäre eine mögliche Ableitung von individuellen Stärken und Schwächen neben dem Vergleich mit einer Altersgruppe wünschenswert.

erfüllt. Kinder mögen es, spannende Rätsel zu lösen. Die Bedingungen der Testsituation sind günstig, die gesellschaftlichen Rahmenbedingungen erschweren jedoch das Testen, denn die Testsituation ist auch eine Leistungssituation und Kinder assoziieren zuweilen mit Leistungssituationen negative Erfahrungen. Ein guter Intelligenztest bietet motivierende Aufgaben und macht dem Kind im besten Fall Spaß, die negativ empfundene Leistungssituation tritt in den Hintergrund. Ein besonders gelungenes Beispiel ist hierfür die IDS, die die Subtests spielerisch und ausgesprochen abwechslungsreich gestaltet hat.

Die KABC-II bietet ebenfalls überwiegend motivierende Stimuli. Die Kinder sind an den Items interessiert, sie sind grafisch ansprechend gestaltet und Frustrationserlebnisse auf Grund ungünstiger Abbruchregeln werden vermieden, es werden sogar praktische Hinweise zur Erhaltung der Kooperation des Kinds im Manual vorgestellt. An dieser Stelle sei kurz erwähnt, dass die eingangs erwähnte Forderung nach rigidem Einhalten aller Regeln ohne jedwede Abweichung vom Manual selbstverständlich nicht bedeutet, dass ein Item an das andere gereiht und der Test „streng“ durchgezogen wird. Es wird auf Seiten des Testleiters mit Lebensfreude in die Testsituation gegangen, Humor findet Anwendung, es werden Scherze ausgetauscht, das Kind wird nach seinen Hobbys gefragt und gerne darf das Kind mit der Stoppuhr auch mal den Tester mit vom Kind ausgedachten Aufgaben testen.

Aspekt ‚Ableitung von Stärken und Schwächen‘

Gerade im pädagogischen Bereich wäre eine mögliche Ableitung von individuellen Stärken und Schwächen neben dem Vergleich mit einer Altersgruppe wünschenswert. Eher eindimensionale Intelligenztests wie der CFT 1-R, der CFT 20R, der SON-R 6-40 oder die WNV sind über die Ermittlung eines Gesamtergebnisses hinaus dazu nicht geeignet.

Es darf aber auch von einem komplexeren Intelligenztest, sei er noch so gut, nicht erwartet werden, dass sich darauf aufbauend umfangreiche Fördermaßnahmen ableiten ließen. Wäre dies der Anspruch, bestünde die Gefahr einer Überinterpretation vorliegender Testergebnisse. Testbatterien wie die WISC-IV, die IDS, die WPPSI-III und die KABC-II sind mit ihren mehrdimensionalen Ansätzen aber durchaus in der Lage, brauchbare Ableitungen von Ressourcen und Defiziten zu ermitteln. Hier sticht die KABC-II noch einmal positiv hervor, denn die theoretische Konzeptionierung (CHC-Modell bzw. Luria-Ansatz; Kaufman & Kaufman, 2015, S. 43 ff.) ist vorbildlich. Besonders das Intelligenzmodell nach Luria (1966; 1970) ist unter Ausblendung des Abfragens von Faktenwissen („kristalline Intelligenz“) für den sonderpädagogischen Kontext sowohl zur Erfassung des intellektuellen Potentials, als auch zur Ableitung von Stärken und Schwächen des getesteten Kinds interessant. Sogar nicht sprechende oder nicht deutsch sprechende Kinder könnten nonverbal mit dem Sprachfrei Index (SFI) getestet werden.

Fazit

Die theoretische Begründung im Sinne der Testtheorie mit der dazugehörigen Prüfung der Testgütekriterien steht in diesem Aufsatz nicht im Zentrum. Eine ausführliche Testvorstellung sind den Rezensionen von Renner (2015) und Irblich (2015) zu entnehmen. Kurz erwähnt sei, dass auch aus testtheoretischer Sicht die KABC-II als gelungen bewertet werden darf und sowohl

Renner (2015), als auch Irblich (2015) sich überwiegend positiv äußern. Die wichtigsten negativ kritischen Anmerkungen von Renner (2015) und Irblich (2015) sind in Tabelle 2 dargestellt.

