

# AUSWAHL VON INTELLIGENZTESTS IN DER BEGABUNGSDIAGNOSTIK<sup>1</sup>

## ERFAHRUNGEN AUS DER PRAXIS

Traditionellen Leistungs- und insbesondere Intelligenztests mangelt es beim Einsatz in der Hoch- oder Minderbegabungsdiagnostik oft an ausreichend schwierigen oder leichten Items. In der entsprechenden Zielpopulation kommt es demnach oft zu Deckeneffekten oder Bodeneffekten und in der Folge zu wenig messgenauen Testwerten (vgl. z.B. Preckel, 2003). Dies ist zweifach problematisch. Erstens muss für eine verbindliche (Förder-)Beratung mehr als nur die ungefähre Feststellung des Fähigkeitsniveaus der untersuchten Person interessieren, v.a. wenn dieses in einen kritischen Grenzbereich fällt. Zweitens führen Decken- oder Bodeneffekte und der damit verbundene Umstand, dass ziemlich viele Items für die Testperson (zu) leicht oder (zu) schwer sind, also für sie zu wenig bzw. zu viel Herausforderung bedeuten, häufig zu reduzierter Leistungsmotivation.

Viele Intelligenztests erbringen die genauesten Messungen im mittleren Begabungsbereich, da der Anteil mittelschwerer Aufgaben im Vergleich zu sehr leichten oder sehr schweren Aufgaben vergleichsweise höher ist. Messungen im über- bzw. unterdurchschnittlichen Begabungsbereich sind oftmals stärker fehlerbehaftet. Sind die Testaufgaben zu schwer für eine Person, löst sie also kaum Aufgaben richtig, so zeigt sich ein sogenannter Bodeneffekt. Enthält ein Test keine oder zu wenige ausreichend schwere Aufgaben, ergibt sich ein sogenannter Deckeneffekt. Boden- bzw. Deckeneffekte verhindern die Abschätzung der wahren Fähigkeit der Person und führen zu ungenauen Schätzungen. Daher müssen bei der Diagnostik von Hoch- oder Minderbegabung Tests vorgegeben werden, die dem Fähigkeitsniveau der zu testenden Person entsprechen und eine verlässliche Aussage über deren intellektuelle Begabung erlauben, auch wenn diese im Extrembereich liegt.

In diesem Beitrag soll eine Auswahl der derzeit gängigen Verfahren zur Messung intellektueller Begabung im Hinblick auf ihre Eignung für die Hoch-, aber auch für die Minderbegabungsdiagnostik vorgestellt werden. Eine ausführliche Beschreibung aller gängigen und derzeit erhältlichen Testverfahren findet sich in Kapitel 3 der ÖZBF-Broschüre „Psychologische Testverfahren zur Messung intellektueller Begabung“ (siehe Infokästchen S. 19).

### ADAPTIVES INTELLIGENZ DIAGNOSTIKUM 2; AID 2

Das AID 2 bietet als einziger deutschsprachiger Test die Möglichkeit des adaptiven Testens, die gerade im Bereich der Hochbegabungsdiagnostik sehr viele Vorteile (Erhöhung der Messgenauigkeit, Verringerung der Testlänge, Vermeidung von Deckeneffekten, Optimierung

der Entscheidungsstrategie) mit sich bringt. Das im AID 2 umgesetzte Konzept bietet abweichend von der standardmäßigen Vorgabe eine attraktive Testadministration für den Fall an, dass zwischen „hochbegabt“ und „nicht hochbegabt“ unterschieden werden soll und eignet sich hervorragend gerade für die Testung (hoch)begabter Kinder.

