

Redaktion

D. Reinhardt, München

Leistungen des HAWIK-IV in der Intelligenz- diagnostik im Schulalter

Der Erwerb von Kenntnissen, Erfahrungen und Fertigkeiten in der Beurteilung der sozialen, psychischen und seelischen Entwicklung steht in der Weiterbildungsordnung zum Kinder- und Jugendarzt [1] an erster Stelle des Curriculums. Einen hohen Stellenwert hat diese Aufgabe bei der Durchführung von Vorsorgeuntersuchungen im täglichen Routinebetrieb. Andererseits wird die Fragestellung nach dem kognitiven Entwicklungsstand auch häufig gezielt durch die Eltern an den Pädiater herangetragen, wenn der Verdacht auf Entwicklungsstörungen besteht. Dies betrifft alle Altersgruppen vom Neugeborenen bis zum Heranwachsenden, Letztere insbesondere dann, wenn Leistungs- und Verhaltensprobleme im Schulalltag auftreten. In diesem Zusammenhang spielen gerade auch die umschriebenen Entwicklungsstörungen schulischer Fertigkeiten wie die Lese-Rechtschreib-Störung eine große Rolle. Verhaltensprobleme können aber auch aus schulischer Über- oder Unterforderung im Rahmen von Minder- oder Hochbegabung resultieren.

Wechsler-Skalen als Konzept für die Intelligenzdiagnostik im Schulalter

Für das Schulalter stellen die Wechsler-Skalen seit Jahrzehnten das weltweit am häufigsten eingesetzte Testverfahren dar. Die von David Wechsler Mitte des vergangenen Jahrhunderts etablierte Form der Erfassung eines kognitiven Fähigkeitsprofils (Wechsler Intelligence Scale for Children, WISC) firmiert im deutsch-

sprachigen Raum seit vielen Jahren unter dem Namen Hamburg-Wechsler-Intelligenztest für Kinder (HAWIK).

Intelligenztests unterliegen aufgrund des Flynn-Effekts [12] einem Alterungsprozess, der dazu führt, dass im Abstand von 10–15 Jahren neue Vergleichswerte (Normierungen) erstellt werden müssen. Dies kann als einer der Gründe angesehen werden, warum der HAWIK-III im August 2007 durch den HAWIK-IV [21], eine Bearbeitung und Normierung der WISC-IV, ersetzt wurde. Der HAWIK-IV stellt allerdings gleichzeitig auch eine sowohl konzeptionell als auch strukturell veränderte Revisionsfassung dar (s. „Aufbau des HAWIK-IV“). Er kommt wie seine vorangegangenen Versionen als umfassendes Einzeltestverfahren zur Beurteilung der kognitiven Fähigkeiten von Kindern und Jugendlichen im Alter von 6;0–16;11 Jahren zum Einsatz. Er wurde in den Jahren 2005 und 2006 in Deutschland, Österreich und der deutschsprachigen Schweiz an 1650 Kindern und Jugendlichen normiert. Die Durchführungsdauer liegt in Abhängigkeit von der Anzahl der jeweils durchgeführten Untertests bei 60–90 Minuten.

Hamburg-Wechsler-Intelligenztest für Kinder – IV

Aufbau des HAWIK-IV

Das klassische Konzept der Wechsler-Skalen mit einem Gesamt-IQ (Gesamtingelligenzquotient), einem Verbal-IQ und einem Handlungs-IQ wurde mit dem HAWIK-IV verlassen. Es werden errechnet:

- Gesamt-IQ,
- Index Sprachverständnis,
- Index Wahrnehmungsgebundenes Logisches Denken,
- Index Arbeitsgedächtnis und
- Index Verarbeitungsgeschwindigkeit.

Diese Veränderungen sind theoretisch begründet und werden durch klinische Forschung und faktorenanalytische Ergebnisse gestützt [6, 21].

Der HAWIK-IV besteht aus 15 Untertests, 10 davon wurden vom HAWIK-III übernommen, 5 wurden neu entwickelt [22]. Für die Berechnung des Gesamt-IQ und der 4 Indizes sind 10 Untertests erforderlich. Den 5 optionalen Untertests können zusätzliche Informationen entnommen werden. Bei bestimmten Fragestellungen können einzelne Kerntests durch einen optionalen Untertest ersetzt werden, z. B. kann der Untertest „Bilder ergänzen“ den „Mosaiktest“ ersetzen, wenn ein Kind feinmotorische Schwierigkeiten hat. Eine Berechnung der Indexwerte und des Gesamt-IQ ist in diesen Fällen nur noch unter Vorbehalt möglich.

Das Gesamtkonstrukt des HAWIK-IV ist in **Abb. 1** dargestellt. Das Sprachverständnis wird durch 3 Kerntests und 2 optionale Tests repräsentiert, das wahrnehmungsbezogene logische Denken durch 3 Kerntests und 1 optionalen Untertest, das Arbeitsgedächtnis durch 2 Kerntests und 1 optionalen Untertest, die Verarbeitungsgeschwindigkeit ebenfalls durch 2 Kernverfahren und 1 optionalen Untertest (**Tab. 1**).

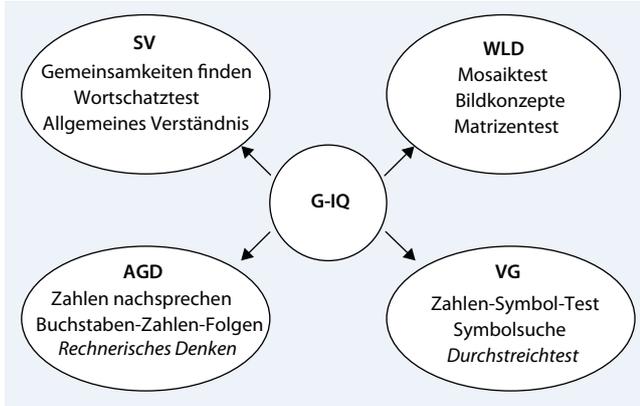


Abb. 1 ◀ Struktur des HAWIK-IV, Möglichkeit der Berechnung des Gesamt-IQ sowie der 4 Indizes für Sprachverständnis (SV), Wahrnehmungsgebundenes Logisches Denken (WLD), Arbeitsgedächtnis (AGD) und Verarbeitungsgeschwindigkeit (VG), kursiv optionale Untertests. (Mod. nach [6])

Für jeden der 4 Indizes gilt, dass jeweils nur ein Kerntest durch einen optionalen Untertest ersetzt werden darf. Der Gesamt-IQ errechnet sich aus den 10 Kerntests. Die vorliegenden Umrechnungstabellen basieren auf einer festen Reihenfolge und Zusammenstellung der Untertests. Werden Änderungen vorgenommen (Reihenfolge, Ersetzungen von Kerntests), sind die abgeleiteten IQ- und Indexwerte nur noch eingeschränkt interpretierbar.