Im Zentrum dieses Artikels steht auch nicht die angemessene Ableitung von Interpretationen aus vorliegenden Testergebnissen, die zwingend eine intensive Beschäftigung mit den zwei Intelligenzmodellen voraussetzt. In diesem Zusammenhang ist ein kritisches Hinterfragen der Überführung von Lurias Ansätzen in einen Intelligenztest lohnend, da bei näherer Betrachtung dies eigentlich nicht möglich scheint und eine Überführung seiner Theorien in einen normierten Test nicht vorgesehen war (Luria, 1966; 1970). Lohnenswert ist zudem eine kritische Betrachtung der postulierten Symbiose von zwei sich auf dem ersten Blick wenig ähnelnden Intelligenzansätzen, dem Cattell-Horn-Carroll Modell (CHC-Modell; Schneider & McGrew, 2012) und dem Ansatz nach Luria in vier Indices mit identischen Subtests.

Im Vordergrund dieses Artikels stehen die Durchführungsbedingungen und praktische Überlegungen. Die Anwendung eines Intelligenztests besteht aus der Abfolge Durchführung, Auswertung und Interpretation. Ist der erste Schritt, die Durchführung, nicht nach den Regeln der Kunst vorgenommen worden, wird es in den folgenden Schritten zu Fehlern kommen, da die Auswertung und die Interpretation auf der Durchführung basieren.

Es stellt sich also unter Berücksichtigung der Aspekte, die bei der praktischen Anwendung eines Intelligenztests beachtet werden sollten die Frage, ob die KABC-II für die Sonderpädagogik geeignet ist und die Durchführung und dementsprechend die Durchführungsobjektivität unter den derzeit aktuellen Bedingungen, unter denen Tests häufig in der Sonderpädagogik durchgeführt werden, gewahrt sind.

Ja, die KABC-II ist geeignet für den Einsatz für Kinder mit sonderpädagogischem Unterstützungsbedarf. Es gibt kaum einen Test, der so vielschichtige Hinweise gibt, so ausgezeichnet konzeptioniert wurde, der alle Altersbereiche abdeckt, der weitgehend intuitiv von den Kindern verstanden wird, der sogar in Teilen nonverbal durchgeführt werden könnte, der den Kindern auch noch Spaß macht und der die Besonderheiten von Kindern mit Unterstützungsbedarfen angemessen berücksichtigt.

Nein, die KABC-II ist nicht geeignet für den Einsatz in der Sonderpädagogik. Sie ist so überkomplex, dass eine seltene Anwendung ziemlich sicher zu Durchführungsfehlern, im schlimmeren Fall sogar zu Interpretationsfehlern führen muss, und ist somit bei einer für den sonderpädagogischen Kontext tendenziell seltenen Anwendung ungeeignet.

Die Kriterien für ein Dilemma sind erfüllt, denn es gibt zwei unbefriedigende Lösungen für eine Fragestellung. Doch sind Lösungen denkbar, die sowohl die Anwendung der KABC-II ermöglichen als auch die Durchführungsobjektivität gewährleisten.

Es ist kaum vorstellbar, dass bei einer seltenen Anwendung der KABC-II viele hundert Regeln erinnert werden können, aber auch kaum vorstellbar, dass vor jeder Anwendung eine ausführliche und mehrstündige Vorbereitung stattfindet.

- keine Angaben zur nonverbalen Anleitung der Subtests des Sprachfrei-Index (Renner, 2015; Irblich, 2015)
- Übungseffekte bei zu rascher Testwiederholung (Renner, 2015; Irblich, 2015)
- keine Hinweise zum Normierungsvorgehen (Renner, 2015; Irblich, 2015)
- Kinder aus Familien mit niedrigem Bildungsabschluss in der Normierungsstichprobe unterrepräsentiert (Renner, 2015; Irblich, 2015)
- leichte Bodeneffekte in einigen Subtests, vor allem bei jungen Kindern (Renner, 2015; Irblich, 2015)
- leichte Deckeneffekte in Ergänzungssubtests für Kinder ab 13 Jahren (Irblich, 2015)
- nicht geeignet für den Förderschwerpunkt Sehen (Renner, 2015; Irblich, 2015)
- in Einzelfällen nicht nachvollziehbare Lösungsvorgaben (Renner, 2015)
- trotz Anweisungen, sich nicht zu beeilen, kann schnelles Arbeiten das Ergebnis erheblich beeinflussen (Renner, 2015; Irblich, 2015)
- Konstruktionsmängel beim Subtest ‚Symbole‘ (Renner, 2015; Irblich, 2015)
- deutlich mangelhafte Materialqualität (Renner, 2015; Irblich, 2015)
- Reliabilitätsberechnung auf Untertestebene unklar beschrieben (Irblich, 2015)
- Indice ‚Planung/Gf‘ eher aus statistischen als aus inhaltlichen Gründen zusammengestellt (Irblich, 2015)
- Mangelnde Erläuterungen von Faktorenanalysen, dadurch sind Rückschlüsse auf Ladungen schwer möglich (Irblich, 2015)
- Fehlende Angaben zur Kriteriumsvalidität und prognostischen Validität (Irblich, 2015)
- Überflüssige Hintergrundinformationen auf den Instruktionseiten (Irblich, 2015)
- Indice ‚Wissen/Gc‘ bildet kristalline Intelligenz nur bedingt ab (Irblich, 2015)
- Postulierte Kulturfairness für den Gesamttest fragwürdig (Irblich, 2015)