### ADVANCED PROGRESSIVE MATRICES; APM

Bezogen auf die Hochbegabungsdiagnostik ist zu beachten, dass die APM nur eine Intelligenzfacette – und dies mit zudem nur einem Aufgabentyp – erfassen. Beim Einsatz mit hochbegabten Schülerinnen und Schülern findet sich bei den APM, trotz Konstruktion für begabtere Schüler/innen, ein deutlicher Deckeneffekt (Preckel, 2003). Bei Schülerinnen und Schülern in höheren Schulen der Klassenstufen 9 und 10 differenzieren die APM zudem nicht hinreichend (Heller, Kratzmeier & Lengfelder, 1998). Vielfach werden Geschlechtereffekte zugunsten männlicher Testpersonen dokumentiert (z.B. Irwing & Lynn, 2005). Dieses Phänomen lässt sich durch den Vorteil von männlichen Schülern im räumlichen Denken erklären (Colom, Escorial & Rebollo, 2004), welches positiv mit der Testleistung der APM korreliert ist (Schweizer et al, 2007). Geschlechterspezifische Normen werden bei den APM nicht angeboten, zudem sind die Normen beider Tests bereits mehr als 10 Jahre alt, was möglicherweise leichte Fähigkeitsüberschätzungen mit sich bringt („Flynn-Effekt“<sup>2</sup>). Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den APM vorgezogen werden. Der alleinige Einsatz figuraler Reasoning-Tests wie der APM ist für die Begabungsdiagnostik im Kontext schulischer Begabtenförderung nicht optimal (siehe auch Preckel, 2010).

### BERLINER-INTELLIGENZSTRUKTUR-TEST; BIS-4

Grundsätzlich eignet sich der BIS-4 sehr gut für eine umfassende Diagnose intellektueller Begabung auch im höheren Bereich. Im Spitzenbereich können jedoch auch hier höhere Messfehler auftreten. Zu bedenken ist bei einer Diagnose allerdings, dass zwar die Gesamtreliabilitäten weitestgehend zufriedenstellen (die Werte liegen zwischen .77 für die Merkfähigkeit und .93 für Generalfaktor g), die Konstruktrelia-bilitäten der spezifischen kognitiven Fähigkeiten unabhängig vom verwendeten Koeffizienten aber nicht zufriedenstellend sind (die Werte liegen zwischen .17 für die numerische Fähigkeit und .67 für die Verarbeitungskapazität; siehe Brunner & Süß, 2007). Aus diesem Grund sollte – wenn eine spezifische Diagnose gestellt werden und die Testung über eine allgemeine Einschätzung hinausgehen soll – ein weiterer Test durchgeführt werden.

<sup>1</sup> Ein Langtext dieses Beitrags findet sich in der Zeitschrift „Psychologie in Österreich“ (Nr. 2-3, 2011, S. 102–109).

<sup>2</sup> Der Flynn-Effekt (James R. Flynn) bezeichnet das Phänomen, dass bis in die 1990er Jahre die Ergebnisse von IQ-Tests in Industrieländern im Mittel immer höhere Werte erbrachten. Zurückgeführt wird dieses Phänomen großteils auf veränderte Umweltbedingungen (z.B. längere Schulbildung, gesteigerter Wohlstand). Der Flynn-Effekt macht eine regelmäßige Normierung von Intelligenztests notwendig.



## **BERLINER INTELLIGENZSTRUKTUR-TEST FÜR JUGENDLICHE: BEGABUNGS- UND HOCHBEGABTENDIAGNOSTIK; BIS-HB**

In einer Rezension folgern Frenzel & Nett (2008, S. 225): „Mit dem BIS-HB legen die Autoren einen empfehlenswerten Test vor, um Intelligenz auf der Basis eines theoretisch und empirisch ausgereiften Intelligenzmodells bei überdurchschnittlich bis weit überdurchschnittlich begabten Jugendlichen objektiv, reliabel und valide zu erfassen. Durch die Abdeckung eines breiten Fähigkeitsspektrums erlaubt der BIS-HB sowohl eine allgemeine als auch eine bereichsspezifische Begabungs- und Underachievementdiagnostik. Jedoch ist eine nicht unerhebliche Unterschätzung des IQ in Rechnung zu stellen.“ Dieser Meinung kann größtenteils beigepflichtet werden, zur Unterschätzung ist jedoch zu sagen, dass der Test zwar eher konservative Schätzungen liefert, es sich jedoch um keine „erhebliche Unterschätzung“ (wie von Frenzel & Nett, 2008, S. 225 postuliert) handelt. Für die Hochbegabungsdiagnostik ist er durchaus empfehlenswert.

### **GRUNDINTELLIGENZTEST; CFT 1; SKALA 1**

Praktische Erfahrungen mit dem CFT 1 weisen darauf hin, dass sich dieser Test eher für eine Intelligenzdiagnostik im mittleren als im höheren Begabungsbereich eignet. Etliche Items haben für begabtere Schüler/innen (IQ>120) einen geringeren Informationsgehalt als für weniger begabte Gruppen (niedrige Trennschärfe bei Begabten).

