

Psychologische Testverfahren zur Messung intellektueller Begabung

Kipman | Kohlböck | Weilguny



Psychologische Testverfahren zur Messung intellektueller Begabung

MMag. DDr. Ulrike Kipman, B.Sc.

Mag. Dr. Gabriele Kohlböck

Mag. Dr. Walburga Weilguny

Hrsg.: Österreichisches Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF)

Herausgeber
Österreichisches Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF)

Schillerstraße 30, Techno 12
A-5020 Salzburg
Internet: www.begabtenzentrum.at

Tel.: +43/662-439581
Fax: +43/662-439581-310
e-mail: info@begabtenzentrum.at

© 2012

Lektorat/Korrektorat: MMag. Dr. Claudia Resch, Mag. Dr. Astrid Fritz
Grafik/Layout: Mag. Christina Klaffinger
Druck: Laber Druck
Eigenverlag: Österreichisches Zentrum für Begabtenförderung und Begabungsforschung

ISBN: 978-3-9502683-8-6

Alle Inhalte sind urheberrechtlich geschützt. Die Nutzungsrechte liegen bei den Autorinnen und dem ÖZBF.

Inhaltsverzeichnis

Vorwort des Herausgebers	8
1 Intelligenz und Begabung	10
1.1 Intelligenztheorien und -modelle	11
1.1.1 Zwei-Faktoren- bzw. Generalfaktoren-Theorie von Spearman	11
1.1.2 Modell mehrerer gemeinsamer Faktoren von Thurstone	12
1.1.3 Theorie der fluiden und kristallinen Intelligenz von Cattell	13
1.1.4 Berliner Intelligenzstrukturmodell von Jäger	15
1.1.5 Carrolls Three-Stratum-Theorie	18
1.1.6 Sternbergs Triarchische Theorie der Intelligenz	20
1.1.7 Zusammenfassung	21
1.2 Hochbegabungstheorien und -modelle	22
1.2.1 Münchner Hochbegabungsmodell	22
1.2.2 Aktiotop-Modell	23
1.2.3 Zusammenfassung	25
2 Auswahl von Intelligenztests für die Begabungsdiagnostik	28
2.1 Objektivität	30
2.2 Reliabilität	30
2.3 Validität	30
2.4 Normierung	31
3 Verfahren zur Messung von intellektuellen Fähigkeiten	34
3.1 AID 2; Adaptives Intelligenz Diagnostikum 2	36
3.2 APM; Advanced Progressive Matrices	40
3.3 AzN 4+; Aufgaben zum Nachdenken 4+	44
3.4 BIS-4; Berliner-Intelligenzstruktur-Test	46
3.5 BIS-HB; Berliner Intelligenzstruktur-Test für Jugendliche: Begabungs- und Hochbegabtendiagnostik	49
3.6 BIVA; Bildbasierter Intelligenztest für das Vorschulalter	52
3.7 BT 1-2; Bildertest 1-2	56
3.8 BT 2-3; Bildertest 2-3	58

3.9	CFT 1; Grundintelligenztest Skala 1	60
3.10	CFT 20-R; Grundintelligenztest Skala 2	63
3.11	CMM 1-3; Columbia Mental Maturity Scale 1-3	66
3.12	CMM 1-4; Columbia Mental Maturity Scale 1-4	67
3.13	CPM; Coloured Progressive Matrices	70
3.14	FAT 6-9; Frankfurter Analogietest 6-9	74
3.15	FRT bzw. FRT-J; Figure Reasoning Test	76
3.16	HAWIE-R; Hamburg-Wechsler-Intelligenztest für Erwachsene	78
3.17	HAWIK-III; Hamburg-Wechsler-Intelligenztest für Kinder III	80
3.18	HAWIK-IV; Hamburg-Wechsler-Intelligenztest für Kinder IV	83
3.19	HAWIVA-III; Hannover-Wechsler-Intelligenztest für das Vorschulalter III	87
3.20	HIT 1-2; Heidelberger Intelligenztest 1-2	91
3.21	HIT 3-4; Heidelberger Intelligenztest 3-4	93
3.22	IBF; Intelligenz-Basis-Faktoren	95
3.23	IDS; Intelligence and Development Scales	97
3.24	ISA; Intelligenz Struktur Analyse	101
3.25	I-S-T 2000 R; Intelligenz-Struktur-Test 2000 R	103
3.26	K-ABC; Kaufman-Assessment Battery for Children	105
3.27	KFT 1-3; Kognitiver Fähigkeits-Test (Grundschulform)	109
3.28	KFT 4-12+R; Kognitiver Fähigkeitstest 4-12+R	111
3.29	KFT-K; Kognitiver Fähigkeits-Test (Kindergarten)	115
3.30	KLI 4-5 / KLI 4-5 R; Kombiniertes Lern- und Intelligenztest für 4. und 5. Klassen	117
3.31	K-TIM; Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene	119
3.32	MHBT-P; Münchner Hochbegabungstestbatterie für die Primarstufe	121
3.33	MHBT-S; Münchner Hochbegabungstestbatterie für die Sekundarstufe	124

3.34	MIT; Mannheimer Intelligenztest	127
3.35	MWT-A; MWT-B; Mehrfachwahl-Wortschatz-Test	129
3.36	NNAT; Naglieri Nonverbal Ability Test	131
3.37	PPVT; Peabody Picture Vocabulary Test	133
3.38	PSB-R 4-6; Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen	135
3.39	PSB-R 6-13; Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen	138
3.40	SCH-F; Schlauchfiguren	141
3.41	SON-R 2½-7; Snijders-Oomen Non-verbaler Intelligenztest	143
3.42	SON-R 5½-17; Snijders-Oomen Non-Verbaler Intelligenztest	145
3.43	SPM; Standard Progressive Matrices	148
3.44	WIE; Wechsler Intelligenztest für Erwachsene	151
3.45	WIT-2; Wilde-Intelligenz-Test 2	155
3.46	WMT; Wiener Matrizen-Test	159
3.47	WNV; Wechsler Nonverbal Scale of Ability	161
3.48	WPPSI-III; Wechsler Preschool and Primary Scale of Intelligence	164
3.49	WST A-P; Wortschatztest – aktiv und passiv	167
3.50	ZVT; Zahlen-Verbindungs-Test	168
4	Spezielle Aspekte in der Hochbegabungsdiagnostik	172
	Gesamt-Literaturverzeichnis	176
	Anhang 1 – Mustergutachten	188
	Vorzeitige Einschulung	190
	Überspringen einer Klasse	199
	Intelligenz (Eignungsuntersuchung)	206
	Anhang 2 – Testübersicht	214

Abbildungsverzeichnis

Abb. 1: Beispielaufgaben zur Erfassung von g_r (nach Penrose & Raven, 1936) (entnommen aus Holling, Preckel & Vock (2004), S. 22)	18
Abb. 2: Beispielaufgaben zur Erfassung von g_c (entnommen aus Holling, Preckel & Vock (2004), S. 23)	18
Abb. 3: Berliner Intelligenzstrukturmodell nach Jäger (entnommen aus Holling, Preckel & Vock (2004), S. 24)	21
Abb. 4: Beispielaufgaben zur Erfassung der Zellen KF, MN und EV (entnommen aus Holling, Preckel & Vock (2004), S. 26)	22
Abb. 5: Three-Stratum-Theorie nach Carroll (1993) (entnommen aus Holling, Preckel & Vock (2004), S. 28)	23
Abb. 6: Münchner Hochbegabungsmodell (entnommen aus Joswig, 2000, S. 23; Bearbeitung: Margarethe Kainig-Huber)	27
Abb. 7: Komponenten eines Aktiotops (Ziegler, 2009, S. 15)	28
Abb. 8: Gängige Normskalen im Vergleich (entnommen aus Holling, Preckel & Vock (2004), S. 83)	36



Vorwort des Herausgebers (ÖZBF)

Liebe Leserin! Lieber Leser!

Intelligenz, als Teilaspekt von Begabung, steht oft im Mittelpunkt einer individuellen Förderdiagnostik. Angesichts der großen Vielfalt an verfügbaren Verfahren zur Testung von Intelligenz stellen sich dabei immer wieder folgende Fragen: Wie gut eignen sich verschiedene Testverfahren für spezifische Fragestellungen im Bereich der Intelligenzabklärung? Welche Tests eignen sich zur Diagnostik von Hochintelligenz? Wie verlässlich und valide sind die am Markt gängigen Testverfahren und auf welchen Intelligenzkonzepten basieren sie?

Auf diese Fragen soll mit der vorliegenden Broschüre eingegangen werden. Das Kernstück dieser Broschüre widmet sich den derzeit gängigen Intelligenztestverfahren und deren Stärken und Schwächen in bestimmten Bereichen. Es werden verschiedene Intelligenztheorien und –modelle vorgestellt sowie allgemeine Informationen zu Testgütekriterien und Normierung angegeben. Drei Mustergutachten im Anhang beschreiben exemplarische Fälle der Begabungsdiagnostik.

Die Broschüre liefert damit wertvolle Informationen, die den Verantwortlichen helfen können, die geeigneten Testverfahren für ihre Fragestellung zu wählen und die Testergebnisse auch in Bezug auf das zugrundeliegende Intelligenzmodell richtig beurteilen zu können. Ein besonderes Augenmerk dieser Broschüre liegt auf der Frage, welche Verfahren sich am besten für die Diagnose von hoher Intelligenz bzw. Begabung eignen.

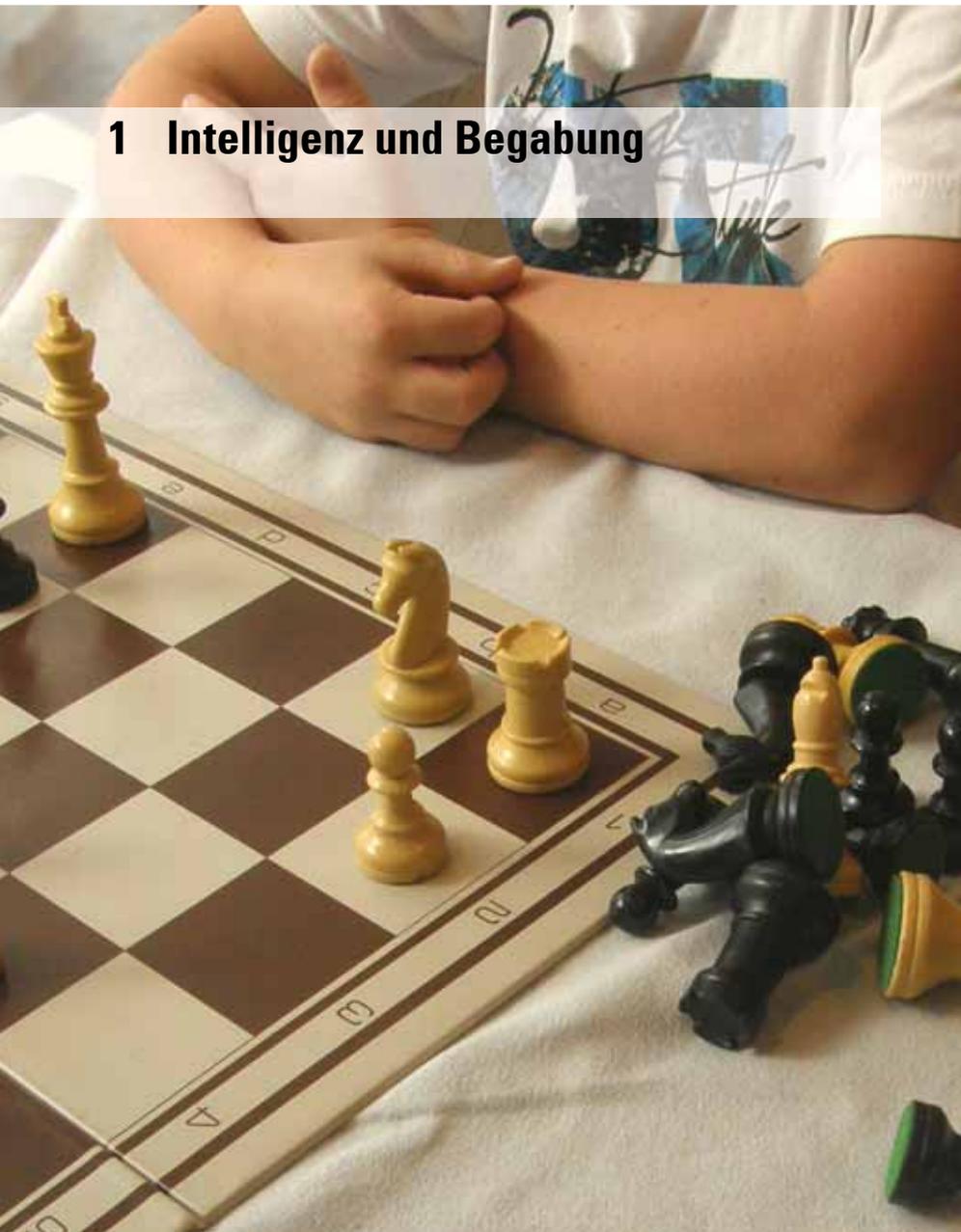
Wir möchten all jenen danken, die durch ihre Mitwirkung diese Broschüre erst ermöglicht haben. An erster Stelle sind hier die Kinder, Jugendlichen und Erwachsenen zu nennen, die durch ihre Teilnahme an den Tests diese Broschüre mit Leben befüllen. Einen Dank möchten wir auch den Eltern der Kinder und Jugendlichen aussprechen, die den Testungen zugestimmt haben sowie den Helferinnen und Helfern, die die Testdurchführungen unterstützten und auch bei den Analysen der Testdaten aktiv mitarbeiteten. Unser Dank gilt auch den engagierten Reviewerinnen, die durch ihre professionelle Arbeit die hohe Qualität dieser Broschüre sicherstellten.

Die Beschreibungen der Intelligenztheorien und -modelle in Kapitel 1 basieren maßgeblich auf dem Buch „Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6“ (2004) von Holling, Preckel & Vock. Den Autorinnen und dem Autor Heinz Holling, Franzis Preckel und Miriam Vock möchten wir herzlich danken, dass wir diese Beschreibungen in Teilen übernehmen durften.

Einen großen Dank möchten wir zum Schluss dem Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) an der Universität Trier aussprechen. Das ZPID betreibt die Referenzdatenbank PSYNDEX Tests.

Das Österreichische Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF) wünscht Ihnen eine spannende Lektüre!

1 Intelligenz und Begabung



1.1 Intelligenztheorien und -modelle

Für die Praxis der Intelligenzdiagnostik, insbesondere im Rahmen der Auswahl bestimmter Intelligenztests für konkrete diagnostische Fragestellungen und die Interpretation von Testergebnissen, ist die Kenntnis des Intelligenzmodells, auf dem der Test aufbaut, unerlässlich. Mit verschiedenen Testverfahren erhält man nicht selten unterschiedliche Ergebnisse für ein- und dieselbe Person. Diese werden in der Regel als Intelligenzquotient (IQ) ausgedrückt, was bei oberflächlicher Betrachtung nahe legt, es handle sich um dieselben erfassten Fähigkeiten. Intelligenztests bauen jedoch auf verschiedenen theoretischen Konzepten auf und verwenden unterschiedliches Aufgabenmaterial. Daraus lässt sich auch die geringe bis moderate Korrelation von Intelligenz(IQ)-Tests untereinander erklären. Um die mit dem Intelligenztest erhaltenen Messwerte inhaltlich füllen zu können, muss man demnach das zugrunde gelegte Intelligenzmodell und dessen Operationalisierung durch die Testaufgaben kennen. Viele – aber nicht alle – Testmanuals geben die entsprechenden Informationen. Zudem erfasst kein Test alle Teilbereiche der Intelligenz. Jeder Test erlaubt also nur Aussagen über einzelne oder mehrere Fähigkeiten, was bei der Interpretation und Rückmeldung zu beachten ist.

Grundsätzlich unterscheidet man zwischen eindimensionalen (Intelligenz = Begabung) und mehrdimensionalen Modellen: Eindimensionale Modelle definieren Hochbegabung ausschließlich über den Aspekt der allgemeinen Intelligenz (Rost, 1991; Feldhusen, 1986; Humphreys, 1985); mehrdimensionale Modelle konzipieren intellektuelle Hochbegabung als mehrdimensionales Fähigkeitskonstrukt und ordnen die intellektuellen Fähigkeiten in eine Reihe verschiedener anderer gleichberechtigter Begabungsfaktoren ein (z.B. Heller, Perleth & Hany, 1994; Marland, 1971). Mehrdimensionale Konzepte jenseits der klassischen IQ-Definition beschreiben intellektuelle Hochbegabung wiederum durch kognitive und metakognitive Prozesskomponenten (z.B. Sternberg, 1986), in Form multipler Intelligenzen (Gardner, 1983) oder durch Variablen der operativen Intelligenz (Dörner, 1986). Trotz der Heterogenität in der theoretischen Definition von Hochbegabung als eindimensionales vs. multidimensionales, fähigkeitsspezifisches Konstrukt wird in der Praxis der Hochbegabungsdiagnostik nach wie vor oftmals ein eindimensionaler Ansatz favorisiert (Feldhusen & Jarvan, 2000).

Im Folgenden sollen einige der einflussreichsten und bekanntesten Intelligenzmodelle, auf welchen die Intelligenztestverfahren z.T. basieren, beschrieben werden¹:

1.1.1 Zwei-Faktoren- bzw. Generalfaktoren-Theorie von Spearman

Das erste explizite Intelligenzmodell stammt vom britischen Forscher Charles Spearman (1904). Weil die Leistungen in verschiedenen kognitiven Tests im Normalfall positiv miteinander korre-

¹ Die Beschreibungen der Modelle wurden teilweise aus Holling, Preckel & Vock (2004) entnommen.

lierten, vermutete er, dass den kognitiven Leistungen ein gemeinsamer Faktor zu Grunde liegt. Spearman bezeichnete diesen als „allgemeine Intelligenz“ und bezeichnete ihn mit „g“ (general factor). Die nicht durch g erklärbare Varianz der Tests erweiterte er durch zusätzliche spezifische Faktoren (abgekürzt: „s“). Für jede einzelne Aufgabenart wird neben einem allgemeinen Faktor g ein spezifischer Faktor s angenommen. Somit geht nach Spearman jede Testleistung auf zwei Faktoren zurück und zwar auf die allgemeine Intelligenz (g-Faktor) und auf eine spezifische Intelligenz (s-Faktor) – daher die Bezeichnung „Zwei-Faktoren-Theorie“. Insgesamt gibt es – folgt man Spearman – so viele spezifische Faktoren wie Tests bzw. Aufgabentypen zur Messung intellektueller Fähigkeiten angewendet werden. Die einzelnen Aufgabentypen erfordern zu ihrer Lösung unterschiedlich hohe g-Anteile, ihr spezifischer Anteil s verändert sich entsprechend.

Empirisch zeigte sich allerdings, dass bestimmte Aufgabentypen stärker miteinander korrelieren als durch ihre jeweiligen Ladungen auf g zu erwarten wäre. Die Erklärung der Leistungsvarianz durch nur zwei Faktoren (g und s) scheint daher nicht ausreichend.

Fraglich ist bei Spearmans Modell auch die Rolle der spezifischen Faktoren, da rein aufgabenspezifische Varianzen keine klare Aussage über die Fähigkeiten bei Menschen zulassen, da zwischen einzelnen spezifischen Faktoren Korrelationen bestehen. Er bezeichnete diese selbst als „spezielle Generalfaktoren“.

Aufgabentypen mit hohem g-Gehalt sind beispielsweise figurale Matrizenaufgaben. Ein Schüler Spearmans, John Raven, entwickelte gemeinsam mit dem Genetiker Lionel Penrose 1938 den ersten Matrizenstest (vgl. Holling, Preckel & Vock, 2004), dessen Nachfolgeversionen noch heute weit verbreitet sind (in den Varianten CPM, SPM und APM, s. Teil 3). Auch Wechsler konstruierte seine Tests (aktuelle deutsche Versionen: HAWIK-III und HAWIE-R, s. Teil 3) explizit mit dem Ziel, den g-Faktor im Sinne Spearmans zu erfassen.

1.1.2 Modell mehrerer gemeinsamer Faktoren von Thurstone²

Der US-Amerikaner Louis Thurstone (1938) versuchte – wie auch viele andere Forscher/innen – Korrelationen intellektueller Leistungen nach Extraktion eines g-Faktors (nach Spearman) zu erklären. Er war der Meinung, dass sich Denkleistungen nicht nur durch einen g- und jeweils einen s-Faktor darstellen ließen, sondern dass es mehrere nebeneinander stehende generelle Faktoren geben müsse. Diese bezeichnete er als „Gruppenfaktoren“, die jeweils eine „primäre“ (grundlegende) Fähigkeit widerspiegeln. Die Berechnung eines Wertes für die Intelligenz (im Sinne eines g-Faktors) schied daher nach Thurstones Auffassung aus. Die Intelligenz einer Person konnte seiner Meinung nach nur als Profil der Ausprägungsgrade einzelner Primärfaktoren dargestellt werden. In seinen Analysen fand er mehrere „primäre Fähigkeiten“, von denen sich sieben wiederholt belegen ließen:

² Beschreibung wurde teilweise aus Holling, Preckel & Vock (2004) entnommen.

1. verbales Verständnis
2. Wortflüssigkeit
3. schlussfolgerndes Denken, Erkennen von Regelmäßigkeiten
4. räumliches Vorstellungsvermögen
5. Merkfähigkeit, Kurzzeitgedächtnis
6. Rechenfähigkeit
7. Wahrnehmungsgeschwindigkeit

Thurstone nahm an, dass die Primärfaktoren wesentlich allgemeiner sind als Spearmans spezifische (s) Faktoren, das heißt es gibt weniger Primärfaktoren als verwendete Testverfahren. Sein Modell stellt somit ein ausdrückliches Gegenmodell zu Spearmans Generalfaktoren-Modell dar.

Thurstone ging davon aus, dass es sich bei den Primärfähigkeiten zwar um verschiedene, nicht aber um völlig unabhängige Fähigkeiten handeln würde, was zur Folge hatte, dass er für seine Berechnungen eine Variante der Faktorenanalyse wählte, die Interkorrelationen zwischen den Faktoren erlaubt. Die einzelnen Primärfähigkeiten korrelieren auch tatsächlich positiv miteinander ($r = .35$; Amelang & Bartussek, 1997, S. 212). Sobald jedoch Korrelationen zwischen Faktoren bestehen, können weitere Faktoren höherer Ordnung extrahiert werden. Dies wurde von anderen Forscherinnen und Forschern so interpretiert, dass auch Thurstones Daten letztlich einen g-Faktor belegen (z.B. Eysenck, 1979). Das Modell mehrerer gemeinsamer Faktoren von Thurstone wurde ebenfalls verschiedenen Intelligenztests zu Grunde gelegt. Bekannteste Beispiele sind die älteren Versionen des Intelligenz-Struktur-Tests von Amthauer (I-S-T, Amthauer, 1955 und I-S-T 70, Amthauer, 1973; der I-S-T 2000 von Amthauer et al., 1999 und die Revision zum I-S-T 2000-R, 2001 basieren auf einem erweiterten Modell).³

1.1.3 Theorie der fluiden und kristallinen Intelligenz von Cattell⁴

Eine weitere – bis heute bekannte und einflussreiche – Modellvorstellung von Intelligenz stammt von Raymond B. Cattell (1954). Sein Modell kann als eine Synthese aus Spearmans Zwei-Faktoren-Theorie und dem Modell mehrerer gemeinsamer Faktoren von Thurstone angesehen werden.

Cattell postulierte, dass sich der g-Faktor Spearmans aus zwei (voneinander unabhängigen) Komponenten zusammensetzt. Die beiden generellen Faktoren ersetzen Spearmans „g“ vollständig. Auf einer ersten Ebene siedelte Cattell – vergleichbar mit Thurstones Modell – mehrere miteinander korrelierte Primärfaktoren an, die sich in Testleistungen bei verschiedenen

³ Abweichend von Thurstones ursprünglichem Modell erlauben die I-S-T-Tests jedoch die Berechnung eines Kennwertes für die allgemeine Intelligenz.

⁴ Beschreibung wurde teilweise aus Holling, Preckel & Vock (2004) entnommen.

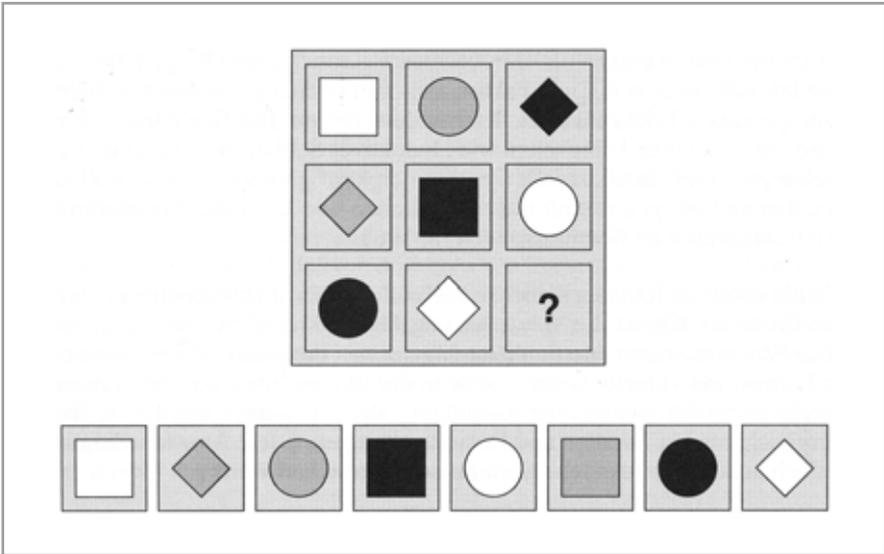


Abb. 1: Beispielaufgaben zur Erfassung von g_i (nach Penrose & Raven, 1936)
 (entnommen aus Holling, Preckel & Vock (2004), S. 22)

Wortschatztest

Welches der Wörter passt in seiner Bedeutung nicht zu den anderen?

Hose	<u>Wolle</u>	Kleid	Jacke
Vogel	Wolke	Ballon	<u>Ball</u>

Verbale Analogien

Blatt zu Baum wie Haar zu ?

Wurzel	<u>Kopf</u>	Fell	Haut
--------	-------------	------	------

Saft zu Milch wie Apfel zu ?

Frucht	Obst	<u>Banane</u>	Boskop
--------	------	---------------	--------

Abb. 2: Beispielaufgaben zur Erfassung von g_c
 (entnommen aus Holling, Preckel & Vock (2004), S. 23)

Aufgaben zeigen. Die beiden generellen Faktoren befinden sich nach Cattell auf einer zweiten übergeordneten Ebene. Den einen generellen Faktor bezeichnete Cattell als „fluide“ allgemeine Intelligenz (abgekürzt: g_f). Die fluide Intelligenz ist die Fähigkeit, sich neuen Situationen anzupassen und neuartige Probleme zu lösen, ohne dass gelerntes Wissen eine bedeutsame Rolle spielt. Sie ist vermutlich weitgehend von Geburt an angelegt und von kulturellen und gesellschaftlichen Einflüssen unabhängig.

Den zweiten Faktor nannte er „kristalline“ allgemeine Intelligenz (abgekürzt: g_c). Die kristalline Intelligenz bezeichnet kognitive Fertigkeiten, die durch die Kumulierung (oder „Kristallisierung“) von Lernerfahrungen seit der Geburt entwickelt wurden. Während die fluide Intelligenz also neuartige Informationen verarbeitet, beschäftigt sich die kristalline Intelligenz mit der Verarbeitung vertrauter Informationen und der Anwendung von Wissen (Berg, 2000). Es handelt sich bei g_f und g_c um Faktoren zweiter Ordnung, die durch mehrere untergeordnete Primärfaktoren beeinflusst werden.

Die fluide Intelligenz wird vor allem über Fähigkeiten wie „Induktives Schließen“, „Figurale Beziehungen“ und „Intellektuelle Geschwindigkeit“ gemessen, die kristalline Intelligenz hingegen stärker über Fähigkeiten wie „Verbales Verständnis“ oder „Mechanische Kenntnisse“ (Horn, 1968).

Nach Cattell ist die fluide Intelligenz die Voraussetzung für die kristalline Intelligenz. Individuelle Unterschiede in der fluiden Intelligenz bestimmen somit mehr oder weniger die Ausprägung der kristallinen Intelligenz (bei vergleichbaren kulturellen Lernumgebungen). Zu dieser Theorie liegen jedoch noch keine empirisch abgesicherten Werte vor.

Die beiden allgemeinen Faktoren g_f und g_c erwiesen sich in verschiedenen Studien als nicht unabhängig voneinander (vgl. Holling, Preckel & Vock, 2004, S. 22). Cattell (1971) erweiterte daher sein Modell um einen zusätzlichen – g_f und g_c übergeordneten – Faktor. Dieser weist einen engeren Zusammenhang mit g_f als mit g_c auf. Diesen Faktor dritter Ordnung interpretierte Cattell jedoch explizit nicht als Spearman's g .

1.1.4 Berliner Intelligenzstrukturmodell von Jäger⁵

Adolf Otto Jäger legte seiner Modellentwicklung (Berliner Intelligenzstrukturmodell – „BIS-Modell“) in den 70er Jahren ein sehr umfassendes Inventar der Typen von Intelligenz- und Kreativitätsaufgaben, die in der Forschung und Testpraxis verwendet wurden, zu Grunde (Jäger, 1967, 1982). Sein Modell wurde als integratives Modell konzipiert, das ein möglichst breites Spektrum intellektueller Fähigkeitskonstrukte abdeckt.

Das Berliner Intelligenzstrukturmodell nach Jäger (s. Abb. 3) hat die Form einer Raute und strukturiert Intelligenz hierarchisch. Er geht davon aus, dass sich die allgemeine Intelligenz aus sieben verschiedenen Fähigkeiten zusammensetzt.

⁵ Beschreibung wurde teilweise aus Holling, Preckel & Vock (2004) entnommen.

Nach dem BIS-Modell ist die allgemeine Intelligenz am besten durch eine möglichst breite Stichprobe aus der Vielfalt intellektueller Prozesse erfassbar. Die sieben spezielleren Fähigkeiten hingegen beeinflussen seiner Meinung nach mehr die Leistungen in enger begrenzten Bereichen als die allgemeine Intelligenz, weisen hier jedoch relativ große Varianzanteile auf. Sie werden unterteilt in so genannte „Operative Fähigkeiten“ und in „Inhaltliche Fähigkeiten“. Es handelt sich bei diesen beiden Arten von Fähigkeiten um zwei Modalitäten, unter denen Leistungen und Fähigkeiten klassifiziert werden können (vergleichbar mit Form, Farbe oder Gewicht als unterschiedliche Qualitäten von Gegenständen).

Die operativen Fähigkeiten beschreiben verschiedene Denkopoperationen⁶:

- B *Bearbeitungsgeschwindigkeit*: Arbeitstempo, Auffassungsleichtigkeit und Konzentrationskraft beim Lösen einfach strukturierter Aufgaben von geringem Schwierigkeitsniveau.
- M *Merkfähigkeit*: Aktives Einprägen und kurzfristiges Wiedererkennen oder Reproduzieren von Informationen.
- E *Einfallsreichtum*: Flexible Produktion von vielfältigen Ideen und Lösungen für eine vorgegebene Problemstellung; erforderlich ist hierfür die Verfügbarkeit vielfältiger Informationen, ein Reichtum an Vorstellungen und das Sehen vieler verschiedener Seiten, Varianten, Gründe und Möglichkeiten von Gegenständen und Problemen.
- K *Verarbeitungskapazität*: Verarbeitung komplexer Informationen bei Aufgaben, die nicht auf Antrieb zu lösen sind, sondern Heranziehen, vielfältiges Beziehungsstiften, formallogisch exaktes Denken und sachgerechtes Beurteilen von Informationen erfordern.

Neben den operativen Fähigkeiten werden drei inhaltsgebundene Fähigkeiten unterschieden⁷:

- F *Anschauungsgebundenes, figural-bildhaftes Denken*: Fähigkeit zum Umgang mit Aufgabenmaterial, dessen Bearbeitung figural-bildhaftes und/oder räumliches Vorstellen erfordert.
- V *Sprachgebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Symbolsystems Sprache.
- N *Zahlengebundenes Denken*: Grad der Aneignung und der Verfügbarkeit des Symbolsystems Zahlen.

Dem Modell liegt die Annahme zugrunde, dass jede intelligente Leistung durch alle diese Fähigkeiten bestimmt wird, jedoch zu verschiedenen Anteilen. Bei vielen Leistungen dominiert je eine operative und eine inhaltliche Fähigkeit. Zum Beispiel erfordert das Lösen von Kopfrechenaufgaben vor allem die Fähigkeiten K und N, obwohl z.B. die Fähigkeit M auch in einem geringen Ausmaß relevant ist, etwa wenn Zwischenergebnisse im Kopf behalten werden müssen.

⁶ entnommen aus Holling, Preckel & Vock (2004)

⁷ entnommen aus Holling, Preckel & Vock (2004)

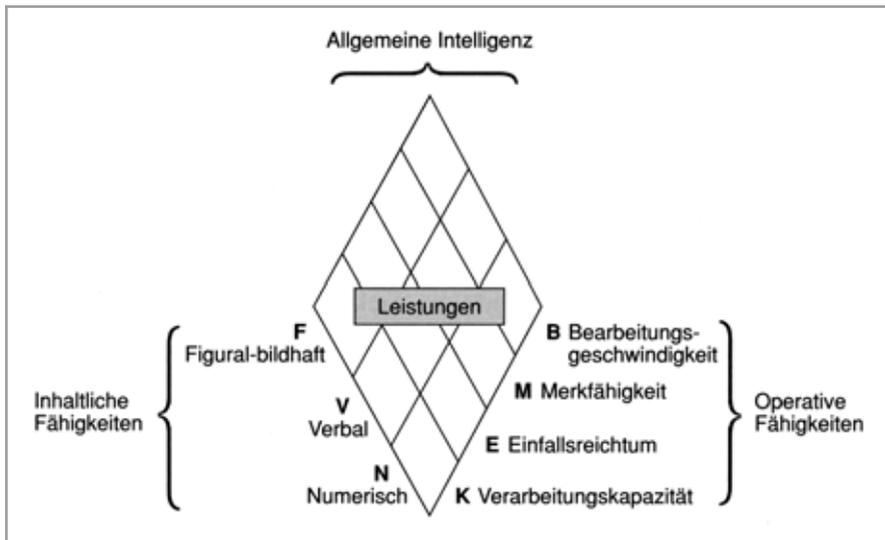


Abb. 3: Berliner Intelligenzstrukturmodell nach Jäger
 (entnommen aus Holling, Preckel & Vock (2004), S. 24)

Entsprechend ist das Modell so konzipiert, dass jeweils eine operative und eine inhaltliche Fähigkeit miteinander kombiniert werden, um verschiedene spezifische Leistungen zu erklären. Die Kombination der vier operativen und der drei inhaltsgebundenen Fähigkeiten in Rautenform ergibt zwölf Zellen. Diese entsprechen zwölf Arten von kognitiven Leistungen, die mit den sieben Fähigkeiten erklärt werden.

Die Leistung beim Lösen verbaler Analogieaufgaben wird in der Zelle „Verarbeitungskapazität verbal“ positioniert (inhaltliche Fähigkeit V, operative Fähigkeit K). Gleichzeitig wird die Leistung, Analogieaufgaben zu lösen, auch durch die allgemeine Intelligenz der Person bestimmt. Die Leistung, viele unterschiedliche Bilder zu einem Thema zu zeichnen, wird im BIS-Modell der Zelle „Einfallsreichtum figural“ zugeordnet.

Das BIS-Modell enthält nicht alle vorstellbaren unterscheidbaren Intelligenzkomponenten, die dazu nötig sind, um intelligente Leistungen vollständig zu erklären und wurde daher von Jäger ausdrücklich als ein Modell verstanden, das für zukünftige Erweiterungen offen ist. Die Validität der bisherigen Form des BIS-Modells konnte inzwischen in verschiedenen Untersuchungen bestätigt werden (z.B. Jäger et al., 2004; Süß, Oberauer, Wittmann, Wilhelm & Schulze, 2002).

Das BIS-Modell ist ein integratives Modell und eignet sich damit für die Einordnung von verschiedenen Intelligenztests. Beispielsweise lassen sich die in Kapitel 3 angeführten figuralen Matrizentests von Raven und der CFT 20 in die Zelle KF des Modells einordnen. Der Untertest

Beispielaufgabe für KF (Verarbeitungskapazität figural)

Welche der Figuren A bis E kann man aus den Teilen im umrandeten Kasten genau zusammensetzen? Nur eine Lösung ist richtig. Streichen Sie den Buchstaben unter der richtigen Lösung durch!



Beispielaufgabe für MN (Merkfähigkeit numerisch)

Sie sehen gleich eine Liste mit zweistelligen Zahlen vor sich. Ihre Aufgabe wird es sein, sich diese Zahlen gut einzuprägen. Dafür haben Sie 1 Minute Zeit! Später sollen Sie alle Zahlen, die Sie sich merken konnten, aufschreiben.

12 44 56 82 34 25 17

Beispielaufgabe für EV (Einfallsreichtum verbal)

Für viele Berufe braucht man bestimmte Eigenschaften und Fähigkeiten. Es gibt aber auch Eigenschaften und Fähigkeiten, die für einen bestimmten Beruf ungünstig sind. Sie sollen nun möglichst viele und sehr verschiedene Eigenschaften und Fähigkeiten nennen, die ein Vertreter eines bestimmten Berufes nicht haben sollte. Stichworte genügen!

Beispielsweise sind für den Beruf des Richters mögliche Lösungen: parteisch, bestechlich, oberflächlich (etc.)

Abb. 4: Beispielaufgaben zur Erfassung der Zellen KF, MN und EV
(entnommen aus Holling, Preckel & Vock (2004), S. 26)

„Zahlen nachsprechen“ des HAWIK-III kann der Zelle MN zugeordnet werden. Derzeit existieren zwei Testverfahren, die auf dem BIS-Modell aufbauen und mit denen alle im Modell spezifizierten Fähigkeiten erfasst werden können: Der BIS-4 von Jäger, Süß & Beauducel (1997) sowie der BIS-HB von Jäger et al. (2004; beide siehe Kap. 3).

1.1.5 Carrolls Three-Stratum-Theorie⁸

Wie sich anhand des BIS-Modells zeigte, ist für die Entwicklung von Intelligenzmodellen die Verwendung eines möglichst breiten Spektrums unterschiedlicher kognitiver Aufgaben erforderlich. Eine in diesem Sinne beispielhafte Intelligenzmodellentwicklung führte auch John B. Carroll (1993) durch. Er entwarf ein Strukturmodell der Intelligenz mit drei Hierarchieebenen, benannt als Three-Stratum-Theorie (Drei-Schichten-Theorie).

⁸ Beschreibung wurde teilweise aus Holling, Preckel & Vock (2004) entnommen.

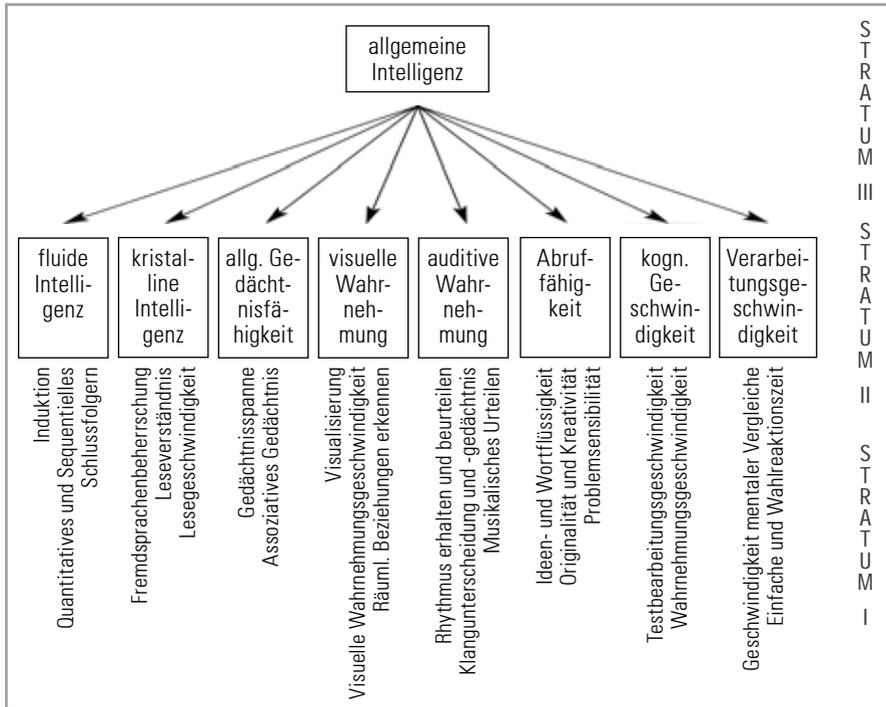


Abb. 5: Three-Stratum-Theorie nach Carroll (1993)
(entnommen aus Holling, Preckel & Vock (2004), S. 28)

Auf der Ebene höchster Generalität (Stratum III) befindet sich die allgemeine Intelligenz, die durch komplexe kognitive Prozesse höherer Ordnung bestimmt wird. Carroll meinte, dass die allgemeine Intelligenz eine Konstellation in der Fähigkeitsstruktur einer Person darstelle, die eine große Bandbreite von Verhaltensweisen und Leistungen beeinflusst (Carroll, 1991).

Auf der zweiten Ebene sind acht Fähigkeiten mittlerer Generalität angesiedelt (Stratum II), die sich wie folgt beschreiben lassen⁹:

1. *Fluide Intelligenz*: Basale Prozesse des schlussfolgernden, logischen Denkens und andere mentale Aktivitäten, die nur minimal durch Lernen und Akkulturation beeinflusst werden.
2. *Kristalline Intelligenz*: Mentale Prozesse, die nicht nur von der fluiden Intelligenz, sondern auch von Erfahrung, Lernen und Akkulturation abhängen.

⁹ entnommen aus Holling, Preckel & Vock (2004)



3. *Allgemeine Gedächtnisfähigkeit*: Fähigkeit zum Lernen und Behalten neuer Inhalte oder neuen Verhaltens.
4. *Visuelle Wahrnehmung*: Fähigkeit, die bei allen Aufgaben eine Rolle spielt, die die Wahrnehmung visueller Formen erfordern.
5. *Auditive Wahrnehmung*: Fähigkeit zur Wahrnehmung oder Unterscheidung auditiver Klangmuster oder gesprochener Sprache.
6. *Abruffähigkeit*: Fähigkeit, Konzepte, Inhalte oder Ideen aus dem Langzeitgedächtnis abzurufen.
7. *Kognitive (Verarbeitungs-)Geschwindigkeit*: Geschwindigkeit der kognitiven Verarbeitung von Informationen.
8. *Verarbeitungsgeschwindigkeit*: Entscheidungsgeschwindigkeit in verschiedenen Reaktionszeitaufgaben.

Den acht Fähigkeiten mittlerer Generalität sind 69 relativ spezifische Fähigkeiten auf Stratum I zugeordnet, von denen in obiger Abbildung aus Platzgründen nur einige exemplarisch dargestellt sind. Die auf Stratum I benannten Fähigkeiten werden häufig von mehreren Stratum-II-Fähigkeiten beeinflusst, stellen also Mischformen aus den übergeordneten Fähigkeiten dar. Als Beispiel sei hier die Wahrnehmungsgeschwindigkeit angeführt, die sowohl von allgemeinen visuellen Fähigkeiten als auch von der kognitiven Geschwindigkeit abhängt. Die Zuordnungen der spezifischen Fähigkeiten erfolgen danach, welche Fähigkeit von Stratum II die spezifische Fähigkeit (Stratum I) anteilmäßig am meisten bestimmt. Diese Ausführungen verdeutlichen, dass intellektuelle Fähigkeiten komplex und nicht voneinander unabhängig sind.

1.1.6 Sternbergs Triarchische Theorie der Intelligenz¹⁰

Robert J. Sternbergs Triarchische Theorie der Intelligenz berücksichtigt neben interindividuellen Unterschieden auch die Interaktion zwischen intelligentem Verhalten und Umgebungsvariablen. Intelligenz wird hier als dynamisches Konzept verstanden, das nicht unabhängig von Umgebungsbedingungen erforscht und verstanden werden kann. Sternbergs Modell weist eher einen kognitionspsychologischen als einen differenziellen Hintergrund auf. Aus diesem Grund werden vorwiegend mentale Prozesse, die bei Aufgabenlösungen bedeutsam sind und deren Prozesskomponenten in Augenschein genommen.

Sternberg (1997) nennt drei interagierende Aspekte, aus denen sich Intelligenz zusammensetzt: die Informationsverarbeitungsfähigkeiten einer Person, die intelligentem Verhalten zu Grunde

¹⁰ Beschreibung wurde teilweise aus Holling, Preckel & Vock (2004) entnommen.

liegen (interner Aspekt), das Verhältnis von Intelligenz und Erfahrung (Erfahrungsaspekt) und die praktische Anwendung der Intelligenz in der externen Welt (externer Aspekt).

Beim internen Aspekt unterscheidet Sternberg die folgenden Komponenten:

- *Metakomponenten*: Mentale Prozesse höherer Ordnung, die in allen Aufgabenstellungen zur Planung, Überwachung und Bewertung von Problemlösungen erforderlich sind.
- *Performanz-Komponenten*: Mentale Prozesse niedrigerer Ordnung, die in der Regel aufgabenspezifisch zur Ausführung der Anweisungen der Metakomponenten benötigt werden.
- *Wissenserwerbskomponenten*: Mentale Prozesse niedrigerer Ordnung, die Lernen und Wissenserwerb steuern (z.B. selektives Enkodieren oder Vergleichen von Informationen).