*Abb. 3:
Eine Auswahl negativ kritischer
Anmerkungen aus unabhängig
voneinander publizierten
Testrezensionen zur KABC-II
von Renner (2015)
und Irblich (2015)*

Ausblick

Schlüsselwörter

KABC-II, Durchführungsobjektivität, Intelligenztest, Intelligenzdiagnostik

Abstract

The use of an intelligence test requires careful preparation, test diagnostic skills, and in particular knowledge of the theories of intelligence that underpins the test.

Only this guarantees that misinterpretations are avoided. Yet these claims could preclude the reality of special needs school days as soon as during the execution of the test.

Difficulties in dealing with intelligence tests in the special education context are described using the example of the now widespread KABC-II, presenting possible solutions and proposing criteria proving the usefulness of an intelligence test for children with special educational needs beyond the actual quality criteria of the test.

Keywords

KABC-II, examiner errors, assessment of intelligence, objectivity

Es bietet sich an, dass nicht alle Sonderpädagogen selten, sondern wenige Sonderpädagogen häufig testen.

Dies könnte vordergründig dem Anspruch widersprechen, dass psychologische Tätigkeiten zur Stellenbeschreibung von Sonderpädagogen gehören. Dieser Anspruch wäre allerdings nicht gefährdet bei einer Spezialisierung weniger Sonderpädagogen auf mehrdimensionale Intelligenztests (WISC-IV, WPPSI-III, IDS, KABC-II). Die Interpretation vorliegender und evtl. „zugearbeiteter“ Testergebnisse sowie weitere Bausteine der sonderpädagogischen Diagnostik (Beobachtung, Elterngespräche etc.) blieben weiterhin psychologische Tätigkeiten.

Es gibt bereits Bestrebungen in diese Richtungen, sowohl auf schulischer, als auch auf Länderebene. Einige Schulen bzw. Schulzentren qualifizieren z. B. wenige Sonderpädagogen für spezielle Testverfahren. Es wäre z. B. möglich, zwei Kollegen von der Verteilung der sonderpädagogischen Gutachten auszunehmen. Diese führen keine Gutachten mehr durch, dafür wird für die Gutachten der anderen Lehrkräfte die KABC-II durchgeführt.

Möglich wäre auch die Bildung eines Diagnostikteams: Lediglich wenige Sonderpädagogen führen alle Gutachten durch, müssen dafür natürlich

entsprechend Zeit zur Verfügung gestellt bekommen. Es sind also strukturell-organisatorische Fragestellungen, die im Sinne einer angemessenen Testdurchführung zu diskutieren sind. Aus zeitökonomischer Sicht ist dies sinnvoll, denn letztlich benötigen Testspezialisten weniger Vorbereitungszeit und diagnostische Kompetenzen wären gebündelt. Was der Testspezialist an Zeit für die Testungen „verliert“, gewinnen die Kollegen, die nicht mehr testen. Eine Freistellung des Testspezialisten würde aufgefangen werden von den Kollegen, die Zeit gewinnen, da sie nicht mehr testen. Müller (2009) kommt zu dem Schluss, dass die aus einer seltenen Anwendung von Intelligenztests resultierenden Schwierigkeiten zu einem ‚schulinternen Diagnostikansprechpartner‘ führen könnten, der nicht nur für die Anwendung von komplexen Testverfahren zuständig ist, sondern auch beratend in Fragen der Diagnostik zur Verfügung steht.