Ziel der umfangreichen Überarbeitung waren die Optimierung der klinischen Anwendbarkeit, die Anpassung an den Entwicklungsstand der Kinder, die Verbesserung der psychometrischen Eigenschaften, die Verbesserung der Anwenderfreundlichkeit und die Aktualisierung der theoretischen Grundlagen. Um die Anwenderfreundlichkeit zu verbessern, wurden auch die Durchführungsprozeduren teilweise verändert. So wurden die sprachlichen Anleitungen überarbeitet sowie zusätzliche Lern- und Übungsaufgaben eingeführt. Durch mehrfache Wiederholung der Aufgabenstellung sowie die Rückmeldung von richtigen Lösungen und Bearbeitungsweisen sollen insbesondere intelligenzgeminderte Kinder in die Lage versetzt werden, ein Verständnis für den jeweiligen Aufgabentyp zu entwickeln.

Testverfahren, die für einen breiten Alters- und Fähigkeitsbereich konzipiert sind, weisen in den jeweiligen Randbereichen so genannte Boden- und Deckeneffekte auf. Dies zeigt sich insbesondere bei jungen intelligenzgeminderten Kindern sowie bei Jugendlichen mit kognitiver Hochbegabung. Um diese Effekte zu reduzieren, wurden für den neuen Test deutlich leichtere und schwierigere Aufgaben eingebunden, die die Differenzierungsfähigkeit erhöhen. Somit lässt sich der HAWIK-IV auch

in der Diagnostik von Hoch- und Minderbegabung einsetzen. Für den HAWIK-III war dies nur begrenzt möglich. Mit dem HAWIK-IV wird der IQ-Bereich von 40–160 abgedeckt, Werte über diese Grenzen hinaus können nicht berechnet werden. In den Randbereichen der IQ-Verteilung (Hochbegabung, Intelligenzminderung) ist zudem von einer geringeren Differenzierungsfähigkeit auszugehen. Dies beruht auf den Annahmen zur Normalverteilung im Rahmen der klassischen Testtheorie, da nur jeweils etwa 2,3% der Stichprobe in die genannten Extrembereiche (IQ 130 und IQ 70) fallen. Das Differenzierungsoptimum liegt bei Tests, die auf einen mittleren IQ von 100 ± 15 standardisiert wurden, im mittleren Leistungsbereich.

Neukonzeption der Indizes

Bei der Aktualisierung der theoretischen Grundlagen der WISC-IV wurde der Messung

- des fluiden Denkens,
- des Arbeitsgedächtnisses und
- der Verarbeitungsgeschwindigkeit

mehr Raum gegeben. Verschiedene Theorien zur kognitiven Funktionsfähigkeit betonen die Bedeutung des fluiden Denkens [3, 4, 5]. Dabei handelt es sich allgemein um die Fähigkeit, neue Probleme zu lösen oder Schlussfolgerungen zu ziehen [4]. Um diese Fähigkeit zu erfassen, wurden für die WISC-IV 3 neue Untertests entwickelt: Matrizentest, Bildkonzepte und Begriffe erkennen.

Die aktuelle Forschung zeigte zudem, dass das Arbeitsgedächtnis eine grundlegende Komponente des fluiden Denkens und anderer kognitive Prozesse höherer Ordnung darstellt [13, 20, 26]. Für sei-

ne Messung wurden der Untertest Buchstaben-Zahlen-Folgen aus der Wechsler-Adult-Intelligence-Scale-III (WAIS-III) adaptiert, das Rechnerische Denken überarbeitet und das Zahlennachsprechen mit zusätzlichen Auswertungsmodalitäten versehen. So ist es nun möglich, die Leistungen in den beiden Teilbereichen Zahlennachsprechen vorwärts und rückwärts miteinander zu vergleichen.

Die Verarbeitungsgeschwindigkeit wiederum ist mit verschiedenen mentalen Prozessen wie der Leseleistung oder dem effizienten Gebrauch des Arbeitsgedächtnisses verknüpft. Sie wurde in faktorenanalytischen Studien als eine wichtige kognitive Funktion identifiziert [3, 17]. Für diese Intelligenzkomponente wurde ein neuer Untertest, der Durchstreichtest, entwickelt.

Messtechnische Güte des HAWIK-IV

Der HAWIK-IV weist eine hohe Reliabilität auf, die über die Testhalbierungsmethode und für die Speedtests über eine Retestung ermittelt wurden. Für den Gesamt-IQ wird die Reliabilität mit $r=0,97$ angegeben, auf Ebene der Indizes liegen die Werte zwischen 0,88 für die Verarbeitungsgeschwindigkeit und 0,94 für das Sprachverständnis. Auf Untertestebene kann den sprachlichen Untertests mit Werten zwischen 0,81 und 0,90 eine hohe Zuverlässigkeit zugesprochen werden. Standardmessfehler und Vertrauensintervalle werden im Manual dezidiert tabellarisch beschrieben.

Die Validität wurde durch Interkorrelationsstudien und faktorenanalytische Studien dargestellt. Auf einzelne Ergebnisse wird im Folgenden noch näher eingegangen.

Vergleich von HAWIK-III und HAWIK-IV

Bei jeder Testrevision muss die Frage beantwortet werden, inwieweit die abgeleiteten Werte der neuen Version, wie der IQ, mit denen der alten übereinstimmen bzw. vergleichbar sind. Daher wurden auch zum Vergleich von HAWIK-III und HAWIK-IV entsprechende Studien durchgeführt; beide Tests wurden in ausbalancierter Reihenfolge insgesamt 110 Kindern vorgegeben [21]. Die höchsten Zusammenhänge ergaben sich in Abhän-

gigkeit von der Testreihenfolge zwischen dem Verbalteil (0,97/0,79 für HAWIK-IV vs. HAWIK-III/HAWIK-III vs. HAWIK-IV) bzw. dem Sprachverständnisindex (0,75/0,80) des HAWIK-III einerseits und dem neuen Index Sprachverständnis im HAWIK-IV andererseits. Der Gesamt-IQ korreliert mit 0,73/0,63.