Sternberg konnte in verschiedenen Untersuchungen aufzeigen, dass intelligentere Personen schneller und effizienter bei der Lösung neuartiger Probleme sind und im Vergleich zu weniger intelligenten Personen Zeit und Kapazität gewinnen, um Neues zu automatisieren, was wiederum Platz für die Verarbeitung weiterer Informationen schafft (z.B. Davidson & Sternberg, 1984). Der externe Aspekt der Intelligenz beschäftigt sich laut Sternberg mit der praktischen Anwendung der Metakomponenten, der Performanz- und der Wissenserwerbskomponenten im Umfeld einer Person. Intelligente Personen wissen demnach, auf welche Weise und zu welchem Zeitpunkt sie sich an eine bestimmte Umgebung anpassen können. Falls eine Anpassung nicht gelingt, äußert sich intelligentes Verhalten entweder darin, die Umwelt zu verändern oder eine neue Umgebung aufzusuchen. Der externe Intelligenzaspekt zeigt sich also in der Aneignung von Handlungswissen, das nicht explizit vermittelt, sondern anderweitig erworben wird.

Intelligenz im Sinne Sternbergs, der darunter die Fähigkeit versteht, aus Erfahrungen zu lernen und sich an die Umgebung anzupassen, geht über die akademische Intelligenz, die mit Intelligenztests erfasst wird, hinaus. Ein Test zur Erfassung der in Sternbergs Theorie spezifizierten Aspekte wurde bis heute noch nicht entwickelt. Seine Theorie stellt vielmehr eine Ergänzung zu den Intelligenz(struktur)-Modellen dar, die aus der Differenziellen Psychologie erwachsen sind.

1.1.7 Zusammenfassung

Die obigen Ausführungen verdeutlichen, dass Intelligenzdiagnostik mit der für offene Konstrukte typischen Situation konfrontiert ist, dass keine allgemein anerkannte verbindliche Definition von Intelligenz existiert. Dieses Fehlen einer einheitlichen Definition von Intelligenz birgt das Dilemma für die Praxis der Diagnostik, dass Intelligenz durch unterschiedliche Intelligenztests erfasst werden kann, denen unterschiedliche Intelligenzmodelle zu Grunde liegen und die unterschiedliche oder auch dieselben Fähigkeitsdimensionen aber mit unterschiedlicher Validität¹¹ erfassen. Die Auswahl eines Testverfahrens muss daher immer situationsangemessen und zielorientiert auf die diagnostische Fragestellung hin erfolgen.

¹¹ Zum Begriff Validität siehe Kapitel 2.3

1.2 Hochbegabungstheorien und -modelle

Der Einfluss des individuellen Intelligenzprofils auf die Wahrscheinlichkeit, herausragende Leistungen zu vollbringen, wird in der Literatur sehr unterschiedlich und z.T. auch domänenabhängig beschrieben. Während manche Wissenschaftler/innen aus der Begabungsforschung Intelligenz als den wichtigsten Faktor für Leistungsprognosen betrachten (vgl. Terman et al. 1925; Rost, 1993), sprechen besonders Expertiseforscher/innen der Intelligenz einen besonderen Einfluss auf exzellente Leistungen ab und betonen die Bedeutung von Lernen, Üben und Training für die Entwicklung von Leistungsexzellenz (vgl. Ericsson et al. 1993; Gruber, 2007; Ziegler, 2008). Wissenschaftliche Meinungen differieren auch darüber, mit welchen weiteren Faktoren die Intelligenz eines Individuums bei der Entfaltung von vorhandenen Potenzialen zu sichtbaren Leistungen interagiert.

Im Folgenden werden zwei Modelle vorgestellt, die im deutschsprachigen Raum allgemeine Anerkennung und Verbreitung gefunden haben und die insbesondere für die Begabungsdiagnostik und -förderung eine breite Palette an Beratungs- und Interventionsansätzen bieten.

1.2.1 Münchner Hochbegabungsmodell

Das Münchner Hochbegabungsmodell nach Kurt Heller, Christoph Perleth und Ernst Hany (1994) geht von einem mehrdimensionalen Prädiktoren-Moderatoren-Modell der Begabungsentwicklung aus und versucht, die Bedingungen für außergewöhnliche Leistungen in verschiedenen Bereichen abzubilden.

Im Münchner Hochbegabungsmodell wird Wert auf die Unterscheidung von Prädiktoren und Moderatoren gelegt. Beide – Prädiktoren wie Moderatoren – sind für die Erklärung außergewöhnlicher Leistungen von Bedeutung:

- Unter *Prädiktoren* werden Begabungsfaktoren verstanden, die – in unterschiedlicher Gewichtung – wichtige Voraussetzungen für außergewöhnliche Leistungen in unterschiedlichen Domänen darstellen: intellektuelle Fähigkeiten, kreative Fähigkeiten, soziale Kompetenz, Musikalität, künstlerische Fähigkeiten, Psychomotorik und praktische Intelligenz.
- Unter *Moderatoren* werden von Heller & Perleth (2007) nicht-kognitive Persönlichkeitsmerkmale und Umweltmerkmale verstanden. Nicht-kognitive Persönlichkeitsmerkmale sind z.B. Interessen, Leistungsmotivation, Arbeitsverhalten, Prüfungssorgen/Ängstlichkeit und Kausalattributionen. Umweltmerkmale im Münchner Hochbegabungsmodell sind u.a. familiäre Lernumwelt, Familienklima, Instruktionsqualität, Schulklima und kritische Lebensereignisse.

Die Autoren vertreten die Auffassung, dass sich eine umfassende Begabungsdiagnose nicht auf eine reine Messung des Intelligenzquotienten (IQ) beschränken dürfe, sondern, je nach Problem- und Fragestellung, weitere relevante Begabungs- und unterschiedliche Moderatoraus-

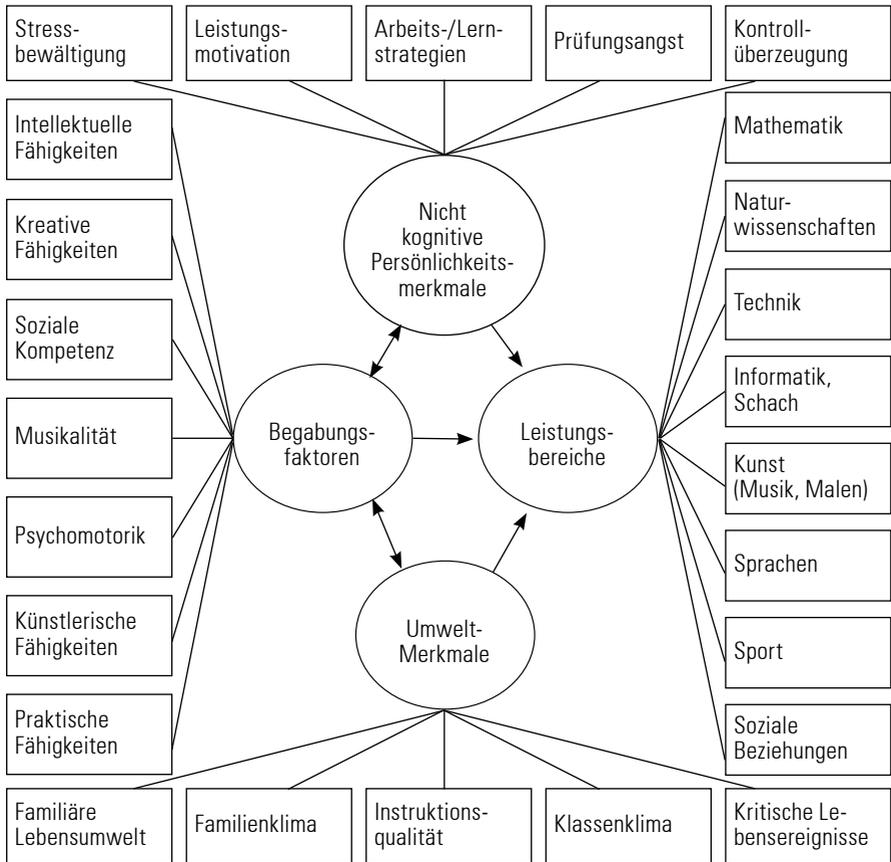


Abb. 6: Münchner Hochbegabungsmodell (entnommen aus Joswig, 2000, S. 23; Bearbeitung: Margarethe Kainig-Huber)

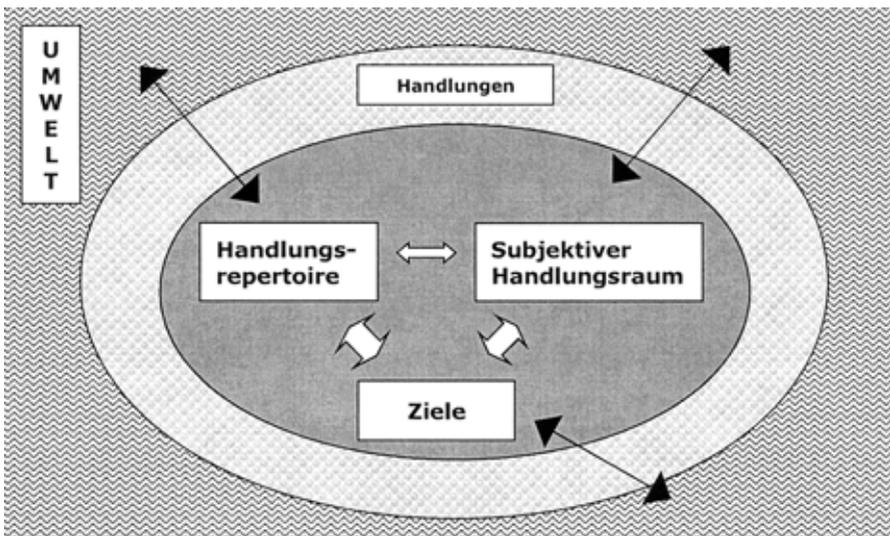
prägungen berücksichtigen müsse, um eine differenzielle Beratung und Intervention anbieten zu können. Leistungen sind nicht nur das Produkt eines entsprechenden Begabungspotenzials, sondern sind auch von nicht-kognitiven Persönlichkeits- und Umweltmerkmalen geprägt.

1.2.2 Aktiotop-Modell

Das Aktiotop-Modell von Albert Ziegler stammt aus der Expertiseforschung und folgt einem systemischen Ansatz. Begabung wird darin dem gesamten System aus Person und der Umwelt, in der sie handelt, zugeschrieben (Ziegler, 2009). Diagnostik und Intervention richten sich auf alle Komponenten des Aktiotops einer Person.

Das Aktiotop umfasst fünf Komponenten: Ziele, Handlungsrepertoire, Handlungsraum, Handlungen und Umwelt:

- Die *Ziele* einer Person geben ihrem Handeln eine Richtung und haben eine steuernde Funktion im Lern- und Entwicklungsprozess.
- Das *Handlungsrepertoire* umfasst alle objektiv feststellbaren Handlungsmöglichkeiten einer Person, also alle jene Handlungen, zu denen die Person grundsätzlich fähig wäre.
- Der subjektive *Handlungsraum* schließt jene Handlungsschritte ein, die eine Person bei ihrer Handlungsplanung und -regulation in Erwägung zieht. Subjektiv ist dieser Handlungsraum, weil die tatsächlich möglichen Handlungen (das gesamte vorhandene Handlungsrepertoire) darin nicht unbedingt erfasst werden.
- *Handlungen* bestehen stets aus mehreren Teilhandlungen, die Regulationsprozesse und Strategien auf verschiedensten Ebenen verlangen, z.B. die korrekte Durchführung einer Aktivität, die Steuerung des Aufwands, den Umgang mit negativen Affekten usw.
- Die *Umwelt* schließlich besteht aus vielen verschiedenen Entitäten, die miteinander interagieren und (Lern-)Handlungen ermöglichen oder behindern, z.B. soziale Akteure, Lernressourcen, Informationen, Lernsettings usw.



Das Aktiotop einer Person

Abb. 7: Komponenten eines Aktiotops (Ziegler, 2009, S. 15)

Das Erreichen von Leistungsexzellenz wird nach dem Aktiotop-Modell nicht sosehr durch die Höhe der Intelligenz bestimmt, sondern durch die erfolgreiche Bewältigung von aufeinander folgenden Lernstufen. Durch Lernen und Entwicklung wird das individuelle Handlungsrepertoire ständig erweitert und es werden dadurch immer höhere Leistungsniveaus ermöglicht.

Voraussetzung ist, dass sich die einzelnen Komponenten des Aktiotops in Abstimmung miteinander weiterentwickeln. Die Erweiterung des Handlungsrepertoires sollte z.B. den Zielen entsprechen und im subjektiven Handlungsraum abgebildet werden. Die Umwelt sollte kontinuierlich so angepasst oder ausgewählt werden, dass sie die Handlungen bestmöglich unterstützen kann.

1.2.3 Zusammenfassung

Obwohl die hier beschriebenen Hochbegabungsmodelle insbesondere für die Praxis eine hohe Attraktivität aufweisen, lassen sich einige Schwächen nicht leugnen. Beide Modelle haben ihre empirische Bewährungsprobe noch nicht bestanden; es fehlen einschlägige Evaluationsstudien. Die ausgewählten Variablen sind nicht unumstritten. So ist z.B. die Kreativität ein psychologisches Konstrukt, dessen Brauchbarkeit von vielen angezweifelt wird; reliable Messverfahren zur Kreativität existieren (noch) nicht. Auch die Variablen sind nicht überschneidungsfrei und ihre gegenseitigen Wechselbeziehungen weitgehend ungeklärt.

Eine für die Begabungsdiagnostik wichtige Schlussfolgerung aus den beiden Modellen zur Hochbegabung und deren Entwicklung ist jedoch, dass Intelligenz nicht identisch mit Begabung ist, sondern lediglich als Teil des Begabungsspektrums anzusehen ist. Begabungsdiagnostik hat mehr als die Messung der Intelligenz zu umfassen und wird auch verschiedene nicht-kognitive Persönlichkeitsmerkmale und Umweltvariablen einschließen. Das sei hier besonders im Hinblick darauf betont, dass sich die folgenden Ausführungen ausschließlich auf Testverfahren zur Messung der intellektuellen Fähigkeiten beschränken.

Eine Erörterung weiterer Verfahren zur Begabungsdiagnostik finden Sie z.B. in den ÖZBF-Broschüren „Psychologische Diagnostik moderierender Persönlichkeitsmerkmale bei Kindern und Jugendlichen“ (2012, in Druck) und „Psychologische Diagnostik von Konzentrationsfähigkeit bzw. Aufmerksamkeit im Kindergarten- und Schulalter“ (2012, in Druck).

Literatur

- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. Auflage). Stuttgart: Kohlhammer.
- Amthauer, R. (1955). *Intelligenz-Struktur-Test (I-S-T)*. Göttingen: Hogrefe.
- Amthauer, R. (1973). *Intelligenz-Struktur-Test (I-S-T 70)*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (1999). *Intelligenz-Struktur-Test 2000 (I-S-T 2000)*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R (I-S-T 2000 R)*. Göttingen: Hogrefe.
- Berg, C. A. (2000). *Intellectual development in adulthood*. In R. J. Sternberg (Hrsg.), *Handbook of intelligence* (S. 117-137). Cambridge: Cambridge University Press.

- Carroll, J. B. (1991). *No demonstration that g is not unitary, but there is more to the story: Comment on Kranzler and Jensen*. *Intelligence*, 15, 423-436.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Cattell, R. B. (1954). *Culture Fair Intelligence Test, Scales 1, 2 and 3*. Champaign, IL: IPAT.
- Cattell, R. B. (1971). *Abilities: their structure, growth and action*. Boston: Houghton Mifflin.
- Davidson, J. E. & Sternberg, R. J. (1984). *The role of insight in intellectual giftedness*. *Gifted Child Quarterly*, 28(2), 58-64.
- DeShon, R. P., Chan, D. & Weissbein, D. A. (1995). *Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants*. *Intelligence*, 21(2), 135-155.
- Dörner, D. (1986). *Diagnostik der operativen Intelligenz*. *Diagnostica*, 32, 209-308.
- Eysenck, H. J. (1979). *The structure and measure of intelligence*. Berlin: Springer.
- Feldhusen, J. F. (1986). *A conception of giftedness*. In R. J. Sternberg & J. E. Davidson (Hrsg.), *International handbook of giftedness* (S.112-127). Cambridge: Cambridge University Press.
- Feldhusen, J. F. & Jarvan, F. A. (2000). *Identification of gifted and talented youth for educational programmes*. In K. A. Heller, F. J. Mönks, A. H. Passow & R. F. Subotnik (Hrsg.), *International handbook of giftedness and talent* (S. 233-252) (2nd edition). Kidlington: Elsevier Science.
- Gardner, H. (1983). *Frames of Mind: The theory of multiple intelligences*. New York: Basic Books.
- Heller, K. A. (2001). *Hochbegabung im Kindes- und Jugendalter* (2. Auflage). Göttingen: Hogrefe.
- Heller, K. A., Gädike, A.-K. & Weinläder, H. (1985). *Kognitiver Fähigkeits-Test für 4. bis 13. Klassen (KFT 4-13+)* (2., verbesserte und erweiterte Auflage). Weinheim: Beltz.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.
- Heller, K. A. & Perleth, C. (2007). *MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe*. Göttingen: Hogrefe.
- Heller, K. A., Perleth, C. & Hany, E. A. (1994). *Hochbegabung – ein lange Zeit vernachlässigtes Forschungsthema. Einsichten – Forschung der Ludwig-Maximilians-Universität München*, Vol. 3, 1, 18-22.
- Holling, H. & Kanning, U. P. (1999). *Hochbegabung*. Göttingen: Hogrefe.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6*. Göttingen: Hogrefe.
- Holling, H., Wübbelmann, K. & Geldschläger, H. (1996). *Kriterien und Instrumente zur Auswahl von Begabten in der beruflichen Bildung*. In R. Manstetten (Hrsg.), *Begabtenförderung in der beruflichen Bildung* (S. 86-174). Göttingen: Hogrefe.
- Horn, J. L. (1968). *Organization of abilities and the development of intelligence*. *Psychological Review*, 75(3), 242-259.
- Humphreys, L. G. (1985). *A conceptualization of intellectual giftedness*. In F. D. Horowitz & M. O'Brian (Hrsg.), *The gifted and talented: Development perspectives* (S. 331-360). Washington: APA.
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Jäger, A. O. (1982). *Mehrmodale Klassifikation von Intelligenzleistungen. Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells*. *Diagnostica*, 28, 195-225.

- Jäger, A. O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2004). *Berliner Intelligenzstrukturtest für Jugendliche: Begabungs- und Hochbegabungsdiagnostik (BIS-HB)*. Göttingen: Hogrefe.
- Jäger, A. O., Süß, H.-M & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test (BIS-4)*. Göttingen: Hogrefe.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger Publishers.
- Joswig, H. (2000). *Begabungen erkennen – Begabte fördern*. Rostock: Universität Rostock, Philosophische Fakultät.
- Marland, S. P. (1971). *Education of the gifted and talented. Report to the Congress of the United States by the Commissioner of Education*. Washington DC: US Government Printing Office.
- Mönks, F. J. (1995). Hochbegabung – ein Mehrfaktorenmodell. *Grundschule*, 28, 15-18.
- Penrose, L. S. & Raven, J. C. (1936). A new series of perceptual tests: Preliminary communication. *British Journal of Medical Psychology*, 16, 97–104.
- Raven, J. C. (1938). *Progressive Matrices*. London: Lewis.
- Renzulli, J. S. (1978). *What makes giftedness: A re-examination of the definition of the gifted and talented*. Storrs, CT: University of Connecticut, Bureau of Educational Research Report Series.
- Renzulli, J. S. (1986). The three-ring-conception of giftedness. In R. J. Sternberg & J. E. Davidson (Hrsg.), *Conceptions of giftedness* (S. 53-92). New York: Cambridge University Press.
- Renzulli, J. S. (2004). Eine Erweiterung des Begabungsbegriffs unter Einbeziehung co-kognitiver Merkmale. In C. Fischer, F. Mönks & E. Grindel (Hrsg.), *Curriculum und Didaktik der Begabtenförderung* (S. 54-82). Münster: LIT-Verlag.
- Rost, D. H. (1991). Identifizierung von Hochbegabung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 23(3), 197-231.
- Spearman, C. (1904). "General intelligence", objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Sternberg, R. J. (1986). A triarchic theory of intellectual giftedness. In R. J. Sternberg & J. E. Davidson (Hrsg.), *Conceptions of giftedness* (S. 223-246). Cambridge: Cambridge University Press.
- Sternberg, R. J. (1997). The triarchic theory of intelligence. In D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Hrsg.), *Contemporary intellectual assessment: Theories, tests, and issues* (S. 92-104). New York: Guilford Press.
- Süß, H.-M., Oberauer, K., Wittmann, W., Wilhelm, W. & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence*, 30, 261-288.
- Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.
- Wagner, R. K. & Sternberg, R. J. (1985). Practical intelligence in real world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, 49(2), 436-458.
- Ziegler, A. (2008). *Hochbegabung*. München: Reinhardt.
- Ziegler, A. (2009). „Ganzheitliche Förderung“ umfasst mehr als nur die Person: Aktiotop- und Soziotopförderung. *Heilpädagogik online* 02/09, 5-34. Verfügbar unter: www.heilpaedagogik-online.com/2009/heilpaedagogik_online_0209.pdf [31.01.2011].

2 Auswahl von Intelligenztests für die Begabungsdiagnostik



Wie bereits in Kapitel 1 dargelegt, kann ein einzelner Intelligenztest nur eine Auswahl dessen darstellen, was Wechsler (1944) als die „Fähigkeit des Individuums, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinanderzusetzen“ bezeichnete. Testverfahren zur Messung von intellektueller Begabung unterscheiden sich in der Auswahl der Konstrukte, die jeweils gemessen werden sollen und in den Aufgaben, durch die diese Konstrukte zu messen versucht werden. Wie Ricken et al. (2007) betonen, liefern Tests zur Erfassung intellektueller Fähigkeiten zwar eine qualitativ angemessene Informationsmenge über die kognitiven Fähigkeiten, sie können jedoch immer nur Ausschnitte des gesamten intellektuellen Potenzials der getesteten Person erfassen. Angesichts der großen Bandbreite an Definitionen der Intelligenz und der Unterschiede in den Intelligenztests, stellte bereits Boring (1923; vgl. Boring, 1923, S. 35ff. zit. nach Funke & Vaterrodt-Plünnecke, 1998) fest, dass Intelligenz jeweils das ist, „was der Test misst“. Ist sich die Testleiterin/der Testleiter dieser Konstrukt- und Testabhängigkeit der Ergebnisse bewusst, können Intelligenztests wertvolle Dienste leisten.

Vorrangige Kriterien für die Auswahl der Testverfahren für die Intelligenzdiagnostik sollten die so genannten Testgütekriterien Objektivität, Reliabilität und Validität sowie die Normierung der Testwerte sein. Es sollten z.B. nur Tests mit einer möglichst aktuellen Normierung verwendet werden. Aktuell wird üblicherweise eine Normierung innerhalb der letzten zehn Jahre bezeichnet. Weiters ist darauf zu achten, dass die Normen differenziert für die entsprechende Altersgruppe und differenziert nach Geschlecht und anderen relevanten Merkmalen vorliegen. Sofern die Normen nicht optimal sind (keine Geschlechts- oder Altersdifferenzierung, veraltete Normen, Grundgesamtheit nicht repräsentativ¹² o.ä.), ist es für Anwender/innen unabdingbar zu wissen, in welchen Punkten die Normierung nicht entspricht, damit dies bei der Testinterpretation berücksichtigt werden kann.

Selbstverständlich sollte das angewendete Testverfahren an Testperson und Zielformulierung angepasst sein, d.h. bei der Testung auf intellektuelle Hochbegabung sollten Intelligenztests verwendet werden, bei welchen Deckeneffekte ausgeschlossen werden können (Auflistung der Tests und deren Anwendungsbereiche sind in Kapitel 3 und im Anhang dieser Broschüre näher beschrieben). Weiters sollten für eine fundierte Diagnostik Intelligenztests mit mehr als nur einer Aufgabenfacette verwendet werden.

Im Folgenden (2.1, 2.2 und 2.3) sollen die Testgütekriterien der Objektivität, der Reliabilität und der Validität definiert und beschrieben werden:

¹² Siehe auch 2.4

2.1 Objektivität

Objektivität bedeutet, dass das Testergebnis unabhängig von der untersuchenden Person festgestellt werden kann. Die *Objektivität* sollte durch die Standardisierung des Verfahrens, d.h. der strukturierten Vorgangsweise in der Anwendung des Tests gegeben sein. Sie ist im Normalfall durch genaue Angaben zur Vorbereitung, zum Wortlaut der Anweisungen und zu den Durchführungszeiten gewährleistet (standardisierte Instruktion).

2.2 Reliabilität

Reliabilität ist ein Maß für die Zuverlässigkeit eines Testverfahrens, d.h. sie gibt an, wie konsistent das Verfahren das gleiche Konstrukt misst. Alle Items sollten die gleiche Eigenschaft oder Charakteristik messen (*interne Konsistenz*). Ein reliabler Test erbringt auch bei einer Wiederholungsmessung (*Retest-Reliabilität*) bzw. der Messung mit jeweils nur der Hälfte der Aufgaben (*Split-Half-Reliabilität*) vergleichbare Ergebnisse. Der Kennwert für Reliabilität sollte nahe bei 1 liegen, auf jeden Fall über .70.

2.3 Validität

Die Validität (Gültigkeit) zeigt an, wie gut ein Test das erfasst, was er messen soll.

Die Inhaltsvalidität (content validity) beschreibt, inwieweit die Fragen in einem Test den Bereich, den man untersuchen will, abdecken. Will man z.B. das Leseverständnis einschätzen, so werden in dem Test Verständnisfragen vorhanden sein.

Die Konstruktvalidität erfasst, wie gut ein Verfahren den relevanten Bereich abdeckt, aber auch, wie gut es das Konstrukt operationalisieren kann. Beispielsweise wird angenommen, dass sich Kreativität aus Fluidität, Flexibilität, Elaboration und Originalität zusammensetzt. Ein Kreativitätstest sollte demnach geeignet sein, alle diese Faktoren einzuschätzen. Als Maß für die Konstruktvalidität werden u.a. die konvergente, die diskriminante (oder auch divergente) und die faktorielle Validität herangezogen. Die konvergente Validität ist gegeben, wenn die Messdaten von Testverfahren, die dasselbe Konstrukt messen wollen, hoch miteinander korrelieren. Die diskriminante Validität ist gegeben, wenn die Messdaten von Testverfahren, die verschiedene Konstrukte messen, nur gering miteinander korrelieren (im besten Fall mit $r = .00$).

Die faktorielle Validität misst, ob die Items einer als eindimensional konzipierten Skala auf einem gemeinsamen Faktor laden oder ob sich die theoretisch angenommenen Dimensionen in

einer Faktorenanalyse als Faktoren wiederfinden lassen. Dies lässt sich mithilfe explorativer und/oder konfirmatorischer Faktorenanalysen überprüfen. In diesen Verfahren werden die Ladungen der einzelnen Items auf einen oder mehrere Faktoren überprüft: Laden Items hoch auf einen Faktor ($> .4$) können sie diesem zugeordnet werden. In konfirmatorischen Faktorenanalysen werden die Ladungen bestimmter Items auf jeweils einen Faktor überprüft, in explorativen Faktorenanalysen prüft man die Ladungen aller Items auf alle möglichen Faktoren, ohne eine Vorannahme über die Zuordnung der Items zu den Faktoren zu treffen.

Die Kriteriumsvalidität beschreibt den Zusammenhang eines Tests im Vergleich zu einem Außenkriterium, mit dem der Test aufgrund seines Messanspruches korrelieren sollte. Die Kriteriumsvalidität zeigt demnach an, wie stark ein Test mit einem Vergleichskriterium zusammenhängt, z.B. ob der Test eine spätere Leistung vorhersagen kann (prädiktive Validität). Untersuchungen zur Kriteriumsvalidität erfolgen häufig anhand von Korrelationen mit Schulnoten und Fremdeinschätzungen (z.B. von Lehrkräften).

2.4 Normierung

Um relative Aussagen über Leistungen, Fähigkeiten, Eigenschaften oder Merkmale von Personen machen zu können, werden die Testergebnisse einer Person mit denen anderer Personen verglichen. Will man beispielsweise die Frage beantworten, ob eine 12-jährige Testperson intellektuell hochbegabt ist, vergleicht man ihre Testergebnisse mit einer sogenannten Normstichprobe. Damit lässt sich feststellen, wie stark der IQ der Person vom Normal-IQ der Zwölfjährigen (gleichen Geschlechts und vergleichbarer Bildung) abweicht. Wenn er im Vergleich zur Norm besonders hoch ausfällt und unter die 2,2% der höchsten Werte fällt, spricht man in der Regel von intellektueller Hochbegabung.

Im optimalen Fall steht für die Normierung eines Tests eine repräsentative Stichprobe zur Verfügung. Repräsentativ bedeutet, dass die Stichprobe bezüglich der Merkmale (z.B. Alter, Geschlecht, Bildung) und der Merkmals- bzw. Fähigkeits- oder Eigenschaftsausprägung mit der Grundgesamtheit vergleichbar ist. Von dieser Normstichprobe liegen die Ergebnisse des Tests im Normalfall in Form von Summenwerten vor, diese Summenwerte („Rohwerte“) werden auf Normalverteilung¹³ geprüft und – sofern keine signifikante Abweichung von der Normalver-

¹³ In der Normalverteilung gilt, dass im Bereich einer Standardabweichung rechts und links vom Mittelwert immer 68,28 Prozent aller Messwerte liegen (siehe Abbildung zur Gauß'schen Glockenkurve Abb. 8). Im Bereich von zwei Standardabweichungen links und rechts vom Mittelwert befinden sich 95,46 Prozent aller Messwerte und im Bereich von drei Standardabweichungen 99,74 Prozent. Untersucht man eine genügend große Stichprobe, so liegen die Mehrzahl der Messwerte im mittleren Bereich und nur wenige Messwerte in den Extrembereichen (z.B. IQ-Mittelwert= 100; IQ > 130 nur etwa 2% der Population weisen diesen Wert auf).

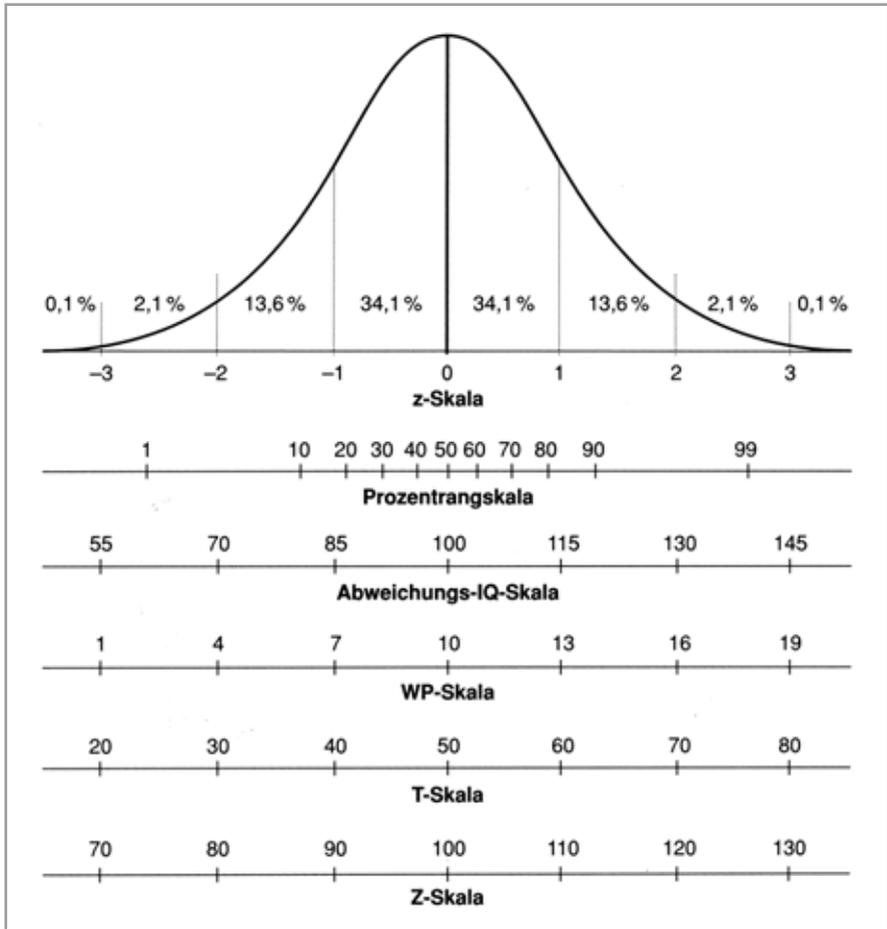


Abb. 8: Gängige Normskalen im Vergleich
 (entnommen aus Holling, Preckel & Vock (2004), S. 83)

teilung festgestellt werden kann – in z-Werte umgerechnet (standardisiert). Die z-Werte können wieder mittels linearer Transformation in verschiedene andere Normskalen transformiert werden. Gängig sind bei Leistungstests die z-Metrik (Mittelwert 0, Standardabweichung 1), die T-Metrik (Mittelwert 50, Standardabweichung 10) oder die allgemein bekannte IQ-Metrik (Mittelwert 100, Standardabweichung 15). Anhand des erreichten Testergebnisses kann man den z-, T- oder IQ-Wert der fraglichen Person ermitteln und somit feststellen, ob und inwieweit die Leistungen durchschnittlich, über- oder unterdurchschnittlich sind.

Oftmals werden zusätzlich zu z-, T- oder IQ-Werten auch die korrespondierenden Prozentränge (Perzentile) in den Testmanuals angeboten, aus welchen hervorgeht, wieviel Prozent der Normpopulation weniger oder gleich viele Punkte erreichen als bzw. wie die betreffende Person. Ein Prozentrang von 90 bedeutet beispielsweise, dass 90% der Vergleichspopulation niedrigere Werte im Test erreichen als die getestete Person. Oft werden neben Prozenträngen auch Staninewerte angegeben, um die Fähigkeit einer Person einschätzen zu können: Die Staninewerte sind Werte zwischen 1 und 9 (Mittelwert 5, Standardabweichung 2).

Im Gegensatz zu anderen Normen eignen sich Prozentränge und Staninewerte auch zur Beschreibung schief verteilter Testwerte. Sie stellen keine lineare Transformation der Itemrohwerte dar. Werden bei normal verteilten Merkmalen keine Prozentränge angeboten, können diese in den entsprechenden Tabellen nachgeschlagen werden.

Literatur

- Funke, J. & Vaterrodt-Plünnecke, B. (1998). Was ist Intelligenz? München: Beck.
- Heller, K. A. & Perleth, C. (2007). MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe. Göttingen: Hogrefe.
- Holling, H., Preckel, F. & Vock, M. (2004). Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6. Göttingen: Hogrefe.
- Ricken, G., Fritz, A., Schuck, K. D. & Preuss, U. (2007). HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter – III. Bern: Huber.
- Wechsler, D. (1944). The measurement of adult intelligence (3rd ed.). Baltimore: Williams & Wilkins.

3 Verfahren zur Messung von intellektuellen Fähigkeiten



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 des gewählten Verfahrens zu schwer für eine Testperson, löst sie also kaum Aufgaben richtig, so zeigt sich ein so genannter Bodeneffekt. Enthält ein Test keine oder zu wenige ausreichend schwere Aufgaben, ergibt sich ein so genannter Deckeneffekt. Boden- und Deckeneffekte verhindern die Abschätzung der wahren Fähigkeit der Person und erbringen ungenaue Schätzungen.

Will man Intelligenzdiagnostik im weit unter- bzw. überdurchschnittlichen Bereich durchführen, müssen Testverfahren ausgewählt werden, deren Eignung für diese Extrembereiche nachgewiesen ist. Für alle in dieser Broschüre vorgestellten Verfahren werden deshalb Hinweise und Kommentare bezüglich der Eignung für die Hochbegabungsdiagnostik bzw. Minderbegabungsdiagnostik angeboten.

Zur besseren Übersichtlichkeit werden die Tests alphabetisch in der Reihenfolge ihrer Testkürzel vorgestellt.

Die Gliederung ist für alle vorgestellten Tests gleich. Nach einer Beschreibung (Anwendungsbereich, Intelligenzmodell etc.) werden Angaben zur Durchführung sowie zur Auswertung und Interpretation dargeboten (beides entnommen aus dem *Psyndex Test Review*¹⁴ und in einigen Fällen zusätzlich aus den Manualen der Tests). Nach dieser allgemeinen Vorstellung der Tests sind die Testgütekriterien¹⁵ überblicksartig angeführt. Abgeschlossen wird jede Testbeschreibung durch einen Kommentar der Autorinnen dieser Broschüre sowie einem Hinweis auf besondere Aspekte der Begabungsdiagnostik (Hochbegabung, Minderbegabung, Underachievement).

¹⁴ PSYINDEX Tests: Ziel des Datenbanksegments PSYINDEX Tests ist es, umfassende Informationen über Testverfahren für den Bereich der Testdiagnostik zur Verfügung zu stellen. Dazu werden in den deutschsprachigen Ländern angewandte Tests, Skalen, Fragebögen, Interviewmethoden, Beobachtungsmethoden, apparative Testverfahren, Methoden der computergestützten Diagnostik und andere diagnostische Instrumente aus allen Bereichen der Psychologie und der Pädagogik nach einem standardisierten Raster beschrieben und bewertet. Gegenwärtig enthält PSYINDEX Tests mehr als 6000 Testnachweise.

PSYINDEX Tests ist Teil der kostenpflichtigen Referenzdatenbank PSYINDEX, die aus den Segmenten PSYINDEX Literatur und AV-Medien sowie PSYINDEX Tests besteht. PSYINDEX, die keine Volltexte enthält, wird hergestellt und vertrieben vom Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). Das ZPID ist die überregionale Dokumentations- und Informationseinrichtung für das Fach Psychologie in den deutschsprachigen Ländern. Es ist das Forschungsdatenzentrum für die Psychologie. Das ZPID informiert Wissenschaft und Praxis aktuell und umfassend über psychologisch relevante Literatur, Testverfahren, audiovisuelle Medien, Forschungsdaten und Psychologie-Ressourcen im Internet.

Kontaktdaten: Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID)

Universität Trier

D-54286 Trier

Telefon: +49 (0)651 201-2877 Fax: +49 (0)651 201-2071

E-Mail: info@zpid.de Internet: www.zpid.de

¹⁵ siehe Definitionen in Kapitel 2 dieser Broschüre

3.1 AID 2; Adaptives Intelligenz Diagnostikum 2

Kubinger, K. D. & Wurst, E. (2000). AID 2. Adaptives Intelligenz Diagnostikum 2. Göttingen: Beltz.

Beschreibung

Das AID 2 dient der Erfassung komplexer und basaler Kognition von Kindern und Jugendlichen zwischen 6 und 15 Jahren und eignet sich für eine förderungsorientierte Diagnostik. Es handelt sich um einen außerordentlich sorgfältig konstruierten Test. Nach einer sicher längeren Einarbeitungsphase kann das AID 2 ökonomisch angewendet werden. Für praktisch tätige Psychologinnen/Psychologen und Lehrer/innen wird es von großem Nutzen sein. Das AID 2 ist gut geeignet für die Diagnostik von (Hoch)Begabten.

Das AID 2 stellt die erweiterte und neu normierte Testfassung des AID dar (Kubinger & Wurst, 1991). Inhaltlich orientiert sich das AID 2 am Testkonzept von Wechsler (1939) und weist diesbezüglich inhaltlich an den HAWIK angelehnte Subtests auf. Intelligenztheoretisch liegt dem AID 2 die Position zugrunde, „ziemlich viele (komplexe und basale) Fähigkeiten, die für ‚intelligentes‘ Verhalten verantwortlich scheinen, zu erfassen“ (Kubinger & Wurst, 2000, S. 14).

„Das AID 2 verfolgt das Ziel einer förderungsorientierten Diagnostik und erteilt der Generalfaktortheorie im Sinne von Spearman eine Absage: Ein Intelligenzquotient, definiert als der Durchschnitt aller geprüften Fähigkeiten, ist intelligenztheoretisch nicht mehr vertretbar und förderungsorientiert nicht zielführend“ (Kubinger & Wurst, 2000, S. 14).

Kubinger & Wurst (1991) versuchen mit den Testergebnissen des AID Hinweise auf entsprechende Fördermöglichkeiten abzuleiten, indem besonderer Wert auf die Interpretation der einzelnen Untertests zueinander und die damit verbundene differenzialdiagnostische Bedeutung gelegt wird.

Ausgangspunkt zur Konzeption der einzelnen Untertests waren operationale Definitionen, wozu u.a. Expertenratings und empirische Analysen herangezogen wurden. Um eine möglichst große Unabhängigkeit von der Zugehörigkeit zu einer Sozialschicht zu erreichen und nicht nur schulisch vermitteltes Wissen abzutesten, wurden Themenbereiche gesammelt, die ein Kind/Jugendlicher durch interessiertes, forschendes Verhalten wissen kann. Hierzu wurde auf Lern- und Quizspiele, Rundfunk- und Fernsehsendungen, Kinderliteratur und Lehrpläne der Schulen zurückgegriffen.

Das AID 2 besteht aus 11 obligatorischen Untertests und drei fakultativ vorzugebenden Zusatztests, die zwei Bereichen zugeordnet werden:

(I) *Verbal-akustische Fähigkeiten*

- 1 Alltagswissen (Fähigkeit, sich Sachkenntnisse über Inhalte anzueignen)

- 3 Angewandtes Rechnen (passende Rechenoperationen zur Problemlösung alltäglicher Aufgabenstellungen)
- 5 Unmittelbares Reproduzieren numerisch (richtige Wiedergabe einer vorgegebenen Zahlenfolge)
- 6 Synonyme finden (zu einem vorgegebenen Wort ein anderes Wort mit möglichst gleicher Bedeutung finden)
- 9 Funktionen abstrahieren (aus zwei Objekten die gemeinsame wesentliche Funktion erkennen)
- 11 Soziales Erfassen und Reflektieren (Sachzusammenhänge der sozialen Umwelt begreifen; Wissen über sozial angepasste Verhaltensweisen und gesellschaftliche Bedingungen)

(II) *Manuell-visuelle Fähigkeiten*

- 2 Realitätssicherheit (auf Bildern einen Gegenstand mit fehlendem Detail entdecken)
- 4 Soziale und sachliche Folgerichtigkeit (eine Bildergeschichte logisch richtig ordnen)
- 7 Kodieren und Assoziieren (zunächst nach einer Vorlage und später aus dem Gedächtnis heraus Symbole kodieren)
- 8 Antizipieren und Kombinieren – figural (Teile einer Figur zusammensetzen)
- 10 Analysieren und Synthetisieren – abstrakt (geometrische Muster mit verschiedenfarbigen Seiten von Würfeln nachlegen)

Durch Zusatztests gemessene Fähigkeiten

- 5a Unmittelbares Reproduzieren – figural/abstrakt
- 5b Merken und Einprägen
- 10a Strukturieren – visomotorisch

Beobachtungen während der Testung können durch das „Beiblatt für Beobachtungen der Arbeitshaltungen“ systematisiert werden.

Altersgruppe: 6;0 bis 15;11 Jahre

Hinweise zur Durchführung

- Die Testdauer beträgt 20–60 Minuten.
- Der Test wird in Form einer **Einzeltestung** durchgeführt.
- Adaptive Testung:
Der Testperson werden nur die Aufgaben gestellt, die ihrem Leistungsniveau entsprechen. Die Testleiterin/Der Testleiter soll in der Durchführung von psychologischen Einzelverfahren bei Kindern gut ausgebildet sein (z.B. Diplompsychologinnen/-psychologen, Sonderschullehrer/innen, Beratungslehrer/innen).
- Es existiert eine Vorgabevariante für hochbegabte Kinder und Jugendliche.

Hinweise zu Auswertung und Interpretation

Die von 1995 bis 1997 erhobene repräsentative Normierungsstichprobe von 977 Kindern und Jugendlichen im Alter von 6 bis 15 Jahren aus Österreich und Deutschland wurde 2006/07 einer Eichungskontrolle unterzogen, auch für die Muttersprache Türkisch.

Angegeben werden T-Werte und Prozentränge nach Altersjahrgängen und nach Alters- und Geschlechtsverteilung für alle 16 Testkennwerte getrennt nach Hauptschule und Gymnasium. Als globales Maß der kognitiven Fähigkeiten einer Testperson wird wie im AID die sogenannte „(untere Grenze der) Intelligenzquantität“ (niedrigster Untertestwert als „Mindestfähigkeit“ einer Person) definiert (AID 2-Manual, 2000, S. 32). Die inhaltliche Interpretation der Testergebnisse soll sich an den operationalen Definitionen der einzelnen Unter- bzw. Zusatztests orientieren. Teilleistungsstörungen können in der Form von intraindividuellen Leistungsdefiziten im Protokollbogen erkannt werden. Die Testleistungen in den eher förderabhängigen Untertests (Alltagswissen, Synonyme finden, soziales Erfassen und sachliches Reflektieren) können als Hinweis für das Ausmaß an Förderung, die einem Kind bisher zuteil wurde, gelten und erlauben eine Prognose für künftige Förderung.

Testgüte

Reliabilität

Die innere Konsistenz des AID 2 ist durch die Rasch-Modellierung gegeben. Die Split-Half-Reliabilitäten der Subtests liegen für das ursprüngliche AID zwischen $r = .70$ (Realitätssicherheit) und $r = .95$ (u.a. Alltagswissen), die Retest-Reliabilität nach 4 Wochen beträgt zwischen $r = .67$ und $r = .95$, nach 1 Jahr zwischen $r = .39$ und $r = .89$. Die Retest-Reliabilitäten einiger AID 2-Untertests lagen zwischen $r = .57$ und $r = .89$.

Validität

- Faktorenanalysen zum AID und AID 2 ergaben eine Vierfaktorenstruktur (Kubinger & Wurst, 2000, S. 30 ff.). Die Faktoren werden als „Informationsverarbeitung in der gesellschaftlichen Umwelt“, „Informationsverarbeitung neuer Inhalte“, „Auffassungskapazität“ und „(Re-)Produktionsfähigkeit durch Strukturierung“ interpretiert.
- Die inhaltliche Gültigkeit ist auf Grund von Expertenratings gegeben.
- Konstruktvalidität: Die Vermutung, dass der Wiener Matrizentest, die Standard Progressive Matrices (SPM) und der CFT 20 grundsätzlich andere Fähigkeiten messen als das AID, bestätigte sich weitgehend. Korrelationen der AID-Subskalen mit CFT 20: AID Angewandtes Rechnen $r = .37$; AID Alltagswissen $r = .49$, AID Funktionen abstrahieren $r = .22$; Korrelationen der AID-Testkennwerte mit WMT: AID Alltagswissen $r = .20$; AID Realitätssicherheit $r = .27$; AID Angewandtes Rechnen $r = .21$; AID Analysieren und Synthetisieren – abstrakt $r = .48$.
- Die diskriminante Validität konnte durch zahlreiche Leistungstests und etliche Persönlichkeitsfragebogen belegt werden.
- Eine Kriteriumsvalidierung wurde nicht vorgenommen.