Berlin, Brandenburg und Hamburg haben bereits sonderpädagogische Diagnostikteams eingerichtet. Zu diesen Teams gehören speziell in Diagnostik geschulte und kontinuierlich vom Landesinstitut für Lehrerbildung (LI Hamburg) bzw. vom Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM) fortgebildete Sonderpädagogen, die häufig komplexe Intelligenztests anwenden.

Die Anwendung von Intelligenztests hat auch etwas mit Handwerk und Erfahrung zu tun. Ab einem gewissen Punkt ist bei einer häufigen Anwendung der Tests die Durchführung inklusive aller Regeln kein Problem mehr und die Wahrscheinlichkeit ist höher, dass man sich in der Testsituation überwiegend auf das Kind konzentriert und nicht überwiegend auf den Test.

Literatur

- Alfonso, V. C., Johnson, A., Patinella, L. & Rader, D. E. (1998). Common WISC-III examiner errors: Evidence from graduate students in training. *Psychology in the Schools*, 35, S. 119–125.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 95, S. 29–51.
- 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: Huber.
- Huber, C. (2000). Sonderpädagogische Diagnostik im Spannungsfeld traditioneller und gegenwärtiger Sichtweisen. Ergebnisse und kritische Reflexion einer Praxisuntersuchung an Schulen für Körperbehinderte. *Zeitschrift für Heilpädagogik*, 10, S. 411–416.

- Irblich, D. (2015). Neuere Testverfahren. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 64, S. 635–648.
- Kaufman, A. S. & Kaufman, N. L. (2015). *KABC-II. Kaufman Assessment Battery for Children-II*. Deutschsprachige Fassung von P. Melchers und M. Melchers. Frankfurt am Main: Pearson.
- Lipsius, M., Petermann, F. & Daseking, M. (2008). Wie beeinflussen Testleiter die HAWIK-IV-Befunde? *Kindheit und Entwicklung*, 17 (2), S. 107–117.
- Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper and Row.
- Luria, A. R. (1970). The functional organization of the brain. *Scientific American*, 222, S. 66–78.
- Melchers, P. & Preuss, U. (1991). *Kaufman Assessment Battery for Children (K-ABC)*. Deutsche Bearbeitung. Lisse: Swets & Zeitlinger.
- Müller, C. M. (2009). Schulinterne Diagnostikansprechpartner – Eine Praxiskonzeption zur Umsetzung systematischer Diagnostik im Schulalltag. *Zeitschrift für Heilpädagogik*, 5, S. 180–187.
- Pearsonclinical (o. J.). *Qualifications Policy*. <http://www.pearsonclinical.com/psychology/qualifications.html>. [Eingesehen am 29.03.2016]
- Petermann, F. (Hrsg.) (2009). *Wechsler Preschool and Primary Scale of Intelligence-III, deutsche Version (WPPSI-III)*. Frankfurt am Main: Pearson.
- Petermann, F. & Petermann, U. (2011). *Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)*. Deutsche Bearbeitung. Frankfurt a. M.: Pearson Assessment.
- Petermann, F. (2014). *Wechsler Nonverbal Scale of Ability (WNV): Manual zur Durchführung und Auswertung*. Frankfurt: Pearson.
- Renner, G. (2015). Neue diagnostische Verfahren für die Sonderpädagogik. *Sonderpädagogische Förderung*, 4, S. 439–444.
- Schneider, W. J. & McGrew, K. S. (2012). *The Cattell-Horn-Carroll Model of Intelligence*. In D. P. Flanagan & P. L. Harrison (Hrsg.), *Contemporary intellectual assessment* (S. 99–144). New York: Guilford.
- Tellegen, P. J. & Laros, J. A. (2005). *SON-R 5½-17. Non-verbaler Intelligenztest* (3. korr. Aufl.). Göttingen: Hogrefe.
- Tellegen, P. J., Laros, J. A. & Petermann, F. (2012). *SON-R 6-40 Non-verbaler Intelligenztest. 4. Überarbeitung des Snijders-Oomen non-verbale Intelligenztests*. Technisches Manual. Göttingen: Hogrefe.
- Weiss, R. H. (2006). *CFT 20-R. Grundintelligenztest Skala 2-Revision*. Göttingen: Hogrefe.
- Weiss, R. H. & Osterland, J. (2013). *CFT 1-R. Grundintelligenztest Skala 1-Revision*. Göttingen: Hogrefe.

Dipl.-Psych. Torsten Joél
Bismarckallee 2
24105 Kiel
joel@tojoel.de