### **GRUNDINTELLIGENZTEST; CFT 20-R; SKALA 2**

Einige praktische Erfahrungen mit dem CFT 20-R weisen darauf hin, dass sich dieser Test eher für eine Intelligenzdiagnostik im mittleren als im höheren Begabungsbereich eignet. „Bei der Hochbegabungsdiagnostik hat sich in der Praxis häufig gezeigt, dass bei nach CFT 20-R hochbegabten Kindern bei IQ-Messung mit einem komplexeren Verfahren keine Hochbegabung vorlag.“ (Jacobs & Petermann, 2007, S. 113) Empirische Analysen mit Schülerinnen und Schülern der Klassenstufen 7 bis 9 belegen dieses auch für die Kurzform des CFT 20-R. Kuhn, Holling & Freund (2008) dokumentieren, dass etliche Testitems für begabtere Schüler/innen (IQ>120) einen geringeren Informationsgehalt haben als für weniger begabte Gruppen (niedrige Trennschärfe bei Begabten) und damit verbunden die Messgenauigkeit der Subtests für Begabte verringert ist. Positiv zu bewerten ist, dass fast alle Testaufgaben in allen Begabungsgruppen eine vergleichbare relative Schwierigkeit aufweisen und sich Bestätigung für die Strukturannahme eines g-Faktors in allen Fähigkeitsgruppen finden lässt. Testanwender/innen müssen beim Einsatz des CFT 20-R für die Hochbegabungsdiagnostik mit relativ großen Messfehlern rechnen.

### **COLUMBIA MENTAL MATURITY SCALE; CMM 1-3 UND 1-4**

Die CMM 1-3 und 1-4 eignet sich nur bedingt zur Testung von intellektueller (Hoch-)Begabung; vielmehr kann eine Minderbegabung oder eine Begabung im unteren Durchschnittsbereich mit großer Wahrscheinlichkeit festgestellt werden. Deckeneffekte treten schon bei Kindern mit einem IQ>120 auf. Die

## **PSYCHOLOGISCHE TEST-VERFAHREN ZUR MESSUNG INTELLEKTUELLER BEGABUNG**

In der ÖZBF-Broschüre „Psychologische Testverfahren zur Messung intellektueller Begabung“ werden nach einem Überblick über verschiedene Intelligenztheorien und -modelle (z.B. Spearman, Thurstone, Cattell, Jäger, Carroll, Sternberg) und einer Klärung der wichtigsten Begriffe im Zusammenhang mit dem Testen von Intelligenz (u.a. Normierung, Testgüte) die derzeit gängigen Intelligenztestverfahren ausführlich beschrieben.

Neben einer kurzen Testbeschreibung werden Angaben zur Durchführung sowie zur Auswertung und Interpretation dargeboten (inklusive Testgütekriterien). Auf die Eignung für bestimmte Anwendungsgebiete (Hochbegabung, Minderbegabung, Underachievement) wird gesondert eingegangen. Ein kurzes Kapitel zum Thema „Adaptives Testen und Automatische Itemgenerierung“ schließt die Broschüre ab. Im Anhang finden sich neben einem Überblickskatalog mehrere Mustergutachten zu den Themenbereichen vorzeitige Einschulung, Überspringen einer Klasse und allgemeine Intelligenzdiagnostik.

KIPMAN, U., KOHLBÖCK, G. & WEILGUNY, W. (2011). Psychologische Testverfahren zur Messung intellektueller Begabung (im Druck). ÖZBF: Eigenverlag.

Diese Broschüre wird Ende 2011 veröffentlicht und wird als Download auf der Website des ÖZBF allen interessierten Leserinnen und Lesern zur Verfügung stehen.

Aufgaben sind zudem wenig abwechslungsreich, weshalb Ermüdungseffekte nicht auszuschließen sind.

#### **HAMBURG-WECHSLER INTELLIGENZTEST FÜR ERWACHSENE; HAWIE-R**

Dem HAWIE-R sind m.E. andere Verfahren der Intelligenzmessung vorzuziehen. Zwar sind die Testaufgaben gut normiert und abwechslungsreich, dennoch kommen Testpersonen mit anderen Testmaterialien oft besser zurecht.