- Ein Extremgruppenvergleich zwischen 40 lernbehinderten Kindern und 22 Kindern von Eltern, die Mitglieder des internationalen Vereins MENSA¹⁶ waren, ergab signifikante Unterschiede. Während die Gruppe der Kinder von MENSA-Mitgliedern in sämtlichen Untertests kaum auf T-Werte unter 45 kam, erreichte die Gruppe der lernbehinderten Kinder kaum T-Werte über 50 (Kubinger & Wurst, 2000, S. 24).

Autorinnenkommentar

Das AID 2 bietet als einer der wenigen deutschsprachigen Tests die Möglichkeit des adaptiven Testens, die gerade im Bereich der Intelligenzdiagnostik sehr viele Vorteile (Erhöhung der Messgenauigkeit, Verringerung der Testlänge, Optimierung der Entscheidungsstrategie) mit sich bringt (siehe auch Kapitel 4). Das dort 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 deshalb hervorragend für die Testung (hoch)begabter Kinder.

Literatur

- Cattell, R. B. (1987). Intelligence: its structure, growth, and action. Amsterdam: Elsevier.*
- Kubinger, K. D. & Wurst, E. (1991). AID. Adaptives Intelligenz Diagnostikum (3., ergänzte Auflage). Weinheim: Beltz.*
- Kubinger, K. D & Wurst, E. (2000). AID 2. Adaptives Intelligenz Diagnostikum 2. Göttingen: Beltz.*
- Wechsler, D. (1939). The measurement of adult intelligence. Baltimore: Williams & Wilkins.*

¹⁶ Alle Mitglieder von MENSA müssen einen IQ von mind. 130 nachweisen.



3.2 APM; Advanced Progressive Matrices

Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998). Matrizen-Test-Manual, Band 2. Ein Handbuch mit deutschen Normen zu den Advanced Progressive Matrices von J. C. Raven. Göttingen: Beltz-Test.

Beschreibung

Die APM dienen der Testung des sprachunabhängigen logischen Denkens bzw. logischen Schlussfolgerns überdurchschnittlich begabter Jugendlicher ab 12 Jahren und Erwachsener. Sie gehören zur Verfahrensgruppe der Raven-Matrizen-Tests (RPM; Raven, 1978; Raven, Court & Raven, 1994; Raven, Raven & Court, 2001) und stellen in diesem Rahmen eine Version für ein überdurchschnittliches kognitives Leistungsniveau dar. Durch den Einsatz eines Zeitlimits kann die intellektuelle Effizienz der Probandinnen und Probanden beurteilt werden.

Der APM-Test ist ein Verfahren zur sprachfreien Erfassung des Intelligenzfaktors g in Spearman's Intelligenzmodell (Spearman, 1938). Laut Carpenter, Just & Shell (1990) wird mit den APM die analytische Intelligenz im Sinne von Cattells Konzept der „fluid intelligence“ erfasst. Ausgangslage für die Entwicklung der APM war die Bewährung und die breite Anwendung der Standard Progressive Matrices (SPM). Die Aufgabenschwierigkeit der APM ist höher als bei der SPM- und der Coloured Progressive Matrices-Testreihe (Beschreibungen der SPM und CPM siehe S. 148 bzw. S. 70). Testaufbau und Anspruchsniveau heben sich von den SPM ab und sollen im oberen Testleistungsbereich Deckeneffekte auffangen.

Die APM bestehen aus Set I mit 12 Items (Übungstest) und Set II mit 36 Items, den eigentlichen Testaufgaben. Die Aufgaben enthalten geometrische Figuren oder Muster, die aus jeweils sechs dargebotenen Antwortalternativen ergänzt werden sollen (multiple choice). Die Testitems sind nach aufsteigendem Schwierigkeitsgrad angeordnet.

Die Aufgaben in Set II sollen analytische und integrierende Operationen erfassen, die an höheren Denkprozessen beteiligt sind und dadurch eine klare Unterscheidung zwischen Personen mit überdurchschnittlichen und durchschnittlichen kognitiven Fähigkeiten ermöglichen, was aufgrund der Deckeneffekte bei den SPM nicht möglich ist.

Der Test kann auch bei Personen mit geringen Deutschkenntnissen eingesetzt werden. Nach der Durchführung des Set I besteht je nach Bedarf die Möglichkeit, anstelle von APM Set II, die CPM oder die SPM für die weitere Test-Untersuchung einzusetzen. Die Autoren geben an, dass die Computer- und Papier-Bleistift-Testversionen vergleichbar sind, wenn am Computer eine Korrekturmöglichkeit besteht. Für jüngere Kinder stehen die CPM zur Verfügung.

Altersgruppe: 12 bis 60 Jahre

Hinweise zur Durchführung

Es wird empfohlen, beide Aufgabensets ohne zeitliche Begrenzung vorzugeben. Sollte dennoch ein Zeitlimit gesetzt werden, wird für das APM Set I eine Begrenzung von 10 Minuten, für das APM Set II eine Begrenzung von 40 Minuten vorgeschlagen.

Zur differenzierten Analyse der individuellen kognitiven Leistungsfähigkeit sollte das APM Set I nicht ohne das APM Set II eingesetzt werden, da die Gütekriterien des 12 Items umfassenden Instruments für den alleinigen Einsatz nicht zufriedenstellend erfüllt sind. Bei einer Durchführung als Speedtest sollten ca. 50 Min. eingeplant werden, entsprechend länger bei einer Durchführung als Niveautest. Der Aufbau des Verfahrens ermöglicht den Einsatz auch bei nicht-deutschsprachigen Jugendlichen und Erwachsenen.

Die Papier-Bleistift-Form kann als **Einzel- oder Gruppentest** durchgeführt werden. Auch eine Form für Personen mit beeinträchtigtem Sprachverständnis liegt vor.

Hinweise zu Auswertung und Interpretation

Das Manual von Raven, Raven & Court (1998, S. 84–97) enthält Daten zu Normierungsstichproben aus Australien, China, Großbritannien, Neuseeland, Nordirland, Polen sowie den USA. Die deutschen Referenzwerte basieren auf einer Stichprobe von $N = 1142$ Probandinnen und Probanden (48,5% Frauen, 51,5% Männer, 13–75 Jahre, Erhebung 1997).

Die Auszählung der Anzahl richtiger Antworten wird auf dem Durchschlagblatt durchgeführt und beansprucht ca. 3 Minuten. Die addierten Rohwerte werden in Prozenträngen dargestellt. Die Interpretation des Testresultates geschieht anhand der Prozentränge. Im Weiteren enthält das Manual umfassende Hinweise zur Interpretation der Ergebnisse, zur Verwendung der Normen, zur Auswahl und Anwendung der Normtabellen sowie eine Auswahl von Fallbeispielen.

Testgüte

Reliabilität

Eine ausreichend hohe innere Konsistenz wird für den Gesamttest sowie für das APM Set II bestätigt. Die Retest-Reliabilität liegt gemäß einer Überprüfung der Autoren bei $r = .80$ (3-Monatsintervall, $N = 128$). Der Wert liegt bei Gymnasiastinnen und Gymnasiasten etwas höher ($r = .82$) als bei Hauptschülerinnen und -schülern ($r = .75$). Für hoch begabte Gymnasiastinnen und Gymnasiasten ergab eine Retestung nach einem Jahr einen Koeffizienten um $r = .70$.

Validität

Die Inhaltsvalidität ist schwer nachzuweisen, da das zugrunde liegende Konstrukt „eductive ability“ der APM nur schwer nachzuweisen ist. Im Manual der APM steht: „Unbestreitbar handelt es sich bei den APM-Matrizen ebenso wie bei den SPM-Matrizen um nonverbale Stimuli, die keine verbale Reaktion verlangen. Die Testdurchführung kann auch ohne verbale Instruktion erfolgen. Die postulierte Unabhängigkeit von verbal-mentalenen Prozessen wurde aber mehrfach



widerlegt (Carlson, Goldman, Bollinger & Wiedl, 1974; DeShon, Chan & Weissbein, 1995; Dörner, 1979; Hussy, 1987; Radford, 1966). Vielmehr scheint verbal-analytisches Denken bei einem Teil der Aufgaben funktional zu sein.“

Carpenter, Just & Shell (1990) weisen im Vergleich mit den Lösungsprozessen zum „Turm von Hanoi“ (klassisches Verfahren der Denkpsychologie) darauf hin, dass die APM-Skalen nicht in der Lage sind, jene kognitiven Fähigkeiten zu messen, die allen Menschen gemeinsam sind.

Zur Bestimmung der Konstrukt- und Kriteriumsvalidität wurden verschiedene Intelligenztests und Schulnoten herangezogen. Es bestehen Kor-

relationen zum Kognitiven Fähigkeitstest (KFT, Heller & Perleth, 1998), insbesondere mit der nonverbalen Skala ($r = .51$) und der quantitativen Skala ($r = .40$) in der 9. und 10. Schulstufe einer hochbegabten Gymnasialgruppe bzw. von $r = .26$ und $r = .44$ in der 5. Schulstufe eines Regelgymnasiums. Zwischen dem APM Set II und dem Matrizentest CFT 20 (Weiß, 1978) wird eine Korrelation von $r = .56$ berichtet, für den Zahlenverbindungstest (ZVT; Oswald & Roth, 1978) eine Korrelation von $r = .33$.

Hinsichtlich der externen kriterienbezogenen Validität über die Schulleistung zeigt sich eine signifikante Korrelation mit der Mathematiknote ($r = .35$) und mit der Deutschnote ($r = .23$), wenn die Datenerhebung zeitgleich durchgeführt wird.

Weitere Hinweise

Die Schwierigkeitsindices der meisten APM-Testitems liegen größtenteils im erwünschten Rahmen. Im Manual wird darauf hingewiesen, dass die Trennschärfe eines Teils der Items der APM trotz guter Reliabilität des Gesamttests zu schwach ist.

Autorinnenkommentar

Die APM sind populäre Verfahren zur Erfassung der fluiden Intelligenz und gelten als bekannteste Vertreter figuraler Matrizentests. Bezogen auf die Intelligenzdiagnostik 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 Schulstufen 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 (Schweizer, Goldhammer, Rauch & Moosbrugger, 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“). Mit den Schulnoten zeigen sich bei den APM schwache bis mittlere Zusammenhänge, wobei die höchsten Korrelationen mit der Schulnote Mathematik errechnet werden konnten ($r = .34$ bis $r = .40$).

Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den APM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der APM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe auch Preckel, 2010).

Literatur

- Carlson, J., Goldman, R., Bollinger, J. & Wiedl, K. H. (1974). Der Effekt von Problemverbalisation bei verschiedenen Aufgabengruppen und Darbietungsformen des „Raven Progressive Matrices Test“. *Diagnostica*, 20, 133-141.
- Carpenter, P. A., Just, M. A. & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404-431.
- Colom, R., Escorial, S. & Rebollo, I. (2004). Sex differences on the progressive matrices are influenced by sex differences on spatial ability. *Personality and Individual Differences*, 37(6), 1289-1293.
- DeShon, R. P., Chan, D. & Weissbein, D. A. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants. *Intelligence*, 21(2), 135-155.
- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Heller, K. A. & Perleth, C. (1998). *Kognitiver Fähigkeitstest für 5. bis 13. Klassen*. Göttingen: Hogrefe.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998a). *Matrizen-Test-Manual, Band 1. Ein Handbuch mit deutschen Normen zu den Standard Progressive Matrices von J. C. Raven*. Weinheim: Beltz.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998b). *Matrizen-Test-Manual, Band 2. Ein Handbuch mit deutschen Normen zu den Advanced Progressive Matrices von J. C. Raven*. Göttingen: Beltz-Test.
- Hussy, W. (1987). Zur Steuerfunktion der Sprache beim Problemlösen. *Sprache & Kognition*, 1, 14-22.
- 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(4), 505-524.
- Kratzmeier, H. (1980). *APM. Manual RAVEN-Matrizen-Test*. Weinheim: Beltz.
- Oswald, W. D. & Roth, E. (1978). *Der Zahlenverbindungstest (ZVT)*. Göttingen: Hogrefe.
- Preckel, F. (2003). *Diagnostik intellektueller Hochbegabung. Testentwicklung zur Erfassung der fluiden Intelligenz*. Göttingen: Hogrefe.
- Preckel, F. (2010). Intelligenztests in der Hochbegabungsdiagnostik. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 19-44). Göttingen: Hogrefe.

- Radford, J. (1966). Verbalisation effects in a „non-verbal“ intelligence test. *The British Journal of Educational Psychology*, 36(1), 33-38.
- Raven, J. C. (1978). *Advanced Progressive Matrices: Set II (1962 Revision)*. London: H.K. Lewis & Co.
- Raven, J. C., Court, J. H. & Raven, J. (1994). *Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales. Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Raven, J., Raven, J. C., & Court, J. H. (2001). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Oxford, England: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.
- Schweizer, K., Goldhammer, F., Rauch, W. & Moosbrugger, H. (2007). On the validity of Raven's Matrices Tests: Does spatial ability contribute to performance? *Personality and Individual Differences*, 43, 198-201.
- Spearman, C. (1938). Measurement of intelligence. *Scientia*, 64, 75-82.
- Weiß, R. (1971). *Grundintelligenzskala 3 (CFT 3)*. Braunschweig: Westermann.
- Weiß, R. (1978). *Grundintelligenztest Skala 2 (CFT 20)*. Braunschweig: Westermann.

3.3 AzN 4+; Aufgaben zum Nachdenken 4+

Hylla, E., Kraak, B., Horn, H., Schwarz, E. & Kühn, R. (1993). Aufgaben zum Nachdenken 4+. Göttingen: Hogrefe. Herausgegeben vom Deutschen Institut für Internationale Pädagogische Forschung. 1993 ist der Test in einer 2. Auflage erschienen.

Beschreibung

Die AzN 4+ eignen sich als Begabungstests für den Übergang auf weiterführende Schulen. Sie wurden entwickelt, um für die Übergangsauslese auf weiterführende Schulen ein objektives Verfahren zur Erfassung der intellektuellen Begabung zur Verfügung zu haben. Sie werden für die folgenden Anwendungsbereiche empfohlen: Intelligenzdiagnose, Übergangsempfehlung, Prognose von Schulerfolg und Schulversagen und Bestimmung von so genannten Under- und Overachievern.

Der Test erfasst die intellektuelle Begabung – die für die Autoren mit Intelligenz synonym ist – von Kindern im 4. und 5. Schuljahr. Er besteht aus fünf Subtests, die jeweils 12 bis 20 Einzelaufgaben umfassen. Die Subtests „Rechnen“ und „Zahlenreihen“ erfassen das rechnerische und logische Denkvermögen, die Subtests „Analogien“ und „Satzergänzung“ erfassen verbale und sprachlogische Fähigkeiten. Der Subtest „Instruktionsverständnis“ prüft die Anpassungsfähigkeit des Kindes an verschiedene neuartige Aufgabenstellungen.

Die AzN 4+ dienen der Schullaufbahnberatung, zu deren Verbesserung dieser Test durch die Objektivierung der Begabungseinschätzung beitragen kann. Er ist als Gruppentest konzipiert. Ins-

gesamt gehen 76 Aufgaben in die Auswertung ein, daneben enthält der Test noch weitere 17 Übungsbeispiele (vgl. <http://testzentrale.de/programm/aufgaben-zum-nachdenken.html> [Stand 28.06.2010]).

Altersgruppe: letztes Drittel der 4. und Anfang der 5. Schulstufe

Hinweise zur Durchführung

- Die Testdauer beträgt insgesamt etwa 110 Minuten.
- Die AzN 4+ können als **Einzel- oder Gruppentest** verwendet werden.
- Die Aufgaben sollten von einer geschulten Lehrkraft oder einer Psychologin/einem Psychologen durchgeführt werden. Eine Selbstanwendung wird nicht empfohlen. Eine intensive Einarbeitung für die Durchführung und Auswertung ist notwendig.

Hinweise zur Auswertung und Interpretation

Die Normierungsstichprobe aus dem Jahr 1976 besteht aus ca. 4.700 Schülerinnen und Schülern der 4. und 5. Schulstufe. Es sind detaillierte Auswertungsrichtlinien und Beispiele für die Interpretation der Testergebnisse vorhanden. Die Rohwerte werden in Prozentrang- und T-Werte transformiert. Die Normierung erfolgte für die 4. und 5. Schulstufe und die unterschiedlichen Schultypen.

Testgüte

Reliabilität

Die Split-Half-Test-Reliabilität beträgt $r = .95$; die Retest-Reliabilität innerhalb von 4 Wochen beträgt $r = .90$.

Validität

Die faktorenanalytische Überprüfung der AzN 4+ ergab eine zweidimensionale Struktur des Tests: „Rechnerisch-logisches Denken“ und „Sprachlogische Fähigkeit“. Die AzN 4+ korrelieren mit dem LPS (Leistungsprüfsystem, Horn, 1983) mit $r = .57$, mit dem PSB (Prüfsystem für Schul- und Bildungsberatung; Horn, 1969) mit $r = .67$ (Kinder mit hohem sozioökonomischen Status, $N = 207$) bzw mit $r = .66$ (Kinder mit niedrigem sozioökonomischen Status, $N = 1.872$).

Die AzN 4+ korrelieren mit den Schulnoten mit $r = .61$ für Sachkunde, $r = .63$ für Deutsch und $r = .71$ für Mathematik (4. Schulstufe, $N = 239$).



Die Vorhersagevalidität des schulischen Erfolgs (Kriterium „versetzt/nicht versetzt“) ist mit $r = .62$ nach einem halben Jahr bzw. $r = .60$ nach sechs Jahren hoch ($N = 218$ Realschüler/innen und Gymnasiastinnen/Gymnasiasten).

Weitere Hinweise

Die zweimal 76 Aufgaben der AzN 4+ sind Endprodukt einer mehrmaligen Itemselektion aus insgesamt 525 Aufgaben, die nach Richtlinien der klassischen Testtheorie erfolgte. Für die beiden Parallelformen wurden Aufgaben beibehalten, deren Trennschärfe mindestens $r = .30$ betrug.

Autorinnenkommentar

Die AzN 4+ eignen sich sehr gut zur Prognose von Schulerfolg bzw. -versagen, können aber eine intellektuelle Hochbegabung nicht hinreichend gut prüfen, da oft nur gelernte Inhalte abgefragt werden, die nicht ausreichend zwischen Hochbegabung und Fleiß differenzieren. Eine zusätzliche Individualtestung mit einem weiteren Test (Reasoning) erscheint sinnvoll.

Literatur

- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.
- Horn, W. (1983). *Leistungsprüfsystem (LPS). Handanweisung (2., erweit. und verbesserte Aufl.)*. Göttingen: Hogrefe.
- Hylla, E., Kraak, B., Horn, H., Schwarz, E. & Kühn, R. (1993). *Aufgaben zum Nachdenken 4+*. Göttingen: Hogrefe.

3.4 BIS-4; Berliner-Intelligenzstruktur-Test

Jäger, A. O., Süß, H.-M. & Beauducel, A. (1997). Berliner Intelligenzstruktur-Test (BIS-4). Göttingen: Hogrefe.

Beschreibung

Der BIS-4 ist für die Differenzialdiagnostik der Intelligenz bei 16- bis 19-jährigen Schülerinnen und Schülern geeignet.

Der BIS-Test testet operative und inhaltsgebundene Fähigkeiten. Als operative Fähigkeiten werden getestet: Verarbeitungskapazität (K), Einfallsreichtum (E), Bearbeitungsgeschwindigkeit

keit (B) und Merkfähigkeit (M); als inhaltsgebundene Fähigkeiten werden getestet: sprachgebundenes Denken (V), zahlengebundenes Denken (N), anschauungsgebundenes und figural-bildhaftes Denken (F); und als deren Gesamtwert wird die Allgemeine Intelligenz (AI) ermittelt (siehe auch 1.1.4). Der BIS-Test umfasst 45 sehr verschiedene repräsentativ ausgewählte Aufgabentypen.

Die vier *operativen Fähigkeiten*:

- B *Bearbeitungsgeschwindigkeit*: Diese Fähigkeit gibt Auskunft über das Arbeitstempo, die Auffassungsleichtigkeit und die Konzentrationskraft beim Lösen einfacher, gut strukturierter Aufgaben.
- M *Merkfähigkeit*: Das aktive Einprägen und kurzfristige Wiedererkennen oder das Reproduzieren von Informationen.
- E *Einfallsreichtum*: Gekennzeichnet durch die flexible Produktion von vielfältigen Ideen und Lösungen für eine vorgegebene Problemstellung. Es ist dabei erforderlich, vielfältige Informationen und Vorstellungen kognitiv verfügbar zu haben sowie viele verschiedene Seiten, Varianten, Gründe und Möglichkeiten von Gegenständen und Problemen gleichzeitig zu erkennen.
- K *Verarbeitungskapazität*: Fähigkeit zur Verarbeitung komplexer Informationen bei Aufgaben, die nicht auf Antrieb zu lösen sind, sondern das Heranziehen, das vielfältige Beziehungsstiften, das formallogisch exakte Denken und das sachgerechte Beurteilen von Informationen erfordern.

Die drei *inhaltsgebundenen Fähigkeiten*:

- F *Anschauungsgebundenes, figural-bildhaftes Denken*: Die Fähigkeit zum Umgang mit bestimmtem Aufgabenmaterial, dessen Bearbeitung das figural-bildhafte und/oder das räumliche Vorstellungsvermögen fordert.
- V *Sprachgebundenes Denken*: Der Grad der Aneignung und der Verfügbarkeit des Symbolsystems Sprache.
- N *Zahlengebundenes Denken*: Der Grad der Aneignung und der Verfügbarkeit des Symbolsystems Zahlen.

Altersgruppe: 16- bis 19-jährige Schüler/innen

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt 2,5 Stunden für die Langform und 45 Minuten für die Kurzform. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Im Manual (Jäger et al., 1997, S. 27) ist bezüglich der Anwendungsvoraussetzungen empfohlen, dass vor der Erstanwendung die Untersuchungsleiter/innen den Test im Selbstversuch und in einem Vorversuch mit



einer Gruppe vollständig durchführen sollen. Für eine korrekte Anwendung haben sich diese Vorbereitungen als unverzichtbar erwiesen.

Hinweise zu Auswertung und Interpretation

Die Normierungsstichprobe ist klein: $N = 478$; Alter 16–19 Jahre; im Manual sind keine Jahresangaben hinsichtlich der Durchführungszeit der Normierung ersichtlich.

Jede Aufgabe geht in der Auswertung in drei verschiedene Skalen ein (in die Skala Allgemeine Intelligenz sowie in eine Operations- und in eine Inhaltsskala kognitiver Leistungen). Vorgelegt werden getrennte Standardnormwert-Tabellen für die

Subgruppen der 16- bis 17-Jährigen und der 18- bis 19-Jährigen.

Testgüte

Reliabilität

Die interne Konsistenz der Skalen variiert zwischen $\alpha = .75$ und $\alpha = .89$.

Die Retest-Reliabilitäten bewegen sich zwischen $r = .65$ (Skala E) und $r = .90$ (Skala K) nach vier Monaten, einem Jahr und vier Jahren für die Vorversionen des BIS.

Validität

Faktorenanalysen zeigen eine gute Übereinstimmung der Struktur der Testleistungen mit der Struktur des BIS-Modells. Diese faktorenanalytischen Befunde konnten auch von Huldi (1992, 1997) in einer Untersuchung einer anderen Stichprobe repliziert werden.

Kriteriumsvalidität: Die Korrelation mit den Schulnoten in naturwissenschaftlichen Fächern beträgt $r = .55$, mit Noten in den geisteswissenschaftlichen Fächern beträgt sie $r = .59$.

Autorinnenkommentar

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 zufriedenstellend sind (die Werte liegen zwischen $r = .77$ für die Merkfähigkeit und $r = .93$ für die Allgemeine Intelligenz), die Konstruktreliabilitäten der spezifischen kognitiven Fähigkeiten unabhängig vom verwendeten Koeffizienten aber nicht zufriedenstellend sind. Wenn eine spezifische Diagnose gestellt werden soll, die Testung also über eine allgemeine Einschätzung hinausgehen und wichtige Entscheidungsgrundlage sein soll, sollte ein weiterer Test durchgeführt werden.

Literatur

- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. Auflage). Stuttgart: Kohlhammer.
- Brunner, M. & Süß, H. M. (2007). Wie genau können kognitive Fähigkeiten gemessen werden? Die Unterscheidung von Gesamt- und Konstruktreliabilitäten in der Intelligenzdiagnostik für den Berliner Intelligenzstrukturtest. *Diagnostica*, 53, Heft 4, 184-193.
- Hermann, T. (1976). *Lehrbuch der empirischen Persönlichkeitsforschung* (3. Auflage). Göttingen: Hogrefe.
- Härmann, H. (1964). *Aussagemöglichkeiten psychologischer Diagnostik*. Göttingen: Hogrefe.
- Huldi, M. (1992). Replikation des Berliner Intelligenzstrukturmodells bei Schweizer Mittelschülern. In L. Montada (Hrsg.), *Bericht über den 38. Kongress der Deutschen Gesellschaft für Psychologie in Trier 1992* (Band 1, S. 207-208). Göttingen: Hogrefe.
- Huldi, M. (1997). *Die Intelligenzstruktur von Mittelschülern: Zur Gültigkeit des Berliner Intelligenzstrukturmodells (BIS)*. Bern: Lang.
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*: Göttingen: Hogrefe.
- Jäger, A. O., Süß, H.-M. & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test (BIS-4)*. Göttingen: Hogrefe.
- Kleine, D. & Jäger, A. O. (1987). Replikation des Berliner Intelligenzstrukturmodells (BIS) bei brasilianischen Schülern und Studenten. *Diagnostica*, 33, 14-29.
- Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.
- Westmeyer, H. (1972). *Logik der Diagnostik*. Stuttgart: Kohlhammer.

3.5 BIS-HB; Berliner Intelligenzstruktur-Test für Jugendliche: Begabungs- und Hochbegabtendiagnostik

Jäger, A. O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2006). BIS-HB. Berliner Intelligenzstruktur-Test für Jugendliche: Begabungs- und Hochbegabungsdiagnostik. Göttingen: Hogrefe.

Beschreibung

Der BIS-HB ist ein Intelligenztest für 12- bis 16-Jährige mit besonderem Fokus auf die Diagnostik von Hochbegabung.

Der BIS-HB ist eine überarbeitete und für den entsprechenden Altersbereich der 12- bis 16-Jährigen angepasste Version des BIS-4. Als solcher beruht er ebenso wie die Vorgängerversionen des BIS-4 (BIS-3, BIS-2 und BIS-V) auf dem BIS-Modell (siehe Testbeschreibung BIS-4). Vorversionen des BIS-HB wurden mehrfach getestet und das Verfahren angepasst. Um einen Deckeneffekt bei Hochbegabten zu verhindern, wurden intellektuell besonders begabte Schüler/innen mit in die Normierungsstichprobe aufgenommen.

Der sprachgebundene Test erfasst mit 45 Aufgaben neben der *Allgemeinen Intelligenz* (AI) vier verschiedene *operative Fähigkeiten*:

- (B) Bearbeitungsgeschwindigkeit
- (M) Merkfähigkeit
- (E) Einfallsreichtum
- (K) Verarbeitungskapazität

Zudem unterscheidet er drei *inhaltsgebundene Fähigkeiten*:

- (F) Anschauungsgebundenes, figural-bildhaftes Denken
- (V) Sprachgebundenes Denken
- (N) Zahlengebundenes Denken

Auf der obersten Hierarchiestufe befindet sich die Allgemeine Intelligenz, eine sehr allgemeine Fähigkeit, die allen intelligenten Leistungen zugrunde liegt (Details siehe BIS-4).

Altersgruppe: 12- bis 16-jährige Schüler/innen

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt für die Langform des BIS-HB ca. 170 Minuten, für die Kurzform 60 Minuten. Der Test kann als **Einzel- oder Gruppentest** eingesetzt und sollte unbedingt von geschulten Personen, am besten Psychologinnen und Psychologen, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Der BIS-HB wurde an 1.328 Schüler/innen an Haupt- und Realschulen, Gymnasien und speziellen Schulen für intellektuell besonders begabte Schüler/innen von 2002 bis 2003 normiert. Alle Gesamtpunktwerte der Skalen und die Kennwerte der Skala AI können in Normwerte und Prozentränge transformiert werden. Die Rohwerte einer Testperson können zum Vergleich mit den Leistungen einer Gruppe ausgelesener Jugendlicher, die Schulen für überdurchschnittlich Begabte besuchen, herangezogen werden.

Testgüte

Reliabilität

Die innere Konsistenz der Skalen des BIS-HB variiert zwischen $\alpha = .80$ und $\alpha = .95$. Die Split-Half-Reliabilitäten der einzelnen Skalen variieren zwischen $r = .79$ und $r = .92$. Die Retest-Reliabilitäten nach sechs Monaten variieren zwischen $r = .71$ und $r = .84$. Bei keiner der Skalen ergaben sich Deckeneffekte bei der Testung hochbegabter Schüler/innen.

Validität

Faktorenanalysen bestätigten das BIS-Modell. Die Strukturannahmen des BIS konnten sowohl in durchschnittlich begabten und hochbegabten Stichproben als auch in den vier verschiedenen Altersgruppen nachgewiesen werden.

Die höchsten Zusammenhänge ergaben sich zwischen dem Gesamtschulnotenschnitt und der Allgemeinen Intelligenz mit $r = .54$. Der Zusammenhang zwischen Leistungen in mathematisch-naturwissenschaftlichen Fächern und der Verarbeitungskapazität beträgt $r = .47$, jener zwischen Leistungen in sprachlichen Fächern und sprachgebundenen Denkfähigkeiten beträgt $r = .52$.

Die Skala Verarbeitungskapazität des BIS-HB korreliert mit dem CFT 20 mit $r = .74$; die Skala Anschauungsgebundenes, figural-bildhaftes Denken korreliert mit dem CFT 20 mit $r = .63$.

Autorinnenkommentar

Der BIS-HB ermöglicht eine differenzierte und theoriegeleitete Erfassung hoher Begabung. Bei der Testung Hochbegabter fanden sich keine Deckeneffekte. Beim BIS-HB ist über Gruppenvergleiche der Intelligenzstruktur mittels konfirmatorischer Faktorenanalysen nachgewiesen, dass das zugrunde liegende Modell auch in Begabtenstichproben Gültigkeit besitzt. In einer Rezension folgern Frenzel & Nett (2008): „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.“ Letzteres wird auf die Überrepräsentation von Gymnasiastinnen/Gymnasiasten und Akademikerkindern in der Normstichprobe zurückgeführt, welche keine Schüler/innen aus Förder- und Gesamtschulen enthält. Frenzel & Nett (2008) verweisen auf deutlich höhere Testwerte im CFT 20 bei einer Teilstichprobe der BIS-HB-Eichpopulation im Vergleich zum Normwert für Verarbeitungskapazität aus dem BIS-HB. Der CFT 20 erfasst allerdings nur figurale Verarbeitungskapazität, während der Wert für die Kapazität im BIS-HB alle 3 Inhaltsklassen (figural, verbal, numerisch) erfasst. Beide Konstrukte sind trotz hoher Korrelationen nicht gleichzusetzen. Weiters liefert der CFT 20 aufgrund des Flynn-Effekts vermutlich eine Überschätzung der figuralen Verarbeitungskapazität. Der BIS-HB liefert demnach eher konservative Schätzungen und keine „erhebliche Unterschätzung“.

Literatur

- Frenzel, A. C. & Nett, U. (2008). *Berliner Intelligenzstrukturtest für Jugendliche: Begabungs- und Hochbegabungsdiagnostik (BIS-HB) von A. O. Jäger et al. Diagnostica, 54, 221-225.*
- Jäger, A. O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2006). *BIS-HB. Berliner Intelligenzstruktur-Test für Jugendliche: Begabungs- und Hochbegabungsdiagnostik. Göttingen: Hogrefe.*

3.6 BIVA; Bildbasierter Intelligenztest für das Vorschulalter

Schaarschmidt, U., Ricken, G., Kieschke, U. & Preuss, U. (2004). BIVA. Bildbasierter Intelligenztest für das Vorschulalter. Göttingen: Hogrefe.

Beschreibung

Der BIVA dient der prozessorientierten Diagnostik quantitativer und qualitativer Aspekte der kognitiven Entwicklung im Vorschulalter. Der Test ist das einzige derzeit publizierte Intelligenztestverfahren im Altersbereich 3;6 bis 7;6 Jahren, welches auch die prozessbezogene Diagnostik mit einbezieht.

Das Verfahren stellt eine grundlegende Überarbeitung des *Diagnosticums für bildlich angelegte kognitive Leistung* (BILKOG; Berg & Schaarschmidt, 1989) dar. Die Konstruktion des BILKOG bzw. BIVA basiert auf den seit Mitte der 80er Jahre unternommenen Bemühungen, die Intelligenzdiagnostik mit der experimentellen Kognitionsforschung zu verbinden (z.B. Klix, 1983; Sternberg, 1985; van der Meer, 1998). Ziel des BIVA ist, nicht nur auf der Resultatebene, sondern auch schon auf der Prozessebene Besonderheiten und Unterschiede in der kognitiven Leistungsfähigkeit der Kinder aufzuspüren und das Zusammenwirken einzelner Teilprozesse zum Gegenstand der Interpretation zu machen.

Von den Autorinnen und Autoren des BIVA wird auf das „Vereinfachungsprinzip“ als ein wichtiges Wirkprinzip menschlicher Intelligenz hingewiesen: Intelligente Informationsverarbeitung „vermindert die Fülle des Informationsangebots auf das entscheidungsnotwendige Maß. Dort, wo bei gleicher Problemlage die einfachere Repräsentation gelingt, liegt die höhere geistige Kraft, liegt der höhere Intelligenzgrad vor“ (Klix, 1983, S. 362; zit. nach Schaarschmidt et al., 2004, S. 12).

Der BIVA besteht aus acht Untertests, von denen die Kinder in Abhängigkeit ihres Alters in der Regel jeweils nur vier bis fünf bearbeiten. Jeder Untertest umfasst sechs Aufgaben sowie bei den komplexeren Untertests ggf. zwei Zusatzaufgaben für hoch leistende Kinder.

Untertest 1+2: Objekte-Herauslösen (positiv/negativ)

Gefordert ist das Herauslösen eines Merkmals, z.B. ist auf ein Tier zu zeigen, das sowohl auf dem blauen als auch auf dem grünen Feld vorkommt.

Untertest 3+4: Wort-Bild-Vergleich (positiv/negativ)

Vergleich zwischen unterschiedlich repräsentierten Strukturen. Die zu Grunde liegende semantische Relation soll erkannt werden, z.B. die Bilder zeigen, die zum Wort „weich“ passen.

Untertest 5: Geschichten-Folgen

Drei bis sieben Bilder werden in ungeordneter Reihenfolge vorgegeben und sind so zu legen, dass eine Geschichte entsteht. Der Aufbau der Geschichten verlangt die Beachtung von Ursache-Folge-Relationen.

Untertest 6: Reihen Fortsetzen

Es werden drei Muster mit vier verschiedenen Tieren vorgegeben, das vierte Muster sollte von der Testperson nach der Regel der vorhergegangenen Muster herausgefunden werden. Die Fortsetzung der Reihen verlangt die Beachtung von räumlichen und Mengenrelationen.

Untertest 7: Geschichten-Analogien

Auf Vorlagebildern wird eine Geschichtenabfolge vorgegeben. Die Testperson soll ungeordnete Bilder so legen, dass eine analoge Geschichte zu den Vorlagebildern entsteht. Erfasst wird das Erkennen der Ursache-Folge-Reaktionen über Vergleichsprozesse.

Untertest 8: Reihen-Analogien

Es werden drei Reihen mit jeweils vier Tieren vorgegeben. Die Testperson soll eine neue Tierreihe so aufbauen, dass die Relation der Reihe den vorgegebenen drei Reihen entspricht. Gefordert ist das Erkennen von Invarianten in der jeweiligen Folge.

Der BIVA beschränkt sich auf die Verwendung bildlichen Materials, verbale Anforderungen sind auf ein Mindestmaß beschränkt. Es werden Rückmeldungen und abgestufte Hilfen an die Kinder gegeben, um das Leistungspotenzial besser erschließen zu können. Zur Unterstützung der prozessorientierten Diagnostik werden untersuchungsbegleitende Verhaltensbeobachtungen erfasst. Speziell in der sonderpädagogischen Literatur wird diese Forderung seit Jahren mit Nachdruck erhoben (vgl. u.a. Schuck, 2001; Klauer, 2003).

Die Grundidee des Verfahrens besteht darin, möglichst klar herauszuarbeiten, welche kognitiven Leistungen beim Lösen einer Aufgabe erbracht bzw. beim Nichtlösen nicht erbracht wurden. Auf diese Weise sollen nicht nur auf der Resultatebene, sondern auch auf der Ebene einzelner Teilprozesse und ihres Zusammenwirkens Unterschiede in der kognitiven Leistungsfähigkeit ermittelt werden.

Schaarschmidt et al. (2004) differenzieren die kognitiven Teilprozesse in:

- (1) Herauslösen von Merkmalen, das zur Strukturbildung in der Wahrnehmung und zur Aktivierung von Wissen führt,
- (2) Vergleichen von Merkmalen und Merkmalssätzen, um Relationen zu erkennen bzw. herzustellen,
- (3) Übertragen von Relationen von einer Struktur auf eine andere, um Analogien zu erkennen oder zu generieren.





Diese kognitiven Teilprozesse werden in drei Untertests in Abhängigkeit vom Alltagswissen erfasst, d.h. das Kind kann bei der Aufgabenlösung unmittelbar auf seine Alltagserfahrungen zurückgreifen. In drei Untertests muss die Lösung allein aus dem Material heraus erfolgen, d.h. die Aufgaben sind unabhängig vom Alltagswissen.

Altersgruppe: 3;6 bis 7;6 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt für jüngere Kinder ca. 30 Minuten, für ältere Kinder ca. 60 Minuten.

Der BIVA kann als **Einzel- oder Gruppentest** durchgeführt werden. Die Testleiter/innen sollten sich in das Verfahren gut eingearbeitet haben. Hierfür wird die Nutzung des Demonstrationsvideos empfohlen.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte an 2.287 Kindern aus Deutschland, Österreich und der Schweiz, der Erhebungszeitraum ist nicht dokumentiert.

Die Berechnung der Testergebnisse erfolgt einerseits an der Lösung der Aufgabe, andererseits an den für die gefundene Lösung notwendigen Hilfestellungen durch die Testleiterin/den Testleiter. Es existieren Geschlechts- und altersspezifische Normen für 14 Altersgruppen in T-Werten und Prozenträngen.

Für die Zusatzaufgaben zur besseren Differenzierung bei hoher Begabung kann die vom Testkind erzielte Zusatzleistung in drei Stufen (oberes, mittleres oder unteres Drittel) eingeordnet werden.

Testgüte

Reliabilität

Die innere Konsistenz (Cronbachs Alpha) der einzelnen Untertests liegt zwischen $\alpha = .65$ und $\alpha = .88$. Es sind keine Angaben zur Retest-Reliabilität des BIVA vorhanden.

Validität

Die Korrelationen der BIVA-Untertests mit inhaltlich ähnlichen KFT-Skalen: Die Korrelationen für das Beziehungs-Erkennen haben Werte zwischen $r = .54$ und $r = .64$, für rechnerisches Denken betragen die Korrelationen zwischen $r = .32$ und $r = .46$, für die KFT-Skala Schlussfolgerndes

Denken betragen die Korrelationen zwischen $r = .11$ und $r = .25$, für die KFT-Sprachverständnis-Skala ergibt sich eine Korrelation von $r = .36$ mit dem BIVA-Subtest GF.

Mit der K-ABC ergeben sich geringe bis mittlere Korrelationen zwischen $r = .22$ und $r = .61$. Prognosen der Lehrer/innen und Erzieher/innen hinsichtlich Sprachen und Mathematik korrelieren mit dem BIVA zwischen $r = .16$ und $r = .29$.

Autorinnenkommentar

Der BIVA eignet sich nicht für die Begabungsdiagnostik über dem Durchschnittsbereich. Es treten schon bei Intelligenzleistungen, die mehr als eine Standardabweichung über dem Mittelwert liegen, Deckeneffekte auf.

Literatur

- Berg, M. & Schaarschmidt, U. (1989). *Diagnosticum für bildlich angeregte kognitive Leistung – BILKOG. Intelligenztest für das Kindergarten- und frühe Schulalter*. Berlin: Psychodiagnostisches Zentrum der Humboldt-Universität.
- Cattell, R. B. (1987). *Intelligence: its structure, growth, and action*. Amsterdam: Elsevier.
- Klauer, K. J. (2003). *Über internationale Entwicklungstendenzen sonderpädagogischer Diagnostik unter förderdiagnostischem Aspekt*. In G. Ricken, A. Fritz & C. Hofmann (Hrsg.), *Förderbedarf der sonderpädagogischen Diagnostik* (S. 127-142). Lengerich: Pabst Science Publisher.
- Klix, F. (1983). *Begabungsforschung – ein neuer Weg in der kognitiven Intelligenzdiagnostik?* *Zeitschrift für Psychologie*, 191 (4), 360-387.
- Schaarschmidt, U., Ricken, G., Kieschke, U. & Preuss, U. (2004). *BIVA. Bildbasierter Intelligenztest für das Vorschulalter*. Göttingen: Hogrefe.
- Schuck, K. D. (2001). *Diagnostik*. In J. Borchert (Hrsg.), *Handbuch der Sonderpädagogischen Psychologie* (S. 233-248). Göttingen: Hogrefe.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Van der Meer, E. (1998). *Intelligenz als Informationsverarbeitung*. In E. Roth (Hrsg.), *Intelligenz* (S. 161-184). Stuttgart: Kohlhammer.

3.7 BT 1-2; Bildertest 1-2

Horn, H. & Schwarz, E. unter Mitarbeit von G. Vieweger (1977). Bildertest BT 1-2. Deutsche Bearbeitung des Moray House Picture Intelligence Test 1. Neubearbeitung von R. Kühn und R. Heck-Möhling (2. Auflage, 1. Auflage 1967). Weinheim: Beltz.

1994 erschien ein unveränderter Nachdruck der 2., überarbeiteten Auflage: Deutsches Institut für Internationale Pädagogische Forschung (Hrsg., 1994). BT 1-2. Bildertest 1-2 (unveränderter Nachdruck der 2., überarbeiteten Auflage) [Testmappe mit Beiheft, Mustertestheft, Auswertungsschlüssel und Klassenliste]. Weinheim: Beltz.

Beschreibung

Der BT 1-2 ist ein sprachfreier Test zur Erfassung der allgemeinen intellektuellen Leistungsfähigkeit, der bereits in den ersten Schultagen anwendbar ist. Er ermöglicht einen Überblick über die intellektuellen Fähigkeiten und den allgemeinen Begabungsstand von 6- bis 8-jährigen Schülerinnen und Schülern. Es werden verschiedene Aspekte der Intelligenz erfasst. Die Items in Bilderform erhöhen die Testmotivation.

Der BT 1-2 besteht aus acht Subtests, mit jeweils zehn bis 13 Aufgaben, insgesamt umfasst er 91 Aufgaben zu folgenden Bereichen:

- | | |
|--------------------|--|
| (1) Instruktionen | Verstehen und Durchführen von Anweisungen |
| (2) Nichtpassendes | Aufgaben zur Begriffsbildung, Oberbegriffe finden |
| (3) Ergänzungen | Genauigkeit und Differenziertheit der optischen Wahrnehmung |
| (4) Unsinniges | Genauigkeit und Differenziertheit der optischen Wahrnehmung |
| (5) Spiegelbilder | zusätzliche Unterscheidung von Spiegelbildern und Nicht-Spiegelbildern |
| (6) Folgen | richtige Anordnung zeitlicher Vorgänge von Größen und Mengen |
| (7) Wesentliches | Umweltverständnis |
| (8) Reihen | Gesetzmäßigkeiten finden und nachvollziehen |

Altersgruppe: 6 bis 8 Jahre

Hinweise zur Durchführung

Die Durchführung nimmt in etwa 110 Minuten in Anspruch, daher sollte der Test an 2 Tagen durchgeführt werden.

Er kann als **Einzel- oder Gruppentest** von Psychologinnen/Psychologen oder Lehrerinnen/Lehrern durchgeführt werden. Eine Selbstanwendung ist nicht zu empfehlen. Für die Durchführung und Auswertung ist eine intensive Einarbeitung notwendig.

Hinweise zu Auswertung und Interpretation

Die Normierungsstichprobe (N = 2.500) aus den Jahren 1975/76 wird als repräsentativ für Grundschüler/innen bezeichnet. Eine Neunormierung liegt nach Angaben der Autorinnen und Autoren vor. Es liegen Prozentrang- und T-Wert-Normen sowie die Umwandlung von T-Werten in Abweichungsintelligenzquotienten in Halbjahres- und Vierteljahresabschnitten vor.

Testgüte

Reliabilität

Die interne Konsistenz der BT 1-2 Subtests ist mit $\alpha = .34$ gering und weist auf die Messung unterschiedlicher Aspekte der intellektuellen Begabung hin. Gute Retest-Reliabilität nach 4–6 Wochen: 1. Schuljahr (N = 278): $r = .75$; 2. Schuljahr (N = 259): $r = .85$; Hohe Split-Half-Reliabilität von $r = .95$.

Validität

Die Konstruktvalidität zwischen dem BT 1-2 und der Columbia Mental Maturity Scale 1-3 (CMM 1-3; Schuck, Eggert & Raatz, 1975) beträgt $r = .62$ (N = 372). Untersuchungen mit anderen Intelligenztests (Primary Mental Abilities von Thurstone (1938) und PSB von Horn (1969)) ergaben Zusammenhänge zwischen $r = .58$ und $r = .67$.