Im Mittel ergeben sich höhere IQ-Werte im HAWIK-III. Das entspricht den durch den Flynn-Effekt induzierten Erwartungen, die von einem mittleren Anstieg der IQ-Werte von 3 Punkten innerhalb von 10 Jahren ausgehen. Interessanterweise betrifft dieser Effekt stärker Testverfahren und Untertests, die die fluide Intelligenz erfassen; weniger deutlich wirkt sich der Flynn-Effekt auf erworbene, sprachbezogene Leistungen (kristalline Intelligenz) aus [12, 27].

Interpretationsmöglichkeiten auf Gesamt-IQ- und Indexebene

Die Einführung von verschiedenen Indizes ermöglicht es, die individuellen Leistungen eines Kindes unterhalb eines allgemeinen Gesamtwertes (G-IQ) differenziert als Stärken und/oder Schwächen herauszuarbeiten. Dieses Vorgehen wurde bereits für den WISC-III entwickelt.

Auf oberster Ebene repräsentiert der Gesamt-IQ zunächst eine globale Fähigkeit, den so genannten g-Faktor, die sich in der Bearbeitung und Lösung sehr unterschiedlicher kognitiver Aufgaben und Anforderungen widerspiegelt. Viele diagnostische Fragestellungen werden unter Verwendung dieses allgemeinen Maßes kognitiver Fähigkeiten beantwortet.

Die Praxis der Profilinterpretation muss dabei durchaus kritisch hinterfragt werden. Leistungsunterschiede zwischen einzelnen Untertests oder Skalen können dann umso zuverlässiger interpretiert werden, je weniger die Untertests inhaltlich und statistisch miteinander zusammenhängen [18]. Außerdem kann eine Bewertung von Differenzen nur unter der Voraussetzung einer guten Reliabilität erfolgen, da hohe Messfehler (geringe Reliabilität) die Interpretation erhobener Testwerte und deren Differenzen deutlich beeinträchtigen. So basiert insbesondere die Prozessanalysen zum Durchstreichtest (DT strukturiert vs. DT unstrukturiert) auf verhältnismäßig nied-

Zusammenfassung · Abstract

Monatsschr Kinderheilkd 2009 · 157:587–594 DOI 10.1007/s00112-008-1909-2
© Springer Medizin Verlag 2008

R.G. Schmid · U. Petermann · F. Petermann · C. Weiß · M. Daseking

Leistungen des HAWIK-IV in der Intelligenzdiagnostik im Schulalter

Zusammenfassung

Der HAWIK-IV [Hamburg-Wechsler-Intelligenztest für Kinder IV, deutsche Bearbeitung der WISC-IV (Wechsler Intelligence Scale for Children)] stellt im Vergleich zum Vorgänger HAWIK-III eine grundlegende Revision dar, die durch die Entwicklung neuer Untertests und die Einführung von 4 Indizes zur Interpretation des Leistungsprofils geprägt ist. Dadurch ergibt sich die Frage nach der Anwendbarkeit bei klinischen und pädagogischen Fragestellungen. Dem durch Diskrepanzanalysen ermittelten individuellen Stärken- und Schwächenprofil können bereits wichtige

Hinweise auf zugrunde liegende Störungsbilder und Komorbiditäten sowie Förderansätze entnommen werden. Der Einsatz des HAWIK-IV bei verschiedenen pädiatrischen und kinderpsychiatrischen Fragestellungen wird am Beispiel von Kindern mit Lese- und Rechtschreibstörung, Aufmerksamkeitsdefizit-Hyperaktivitäts-Störung (ADHS) und Hochbegabung diskutiert.

Schlüsselwörter

Intelligenzdiagnostik · HAWIK-IV · Lese- und Rechtschreibstörung · Hochbegabung · ADHS

Benefits of intelligence assessment using the HAWIK-IV in school-aged children

Abstract

The HAWIK-IV [Hamburg-Wechsler intelligence test for children IV, the German version of the WISC-IV (Wechsler Intelligence Scale for Children)], is a fundamental revision of its predecessor, HAWIK-III, characterized by the development of new sub-tests and the introduction of 4 indices for the interpretation of intelligence profiles. Thus, the question of its application in clinical and pedagogic settings arises. Individual strengths and weaknesses identified by means of discrepancy analyses are already able to provide informa-

tion pointing to underlying disorders and comorbidities, as well as promotion approaches. The use of HAWIK-IV in pediatrics and pediatric psychiatry is discussed using children with dyslexia or attention deficit-hyperactivity syndrome (ADHD), as well as highly gifted children as examples.

Keywords

Assessment of intelligence · HAWIK-IV (German version of WISC-IV) · Dyslexia · Giftedness · ADHD

Tab. 1 Untertests des HAWIK-IV mit den dadurch erfassten Funktionen. (Mod. nach [6])