#### **HAMBURG-WECHSLER-INTELLIGENZTEST FÜR KINDER III; HAWIK-III**

Der HAWIK-III ist ein Standardinstrument der allgemeinen Intelligenzdiagnostik. Er erfährt einen weiten Anwendungsbereich, seine Differenzierungsstärke liegt allerdings im mittleren Leistungsbereich. Somit ergeben sich in den Extrembereichen der kognitiven Leistungsfähigkeit (Geistige Behinderung – Hochbegabung) Boden- bzw. Deckeneffekte. Das heißt, die Items sind größtenteils zu schwer oder zu leicht, daher können keine präzisen Aussagen mehr getroffen werden. Im überarbeiteten Manual werden verschiedene Kurzformen vorgestellt, die sich allerdings nur im Rahmen eines Screenings zur Einschätzung des Gesamt-IQs eignen. In der klinischen Praxis ist eine solche Anwendung möglich, wenn das allgemeine Begabungsniveau als Bezugspunkt für eine Diagnosestellung notwendig ist. Dies ist zum Beispiel für die Abklärung von umschriebenen Entwicklungsstörungen schulischer Fertigkeiten der Fall.

#### **HAMBURG-WECHSLER-INTELLIGENZTEST FÜR KINDER IV; HAWIK-IV**

Relevant für die Hochbegabungsdiagnostik ist, dass im Bereich hoher Begabung nicht selten extreme Leistungsunterschiede zwischen den 4 Indexwerten bzw. den beiden Indexpaaren (Sprachverständnis – logisches Denken vs. Arbeitsgedächtnis – Verarbeitungsgeschwindigkeit) vorkommen, was die Interpretierbarkeit des Gesamt-IQ erschwert (Daseking, Petermann & Petermann, 2007). Zudem zeigt sich in Validierungsstudien zur Hochbegabung, dass Kinder, welche in anderen Verfahren einen  $IQ > 130$  erreichten, im HAWIK-IV einen niedrigeren IQ erreichten. Bei Hochbegabten empfiehlt sich daher die Verwendung des Allgemeinen Fähigkeitsindex alternativ zum Gesamt-IQ, da für zwei Drittel der Begabten ( $IQ > 120$ ) der AFI höher ausfällt als der Gesamt-IQ. Der Zusammenhang mit der fluiden Intelligenz<sup>3</sup> scheint hier besser gegeben zu sein. Bei den Aufgaben zum Arbeitsgedächtnis können Deckeneffekte nachgewiesen werden, da die Aufgaben für ältere hochbegabte Kinder/Jugendliche nicht hinreichend schwer sind.

#### **HANNOVER WECHSLER INTELLIGENZTEST FÜR DAS VORSCHULALTER III; HAWIVA-III**

Die Subtests des HAWIVA-III erlauben eine differenzierte Diagnose sprachlich-kognitiver Fähigkeiten, also der kristallinen Intelligenz. Dies lässt sich aus den geringen Interkorrelationen mit dem CFT1-Gesamtscore und den relativ höheren Korrelationen mit Subtests der Fertigkeitenskala aus der K-ABC ableiten. Zugleich lassen diese Validitätsbefunde aber den Schluss zu, dass ein Ziel der Testrevision, nämlich eine verbesserte Erfassung der fluiden Intelligenz, nicht erreicht wurde (vgl. Kastner-Koller & Deimann (2008)).

#### **INTELLIGENZ-STRUKTUR-TEST 2000R; I-S-T 2000R**

Der I-S-T 2000 R ist zwar zeitaufwändig, erlaubt aber eine verlässliche, umfassende und objektive Aussage über die intellektuellen Fähigkeiten der getesteten Personen und ist zur Diagnostik auch im oberen Begabungsbereich zu empfehlen.

#### **KAUFMAN-ASSESSMENT BATTERY FOR CHILDREN; K-ABC**

Die K-ABC wird nach einer Umfrage aus dem Jahr 2009 an verschiedenen begabungspsychologischen Stellen in Österreich und Deutschland häufig für die Intelligenzdiagnostik von Hochbegabung eingesetzt (Weiß, 2008), obwohl die Normen mit mehr als 20 Jahren deutlich veraltet sind. Möglicherweise kann die Einsatzhäufigkeit der K-ABC dadurch erklärt werden, dass es für den Vorschulbereich nur sehr wenige Verfahren gibt. Die K-ABC ist im Hinblick auf die theoretische Fundierung und die Anwenderfreundlichkeit als sehr gut zu bewerten, die Gütekriterien sind hinreichend erfüllt, beziehen sich jedoch nicht auf aktuelle Stichproben. Das Itemmaterial ist zum Teil veraltet (Horn, 2003). Studien zur Bewertung der Eignung der K-ABC zum Zweck der Feststellung von Hochbegabung gibt es bislang keine. In einer vergleichenden Bewertung von K-ABC, HAWIK-III und AID wird jedoch die Stärke der K-ABC eher im Bereich der Diagnostik von Minderbegabung gesehen (Preusche & Leiss, 2003).