Autorinnenkommentar

Der BT 1-2 ist ein ansprechender Test für Kinder im Grundschulalter. Erfahrungsgemäß machen ihn die Kinder sehr gerne und haben Spaß an den Aufgaben. Intellektuelle Hochbegabung kann mit ihm allerdings nicht abgeprüft werden, da die Testaufgaben zumeist sehr einfach sind („Deckeneffekte“) und auch die Trennschärpen der Items in vielen Fällen nicht hinreichend groß sind. Zudem ist die Reliabilität (interne Konsistenz) des Tests gering, weshalb eine „allgemeine Intelligenz“ im Sinne eines Gesamt-IQs ohnehin nicht festgestellt werden kann. Auch eine Mehrdimensionalität des Tests ist nicht auszuschließen (vgl. auch Kühn, 1974). Spezifische Neigungen und auch Begabungen können aufgrund der Testergebnisse angenommen werden. Bevor eine intellektuelle Hochbegabung diagnostiziert wird, sollte aber unbedingt ein zweiter unabhängiger Test durchgeführt werden.

Literatur

Deutsches Institut für Internationale Pädagogische Forschung (Hrsg., 1994). *BT 1-2. Bildertest 1-2 (unveränderter Nachdruck der 2., überarbeiteten Auflage) [Testmappe mit Beiheft, Mustertestheft, Auswertungsschlüssel und Klassenliste]*. Weinheim: Beltz.

Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.

Kühn, R. (1974). *Untersuchungen zur Dimensionalität des BT 1-2. Mitteilungen und Nachrichten Nr.*

75/76, Deutsches Institut für international pädagogische Forschung, 40-73.

Mellone, M. A. & Thomson, G. H. (1958). *Moray House Picture Intelligence Test 1*. Edinburgh: Darien Press.

Horn, H. & Schwarz, E. unter Mitarbeit von G. Vieweger. (1977). *Bildertest BT 1-2. Deutsche Bearbeitung des Moray House Picture Intelligence Test 1. Neubearbeitung von R. Kühn und R. Heck-Möhling* (2. Auflage., 1. Auflage 1967). Weinheim: Beltz.

Schuck, K. D., Eggert, D. & Raatz, U. (1975). *Columbia Mental Maturity Scale CMM 1-3. Sprachfreier Gruppenintelligenztest für die Grundschule*. Weinheim: Beltz.

Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.

3.8 BT 2-3; Bildertest 2-3

**Ingenkamp, K. (1966). *Bildertest (BT 2-3)*. Weinheim: Beltz.
Seit 1976 liegt eine 2. Auflage vor.**

Beschreibung

Der BT 2-3 ist ein Test zur Erfassung der allgemeinen intellektuellen Leistungsfähigkeit bei Kindern im Alter von 7 bis 9 Jahren. Er ist als Hilfsmittel zur Beurteilung der intellektuellen Leistungsfähigkeit geeignet.

Der BT 2-3 in der Neuauflage von 1976 setzt sich aus fünf Subtests zusammen:

- (1) Anweisungsverständnis
- (2) Unterscheidungsfähigkeit
- (3) Verständnis von Folgen
- (4) Raumorientierung
- (5) Analogiebildung

Insgesamt besteht der Test aus 100 Aufgaben, in denen Bilder oder grafische Symbole vorgegeben werden, aus denen ein oder zwei durch Ankreuzen ausgewählt werden müssen.

Altersgruppe: 7;0 bis 9;3 Jahre

Hinweise zur Durchführung

Die Durchführung nimmt in etwa 80 Minuten in Anspruch.

Der Test kann als **Einzel- oder Gruppentest** von Lehrpersonen oder wissenschaftlichen und studentischen Hilfskräften durchgeführt werden. Eine Selbstanwendung ist nicht zu empfehlen. Für die Durchführung und Auswertung ist eine intensive Einarbeitung notwendig.

Hinweise zu Auswertung und Interpretation

Es liegen Prozentrang- und T-Wert-Normen in Halbjahres- und Vierteljahresabschnitten für einen Altersbereich von 7;0 bis 9;3 Jahren vor. Die Normierung erfolgte 1975.

Testgüte

Reliabilität

Die 5 Subtests zeigen mittelhohe Item-Interkorrelationen zwischen $r = .41$ und $r = .52$ (3. Schulstufe, $N = 326$). Eine Faktorenanalyse ergab einen Faktor „Allgemeine intellektuelle Leistungsfähigkeit“, der 89% der Varianz aufklärte. Die Split-Half-Reliabilität beträgt $r = .95$; die Retest-Parallelform-Reliabilität nach 3 bis 10 Tagen $r = .79$.

Validität

Bereits die 1. Auflage des BT 2-3 zeigte hohe Korrelationen mit anderen Intelligenztests: $r = .78$ mit der WPPSI (Wechsler, 1967/2002), $r = .80$ mit dem HAWIK (Tewes et al., 1999), $r = .83$ mit dem PSB (Horn, 1969) und $r = .80$ mit den AzN 4+ (Hylla et al., 1993).

Dem BT 2-3 wird eine hohe prognostische Validität für den allgemeinen Schulerfolg zugesprochen. Die prognostische Validität des BT 2-3 gemessen in der 3. Schulstufe mit den Schulnoten der 6. Schulstufe beträgt für Deutsch $r = .49$, für Rechnen $r = .58$ und für Fremdsprachen $r = .24$.

Autorinnenkommentar

Der BT 2-3 ist ebenso wie der BT 1-2 ein sehr ansprechender Test für Kinder im Grundschulalter. Positiv sind die Verschiedenartigkeit der Aufgaben und die hohe prognostische Validität für den Schulerfolg. Für differenzielle Intelligenzdiagnosen ist er nicht geeignet. Intellektuelle Hochbegabung kann mit ihm nicht abgeprüft werden, da die Testaufgaben zumeist sehr einfach sind („Deckeneffekte“).

Literatur

- Emmet, W. G. (1956). *Deeside Picture Test for Seven-Years-Old*. London: Harrap.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.



- Hylla, E., Kraak, B., Horn, H., Schwarz, E. & Kühn, R. (1993). *Aufgaben zum Nachdenken 4+*. Göttingen: Hogrefe.
- Ingenkamp, K. (1966). *Bildertest (BT 2-3)*, Weinheim: Beltz.
- Tewes, U., Rossmann, P. & Schallberger, U. (Hrsg., 1999). *HAWIK-III. Hamburg-Wechsler-Intelligenz-Test für Kinder – Dritte Auflage. Übersetzung und Adaptation der WISC-III Wechsler Intelligence Scale for Children – Third Edition von David Wechsler*. Bern: Huber.
- Wechsler, D. (1967/2002). *Wechsler Primary and Preschool Scale of Intelligence™ -Third edition (WPPSI™-III)*. San Antonio, TX: Harcourt Assessment.

3.9 CFT 1; Grundintelligenztest Skala 1

Weiß, R. H. & Osterland, J. (1976). Grundintelligenztest CFT 1. Skala 1. Braunschweig: Westermann.
Seit 1997 liegt eine 5., revidierte Auflage vor: Cattell, R. B., Weiß, R. H. & Osterland J. (1997). Grundintelligenztest CFT 1. Skala 1. Göttingen: Hogrefe.

Beschreibung

Der CFT 1 erfasst die Grundintelligenz bei Kindern. Insgesamt handelt es sich beim CFT 1 um ein bewährtes und einfach einsetzbares Verfahren, das für den Einsatz in der diagnostischen Praxis gut geeignet scheint. Typische Einsatzmöglichkeiten bieten sich bei der Überprüfung kognitiver Komponenten bei Schulschwierigkeiten und in der Diagnostik von Legasthenie.

Dem CFT 1 liegt die Intelligenztheorie von Cattell zugrunde. Cattell geht davon aus, dass der Bereich der allgemeinen intellektuellen Leistungsfähigkeit („general ability“, abgekürzt g) sich in einen fluiden Intelligenzfaktor und in einen kristallisierten Intelligenzfaktor aufgliedern lässt. Die kristallisierte Intelligenz ist am CFT 1 in geringerem Umfang beteiligt (siehe auch Kapitel 1).

Der CFT 1 erfasst im Wesentlichen den Faktor der „General Fluid Ability“, d.h. der Fähigkeit des Kindes, in neuartigen Situationen und anhand von sprachfreiem, figuralem Material Denkprobleme zu erfassen, Beziehungen herzustellen, Regeln zu erkennen, Merkmale zu identifizieren und rasch wahrzunehmen (Weiß & Osterland, 1997, S. 4). Der Test gibt darüber Aufschluss, bis zu welchem Komplexitätsgrad das Kind bereits in der Lage ist, insbesondere nonverbale Problemstellungen zu erfassen und zu lösen. Der kulturelle Einfluss auf die Testleistungen betrifft primär motivationale Aspekte: Die Probandinnen und Probanden müssen nach einem Zeitplan mit Ausdauer und ohne sofortige Belohnung arbeiten.

Die deutsche Version des CFT 1 basiert auf der Grundkonzeption des Culture Fair Intelligence Tests von Cattell (1966), wurde jedoch partiell adaptiert.

Er setzt sich aus fünf Subtests zusammen (Weiß & Osterland, 1997):

- (1) Substitutionen: Unter die Bilder müssen zugehörige Symbole gesetzt werden; eher reproduktiver Aspekt der visuellen Wahrnehmung.
- (2) Labyrinth: Labyrinthreihen müssen richtig durchfahren werden; erfasst eher einen produktiven Aspekt der optischen Wahrnehmung.
- (3) Klassifikationen: In einer Reihe von fünf Figuren ist die jeweils nicht zu den anderen passende Figur zu identifizieren; sprachfreie Allgemeinintelligenz.
- (4) Ähnlichkeiten: Wiedererkennen einer vorgegebenen Figur aus fünf Zeichnungen, von denen vier in Details verändert sind; vergleichendes Inbeziehungsetzen bei figuralen Vorgaben.
- (5) Matrizen: Eine Ergänzungsfigur muss aus 5 unterschiedlichen Figuren ausgewählt werden; Fähigkeit, Regeln und Zusammenhänge erkennen zu können.

Altersgruppe: 5;3 bis 9;5 Jahre

Hinweise zur Durchführung

Die Einzeltestung nimmt ca. 30 Minuten in Anspruch, Gruppentests dauern zwischen 45 und 60 Minuten.

Der CFT1 kann als **Einzel- oder Gruppentest** durchgeführt werden (Kindergarten max. 8 Kinder, Grundschule max. 20 Kinder).

Geeignete Testleiter/innen sind Psychologinnen/Psychologen, Beratungslehrer/innen, Sonderschullehrer/innen, LRS-Trainer/innen und andere in Testdurchführung und Interpretation geübte Lehrpersonen. Eine über die Anwendung der Testanleitung (einschließlich Interpretationshilfen) hinausgehende Einarbeitung ist nicht notwendig.

Hinweise zu Auswertung und Interpretation

Die Normierung des Tests fand 1976 statt (N = 6.078 Kindergarten- und Vorschulkinder sowie Schüler/innen). Eine Überprüfung der Normwerte wurde mit der revidierten Fassung des CFT 1 1995 an N = 1.500 Schülerinnen und Schülern durchgeführt.

Für den CFT 1 liegen Altersnormen für 5;6 bis 9;5-jährige Kinder in Form von Prozenträngen, T-Werten und Intelligenzquotienten vor. Klassennormen bestehen für die 1. bis 3. Klasse Grundschule, zusätzlich existieren Normen für Kindergarten und Vorschule.

Testgüte

Reliabilität

Die Split-Half-Reliabilitätskoeffizienten liegen zwischen $r = .65$ und $r = .86$ und werden als zuverlässig eingeschätzt. Die Retest-Reliabilität nach 7–24 Wochen liegt zwischen $r = .61$ und $r = .86$.



Validität

Faktorenanalysen ergeben einen dominanten Hauptfaktor (Generalfaktor), der bei den jüngeren Kindern einen Varianzanteil von 82%, bei den älteren Kindern einen Varianzanteil von 75% aufklärt. Die Korrelation zwischen der CFT 1-Gesamtleistung und dem HAWIK-Gesamtergebnis (HAWIK, Tewes et al., 1999) liegt bei $r = .66$ (mit dem HAWIK-Handlungsteil bei $r = .66$, mit dem HAWIK-Verbalteil bei $r = .48$). Ebenso werden Korrelationen mit dem BT 1-2 (Horn & Schwarz, 1977) von $r = .51$ und dem BT 2-3 (Ingenkamp, 1966) von $r = .55$ berichtet. Der CFT 1-Gesamtwert korreliert im Fach Mathematik in allen Klassenstufen mit $r = .44$ am höchsten, in Sachkunde beträgt $r = .39$, in Deutsch beträgt $r = .35$.

Autorinnenkommentar

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). Testanwender/innen müssen beim Einsatz des CFT 1 für die Intelligenzdiagnostik mit relativ großen Messfehlern rechnen.

Literatur

Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest CFT 1. Skala 1*. (5. revidierte Auflage). Göttingen: Hogrefe.

Cattell, R. B. (1966). *Culture Fair (or Free) Intelligence Test, Scale 1 (Manual)*. Champaign, Ill.: IPAT.

Horn, H. & Schwarz, E. unter Mitarbeit von G. Vieweger. (1977). *Bildertest BT 1-2. Deutsche Bearbeitung des Moray House Picture Intelligence Test 1. Neubearbeitung von R. Kühn und R. Heck-Möhling* (2. Auflage, 1. Auflage 1967). Weinheim: Beltz.

Ingenkamp, K. (1966). *Bildertest (BT 2-3)*. Weinheim: Beltz.

Tewes, U., Rossmann, P. & Schallenberger, U. (Hrsg., 1999). *Hamburg-Wechsler-Intelligenztest für Kinder III (HAWIK-III)*. Bern: Huber.

Weiß, R. & Osterland, J. (1997). *Grundintelligenztest Skala 1. CFT 1 (Westermann Test)*. 5., revidierte Auflage. Göttingen: Hogrefe.

Weiß, R. H. & Osterland, J. (1976). *Grundintelligenztest CFT 1. Skala 1*. Braunschweig: Westermann.

3.10 CFT 20-R; Grundintelligenztest Skala 2

Weiß, H. R. & Weiß, B. (2006). CFT 20-R. Grundintelligenztest Skala 2 – Revision. Göttingen: Hogrefe.

Beschreibung

Der CFT 20-R ist ein Intelligenztest zur Erfassung der sprachfreien allgemeinen geistigen Leistungsfähigkeit. Er kann zum Screening der logischen Fähigkeiten im Alter von 8–60 Jahren verwendet werden. Es handelt es sich um ein theoretisch fundiertes, nach Maßgabe der klassischen Testtheorie konstruiertes, valides, objektives und reliables Instrumentarium zur Erfassung einer sprachfreien Grundintelligenz. Das Verfahren ist mit geringem Aufwand einsetzbar, wenn es um die Diagnose einer Diskrepanz zwischen Intelligenzpotenzial und tatsächlicher Leistung geht.

Der CFT 20-R stellt eine Weiterentwicklung des CFT 20 von R. B. Cattell (1960) und R. H. Weiß (1978) dar und erfasst das allgemeine intellektuelle Niveau (Grundintelligenz) im Sinne der Cattellschen „General Fluid Ability“. Diese kann als Fähigkeit, komplexe Beziehungen in neuartigen Situationen wahrnehmen und erfassen zu können, umschrieben werden (Weiß & Weiß, 2006). Durch sprachfreie und anschauliche Testaufgaben werden Personen mit schlechten Kenntnissen der deutschen Sprache und mangelhaften Kulturtechniken nicht benachteiligt. Kritisch anzumerken ist allerdings, dass die im CFT 20-R verlangten Fähigkeiten, wie Vergleiche machen, Gemeinsamkeiten finden, Gesetzmäßigkeiten erkennen, Reihenfolgen fortsetzen etc., stark von der Sozialisation und der Schulbildung abhängig sind. Oberstes Ziel zur Weiterentwicklung des CFT 20 (Weiß, 1978) zum CFT 20-R war eine verbesserte Differenzierung im oberen Leistungsbereich mit einer Erhöhung der „Testdecke“ unter Beibehaltung des bewährten Aufbaus und der Struktur des Grundintelligenztests CFT 20 (Weiß & Weiß, 2006).

Der CFT 20-R ist ein vielseitig einsetzbarer und ökonomischer Intelligenztest. Er besteht aus zwei gleichartig aufgebauten Testteilen mit je vier Untertests:

- (1) Reihen fortsetzen: Es soll die Figur gefunden werden, die eine Reihe logisch fortsetzt.
- (2) Klassifikationen: Welche Figur passt nicht zu einer Reihe übriger Figuren?
- (3) Matrizen: Die untere rechte Fläche innerhalb eines Vier- oder Neunfelder-Schemas soll durch eine Figur sinnvoll ergänzt werden.
- (4) Topologische Schlussfolgerung: Die Lage eines Punktes wird innerhalb sich überschneidender geometrischer Formen analysiert.

Seit der 3. Auflage 1987 wurde das Testkonzept durch Aufnahme zweier schulnaher Ergänzungstests erweitert: Durch Wortschatz- (WS) und Zahlenfolgen-Aufgaben (ZF) können verbale und numerische Elemente des Faktors Verarbeitungskapazität erfasst werden. WS und ZF stellen eigenständige Teiltests dar.

Altersgruppe: 8;5 bis 60 Jahre

Hinweise zur Durchführung

Inklusive der allgemeinen Testeinführung dauert der CFT 20-R rund 60 Minuten, bei Durchführung der Kurzform (1. Testteil) rund 35–40 Minuten. Bei Einzeltestung werden in etwa 10 Minuten weniger benötigt.

Er ist als **Einzel- oder Gruppentest** durchführbar. Die Testleiter/innen sollten sich gut mit dem Test, den Zeitbemessungen und der Instruktion auskennen. Bei Gruppenuntersuchungen ab 20 Personen wird die Unterstützung durch eine weitere mit dem Test vertraute Person als hilfreich empfohlen. Als Testanwender/innen kommen vor allem Psychologinnen/Psychologen in der Schul-, Berufs- und Erziehungsberatung sowie Beratungslehrer/innen in Frage; eine besondere Qualifikation ist nicht explizit gefordert.

Hinweise zur Auswertung und Interpretation

Die Normierung wurde 2003 mit rund 4.400 Schüler/innen, schulartenrepräsentativ über sechs deutsche Bundesländer, durchgeführt. Für den Grundintelligenztest CFT 20-R wurden repräsentative Altersnormen von 8;5 bis 19 Jahren in Halbjahres- bzw. Jahresschritten sowie Klassenstufen-Normwerte für das 3. bis 10./13. Schuljahr erstellt. Die Rohwerte können anhand von Normtabellen in IQ-Werte, T-Werte oder Prozentränge umgewandelt werden und es kann ein Testprofil erstellt werden.

Jacobs & Petermann (2007) weisen auf Grenzen des Verfahrens hin, die ihrer Ansicht nach im Bereich möglicher Fehlinterpretationen über- oder unterdurchschnittlicher Intelligenztestwerte liegen. Bei der Diagnose von Hochbegabung kann es zu Fehleinstufungen kommen, wenn nur der CFT 20-R eingesetzt wird. Im Rahmen der Diagnostik von Hochbegabung sollte es allerdings selbstverständlich sein, dass weitere Intelligenztests eingesetzt werden, um die Diagnose abzusichern oder ggf. akzelerationsorientiert zu testen (vgl. Holling, Preckel & Vock, 2004, S. 146-147).

Testgüte

Reliabilität

Für den CFT 20-R werden Testreliabilitätskoeffizienten von $\alpha = .96$ (Gesamttest) und $\alpha = .91$ für den 1. Teil sowie von $\alpha = .91$ für den 2. Teil berichtet.

Validität

Die *faktorielle Validität* zeigt sich in hohen Ladungen aller vier Subtests auf dem Faktor „General Fluid Ability“. Die Faktorenstruktur der CFT 20-R-Subtests mit Hauptkomponenten-Analyse (deutsche Normierungsdaten) erbrachte mit drei extrahierten Faktoren eine Varianzerklärung von 69%: Faktor 1 „Gesetzmäßigkeiten und Regelmäßigkeiten erkennen“ („Reasoning-Faktor“) klärt rund 50% des Varianzanteils auf.

Der CFT 20-R korreliert mit dem konstruktnahen Verfahren PSB-R 4-6 zwischen $r = .60$ und $r = .63$. Weiß & Weiß (2006) beschreiben, dass die Korrelationen zwischen dem CFT 20-R und den Schulnoten im Fach Mathematik insgesamt über alle Klassen, Schularten und Schulstufen hinweg sehr zufriedenstellend ausfallen ($r = .49$). Im Fach Deutsch liegen sie mit $r = .35$ wesentlich niedriger. Für den CFT 20-R liegen noch keine prognostischen Validitätsuntersuchungen vor.

Autorinnenkommentar

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 Schulstufen 7 bis 9 belegen dies 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 eine Bestätigung für die Strukturannahme eines g-Faktors in allen Fähigkeitsgruppen finden lässt.¹⁷ Testanwender/innen müssen beim Einsatz des CFT 20-R für die Intelligenzdiagnostik mit relativ großen Messfehlern rechnen. Auch die Ergänzung schwierigerer Items im CFT 20-R im Vergleich zum CFT 20 hat das Problem nicht lösen können.

Literatur

- Cattell, R. B. (1960). *Culture Fair Intelligence Test, Scale 2*. Champaign, IL: IPAT.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6*. Göttingen: Hogrefe.
- Jacobs, C. & Petermann, F. (2007). *Grundintelligenztest (CFT-20R) von Rudolf Weiss (2006) [Testinformation]*. *Diagnostica*, 53, 109-113.
- Kuhn, J. T., Holling, H. & Freund, A. (2008). *Begabungsdiagnostik mit dem Grundintelligenztest (CFT 20-R). Psychometrische Eigenschaften und Messäquivalenz*. *Diagnostica*, 54, 184-192.
- MacArthur, R. S. & Elley, W. B. (1963). *The reduction of socio-economic bias in intelligence testing*. *British Journal of Educational Psychology*, 33, 107-119.
- Weiß, H. R. & Weiß, B. (2006). *CFT 20-R. Grundintelligenztest Skala 2 – Revision*. Göttingen: Hogrefe.
- Weiß, R. (1978). *Grundintelligenztest Skala 2 (CFT 20)*. Braunschweig: Westermann.

¹⁷ Diese Strukturäquivalenz ist lediglich noch für den BIS-HB nachgewiesen.

3.11 CMM 1-3; Columbia Mental Maturity Scale 1-3

Schuck, K. D., Eggert, D. & Raatz, U. (1975). Columbia Mental Maturity Scale CMM 1-3. Sprachfreier Gruppenintelligenztest für die Grundschule. Weinheim: Beltz.

Beschreibung

Mit der CMM 1-3 liegt ein ökonomischer und dabei zuverlässiger Gruppenintelligenztest für Grundschüler/innen vor, der als Niveautest eine Abschätzung der allgemeinen Intelligenz erlaubt. Die CMM 1-3 ist ein sprachfreies Verfahren.

Die CMM 1-3 als Gruppenintelligenztest für Grundschüler/innen erster bis dritter Klassen entstand als letzte Stufe deutscher Weiterentwicklungen der CMM aus dem Wunsch von Grund- und Sonderschullehrerinnen und -lehrern heraus, die kognitiven Leistungen von schulleistungsschwachen Grundschülerinnen und -schülern mit denen ihrer Klassenkameradinnen und -kameraden bzw. mit altersgleichen Kindern der Grundschule und nicht nur mit Sonderschülerinnen und -schülern vergleichen zu können. Die Kinder wählen nach dem Multiple-choice-Prinzip bei jeder Aufgabe das Bild aus, das nicht zu den übrigen passt.

Altersgruppe: 6;0 bis 10;0 Jahre

Hinweise zur Durchführung

Die Durchführung nimmt in etwa 1 Unterrichtsstunde in Anspruch.

Es handelt sich um einen **Gruppentest**. Eine besonders intensive Einarbeitung in die Durchführung ist nicht notwendig. Dringend erwünscht ist allerdings Erfahrung mit Kindern.

Hinweise zu Auswertung und Interpretation

Die Normierung der CMM 1-3 wurde in den Jahren 1971 und 1972 durchgeführt. Daran waren insgesamt 8.437 Kinder aus Deutschland beteiligt. Die Eichstichprobe kann als repräsentativ gelten. Für die CMM 1-3 liegen Prozentrang-Normen für den Gesamttestwert für Grundschulkinder vor und zwar als Klassennormen und Altersnormen.

Testgüte

Reliabilität

Bei dieser kürzesten aller CMM-Versionen ergeben sich die niedrigsten Reliabilitätskoeffizienten: für Grundschüler/innen aus 2. Klassen lassen sich Split-Half-Koeffizienten von $r = .88$ bzw. $r = .87$ errechnen, für Kinder an Sonderschulen ergibt sich ein r von $.91$. Die Retest-Relia-

bilitäten betragen nach 14 Tagen zwischen $r = .84$ und $r = .91$ (Sonderschulen).

Validität

Bei Haupt- und Sonderschüler/innen ergeben sich zwischen CMM 1-3 und FAT 4-6 bzw. 7-9 Korrelationen zwischen $r = .55$ und $r = .61$. Bei Kindern aus der 2. Schulstufe lassen sich Zusammenhänge zwischen der CMM 1-3 und dem BT 1-2 zwischen $r = .60$ und $r = .66$ errechnen, zwischen CMM 1-3 und CPM reichen sie von $r = .54$ bis $r = .59$.

Autorinnenkommentar

Die CMM 1-3 eignet sich nur bedingt zur Testung von intellektueller (Hoch-)Begabung. Deckeneffekte treten schon bei Kindern mit einem $IQ > 120$ auf. Die Aufgaben sind zudem wenig abwechslungsreich, weshalb Ermüdungseffekte nicht auszuschließen sind.

Literatur

Schuck, K. D., Eggert, D. & Raatz, U. (1975). *Columbia Mental Maturity Scale CMM 1-3. Sprachfreier Gruppenintelligenztest für die Grundschule*. Weinheim: Beltz.

3.12 CMM 1-4; Columbia Mental Maturity Scale 1-4

Schuck, K. D., Eggert, D. & Raatz, U. (1976). Columbia Mental Maturity Scale. CMM 1-4. Sprachfreier Einzel- und Gruppenintelligenztest für die Grundschule und Allgemeine Sonderschule. Wien: Ketterl.

Beschreibung

Die Columbia Mental Maturity Scale ist ein Testverfahren, mit dem die allgemeine Intelligenzleistung von Kindern beschrieben werden kann. Sie erlaubt bei unterschiedlichen Zielgruppen mit gleichem Material und in variierbaren Untersuchungssituationen eine schnelle, dabei reliable und valide Abschätzung der allgemeinen Intelligenz.

Der ursprüngliche Anwendungsbereich der CMM war die Diagnose der geistigen Behinderung und – in der Erweiterung – die Erfas-



sung kognitiver Leistungen im Vorschulalter. Burgemeister, Blum & Lorge (1954) entwickelten die CMM zunächst als Einzeltest für normalentwickelte und hirngeschädigte 3- bis 12-jährige US-amerikanische Kinder. Die einzelnen Aufgaben sollten Denkfähigkeit, logisch-schlussfolgerndes Denken und Abstrahierungsfähigkeit erfassen und nach dem Gesamtwert eine Klassifikation der allgemeinen Intelligenz erlauben (Schuck et al., 1976). Die Aufgaben der CMM 1-4 bestehen aus jeweils nach einem bestimmten Prinzip zusammengestellten Einzelbildern (Schuck et al., 1976). Eines der fünf auf den Bildern dargestellten Objekte genügt diesem Prinzip nicht und muss von den Kindern als nicht zu den anderen Bildern passend gekennzeichnet werden. Zur Aufgabenlösung sind dabei folgende kognitive Operationen erforderlich: Analogienbildung, Bildung und Anwendung von Oberbegriffen, Erkennen von Identitäten.

Altersgruppe: 3;0 bis 10;0 Jahre

Hinweise zur Durchführung

Die Durchführungszeit variiert zwischen 10 und 30 Minuten.

Es handelt sich um einen **Einzeltest**, für die Gruppentestung ist die CMM 1-3 gedacht. Eine besonders intensive Einarbeitung in die Durchführung ist nicht notwendig. Dringend erwünscht ist allerdings Erfahrung im Umgang mit Kindern dieser Altersstufe bzw. mit lernbehinderten Schülerinnen und Schülern.

Hinweise zu Auswertung und Interpretation

Die Normierung der CMM 1-4 wurde in den Jahren 1968 bis 1972 an einer repräsentativen Stichprobe vorgenommen. Es liegen Normtabellen (T-Werte, Prozentränge) des Gesamtwertes vor für

1. geistig behinderte Kinder im Alter von 7–12 Jahren,
2. lernbehinderte Kinder im Alter von 9–12 Jahren und
3. normalentwickelte Kinder im Alter von 3–10 Jahren.

Testgüte

Reliabilität

Für Kinder der 2. Schulstufe (N = 771) wurde eine Split-Half-Reliabilität von $r = .88$ ermittelt. Die Retest-Reliabilität nach einem Jahr beträgt $r = .60$.

Validität

Es werden Korrelationen zwischen Ergebnissen der CMM 1-4 und Ergebnissen aus dem HAWIK (Wechsler, 1956) zwischen $r = .29$ (Verbale Intelligenz) und $r = .68$ (Handlungsintelligenz) berichtet. Für die Gesamtintelligenz beträgt der Zusammenhang $r = .52$. Das PSB (Horn, 1969) korreliert

mit der CMM 1-4 mit $r = .59$, der BT 2-3 (Ingenkamp, 1966) mit $r = .57$. Die Korrelationen mit den Zeugnisnoten am Ende der 4. Schulstufe der Grundschule betragen im Sachunterricht $r = .40$, im Fach Mathematik $r = .67$, im Fach Deutsch $r = .21$ und mit dem Lesen $r = .62$. In verschiedenen Faktorenanalysen der Ergebnisse von Kindern im Vorschulalter konnte gezeigt werden (Eggert & Schuck, 1975; Schuck, 1976), dass der Summenwert der CMM 1-4-Elemente sprachlichen und sprachfreien Problemlösungsverhaltens enthält und demgemäß nicht als „nonverbaler Intelligenztest“ zu klassifizieren ist. So korreliert die CMM 1-4 mit dem Verbalfaktor des HAWIVA bei Vierjährigen mit $r = .63$ und mit dem Handlungsfaktor (Organisation der Wahrnehmung) des HAWIVA mit $r = .35$ (Eggert & Schuck, 1975).

Autorinnenkommentar

Die CMM 1-4 eignet sich wie auch die CMM 1-3 nur bedingt zur Testung von intellektueller (Hoch-)Begabung, Eine Minderbegabung oder eine Begabung im unteren Durchschnittsbereich kann mit großer Wahrscheinlichkeit festgestellt werden. Deckeneffekte treten schon bei Kindern mit einem $IQ > 120$ auf. Die Aufgaben sind zudem wenig abwechslungsreich, weshalb Ermüdungseffekte nicht auszuschließen sind.

Literatur

- Burgemeister, B. B., Blum, C. H. & Lorge, I. (1954). *Columbia Mental Maturity Scale*. New York: Harcourt, Brace and World.
- Eggert, D. & Schuck, K. D. (1975). Ein empirischer Beitrag zum Zusammenhang von Intelligenz, Motorik und Sozialstatus im Vorschulalter. *Sonderpädagogik*, 1(5), 57-62 und 121-135.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.
- Ingenkamp, K. (1966). *Bildertest (BT 2-3)*. Weinheim: Beltz.
- Schuck, K. D. (1976). *Untersuchungen zur Erfassung von Umweltkorrelaten des verbalen und nicht-verbalen kognitiven Leistungsniveaus vier- bis sechsjähriger Kinder*. Frankfurt: Universität, Dissertation.
- Schuck, K. D., Eggert, D. & Raatz, U. (1976). *Columbia Mental Maturity Scale. CMM 1-4. Sprachfreier Einzel- und Gruppenintelligenztest für die Grundschule und Allgemeine Sonderschule*. Wien: Ketterl.
- Wechsler, D. (1956). *Intelligenztests für Kinder (HAWIK)*. Bern: Huber.

3.13 CPM; Coloured Progressive Matrices

Raven, J. C., Raven, J. & Court, J. H. (2002). Raven's Progressive Matrices und Vocabulary Scales. Coloured Progressive Matrices mit der Parallellform des Tests und der Puzzle-Form. Deutsche Bearbeitung und Normierung Stephan Bulheller und Hartmut Häcker [Manual]. Frankfurt/M.: Swets.

Beschreibung

Die CPM messen das allgemeine Intelligenzpotenzial von Kindern weitgehend sprachfrei. Es handelt sich um ein weit verbreitetes und mittlerweile seit ca. 50 Jahren interkulturell bewährtes Untersuchungsverfahren.

Die CPM wurden für eine möglichst genaue Leistungsbeurteilung der kognitiven Entwicklung bis zum Stadium voll entwickelter intellektueller Fähigkeiten konstruiert. Die Erkenntnis, dass die Itemschwierigkeiten der Standard Progressive Matrices nach oben ausgebaut werden mussten, um genaue Differenzierungen bei Jugendlichen und jüngeren Erwachsenen vornehmen zu können, führte zur Entwicklung einer erweiterten Version der Standard Progressive Matrices (SPM): den Coloured Progressive Matrices (CPM; Raven, Raven & Court, 2002). Für die Konzeption des CPM wurde eine Itemanalyse der Antworten der SPM, die 1947 vorlagen, durchgeführt und für die Zusammenstellung der CPM verwendet. 1956 wurden die CPM neu überarbeitet. 1998 wurde eine Parallellform der CPM, die CPM-P, entwickelt, um Lerneffekte durch mehrmalige Testung mit den CPM ausschalten zu können.

Die CPM bestehen aus drei Sets (A, Ab, B) mit je 12 Aufgaben. Die Vorgangsweise bei den 36 Aufgaben ist immer gleich und besteht darin, zu einer unvollständigen farbigen geometrischen Figur die richtige Ergänzung aus sechs möglichen Vorgaben auszuwählen. Nach den Angaben der Testautoren (Raven, Raven & Court, 2002) sind die Items der CPM so angeordnet, dass die kognitive Entwicklung einer Testperson bis zu dem Stadium beurteilt werden kann, in dem die Person in der Lage ist, in Analogien zu denken und sich dieses Vorgehen als konsistente Methode für schlussfolgerndes Denken anzueignen.

Aus den Ergebnissen der experimentellen Arbeiten, die zur Entwicklung der CPM führten, lassen sich fünf qualitative Entwicklungsstufen unterscheiden:

- (1) Kinder sind zunächst in der Lage, gleiche Formen von verschiedenen Formen zu unterscheiden, später auch ähnliche von unähnlichen.
- (2) Nach einiger Zeit können Kinder die Anordnung einer Form in Bezug auf sich selbst und in Bezug auf andere Objekte im Wahrnehmungsbereich erkennen.
- (3) Kinder können analoge Veränderungen der wahrgenommenen Merkmale vergleichen und sich dies als logische Denkmethode aneignen.
- (4) Kinder können aus der Analyse des Ganzen dessen konstituierende Elemente oder „Merkmale“ erkennen und unterscheiden, was gegeben ist und was sie selbst beitragen.

- (5) Schließlich können Kinder wahrnehmen, ob zwei oder mehr diskrete Figuren ein bestehendes Ganzes bilden oder als individuelle Einheiten organisiert sind (Raven, Raven & Court, 2002).

Die CPM eignen sich gut zur Anwendung bei Personen, die die deutsche Sprache weder ausreichend verstehen noch sprechen.

Altersgruppe: 3;9 bis 11;8 Jahre

Hinweise zur Durchführung

Die Durchführung dauert etwa 20 bis 30 Minuten (eine Zeitbegrenzung ist nicht vorgesehen).

Die CPM können als **Einzel- oder Gruppentest** (nicht mehr als acht bis neun Personen) durchgeführt werden.

Der Test kann in der Version als Testheft vorgelegt werden oder in der „Board-Form“, das ist eine Version des Tests, die das Einfügen der Antwortmöglichkeiten nach Art eines Puzzles erlaubt. Bei der Board-Form befinden sich die zur Auswahl stehenden Antwortmöglichkeiten auf beweglichen Teilen, die in das Muster eingefügt werden können. Die Board-Form ermöglicht unabhängig von ethnischer Zugehörigkeit und Sprache eine verständliche Demonstration der Testaufgaben und ist darüber hinaus einer der wenigen Tests, die mit zufriedenstellenden Ergebnissen bei Personen mit Teillähmungen, Taubheit oder Sprachstörungen angewendet werden können. Als sprachunabhängig können diese Tests aber nur insofern gelten, als die Aufgaben figuraler Art sind. Die Aufgabenlösung hängt aber auch von der Verbalisierung der Problemstellung ab.

Eine testspezifische, über eine übliche Testleiterschulung hinausgehende Einarbeitung ist der Handanweisung zu Folge nicht notwendig. Aufgrund der Beeinflussbarkeit des Ergebnisses durch spezifische Testungsbedingungen sollte jedoch auch dieses Verfahren in erster Linie von Psychologinnen und Psychologen angewandt werden.

Hinweise zu Auswertung und Interpretation

In der Neunormierung 1998/1999 in Deutschland und Frankreich (N = 1.218) sind Normtabellen und Prozentränge für die Altersgruppe von vier bis elf Jahren ersichtlich. Für die Beurteilung der Prozentränge kann auf fünf Kategorien der intellektuellen Leistungsfähigkeit (von Leistungsstufe I „Überragende intellektuelle Leistungsfähigkeit



Prozentrang >95“ bis Leistungsstufe V „Geistig behindert Prozentrang <5“) zurückgegriffen werden.

Die CPM sind Mills, Ablard & Brody (1993) zufolge als hervorragend geeignetes Testinstrument für die allgemeine frühzeitige Erkennung von Hochbegabung anerkannt. Ebenso beschreiben Saccuzzo, Johnson & Guertin (1994) die CMP als geeignetes Instrument für die Erkennung von hochintelligenten Erwachsenen. Dies wird von den Autorinnen und Autoren auf die umfangreiche Datensammlung zurückgeführt, in der zahlreiche Informationen über Variablen wie Alter, ethnische Zugehörigkeit, sozioökonomischer Status sowie Normdaten aus vielen verschiedenen Ländern und Regionen der Erde zu finden sind (Raven, Raven & Court, 2002). Allerdings wird im Psyndex-Testreview darauf hingewiesen, dass die CPM lediglich bis zum Alter von 8;2 Jahren eine hohe Differenzierungsfähigkeit über den gesamten Leistungsbereich hinweg zeigen. Bei älteren Kindern mit (mutmaßlich) hohen Intelligenzausprägungen sind Deckeneffekte zu erwarten. Bei jüngeren Kindern (< 6 Jahre) erwiesen sich die CPM hingegen als zu schwierig und sind daher ebenfalls nur bedingt brauchbar (z.B. Knaack, 1978).

Testgüte

Reliabilität

Im Manual (Raven, Raven & Court, 2002) sind keine aktuellen Reliabilitätsschätzungen für die deutsche Normierungsstichprobe ersichtlich. Es werden interne Konsistenzen über verschiedene Stichproben und Altersgruppen zwischen $\alpha = .67$ und $\alpha = .93$ berichtet (Schmidtke, Schaller & Becker, 1980). Für die deutsche Eichstichprobe aus den Jahren 1971–1976 (4;9–11;0 Jahre) ergab sich eine Split-Half-Reliabilität über die verschiedenen Altersgruppen zwischen $r = .69$ und $r = .90$. Die Retest-Reliabilität (ein bis zwei Wochen) ($N = 1.029$) liegt zwischen $r = .58$ und $r = .82$. Verschiedene Untersuchungen belegen Split-Half-Reliabilitäten zwischen $r = .85$ und $r = .90$ für unterschiedliche Länder und Altersstufen der Kinder. Die höheren Werte wurden bei älteren, etwas niedrigere bei jüngeren Kindern gefunden. Die Untersuchungen zur Testwiederholungsreliabilität von Wenke & Müller (1966) belegen Werte zwischen $r = .81$ und $r = .87$ bei Stichproben aus 5–8 Jahre alten Kindern.

Validität

Nach den Untersuchungen von Das & Jarman (1991) erfassen die CPM den Faktor „simultanes Verarbeiten“ mit Ladungen zwischen $r = .75$ und $r = .85$. In einer Studie ($N = 702$) untersuchten Wilkes & Weigel (1998) 10- bis 11-jährige Kinder. Sie ermittelten eine Korrelation von $r = .67$ zwischen CPM und HAWIK-R. Zugleich ergaben sich jedoch hochsignifikante Mittelwertsunterschiede zwischen den ermittelten Intelligenzleistungen, die dafür sprechen, dass die Verwendung der CPM in diesem Altersbereich zu einer erheblichen Überschätzung der IQ-Werte führen kann. Der mittlere IQ bei den CPM lag um nahezu 10 Punkte über dem im HAWIK-R und ca. 16 Punkte über dem, der mittels SPM geschätzt worden war. Zur Bestimmung der prognostischen Validität setzte Knaack (1978) die CPM an einer Stichprobe von 5- bis 6-jährigen Vorschulkindern ein und ermittelte Zusammenhänge der CPM-Werte mit 3 Jahre später erhobenen Leistungstests und Lehrer/innenbeurteilungen: Allgemeiner Schulleistungstest AST 2: Lese-

verständnis $r = .17$, Rechtschreibung $r = .10$, Sachwissen $r = .30$, Lehrerurteil (Zensur) Deutsch $r = .29$, Mathematik $r = .33$, Sachkunde $r = .41$.

Autorinnenkommentar

Die CPM sind populäre Verfahren zur Erfassung der fluiden Intelligenz. Bezogen auf die Intelligenzdiagnostik ist zu beachten, dass die CPM wie auch die APM nur eine Intelligenzfacette – und dies zudem mit nur einem Aufgabentyp – erfassen. Vielfach werden Geschlechtereffekte zugunsten männlicher Testpersonen dokumentiert. Viele Studien belegen außerdem, dass die CPM nicht für alle Gruppen, in denen sie laut Handanweisung eingesetzt werden sollen, eine valide Messung ermöglichen. Dies gilt für die Testung von Vorschulkindern und für geistig behinderte Menschen, für die das Verfahren im Regelfall zu schwer ist, ebenso wie für ältere Kinder sowie für gesunde Erwachsene im höheren Lebensalter, bei denen die CPM-Werte einen Deckeneffekt zeigen.

Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den CPM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der CPM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe auch Preckel, 2010).

Literatur

- Das, J. P. & Jarman, R. (1991). *Cognitive integration. Alternative model for intelligence*. In H. A. H. Rowe (Hrsg.), *Intelligence: Reconceptualization and Measurement*. Hillsdale, NJ: Lawrence Erlbaum & Associates.
- Knaack, R. (1978). *Zur Brauchbarkeit der Coloured Progressive Matrices (CPM) von Raven bei Kindern im Vorschulalter*. *Psychologie in Erziehung und Unterricht*, 25, 159-167.
- Mills, C., Ablard, K. & Brody, L. (1993). *The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students*. *Roeper Review*, 15(3), 183-186.
- Preckel, F. (2010). *Intelligenztests in der Hochbegabungsdiagnostik*. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 19-44). Göttingen: Hogrefe.
- Raven, J. C., Raven, J. & Court, J. H. (2002). *Raven's Progressive Matrices und Vocabulary Scales. Coloured Progressive Matrices mit der Parallelförm des Tests und der Puzzle-Förm*. Deutsche Bearbeitung und Normierung Stephan Bulheller und Hartmut Häcker [Manual]. Frankfurt/M.: Swets.
- Saccuzzo, D. P., Johnson, N. E. & Guertin, T. L. (1994). *The Raven's Progressive Matrices Test as an Equitable Measure of IQ in an Ethnically Diverse Population*. Manuscript submitted for publication.
- Schmidtke, A., Schaller, S. & Becker, P. (1980). *CPM. RAVEN-Matrizen-Test. Coloured Progressive Matrices. Manual* (2. verbesserte Auflage). Weinheim: Beltz.
- Wenke, W. & Müller, U. (1966). *Möglichkeiten und Grenzen des Einsatzes einzelner diagnostischer*



Kurzverfahren bei der Schülerauslese. Zeitschrift für Psychologie, 172, 82-116.
Wilkes, J. & Weigel, A. (1998). Vergleich von HA-WIK-R und den Progressiven Matrizentests (Raven) in einer klinischen Inanspruchnahmepopulation. Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 26, 261-265.

3.14 FAT 6-9; Frankfurter Analogietest 6-9

Barth, N. (1989). Frankfurter Analogietest 6-9. Test des logisch-schlussfolgernden Denkens für die 6. bis 9. Schulstufe [Testmappe mit Handanweisung, je einem Musteraufgabenheft Form A und B und einem Musterantwortbogen Form A und B]. Weinheim: Beltz.

Beschreibung

Der FAT ist ein Schultest, der das logisch-schlussfolgernde Denken erfassen soll. Im oberen Leistungsbereich ist er häufig zu leicht, weshalb hier zur exakten Leistungsmessung andere Testverfahren verwendet werden sollten.

Der FAT 6-9 dient zur Ermittlung des logisch-schlussfolgernden Denkens in der Schulpraxis. Zielgruppen sind Schüler/innen der Schulstufen 6–9 der Schultypen Haupt-, Realschule und Gymnasium. Der FAT 6-9 wurde auf Basis des FAT 4-6 (Belser et al., 1973) entwickelt. Der FAT 6-9 weist nunmehr eine verkürzte Bearbeitungsdauer (30 Minuten anstatt 60 Minuten) sowie eine Neunormierung auf. Der FAT 6-9 erfasst die allgemeine Intelligenz (logisch-schlussfolgerndes Denken) und differenziert am besten bei leistungsschwächeren Schülerinnen und Schülern. Das im FAT 6-9 geforderte analoge Denken schließt viele unterschiedliche induktive und deduktive Denkleistungen ein. Der FAT 6-9 enthält jeweils 50 Aufgaben mit vier unterschiedlichen Aufgabentypen: Wortanalogien, gegenständliche und abstrakte Bildanalogien sowie gemischte Analogien. Aus fünf Wahlmöglichkeiten ist die passende Lösung zu finden.