Index Sprachverständnis	Gemeinsamkeiten finden	Verbale Schlussfolgerungen und Konzeptbildung Auditives Verständnis Gedächtnis Verbaler Ausdruck
	Wortschatztest	Wortwissen Lernfähigkeit Langzeitgedächtnis und Stand der Sprachentwicklung
	Allgemeines Verständnis	Verbale Schlussfolgerungen Sprachliches Verständnis Sprachlicher Ausdruck Soziales Urteil Wissen um Verhaltensstandards
	<i>Allgemeines Wissen (optional)</i>	Kristalline Intelligenz Langzeitgedächtnis Faktenwissen
	<i>Begriffe erkennen (optional)</i>	Sprachliches Verständnis Sprachliches schlussfolgern Verbale Abstraktion Bereichswissen Entwicklung von alternativen Konzepten Kognitive Flexibilität
Index wahrnehmungs- gebundenes logisches Denken	Mosaiktest	Analyse und Synthese abstrakter visueller Reize Visuomotorische Koordination
	Bildkonzepte	Abstraktes und kategoriales Denken
	Matrizentest	Allgemeine intellektuelle Fähigkeit Fluide Intelligenz
	<i>Bilder ergänzen (optional)</i>	Visuelle Wahrnehmung und visuelle Organisation
Index Arbeitsgedächtnis	Zahlennachsprechen	Auditives Kurzzeitgedächtnis Fähigkeit zur Reihenbildung Aufmerksamkeit Zahlennachsprechen vorwärts: Automatisiertes Lernen und Gedächtnis Umformung von Informationen Aufmerksamkeit Auditive Verarbeitung Zahlennachsprechen rückwärts: Arbeitsgedächtnis Mentale Rotation Visuell-räumliches Vorstellungsvermögen Wechsel von Zahlennachsprechen vorwärts zu Zahlennachsprechen rückwärts: Kognitive Flexibilität
	Buchstaben-Zahlen-Folgen	Reihenfolgebildung (Serialität) Mentale Rotation Aufmerksamkeit Auditives Kurzzeitgedächtnis Verarbeitungsgeschwindigkeit
	<i>Rechnerisches Denken (optional)</i>	Konzentration Aufmerksamkeit Kurz- und Langzeitgedächtnis Rechenfähigkeit
Index Verarbeitungsgeschwindigkeit	Zahlen-Symbol-Test	Kurzzeitgedächtnis Lernfähigkeit Visuelle Wahrnehmung Visuomotorische Koordination Aufmerksamkeit
	Symbolsuche	Verarbeitungsgeschwindigkeit Visuelles Kurzzeitgedächtnis Visuomotorische Koordination Kognitive Flexibilität
	<i>Durchstreichtest (optional)</i>	Verarbeitungsgeschwindigkeit Visuelle selektive Aufmerksamkeit

rigen Reliabilitätskennwerten (DT-U=0,62, DT-S=0,76) und sollte daher nur vorsichtig interpretiert werden. Mit Ausnahme der Untertests Symbolsuche (0,79) und Begriffe erkennen (0,79) weisen alle übrigen Untertests gute bis sehr gute Kennwerte auf.

Einsatzmöglichkeiten des HAWIK-IV bei besonderen Fragestellungen

Um den Entwicklungsstand eines Kindes oder Jugendlichen erfassen zu können, ist u. a. der Einsatz wissenschaftlich fundierter und aktuell normierter Testverfahren notwendig. Dabei muss zwischen Entwicklungsdiagnostik im allgemeinen Sinn und Intelligenzdiagnostik im Speziellen unterschieden werden. Entwicklungs- und Intelligenztests kommen bei einer Vielzahl von pädiatrischen, kinderpsychologischen und kinderpsychiatrischen Fragestellungen zum Einsatz [5, 6, 8]. Der Testanwender steht dabei immer wieder auch vor der Frage, welchen Test er bei welcher Fragestellung einsetzen kann bzw. anhand welcher Kriterien er beurteilen kann, ob der ihm zur Verfügung stehende Test für die jeweilige Fragestellung geeignet ist. Diese Frage stellt sich insbesondere dann, wenn ein neuer Test publiziert wurde. Zur Untersuchung klinischer Fragestellungen wurde der HAWIK-IV eingesetzt bei:

- Kindern mit Hochbegabung,
- Kindern mit leichter und mittelgradiger Intelligenzminderung,
- Kindern mit Aufmerksamkeitsdefizit-Hyperaktivitäts-Störungen (ADHS) und
- Kindern mit Lese-Rechtschreib-Störung.

Dabei ist zu berücksichtigen, dass ein Intelligenztest per se – mit Ausnahme von Hochbegabung und Intelligenzminderung – keine Diagnosestellung erlaubt, sondern dass Profilanalysen Hinweise auf einen weiteren differenzialdiagnostischen Prozess geben können.

Diagnostik von umschriebenen Entwicklungsstörungen schulischer Fertigkeiten

Etwa 10–15% der Grundschul Kinder haben deutliche Probleme beim Erwerb der Schriftsprache; die Prävalenz für eine kli-

nisch bedeutsame Lese-Rechtschreib-Störung wird über alle bekannten Sprachsysteme hinweg in Abhängigkeit von der Definition mit 2–8% angegeben [23, 28]. Die Leitlinien der Deutschen Gesellschaft für Kinder- und Jugendpsychiatrie und -psychotherapie (DGKJP) und die Qualitätskriterien der Sozialpädiatrie (DGSP) ordnen die Intelligenzdiagnostik neben den leistungsspezifischen Tests (Rechtschreib-, Rechen-, Lesetest) als unverzichtbaren Bestandteil im Rahmen einer mehrdimensionalen Bereichsdiagnostik bzw. multiaxial konzipierten Diagnostik umschriebener Entwicklungsstörungen ein [8, 24, 28].

HAWIK-IV und Lese- und Rechtschreibstörung

Das bestehende Leitlinienverfahren (insbesondere die Diskrepanzannahme von Fähigkeiten zu Fertigkeiten) führt in vielen Fällen dazu, dass eine Diagnose, die zu einer institutionalisierten (und nicht durch Eltern selbst zu finanzierenden, s. Sozialgesetzbuch (SGB) VIII) Förderung berechtigt, im Entwicklungsverlauf häufig erst verspätet gestellt wird [30]. Dabei wird ein Leistungsunterschied von mindestens 12 T-Wert-Punkten oder 1,2 Standardabweichungen [8, 24] zwischen allgemeinem kognitiven Niveau und der beeinträchtigten Teilleistung (Lesen, Schreiben, Rechnen) vorausgesetzt, wobei die Teilleistung selbst gleichzeitig schwächer als Prozentrang 10 ausfallen muss. Als Grenzwert für die Normalitätsannahme wird ein IQ > 70 vorgegeben. Durch diese beiden Kriterien werden einerseits Kinder mit niedrigem Gesamt-IQ benachteiligt, da sie das Diskrepanzkriterium trotz niedriger Lese- oder Rechtschreibkompetenz nicht einhalten können, andererseits unterschreiten Kinder mit einem überdurchschnittlichen IQ trotz Diskrepanzen von 1,5 SD und mehr häufig den Grenzwert von PR=10 nicht.