#### **KOGNITIVER FÄHIGKEITSTEST (KINDERGARTEN); KFT-K**

Der KFT-K ist ein geeignetes Instrument zur Feststellung der kognitiven Fähigkeiten von Kindergartenkindern und eignet sich auch gut zur Beantwortung der Frage nach einer früheren Einschulung. Es sollte unbedingt ein bis zwei Jahre später ein weiterer Test durchgeführt werden, da eine verlässliche Aussage über die intellektuelle Begabung in diesem Alter noch schwer möglich ist.

<sup>3</sup> Fluide Intelligenz bezeichnet die Fähigkeit, sich neuen Situationen anzupassen und neuartige Probleme zu lösen, ohne dass gelerntes Wissen eine bedeutende Rolle spielt. Es wird angenommen, dass die fluide Intelligenz weitgehend von Geburt an angelegt und von kulturellen und gesellschaftlichen Einflüssen unabhängig ist. Die kristalline Intelligenz bezeichnet kognitive Fertigkeiten, die durch die Kumulierung (oder „Kristallisierung“) von Lernerfahrungen seit der Geburt entwickelt wurden. Das Konzept der fluiden und kristallinen Intelligenz basiert auf Raymond Cattell.

### KOGNITIVER FÄHIGKEITS-TEST (GRUNDSCHULFORM); KFT 1-3

Der KFT 1-3 eignet sich gut für die Erstellung von Intelligenzdiagnosen im Durchschnittsbereich. In Bereichen über 1,5 Standardabweichungen über dem Mittelwert sind Deckeneffekte zu beobachten.

### KOGNITIVER FÄHIGKEITSTEST; KFT 4-12+R

Die MHBT enthält eine Version des KFT zur Hochbegabungsdiagnostik. Die nicht speziell für die Hochbegabungsdiagnostik vorgesehenen KFT-Versionen, insbesondere der KFT 4-12+R, werden ebenfalls für die Hochbegabungsdiagnostik eingesetzt (Weiß, 2008). Jedoch sollte in diesem Bereich den KFT-Versionen aus dem MHBT aufgrund der höheren Aufgabenschwierigkeit und den aktuelleren Normdaten der Vorzug gegeben werden.

### MÜNCHNER HOCHBEGABUNGSTESTBATTERIE FÜR DIE PRIMARSTUFE; MHBT-P

### MÜNCHNER HOCHBEGABUNGSTESTBATTERIE FÜR DIE SEKUNDARSTUFE; MHBT-S

Gerade im Bereich über dem Durchschnitt und im oberen Extrembereich eignet sich die MHBT sehr gut für die Erstellung einer verlässlichen Intelligenzdiagnostik. Deckeneffekte treten nicht auf, es gibt hinreichend schwere Aufgaben und die Messgenauigkeit ist deutlich höher als bei anderen Verfahren zur Messung intellektueller Fähigkeiten.

Das Material ist ansprechend und abwechslungsreich. Die Kinder haben erfahrungsgemäß Spaß an den Aufgaben, lösen diese gerne und langweilen sich nicht.