Altersgruppe: 6. bis 9. Schulstufe

Hinweise zur Durchführung

Die Bearbeitungszeit beträgt ca. 20–30 Minuten.

Der Test kann als **Einzel- oder Gruppentestung** durchgeführt werden. Der FAT 6-9 wird in der Regel von einer Lehrperson durchgeführt, die sich zuvor mit dem Verfahren vertraut machen sollte.

Hinweise zu Auswertung und Interpretation

Die Eichstichprobe aus dem Jahr 1982 umfasste $N = 13.245$ Schüler/innen und wird als repräsentativ bezeichnet. Für die Schulstufen 6 bis 9 und über alle Schultypen hinweg bestehen Normwerte in Form von Prozenträngen und T-Werten.

Testgüte

Reliabilität

Die interne Konsistenz des FAT 6-9 liegt zwischen $\alpha = .67$ und $\alpha = .94$, die Retest-Reliabilitäten (Testwiederholung nach ca. 2,5 Monaten) zwischen $r = .44$ und $r = .92$.

Validität

Faktorenanalytische Untersuchungen der Konstruktvalidität ergaben vier interpretierbare Faktoren, die zusammen 55,7% der Varianz erklären: „verbal-logische Fähigkeiten“, „Raumvorstellung“, „allgemeine Schulleistung mit deutsch-musischem Schwerpunkt“ und „naturwissenschaftliches Verständnis vs. Gedächtnis oder Verstehen vs. Wissen“. Diese Faktoren werden als „Generalfaktor“ allgemeiner Intelligenz im Sinne von Spearman aufgefasst. Hinsichtlich der Konstruktvalidität wird eine Korrelation von $r = .66$ (Verbaltest 4 KFT) bzw. $r = .60$ (Analogienfinden, IST) berichtet. Der Zusammenhang zwischen den Ergebnissen des FAT 6-9 und Schulnoten ist gering, der höchste Wert zeigte sich in der Deutschnote ($r = .29$).

Autorinnenkommentar

Der FAT ist für die Testung von Hochbegabung ungeeignet, da die Aufgaben im oberen Leistungsbereich zu leicht sind und Deckeneffekte schon im oberen Durchschnittsbereich auftreten können. Die Validität ist zudem eher gering, weshalb er für die Testung von Begabung nur bedingt geeignet erscheint. Eine Kombination mit anderen Verfahren wird – um eine Diagnose stellen zu können – unerlässlich sein.

Literatur

Barth, N. (1989). *Frankfurter Analogietest 6-9. Test des logisch-schlussfolgernden Denkens für 6. bis 9. Klassen [Testmappe mit Handanweisung, je einem Musteraufgabenheft Form A und B und einem Musterantwortbogen Form A und B]*. Weinheim: Beltz.

Barth, N. (1996). *Zensuren und logisches Denken – eine empirische Untersuchung zur Übereinstimmungsgültigkeit zwischen 17 Schulzensuren und den Leistungen im Frankfurter Analogietest*. *Zeitschrift für internationale erziehungs- und sozialwissenschaftliche Forschung* 13(1), 139-155.

Belser, H., Anger, H., Bargmann, R. (1965). *Frankfurter Analogietest. FAT 7-8. Begabungstest für 7. und 8. Klassen (2. Auflage)* [Testmappe mit „Beiheft mit Anleitung und Normentabellen“, Testheften Form A und Form B sowie Klassenliste]. Weinheim: Beltz.

Belser, H., Anger, H., Bergmann, R. & Raatz, U., (1973). *Frankfurter Analogietest FAT 4-6. Begabungstest für 4. bis 6. Klassen (2. Auflage, Bearbeitung und Normierung 1972. 1. Auflage 1965. Testmappe mit „Beiheft mit Anleitung und Normentabellen“, Testheften Form A und Form B sowie Klassenliste)*. Weinheim: Beltz.

3.15 FRT bzw. FRT-J; Figure Reasoning Test

Daniels, J. C. (2004). Figure Reasoning Test. Frankfurt/M.: Swets.

Beschreibung

Der FRT bzw. der FRT-J dienen zur Abklärung der allgemeinen Intelligenz ab zehn Jahren. In der gegenwärtigen Form sollte das Verfahren nur als Forschungsinstrument und nicht als Individualdiagnostikum angewendet werden.

Das Verfahren ist ähnlich konzipiert wie die SPM oder die APM (siehe Testbeschreibungen). In den Testaufgaben geht es um die Erkennung von logischen Zusammenhängen bei Figuren und der Auswahl eines dazupassenden Elements nach dem Multiple-choice-Prinzip. Die 45 Items sind progressiv nach Schwierigkeitsgrad angeordnet. Aus einer Vorgabe von sechs Auswahlfiguren soll die passende Figur zu einem unvollständigen Muster aus acht Figuren gewählt werden. Wegen der weitgehenden Sprachfreiheit des Verfahrens können auch Personen mit Hör- und Sprechbehinderungen problemlos getestet werden.

Altersgruppe: FRT-J – Kinder und Jugendliche von 10 bis 15 Jahren
FRT – Jugendliche ab 14 Jahren und Erwachsene

Hinweise zur Durchführung

Der Test dauert ca. 45 Minuten.

Er kann als **Einzel- oder Gruppentest** von jeder Fachkraft, die mit den Prinzipien der Anwendung und Auswertung von Tests vertraut ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Eine Neunormierung erfolgte im Zuge der 2004 herausgegebenen Neubearbeitung des Tests. Die ursprüngliche Normierung 1962 erfolgte an ca. 4.800 Schülerinnen und Schülern eines mittelenglischen Verwaltungsbezirks. Für die Normierung stehen IQ-Werte für 10- bis 16-Jährige (in Jahresschritten) und Erwachsene (ohne Differenzierung) zur Verfügung.

Testgüte

Reliabilität

Es wurde ein Split-Half-Reliabilitätskoeffizient von $r = .96$ ermittelt; die Retest-Reliabilität nach zwei Wochen betrug $r = .97$, nach einem Jahr $r = .89$.

Validität

Ein g-Faktor-Ladungsanteil von über 80% wurde gefunden.

Die Korrelationen zwischen den FRT- und den SPM-Ergebnissen beträgt $r = .93$; zwischen FRT und Stanford-Binet beträgt die Korrelation $r = .86$; zwischen dem FRT und dem WISC beträgt die Korrelation $r = .87$; mit der Abschlussnote errechnet sich ein Korrelationskoeffizient von $r = .33$.

Autorinnenkommentar

Der FRT enthält nur einen Aufgabentyp, was die Diagnostik einer allgemeinen Intelligenz bzw. einer intellektuellen Begabung erschwert. Deckeneffekte im oberen Begabungsbereich können vielfach beobachtet werden. Oft werden auch Geschlechtereffekte zugunsten männlicher Testpersonen beobachtet. Tests mit unterschiedlichen Aufgabenfacetten sollten dem FRT vorgezogen werden. Will man lediglich das Reasoning überprüfen, so eignen sich andere figurale Matrizen tests besser als der FRT.

Literatur

Daniels, J. C. (1962). Figure Reasoning Test. A nonverbal intelligence test. London: Crosby Lockwood & Son Ltd.

Daniels, J. C. (2004). Figure Reasoning Test. Frankfurt/M.: Swets.

3.16 HAWIE-R; Hamburg-Wechsler-Intelligenztest für Erwachsene

Tewes, U. (1991). HAWIE-R. Hamburg-Wechsler-Intelligenztest für Erwachsene. Revision 1991. Bern: Huber.

2006 erschien eine deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler als „Wechsler Intelligenztest für Erwachsene“ (WIE; Aster, Neubauer & Horn, 2006 – siehe Testbeschreibung WIE).

Beschreibung

Der HAWIE-R dient zur Erfassung der Intelligenz bei Jugendlichen und Erwachsenen. Er ist stärker als sein Vorläufer auf die Untersuchung älterer Probandinnen und Probanden ausgerichtet.

Beim HAWIE-R handelt es sich um die Revision (1991) des Hamburg-Wechsler-Intelligenztests für Erwachsene, welche auf der Wechsler Adult Intelligence Scale-R (WAIS-R; 1981) von Wechsler beruht. Wechsler verstand Intelligenz als „die zusammengesetzte oder globale Fähigkeit des Individuums, zielgerichtet zu handeln, rational zu denken und sich wirkungsvoll mit seiner Umwelt auseinanderzusetzen. Sie ist zusammengesetzt oder global, weil sie aus Elementen oder Fähigkeiten besteht, die, obwohl nicht völlig unabhängig, qualitativ unterscheidbar sind“ (Matarazzo, 1982, S. 121).

Der HAWIE-R besteht wie der HAWIE aus 11 Untertests (6 Verbaltests und 5 Handlungstests). Die sprachliche Intelligenz wird dabei über die Untertests Allgemeines Wissen, Zahlen nachsprechen, Wortschatz-Test, Rechnerisches Denken, Allgemeines Verständnis und Gemeinsamkeiten finden geprüft, die praktische Intelligenz über die Untertests Bilder ergänzen, Bilder ordnen, Mosaik-Test, Figuren legen und den Zahlen-Symbol-Test. Neben der Erhebung des Gesamt-IQs erlaubt der HAWIE-R eine Profilanalyse über alle Untertests. Die Tests haben durch Zeitbegrenzungen eine Speed-Test-Komponente.

Altersgruppe: 16 bis 74 Jahre

Hinweise zur Durchführung

Die Bearbeitungszeit beträgt 60–90 Minuten.

Der Test kann als **Einzeltest** durchgeführt werden. Testleiter/innen sollten mit den Grundsätzen der Anwendung von psychologischen Tests vertraut sein.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte an N = 2.000 Personen im Alter von 16 bis 74 Jahren. Rohpunktwerte werden zu standardisierten Wertpunkten transformiert. Es existieren altersgestaffelte sowie

schulbezogene Wertpunkt-Normen, die in IQ-Werte transformiert und in Verbal-IQ, Handlungs-IQ und Gesamt-IQ ausgegeben werden.

Testgüte

Reliabilität

Für die HAWIE-R-Untertests variieren die inneren Konsistenzen zwischen $\alpha = .71$ und $\alpha = .96$, für den Handlungsteil liegt die Konsistenz bei $\alpha = .90$. für den Verbal- und den Gesamtteil bei $\alpha = .96$.

Validität

Die Wechsler-Intelligenztests stellen einen der grundlegenden Ansätze zur Messung der Intelligenz dar und verfügen so über inhaltlich-logische Validität. Per Faktorenanalyse konnte man einen relativ starken ersten Faktor finden. Bei Rotation findet man die Faktoren für Handlungs- und Verbalteil.

Autorinnenkommentar

Dem HAWIE-R sind andere Verfahren der Intelligenzmessung vorzuziehen. Zwar sind die Testaufgaben gut normiert und abwechslungsreich, dennoch zeigt die Erfahrung, dass Testpersonen mit anderen Testmaterialien besser zurechtkommen. Für die Diagnostik im oberen Bereich ist er nur bedingt geeignet.

Literatur

- Aster, M. v., Neubauer, A. & Horn, R. (Hrsg., 2006). *WIE. Wechsler Intelligenztest für Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler.* Frankfurt/M.: Harcourt.
- Matarazzo, J. D. (1982). *Die Messung und Bewertung der Intelligenz Erwachsener nach Wechsler.* Bern: Huber.
- Tewes, U. (1991). *Hamburg-Wechsler-Intelligenztest für Erwachsene. Revision 1991. Handbuch und Testanweisung.* Stuttgart: Huber.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised (WAIS-R).* New York: Psychological Corporation.



3.17 HAWIK-III; Hamburg-Wechsler-Intelligenztest für Kinder III

Tewes, U., Rossmann, P. & Schallberger, U. (Hrsg., 1999). HAWIK-III. Hamburg-Wechsler-Intelligenz-Test für Kinder – Dritte Auflage. Übersetzung und Adaptation der WISC-III Wechsler Intelligence Scale for Children – Third Edition von David Wechsler. Bern: Huber.

Beschreibung

Der HAWIK-III dient zur Untersuchung der kognitiven Entwicklung bei Kindern und Jugendlichen. Er ist zur Einzeltestung mit vielfältigem Anwendungsgebiet im Bereich der klinischen und pädagogischen Diagnostik geeignet. Der HAWIK-III wurde für den Normalbereich geeicht, bei Kindern mit Hochbegabung wird von den Autoren auf den Deckeneffekt im Sinne einer Unterschätzung ihrer tatsächlichen Begabung hingewiesen.

Es handelt sich beim HAWIK-III um die deutschsprachige Version der WISC-III (Wechsler, 1991), die gegenüber dem HAWIK-R (Tewes, 1983) um zwei Untertests erweitert wurde. Das Grundkonzept wurde von Wechsler in den USA als WISC (Wechsler Intelligence Scale for Children, 1949; 1974) entwickelt. Wechsler fasste Intelligenz auf als „die zusammengesetzte oder globale Fähigkeit des Individuums, zweckvoll zu handeln, vernünftig zu denken und sich mit seiner Umgebung wirkungsvoll auseinander zu setzen“ (Wechsler, 1956, S. 13). Wechsler stellte für seine Tests eine Reihe von Skalen zusammen, um eine möglichst breit gefächerte Erfassung der unterschiedlichen Fähigkeiten des Kindes in seiner geistigen Entwicklung zu erreichen. Der Test soll einen Blick auf den geistigen Entwicklungsstand des Kindes aus möglichst vielen verschiedenen Perspektiven ermöglichen. Die Testaufgaben sollen an alltägliche Situationen des Kindes anknüpfen und seinem gewohnten Zugang zu Wissensinhalten gerecht werden.

Erfasst werden über 13 verschiedene Untertests die praktische, die verbale und die allgemeine Intelligenz im Sinne des Globalkonzepts von Wechsler. Dies ermöglicht die Darstellung eines Leistungsprofils. Weiters werden die vier Indizes Sprachliches Verständnis, Wahrnehmungsorganisation, Unablenkbarkeit und Arbeitsgeschwindigkeit berechnet.

Die HAWIK-III Verbal-IQ-Skala besteht aus folgenden Untertests:

- (1) Allgemeines Wissen: Das Kind soll eine Serie mündlich gestellter Wissensfragen beantworten.
- (2) Gemeinsamkeiten finden: Das Kind soll das Gemeinsame bei zwei Begriffen herausfinden, welche sich auf Gegenstände oder Konzepte des Alltags beziehen.
- (3) Rechnerisches Denken: Das Kind soll eine Serie von Rechenaufgaben im Kopf lösen und mündlich beantworten.
- (4) Wortschatz-Test: Das Kind soll eine Serie von mündlich vorgegebenen Wörtern definieren.

- (5) Allgemeines Verständnis: Das Kind soll eine Serie von mündlichen Fragen, die sich auf Alltagsprobleme, soziale Regeln und Konzepte beziehen, beantworten.
- (6) Zahlen nachsprechen: Das Kind soll eine Serie von Ziffernfolgen in derselben oder in der entgegengesetzten Reihenfolge nachsprechen.

Die HAWIK-III Handlungs-IQ-Skala besteht aus folgenden Untertests:

- (7) Bilder ergänzen: Das Kind soll aus einer Serie von Abbildungen aus dem Lebensalltag das fehlende wichtige Detail herausfinden.
- (8) Zahlen-Symbol-Test: Das Kind soll unter jede Zahl das passende Symbol zeichnen, das oberhalb in einem Muster zugeordnet ist.
- (9) Bilder ordnen: Das Kind soll eine Bildergeschichte in die logisch richtige Reihenfolge bringen.
- (10) Mosaik-Test: Das Kind soll eine Serie von vorgegebenen geometrischen Mustern mit Hilfe von zweifarbigen Würfeln nachbauen.
- (11) Figuren legen: Das Kind soll eine sinnvolle Figur aus Puzzle-Teilen zusammensetzen.
- (12) Symbolsuche: Das Kind soll das gemeinsame Symbol aus Gruppen von abstrakten Formen und Symbolen herausfinden.
- (13) Labyrinth-Test: Das Kind soll mit einem Bleistift eine Linie vom Zentrum zum Ausgang eines Labyrinths ziehen.

Der Test beantwortet klinische und pädagogische Fragestellungen, beispielsweise, ob im Einzelfall die praktische oder die verbale Intelligenz stärker ausgeprägt ist, ob unter Zeitdruck schlechter gearbeitet wird als ohne Zeitdruck, ob die mittelfristige Merkfähigkeit im Vergleich zum sonstigen Leistungsvermögen beeinträchtigt ist oder ob Probleme in der figuralen Wahrnehmung vorliegen.

Der Test ist außerdem im Bereich der Lernbehinderung und der Hochbegabungsdiagnostik einsetzbar. Tewes et al. (1999, S. 29) weisen jedoch darauf hin, dass die Testaufgaben für unterdurchschnittlich intelligente Sechsjährige zu schwer, für hochbegabte 16-Jährige zu leicht sein können.

Altersgruppe: 6;0 bis 16;11 Jahre

Hinweise zur Durchführung

Die Durchführungszeit für die Standardtests beträgt 50 bis 70 Minuten.

Es handelt sich um einen **Einzeltest**. Der HAWIK-III sollte nur von entsprechend geschulten Personen eingesetzt werden. Die Einarbeitung in das Manual und die Beherrschung der Anfangs- und Abbruchkriterien sind Voraussetzung für die Anwendung des HAWIK-III. Auch die Interpretation der Ergebnisse setzt entsprechende Sachkenntnis voraus.



Hinweise zu Auswertung und Interpretation

Die Normierung für die Altersgruppen 6;0 bis 16;11 für Deutschland, Österreich und die deutschsprachige Schweiz (N = 1.600) wurde von 1995 bis 1998 durchgeführt. Für die Rohwerte werden in den Tabellen Wertpunkte erhoben und in weiterer Folge IQ-Äquivalente für Verbal- und Handlungsteil sowie für den Gesamttest ermittelt.

Testgüte

Reliabilität

Die Split-Half-Reliabilität liegt für den Verbalteil bei $r = .95$, für den Handlungsteil bei $r = .91$.

Die Split-Half-Reliabilität für den Gesamtwert beträgt $r = .96$.

Validität

Auch für den HAWIK-III bestätigte sich die Zwei-Faktorenlösung der sprachlichen und praktischen Intelligenz des WISC-III. Die Korrelationen zwischen Schulnoten und HAWIK-III-Leistungen betragen für Mathematik $r = .40$ (Verbal-IQ) bzw. $r = .43$ (Gesamt-IQ), für Deutsch $r = .29$ (Verbal-IQ) bzw. $r = .31$ (Gesamt-IQ).

Autorinnenkommentar

Der HAWIK-III ist ein Standardinstrument der allgemeinen Intelligenzdiagnostik. Er erfährt einen weiten Anwendungsbereich, seine Differenzierungsstärke liegt allerdings im mittleren Leistungsbereich. In den Extrembereichen der kognitiven Leistungsfähigkeit (Geistige Behinderung – Hochbegabung) ergeben sich Boden- bzw. Deckeneffekte. Das heißt, die Items sind größtenteils zu schwer oder zu leicht, daher können keine präzisen Aussagen mehr getroffen werden. Der HAWIK-III wird häufig im Rahmen der neuropsychologischen Diagnostik eingesetzt. Es ist zu beachten, dass dieser Test hier im Sinne eines Screenings angewendet werden kann, um Richtungshinweise zu geben. Eine differenzialdiagnostische Abklärung von Funktionsdefiziten kann dieses Verfahren nicht leisten. Die Möglichkeit zur gezielten Verhaltensbeobachtung ist ein enormer Vorteil des HAWIK-III gegenüber Leistungstests, die von der Probandin/vom Probanden selbst ausgefüllt werden. Die hohe Kontrolle der Testleiterin/des Testleiters über die serielle Darbietung von Materialien und die aufgabenspezifischen Zeitmessungen führen zu einem individuell angepassten, fairen Vorgehen und ermöglichen eine integrative Diagnostik.

Die Zusammenfassung der Untertestleistungen zu IQ-Werten führt in manchen Fällen zu widersprüchlichen Ergebnissen. So kann es vorkommen, dass die Wertpunkte der Untertests alle noch im (untersten) Normbereich liegen, der Gesamtwert jedoch als unterdurchschnittlich zu bewer-

ten ist. Ähnlich ergibt sich bei einem jeweils niedrigen, aber noch durchschnittlichen Verbal- und Handlungs-IQ ein Gesamt-IQ, der nicht mehr im Normbereich liegt. In solchen Fällen sollte die Interpretation der Gesamtwerte zurückhaltend erfolgen. Wechsler entwickelte seine Verfahren besonders für den klinischen Einsatz und orientierte sich im Rahmen der Testkonstruktion an praxisnahen, weniger an theoretischen Belangen. Daraus resultiert der häufig angemerkte Mangel einer fehlenden theoretischen Untermauerung des Intelligenzkonzeptes, insbesondere in Bezug auf die Einteilung in einen Verbal- und Handlungs-IQ. Eine Interpretation der Ergebnisse (v.a. auf Untertestebene) sollte daher immer auf bereits spezifizierte klinische Fragestellungen Bezug nehmen. Im überarbeiteten Manual (S. 271–277) 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.

Literatur

- Tewes, U. (1983). *Hamburg-Wechsler-Intelligenztest für Kinder – Revision 1983*. Bern: Huber.
- Tewes, U., Rossmann, P. & Schallberger, U. (Hrsg., 1999). *HAWIK-III. Hamburg-Wechsler-Intelligenz-Test für Kinder – Dritte Auflage. Übersetzung und Adaptation der WISC-III Wechsler Intelligence Scale for Children – Third Edition von David Wechsler*. Bern: Huber.
- Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.
- Wechsler, D. (1956). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children – Revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children – Third Edition (WISC-III)*. San Antonio, TX: Psychological Corporation.

3.18 HAWIK-IV; Hamburg-Wechsler-Intelligenztest für Kinder IV

Petermann, F. & Petermann, U. (2007). *HAWIK-IV. Hamburg-Wechsler-Intelligenztest für Kinder – IV. Übersetzung und Adaptation der WISC-IV von David Wechsler*. Bern: Huber.

Beschreibung

Der HAWIK-IV erfasst die kognitive Leistungsfähigkeit (Intelligenz) bei Kindern und Jugendlichen. Mit diesem Test liegt die Komplettüberarbeitung des HAWIK-III vor, eines in der Praxis (wohl wie kaum ein anderer Test) etablierten Verfahrens zur Intelligenzdiagnostik bei Kindern

und Jugendlichen, welche die neueren theoretischen Entwicklungen berücksichtigt. Mit dem Allgemeinen Fähigkeits-Index (AFI; Daseking, Petermann & Waldmann, 2008) ist überdies die Möglichkeit des Einsatzes in der Hochbegabungsdiagnostik gegeben. Die Nutzerfreundlichkeit konnte im Vergleich zur Vorgängerversion deutlich verbessert werden. Insgesamt kann die Neufassung als äußerst gelungen bezeichnet werden.

Der WISC-IV (Wechsler, 2003) auf dem der HAWIK-IV beruht, stellt eine vollständige Neubearbeitung der Vorgängerversion dar. Der WISC-IV ist insgesamt eines der weltweit meist verbreiteten Verfahren zur Intelligenzdiagnostik bei Kindern. Im WISC-IV bzw. HAWIK-IV wurde die in den früheren Testverfahren erfolgte Unterteilung in Handlungs- und Verbal-IQ aufgegeben, da neuere theoretische Erkenntnisse eine stärkere Betonung von grundlegenden kognitiven Fähigkeiten nahelegen.

Der kognitive Entwicklungsstand wird nun mit den fünf Intelligenzwerten Sprachverständnis, Wahrnehmungsgebundenes Logisches Denken, Arbeitsgedächtnis, Verarbeitungsgeschwindigkeit sowie dem Gesamt-IQ gemessen. Zusätzlich können auf der Untertestebene weitere Analysen sowie eine Profilanalyse über Stärken und Schwächen eines Kindes vorgenommen werden.

Anwendung findet der HAWIK-IV beispielsweise bei der individuellen Beurteilung von kognitiven Stärken und Schwächen sowie bei der Diagnostik von Intelligenzminderung und intellektueller Hochbegabung, wobei die Ergebnisse Informationen bei der Entscheidung für Interventionen (spezielle Förderung, Platzierungsentscheidungen o.ä.) und zur Überprüfung ihrer Wirksamkeit liefern können.

Der HAWIK-IV besteht aus insgesamt 10 Kern-Untertests und 5 optionalen Untertests, die sich folgenden vier Indizes zuordnen lassen (Beschreibung der Untertests siehe HAWIK-III):

I. Sprachverständnis

- (1) Gemeinsamkeiten finden
- (2) Wortschatz-Test
- (3) Allgemeines Verständnis
- (4) Allgemeines Wissen
- (5) Begriffe erkennen

II. Wahrnehmungsgebundenes Logisches Denken

- (1) Mosaik-Test
- (2) Bildkonzepte
- (3) Matrizen-Test
- (4) Bilder ergänzen

III. Arbeitsgedächtnis

- (1) Zahlen nachsprechen

- (2) Buchstaben-Zahlen-Folgen
- (3) Rechnerisches Denken

IV. Verarbeitungsgeschwindigkeit

- (1) Zahlen-Symbol-Test
- (2) Symbolsuche
- (3) Durchstreich-Test

Altersgruppe: 6;0 bis 16;11 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt zwischen 60 und 90 Minuten. Es handelt sich um einen **Einzeltest**.

Als explizite Voraussetzung zur Durchführung des HAWIK-IV nennen Petermann & Petermann (2007), dass Testleiter/innen in der Anwendung des Verfahrens geübt sein müssen. Die Relevanz des Verfahrens im Bereich der Intelligenz- und Begabungsdagnostik legt eine Anwendung des Tests durch Psychologinnen/Psychologen, Psychotherapeutinnen/-therapeuten sowie u.U. pädagogisches Fachpersonal und Ärztinnen/Ärzte nahe.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte an einer Stichprobe von $N = 1.650$ Kindern aus Deutschland, Österreich und der deutschsprachigen Schweiz in den Jahren 2005 bis 2006.

Für jeden Subtest werden die Rohwerte in Wertpunkten und in Index-Wert-Äquivalenten berechnet. Zur Ermittlung des Gesamt-IQ werden die Wertpunktsummen in ein IQ-Äquivalent umgerechnet. Eine interessante Ergänzung zur Auswertung schlagen Daseking, Petermann & Petermann (2007, zitiert nach Psyndex-Test-Review) mit dem „Allgemeinen Fähigkeits-Index (AFI)“ vor, der insbesondere bei der Intelligenzdagnostik die eigentliche kognitive Fähigkeit genauer zu schätzen vermag als der Gesamt-IQ. Er setzt sich lediglich aus den Indizes „Sprachverständnis“ und „Wahrnehmungsgebundenes Logisches Denken“ zusammen, die nach Ansicht der Autorinnen und Autoren bei Hochbegabten eher repräsentativ für das kognitive Leistungsniveau sind.

Testgüte

Reliabilität

Die Split-Half-Reliabilität der Untertests variiert zwischen $r = .79$ und $r = .91$, auf Indexebene zwischen $r = .87$ und $r = .94$. Für den Gesamttest beträgt die Split-Half-Reliabilität $r = .97$.

Validität

Die explorativen Faktorenanalysen bestätigten die Vier-Faktoren-Struktur des HAWIK-IV. Die vier Faktoren entsprechen den vier Indexwerten.



Konstruktvalidität:

Die Korrelation des HAWIK-IV mit dem Vorgängerverfahren HAWIK-III beträgt beim Gesamt-IQ $r = .73$, bei den Untertests zwischen $r = .34$ und $r = .79$, bei den Indizes zwischen $r = .53$ und $r = .75$, beim Handlungs-IQ $r = .50$ und beim Verbal-IQ $r = .97$. Die Korrelationen der Untertests mit den Noten in sieben verschiedenen Schulfächern variieren zwischen $r = .39$ und $r = .73$.

Autorinnenkommentar

Während die Wechsler-Skalen traditionell auf dem eher pragmatischen Intelligenzkonzept nach David Wechsler basieren und eine Anbindung an neuere Konzepte bisher fehlte, erfolgte nun offenbar eine Orientierung am empirisch abgesicherten Cattell-Horn-Modell (siehe Kapitel 1 in diesem Buch)¹⁸ der Intelligenz (CHC-Modell), auch wenn im Manual nicht explizit darauf Bezug genommen wird (Daseking, Petermann & Petermann, 2007). Das CHC-Modell wird allerdings nicht vollständig abgebildet.

Mit dem HAWIK-IV kann zusätzlich zu den vier Indizes ein Wert für die allgemeine Intelligenz im Sinne des g-Faktors nach Spearman gebildet werden, dieser g-Faktor fehlt im CHC-Modell. Relevant für die Intelligenzdiagnostik ist, dass im Bereich hoher Begabung nicht selten extreme Leistungsunterschiede zwischen den vier Indexwerten bzw. den beiden Indexpaaren (Sprachverständnis – logisches Denken vs. Arbeitsgedächtnis – Verarbeitungsgeschwindigkeit) vorkommen, was die Interpretierbarkeit des Gesamt-IQs 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 scheint hier besser gegeben zu sein.

Bei den Aufgaben zum Arbeitsgedächtnis können Deckeneffekte nachgewiesen werden: Für ältere hochbegabte Kinder und Jugendliche sind die Aufgaben nicht hinreichend schwer. Wichtige Aspekte der Kriteriumsvalidierung (Zusammenhänge mit konstrukt-nahen Testverfahren/Schulnoten) stehen beim HAWIK-IV noch aus.

¹⁸ Zwei Ebenen der Intelligenz werden postuliert; auf der oberen Ebene befinden sich 16 Faktoren, denen wiederum mehr als 80 basale Fähigkeiten untergeordnet sind.

Literatur

- Daseking, M., Petermann, U. & Petermann, F. (2007). *Intelligenzdiagnostik mit dem HAWIK-IV. Kindheit und Entwicklung*, 16(4), 250-259.
- Daseking, M., Petermann, F. & Waldmann, H.-C. (2008). *Der allgemeine Fähigkeitsindex (AFI) – eine Alternative zum Gesamt-Intelligenzquotienten (G-IQ) des HAWIK-IV? Diagnostica*, 54(4), 211-22.
- Petermann, F. & Petermann, U. (2007). *HAWIK-IV. Hamburg-Wechsler-Intelligenztest für Kinder – IV. Übersetzung und Adaptation der WISC-IV von David Wechsler*. Bern: Huber.
- Wechsler, D. (2003). *The WISC-IV technical and interpretive manual*. San Antonio, TX: Psychological Corporation.

3.19 HAWIVA-III; Hannover-Wechsler-Intelligenztest für das Vorschulalter III

Ricken, G., Fritz, A., Schuck, K. D. & Preuss, U. (2007). HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter – III. Bern: Huber.

Beschreibung

Der HAWIVA-III ist ein Intelligenztest zur Erfassung allgemeiner und spezifischer Fähigkeiten bei Kindern im Vorschulalter. Im psychologisch-pädagogischen Anwendungsbereich kann der HAWIVA-III zur Diagnostik hochbegabter junger Vorschulkinder eingesetzt werden, es zeigte sich aber ein Deckeneffekt bei älteren Kindern.

Der HAWIVA-III ist ein individuell durchzuführender Intelligenztest für das Vorschulalter und die ersten Jahre des Schulalters. Er orientiert sich in seiner Grundkonzeption an der Wechsler-Tradition sowie an neuen Erkenntnissen aus der kognitions- und neuropsychologischen Forschung. Der HAWIVA-III wurde auf der Basis der Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI-III, Wechsler, 2002) adaptiert. Neben den allgemeinen intellektuellen Fähigkeiten eines Kindes werden Werte für die Skalen Verbalteil, Handlungsteil, Verarbeitungsgeschwindigkeit und Allgemeine Sprache ermittelt. Im Vorschulalter unterliegen die kognitiven Prozesse und Strukturen starken Veränderungen (White, 1996; Sameroff & Haith, 1996). Piaget (1952) charakterisierte die kognitive Entwicklung als eine Folge qualitativer Veränderungen im logischen Denken: sensumotorische Stufe, präoperatives, konkret-operatorisches und formal-operatorisches Denken. Neueren Erkenntnissen zufolge (z.B. Siegler, 1996) vollzieht sich die Entwicklung in überlappenden Wellen, so dass wohl eine Strategie oder ein Ansatz zur Problemlösung dominiert, andere aber grundsätzlich nicht ausgeschlossen werden. Kinder im Alter von drei Jahren ziehen bereits einfache kausale Schlussfolgerungen, machen aber im Unterschied zu Vierjährigen Fehler, wenn sie gleichzeitig verschiedene Aspekte einer Sache berücksichtigen müssen (Frye, Zelazo, Brooks & Samuels, 1996; zitiert nach Ricken et al., 2007, S. 26).

Um den großen Veränderungen in der Leistungsfähigkeit der Kinder gerecht zu werden, wurden für die Altersbereiche 2;6 bis 3;11 und 4;0 bis 6;11 Jahre zwei unterschiedliche Zusammenstellungen der Subtests gewählt.

Bei den jüngeren Kindern können durch vier Kernuntertests Aussagen über den Verbal- und Handlungsteil gemacht werden.

Die Kernuntertests des Verbalteils sind:

(1) Passiver Wortschatz:

Das Kind zeigt auf das von der Versuchsleiterin/vom Versuchsleiter genannte Bild aus einer Vorlage von vier Bildern.

(2) Allgemeines Wissen:

Das Kind beantwortet mündlich gestellte Fragen zu allgemeinem Wissen und zu Bildern.

Die Kernuntertests des Handlungsteils sind:

(3) Mosaik-Test:

Das Kind baut vorgegebene Muster mit Klötzchen oder gezeichneten Vorlagen nach.

(4) Figuren legen:

Das Kind soll Puzzleteile innerhalb von 90 Sekunden zu einem Bild zusammensetzen.

Für die älteren Kinder gibt es sieben Kernuntertests.

Die Kernuntertests des Verbalteils sind:

(1) Allgemeines Wissen:

Das Kind beantwortet mündlich gestellte Fragen zu allgemeinem Wissen und Bildern.

(2) Begriffe erklären:

Das Kind soll Begriffe definieren, die zuvor laut vorgelesen wurden.

(3) Begriffe erkennen:

Das Kind soll Begriffe erkennen, die durch immer spezifischer werdende Merkmale beschrieben werden.

Den Handlungsteil stellen:

(4) Mosaik-Test:

Das Kind baut vorgegebene Muster mit Klötzchen oder gezeichneten Vorlagen nach.

(5) Matrizen-Test:

Das Kind vervollständigt Muster, indem es das fehlende Teil aus 4 oder 5 Antwortmöglichkeiten herausucht.

(6) Klassen bilden:

Dem Kind werden zwei Bilderreihen vorgegeben. Das Kind wählt aus jeder Reihe jenes Bild aus, das sich mit den anderen zu einer „Klasse“ zusammenfassen lässt.

(7) Kodieren:

Dieser Untertest wird für die Erfassung der Verarbeitungsgeschwindigkeit verwendet. Das Kind malt Striche oder Kreise nach einer gezeichneten Vorlage.

Altersgruppe: 2;6 bis 6;11 Jahre

Hinweise zur Durchführung

Die Bearbeitungszeit liegt zwischen 30–90 Minuten (abhängig von den verwendeten Untertests). Das Verfahren ist als **Einzeltest** konzipiert. Von der Anwenderin/vom Anwender werden gute psychodiagnostische Grundkenntnisse und hinreichende Erfahrungen in der Anwendung und Interpretation standardisierter Testinstrumente vorausgesetzt (Ricken et al., 2007).

Hinweise zu Auswertung und Interpretation

Die Normierung des HAWIVA-III basiert auf einer Stichprobe von insgesamt $N = 1.322$ Kindern im Alter von 2;6 bis 6;11 Jahren und wurde in den Jahren 2004–2005 durchgeführt. Normdaten liegen für Vierteljahresstufen vor. Die Rohwerte werden in Wertpunkte transformiert, die Berechnung der IQ-Werte der Skalen (Verbal-, Handlungsteil, Verarbeitungsgeschwindigkeit, Gesamt-IQ, Allgemeine Sprachskala) erfolgt auf Basis der Wertpunkte der Untertests.

Testgüte

Reliabilität

Für die Gesamtskala konnte eine interne Konsistenz über die verschiedenen Altersgruppen zwischen $\alpha = .89$ und $\alpha = .95$ gezeigt werden. Für die Untertests liegen die mittleren Reliabilitäten der internen Konsistenz zwischen $\alpha = .74$ und $\alpha = .90$. Die Retest-Reliabilität nach 2,5 Monaten für die Untertests liegt für den gesamten Altersbereich bei Werten zwischen $r = .66$ und $r = .92$.

Validität

Explorative Faktorenanalysen erbrachten den Nachweis einer Zwei-Faktorenlösung für den Altersbereich 2;6 bis 3;11 Jahre und einer Drei-Faktorenlösung für den Altersbereich 4;0 bis 7;3 Jahre. Die Zwei- bzw. Drei-Faktorenlösung bestätigte sich auch in den konfirmatorischen Faktorenanalysen für den WPPSI-III.

Die Korrelationen der HAWIVA-III-Untertests mit Untertests aus dem CFT 1 und der K-ABC liegen zwischen $r = .39$ (Allgemeines Wissen und CFT-IQ) und $r = .80$ (Begriffe erklären und K-ABC, Rätsel). Erzieher/innenurteile korrelieren signifikant mit den HAWIVA-III-Werten (z.B. Gesamtskala mit der Einschätzung der Fähigkeit, Neues zu Lernen: $r = .43$). Verhaltensprobleme (beurteilt mit dem Strengths and Difficulties Questionnaire (SDQ)) korrelieren ebenfalls signifikant positiv mit einigen HAWIVA-Untertests und -Indizes.

Autorinnenkommentar

Wie die Testanalysen belegen, gelingt es dem HAWIVA-III in überzeugender Weise, sprachliche Intelligenz und erworbenes Wissen zu erfassen. Damit bietet er eine Ergänzung zu den Intelligenztestverfahren, die bisher für das Kindergarten- und Vorschulalter vorlagen. Gut



eingeführte Intelligenztests wie die K-ABC, der CFT 1, aber auch neuere Verfahren wie der BIVA, prüfen eher sprachunabhängige intellektuelle Fähigkeiten, die der fluiden Intelligenz zuzurechnen sind.

Die Subtests des HAWIVA-III erlauben dagegen eine differenzierte Diagnose sprachlich-kognitiver Fähigkeiten, also der kristallinen Intelligenz, wie aus den geringen Interkorrelationen mit dem CFT 1-Gesamtscore und den relativ höheren Korrelationen mit Subtests der Fertigkeitenskala aus der K-ABC abzuleiten ist.

Zugleich lassen diese Validitätsbefunde aber den Schluss zu, dass ein Ziel der Testrevision, nämlich eine verbesserte Erfassung der fluiden Intelligenz, nicht erreicht wurde.

Diskussionswürdig sind die empirischen Befunde zum Handlungsteil, die darauf hindeuten, dass v.a. die neu entwickelten Subtests *Matrizen-Test* und *Klassen bilden* nicht Handlungsintelligenz untersuchen, sondern ebenfalls maßgeblich von sprachlichen Fähigkeiten beeinflusst werden. Inwieweit dies ein Mangel des Tests ist oder ein Vorzug, der es erlaubt, den HAWIVA-III besonders deutlich als Test für sprachlich-intellektuelle Fähigkeiten zu positionieren, müssen weitere Validitätsuntersuchungen klären.

Anwender/innen, die noch den „alten“ HAWIVA gekannt und verwendet haben, wird besonders verwundern, dass zwar ein Ziel in der Verbesserung der Entwicklungsangemessenheit lag, tatsächlich aber Elemente eingeführt wurden, die den Besonderheiten von Klein- und Vorschulkindern wenig entgegenkommen. Darunter fallen die 30-Sekunden-Zeitgrenzen für verbale Antworten schon bei der jüngsten Altersgruppe (2;6 bis 3;11), das Einbauen von Zeit-Bonus-Punkten, um Deckeneffekte abzufangen und die rigide Sitzordnung.

Beim Ausfüllen der Arbeitsblätter für das Kodieren und die Symbol-Suche sind grafomotorische Fähigkeiten notwendig, die erst bei Schulanfänger/innen sicher zu erwarten wären. Das Material scheint gerade für die jüngste Altersgruppe wenig attraktiv, besteht es doch größtenteils aus Bildtafeln, auf denen etwas gezeigt oder benannt werden muss (zur Testbesprechung siehe Kastner-Koller & Deomann, 2008).

Insgesamt gewinnt man den Eindruck, dass das Wechsler-Testkonzept für das Schulalter auf Vorschulkindern umgelegt wurde, ohne auf die Besonderheiten der Altersgruppe Rücksicht zu nehmen. Es bleibt abzuwarten, welche Erkenntnisse die von den Testautorinnen und -autoren angekündigten weiterführenden Validitätsuntersuchungen liefern. Eine Erweiterung der Normierungsstichprobe auf österreichische Kinder wäre jedenfalls wünschenswert. Zudem sollte in der nächsten Überarbeitung auf die Altersgruppe der Vorschulkindern noch mehr Rücksicht genommen werden.

Literatur

- Frye, D., Zelazo, P. D., Brooks, P. J. & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology*, 32 (1), 120-131.
- Kastner-Koller, U. & Deimann, P. (2008). HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter-III. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40 (1), 49-53.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Ricken, G., Fritz, A., Schuck, K. D. & Preuss, U. (2007). HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter – III. Bern: Huber.
- Sameroff, A. J. & Haith, M. M. (1996). *The five to seven year shift: The age of reason and responsibility*. Chicago, IL: The University of Chicago Press.
- Siegler, R. S. (1996). Unidimensional thinking, multidimensional thinking, and characteristic tendencies of thought. In A. J. Sameroff & M. M. Haith (Hrsg.), *The five to seven year shift: The age of reason and responsibility* (S. 63-84). Chicago, IL: The University of Chicago Press.
- Wechsler, D. (2002). *Wechsler Primary and Preschool Scale of Intelligence – Third edition (WPPSI-III)*. San Antonio, TX: The Psychological Corporation.
- White, S. H. (1996). The child's entry into the „Age of Reason“. In A. J. Sameroff & M. M. Haith (Hrsg.), *The five to seven year shift: The age of reason and responsibility* (S. 17-32). Chicago, IL: The University of Chicago Press.

3.20 HIT 1-2; Heidelberger Intelligenztest 1-2

Kratzmeier, H. (1977). Heidelberger Intelligenztest HIT 1-2. Weinheim: Beltz.
(Anmerkung: 1994 ist das Verfahren in 2., unveränderter Auflage erschienen.)

Beschreibung

Der HIT 1-2 ist ein sprachfreies Verfahren zur Erfassung schulrelevanter Intelligenzkomponenten in den ersten beiden Schuljahren. Er ist zur Aufklärung individueller Lernprobleme geeignet und liefert damit Hinweise für den Einsatz spezifischer Fördermaßnahmen. Die Interpretation der Ergebnisse ist durch die veraltete Normierung eingeschränkt. Positiv zu erwähnen ist die Erfassung eines Kreativitätsindikators.

Der HIT 1-2 ist ein Intelligenztest, mit dem sich Lehrer/innen ein Bild über eine der kognitiven Lernvoraussetzungen ihrer Schüler/innen der 1. und 2. Schulstufen machen können. Der HIT verlangt von der Schülerin/vom Schüler keine verbalen Äußerungen. Durch den hohen Aufforderungsgehalt der Testaufgaben muss nur ein Minimum an verbaler Instruktion gegeben werden. Der Ansatz im HIT geht davon aus, dass Intelligenz durch die mit mehreren Operationalisierungen gemessene Differenzierungsfähigkeit erfasst werden kann. Zusätzlich wird mit dem Verfahren auch das Konstrukt Kreativität erfasst.

Der Test setzt sich aus folgenden 5 Teiltests mit insgesamt 71 Aufgaben zusammen:

Teilttest 1: Relationen

Jeweils 2 zusammengehörende Figuren, Buchstaben, geometrische Gebilde etc. von 5 sollen herausgefunden werden.

Teilttest 2: Logik

Ein vorgegebenes Muster soll entsprechend seinen Aufbaueregeln fortgesetzt werden.

Teilttest 3: Reproduktion

Vorgegebene geometrische Muster und Punktmengen müssen nachgezeichnet werden.

Teilttest 4: Differenzierung

Teile, die zu einem Musterbild gehören, müssen erkannt werden.

Teilttest 5: Kreativität

Ein begonnenes Bild muss fertig gezeichnet werden.

Altersgruppe: 7 bis 8 Jahre

Hinweise zur Durchführung

Die Durchführung ist innerhalb von zwei Unterrichtsstunden (90 bis 100 Min.) in Form einer **Gruppentestung** möglich. Das Verfahren kann von jeder Lehrperson, die in die Prinzipien der Testdurchführung und -auswertung eingewiesen ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Die Eicherhebung wurde 1976 an einer Stichprobe von 2.446 Kindern der 1. und 2. Schulstufe durchgeführt. Zur Umrechnung der Rohwerte in IQ- und Prozentrangwerte finden sich Tabellen für alle Subtests, getrennt für die 1. und 2. Schulstufe. Der Teilttest Kreativität wird gesondert ausgewertet (Zahl der Ergänzungen, Ausgestaltungen, Erweiterungen, Verbindungen werden gewertet und interpretiert). Außerdem sind zur Kreativität Interpretationsbeispiele angeführt.

Testgüte

Reliabilität

Die innere Konsistenz für den Subtest Kreativität liegt bei $\alpha = .85$. Die Split-Half-Reliabilität für den Gesamttest liegt bei $r = .95$, für Relationen beträgt die Korrelation $r = .76$, für Logik ergibt sich ein Wert von $r = .24$, für die Reproduktion $r = .64$ und für Differenzierung $r = .83$.

Validität

Eine Faktorenanalyse nach der Hauptkomponentenmethode ergab einen varianzstarken Generalfaktor, auf dem nur der Kreativitätstest nicht substantiell lädt, sowie einen zweiten (Schul-

leistungs-)Faktor und drei weitere Faktoren, auf denen vereinzelte Subtests eine wesentliche Ladung aufweisen.