Oft dauert es bis zur 3. Klasse, ehe Kinder mit Lese- oder Rechtschreibschwierigkeiten die Diskrepanzkriterien für eine Förderung erfüllen, auch wenn frühzeitig Anzeichen für Leistungsschwächen bestanden haben [19]. Dies kann auf die Variabilität von Lehrplänen in den ersten beiden Schuljahren zurückgeführt

werden (Abschluss des Lese- und Rechtschreiblehrgangs erfolgt erst Ende der 2. Klasse), andererseits aber auch durch individuelle Kompensationsversuche (wie das Auswendiglernen von Texten) mitbegründet sein. Eine Förderung setzt dann u. U. erst zu einem Zeitpunkt ein, an welchem betroffene Kinder bereits ein hohes Maß an Frustration erlebt haben. Hier wird von verschiedenen Arbeitsgruppen der Vorschlag gemacht, dass im Fall eines Verdachts auf Lese-Rechtschreib-Störung bei jüngeren Kindern basale Fähigkeiten geprüft werden: phonologische Bewusstheit, Benennungsgeschwindigkeit oder Kurzzeitgedächtnis. Bei älteren Kindern sollte dies ergänzt werden durch grammatisches Grundwissen und andere für das Lesen und das Leseverständnis relevanten Aspekte [30]. Die Fähigkeiten-Fertigkeiten-Diskrepanz spielt dabei für die Identifikation und Intervention von lese- und rechtschreibschwachen Kindern nur eine untergeordnete Rolle.

Kinder aller IQ-Bereiche können unter einer Lese- und Rechtschreibstörung leiden. Da das Diskrepanzmodell in den Randbereichen der IQ-Verteilung (niedrige oder hohe Intelligenz) dem klinischen Bild nicht immer gerecht wird, wird für diese Fälle in den Leitlinien zur Diagnostik und Therapie von psychischen Störungen [8] vorgeschlagen, einen mittels Regression errechneten Erwartungswert für den Prozentrang der Leistung im Lese- oder Rechtschreibtest im Verhältnis zum IQ einzusetzen. Liegt die Lese- und/oder Rechtschreibleistung unterhalb dieses Werts, kann eine entsprechende Störung diagnostiziert werden. Dies gilt in ähnlicher Weise für die Diagnostik einer Rechenstörung ([8], S 213).

Zusätzlich ist zu berücksichtigen, dass bei vielen Kindern mit Lese-Rechtschreib-Störung komorbid eine Aufmerksamkeitsstörung vorliegt.

Zur Bestimmung des kognitiven Niveaus eines Kindes und insbesondere zum Ausschluss einer allgemeinen Intelligenzminderung wurde in den Leitlinien mit der Begründung einer im Vergleich zu eindimensionalen Verfahren höheren Messgenauigkeit bislang der Einsatz des HAWIK-III oder eines vergleichbaren, breit angelegten Intelligenztests empfohlen. Da jedoch der Gesamt-IQ durch schwächere

Leistungen in den sprachlichen Untertests konfundiert sein könnte, soll die Entscheidung über das Vorliegen einer Lese-Rechtschreib-Störung unter Bezugnahme auf den Handlungsteil getroffen werden. Um die Sensitivität und Spezifität von Entscheidungen zu erhöhen, sollte daher zusätzlich eine gezielte Analyse auf Index- und Untertestebene vorgenommen werden. Während sich beispielsweise bei intelligenzgeminderten Kindern homogene Profile im unterdurchschnittlichen Leistungsbereich ergeben, kann die Interpretation eines Gesamt-IQ gerade im Fall der umschriebenen Entwicklungsstörungen zu einer Fehldiagnose führen [10]. Durch den Wegfall der Aufteilung des HAWIK-IV in Verbal- und Handlungsteil ergibt sich nunmehr die Frage, welcher der insgesamt 5 Indizes (SV, WLD, AGD, VG, G-IQ) in Zukunft als Referenzwert herangezogen werden kann. Der WLD-Index scheint am ehesten geeignet, diese Funktion zu übernehmen. Der HAWIK-IV erfasst in 2 Indizes explizit Leistungen des Arbeitsgedächtnisses und der Verarbeitungsgeschwindigkeit, beides Bereiche, die bereits als zentrale Komponenten im diagnostischen Prozess der Lese-Rechtschreib-Störung benannt wurden. Diskrepanzanalysen können bestätigen, ob es sich bei individuellen Leistungsunterschieden im Indexprofil tatsächlich auch um klinisch relevante Differenzen handelt. Darüber hinaus können weitere Analysen (z. B. Vergleich von Zahlennachsprechen vorwärts vs. rückwärts) zusätzliche Informationen über kognitive Verarbeitungsprozesse des Kindes liefern.

Vorläufige Ergebnisse zum HAWIK-IV bei Kindern mit Lese-Rechtschreib-Störung ([21], S 142) zeigen, dass insbesondere die Differenz zwischen den Indizes WLD und AGD statistisch signifikant ausfällt. Der HAWIK-IV kann also über die Profilanalyse zusätzliche diagnostische Informationen zur Verfügung stellen.

HAWIK-IV und Aufmerksamkeitsstörung

Auch im Rahmen einer Diagnostik bei Verdacht auf eine Aufmerksamkeitsstörung wird der Einsatz von Intelligenztests kontrovers diskutiert. Bei Schulkindern wird zum Ausschluss von Lernstörungen

oder einer (möglicherweise auch komorbiden) Intelligenzminderung in Form einer Lernbehinderung die Durchführung eines Intelligenztests empfohlen, v. a. wenn es Hinweise auf schulische Über- oder Unterforderung gibt ([8], S 243). Die Indexstruktur des HAWIK-IV ermöglicht es wiederum, spezifische Teilleistungsdefizite zu identifizieren. Bei Kindern mit ADHS wurden schwächere Leistungen im Index Verarbeitungsgeschwindigkeit gefunden und damit anderweitige Studienergebnisse bestätigt [2, 9]. Niedrige Werte erzielten Kinder mit ADHS auch im Arbeitsgedächtnis. Calhoun u. Dickerson Mayes [2] gingen davon aus, dass der Index VG des HAWIK-IV durch die strukturellen Veränderungen deutlich sensibler gegenüber Aufmerksamkeitsstörungen ist als der entsprechende Index Arbeitsgeschwindigkeit des HAWIK-III. Ähnliches wird für den Index Arbeitsgedächtnis des HAWIK-IV im Vergleich zum Index Unablenkbarkeit des HAWIK-III angenommen.