### NON-VERBALER INTELLIGENZTEST; SON-R 2 1/2-7

Der SON-R 2 1/2-7 ist eines der wenigen Intelligenztestverfahren für das Vorschulalter im deutschsprachigen Raum. Er ist solide konstruiert, die Gütekriterien wurden sorgfältig geprüft und die Normen sind aktuell. Für die Hochbegabungsdiagnostik relevant ist, dass der SON-R 2 1/2-7 insbesondere für die Anwendung mit sprachbeeinträchtigten, entwicklungsverzögerten oder schwach begabten Kindern konstruiert ist. Daher sind Deckeneffekte bei der Testung – vor allem älterer – hochbegabter Kinder zu erwarten. Beim Einsatz mit jüngeren Kindergartenkindern ist zudem zu beachten, dass in diesem Alter erhobene Intelligenzschätzungen eher instabil sind und eine Hochbegabungsdiagnostik vor einem Alter von 5 Jahren wenig sinnvoll ist (Rost, 2000). Wenn es um Einschätzung bei Fragen der (früheren) Einschulung geht, sollte zusätzlich ein Test zur Erfassung sprachlicher Fähigkeiten eingesetzt werden. Weil räumliche Fähigkeiten in der Schule kaum benötigt werden, ist nach Vock (2008) die Handlungsskala zur Abschätzung des Schulerfolgs wahrscheinlich nur wenig aussagekräftig. Ein Vorzug des



Foto: Christina Klafinger

SON-R 2 1/2-7 ist, dass der Test auch rein nonverbal instruiert werden kann, so dass man ihn bei Kindern mit Migrationshintergrund oder auch gehörlosen Kindern problemlos einsetzen kann.

### NON-VERBALER INTELLIGENZTEST; SON-R 5 1/2-17

Auch in der Altersvariante für das Schulalter liegt mit dem SON-R 5 1/2-17 ein Diagnostikum vor, das für den deutschen Sprachraum in einzigartig differenzierter Weise sprachfreie Intelligenzanteile erfasst. Kritisch zu vermerken ist jedoch der Umstand, dass eine vertiefende Auswertung vom Anwender ein über das übliche Maß hinaus gehendes statistisches Hintergrundwissen erfordert. Zwar dürfte in der Mehrzahl der Fälle eine einfache Auswertung ausreichend sein, wenn jedoch Leistungsschwankungen interpretiert werden müssen, wirkt das hier umgesetzte Bemühen um Exaktheit der Praktikabilität entgegen. Wie auch bei der Testversion für das Vorschulalter, liegen hier bislang nur Normen aus den Niederlanden vor. Schwerer wiegen dürfte jedoch das Alter der Normen (1984/85), so dass vor dem Hintergrund des Flynn-Effekts aktuell eine faire Beurteilung erschwert scheint. Dem ist jedoch gegenüber zu stellen, dass im Gegensatz zu den sprachfreien Teilbereichen anderer allgemeiner Intelligenztests mit dem SON-R eine inhaltlich validere Überprüfung sprachfreier Intelligenzanteile vorgenommen wird. Zur Diagnostik von Hochbegabung ist der SON-R 5 1/2-7 alleine eher nicht zu empfehlen.



## VERANSTALTUNGS- KALENDER

Das ÖZBF informiert auf seiner Homepage regelmäßig und aktuell über Veranstaltungen zum Thema Begabtenförderung und Begabungsforschung. Leser/innen können sich aktuell über Termine von Kongressen & Tagungen sowie österreichischen Sommerakademien, Olympiaden und Schüler/innen-Wettbewerben informieren.

[www.begabtenzentrum.at](http://www.begabtenzentrum.at) > Kalender 

The screenshot shows the website interface for the 'Veranstaltungskalender' (Event Calendar) for November 2011. The calendar grid highlights the 11th of the month. To the right, a list of events is displayed:

Datum	Termin / Ereignis
04.11.2011	100th Geburtstag von Otto von Guericke
08.11.2011	100th Geburtstag von Otto von Guericke
10.11.2011	10.11.2011 - 11.11.2011
21.11.2011	100th Geburtstag von Otto von Guericke
22.11.2011	100th Geburtstag von Otto von Guericke
21.11.2011	100th Geburtstag von Otto von Guericke
22.11.2011	100th Geburtstag von Otto von Guericke
26.11.2011	100th Geburtstag von Otto von Guericke
29.11.2011	100th Geburtstag von Otto von Guericke

### STANDARD PROGRESSIVE MATRICES; SPM

Bezogen auf die Hochbegabungsdiagnostik ist zu beachten, dass die SPM, wie auch die APM, nur eine Intelligenzfacette mit nur einem Aufgabentyp erfassen. Beim Einsatz mit hochbegabten Schülerinnen und Schülern findet sich bei den SPM noch mehr als bei den APM ein deutlicher Deckeneffekt. Geschlechtereffekte zugunsten männlicher Testpersonen sind dokumentiert. Kubinger & Farkas (1991) stellten anhand von Analysen nach dem Rasch-Modell (mathematisch-psychologisches Modell der probabilistischen Testtheorie von Georg Rasch) zudem fest, dass von insgesamt 60 Aufgaben 43 nicht eindimensional zu den anderen skaliert sind. Um die testspezifische Fehlervarianz gering zu halten, sollen Tests mit unterschiedlichen Aufgabenfacetten den SPM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der SPM ist für die Begabungsdiagnostik im Kontext schulischer Begabtenförderung nicht optimal (siehe Preckel, 2010).