Konstruktvalidität: Die Korrelationen mit der CMM 1-3 beträgt $r = .75$, mit dem SON-R $r = .67$ und mit den SPM $r = .42$.

Für den Gesamtwert der Subtests des HIT 1-2 und Schulnoten im 1. Schuljahr fanden sich Korrelationen zwischen $r = .39$ (Sachkunde) und $r = .45$ (Rechnen). Der Subtest Kreativität zeigt keinen systematischen Zusammenhang zu Schulnoten.

Autorinnenkommentar

Der HIT 1-2 eignet sich gut für ein Intelligenzscreening in Form einer Gruppentestung im Schulkontext, lässt aber Individualdiagnosen nur bedingt zu. Um eine verlässliche Intelligenzdiagnostik machen zu können, ist es sowohl für den oberen als auch für den unteren Leistungsbereich unabdingbar, weitere Tests einzusetzen.

Literatur

Kratzmeier, H. (1977). Heidelberger Intelligenztest HIT 1-2. Weinheim: Beltz.

(Anmerkung: 1994 ist das Verfahren in 2., unveränderter Auflage erschienen.)

3.21 HIT 3-4; Heidelberger Intelligenztest 3-4

Kratzmeier, H. (1982). Heidelberger Intelligenztest HIT 3-4. Weinheim: Beltz.

(Anmerkung: 1994 ist das Verfahren in 2., unveränderter Auflage erschienen.)

Beschreibung

Der HIT 3-4 ist für die sprachfreie Erfassung von Intelligenzkomponenten bei 9-10-Jährigen konzipiert. Er ermöglicht Lehrerinnen und Lehrern einen Überblick über die kognitive und kreative Leistungsfähigkeit ihrer Schulklassen und eignet sich für die Schullaufbahnberatung.

Der HIT 3-4 ist ein sprachfreier Intelligenztest, mit dem sich Lehrer/innen ein Bild über eine der kognitiven Lernvoraussetzungen ihrer Schüler/innen der 3. und 4. Schulstufen machen können. Der HIT verlangt von der Schülerin/vom Schüler keine verbalen Äußerungen. Durch den hohen Aufforderungsgehalt der Testaufgaben muss nur ein Minimum an verbaler Instruktion gegeben werden. Der Ansatz im HIT geht davon aus, dass Intelligenz durch die mit mehreren Operationalisierungen gemessene Differenzierungsfähigkeit erfasst werden kann. Zusätzlich wird mit dem Verfahren Kreativität gemessen.

Wie der HIT 1-2 besteht der HIT 3-4 aus folgenden fünf Teiltests:

Teilttest 1: Relationen

Aus 5 vorgegebenen Aufgaben sollen jeweils 2 zusammengehörende Figuren, Buchstaben, geometrische Gebilde etc. herausgefunden werden.

Teilttest 2: Logik

Ein vorgegebenes Muster soll entsprechend seiner Gesetzmäßigkeit fortgesetzt werden.

Teilttest 3: Reproduktion

Vorgegebene geometrische Muster müssen nachgezeichnet werden.

Teilttest 4: Differenzierung

Teile, die zu einem Musterbild gehören, müssen erkannt werden.

Teilttest 5: Kreativität

Ein begonnenes Bild muss fertig gezeichnet werden.

Altersgruppe: 9 bis 10 Jahre

Hinweise zur Durchführung

Die Durchführung ist innerhalb von zwei Unterrichtsstunden (90 bis 100 Min.) in Form einer **Gruppentestung** möglich. Das Verfahren kann von jeder Lehrperson, die in die Prinzipien der Testdurchführung und -auswertung eingewiesen ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Die Normierung stammt aus dem Jahr 1980 mit einer Stichprobe von $N = 3.225$ Kindern aus 3. und 4. Schulstufen. Der Teilttest Kreativität wird im Gegensatz zu den anderen Teiltests gesondert ausgewertet (Zahl der Ergänzungen, Ausgestaltungen, Erweiterungen, Verbindungen werden gewertet und interpretiert). Außerdem sind zur Kreativität Interpretationsbeispiele angeführt. Normwerte bestehen in Form von schulstufenbezogenen (nicht altersnormierten) Prozentrang- und T-Werten sowie IQ-Normen für den Gesamttestwert und für die fünf Teiltests.

Testgüte

Reliabilität

Die Reliabilität des Gesamttests liegt bei $\alpha = .92$ und ist damit als sehr messgenau zu bezeichnen. Die Konsistenz der 4 Teiltests (ausgenommen Kreativitätsteil) variiert zwischen $\alpha = .74$ und $\alpha = .84$.

Validität

Der Gesamttest und der Kreativitätswert sind nur gering miteinander verbunden ($r = .12$), d.h.

es werden damit zwei im Wesentlichen voneinander unabhängige Begabungskomponenten erfasst. Die Korrelation des Gesamttests mit Schulnoten variiert zwischen $r = .37$ (Sachkunde) und $r = .43$ (Rechnen), der Kreativitätstest korreliert mit den Schulnoten zwischen $r = .05$ (Rechnen) und $r = .11$ (Deutsch).

Autorinnenkommentar

Der HIT 3-4 eignet sich gut für ein Intelligenzscreening in Form einer Gruppentestung im Schulkontext, lässt aber Individualdiagnosen nur bedingt zu. Um eine verlässliche Intelligenzdiagnostik machen zu können, ist es sowohl für den oberen als auch für den unteren Leistungsbereich unabdingbar, weitere Tests einzusetzen.

Literatur

Kratzmeier, H. (1982). *Heidelberger Intelligenztest HIT 3-4*. Weinheim: Beltz.

(Anmerkung: 1993 ist das Verfahren in einer 2., unveränderten Auflage erschienen.)

3.22 IBF; Intelligenz-Basis-Faktoren

Ibrahimovic, N., Bulheller, S., Horn, R., Gittler, G., Institut für Test- und Begabungsforschung-GmbH (2006). IBF. Intelligenz-Basis-Faktoren. FrankfurtM.: Harcourt.

Beschreibung

Der IBF erfasst die intellektuelle Leistungsfähigkeit von Jugendlichen und Erwachsenen.

Der IBF versucht mit seinen beiden Versionen IBF-L (leicht) und IBF-S (schwer) die Verwirklichung der Idee, einerseits ein Verfahren zur Diagnose von Begabungsschwerpunkten und damit auch der intellektuellen Leistungsfähigkeit bereit zu stellen, andererseits die Forderung an eine Steigerung der Testeffizienz durch eine Verkürzung der Bearbeitungszeit zu erfüllen.



IBF-L und IBF-S unterscheiden sich durch den unterschiedlichen Schwierigkeitsgrad und verfügen jeweils über 7 Untertests zu je 12 Aufgaben. Darüber hinaus wurde durch eine umfangreiche und über eine breite Altersspanne reichende Normierung angestrebt, ein attraktives Verfahren für eine Vielzahl von diagnostischen Fragestellungen zur Verfügung zu stellen. Die beiden Tests IBF-L und IBF-S liefern mit sieben Untertests (drei sprachlichen, zwei mathematisch-numerischen, einem für das räumliche Vorstellungsvermögen und einem Gedächtnistest) eine Aussage über individuelle Stärken und Schwächen der Testperson.

Altersgruppe: 14 bis 60+ Jahre

Hinweise zur Durchführung

Die Bearbeitungszeit beträgt für den IBF-L etwa 70 Minuten und für den IBF-S-Test etwa 60 Minuten. Er kann als **Einzel- oder Gruppentest** durchgeführt werden. Es gibt keine Angaben zu notwendigen Vorerfahrungen bzw. Diagnostikkenntnissen der Testleiter/innen.

Hinweise zu Auswertung und Interpretation

Mit einer Stichprobe von $N = 3.000$ im Altersbereich von 14 bis über 60 Jahren, die 2004 erhoben wurde, stellt das Verfahren aktuelle Normwerte für die Anwenderin/den Anwender zur Verfügung. Die Zusammensetzung der Normierungsstichprobe ist repräsentativ im Hinblick auf Schulbildung und Geschlecht. Zur Interpretation werden Prozentrang- und T-Werte angeboten.

Testgüte

Reliabilität

Die IBF-L Split-Half-Reliabilität der Untertests variiert zwischen $r = .73$ und $r = .91$.

Validität

Eine Faktorenanalyse bestätigte die Vier-Faktorenstruktur des IBF und konnte 70 Prozent der Varianz aufklären.

Autorinnenkommentar

Der IBF-Test liefert Informationen zu individuellen Stärken und Schwächen der Testperson. Als alleiniges Intelligenzdiagnostikum eignet er sich allerdings nur bedingt. Für die Hochbegabungsdiagnostik sollten jedenfalls zusätzlich andere Verfahren durchgeführt werden, da Deckeneffekte beobachtbar sind.

Literatur

Ibrahimovic, N., Bulheller, S., Horn, R., Gittler, G., Institut für Test- und Begabungsforschung-GmbH (2006). IBF. Intelligenz-Basis-Faktoren. Frankfurt/M.: Harcourt.

3.23 IDS; Intelligence and Development Scales

Grob, A., Meyer, C. S. & Hagemann-von Arx, P. (2009). *Intelligence and Development Scales (IDS). Intelligenz- und Entwicklungsskalen für Kinder von 5–10 Jahren. Bern: Huber.*

Beschreibung

Die IDS erlauben eine differenzierte, entwicklungspsychologische Standortbestimmung für Kinder. Sie sind das erste Verfahren, das Kognition und Entwicklung mit Schwerpunkt Einschulung verbindet. Positiv hervorzuheben ist die Einbeziehung der nicht-kognitiven Funktionen „Sozial-Emotionale Kompetenz“ und „Leistungsmotivation“. Ein Vorteil der IDS besteht in der Flexibilität des modularen Testaufbaus, welcher ein individuell angepasstes Testen erlaubt. Mit 45 Minuten für eine Einschätzung der Intelligenz und jeweils 10–15 Minuten zur Einschätzung der weiteren Funktionsbereiche handelt es sich zugleich um ein zeitökonomisches Verfahren.

Die Intelligence and Development Scales (IDS) sind ein neu entwickeltes Verfahren, deren Konstruktion teilweise auf dem Kramer-Intelligenztest (KIT; Kramer, 1972) beruht. Ausgangspunkt der Entwicklung der IDS war das Anliegen einer Gruppe erfahrener Schulpsychologinnen und -psychologen, den leistungsdagnostischen Ansatz sowie die zeitökonomische und valide Verwendung des Kramer-Intelligenztests insbesondere für die Einschulungsdiagnostik auf einen zeitgemäßen Stand zu bringen (Grob et al., 2009).

Die IDS basieren auf der Tradition von Alfred Binet (Binet & Simon, 1905). Binet (zitiert nach Psyndex-Test-Review IDS) ging bei der Entwicklung von einer konstruktivistischen Perspektive aus. Die IDS verknüpfen diese jedoch mit der modernen kompetenzorientierten Forschung (Stone, Smith & Murphy, 1973; zitiert nach Psyndex-Test-Review IDS), die anstelle von Reifungsprozessen oder der Stufentheorie von Piaget von aufeinander aufbauenden Lebensphasen ausgeht. Zudem bezieht es eine universelle, differenzielle und individuelle Sichtweise mit ein.

Die universelle Perspektive versucht, Gesetzmäßigkeiten in der menschlichen Ontogenese aufzuspüren. Gesucht werden also für die Entwicklung relevante Konzepte und Normen für den entsprechenden Altersbereich. Die differenzielle Perspektive beleuchtet interindividuelle Entwicklungsunterschiede bei motorischen, kognitiven, sozial-emotionalen und motivationalen Eigenschaften.



Bei der individuellen Perspektive werden die Entwicklung des Individuums und das Zusammenspiel seiner verschiedenen Funktionsbereiche beschrieben und erklärt. In den IDS erfolgte die Umsetzung dieser entwicklungspsychologischen Konzepte und Erkenntnisse durch die Betrachtung des Zusammenspiels von Kognition, Psychomotorik, sozial-emotionaler Kompetenz, Sprache, Mathematik und Leistungsmotivation.

Gemäß den Angaben der Autorinnen und Autoren (Grob et al., 2009) bieten die IDS als erstes deutschsprachiges Verfahren eine direkte, fundierte und multidimensionale Einschätzung sozial-emotionaler Kompetenzen. Über die Einschätzung

von Rechenfertigkeiten hinausgehend bieten die IDS ein Screening für logisch-mathematisches Denken. Sprachliche Kompetenzen können explizit mit dem Funktionsbereich Sprache eingeschätzt werden. Mit einem Fragebogen zur standardisierten Einschätzung der Leistungsmotivation kann die Motivation des Kindes während der Testung gemessen werden.

Die IDS beinhalten insgesamt sieben kognitive Untertests, die sich vier Bereichen zuordnen (Wahrnehmung, Aufmerksamkeit, Gedächtnis und Denken) und zu einem Intelligenzwert verrechnen lassen (Grob et al., 2009):

(1) Wahrnehmung visuell

Unterschiedlich lange Striche sollen in eine Rangreihe gebracht werden.

(2) Aufmerksamkeit selektiv

Drei Entenarten mit spezifischen Merkmalen sollen schnell und richtig aus Reihen verschiedener Enten durchgestrichen werden.

(3) Gedächtnis phonologisch

Zahlen- und Buchstabenreihen sollen vorwärts nachgesprochen werden.

(4) Gedächtnis räumlich-visuell

Verschiedene geometrische Figuren sollen aus einer Auswahl an ähnlichen Figuren wieder erkannt werden.

(5) Gedächtnis auditiv

Eine zu Beginn der Testung gehörte Geschichte soll nach 20 Minuten frei und gestützt erinnert werden.

(6) Denken bildlich

Geometrische Figuren sollen mit drei- und viereckigen Klötzchen nachgelegt werden.

(7) Denken konzeptuell

Ein Konzept von drei Bildern soll erkannt und anschließend zweimal aus einer Auswahl anderer Bilder ausgewählt werden.

Die allgemeine Entwicklung wird in den IDS mit den folgenden fünf Funktionsbereichen erhoben (Grob et al., 2009):

(1) Psychomotorik

Grobmotorik: Im Rahmen von drei Turnübungen sollen die Kinder über ein Seil balancieren, einen Tennisball fangen und werfen sowie beidbeinig seitlich über ein Seil hüpfen.

Feinmotorik: Es sollen Perlen und Würfel schnellstmöglich aufgefädelt werden.

Visumotorik: Es sollen geometrische Figuren abgezeichnet werden.

(2) Sozial-Emotionale Kompetenz

Emotionen Erkennen: Auf Fotos sollen Basisemotionen von Kindern anhand ihrer Gesichtsausdrücke erkannt werden.

Emotionen Regulieren: Es sollen Strategien zur Regulierung von negativen Emotionen angegeben werden.

Soziale Situationen Verstehen: Auf Bildern sollen die sozialen Situationen verstanden und erklärt werden.

Sozial Kompetent Handeln. Zu bildlich präsentierten sozialen Situationen sollen kompetente Verhaltensweisen genannt werden.

(3) Mathematik

Logisch-Mathematisches Denken: Es sollen qualitativ unterschiedliche mathematische Verständnisaufgaben gelöst werden.

(4) Sprache

Sprache Rezeptiv: Vorgesprochene Sätze sollen mit Holzfiguren nachgespielt werden.

Sprache Expressiv: Aus Wörtern soll ein sinnvoller Satz gebildet werden.

(5) Leistungsmotivation

Durchhaltevermögen: Die Testleiterin/Der Testleiter schätzt das Durchhaltevermögen am Ende der Testung mit einem standardisierten Fragebogen ein.

Leistungsfreude: Die Testleiterin/Der Testleiter schätzt die Leistungsfreude am Ende der Testung mit einem standardisierten Fragebogen ein.

Altersgruppe: 5 bis 10 Jahre

Hinweise zur Durchführung

- Die Testdauer beträgt 90 bis 120 Minuten.
- Der Test ist als **Einzeltestung** durchzuführen.
- Eine gute Schulung der Testleiter/innen ist wünschenswert.
- Die Funktionsbereiche Kognition, Psychomotorik, Sozial-Emotionale Kompetenz, Mathematik und Sprache sind auch einzeln einsetzbar.
- Eine überwiegend sprachfreie Intelligenztestung ermöglicht eine gerechte Einschätzung der Intelligenz auch bei fremdsprachigen Kindern.

Hinweise zu Auswertung und Interpretation

Normwerte existieren von einer in den Jahren 2007 bis 2008 erhobenen Normierungsstichprobe von $N = 1.330$ Kindern aus Deutschland, Österreich und der Schweiz. Die Rohwerte werden in altersentsprechende Wertpunkte umgewandelt und in ein Entwicklungsprofil eingetragen. Das Ergebnis kann sowohl auf der Ebene der Funktionsbereiche untertestspezifisch als auch auf Ebene der Gesamtentwicklung entwicklungsbereichsspezifisch betrachtet werden. Es ergibt sich eine differenzierte Profilanalyse mit individuellen Stärken und Schwächen des Kindes.

Testgüte

Reliabilität

Die interne Konsistenz der kognitiven Untertests variiert zwischen $\alpha = .68$ und $\alpha = .96$, die Konsistenz der Untertests zur allgemeinen Entwicklung zwischen $\alpha = .57$ und $\alpha = .88$. Für den Intelligenzwert beträgt die interne Konsistenz $\alpha = .92$.

Die Retest-Reliabilität nach 15 Monaten variiert bei den kognitiven Untertests zwischen $r = .45$ und $r = .81$, bei den Untertests der allgemeinen Entwicklung zwischen $r = .34$ und $r = .88$. Beim Intelligenzwert beträgt die Korrelation $r = .83$.

Validität

Mit den IDS konnte eine Leistungssteigerung mit zunehmendem Alter festgestellt und somit bestätigt werden, dass die IDS Entwicklungsschritte abzubilden vermögen. Faktorenanalysen bestätigten die Aufteilung der Untertests in vier Gruppen: Kognition, Sozial-Emotionale Kompetenz, Leistungsmotivation und Psychomotorik. Mit Ausnahme eines Untertests luden die Untertests auch am meisten auf den ihnen zugeordneten Faktor.

Die Korrelationen des Intelligenzwerts aus den IDS mit dem Wert aus dem HAWIK-IV betragen $r = .69$. Die Korrelation mit Schulleistungstests beträgt $r = .41$.

Autorinnenkommentar

Die IDS können für die Leistungsdiagnostik im Alter von 5 bis 8 Jahren durchaus empfohlen werden. Für die Diagnostik im oberen Extrembereich eignen sie sich weniger.

Literatur

- Binet, A. & Simon, T. (1905). *Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux*. *L'Annee Psychologique*, 11, 191-244.
- 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.
- Kramer, J. (1972). *Intelligenztest. Mit einer Einführung in Theorie und Praxis der Intelligenzprüfung* (4. revidierte Auflage). Solothurn: Antonius.

Stone, L. J., Smith, H. T. & Murphy, L. B. (Hrsg., 1973). *The competent infant – research and commentary*. New York: Basic Books.

3.24 ISA; Intelligenz Struktur Analyse

Institut für Test- und Begabungsforschung (ITB) & Gittler, G. (1998). Intelligenz Struktur Analyse. Ein Test zur Messung der Intelligenz. Frankfurt: Swets. 2001 erschien eine 2., erweiterte Auflage (Institut für Test und Begabungsforschung & Gittler, 2001).

Beschreibung

Die ISA dient zur differenziellen Intelligenztestung für Jugendliche und Erwachsene und ist besonders für Schüler/innen weiterführender Schulen im Alter zwischen 14 und 19 Jahren zu empfehlen. Sie ist ein sorgsam entwickeltes und reliables Verfahren zur Messung der allgemeinen Intelligenz sowie der untergeordneten Fähigkeitsbereiche verbale Intelligenz, numerische Intelligenz, figural-räumliche Intelligenz und Gedächtnisleistung. Insbesondere positiv hervorzuheben ist die Gestaltung des Testmaterials, die der Testleiterin/dem Testleiter eine leichte und unkomplizierte Durchführung sowie eine schnelle Auswertung ermöglicht.

Die ISA basiert auf dem Berliner Intelligenzstrukturmodell von Jäger und Mitarbeiter/innen (siehe Kapitel 1), welches ein hierarchisches Intelligenzmodell postuliert. Mit den neun Subtests sollen die vier Fähigkeitsbereiche „verbale Intelligenz“, „numerische Intelligenz“, „figural-räumliche Intelligenz“ und „Gedächtnisleistung“ gemessen werden, die auf der zweiten Ebene der hierarchischen Intelligenzstruktur angesiedelt sind. Die Aggregation der Ergebnisse aus den Subtests ermöglicht darüber hinaus eine Beurteilung der allgemeinen Intelligenz.

Die ISA besteht aus 177 Aufgaben, die folgenden vier Fähigkeitsbereichen zugeordnet werden:

- (1) Verbale Intelligenz oder Verarbeitungskapazität: Sätze ergänzen, Gemeinsamkeiten finden, Beziehungen erschließen, Begriffe bilden
- (2) Numerische Intelligenz oder Verarbeitungskapazität: Zahlenreihen fortsetzen, praktisches Rechnen
- (3) Figurale Intelligenz oder Verarbeitungskapazität: Würfel erkennen, Figuren zusammensetzen
- (4) Merkfähigkeit oder Gedächtnis: Waren merken

Es wurden zwei Kurzformen entwickelt (ISA-L und ISA-S), welche die Anpassung der Testsituation an die Leistungsfähigkeit der Testperson und die Verkürzung der Testzeit ermöglichen.

ISA-L (leicht) dient der Erfassung von Intelligenzleistungen, die unter dem Durchschnitt liegen, ISA-S (schwer) der Erfassung von Leistungen, die über dem Durchschnitt liegen.

Altersgruppe: 14 bis 55 Jahre

Hinweise zur Durchführung

Die Bearbeitungszeit beträgt 90 Minuten. Bei den Kurzformen ISA-L (leicht) beträgt die Dauer max. 60 Minuten, bei ISA-S (schwer) max. 70 Minuten.

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden.

Testleiter/innen sollten über Erfahrungen in der Durchführung von Leistungstests verfügen und sich im Vorfeld mit dem Test vertraut gemacht haben. Es gibt keine Angaben über diagnostische Kenntnisse der Testleiter/innen.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte von 1997 bis 1999 und setzte sich aus einer Normierungsstichprobe von $N = 3.812$ Schülerinnen und Schülern der Abschlussklassen von Hauptschulen, Realschulen und Gymnasien zusammen, aber auch aus Studierenden und Berufstätigen. Die Rohwerte können anhand von Tabellen in altersstufenbezogene und schultypenbezogene T-Werte oder Prozentränge transformiert werden.

Testgüte

Reliabilität

Die interne Konsistenz für die einzelnen Aufgabengruppen variiert zwischen $\alpha = .78$ und $\alpha = .89$, beim Gesamttest beträgt sie $\alpha = .97$.

Validität

Eine explorative Faktorenanalyse bestätigte die Vier-Faktorenstruktur der ISA, mit welcher 81% der Varianz der Testwerte aufgeklärt werden konnten.

Konstruktvalidität: Die Korrelationen der ISA-Gesamttestleistung mit den Advanced Progressive Matrices (APM) beträgt $r = .52$, mit den Standard Progressive Matrices (SPM) lässt sich eine Korrelation von $r = .66$ errechnen.

Autorinnenkommentar

Die ISA eignet sich sehr gut für die Messung intellektueller Begabung gesamt und in Teilgebieten auch im oberen Bereich. Deckeneffekte können nur in den Extrembereichen beobachtet werden. Insgesamt stellt die ISA ein durchaus empfehlenswertes Diagnostikum im Bereich der Intelligenzdiagnostik dar.

Literatur

Institut für Test- und Begabungsforschung (ITB) & Gittler, G. (1998). *Intelligenz Struktur Analyse. Ein Test zur Messung der Intelligenz.* Frankfurt: Swets. 2001 erschien eine 2., erweiterte Auflage (Institut für Test- und Begabungsforschung & Gittler, 2001).

3.25 I-S-T 2000 R; Intelligenz-Struktur-Test 2000 R

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). I-S-T 2000 R. Intelligenz-Struktur-Test 2000 R. Göttingen: Hogrefe.

Beschreibung

Der I-S-T 2000 R ist ein Verfahren zur Intelligenzdiagnostik, das auf die Erfassung verschiedener Intelligenzbereiche sowie der Generalfaktoren Wissen und schlussfolgerndes Denken zielt.

Der I-S-T 2000 R stellt die Weiterentwicklung des I-S-T 2000 dar. Dieser wiederum war 1999 von den Autoren als revidierte, neu normierte und erweiterte Fassung des I-S-T 70 veröffentlicht worden. Der I-S-T 2000 R erfasst wie seine Vorgänger verschiedene Intelligenzbereiche (fünf der sieben Primärfaktoren Thurstones) und auf einer höheren Hierarchieebene die Generalfaktoren fluide und kristallisierte Intelligenz gemäß dem Cattell-Horn-Modell (siehe Kapitel 1 in diesem Buch).

Demnach können mit dem I-S-T 2000 R folgende 11 Fähigkeiten erfasst werden: verbale Intelligenz, figural-räumliche Intelligenz, rechnerische Intelligenz, figurale Merkfähigkeit, schlussfolgerndes Denken, verbales Wissen, figural-bildhaftes Wissen, numerisches Wissen und Wissen (Gesamt) sowie fluide und kristallisierte Intelligenz.

Das Grundmodul besteht aus 11 Untertests mit je 20 Aufgaben, diese decken fünf Fähigkeitsbereiche ab:

(A) Verbale Intelligenz

Umgang mit sprachlichem Material im Rahmen des schlussfolgernden Denkens.



(B) Numerische Intelligenz

Erfasst die Rechenfertigkeit und die Fähigkeit, logische Beziehungen zwischen Zahlen herzustellen.

(C) Figurale Intelligenz

Erfasst den Umgang mit figural-bildhaftem Material (zwei- und dreidimensionale Figuren).

(D) Schlussfolgerndes Denken (Reasoning):

Erfasst schlussfolgerndes, auch formallogisches Denken.

(E) Merkfähigkeit

Es wird die Fähigkeit zum aktiven Einprägen und kurzfristigen Wiedererkennen von Informationen erfasst.

Das Erweiterungsmodul besteht aus einem Test zur Erfassung des erworbenen Wissens und soll die Generalfaktoren fluide und kristallisierte Intelligenz erfassen (Beschreibung siehe Kapitel 1).

Altersgruppe: 15 bis 60 Jahre

Hinweise zur Durchführung

Die Durchführungszeit beträgt für das Grundmodul ca. 80 Minuten, für das Erweiterungsmodul ca. 40 Minuten.

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden.

Die Testung kann durch ausgewiesene Hilfskräfte (z.B. Fachausbildungskandidatinnen und -kandidaten für klinische Psychologie) erfolgen.

Es wird auf Seminare hingewiesen, die die Grundlagen des Tests und die praktische Handhabung zum Inhalt haben.

Hinweise zu Auswertung und Interpretation

Es liegen Standardwert-Normen auf der Basis von inzwischen mehr als 5.800 Personen im Alter zwischen 15 und 60 Jahren vor (Gymnasiastinnen/Gymnasiasten, Nicht-Gymnasiastinnen/-Gymnasiasten).

Die Rohwerte können anhand von Tabellen in Standardwerte oder Prozentränge umgewandelt und auch in einen IQ-Wert transformiert werden.

Testgüte

Reliabilität

Die innere Konsistenz des Grundmoduls beträgt $\alpha = .96$, für den Wissensteil beträgt sie $\alpha = .93$, für die einzelnen Skalen variiert sie zwischen $\alpha = .87$ und $\alpha = .97$. Die Split-Half-Reliabilität der Skalen variiert zwischen $r = .88$ und $r = .96$.

Validität

Faktorenanalysen der Normierungsstichprobe bestätigten die Drei-Faktorenstruktur. Die Korrelation des I-S-T 2000 R (numerische Intelligenz) mit dem Test d2 (Aufmerksamkeitsverfahren, siehe die ÖZBF-Broschüre „Psychologische Diagnostik von Konzentrationsfähigkeit bzw. Aufmerksamkeit im Kindergarten- und Schulalter“ (2012, in Druck)) beträgt $r = .24$. Der I-S-T 2000 R Verbalteil korreliert mit dem HAWIE-R-Wissen mit $r = .48$, mit dem Mehrfachwortschatztest (MWBT) mit $r = .39$; der I-S-T 2000 R zum schlussfolgernden Denken korreliert mit den CFT 20-Matrizen mit $r = .63$ etc. Die Korrelationen des I-S-T 2000 R Reasoning-Gesamtwerts mit der Deutschnote beträgt $r = .14$, mit der Mathematiknote $r = .45$.

Autorinnenkommentar

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.

Literatur

Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). I-S-T 2000 R. Intelligenz-Struktur-Test 2000 R. Göttingen: Hogrefe.

3.26 K-ABC; Kaufman-Assessment Battery for Children

***Melchers, P. & Preuss, U. (1991). K-ABC. Kaufman-Assessment Battery for Children von Alan S. Kaufman und Nadeen L. Kaufman. Deutschsprachige Fassung. Amsterdam: Swets & Zeitlinger.
(Anmerkung: 2006 erschien die 6. Auflage der K-ABC.)***

Beschreibung

Die K-ABC ist ein Test zur Messung von intellektuellen Fähigkeiten und erworbenen Fertigkeiten bei Kindern. Sie ist v.a. für jüngere Kinder und Volksschüler/innen, die mit dem Lesen und Schreiben Probleme haben, und eher für den unteren Leistungsbereich geeignet. Mit dem K-ABC-Ergebnis können Förderansätze für das untersuchte Kind erstellt werden. Sie ist ein sehr stark auf optische Reize bezogenes Verfahren und trägt damit der Tatsache Rechnung, dass der Mensch ein vornehmlich visuell orientiertes Wesen ist.

Die K-ABC basiert auf Erkenntnissen der Neuropsychologie und Konzepten der kognitiven Psychologie zu Theorien der Informationsverarbeitung. Intelligenz wird in der K-ABC als die

Fähigkeit definiert, Probleme durch geistiges Verarbeiten zu lösen und Informationen zu verarbeiten. Melchers & Preuss (1991) vertreten die Ansicht, dass die Messung intellektueller Fähigkeiten von der Messung des Standes erworbener Fertigkeiten (Lernen und Wissen) getrennt werden muss. Deshalb ist die K-ABC in vier Skalen gegliedert: „Skala einzelheitlichen Denkens“, „Skala ganzheitlichen Denkens“ (intellektuelle Fähigkeiten), „Fertigkeitenskala“ und „Sprachfreie Skala“.

Die Skala einzelheitlichen Denkens umfasst:

Untertest (1) Zauberfenster:

Das Kind soll ein Objekt, dessen Bild auf einer Scheibe in einer Drehbewegung hinter einem kleinen Schlitz gezeigt wird, erkennen und benennen.

Untertest (3) Handbewegungen:

Das Kind macht präzise die Folge von Bewegungen nach, die von der Versuchsleiterin/vom Versuchsleiter vorgeführt werden.

Untertest (5) Zahlennachsprechen:

Das Kind soll die von der Versuchsleiterin/vom Versuchsleiter vorgesprochenen Zahlenfolgen richtig wiederholen.

Untertest (7) Wortreihe:

Es werden einige Objekte genannt, auf die das Kind danach in der gleichen Reihenfolge zeigen soll.

Die Skala ganzheitlichen Denkens umfasst:

Untertest (2) Wiedererkennen von Gesichtern:

Dem Kind werden die Fotografien von jeweils 1–2 Gesichtern kurz vorgelegt, die es danach auf einem gezeigten Gruppenfoto wieder erkennen soll.

Untertest (4) Gestaltschließen:

Das Kind soll Zeichnungen, die unvollständig sind, benennen oder beschreiben.

Untertest (6) Dreiecke:

Dem Kind werden Vorlagen gezeigt, die es unter Verwendung von gelben und blauen Gummidreiecken auf dem Tisch nachbauen soll.

Untertest (8) Bildhaftes Ergänzen:

Das Kind hat diejenige Abbildung zu erkennen, die eine Analogie am besten vervollständigt.

Untertest (9) Räumliches Gedächtnis:

Das Kind soll sich die Anordnung von Bildern auf einer Seite einprägen und diese unmittelbar danach wiedergeben, indem es auf die Kästchen eines nachfolgend dargebotenen Rasters zeigt.

Untertest (10) Fotoserie:

Das Kind soll eine ungeordnete Reihe von Fotos in die chronologisch richtige Reihenfolge bringen.

Die sechs Untertests der Fertigkeitenskala messen primär erworbenes Wissen und Können:

Untertest (11) Wortschatz:

Dem Kind werden Abbildungen von Gegenständen gezeigt, die es richtig benennen soll.

Untertest (12) Gesichter und Orte:

Das Kind sieht in den einzelnen Aufgaben Bilder fiktionaler Personen, berühmter Persönlichkeiten oder bekannter Orte bzw. Gegenstände, die es korrekt benennen soll.

Untertest (13) Rechnen:

Dabei handelt es sich um einen Rechentest, der thematisch in die Erzählung eines Zoobesuchs eingebunden ist.

Untertest (14) Rätsel:

Dem Kind werden in Form eines Rätsels einzelne Merkmale aus der Definition eines abstrakten oder konkreten sprachlichen Konzepts vorgelesen, das das Kind aus dieser Beschreibung heraus herleiten und dann benennen soll.

Untertest (15) Lesen/Buchstabieren (wird nur bei gezielten Fragestellungen angewendet).

Die Aufgabe besteht im Vorlesen einzelner vorgelegter Wörter.

Untertest (16) Lesen/Verstehen:

Dem Kind wird bei den einzelnen Aufgaben jeweils ein kurzer Text mit einem Umfang zwischen einem Wort und zwei Sätzen vorgelegt, der eine Handlungsanweisung beinhaltet, die das Kind gestisch-mimisch („wie ein Schauspieler“) darstellen soll.

Altersgruppe: 2;6 bis 12;5 Jahre

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt zwischen 30 Minuten bei zweijährigen und 90 Minuten bei 12-jährigen Kindern.

Der Test wird in Form einer **Einzeltestung** durchgeführt.

Die K-ABC sollte nur von Psychologinnen und Psychologen angewendet werden. Obwohl Durchführung und Auswertung technisch gesehen sehr einfach sind, erfordert eine vollwertige Interpretation erhebliche Testerfahrung und trainierte Fähigkeiten zur Verhaltensbeobachtung.



Hinweise zu Auswertung und Interpretation

Die Normierung fand zwischen 1986 und 1989 an einer Stichprobe von 3.098 Kindern aus Deutschland, Österreich, der Schweiz und Südtirol statt. Die Repräsentativität der Normierungsstichprobe wird auf Grund der Zufallsauswahl der Kinder kritisiert. Es bestehen Skalenergebnisse für die einzelnen Subtests sowie Standardwerte für die Subtests der Fertigkeitenskala und die vier Gesamtskalen.

Testgüte

Reliabilität

Die Split-Half-Reliabilitäten für die vier Skalen bewegen sich über alle elf Altersstufen im Bereich zwischen $r = .83$ und $r = .98$. Die Retest-Reliabilität bewegt sich zwischen $r = .57$ und $r = .96$ für die Untertests.

Validität

Exploratorische und konfirmatorische Faktorenanalysen bestätigten die zweifaktorielle Lösung der Untertests des Intelligenzteils als auch die dreifaktorielle Lösung über alle Untertests, die den Konstrukten einzelheitliches Denken, ganzheitliches Denken und Fertigkeiten zugeordnet sind. Die Korrelationen der einzelnen K-ABC-Gesamtskalen mit der Gesamtleistung des HAWIK-R variieren zwischen $r = .57$ und $r = .79$, mit dem AID zwischen $r = .50$ und $r = .86$ und mit dem Begabungstestsystem (BTS) von Horn zwischen $r = .54$ und $r = .84$.

Autorinnenkommentar

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 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).

Literatur

Horn, R. (2003). *Intelligenztests für Kinder, Heft 1/03. Keine kritische Anmerkung zum K-ABC. Report Psychologie, 28, 189.*

- Melchers, P. & Preuss, U. (1991). *K-ABC. Kaufman-Assessment Battery for Children* von Alan S. Kaufman und Nadeen L. Kaufman. Deutschsprachige Fassung. Amsterdam: Swets & Zeitlinger. (Anmerkung: 2006 erschien die 6. Auflage der K-ABC.)
- Preusche, I. & Leiss, U. (2003). *Intelligenztests für Kinder. HAWIK-III, AID 2 und K-ABC im Vergleich. Report Psychologie, 28, 12-26.*
- Weiß, C. (2008). *Hochbegabtenberatung: Konzepte und Evaluation. Unveröffentlichte Diplomarbeit. Institut für Pädagogische Psychologie, Julius-Maximilians-Universität Würzburg.*

3.27 KFT 1-3; Kognitiver Fähigkeits-Test (Grundschulform)

Heller, K. & Geisler, H.-J. (1983). *Kognitiver Fähigkeits-Test (Grundschulform).* Weinheim: Beltz.

Beschreibung

Der KFT 1-3 ist ein Testsystem zur Erfassung kognitiver (Lern-)Fähigkeiten in der Grundschule. Die mitgeteilten Daten erlauben seinen Einsatz im Rahmen einer wissenschaftlich fundierten Diagnostik.

Das Verfahren ist die deutsche Bearbeitung des Cognitive Abilities Tests (CAT) von Thorndike, Hagen & Lorge (1968).

Mit dem Verfahren sollen in differenzieller Weise schulisch relevante intellektuelle Lern- und Leistungsvoraussetzungen erfasst werden, die auch Vorhersagequalität für den späteren Schulerfolg haben.

Der KFT 1-3 besteht aus insgesamt 60 Aufgaben, die folgenden vier Untertests zugeordnet sind:

(1) Sprachverständnis

Das Kind benennt anhand bildlicher Vorlagen Gegenstände oder Handlungen.

(2) Beziehungserkennen

Das Kind muss zeitliche, räumliche und größenmäßige Beziehungen aus den bildlichen Vorlagen erkennen.

(3) Schlussfolgerndes Denken

Das Kind muss erkennen, in welchem Aspekt sich jeweils vier von fünf Bildvorlagen gleichen; die nicht dazupassende Abbildung ist anzukreuzen.

(4) Rechnerisches Denken

Es wird erfasst, inwieweit das Kind mit Zahlbegriffen und mengenmäßigen Vorstellungen umgehen kann.

Altersgruppe: 1. bis 3. Schulstufe



Hinweise zur Durchführung

Die Testdauer beträgt ca. 45 bis 60 Minuten. Eine Zeitbegrenzung gibt es nicht.

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Das Verfahren kann aufgrund seiner hohen Standardisierung auch von jeder Lehrperson, die mit den Prinzipien der Testdurchführung und -auswertung vertraut ist, angewendet werden.

Hinweise zu Auswertung und Interpretation

In die Normierung aus den Jahren 1978 bis 1979 wurden 4.351 Schüler/innen einbezogen. Alters-

normen liegen für Kinder zwischen 6;0 und 12;0 Jahren vor.

Für die Subtests und die Gesamtleistung liegen T-Werte differenziert nach Schulstufen, Geschlecht und Alter vor. Außerdem besteht die Möglichkeit der Transformation der T-Werte in IQ-Werte und Prozentränge.

Testgüte

Reliabilität

Die interne Konsistenz variiert bei den Untertests zwischen $\alpha = .56$ und $\alpha = .77$, für die Gesamtleistung liegt die interne Konsistenz bei $\alpha = .85$.

Die Retest-Reliabilität nach fünf Wochen variierte auf Subtestebene zwischen $r = .58$ und $r = .80$, der Gesamtestwert variierte zwischen $r = .76$ und $r = .84$.

Validität

Die Ergebnisse der Faktorenanalyse lassen einen Schulleistungs- und einen Intelligenzfaktor erkennen. Die Untertests Sprachverständnis und Beziehungserkennen werden dem Schulleistungsfaktor zugerechnet, die Untertests Schlussfolgerndes Denken und Rechnerisches Denken dem Intelligenzfaktor.

Die Korrelationen der KFT 1-3-Gesamtleistung mit den CFT 1-Untertests variieren zwischen $r = .16$ und $r = .55$, mit den Bildertest (BT) 2-3-Untertests bewegen sie sich zwischen $r = .40$ und $r = .61$. Mit der mündlichen Deutschnote beträgt die Korrelation $r = .53$, mit der Mathematik-Note errechnet sich eine Korrelation von $r = .44$, mit der Note im Sachunterricht ergibt sich ein Korrelationskoeffizient von $r = .53$. Mit dem Notendurchschnitt am engsten hängen die KFT-Subtests zum Beziehungserkennen ($r = .51$) und zum schlussfolgernden Denken ($r = .44$) zusammen. Um die Prognosevalidität zu überprüfen werden die Korrelationen zwischen dem KFT 1-3 und den erreichten Schulnoten nach einem Jahr berichtet. Die Korrelationen der KFT-Gesamtleistung variieren zwischen $r = .53$ (Mathematik) und $r = .18$ (Sport). Auch bei der Prognosevalidität zeigt

sich, dass die KFT-Subtests 2 und 4 (Beziehungserkennen und schlussfolgerndes Denken) am engsten mit den später erreichten Schulnoten zusammenhängen.

Autorinnenkommentar

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.

Literatur

Heller, K. & Geisler, H.-J. (1983). Kognitiver Fähigkeits-Test (Grundschulform). Weinheim: Beltz.
Thorndike, R. L., Hagen, E. & Lorge, I. (1968). Cognitive Abilities Tests. Boston: Houghton-Mifflin.

3.28 KFT 4-12+R; Kognitiver Fähigkeitstest 4-12+R

Heller, K. A. & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz.

Beschreibung

Der KFT 4-12+R dient der Erfassung des kognitiven Fähigkeitsniveaus von Schülerinnen und Schülern der 4. bis 13. Schulstufen. Er wurde zunächst für den Bereich des allgemeinbildenden höheren Schulwesens entwickelt und hat dort inzwischen seine Tauglichkeit als differenzielles Diagnostikum zur Erfassung schulisch relevanter Lernfähigkeiten vielfach unter Beweis gestellt.

Der KFT 4-12+R ist die deutsche Version des Cognitive Abilities Tests (CAT) von Thorndike & Hagen (1971; 1993). Der Cognitive Abilities Test geht auf den Lorge-Thorndike-Intelligence-Test (Lorge, Thorndike & Hagen, 1964) zurück. Der KFT 4-12+R stellt die Revision des KFT 4-13+ von Heller, Gaedicke & Weinläder (1985) dar. Gegenüber der früheren Version wurde die Normierung aktualisiert, statt dem Subtest Figurensynthese wird der Subtest Faltaufgaben verwendet und das Verfahren um die Subtests Satzergänzen und Textrechenaufgaben gekürzt. Mit dem KFT 4-12+R wird sowohl sprachliches und nonverbal-figurales als auch quantitatives Denken erfasst. Darüber hinaus kann das kognitive Gesamtleistungsniveau ermittelt werden.

Im Sinne des Berliner Intelligenzstrukturmodells von Jäger (1984) (siehe Kapitel 1) wird mit dem KFT 4-12+R v.a. Verarbeitungskapazität erfasst.

Der KFT 4-12+R besteht aus insgesamt neun Subtests und 507 Aufgaben, die sich drei Testteilen (verbal, quantitativ, nonverbal) zuordnen lassen (Heller & Perleth, 2000):

Der Verbalteil:

V1 Wortschatz

Zu einem vorgegebenem Wort soll aus einer Reihe von fünf weiteren Wörtern dasjenige herausgefunden werden, das am ehesten zu dem gegebenen Wort passt.

V2 Wortklassifikationen

Bei drei vorgegebenen Wörtern soll der gemeinsame Oberbegriff gefunden werden und aus fünf weiteren Wörtern dasjenige herausgefunden werden, das unter den gleichen Oberbegriff fällt.

V3 Wortanalogien

Aus einem Wortpaar ist deren Gesetzmäßigkeit herauszufinden und zu einem dritten Wort diejenige Antwortalternative zu wählen, die dazu in gleicher Relation steht wie zu dem Wortpaar.

Der quantitative Teil:

Q1 Mengenvergleiche

Die Relationen „kleiner“, „größer“ oder „gleich“ sollen in verschiedenen schulmathematischen Problemen erkannt werden.

Q2 Zahlenreihen

Eine Zahlenreihe soll nach einer bestimmten Regel folgerichtig fortgesetzt werden.

Q3 Gleichungen bilden

Zahlen sollen durch Operationszeichen so miteinander kombiniert werden, dass ein richtiges Ergebnis erzielt wird (Auswahl aus fünf Antwortalternativen).

Der nonverbale Teil:

N1 Figurenklassifikation

Aus einer Reihe von Figuren ist deren Klassifikationsmerkmal herauszufinden und aus fünf weiteren Figuren diejenige Figur zu wählen, die ebenfalls zum Klassifikationsmerkmal gehört.

N2 Figurenanalogie

ähnlich N1

N3 Faltaufgaben

Ein Blatt Papier wird gedanklich mehrmals gefaltet, anschließend werden ebenfalls virtuell Löcher hineingestanz. Es muss entschieden werden, welches von fünf vorgegebenen Mustern sich nach dem Auffalten ergeben würde.

Altersgruppe: 4. bis 13. Schulstufe

Hinweise zur Durchführung

Der KFT ist ein kombinierter Power-Speed-Test.

Für die Testdurchführung ist inklusive Einführung und Pausen mit drei Schulstunden zu rechnen. Jeder der drei Testteile kann in einer Schulstunde durchgeführt werden. Eine Kurzform umfasst die Subtests V1 und V3, Q1 und Q2 sowie N1 und N2.

Der Test kann in Form von **Einzel- oder Gruppentestungen** durchgeführt werden. Der KFT 4-12+R kann sowohl von Psychologinnen/Psychologen als auch von Lehrerinnen/Lehrern aller Schularten nach entsprechender Ausbildung und Einarbeitung durchgeführt und ausgewertet werden.

Hinweise zu Auswertung und Interpretation

Die Neunormierung erfolgte in den Jahren 1995–1997 mit 6.765 Schülerinnen und Schülern aus Grundschulen in Bayern sowie Haupt-, Realschulen und Gymnasien in Bayern und Baden-Württemberg. Es werden T-Werte und Prozentränge sowohl jahrgangs- als auch schultypenspezifisch angeboten. Die Auswertung erfolgt über die Bestimmung der Subtestscores, die drei Testteilscores und das intellektuelle Gesamtleistungsniveau.