Schwean u. Saklofske [29] wiesen darauf hin, dass auf der Basis eines HAWIK-IV-Befunds weder eine ADHS-Diagnosestellung noch eine Unterscheidung von Subtypen erfolgen können. Dennoch lässt sich aus dem Profil ein individuelles Muster von Stärken und Schwächen ableiten, das wesentliche Impulse für die Therapieplanung geben kann. Bei der Interpretation der Ergebnisse ist unbedingt zu berücksichtigen, dass die Aufmerksamkeitsstörung an sich die Testdurchführung und die Testleistung beeinflussen kann. So kann beispielsweise der impulsive Arbeitsstil zu einer höheren Fehlerzahl führen. Andererseits können sich aber auch eine unstrukturierte Testsituation oder Umgebung oder die fehlende Routine in der Testdurchführung negativ auf das Arbeitsverhalten des Kindes auswirken. Daher stellen Verhaltensbeobachtung und Dokumentation der Testsituation wichtige Momente in der Interpretation des Ergebnisses dar.

HAWIK-IV und Hochbegabung

Die Diagnostik von Hochbegabung genießt inzwischen in Deutschland einen höheren Stellenwert als noch vor wenigen Jahren. An dieser Stelle soll nicht nä-

her auf Modellvorstellungen zur Hochbegabung eingegangen werden. Es ist aber festzustellen, dass allen Modellen eine Nähe zum psychometrischen Intelligenzkonstrukt eigen ist. Als Maß der quantitativen Abgrenzung für Hochbegabung hat sich in vielen Bereichen der IQ-Wert von 130 (entspricht 2 SD über dem Mittelwert) durchgesetzt [25], der sich auf die allgemeine Intelligenz, also auf den g-Faktor nach Spearman bezieht, und mittels eines Intelligenztestes erhoben wird. Dies wird u. a. damit begründet, dass die erfasste Allgemeinbegabung die Leistung in anderen Bereichen wie Schule oder Beruf am besten vorhersagen kann.

Gerade die Diagnostik von Leistungen an den Randbereichen der Normalverteilung stellt besondere Ansprüche an die dafür einzusetzenden Testverfahren. Wurde von den deutschen Herausgebern des HAWIK-III noch empfohlen, diesen Intelligenztest eher im Sinne eines Screenings einzusetzen und den Ausprägungsgrad der Hochbegabung mit einem speziell für diese Fragestellung normierten Verfahren zu differenzieren, kann davon ausgegangen werden, dass der Deckeneffekt im HAWIK-IV durch die Erhöhung der Itemzahlen pro Untertest und die Hinzunahme deutlich schwierigerer Aufgaben nicht mehr so deutlich auffällt. Dennoch bleibt zu berücksichtigen, dass die Standardisierung auf der Basis der Normalverteilung mit einem Mittelwert von 100(±15) nur IQ-Werte von bis 160 vorsieht. Kognitive Leistungen, die darüber hinaus gehen, können also nicht „berechnet“ werden. Zudem fällt die Genauigkeit der Leistungsbeurteilung aufgrund der Randlage testtheoretisch bedingt deutlich niedriger aus als im mittleren Leistungssegment.

Auch auf ein weiteres Phänomen ist an dieser Stelle hinzuweisen: den „IQ-Gewinn“. Dabei handelt es sich um die auch als Flynn-Effekt bezeichnete Normverschiebung des IQ, die sich beispielsweise im HAWIK-R auf etwa 0,6–0,7 IQ-Punkte pro Jahr belief [27]. Das bedeutet, dass bei einer Leistungsverschiebung von 5 IQ-Punkten (tritt etwa 7 Jahre nach einer Neunormierung ein) doppelt so viele Kinder als hochbegabt diagnostiziert werden als es der Normalverteilungsannahme entspricht, bei einem Gewinn von 10 IQ-Punkten würden dann mit dem Test

etwa 9% Hochbegabte „erzeugt“, in dieser Gruppe befinden sich dann nur noch 1/4 an Kindern mit einem $IQ > 130$ ([15], S 104). Daraus ist abzuleiten, dass insbesondere bei diagnostischen Fragestellungen, bei denen ein punktgenauer IQ-Wert zur Zuordnung zu der einen oder anderen Kategorie (hier: hochbegabt oder nicht hochbegabt) führt, ausschließlich Testverfahren mit aktueller Normierung zum Einsatz kommen dürfen.

Zur Diagnostik einer Hochbegabung sollten zudem 2 Intelligenztests herangezogen werden, die ähnliche Fähigkeiten auf verschiedene Weise messen. Rost et al. [25] schlugen ein mehrphasiges Vorgehen vor: Durch einen ökonomischen Gruppenintelligenztest, der im Sinne eines Screenings eingesetzt wird, kann zwischen Kindern unterschieden werden, die hochbegabt oder nicht hochbegabt sind, außerdem ergibt sich eine 3. Gruppe von Kindern, bei denen eine genaue Zuordnung nicht möglich ist. Daher wird durch ein 2. Testverfahren, in diesem Fall ein Individualtest, die Screeningentscheidung (Diagnose) verifiziert. Hier kommen Verfahren wie der HAWIK-IV oder der AID 2 (Adaptives Intelligenzdiagnostikum 2) zum Einsatz. Holocher-Ertl et al. [16] verwiesen darauf, dass der HAWIK-IV dem traditionellen Hochbegabungsmodell (Kriterium $IQ > 130$) gerecht wird, während der AID 2, der im Übrigen keinen Gesamt-IQ im herkömmlichen Sinn vorsieht, mit einer förderdiagnostischen Orientierung eingesetzt werden kann.

Sowohl für den WISC-IV als auch für den HAWIK-IV zeigen sich bei Kindern mit diagnostizierter Hochbegabung in allen Kerntests, Indizes und dem Gesamt-IQ signifikante Unterschiede im Vergleich zu einer gematchten Kontrollstichprobe ([11, 21], S 139). Es muss allerdings auch festgehalten werden, dass der Gesamt-IQ sowohl im amerikanischen Originaltest als auch im HAWIK-IV mit 123,5 bzw. 124,3 deutlich unterhalb der Grenze von 130 liegt. Dies wird neben dem Verweis auf den Flynn-Effekt auch mit der neuen Untertest- und Indexstruktur begründet, insbesondere der stärkeren Gewichtung von Arbeitsgedächtnis und Verarbeitungsgeschwindigkeit [11]. In den Untertests dieser beiden Indizes weisen hochbegabte Kinder und Jugendliche im Durch-

schnitt die intraindividuell schwächsten Leistungen auf.