### WILDE-INTELLIGENZ-TEST; WIT 2

Der WIT ist ein brauchbares Verfahren zur Messung von intellektuellen Fähigkeiten. Am besten misst er im Durchschnittsbereich bis zu Abweichungen von 1,5 Standardabweichungen vom Mittelwert.

### WECHSLER NONVERBAL SCALE OF ABILITY; WNV

Die WNV eignet sich sehr gut zur Abschätzung der intellektuellen Fähigkeiten im Durchschnittsbereich und auch für Intelligenzquotienten bis zu 2 Standardabweichungen über dem Mittelwert, im oberen Extrembereich können jedoch Deckeneffekte auftreten. Grundsätzlich kann dieser Test zur Testung der allgemeinen Intelligenz (auch wenn das Einfaktorenmodell empirisch (noch) nicht bestätigt wurde) und spezifischen Intelligenz empfohlen werden. Für den Extrembereich ist es ratsam, noch weitere Verfahren einzusetzen, da hier die Messfehler größer werden.

### ZUSAMMENFASSUNG UND AUSBLICK

Den richtigen Test zur Intelligenzdiagnostik auszuwählen ist – besonders wenn es um eine verbindliche (Förder-)Beratung geht – essentiell. Es muss, vor allem wenn kritische Grenzbereiche abgetestet werden sollen, mehr als nur die ungefähre Feststellung des Fähigkeitsniveaus der untersuchten Person interessieren. Decken- und Bodeneffekte sollten ausgeschlossen werden, um eine verlässliche Aussage über die Intelligenz der getesteten Person machen zu können, zudem führen zu viele leichte bzw. schwere Items häufig zu reduzierter Leistungsmotivation. In der Praxis bewährt sich in solchen Fällen immer wieder das adaptive Testen: Jeder Testperson werden im Wesentlichen nur diejenigen Items eines Tests (Itempools) gestellt, welche ihrem Leistungsniveau entsprechen, wie es derzeit im deutschsprachigen Raum einzig der AID-2 standardisiert anbietet.