Testgüte

Reliabilität

Die Werte der internen Konsistenz liegen über alle Subtests hinweg bei $\alpha = .79$. Die kombinierte Re- und Parallel-Test-Reliabilität zwischen Form A und einer erneuten Testung nach drei Wochen mit der Parallelform B variiert über die Schulstufen zwischen $r = .71$ und $r = .90$.

Validität

Exploratorische Faktorenanalysen zum KFT 4-12+R ergaben eine Dreifaktorenlösung, die zwischen 63.9% und 73.5% der Varianz aufklärte. In den Strukturgleichungsmodellen von Perleth et al. (1994) sowie Perleth & Sierwald (1996) zur Struktur der KFT-Batterie (KFT 4-13+ und KFT 4-12+R) ließen sich ein verbaler (15% der erklärten Varianz), und ein quantitativer Faktor (40% der erklärten Varianz) nachweisen. Für die nonverbalen Subtests zeigte sich kein gemeinsamer Faktor. Heller & Perleth (2000) beschreiben für die Gesamtstruktur des KFT 4-12+R, dass sich ein Großteil der aufgeklärten Varianz (34–88%) auf die „Allgemeine Intelligenz“ zurückführen lässt. Trotz seiner Konzeption als differenzieller Intelligenztest erfasst der KFT 4-12+R bei jüngeren Sekundarstufenschülerinnen und -schülern eher einen übergeordneten Faktor der allgemeinen Intelligenz.

Konstruktvalidität: Korrelationen des 4-12+R mit dem PSB von Horn (1969) und dem CFT von Weiß (1971): Die verbalen 4-12+R-Subtests zeigen deutliche Zusammenhänge mit den Subtests zur Messung des Sprachverständnisses der zwei oben genannten Tests. Die nonverbalen Subtests korrelierten hoch mit den anderen Reasoningsubtests, auch Reasoninganteile der quantitativen Subtests bestätigten sich. Ferner zeigte sich, dass insbesondere Q1 den Schulleistungs-



aspekt erfasst, während Q2 und Q3 zahlengebundenes Denken auf abstrakterer Ebene erfassen (Heller & Perleth, 2000, S. 39).

Kriteriumsvalidität: Die Korrelationen der KFT 4-12+R-Gesamtleistung mit Deutschnoten variieren über die Schulstufen zwischen $r = .09$ bis $r = .56$; mit Mathematiknoten zwischen $r = .15$ und $r = .65$.

Autorinnenkommentar

Die nicht speziell für die Hochbegabungsdiagnostik vorgesehenen KFT-Versionen, insbesondere der KFT 4-12+R, werden trotzdem häufig für die Hochbegabungsdiagnostik eingesetzt (Weiß, 2008), allerdings gäbe es auch Tests mit einer höheren Aufgabenschwierigkeit und aktuelleren Normdaten (z.B. KFT-HB 3+4 und KFT-HB 4-12+ – siehe MHBT), die stattdessen oder zusätzlich eingesetzt werden könnten.

Literatur

- Heller, K. A., Gädike, A.-K. & Weinläder, H. (1985). *Kognitiver Fähigkeits-Test für 4. bis 13. Klassen (KFT 4-12+R)* (2., verbesserte und erweiterte Auflage). Weinheim: Beltz.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.
- Heller, K. A. & Perleth, C. (2007). *MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe*. Göttingen: Hogrefe.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.
- Jäger, A. O. (1984). *Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven*. *Psychologische Rundschau*, 35, 21-35.
- Lorge, I., Thorndike, R. L. & Hagen, E. (1964). *The Lorge-Thorndike Intelligence Tests (multilevel edition)*. Boston: Houghton-Mifflin.
- Perleth, C., Hofmann, U., Schauer, S. & Wernberger, H. (1994). *Intelligence testing in a Bavarian comprehensive school*. *School Psychology International*, 15, 263-277.
- Perleth, C. & Sierwald, W. (1996). *Stabilität und Veränderungen in einem zweidimensionalen Intelligenzstrukturmodell. Ein Beitrag zur psychometrischen Intelligenzforschung. Vortrag auf der 38. TEAP in Eichstätt*.
- Thorndike, R. L. & Hagen, E. P. (1971). *Cognitive Abilities Test*. Boston: Houghton Mifflin.
- Thorndike, R. L. & Hagen, E. P. (1993). *Form 5 Cog AT. Norms booklet*. Chicago: Riverside.
- Weiß, R. (1971). *Grundintelligenzskala 3 (CFT 3)*. Braunschweig: Westermann.
- Weiß, C. (2008). *Hochbegabtenberatung: Konzepte und Evaluation. Unveröffentlichte Diplomarbeit. Institut für Pädagogische Psychologie, Julius-Maximilians-Universität Würzburg*.

3.29 KFT-K; Kognitiver Fähigkeits-Test (Kindergarten)

Heller, K. & Geisler, H.-J. (1983). Kognitiver Fähigkeits-Test (Kindergarten). Weinheim: Beltz.

Beschreibung

Der KFT-K ist ein Testsystem zur Erfassung kognitiver (Lern-)Fähigkeiten im Vorschulalter. Die mitgeteilten Daten erlauben seinen Einsatz im Rahmen einer wissenschaftlich fundierten Diagnostik.

Der KFT-K basiert auf den Primary Tests I der Cognitive Abilities Tests (CAT) von Thorndike & Hagen (1955, 1977). Ziel dieses Verfahrens ist es, die allgemeine Intelligenz bzw. das kognitive Fähigkeitsniveau von Kindergartenkindern zu erfassen.

Das Verfahren besteht aus vier Subtests mit je 15 Mehrfachwahl-Aufgaben:

(1) Sprachverständnis

Das Kind muss anhand von bildlichen Vorlagen Gegenstände oder Handlungen erkennen.

(2) Beziehungserkennen

Das Kind muss zeitliche, räumliche und größenmäßige Beziehungen aus Bildvorlagen erkennen.

(3) Schlussfolgerndes Denken

Das Kind muss aus fünf vorgegebenen Bildern das jeweils nicht dazu passende Bild herausfinden.

(4) Rechnerisches Denken

Das Kind muss mit Zahlen- und Mengenbegriffen umgehen.

Altersgruppe: 5 bis 6 Jahre

Hinweise zur Durchführung

Es gibt keine Zeitbegrenzung. Mit Instruktionen dauert die Testung ca. 85–105 Minuten. Die Durchführung der Testteile ist auf mehrere Tage verteilt.

Der KFT-K kann als **Einzel- oder Gruppentest** (maximal 8 Kinder) durchgeführt werden.

Der Test eignet sich in der Schuleingangspädagogik und Einzelfallhilfe sowie bei der individuellen Begabungs- und Bildungsförderung in Kindergärten und Vorschulklassen (Förderdiagnostik).

Das Verfahren sollte von Psychologinnen und Psychologen durchgeführt werden. Erfahrung im Umgang mit Kindern der entsprechenden Altersgruppe wird vorausgesetzt.

Hinweise zu Auswertung und Interpretation

Es erfolgte eine Normierung des KFT-K mit einer repräsentativen Stichprobe von 509 Kindern im Alter von 5;1 bis 7;0 Jahren in den Jahren 1979 bis 1981.

Es werden T-Werte und Prozentränge angeboten.

Testgüte

Reliabilität

Die interne Konsistenz liegt für die Subskalen zwischen $\alpha = .53$ und $\alpha = .88$, für die Gesamtleistung beträgt der Alpha-Koeffizient $\alpha = .90$.

Retest-Reliabilität: Der KFT-K besitzt bei den 5- bis 6-jährigen Kindern eine hohe Wiederholungszuverlässigkeit ($r = .80$ für die einzelnen Subtests und $r = .93$ für die Gesamtleistung).

Validität

Die Korrelation der KFT-K-Gesamtleistung mit dem Summenwert des CFT 1 beträgt $r = .64$.

Als Hinweise zur Konstruktvalidität werden Gruppenunterschiede angegeben. Es zeigt sich dabei ein deutlicher Altersfortschritt der Testrohwerte, ebenso treten Sozialschichtdifferenzen auf.

Autorinnenkommentar

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.

Literatur

Heller, K. & Geisler, H.-J. (1983). *Kognitiver Fähigkeits-Test (Kindergarten)*. Weinheim: Beltz.
Thorndike, R. L. & Hagen, E. P. (1977). *Measurement and Evaluation in Psychology and Education*. New York: Wiley.

3.30 KLI 4-5 / KLI 4-5 R; Kombinerter Lern- und Intelligenztest für 4. und 5. Klassen

Schröder, H. (2005). KLI 4-5 R. Kombinerter Lern- und Intelligenztest für 4. und 5. Klassen – Revidierte Form (6., revidierte Auflage). Göttingen: Hogrefe.

Beschreibung

Der KLI 4-5 erfasst die Lernfähigkeit (Gedächtnis) und die Denkfähigkeit (Intelligenz) von 10- bis 11-jährigen Kindern.

Der kombinierte Lern- und Intelligenztest KLI 4-5 R geht von der Annahme aus, dass neben dem fächerübergreifenden Vorwissen das Erfassen und Behalten neuer Inhalte (Lernen) sowie geistige Beweglichkeit und Kreativität (Intelligenz) wichtige Einflussgrößen auf schulische Erfolge darstellen (Schröder, 2005, S. 7 f). Eine verlässliche Prognose des Schulerfolgs dürfe sich nicht allein auf die Ermittlung der intellektuellen Fähigkeiten beschränken. Zwar fördere Intelligenz, verstanden als die Fähigkeit, das Denken den jeweiligen situativen Erfordernissen anzupassen, den schulischen Erfolg. Für die individuelle Entfaltung und den Schulerfolg ist es jedoch in Verbindung mit Motiven, Einstellungen und Selbstkonzepten darüber hinaus erforderlich, Erkenntnisse bewahren und reaktivieren zu können (d.h. zu lernen). Denken und Lernen wirken bei der Bewältigung schulischer Anforderungen interaktiv zusammen: Jedes sinnvolle Lernen basiert auf Denkprozessen und Denken erfordert ein Minimum an Lernfähigkeit.

Gemäß den beiden Kategorien Lernen und Denken teilt sich der Test in zwei Teile: Der *Lernteil* bezieht sich auf das Erlernen und Anwenden von Regeln in der verbalen Kommunikation (Geheimschrift Lernen) und im numerischen Bereich (Zahlen umwandeln). Der *Behaltentest* wird am Ende des Tests als letzter Untertest durchgeführt und ermittelt das Behalten des in Geheimschrift und Zahlenumwandeln Gelernten. Der *Intelligenzteil* ermittelt im Sprachbereich Abstraktionsfähigkeit (Gemeinsamkeiten finden) und Leseverständnis (Satzbestimmung), im Zahlenbereich logisches Schließen (Rechenaufgaben) und Kritikfähigkeit (Reihenkorrektur).

Altersgruppe: 4. und 5. Schulstufe

Hinweise zur Durchführung

Die Testdauer beträgt ca. 60 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Spezielle Durchführungsvoraussetzungen werden nicht genannt. Es findet sich lediglich der Hinweis, dass Testleiter/innen vor der Durchführung die Anweisungen genau durchlesen sollen.



Hinweise zu Auswertung und Interpretation

Das Verfahren wurde in seiner ursprünglichen Form an einer Stichprobe von $N = 3.800$ Schülerinnen und Schülern erprobt. Im Verlauf der Revisionen wurden weitere 9.200 Schüler/innen erfasst. Für die aktuelle Version wurde 2004 erneut eine Stichprobe von $N = 680$ Schülerinnen und Schülern untersucht. Normwerte stehen in der Form von T-Werten, Prozenträngen und Schulerwartungswerten zur Verfügung.

Aus den ermittelten Ergebnissen kann ein individuelles Leistungsprofil verteilt über den Lern- und Intelligenzbereich erstellt werden, aus dem

sich innerhalb der Fähigkeitsstruktur persönliche Schwerpunkte und Fördermöglichkeiten erschließen lassen.

Testgüte

Reliabilität

Die Split-Half-Reliabilität für den Gesamttest liegt bei $r = .95$. Die Paralleltest-Reliabilität der beiden Testformen A und B beträgt $r = .93$, die Retest-Reliabilität nach einer Woche liegt bei $r = .85$.

Validität

Die konvergente Validität zwischen dem KLI-Gesamttest mit der Schulleistung beträgt $r = .95$. Prognostische Validität: Die Korrelationen mit dem nach einem Jahr erzielten Schulerfolg (Notensumme der Fächer Deutsch, Mathematik und erste Fremdsprache) betragen $r = .66$ für den Gesamttest, $r = .56$ für den Lerntest und $r = .62$ für den Intelligenztest.

Autorinnenkommentar

Der KLI 4-5 ist für eine verlässliche Intelligenzdiagnostik nur bedingt geeignet. Er erlaubt einen guten Überblick über die einzelnen Fähigkeiten der Kinder, lässt aber nur wenige Aussagen über die Gesamtleistungsfähigkeit zu. Weiters ist anzumerken, dass die Darstellung des theoretischen Hintergrunds, der Testkonstruktion und der psychometrischen Überprüfung des KLI 4-5 R rudimentär, intransparent und vereinzelt auch fehlerhaft ausgefallen ist (z.B. müsste im Fallbeispiel auf S. 16 der Schüler S statt „20“ Rohwertpunkten „30“ erzielt haben, damit das Beispiel einen Sinn ergibt, und in den Normtabellen fehlen die laut Text vorhandenen IQ-Werte). Zwar scheint der Test insofern seinen Zweck zu erfüllen, als er eine sehr hohe Messgenauigkeit und hohe Validität aufweist und eine relativ gute Prognose des Schulerfolgs ermöglicht. Bei der

testmethodischen Prüfung des KLI 4-5 R wird jedoch meist nicht klar, auf welche Verfahrensversion bzw. Stichproben sich die Analysen beziehen. Belege für die konvergente und divergente Validität des Verfahrens (z.B. Korrelationen mit Lern-, Gedächtnis- und Intelligenztests) fehlen ebenso wie Angaben zu den Subtestinterkorrelationen und zur Profilveriabilität. Schließlich erscheint auch der Anspruch, aus den Ergebnissen ließen sich Aussagen über Fördermöglichkeiten ableiten, angesichts der sehr knappen Hinweise zu hoch gegriffen. Das Manual müsste daher dringend einer systematischen Überarbeitung unterzogen und die angegebenen Kennwerte auf die aktuelle Testversion bezogen werden, bevor das Verfahren im Rahmen bedeutungsvoller Entscheidungsprozesse mit gutem Gewissen eingesetzt werden kann.

Literatur

Schröder, H. (2005). *KLI 4-5 R. Kombiniertes Lern- und Intelligenztest für 4. und 5. Klassen – Revidierte Form (6., revidierte Auflage)*. Göttingen: Hogrefe.

3.31 K-TIM; Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene

Melchers, P., Schürmann, S. & Scholten, S. (2006). K-TIM. Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene. Leiden: PITS.

Beschreibung

Der K-TIM ist ein individuell durchzuführender, umfassender Intelligenztest, der auf der Theorie der fluiden und kristallinen Intelligenz beruht. Er ist ein stark differenzierender und sehr umfassender Intelligenztest, mit dem auch spezifische Intelligenzfunktionen von Erwachsenen und Jugendlichen erfasst werden können. Mit einer Durchführungsdauer von 90 Minuten ist der K-TIM so zeitaufwändig wie andere Testverfahren zu einer umfassenden Individualdiagnostik; durch die integrierte Profilinterpretation ist aber eine sehr differenzierte Beurteilung der untersuchten Person möglich.

Der K-TIM basiert auf Theorien und empirischen Belegen aus der Neuropsychologie und der Kognitiven Psychologie, v.a. auf der Theorie fluider und kristalliner Intelligenzanteile von Horn und Cattell. Der K-TIM ist die deutschsprachige Version des Kaufman-Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993).

Gemäß der zugrunde liegenden Theorie erfasst die Skala fluider Intelligenz diejenigen Intelligenzanteile, die insbesondere bei der Bearbeitung und Bewältigung neuartiger Probleme und Anforderungen entscheidend sind. Kristalline Intelligenzanteile sind hingegen das Produkt

früherer Anwendungen fluider Fähigkeiten auf neuartige Problemstellungen. Im Gegensatz zu fluiden Anforderungen hat kristalline Intelligenz eine enge Beziehung zu höherer Schulbildung und kulturellen Faktoren. Kristalline Intelligenz wird u.a. durch Tests zu allgemeinem Wissen, verbalem Verständnis und der Fähigkeit illustriert, eigene Erfahrungen für die aktuelle Situation nutzbringend auszuwerten und anzuwenden zu können.

Der K-TIM umfasst acht Untertests zur Intelligenzmessung sowie zwei weitere, fakultativ durchzuführende Untertests zur Beurteilung verzögerter bzw. unmittelbarer Gedächtnisleistungen.

Die Untertests 1, 4, 6, 10 werden der kristallinen Intelligenz zugezählt:

- (1) Worträtsel
- (4) Auditives Verständnis
- (6) Doppelte Bedeutungen
- (10) Persönlichkeiten

Fluide Intelligenz wird mit den Untertests 2, 3, 5, 9 erfasst:

- (2) Symbole lernen
- (3) Logische Denkschritte
- (5) Zeichen entschlüsseln
- (9) Figurales Gedächtnis

Die zwei fakultativen Untertests sind:

- (7) Symbole – Abruf nach Intervall
- (8) Auditives Verständnis – Abruf nach Intervall

Altersgruppe: 11 bis über 80 Jahre

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt 70 bis 90 Minuten. Der Test ist als **Einzeltest** durchzuführen. Auswertung und v.a. Interpretation des Verfahrens erfordern eine hohe psychologische Kompetenz, so dass seine Anwendung Psychologinnen und Psychologen vorbehalten sein sollte, sofern nicht Vertreter/innen anderer Berufsgruppen über eine entsprechende, hoch qualifizierte Zusatzausbildung verfügen.

Hinweise zur Auswertung und Interpretation

Die aktuelle Normierungsstichprobe umfasst $N = 2.200$ Versuchspersonen aus Deutschland, der Schweiz, Österreich und Südtirol im Alter zwischen 11 und 80 Jahren.

Die Rohwerte sind in standardisierte T-Werte und in Prozentränge transformiert.

Testgüte

Reliabilität

Die Split-Half-Reliabilität der Untertests liegt im Mittel bei $r = .95$; die Split-Half-Reliabilität variiert für die drei Skalen (fluide Intelligenz, kristalline Intelligenz und Gesamtintelligenz) über die einzelnen Altersstufen zwischen $r = .92$ und $r = .98$.

Validität

Konstruktvalidität: In explorativen und konfirmativen Faktorenanalysen wurde ein Hauptfaktor, der so genannte „g“-Faktor, identifiziert. Aber auch die Zweifaktorenlösung bestätigte die theoretische Struktur des K-TIM in eine fluide und kristalline Intelligenz.

Die Korrelationen zwischen den K-TIM-Skalen mit dem HAWIE-R und dem I-S-T 70 liegen zwischen $r = .65$ und $r = .90$.

Autorinnenkommentar

Der K-TIM ist 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 (nicht im oberen Extrembereich).

Literatur

Kaufman, A. S. & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines, MN: American Guidance Service.

Melchers, P., Schürmann, S. & Scholten, S. (2006). *K-TIM. Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene*. Leiden: PITS.

3.32 MHBT-P; Münchner Hochbegabungstestbatterie für die Primarstufe

Heller, K. A. & Perleth, C. (2007). MHBT-P. Münchner Hochbegabungstestbatterie für die Primarstufe. Göttingen: Hogrefe.

Beschreibung

Die MHBT-P ermöglicht die Erfassung von unterschiedlichen Begabungsdimensionen, Persönlichkeits- und sozialen Umweltmerkmalen in der Grundschule. Das MHBT-Inventar wurde speziell für (Hoch-)Begabte entwickelt. Die mit der MHBT erfassten fünf Begabungsbereiche gehören zu den am häufigsten genannten Begabungsformen, die MHBT beschränkt sich also nicht

nur auf die Erfassung der intellektuellen-kognitiven Fähigkeiten. Die MHBT wird nach Angaben der Autoren den aktuellen Erfordernissen einer möglichst umfassenden, differenziellen (Hoch-) Begabungsdiagnostik konzeptuell gerecht.

In der Münchener Hochbegabungstestbatterie (MHBT) sind Tests und Fragebögen zur Erfassung unterschiedlicher (Hoch-)Begabungsdimensionen sowie von relevanten nicht-kognitiven Persönlichkeits- und sozialen Umweltmerkmalen zusammengefasst.

Nach der theoretischen Konzeption der MHBT ist (Hoch-)Begabung ein mehrdimensionales Fähigkeitskonstrukt (Prädiktoren) in einem Netz von nicht-kognitiven und sozialen Moderatorvariablen sowie kriterialen Leistungsbezugsvariablen.

Die Kernhypothese, so Heller & Perleth (2007, S. 9ff), geht von sieben mehr oder weniger eigenständigen Begabungsbereichen aus, von denen in der MHBT fünf erfasst werden können: Intelligenz, Kreativität, Soziale Kompetenz, Musikalität und Psychomotorik. Diese sind in ein motivationales und soziales Bedingungsgefüge eingebettet, das für die Begabungsentwicklung und die Umsetzung individueller Fähigkeitspotenziale in außergewöhnliches Leistungsverhalten relevant ist.

Die einzelnen Tests und Fragebögen wurden im Rahmen der Münchner Hochbegabungsstudie für die Verwendung bei überdurchschnittlich begabten Kindern und Jugendlichen entwickelt und an diese oder ähnliche Zielgruppen möglichst gut angepasst. Der Einsatz der MHBT-P wird in der Einzelfallberatung zur Abklärung von Hochbegabung und zur Talentsuche für Hochbegabtenförderprogramme empfohlen.

Die MHBT-P beinhaltet Skalen zu folgenden Konstrukten:

- Kognitive Fähigkeiten (KFT-HB 3+4 für Schüler/innen sowie Lehrer/innen-Checkliste)
- Kreativität (Schüler/innen-Test und Fragebogen sowie Lehrer/innen-Checkliste)
- Soziale Kompetenzen (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste)
- Motivation (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste)
- Arbeitsverhalten (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste)
- Psychomotorische Fähigkeiten (Lehrer/innen-Checkliste)
- Musikalische Fähigkeiten (Lehrer/innen-Checkliste)

Altersgruppe: 1. bis 4. Schulstufe

Hinweise zur Durchführung

Die Testdauer beträgt zwischen 120 und 240 Minuten.

Der Test ist als **Einzeltestung** durchzuführen.

Laut den Autoren (Heller & Perleth, 2007, S. 13) sollten die Tests und Fragebögen des MHBT-Inventars nur von geschulten psychologischen und pädagogischen Fachkräften, z.B. Schulpsychologinnen/-psychologen und Erziehungsberaterinnen/-beratern, ausgewertet und

interpretiert werden. Für die Durchführung selbst reiche eine entsprechende Schultesterfahrung aus.

Hinweise zu Auswertung und Interpretation

Die Normwerte für den KFT-HB 3+4 sowie das MHBT-Inventar wurden auf der Grundlage einer unausgelesenen Normierungsstichprobe von Schülerinnen und Schülern der Grundschulklassen 3 und 4 (N = 959) in den Jahren 2002–2003 im Rahmen der bayerischen Grundschulstudie erhoben. Die Profildarstellungen basieren auf Hochbegabungs(HB)-Standards von ca. 1.000 hoch begabten, hoch leistenden Achievern sowie Standards von Underachievern (hoch intelligenten Minderleistern). Mit den Normdaten ist eine gute Differenzierung für die einzelnen Skalen im oberen Leistungsbereich möglich.

Die MHBT-Testleistungen der Probandinnen und Probanden müssen für deren Auswertung in ein Computer-Auswertungsprogramm eingegeben werden. Es werden T-Werte (Klassennormen) und/oder Prozentränge mit den entsprechenden Vertrauensintervallen ausgegeben. Über die Profilauswertung erhält man einen guten Überblick zum Testergebnis der Probandin/des Probanden. Im „TRW-Profilblatt“ wird dargestellt, wie viel Prozent des maximalen Rohwertes der betreffenden Skala eine Probandin/ein Proband erreicht hat.

Weiters liegen spezifische Hochbegabtenprofile in verschiedenen Begabungs-Bereichen vor, die für die differenzierte Stärken-Schwächen-Diagnose herangezogen werden können.

Testgüte

Reliabilität

Abhängig vom jeweiligen Test oder Fragebogen ergaben sich in der Normierungsstichprobe Werte von z.T. unter $\alpha = .60$ (z.B. einzelne Fragebogenskalen) bis $\alpha = .95$ (z.B. einzelne KFT-Skalen).

Validität

Die faktorielle Validität des KFT-HB wurde mit Hauptkomponentenanalysen überprüft und rechtfertigte die Bildung einer KFT-Gesamtleistung. Es gelang der Nachweis, dass mit dem KFT-HB zuverlässig relativ stabile Fähigkeitsdimensionen gemessen werden können.

Zwischen der KFT-HB-Gesamtleistung und den Schulnoten beste-



hen Korrelationen zwischen $r = .40$ (Deutsch) und $r = .46$ (Mathematik). Mit den KFT-HB-Skalen konnten die Abiturdurchschnittsnoten im Mittel mit $\beta = .40$ sowie die Leistungskursnoten mit bis zu $\beta = .80$ vorhergesagt werden.

Autorinnenkommentar

Gerade im Bereich über dem Durchschnitt und im oberen Extrembereich eignet sich die MHBT-P sehr gut für die Erstellung einer verlässlichen Hochbegabungsdiagnostik. 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.

Literatur

Heller, K. A. & Perleth, C. (2007). *MHBT-P. Münchner Hochbegabungstestbatterie für die Primarstufe*. Göttingen: Hogrefe.

3.33 MHBT-S; Münchner Hochbegabungstestbatterie für die Sekundarstufe

Heller, K. A. & Perleth, C. (2007). MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe. Göttingen: Hogrefe.

Beschreibung

Die MHBT-S ermöglicht die Erfassung von unterschiedlichen Begabungsdimensionen, Persönlichkeits- und sozialen Umweltmerkmalen ab der 4. Grundschulklasse und in den Sekundarstufen 1 und 2. Sie bietet eine ausgesprochen breite Palette an Testverfahren, die eine differenzierte Hochbegabungsdiagnose erlauben. Positiv hervorzuheben sind auch die umfangreichen Fallbeispiele mit Interpretationsregeln als Hilfestellung bei unterschiedlichen Diagnosebefunden, z.B. im Rahmen einer Talentsuche für Förderprogramme oder in der Einzelfallberatung.

Theoretische Grundlage bildet das Prädiktoren-Moderatoren-Modell zur Erklärung von außergewöhnlichen Leistungen in verschiedenen Bereichen. (Hoch-)Begabung wird als ein mehrdimensionales Fähigkeitskonstrukt (Prädiktoren) in einem Netz von nicht-kognitiven und sozialen Moderatorvariablen sowie kriterialen Leistungsbezugsvariablen verstanden. Das Münchner (Hoch-)Begabungsmodell wurde im Rahmen der Münchner Längsschnittstudien

zur Hochbegabung in den Jahren 1986 bis 1988 und 1995 bis 1997 entwickelt und überprüft. Die Bestandteile der Testbatterie wurden ebenfalls im Rahmen dieser Studien entwickelt und erprobt. Dabei wurden sowohl bestehende Verfahren unverändert oder modifiziert übernommen als auch eigene Neuentwicklungen vorgenommen.

Die MHBT-S enthält folgende Verfahren:

- (1) *Kognitiver Fähigkeits-Test für Hochbegabte (KFT-HB 4-12+)*: Erfassung wichtiger Intelligenzkomponenten (verbale, quantitative und nonverbale Denkfähigkeiten)
- (2) *AW Abwicklungen*: Erfassung des räumlichen Vorstellungsvermögens
- (3) *SP Spiegelbilder*: Erfassung des räumlichen Vorstellungsvermögens
- (4) *APT*: Erfassung physikalisch-technischer Problemlösekompetenzen
- (5) *SK-S Soziale Kompetenzen (Fragebogen)*
- (6) *KRT-S Kreativitätsfragebogen*
- (7) *IFB Interessenfragebogen*
- (8) *FES Fragebogen des Erkenntnisstrebens*
- (9) *LM-S Leistungsmotivations-Fragebogen*
- (10) *AV-S Arbeitsverhaltens-Fragebogen*
- (11) *SCHUL Schulklima-Fragebogen*
- (12) *FAM Familienklima-Fragebogen*

Außerdem enthält die Testbatterie insgesamt fünf Checklisten für Lehrkräfte zu den wichtigsten Hochbegabungsbereichen:

- (13) *Intelligenz*
- (14) *Kreativität*
- (15) *Musikalität*
- (16) *Sozialbegabung*
- (17) *Psychomotorik*

Im Rahmen der Einzelfalldiagnostik kann die MHBT-S zur Diagnose von Hochbegabung, zur Analyse von Begabungsschwerpunkten sowie zur Bedingungsanalyse bei Leistungsproblemen (Underachievement) eingesetzt werden.

Altersgruppe: 4. bis 12. bzw. 13. Schulstufe



Hinweise zur Durchführung

Die Testdauer beträgt insgesamt ca. 165 Minuten.

Der Test ist als **Einzeltest** durchzuführen.

Die erforderliche testdiagnostische Fachkompetenz sehen die Autoren je nach Verfahrensteil unterschiedlich (Heller & Perleth, 2007, S. 13): Der KFT-HB 4-12+ könne von in Diagnostik ausgebildeten Lehrkräften durchgeführt und ausgewertet werden. Dagegen sollten die Tests und Fragebögen des MHBT-Inventars nur von geschulten psychologischen und pädagogischen Fachkräften, z.B. Schulpsychologinnen/-psychologen, Erziehungsberaterinnen/-beratern, ausgewertet und interpretiert werden. Für die Durchführung selbst reiche eine entsprechende Schultesterfahrung aus.

Hinweise zu Auswertung und Interpretation

Die Normierungsstichprobe des MHBT-S-Inventars ($N = 2.053$) stammt aus den Jahren 1996 und 1997 (Perleth, 2006; Heller & Perleth, 2007, S. 15). Es liegen T-Normen für alle Schulstufen vor. Für die Auswertung müssen die Testdaten in das zugehörige Computerprogramm eingegeben werden, eine andere Form der Auswertung ist nicht möglich. Im Manual sind Profile für folgende Gruppen Hochbegabter abgedruckt (S. 120–131): Hochbegabte Achiever, hochbegabte Underachiever, intellektuell Hochbegabte, kreativ Hochbegabte, sozial Hochbegabte, sprachlich Hochbegabte, sprachlich extrem Hochbegabte, sprachlich-kreativ Hochbegabte, mathematisch Hochbegabte, mathematisch extrem Hochbegabte, mathematisch-naturwissenschaftlich Hochbegabte und naturwissenschaftlich-technisch Hochbegabte.

Testgüte

Reliabilität

Es liegen für alle Verfahren der Testbatterie interne Konsistenzkoeffizienten vor. Für den KFT-HB über die Schulstufen werden für die Gesamtleistung Koeffizienten zwischen $\alpha = .92$ und $\alpha = .94$ genannt. Die Split-Half-Reliabilität für die KFT-HB-Gesamtleistung über die Schulstufen variiert zwischen $r = .65$ und $r = .76$.

Für einige Verfahren werden darüber hinaus Paralleltest-Reliabilitäten (KFT-HB 4-12+) sowie Retest-Reliabilitäten angegeben: Die KFT-HB Paralleltest-Reliabilität reicht von $r = .66$ bis $r = .68$. Die Split-Half-Reliabilität variiert beim Subtest Abwicklungen zwischen $r = .71$ und $r = .86$; bei den Spiegelbildern zwischen $r = .56$ und $r = .85$ und bei den Aufgaben aus Physik und Technik zwischen $r = .13$ und $r = .64$.

Validität

Für einzelne Verfahren und Prädiktorguppen wurden Koeffizienten in mittlerer Höhe und z.T. darüber ermittelt, z.B. (in Follow-up-Untersuchungen) mit dem KFT-HB bis zu $r = .79$ mit Studienleistungen bzw. bis zu $r = .80$ mit einzelnen Leistungskursnoten in der Kolleg- bzw. Oberstufe des Gymnasiums.

Autorinnenkommentar

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

Literatur

- Heller, K. A. & Perleth, C. (2007). *MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe*. Göttingen: Hogrefe.
- Perleth, C. (2006). *Die Münchner Hochbegabungs-Testbatterie (MHBT)*. In H. Wagner (Hrsg.), *Intellektuelle Hochbegabung. Aspekte der Diagnostik und Beratung. Tagungsbericht (in Zusammenarbeit mit der Thomas-Morus-Akademie Bensberg; S. 56-69)*. Bad Honnef: Bock.

3.34 MIT; Mannheimer Intelligenztest

Conrad, W., Büscher, P., Hornke, L., Jäger, R., Schweizer, H., Stünzner, W. v. & Wiencke, W. (1986). Mannheimer Intelligenztest. MIT. Weinheim: Beltz.

Beschreibung

Der Mannheimer Intelligenztest dient der Erfassung der intellektuellen Leistungsfähigkeit bei Jugendlichen und Erwachsenen. Er stellt ein sehr ökonomisches Verfahren insbesondere für Gruppenuntersuchungen dar, bei hoher Zuverlässigkeit der erhaltenen Daten.

Der MIT wurde sowohl auf der Basis des globalen Intelligenzmodells als auch nach dem Modell einer Kombination von teilweise unabhängigen Teilfähigkeiten entwickelt. Die Definition von Intelligenz orientierte sich an den Konzepten von Thurstone (1938), Guilford (1964) und Jäger (1967). Das Ziel dieses Verfahrens ist, die bedeutsamen Bereiche intellektuellen Verhaltens zu erfassen.

Dieser theoretischen Konzeption zu Folge setzt sich der MIT aus 10 Untertests mit sprachfreien und sprachgebundenen Aspekten der Intelligenz zusammen:

- (1) Figurenreihen
- (2) Wortbedeutungen
- (3) Dominos



- (4) Buchstabengruppen
- (5) Zahlenreihen
- (6) Wortverhältnisse
- (7) Mosaik
- (8) Sprichwörter
- (9) Zahlensymbole
- (10) Unmöglichkeiten

Altersgruppe: 12 bis 45 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt ca. 60 Minuten.

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden.

Die Testleiter/innen benötigen keine über die bei vergleichbaren Verfahren hinausgehenden Fähigkeiten und Kenntnisse.

Hinweise zu Auswertung und Interpretation

Die Erstnormierung erfolgte in den Jahren 1969–1973. Für die 3. Auflage des MIT (Conrad et al., 1986) stand eine Normierungsstichprobe von $N = 5.236$ Probandinnen und Probanden zur Verfügung.

Für die Interpretation der Testergebnisse können Standard-Werte und Stanine-Werte gesondert nach (1) Altersbereichen, (2) Geschlecht und (3) Schulbildung (Grundschule, Realschule, Gymnasium) ermittelt werden.

Testgüte

Reliabilität

Die interne Konsistenz beträgt für die S-Form $\alpha = .97$, für die T-Form $\alpha = .96$. Die Korrelation zwischen den beiden Parallelformen beträgt $r = .89$. Die Retest-Reliabilität nach einer Woche variiert zwischen $r = .80$ (Form S) und $r = .84$ (Form T).

Validität

Faktorenanalytische Studien belegen einen übergeordneten Faktor einer allgemeinen geistigen Leistungsfähigkeit. Die Korrelation zwischen der MIT-Gesamtleistung und dem Intelligenztest I-S-T 70 beträgt $r = .78$. Die Korrelation der MIT-Gesamtleistung mit dem Lehrer/innenurteil zur Intelligenz beträgt $r = .35$, mit der Deutschnote $r = .34$, mit dem Lehrer/innenurteil zum Selbstvertrauen $r = .27$ und mit der Mathematiknote $r = .27$. Die Einschätzung der eigenen relativen Begabung ist nur gering mit der MIT-Gesamtleistung verbunden ($r = .18$).

Autorinnenkommentar

Der MIT ist ein geeignetes Instrumentarium zur Erfassung der Intelligenz auch im höheren Bereich (bis zu 1,5 Standardabweichungen im Mittelwert).

Literatur

- Conrad, W., Büscher, P., Hornke, L., Jäger, R., Schweizer, H., Stünzner, W. v. & Wiencke, W. (1986). *Mannheimer Intelligenztest. MIT*. Weinheim: Beltz.
- Guilford, J. P. (1964). *Persönlichkeit*. Weinheim: Beltz.
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.

3.35 MWT-A; MWT-B; Mehrfachwahl-Wortschatz-Test

- Form A: Lehrl, S., Merz, J., Erzigkeit, H. & Galster, V. (1974). *Der MWT-A – ein wiederholbarer Intelligenz-Kurztest, der weitgehend unabhängig von seelisch-geistigen Störungen ist. Nervenarzt, 45, 364-369.***
- Form B: Lehrl, S. (1977). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B. Erlangen: Straube.***

Beschreibung

Der MWT ist ein Kurztestverfahren zur Messung des allgemeinen Intelligenzniveaus für Erwachsene.

Er besitzt ein einfaches und zuverlässiges Schema und wird bereits seit langem verwendet. Er kann überall dort angewendet werden, wo eine Grobklassifikation der Intelligenz ausreicht, eine differenzierte Beurteilung der Intelligenzstruktur nicht erforderlich ist und die Intelligenzdiagnose keine weitreichenden Konsequenzen für die Probandin/den Probanden haben wird (Wolfram et al., 1986, S. 64).

Der MWT ist eine Abwandlung des HAWIE-Wortschatztests als Kurztest zur Messung der Intelligenz. Im MWT ist ein umgangssprachliches oder wissenschaftliches Wort aus vier sinnlosen ähnlichen Wörtern zu identifizieren. Der MWT-A besteht aus 33 Aufgaben mit ansteigender Schwierigkeit. Das Konzept der Mehrfachwahl-Wortschatztests provoziert nach Lehrl (1972) keine echten Intelligenzleistungen, sondern ruft im Laufe der Lebensbewältigung angesammelte „Spuren“ ab, weswegen er auch als „Intelligenzspurtest“ zu bezeichnen ist. Die individuelle Leistung kann den Autorinnen und Autoren zufolge als Maß der allgemeinen Intelligenz gelten.

Altersgruppe: Erwachsene ab 18 Jahren

Hinweise zur Durchführung

Die Durchführungszeit beträgt nur 4 bis 6 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden.

Besondere Kenntnisse oder Fähigkeiten sind nicht notwendig für die Durchführung des Verfahrens. Auf Seiten der Probandinnen und Probanden ist lediglich Lesefähigkeit bzw. die Beherrschung der schriftlichen Sprache vorauszusetzen (entspricht einem Mindest-IQ von 75).

Hinweise zu Auswertung und Interpretation

Die Normwerte von Lehrl et al. (1974) sind veraltet, neuere Normwerte wurden von Metzler & Schmidt (1992) veröffentlicht. Normwerte bestehen in der Form von IQ-Standardwerten und Prozenträngen.

Testgüte

Reliabilität

Der Retest-Reliabilitätskoeffizient beträgt $r = .96$.

Validität

Es konnte eine Korrelation des MWT-A mit dem Leistungsprüfsystem (LPS; Horn, 1962) von $r = .74$ berechnet werden.

Autorinnenkommentar

Der MWT eignet sich nicht zur Diagnostik allgemeiner Intelligenz. Er kann diesbezüglich keine zuverlässige und differenzierte Aussage machen. Als Screening ist er durchaus empfehlenswert (die Messgrenzen liegen bei 75 Punkten nach unten und 125 Punkten nach oben, was für ein Intelligenzscreening als ausreichend betrachtet werden kann), als Verfahren zur Intelligenzdiagnostik ist er aber eher abzulehnen.

Literatur

Horn, W. (1962). *Leistungsprüfsystem (LPS)*. Göttingen: Hogrefe.

Lehrl, S. (1972). *Intelligenzspurtests – psychopathologische Messinstrumente der Intelligenz*. *Medizinische Welt*, 23, 167-168.

Lehrl, S. (1977). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B*. Erlangen: Straube.

Lehrl, S., Merz, J., Erzigkeit, H. & Galster, V. (1974). *Der MWT-A – ein wiederholbarer Intelligenz-*

Kurztest, der weitgehend unabhängig von seelisch-geistigen Störungen ist. Nervenarzt, 45, 364-369.

Metzler, P. & Schmidt, K.-H. (1992). RASCH-Skalierung des Mehrfachwahl-Wortschatztests (MWT). Diagnostica, 38, 31-51.

Wolfram, H., Neumann, J. & Wieczorek, V. (1986). Psychologische Leistungstests in der Neurologie und Psychiatrie. Methoden und Normwerte. Leipzig: VEB Georg Thieme.

3.36 NNAT; Naglieri Nonverbal Ability Test

Naglieri, J. A. (1997). Naglieri Nonverbal Ability Test (NNAT). San Antonio, TX: Psychological Corp.

Beschreibung

Der NNAT ist ein nonverbales Verfahren zur Erfassung der allgemeinen Intelligenz bei Kindern und Jugendlichen und wird in den USA als Screening-Test für die Einschätzung von Begabungen und Talenten verwendet. Er wird als ökonomisches und zuverlässiges Verfahren für die Diagnostik der generellen intellektuellen Fähigkeiten beschrieben.

Der Intelligenztest „Naglieri Nonverbal Ability Test“ (NNAT) ist ein neueres Verfahren zur Erfassung der allgemeinen intellektuellen Fähigkeiten und ist auf Grundlage der Matrix Analogies Test-Expanded Form (MAT-Expanded Form; Naglieri, 1985) konzipiert. Der NNAT misst schlussfolgerndes Denken.

Die Aufgaben lassen sich vier unterschiedlichen Itemtypen zuordnen:

(1) Muster komplettieren:

Die Testperson soll das Muster ergänzen.

(2) Analogien: Die Testperson soll Beziehungen zwischen Formen identifizieren.

(3) Seriellles Schlussfolgern:

Die Testperson soll jenes Muster identifizieren, das als nächstes in einer Reihe kommt.

(4) Räumliche Visualisierung:

Die Testperson soll identifizieren können, wie ein rotiertes oder transformiertes Objekt aussieht.

Der NNAT ist durch die Einfachheit in der Anwendung und den minimalen Einsatz der Sprache, um die Aufgaben lösen zu können, für die Testung von Kindern mit divergentem kulturellen und sprachlichen Hintergrund geeignet. Naglieri & Ford (2003) wiesen in einer Studie nach, dass mit Hilfe des Naglieri Nonverbal Ability Tests (NNAT; Naglieri, 1997) eine ähnliche Anzahl von weißen, afro-amerikanischen und hispanischen Kindern als begabt identifiziert werden konnte. Naglieri ist der Ansicht, dass der Test auch Kindern mit Zuwanderungshintergrund oder aus Herkunftsfamilien mit niedrigem Bildungsniveau eine gleichwertige Chance der Testung ihrer Begabungen und Talente ermöglicht. Inwieweit dies auf die deutsche Version des NNAT zutrifft, ist zu prüfen.

Altersgruppe: 5 bis 17 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt ca. 30 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Eine umfangreiche US-Normierung für den Altersbereich von 5 bis 17 Jahren ($N = 66.000$) ist vorhanden, bislang gibt es aber noch keine deutschen Normen. Testnormen liegen in der Form von Standardwerten, Prozenträngen und altersbezogenen Werten vor.

Testgüte

Reliabilität

Die interne Konsistenz für die USA-Normierungsstichprobe reicht von $\alpha = .64$ bis $\alpha = .96$; die Retest-Reliabilität beträgt zwischen $r = .63$ und $r = .93$.

Validität

Die Kriteriumsvalidität liegt zwischen $r = .36$ und $r = .78$ bei der US-Version.

In der Diagnostik zweier klinischer Stichproben wurde der NNAT mit der K-ABC ($N = 107$) und mit dem HAWIK-III ($N = 31$) verglichen. Im Vergleich mit der K-ABC lagen die IQ-Ergebnisse des NNAT etwas höher, die Unterschiede waren allerdings nicht bedeutsam. Die Zusammenhänge des NNAT-IQ mit den K-ABC-IQ-Ergebnissen und auch mit den meisten K-ABC-Untertests waren hoch bis sehr hoch (keine konkreten Werte angegeben).

Autorinnenkommentar

Der NNAT erlaubt eine relativ schnelle und dennoch gute Aussage zur allgemeinen Intelligenz, welche aber oft überschätzt wird. Für die Diagnostik von (Hoch-)Begabung ist daher unbedingt eine Testung mit einem anderen Intelligenztest (z.B. MHB, AID 2) zu empfehlen.

Literatur

- Naglieri, J. A. (1985). *Matrix analogies test expanded form (MAT)*. San Antonio, TX: Psychological Corp.
- Naglieri, J. A. (1997). *Naglieri Nonverbal Ability Test (NNAT)*. San Antonio, TX: Psychological Corp.
- Naglieri, J. A. & Ford, D. (2003). Addressing Underrepresentation of Gifted Minority Children Using the Naglieri Nonverbal Ability Test (NNAT). *Gifted Child Quarterly*, 47(2), 155-160.
- Süss-Burghart, H. (2006). Der Intelligenztest „Naglieri Nonverbal Ability Test (NNAT)“ in der Diagnostik einer klinischen Stichprobe und im Vergleich mit der „Kaufman Assessment Battery for Children (K-ABC)“ und dem „Hamburg-Wechsler Intelligenztest für Kinder (HAWIK-III)“. *Frühförderung interdisziplinär*, 25(3), 113-121.

3.37 PPVT; Peabody Picture Vocabulary Test

Bulheller, S. & Häcker, H. O. (2003). *Peabody Picture Vocabulary Test. PPVT. Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene. Frankfurt/M.: Swets.*

Beschreibung

Der PPVT ist zur Erfassung des Sprachverständnisses für den Einsatz in der Berufseignungsdiagnostik bei Jugendlichen und Erwachsenen gedacht. Es handelt sich um einen Test mit nur einem Aufgabentyp, welcher nicht vorwiegend die Intelligenzleistung erfasst.