HAWIK-IV bei Kindern mit anderem sprachlichen und kulturellen Hintergrund

Der HAWIK-IV wurde nicht nur in Deutschland, sondern auch in Österreich und in der deutschsprachigen Schweiz normiert. Bei der Adaptation wurde viel Wert darauf gelegt, regionale Besonderheiten in der Sprache so in den Test einzubinden, dass für Kinder mit einem anderen sprachlichen Hintergrund (wie Schweizerdeutsch) kein Nachteil entsteht. Die Ergebnisse einer vergleichenden Studie [14] deuten darauf hin, dass zwischen Kindern aus Deutschland und der Schweiz mit Ausnahme des Untertests Mosaiktest keine signifikanten Leistungsunterschiede bestehen. Insbesondere in den sprachlichen Untertests zeigte sich eine hohe Übereinstimmung in Bezug auf die durchschnittliche Leistung.

Die Frage, welcher Intelligenztest bei Kindern eingesetzt werden sollte, die einen Migrationshintergrund aufweisen oder die deutsche Sprache erst als Zweitsprache erlernt haben, wird vielfältig diskutiert. Dabei ist einerseits die Fragestellung zu berücksichtigen, die die Durchführung eines Intelligenztests begründet, andererseits müssen aber die Grenzen des Einsatzes von Intelligenztests bei Kindern, die die Sprache des Tests nur unzureichend beherrschen, beachtet werden.

In einer Studie zum HAWIK-IV konnte gezeigt werden, dass die größten Unterschiede zwischen einsprachig und mehrsprachig aufwachsenden Kindern und Jugendlichen für diejenigen Kinder und Jugendlichen bestehen, die eine Hauptschule besuchen. Dazu kommt, dass nur 1/3 der Eltern von mehrsprachig aufwachsenden Kindern, die eine Hauptschule besuchen, selbst einen Realschulabschluss aufweisen, während die Eltern der anderen Kinder und Jugendlichen entweder keinen oder auch nur einen Hauptschulabschluss haben [7]. Bei Kindern und Jugendlichen, die eine Realschule oder ein Gymnasium besuchen, fallen die Differenzen zwischen beiden Gruppen erheblich niedriger aus. Beim Einsatz des HAWIK-IV ist also zu fragen, welches Ziel

mit der Diagnostik verfolgt wird: Im Fall von Entscheidungen, die bezogen auf die Intelligenz an einen Cut-off-Wert gebunden sind (s. beispielsweise Diagnostik der umschriebenen Entwicklungsstörungen oder Platzierungsentscheidungen im schulischen Setting), sollte unbedingt ein 2., sprachfreier Test durchgeführt werden. Dazu könnten der SON (Snijders-Oomen nichtverbaler Intelligenztest) oder die Raven-Matrizen herangezogen werden. Geht es allerdings um ein differenziertes Leistungsprofil, das auch den Umgang mit Sprache einschließt, kann das Profil des HAWIK-IV wertvolle Informationen über Stärken und Schwächen geben. So kann erfasst werden, ob es sich um ein allgemeines kognitives Defizit handelt (schwaches Abschneiden in allen Indizes) oder ob es aufgrund mangelnder Sprachkompetenz zu eingeschränkten Leistungen besonders im Index Sprachverständnis kommt, die übrigen Indexwerte dagegen aber höher ausfallen. Dieser Befund spricht für eine intensive Sprachförderung, um zu verhindern, dass sich die Unterschiede im weiteren Entwicklungsverlauf noch verschärfen. Zudem lassen gute Leistungen in den Indizes Arbeitsgedächtnis und Verarbeitungsgeschwindigkeit auf ein basales Lernpotenzial schließen, auf dessen Basis das Kind von einer angemessenen Sprachförderung profitieren kann.

Fazit für die Praxis

Die Wechsler-Skalen gehören weltweit zu den meist verwendeten Testverfahren zur Messung der Intelligenz und eines breit angelegten Intelligenzprofils. Mit dem HAWIK-IV liegt mit einer Normierung an 1650 Kindern und Jugendlichen in Deutschland, Österreich und der deutschsprachigen Schweiz eine aktuelle Fassung für den deutschen Sprachraum vor. Das klassische Konzept des HAWIK mit einem Gesamt-IQ, Verbal-IQ und Handlungs-IQ wurde verlassen. Neben dem Gesamt-IQ werden Indizes für Sprachverständnis, Wahrnehmungsgeladenes Logisches Denken, Arbeitsgedächtnis und Verarbeitungsgeschwindigkeit errechnet. Die Konstruktion der 4 Indizes durch Entwicklung neuer bzw. modifizierter Untertests trägt der kindlichen

Entwicklung und testtheoretischen Modellen Rechnung. Beim Vergleich von HAWIK-IV und HAWIK-III mittels einer ausbalancierten Stichprobe von 100 Kindern liegt eine gute Korrelation insbesondere bei den Untertests und Skalen zum Sprachverständnis vor. Der Gesamt-IQ im HAWIK-IV liegt im Mittel um 3,4 IQ-Punkte niedriger als im HAWIK-III. Erste Ergebnisse zu Hochbegabung, Intelligenzminderung, ADHS sowie zur Lese- und Rechtschreibstörung zeigen die Validität und bestätigen damit die theoretische Grundlage des HAWIK-IV. Im Rahmen pädiatrischer und kinderpsychiatrischer Diagnostik kann der HAWIK-IV somit den Platz des Vorgängers HAWIK-III einnehmen. Er bietet über die Diskrepanzanalysen auf Indexebene zusätzliche Auswertungsmodalitäten an, aus denen sich beispielsweise erste Hinweise auf das Vorliegen einer umschriebenen Entwicklungsstörung schulischer Fertigkeiten oder einer Aufmerksamkeitsstörung entnehmen lassen, die dann jeweils differenzialdiagnostisch zu validieren sind. Die Profilanalysen können wichtige Impulse für die Planung von Therapien oder Fördermaßnahmen geben. Dies gilt in gleicher Weise für klinische Störungen, schulische Lernstörungen oder eine überdurchschnittlich hohe Begabung. Mit dem HAWIK-IV steht damit auch im deutschen Sprachraum wieder ein zeitnah normiertes und standardisiertes Verfahren zur Verfügung, mit dem eine internationale Vergleichbarkeit von Entwicklungs- und Intelligenzdiagnostik auch für den deutschen Sprachraum gegeben ist.