## LITERATUR

- Brunner, M. & Süß, H. M. (2007). Wie genau können kognitive Fähigkeiten gemessen werden? Die Unterscheidung von Gesamt- und Konstruktrelia- bilitäten in der Intelligenzdiagnostik für den Berliner Intelligenzstruktur- test. *Diagnostica*, 53, Heft 4, 184-193.
- Colom, R., Escorial, S. & Rebollo, I. (2004). Sex differences on the pro- gressive matrices are influenced by sex differences on spatial ability. *Personality and Individual Differences*, 37, 1289-1293.
- Daseking, M., Petermann, F. & Waldmann, H.-C. (2008). Der allgemeine Fähigkeitsindex (AFI) – eine Alternative zum Gesamt-Intelligenzquoti- enten (G-IQ) des HAWIK-IV? *Diagnostica*, 54 (4), 211-222.
- Daseking, M., Petermann, U. & Petermann, F. (2007). Intelligenzdiagnos- tik mit dem HAWIK-IV. *Kindheit und Entwicklung*, 16 (4), 250-259.
- Frenzel, A. C. & Nett, U. (2008). Berliner Intelligenzstrukturtest für Ju- gendliche: Begabungs- und Hochbegabungsdiagnostik (BIS-HB) von A. O. Jäger et al. *Diagnostica*, 54, 221-225.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998). APM. Matrizen-Test- Manual Band 2. Göttingen: Beltz-Test.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998). Matrizen Test Ma- nual, Bd. 2, Ein Handbuch mit deutschen Normen zu den SPM. Wein- heim: Beltz.
- Horn, R. (2003). Eine kritische Anmerkung zum K-ABC. *Report Psycho- logie*, 28, 189.
- Irwing, P. & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology*, 96, 505-524.
- Jacobs, C. & Petermann, F. (2007). Besprechung des Grundintelli- genztests (CFT-20-R). *Diagnostica* 53, 2, 109-113.
- Kastner-Koller, U. & Deimann, P. (2008). HAWIVA-III. Hannover-Wechs- ler-Intelligenztest für das Vorschulalter-III. *Zeitschrift für Entwicklungs- psychologie und Pädagogische Psychologie*, 40 (1), 49-53.
- Kubinger, K. D. (2003). Adaptives Testen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 1-9). Wein- heim: Beltz/PVU.
- Kubinger, K. D. (2006). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Kubinger K. D. & Farkas, M. G. (1991). Die Brauchbarkeit der Normen von Papier-Bleistift-Tests für die Computer-Vorgabe: Ein Experiment am Bei- spiel der SPM von Raven als kritischer Beitrag. *Zeitschrift für Differenti- elle und Diagnostische Psychologie*, 12, 257-266.
- Kubinger K. & Holocher-Ertl, S. (2010). Die Vorteile adaptiven Testens in der Hochbegabungsdiagnostik. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung* (S. 197-211). Göttingen: Hogrefe.
- Kubinger, K. D. & Wurst, E. (2000). AID 2. *Adaptives Intelligenz Diagnos- tikum 2*. Göttingen: Beltz.
- Kubinger, K. D. & Wurst, E. (1991). AID. *Adaptives Intelligenz Diagnosti- kum* (3., ergänzte Auflage). Weinheim: Beltz.
- Kuhn, J. T., Holling, H. & Freund, A. (2008). Begabungsdiagnostik mit dem Grundintelligenztest (CFT 20R). *Psychometrische Eigenschaften und Messäquivalenz*. *Diagnostica*, 54, 184-192.
- Mittring, G. & Rost, D. (2008). Die verflixten Distraktoren: Über den Nut- zen einer theoretischen Distraktorenanalyse bei Matrizen-Tests (für bes- ser Begabte und Hochbegabte). *Diagnostica*, 54, Heft 4, 193–201.
- Preckel, F. (2003). *Diagnostik intellektueller Hochbegabung*. Testentwick- lung zur Erfassung der fluiden Intelligenz. Göttingen: Hogrefe.
- Preckel, F. (2010). Intelligenztests in der Hochbegabungsdiagnostik. In F. Preckel, W. Schneider & H. Holling, *Diagnostik von Hochbegabung – Tests und Trends* (S. 19–44). Göttingen: Hogrefe.
- Preusche, I. & Leiss, U. (2003). Intelligenztests für Kinder. HAWIK-III, AID 2 und K-ABC im Vergleich. *Report Psychologie*, 28, 12-26.
- Radford, J. (1966). Verbalisation effects in a „non-verbal“ intelligence test. *The British Journal of Educational Psychology*, 36 (1), 33-38.
- Rost, D. H. (2000). *Hochbegabte und hochleistende Jugendliche*. Mün- ster: Waxmann.
- Schweizer, K., Goldhammer, F., Rauch, W. & Moosbrugger, H. (2007). On the validity of Ravens matrices test. *Personality and Individual Diffe- rences*, 43, 1998-2010.
- Vock, M. (2008). Non-verbaler Intelligenztest (SON R 2 1/2-7). *Diagnosti- ca*, 54, 112-115.
- Weiß, C. (2008). *Hochbegabtenberatung: Konzepte und Evaluation*. Un- veröffentlichte Diplomarbeit. Institut für Pädagogische Psychologie, Ju- lius-Maximilians-Universität Würzburg.

MMAG. DDR. ULRIKE KIPMAN, B.SC.  
ÖZBF  
ulrike.kipman@begabtenzentrum.at



### Zur Person

MMag. DDR. Ulrike Kipman,  
B.Sc.

Studium der Psychologie,  
Mathematik, Pädagogik  
und der Rechtswissen-  
schaften

Postgraduateausbildungen  
zur Klinischen Psychologin, Gesundheitspsychologin und zur  
Arbeitspsychologin

Fachgebiete:

Klinische Psychologie (Diagnostik), Familien- und Kinderpsy-  
chologie (Diagnostik), Mathematische Psychologie (Statistik,  
Methoden) sowie Betriebs- und Organisationspsychologie  
(Analysen)