Korrespondenzadresse

Prof. Dr. R.G. Schmid
Zentrum für Kinder und Jugendliche
Vinzenz-von-Paul-Strasse 10–14,
84503 Altötting
mail@kinderzentrum.de

Interessenkonflikt. Keine Angaben

Literatur

1. Bayerische Landesärztekammer (2005) Weiterbildungsordnung für Ärzte Bayerns vom 24.04.2004. Bayerische Landesärztekammer, München, <http://www.blaek.de/weiterbildung/WBO%2024%2004%202004.pdf>
2. Calhoun SL, Dickerson Mayes S (2005) Processing speed in children with clinical disorders. *Psychol Schools* 42: 333–343
3. Carroll JB (1993) Human cognitive abilities: a survey of factor-analytic studies. Cambridge University Press, Cambridge
4. Cattell RB, Horn JL (1978) A check on the theory of fluid and crystallized intelligence with description of new subtest design. *J Educ Meas* 15: 139–164
5. Daseking M, Janke N, Petermann F (2006) Intelligenzdiagnostik. *Monatsschr Kinderheilkd* 154: 314–319
6. Daseking M, Petermann U, Petermann F (2007) Intelligenzdiagnostik mit dem HAWIK-IV. *Kindheit Entwicklung* 16: 250–259
7. Daseking M, Lipsius M, Petermann F et al. (2008) Differenzen im Intelligenzprofil bei Kindern mit Migrationshintergrund: Befunde zum HAWIK-IV. *Kindheit Entwicklung* 17: 76–89
8. Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie (Hrsg) (2007) Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter, 3. überarb. Aufl. Deutscher Ärzteverlag, Köln
9. Donders J, Warschawsky S (1997) WISC-III factor index score pattern after traumatic head injury in children. *Child Neuropsychol* 3: 71–78
10. Fiorello CA, Hale JB, Holdnack JA et al. (2007) Interpreting intelligence test results for children with disabilities: is global intelligence relevant? *Appl Neuropsychol* 14: 2–12
11. Flanagan DP, Kaufmann AS (2004) Essentials of WISC-IV assessment. Wiley, Hoboken
12. Flynn JR (2007) What is intelligence? Beyond the Flynn effect. Cambridge University Press, Cambridge
13. Fry AF, Hale S (1996) Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade. *Psychol Sci* 7: 237–241
14. Grob A, Petermann F, Lipsius M et al. (2008) Differences in Swiss and German children's intelligence as measured by the HAWIK-IV. *Swiss J Psychol* 67: 113–118
15. Hanes P (2000) Stabilität von Hochbegabung. In: Rost DH (Hrsg) Hochbegabte und hochleistende Jugendliche. Waxmann, Münster, S 93–159
16. Holocher-Ertl S, Kubinger KD, Hohensinn C (2008) Hochbegabungsdiagnostik: HAWIK-IV oder AID 2. *Kindheit Entwicklung* 17: 99–106
17. Horn JL, Noll J (1997) Human cognitive capabilities: Gf-Gc theory. In: Flanagan DP, Genshaft JL, Harrison PL (eds) Contemporary intellectual assessment: theories, tests, and issues. Guilford Press, New York, pp 53–91
18. Lienert GA, Raatz U (1998) Testaufbau und Testanalyse, 6. Aufl. Beltz PVU, Weinheim
19. Meyer MS (2000) The ability-achievement discrepancy: Does it contribute to an understanding of learning disabilities? *Educ Psychol Rev* 12: 315–337
20. Perlow R, Jattusa M, Moore DD (1997) Role of verbal working memory in complex skill acquisition. *Hum Perf* 10: 183–302
21. Petermann F, Petermann U (Hrsg) (2007) Hamburg Wechsler Intelligenztest für Kinder – IV (HAWIK-IV). Huber, Bern Stuttgart Toronto
22. Petermann F, Petermann U (2008) HAWIK-IV. *Kindheit Entwicklung* 17: 71–75
23. Plume E, Warnke A (2007) Definition, Symptomatik, Prävalenz und Diagnostik der Lese-Rechtschreib-Störung. *Monatsschr Kinderheilkd* 155: 322–327
24. Rosenkötter H, Axtmann S, Kühne H et al. (2007) Umschriebene Entwicklungsstörungen im SPZ. In: Fricke C, Kretzschmar C, Hollmann H et al. (Hrsg) Qualität in der Sozialpädiatrie. RS Verlag, Altötting, S 91–116
25. Rost DH, Sparfeldt JR, Schilling SR (2006) Hochbegabung. In: Schweizer K (Hrsg) Leistung und Leistungsdiagnostik. Springer, Berlin Heidelberg New York, S 187–222
26. Savage R, Lavers N, Pillay V (2007) Working memory and reading difficulties: what we know and what we don't know about the relationship. *Educ Psychol Rev* 19: 185–221
27. Schallberger U (1991) Das Ausmass des „IQ-Gewinns“ im deutschen Sprachraum von ca. 1956 bis ca. 1983. Zur Diskussion um die HAWIK-R-Normen. *Berichte aus der Abteilung Angewandte Psychologie* 27. Universität Zürich, Zürich.
28. Schmid RG (2004) Lese-Rechtschreibstörung. *Monatsschr Kinderheilkd* 152: 1111–1122
29. Schwan VL, Saklofske DH (2004) Assessment of attention deficit hyperactivity disorder with WISC-IV. In: Prifitera A, Saklofske DH, Weiss LG (eds) WISC-IV clinical use and interpretation. Elsevier, San Diego, pp 236–280
30. Stuebing KK, Fletcher JM, LeDoux JM et al. (2002) Validity of IQ-discrepancy classification of reading disabilities: a meta-analysis. *Am Educ Res J* 39: 469–518