Der PPVT gehört zu den „klassischen“ Verfahren, die sich seit der Erstpublikation in den 50er Jahren bei einer Vielzahl von Untersuchungen bewährt haben. Basierend auf der dritten Ausgabe des PPVT-III (Dunn & Dunn, 1997), liegt nun erstmals eine Version in deutscher Sprache vor. Das Verfahren ermöglicht einen Überblick über den (passiven) Wortschatz der Testperson und erfasst auch deren Sprachverständnis. Ein Einsatz des Verfahrens eignet sich den Testautoren zufolge insbesondere dann, wenn die Schulzeit nur zum Teil in einem deutschsprachigen Land verbracht wurde.

Der PPVT besteht aus einer Serie von 89 Aufgaben. Bei jeder Aufgabe muss ein von der Testleiterin/vom Testleiter genannter Begriff einer von vier einfachen Abbildungen zugeordnet werden. In der Individualform reicht es aus, wenn die Testperson auf die Abbildung zeigt, die den Begriff repräsentiert. In der Gruppenversion müssen von den Testpersonen die Antworten notiert werden.

Altersgruppe: 14 bis etwa 60 Jahre



Hinweise zur Durchführung

Die Testdauer wird mit ca. 20 Minuten veranschlagt. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Bezüglich der benötigten Kenntnisse und der Vertrautheit in der Anwendung des Verfahrens sind keine Informationen vorhanden.

Hinweise zu Auswertung und Interpretation

Der Schwerpunkt der deutschen Normierung liegt zwischen 14 und 18 Jahren ($N = 1.261$, Form A bzw. $N = 1.266$, Form B). Es werden Prozentränge

für Alters- und Bildungsgruppen, aber auch für die Gesamtstichprobe angegeben.

Testgüte

Reliabilität

Die innere Konsistenz (Cronbachs Alpha) beträgt für beide Formen $\alpha = .93$. Die Retest-Reliabilität nach einem Monat variiert zwischen $r = .91$ und $r = .93$.

Validität

Für die deutsche Version wird als Validitätsbeleg auf die unterschiedlichen Testleistungen je nach schulischer Bildung verwiesen. US-amerikanische Studien ergaben Korrelationen des PPVT-III mit der Wechsler Intelligence Scale for Children WISC-III zwischen $r = .82$ und $r = .92$.

Autorinnenkommentar

Der PPVT ist für die Diagnostik von intellektueller Begabung ungeeignet und kann für die psychologische Begabungsdiagnostik nicht empfohlen werden. Die Autoren geben auf Seite 2 des Manuals auch an, dass der PPVT nicht zur Messung der allgemeinen Intelligenz gedacht ist: „It is not, however, a comprehensive test of general intelligence“ (Dunn & Dunn, 1981, S. 2). Da die Normstichproben sich fast ausschließlich aus Jugendlichen zusammensetzen, erscheint der Einsatz des PPVT bei Erwachsenen wegen möglicher Kohorteneffekte fraglich.

Literatur

Bulheller, S. & Häcker, H. O. (2003). *Peabody Picture Vocabulary Test. PPVT. Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene*. Frankfurt/M.: Swets.

- Dunn, L. M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised (PPVT-R). Manual for forms L and M.* Bloomington, MN: Pearson Assessment.
- Dunn, L. M. & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III).* Circle Pines, MN: American Guidance Service.

3.38 PSB-R 4-6; Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen

Horn, W., Lukesch, H., Kormann, A. & Mayrhofer, S. (2002). *PSB-R 4-6. Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen. Revidierte Fassung. Göttingen: Hogrefe.*

Beschreibung

Das PSB-R 4-6 ist ein Intelligenztest für Schüler/innen der 4.–6. Schulstufe. Es ist ein recht ökonomisch anwendbares Verfahren zur Erhebung verschiedener für den Schulerfolg relevanter Leistungsbereiche. Das separate Instruktionshft erleichtert die Durchführung des Verfahrens; hilfreich hierfür sind auch die zahlreichen Interpretationsbeispiele im Manual.

Das PSB-R 4-6 stellt die aktuelle, revidierte Fassung des von Horn (1969) entwickelten Prüfsystems für Schul- und Bildungsberatung (PSB) dar. Der Vorgänger des PSB ist das Leistungsprüfsystem (LPS) von Horn (1983). Da das Leistungsprüfsystem sich als sehr aufwändig erwiesen hatte, wurde es unter dem Gesichtspunkt der Testökonomie und der diagnostischen Bedeutsamkeit für Schullaufbahnentscheidungen optimiert und zum PSB verkürzt (Horn et al., 2002, S. 9). Beide Verfahren basieren auf der Tradition des Thurstone'schen Intelligenzkonzepts (Thurstone, 1938). Das PSB ist eines der gebräuchlichsten Intelligenztestverfahren, jedoch wurde ihm immer wieder Skepsis hinsichtlich der empirischen Überprüftheit und Güte entgegengebracht (Fay, 1997). Dieser Umstand und die fällig gewordene Neunormierung führten zu einer grundlegenden Revision des Verfahrens.

Der Test umfasst 10 Subtests:

- (1) Allgemeinwissen: Bei insgesamt 80 Wörtern ist jeweils ein falscher Buchstabe durchzustreichen.
- (2) Zahlenreihen: Es sollen Zahlenreihen logisch fortgesetzt werden.
- (3) Buchstabenreihen: Es sollen Buchstabenreihen logisch fortgesetzt werden.
- (4) Figurale Reihen: Es sollen Figurenreihen logisch fortgesetzt werden.
- (5) Wortflüssigkeit: Zu jeweils vier Anfangsbuchstaben sollen so viele Wörter wie möglich aufgeschrieben werden.

- (6) Gliederungsfähigkeit: Es wird ein Muster vorgegeben. Die Probandin/der Proband soll aus fünf kleinen Zeichen dasjenige markieren, welches in dem Muster versteckt ist.
- (7) Raumvorstellung: Die Probandin/der Proband soll anhand vorgegebener, in Perspektive gezeichneter Körper, einschätzen, wie viele Flächen der Körper besitzt.
- (8) Gemeinsamkeiten finden: Die Probandin/der Proband soll aus fünf vorgegebenen Wörtern dasjenige herausfinden, welches nicht zu den anderen passt.
- (9) Zahlenaddition: Siebenstellige Zahlen sollen addiert werden. Anschließend ist die letzte Ziffer des Endergebnisses durchzustreichen (Konzentrationstest).
- (10) Zahlenvergleich: In zwei Spalten stehen siebenstellige Zahlen, die sich durch eine Ziffer unterscheiden. Diese soll durchgestrichen werden.

Altersgruppe: 4. bis 6. Schulstufe

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt ca. 45 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Die Testung kann durch eingewiesene Hilfskräfte erfolgen.

Hinweise zu Auswertung und Interpretation

Die Neunormierung fand im Schuljahr 1999/2000 statt. In die Normierungsstichprobe waren 1.559 Schüler/innen der 4.–6. Schulstufen aus Bayern und Baden-Württemberg einbezogen. Der Test bietet schulstufen- und schulartenspezifische Standard- und Prozentrangwerte sowie eine grafische Darstellung auf dem Profilbogen. Die Standardwerte können mit Hilfe einer Tabelle auch in IQ-Werte umgerechnet werden. Bei Schülerinnen und Schülern mit nicht-deutscher Muttersprache werden anschließend Korrekturwerte addiert oder subtrahiert, die den Vergleich der Leistungsfähigkeit mit deutschsprachigen Schülerinnen und Schülern ermöglichen sollen.

Testgüte

Reliabilität:

Die internen Konsistenzen für die Subtests variieren zwischen $\alpha = .72$ und $\alpha = .94$.

Validität

Horn et al. (2002, S. 49) berichten Ergebnisse von Faktorenanalysen, bei denen sich unterschiedliche Faktorenlösungen ergaben. Zwei Faktoren ließen sich nachweisen: ein Faktor des Wissens und der sprachlichen Leistungsfähigkeit sowie ein Faktor des schlussfolgernden Den-

kens.

Zwischen dem PSB-Gesamtwert und dem CFT 20 zeigten sich hohe Korrelationen von $r = .55$ bzw. $r = .63$. Die Gesamtleistungen korrelieren zwischen KFT und PSB mit $r = .69$.

Die PSB-R 4-6-Gesamtleistung korrelierte mit der Deutschnote mit $r = .62$, mit der Mathematiknote mit $r = .61$ sowie mit der Note in Heimat- und Sachkunde mit $r = .38$.

Autorinnenkommentar

Das PSB-R 4-6 ist stark an der Schulleistung orientiert. Für eine differenzierte Begabungsdiagnostik ist dieser Test weniger geeignet.

Literatur

Fay, E. (1997). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18(1/2), 36-38.

Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.

Horn, W. (1983). *Leistungsprüfsystem. L-P-S (2., erweiterte Auflage)*. Göttingen: Hogrefe.

Horn, W., Lukesch, H., Kormann, A. & Mayrhofer, S. (2002). *PSB-R 4-6. Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen. Revidierte Fassung*. Göttingen: Hogrefe.

Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.



3.39 PSB-R 6-13; Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen

Horn, W., Lukesch, H., Mayrhofer, S. & Kornmann, A. (2004). PSB-R 6-13. Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen. Revidierte Fassung. Göttingen: Hogrefe.

Beschreibung

Das PSB-R 6-13 ist ein Intelligenztest, der insbesondere in der Bildungsberatung für Schüler/innen der 6.–13. Schulstufe und für Erwachsene angewendet werden kann. Für Erwachsene, deren Schulzeit bereits einige Jahre zurückliegt, eignet sich dieses Verfahren jedoch nicht. Es ist ein recht ökonomisch anwendbares Verfahren zur Erhebung verschiedener für den Schulerfolg relevanter Leistungsbereiche. Es kann in der schulischen Praxis sicherlich mit Nutzen eingesetzt werden. Hilfreich hierfür sind auch die zahlreichen Interpretationsbeispiele im Manual.

Das PSB-R 6-13 stellt die revidierte Fassung des von Horn (1969) entwickelten PSB dar. Dieses war ebenso wie sein Vorgängerverfahren, das Leistungsprüfsystem (LPS), in der Tradition des Thurstone'schen Intelligenzkonzepts entwickelt worden. Da sich bei der Neubearbeitung des PSB zeigte, dass die Items des ursprünglichen PSB für ältere Probandinnen und Probanden bzw. für solche mit einem gymnasialen Bildungshintergrund viel zu leicht waren, wurden für die Mehrzahl der PSB-Subtests neue Items generiert. Nach Bestimmung ihres Schwierigkeitsgrades wurden diese auf zwei getrennte Testverfahren aufgeteilt, dem PSB-R 4-6 für jüngere Probandinnen und Probanden und dem PSB-R 6-13 für ältere oder auch erwachsene Probandinnen und Probanden.

Das PSB-R 6-13 umfasst neun Untertests:

- (1) Allgemeinwissen:
 - (a) Biologie/Medizin/Psychologie
 - (b) Mathematik/Chemie/Physik
 - (c) Geographie/Astronomie
 - (d) Musik/Kultur/Kunst
 - (e) Sprache/Medien/Kommunikation/Dichtung
- (2) Zahlreihen (Reasoning)
- (3) Buchstabenreihen (Reasoning)
- (4) Figurale Reihen (Reasoning)
- (5) Wortflüssigkeit (Verbal)
- (6) Raumvorstellung (Reasoning)
- (7) Gemeinsamkeiten finden (Verbal)
- (8) Zahlenaddition (Konzentration)
- (9) Zahlenvergleich (Konzentration)

Die neun Untertests werden drei Faktoren (Verbal-, Reasoning- und Konzentrationsfaktor) zugeordnet.

Das PSB-R 6-13 eignet sich für die Beratung von Schülerinnen und Schülern bezüglich des Übertritts in eine schulisch anspruchsvollere Ausbildung der Sekundarstufe II. Auf der Grundlage der Faktorebene (Verbal-, Reasoning- und Konzentrationsfaktor) können die Ergebnisse des Verfahrens gute Hinweise, z.B. zu Potenzialabklärungen der Schüler/innen, liefern.

Altersgruppe: 6. bis 13. Schulstufe

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt ca. 45 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Die Testung kann durch eingewiesene Hilfskräfte erfolgen.

Hinweise zu Auswertung und Interpretation

Die Normierungsstichprobe aus dem Jahr 2001 umfasste $N = 7.373$ bayerische Schüler/innen der 6.–13. Schulstufen aus verschiedenen Schularten. Es liegen Standardwerte und Prozentränge für verschiedene Jahrgangsstufen und Schularten vor. Außerdem werden Normen für Erwachsene (ebenfalls testformspezifisch getrennt) angegeben.

Die Rohwerte werden mittels Normtabellen in Standard- und Prozentwerte umgerechnet. Aus den Subtests wird ein Verbal-, Reasoning-, und Konzentrationsfaktor errechnet und für die Gesamtleistung ein IQ-Wert ermittelt. Die Werte werden in einem Profilblatt grafisch dargestellt.

Testgüte

Reliabilität

Die interne Konsistenz der PSB-R 6-13-Gesamtleistung (Form A) beträgt $\alpha = .93$. Die interne Konsistenz variierte auf der Faktorebene zwischen $\alpha = .79$ und $\alpha = .91$, auf der Subtestebene zwischen $\alpha = .58$ und $\alpha = .93$.

Validität

Eine Faktorenanalyse bestätigte eine Dreifaktorenlösung, die 69% (Form A) bzw. 73% (Form B) der Varianz der Testwerte aufklärte. Der erste Faktor wurde als Dimension des verbalen Denkens (V) bezeichnet. Auf dem zweiten Faktor luden die Reasoning-Untertests sowie der Subtest Raumvorstellung. Der dritte Faktor wurde als Anstrengungsbereitschaft und Konzentrationsfähigkeit (K) interpretiert. Der Gesamtwert des PSB-R 6-13 korreliert mit dem CFT 20 mit $r = .41$ (Form A) bzw. $r = .31$ (Form B). In verschiedenen Untersuchungen (Frohnholzer, 2001; Moser, 2003; Horn et al., 2004, S. 28ff; zitiert nach Psyndex-Testreview PSB-R 6-13) wurde der Zusammenhang zwischen PSB-R 6-13-Ergebnissen und Schulnoten von Hauptschülerinnen und -schülern



untersucht. Es ließen sich für die Schulstufen 6–8 signifikante Zusammenhänge zwischen dem PSB-R 6-13-Gesamtwert und den Schulnoten (Deutsch, Englisch, Mathematik) sichern, nicht aber für die Jahrgangsstufe 9. In einigen Studien zeigten sich erwartungsgemäß hohe Korrelationen zwischen den PSB-R 6-13-Untertests (Allgemeines Wissen, Wortflüssigkeit, Gemeinsamkeiten finden) und der Deutschnote, sowie zwischen den Reasoning-Tests (Raumvorstellung, Gemeinsamkeiten finden, Zahlenaddition) und der Mathematiknote (Frohnholzer, 2001; Horn et al., 2004, S. 29). In anderen Studien konnten diese Befunde jedoch nicht repliziert werden. Auch bei Gymnasiastinnen und Gymnasiasten zeigten sich inkonsistente Befunde.

Autorinnenkommentar

Das PSB-R 6-13 eignet sich als Screeninginstrument zur Erfassung intellektueller Begabung vor allem im mittleren Bereich. Eine verlässliche Intelligenzdiagnostik besonders im oberen Bereich ist aufgrund der gewonnenen Ergebnisse des PSB-R nicht möglich.

Literatur

- Frohnholzer, F. (2001). PSB-R 6-13. Untersuchung an der Hauptschule. Unveröffentlichte Zulassungsarbeit, Universität Regensburg, Lehrstuhl Psychologie VI.*
- Moser, A. (2003). Intelligenz und ihre Beziehung zur Schulleistung und Konzentration. Unveröffentlichte Diplomarbeit: Universität Regensburg.*
- Horn, W. (1969). Prüfungssystem für Schul- und Bildungsberatung (PSB). Göttingen: Hogrefe.*
- Horn, W., Lukesch, H., Mayrhofer, S. & Kornmann, A. (2004). PSB-R 6-13. Prüfungssystem für Schul- und Bildungsberatung für 6. bis 13. Klassen. Revidierte Fassung. Göttingen: Hogrefe.*

3.40 SCH-F; Schlauchfiguren

Stumpf, H. & Fay, E. (1983). Schlauchfiguren – Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens (Testmappe mit Handanweisung, Testheften A und B, 10 Antwortbögen und Schablonen A und B). Göttingen: Hogrefe.

Beschreibung

Die SCH-F testen das räumliche Vorstellungsvermögen. Sie sind als Instrument zur Ergänzung und Differenzierung von Testbatterien in den verschiedensten Bereichen der Diagnostik gedacht (z.B. Berufseignungsdiagnostik, pädagogische Diagnostik, besonders im sekundären und tertiären Bildungssystem).

Die „Schlauchfiguren“ sind ein Aufgabentyp zur Prüfung der Raumvorstellung, einer der am besten abgesicherten Primärfaktoren der Intelligenz (McGee, 1979). Die Schlauchfiguren werden als 3D-Aufgaben vorgegeben und sind per Multiple-choice-Verfahren zu beurteilen.

Es werden Fotos von transparenten Würfeln mit verdrehten Schläuchen vorgelegt. In jedem Würfel befindet sich ein schlauchartiges Gebilde. Auf einem danebenstehenden Bild ist der Würfel von einer anderen Seite abgebildet. Die Testpersonen sollen herausfinden, ob der Würfel von rechts, links, oben, unten oder hinten dargestellt ist.

Altersgruppe:

15 bis 20 Jahre (Die Schlauchfiguren können auch bei älteren Probandinnen und Probanden angewendet werden.)

Durchführung

Die Testdauer beträgt in etwa 16 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Weder für die Durchführung noch für die Auswertung ist eine besondere Einweisung erforderlich.

Hinweise zu Auswertung und Interpretation

Die Normierung fand 1982 statt (N = 727 für Parallelform A und N = 674 für Parallelform B). Normwerte bestehen differenziert nach Schulformen und Geschlecht.

Testgüte

Reliabilität

Die interne Konsistenz beträgt $\alpha = .80$; die Split-Half-Reliabilität beträgt $r = .78$; die Retest-Reliabilität beträgt nach 14 Monaten $r = .72$.

Validität

Die Korrelation der Schlauchfiguren mit der deutschen Version des Mechanical Comprehension Tests (MCT; Bennet, 1969) beträgt $r = .62$. Der MCT ist ein klassischer Indikator der „technischen Begabung“, deren wichtigster Subfaktor wiederum die Raumvorstellung ist.

Es lassen sich Korrelationen der Schlauchfiguren mit Konzentrations- und Schnelligkeitsleistungen von $r = .14$ errechnen, die Korrelation mit der Merkfähigkeit beträgt $r = .35$, mit Wahrnehmungsleistungen beträgt der Korrelationskoeffizient $r = .53$ und mit „mathematical reasoning“ $r = .42$ (Stumpf & Fay, 1981).

Autorinnenkommentar

Bezogen auf die Intelligenzdiagnostik ist zu beachten, dass die Schlauchfiguren nur eine Intelligenzfacette (räumliches Vorstellungsvermögen) mit nur einem Aufgabentyp erfassen. Es gibt eindeutige Geschlechtereffekte zugunsten männlicher Testpersonen, dies lässt sich u.U. durch den Vorteil von männlichen Personen im räumlichen Denken erklären, welches positiv mit der Testleistung der Schlauchfiguren korreliert ist. Dadurch, dass ausschließlich das räumliche Denken abgeprüft wird, sind weibliche Testpersonen stark benachteiligt. Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den Schlauchfiguren vorgezogen werden bzw. dieser Test nur als Zusatzinstrument zur Erhebung der räumlichen Vorstellungskraft eingesetzt werden.

Literatur

- Bennett, G. K. (1969). *Mechanical Comprehension Test*. New York: Psychological Corporation.
- McGee, M. G. (1979). *Human spatial abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences*. *Psychological Bulletin*, 86, 889-918.
- Stumpf, H. & Fay, E. (1981). *Entwicklung und Erprobung eines neuartigen Aufgabentyps zur Erfassung des räumlichen Vorstellungsvermögens*. *Diagnostica*, 27(2), 157-174.
- Stumpf, H. & Fay, E. (1983). *Schlauchfiguren – Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens (Testmappe mit Handanweisung, Testheften A und B, 10 Antwortbögen und Schablonen A und B)*. Göttingen: Hogrefe.

3.41 SON-R 2½-7; Snijders-Oomen Non-verbaler Intelligenztest

Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2007). SON-R 2½-7. Non-verbaler Intelligenztest. Göttingen: Hogrefe.

Beschreibung

Der SON-R 2½-7 ist ein sprachfreier Einzeltest zur Erfassung der allgemeinen, fluiden Intelligenz bei Kindern. Er ist ein weitgehend sprachfreies Verfahren, welches eher im niedrigen Bereich differenziert. Er ist äußerst praktikabel und anwenderfreundlich. Das Verfahren ist auch sehr gut für die Diagnostik bei Kindern mit einem fremd- oder zweisprachigen Hintergrund geeignet.

Die erste Ausgabe des Snijders-Oomen Non-Verbalen Intelligenztests wurde 1943 veröffentlicht und war für die Untersuchung gehörloser Kinder bestimmt (Snijders-Oomen, 1943). Der Test zielte darauf ab, nonverbale Aufgaben zum räumlichen Verständnis und zum abstrakten und konkreten Problemlösen vorzugeben, die möglichst wenig durch Bildung und Erziehung beeinflusst sein sollten.

Beim SON-R 2½-7 handelt es sich um eine deutsche Neunormierung der überarbeiteten letzten holländischen Version der SON-Testreihe. Das Verfahren besteht aus Reasoningtests und Tests zum räumlichen Vorstellungsvermögen (Tellegen, Winkel, Wijnberg-Williams & Laros, 1998, S. 11). Es ist auf der Grundlage des Intelligenzmodells von Cattell (1971), der fluiden allgemeinen Intelligenz konzipiert. Es handelt sich dabei um die Fähigkeit, neue Situationen und Probleme zu meistern, ohne dass Vorwissen von großer Bedeutung ist.

Der SON-R 2½-7 besteht aus sechs Subtests mit insgesamt 91 Testaufgaben: Die Subtests (1) Kategorien und (2) Analogien dienen der Erfassung des abstrakten Denkens. Die Subtests (3) Situationen und (4) Puzzles erfassen konkretes Denken. Die Subtests (5) Mosaike und (6) Zeichenmuster dienen der Erfassung des räumlichen Vorstellungsvermögens.

Altersgruppe: 2;6 bis 7 Jahre

Hinweise zur Durchführung

Die Durchführungszeit des SON-R 2½-7 beträgt etwa 40 bis 60 Minuten für das gesamte Verfahren und fällt für jüngere Kinder etwas kürzer aus.

Der Test ist nur als **Einzeltestung** möglich. Die Instruktionen können sowohl verbal als auch non-verbal gegeben und somit individuell auf das Sprachvermögen des Kindes eingestellt werden.

Eine Besonderheit des Tests ist das Feedback nach jedem Item. Bei einer falschen Aufgabenlösung durch das Kind wird zusätzlich das richtige Vorgehen demonstriert. Laut Manual sollten Testleiter/innen die für psychologische Testuntersuchungen an Kindern benötigten Qualifikationen besitzen. Denkbar ist die Durchführung durch Psychologinnen/Psychologen, Pädagoginnen/Pädagogen, Ärztinnen/Ärzten oder durch sorgfältig geschultes Personal.



Hinweise zu Auswertung und Interpretation

Es liegt eine niederländische Normierungsstichprobe ($N = 1.100$) aus den Jahren 1993–1994 sowie eine deutsche Normierungsstichprobe ($N = 1.027$) aus dem Jahr 2005 vor. Die Leistungen in den Subtests können mit Hilfe von Normierungstabellen zu diesen Faktoren und zu einem Intelligenzscore, dem SON-IQ, zusammengefasst und ausgewertet werden. Mit dem ebenfalls angegebenen Referenzalter und dem Prozentrang wird die Leistung des Kindes im Vergleich zu den Leistungen der Gleichaltrigen bewertet.

Testgüte

Reliabilität

Der Gesamtwert des SON-R 2½-7 hat eine interne Konsistenz von $\alpha = .90$. Die Retest-Reliabilität nach drei Monaten beträgt $r = .79$.

Validität

Der SON-R 2½-7 korreliert mit $r = .75$ mit dem WPPSI, mit $r = .51$ mit der K-ABC-Fertigkeitenskala und mit $r = .67$ mit der sprachfreien Skala der K-ABC. Der hohe Zusammenhang mit der sprachfreien Skala der K-ABC gilt als Evidenz für die Sprachunabhängigkeit des SON-R 2½-7.

Autorinnenkommentar

Der SON 2½-7 ist eines der wenigen Intelligenztestverfahren für das Vorschulalter im deutschsprachigen Raum. Es ist solide konstruiert, die Gütekriterien wurden sorgfältig geprüft und die Normen sind aktuell. Der SON R 2½-7 ist insbesondere für die Anwendung mit sprachbeeinträchtigten, entwicklungsverzögerten oder schwach begabten Kindern konstruiert. Bei der Testung – v.a. älterer – hochbegabter Kinder sind Deckeneffekte zu erwarten. Beim Einsatz mit jüngeren Kindergartenkindern ist zudem zu beachten, dass in diesem Alter erhobene Intelligenzschätzungen eher instabil sind und eine Intelligenzdiagnostik vor einem Alter von 5 Jahren wenig sinnvoll ist (Rost, 2000).

Wenn es um eine 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 Vorteil des SON-R 2 ½-7 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.

Literatur

- Cattell, R. B. (1971). *Abilities: their structure, growth and action*. Boston: Houghton Mifflin.
- Rost, D. H. (2000). *Hochbegabte und hochleistende Jugendliche*. Münster: Waxmann.
- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2007). *SON-R 2 ½-7. Non-verbaler Intelligenztest*. Göttingen: Hogrefe.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doof-stomme kinderen (Intelligenzuntersuchung bei taubstummen Kindern)*. Nijmegen: Berkhout.
- Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. & Laros, J. (1998). *Snijders-Oomen Nonverbal Intelligence Test, SON-R 2 ½-7. Manual and Research Report*. Lisse: Swets & Zeitlinger.
- Vock, M. (2008). *Non-verbaler Intelligenztest (SON R 2½-7)*. *Diagnostica*, 54, 112-115.

3.42 SON-R 5½-17; Snijders-Oomen Non-Verbaler Intelligenztest

Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2005). Non-Verbaler Intelligenztest SON-R 5½-17 (3. korrigierte Auflage). Göttingen: Hogrefe.

Beschreibung

Es handelt sich um ein nonverbales Verfahren zur Messung der allgemeinen Intelligenz von Kindern und Jugendlichen. Sein besonderes Verdienst ist die nonverbale Anwendbarkeit, die seinen Einsatz für Populationen eröffnet, die Schwierigkeiten mit gesprochener oder geschriebener Sprache haben (z.B. gehörlose Kinder oder Kinder mit Sprachstörungen, fremdsprachige Kinder). Die Konzipierung als adaptives Testverfahren spart nachweislich Zeit ohne wesentlichen Verlust der Aussagekraft des Verfahrens. Insgesamt kann der SON-R, der von Kindern aufgrund des interessanten und abwechslungsreichen Materials gerne bearbeitet wird, zur Anwendung empfohlen werden.

Der SON-R 5½-17 ist ein Einzeltest zur Untersuchung der Intelligenz, der ohne Verwendung gesprochener oder geschriebener Sprache bei 5 1/2 bis 17-jährigen Kindern und Jugendlichen durchgeführt werden kann.

Die erste Fassung des SON-Tests, ursprünglich für die Testung gehörloser Kinder konzipiert, erschien 1943 (Snijders-Oomen, 1943). Der von Snijders-Oomen entwickelte Test hatte zum Ziel, ein breites Spektrum von Intelligenzfunktionen zu untersuchen, ohne dabei von Sprache abhängig zu sein. So werden im SON nicht nur Testaufgaben mit minimalem Spracheinfluss, wie geometrische Figuren, verwendet, sondern auch Testaufgaben zur Abstraktions- und Kommunikationsfähigkeit sowie Einsicht in soziale Situationen vorgegeben. Mit dem SON-R 5½-17 liegt eine umfassende Revision von zwei früheren Tests der SON-Testreihe vor. Für die Neubearbeitung wurden fünf Subtests übernommen und zwei Subtests (Suchbilder, Zeichenmuster) neu konstruiert.

Der SON-R 5½-17 setzt sich aus sieben Subtests zusammen, die vier Bereiche prüfen:

(1) *Abstraktes Denken*

Subtest 1 Kategorien: Die Testperson muss durch Auswahl zweier Bilder den Begriff finden, der drei Bilder miteinander verbindet.

Subtest 6 Analogien: Eine geometrische Figur wird verändert, die Testperson muss das Prinzip dieser Veränderung finden.

(2) *Konkretes Denken*

Subtest 5 Situationen: Die Testperson muss eine sinnvolle Ergänzung zu einer Zeichnung einer Situation finden, in der ein oder mehrere Teile fehlen.

Subtest 7 Bildgeschichten: Die Testperson muss die richtige Abfolge von Abbildungen finden, die eine kurze Geschichte darstellen.

(3) *Räumliches Denken*

Subtest 2 Mosaik: Die Testperson muss mit rot-weißen Quadraten verschiedene Muster in einem Rahmen nachlegen.

Subtest 4 Zeichenmuster: Die Testperson muss ein unterbrochenes Linienmuster so vervollständigen, dass das ursprüngliche Muster als Ganzes wiederhergestellt ist.

(4) *Perzeption*

Subtest 3 Suchbilder: Die Testperson muss ein Objekt, das fünfzehnmal in einer Zeichnung verborgen ist, so oft wie möglich innerhalb einer festgesetzten Zeit finden.

Altersgruppe: 5;6 bis 17;0 Jahre

Durchführung

Der Test dauert insgesamt ca. 90 Minuten, für die verkürzte Version (4 Subtests) benötigt man ca. 45 Minuten.

Es handelt sich um einen **Einzeltest**. Die Items sind nach Schwierigkeit auf zwei bis drei Serien verteilt, so dass die Testung adaptiv erfolgen kann, d.h. nicht jeder Testperson wird dieselbe Reihe von Items vorgelegt. Die Wahl der Items hängt von der Leistung bei den bereits bearbeiteten Items und evtl. von bestimmten Eigenschaften der Testperson ab. Laut Manual sollte die Testleiterin/der Testleiter die für psychologische Testuntersuchungen an Kindern und Jugendlichen benötigten Qualifikationen besitzen. Denkbar ist die Durchführung durch Psychologinnen/Psychologen, Pädagoginnen/Pädagogen, Ärztinnen/Ärzte oder durch sorgfältig geschultes Personal.

Hinweise zu Auswertung und Interpretation

Die repräsentative Normierungsstichprobe aus den Jahren 1984–85 umfasste 1.350 hörende Kinder aus den Niederlanden im Alter von 6;6 bis 14;6 Jahren. Snijders et al. (1997, S. 37) weisen darauf hin, dass die Normen auch für andere Länder anzuwenden sind, da eine mit dem SON-

R 5½-17 durchgeführte Untersuchung an deutschen, belgischen und niederländischen Kindern (Snijders & Snijders-Oomen, 1970) große Übereinstimmungen hinsichtlich der Leistungen erbrachte. Die Berechnung der Normwerte basiert auf dem exakten Alter der Kinder; es bestehen Normtabellen für 38 Altersklassen. Das Testergebnis wird in Form von IQ-Wert, Prozentrang und als Referenzalter bekannt gegeben.

Testgüte

Reliabilität

Die interne Konsistenz (Cronbachs Alpha) für den Gesamttest liegt bei $\alpha = .85$.

Validität

Eine Hauptkomponentenanalyse ergab eine Einfaktorenstruktur, wobei der erste Faktor als allgemeiner „non-verbaler Intelligenzfaktor“ interpretiert wird. Die Aufteilung der Tests in vier Gruppen ließ sich nur eingeschränkt bestätigen.

Die Korrelation des Lehrer/innenurteils mit dem SON-IQ-Wert beträgt $r = .33$. Der SON-R unterscheidet sich von anderen, allgemeineren Intelligenztests dadurch, dass Aufgaben, die mehr auf „Kenntnis“ gerichtet sind, fehlen (Snijders et al., 2005, S. 79). Derartige Aufgaben liegen bei anderen Tests oft in verbaler Form vor, wobei sich zeigt, dass diese am stärksten mit schulischen Leistungen korrelieren. Wegen dieser inhaltlichen Unterschiede, so die Autoren, werden die IQ-Werte des SON-R nicht so stark mit der Schulkarriere zusammenhängen wie IQ-Werte anderer Intelligenztests.

Autorinnenkommentar

Auch in der Altersvariante für das Schulalter liegt mit dem SON-R 5½-17 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 von der Anwenderin/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. Schwerer wiegen dürfte jedoch das Alter der Normen, 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 werden kann.

Literatur

Snijders, J. T. & Snijders-Oomen, N. (1970). *Snijders-Oomen nicht-verbale Intelligenztestreihe (S.O.N.)*. Groningen: Wolters-Nordhoff.

- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (1997). *Snijders-Oomen Non-verbaler Intelligenztest SON-R 5½-17* (2. korrigierte Auflage). Frankfurt/M.: Swets.
- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2005). *Non-Verbaler Intelligenztest SON-R 5½-17* (3. korrigierte Auflage). Göttingen: Hogrefe.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen (Intelligenzuntersuchung bei taubstummen Kindern)*. Nijmegen: Berkhout.

3.43 SPM; Standard Progressive Matrices

Heller, K. A., Kratzmeier, H., Lengfelder, A. & Kratzmeier, H. (1998). *Matrizen-Test-Manual, Band 1*. Göttingen: Beltz.

Beschreibung

Die SPM sind ein sprachfreies Verfahren zur Erfassung der allgemeinen Intelligenz bei Kindern, Jugendlichen und Erwachsenen. Sie sollten nicht bei Personen mit höherer Schulbildung verwendet werden, da der Test im oberen Leistungsbereich zu wenig differenziert. Insgesamt können die SPM in Kombination mit anderen Tests als nützliches Verfahren zur Messung der allgemeinen Intelligenz angesehen werden.

Die Standard Progressive Matrices von Raven wurden erstmals 1938 veröffentlicht, um die unterschiedlichsten Grade kognitiver Fähigkeiten sprachfrei und kulturunabhängig zu erfassen. Die SPM messen neben der allgemeinen Intelligenz (g-Faktor) auch Unterscheidungsgenauigkeit und induktives, räumliches Denken (Heller et al., 1998). In seiner Reanalyse verschiedener faktorenanalytischer Untersuchungen kommt Carroll (1993) zum Schluss, dass die SPM ein gutes Maß für den g-Faktor und den Faktor fluide Intelligenz darstellen, jedoch sei die SPM-Messung von induktivem Denken und Visualisierung fraglich.

Die geringe Differenzierungskraft der SPM-Testaufgaben im oberen Leistungsbereich wurde in den späten 1940er Jahren mit der Entwicklung der Advanced Progressive Matrices (APM) kompensiert. Zur genaueren Differenzierung im unteren Leistungsbereich wurden die Coloured Progressive Matrices (CPM) entwickelt.

Die von Raven angestrebte Kulturunabhängigkeit und Sprachunabhängigkeit der SPM konnte nur eingeschränkt bestätigt werden. Eine erste deutsche Version des Tests wurde 1979 von Kratzmeier & Horn veröffentlicht.

Die SPM bestehen aus fünf Sets (A bis E) mit jeweils 12 Aufgaben. Die insgesamt 60 Aufgaben bestehen aus Mustern einfacher Figuren, wobei ein Teil des Musters ausgelassen ist. Die Testpersonen müssen dasjenige Muster aus einer gegebenen Auswahl von Mustern finden, das aufgrund einer oder mehrerer Gesetzmäßigkeiten in diese Lücke passt. Die Aufgabenschwierigkeit nimmt im Verlauf des Tests zu.

Altersgruppe: ab 6 Jahren

Hinweise zur Durchführung

Es besteht keine Zeitbegrenzung (kein Speedtest), die Testdauer beträgt im Regelfall ca. 60 Minuten.

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Die Testung kann durch eingewiesene Hilfskräfte erfolgen. Die Testleiter/innen sollten die Testanleitung vor der Testdurchführung eingehend studieren und einüben.

Sowohl Instruktion als auch Aufgabenbearbeitung können nonverbal erfolgen.

Hinweise zu Auswertung und Interpretation

Die deutsche Neunormierung fand in den Jahren 1996 und 1997 an Schülerinnen und Schülern der Grundschule und Sekundarstufe (Haupt-, Realschule und Gymnasium; N = 2.134) statt. Es bestehen alters- und schultypenspezifische Normen (Prozentränge inkl. Angaben zum Vertrauensintervall).

Testgüte

Reliabilität

Die interne Konsistenz der SPM liegt schultyp- und schulstufenabhängig zwischen $\alpha = .75$ und $\alpha = .97$. Die Retest-Reliabilität nach drei Monaten liegt zwischen $r = .72$ (Sekundarschüler/innen) und $r = .91$ (Realschüler/innen).

Validität

Es bestehen signifikante Zusammenhänge zu verschiedenen anderen Intelligenztests: Nonverbale Skalen des KFT 4-12+R ($r = .66$), PSB-R 4-6 ($r = .56$), CFT 20 ($r = .55$) und ZVT ($r = .33$). Die Korrelation der SPM mit dem ZVT spricht dafür, dass die Geschwindigkeit der Informationsverarbeitung auch eine Komponente der allgemeinen Intelligenz darstellt. Die Korrelationen der SPM mit der Schulnote in Deutsch beträgt $r = .20$, mit Mathematik $r = .40$. Der geringe Zusammenhang der SPM mit der Deutschnote lässt sich dadurch erklären, dass die Anwendung erworbenen Wissens zentraler Bestandteil der Leistungen in sprachlichen Fächern ist, während dies in Mathematik eine untergeordnete Rolle spielt.

Die SPM-Koeffizienten zur Vorhersage von Schulnoten variieren laut Raven et al. (1994) zwischen $r = .33$ und $r = .70$.

Autorinnenkommentar

Die SPM sind wie die APM populäre Verfahren zur Erfassung der fluiden Intelligenz und gelten als bekannteste Vertreter figuraler Matrizentests. Bezogen auf die Intelligenzdiagnostik ist zu beachten, dass die SPM wie auch die APM nur eine Intelligenzfacette – und dies zudem 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, Formann & Farkas (1991) stellten anhand von Analysen nach dem Rasch-Modell zudem fest, dass von insgesamt 60 Aufgaben 43 nicht eindimensional zu den anderen skaliert sind. Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den SPM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der SPM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe auch Preckel, 2010).

Literatur

- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998). *Matrizen-Test-Manual, Band 1. Ein Handbuch mit deutschen Normen zu den Standard Progressive Matrices von J. C. Raven*. Weinheim: Beltz.
- Kratzmeier, H. & Horn, R. (1979). *SPM. Raven-Matrizen-Test. Standard Progressive Matrices*. Weinheim: Beltz.
- Kubinger, K. D., Formann, A. K. & Farkas, M. G. (1991). *Psychometric shortcomings of Raven's Standard Progressive Matrices (SPM) in particular for computerized testing*. *European Review of Applied Psychology*, 41, 295-300.
- Preckel, F. (2010). *Intelligenztests in der Hochbegabungsdiagnostik*. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 19-44). Göttingen: Hogrefe.
- Raven, J. C. (1938). *Progressive Matrices*. London: Lewis.
- Raven, J. C., Court, J. H. & Raven, J. (1994). *Manual for Raven's Standard Progressive Matrices and Mill Hill Vocabulary Scales. Standard Progressive Matrices*. Oxford: Oxford Psychologists Press.

3.44 WIE; Wechsler Intelligenztest für Erwachsene

Aster, M. v., Neubauer, A. & Horn, R. (Hrsg., 2006). WIE. Wechsler Intelligenztest für Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler. Frankfurt/M.: Harcourt.

Beschreibung

Der WIE ist ein Intelligenztest für die diagnostische Einzelfalluntersuchung bei Jugendlichen und Erwachsenen. Gegenüber seinen Vorgängern besitzt der WIE einige Vorzüge, was seine praktische Handhabung angeht (z.B. detailliertere Durchführungs- und Auswertungshinweise im Manual, separates Kapitel mit Interpretationsvorschlägen). Da von den Testleiterinnen und -leitern auch mehr Entscheidungen zu treffen und mehr Informationen zu beachten sind als bei den Testvorgängern (Auswahl einzusetzender Untertests, Beachtung der Umkehrregel etc.), erfordert das Verfahren allerdings nach wie vor eine sehr umfassende und sorgfältige Einarbeitung.

Die Wechsler-Intelligenztests für Erwachsene und für Kinder sind die internationalen Klassiker der IQ-Testungen. Beim WIE handelt es sich um die deutschsprachige Version der Wechsler Adult Intelligence Scale WAIS-III (Wechsler, 1997). Der WIE wurde grundlegend revidiert und weist deutliche Unterschiede zu seiner Vorgängerversion (HAWIE-R; Tewes, 1991) auf. Verglichen mit dem HAWIE-R, betont der WIE weniger die Arbeitsgeschwindigkeit, sondern stärker die Fähigkeit zum abstrakten, sprachfreien, schlussfolgernden Denken. Weiters wurde die Zahl der Interpretationsebenen von drei auf vier erhöht: Während neben Gesamt-IQ sowie Verbal- und Handlungs-IQ ursprünglich nur die einzelnen Subtestergebnisse interpretiert wurden, werden im WIE ausgewählte Subtests aus dem Verbal- und Handlungsteil sowie drei neue Untertests verwendet, um vier Indizes zur Abbildung kognitiver Teilfunktionen zu ermitteln.

- Der Index *Sprachliches Verständnis* (SV) bildet im Vergleich zum Verbal-IQ stärker die rein sprachlichen Kompetenzen der Testperson ab.
- Der Index *Wahrnehmungsorganisation* (WO) soll das nonverbale, flüssige, schlussfolgernde Denken der Testperson erfassen. Verglichen mit dem Handlungs-IQ spielt die Schnelligkeit der Lösung eine geringere Rolle.
- Der Index *Arbeitsgedächtnis* (AGD) erfasst das verbale kurzfristige Merken von Informationen und deren Bearbeitung im Gedächtnis. Seine Entwicklung basiert auf der Konzeption, in der das Arbeitsgedächtnis als Anteil des informationsverarbeitenden Systems beschrieben wird, in der sowohl eine aktive Verarbeitung eingehender Informationen als auch die zeitlich begrenzte Speicherung von Informationen stattfindet (z.B. Baddeley, 2002; Baddeley & Hitch, 1994).
- Der Index *Arbeitsgeschwindigkeit* (AGS) beschreibt die Fähigkeit zur schnellen Verarbeitung visueller Information; er scheint besonders sensibel auf neuropsychologische Einflüsse zu reagieren (Aster et al., 2006, S. 121). Seine Einführung wird mit Befunden der ko-

gnitionspsychologischen und psychophysiologischen Forschung begründet. Diese deuten darauf hin, dass die Arbeitsgeschwindigkeit neben der Arbeitsgedächtniskapazität eine wesentliche Grundlage der menschlichen Intelligenz darstellen dürfte. Die Geschwindigkeit der Informationsverarbeitung besitzt ferner hohe Bedeutung für die Erklärung der ontogenetischen Intelligenzentwicklung (S. 18).

Insgesamt umfasst der WIE 14 Subtests, die teils mit, teils ohne Zeitbegrenzung zu bearbeiten sind:

- (1) Bilder ergänzen:
Gezeigt werden mehrfarbige Bildvorlagen mit Zeichnungen von Alltagsgegenständen. Die Testperson muss angeben, was auf dem Bild fehlt.
- (2) Wortschatz-Test:
Die Testperson muss die Bedeutung von mündlich vorgegebenen Wörtern erklären.
- (3) Zahlen-Symbol-Test:
Die Testperson muss unter Zeitbeschränkung unter jede Zahl das passende Symbol zeichnen, das oberhalb in einem Muster zugeordnet ist.
- (4) Gemeinsamkeiten finden:
Die Testperson muss angeben, welche Gemeinsamkeit zwei Begriffe aufweisen oder welches gemeinsame Konzept sie repräsentieren.
- (5) Mosaik-Test:
Die Testperson muss unter Zeitbeschränkung mit bis zu neun zweifarbigen Würfeln geometrische Muster nachlegen, die mit Hilfe von Würfeln vor ihr aufgebaut oder als Zeichnung vorgelegt werden.
- (6) Rechnerisches Denken:
Die Testperson muss mündlich vorgetragene Rechenaufgaben möglichst schnell im Kopf lösen und mündlich beantworten.
- (7) Matrizen-Test:
Die Testperson muss das Konstruktionsprinzip von Serien geometrischer Muster erkennen und die Serien durch Auswahl eines von fünf vorgegebenen Mustern ergänzen.
- (8) Zahlen nachsprechen:
Die Testperson muss Serien drei- bis neunstelliger Zahlen vorwärts sowie Serien zwei- bis achtstelliger Zahlen rückwärts nachsprechen.
- (9) Allgemeines Wissen:
Mit mündlich gestellten Fragen soll geprüft werden, ob der Testperson bestimmte Ereignisse, Orte, Sachverhalte und Persönlichkeiten bekannt sind.
- (10) Bilderordnen:
Die Testperson muss Serien mit Bildern, die eine kurze Geschichte bzw. einen Handlungsablauf wiedergeben, unter Zeitbeschränkung in die logisch richtige Reihenfolge bringen.
- (11) Allgemeines Verständnis:
Es werden Fragen gestellt, mit deren Beantwortung die Testperson zeigen soll, ob sie über praktisches Urteilsvermögen verfügt und ob sie konventionelle soziale

- Regeln und deren Bedeutung versteht.
- (12) Symbolsuche:
Die Testperson muss unter Zeitbeschränkung jeweils zwei Gruppen von abstrakten Formen und Symbolen daraufhin vergleichen, ob beide Gruppen ein gemeinsames Symbol enthalten.
- (13) Buchstaben-Zahlen-Folgen:
Der Testperson werden unterschiedlich lange Folgen von Zahlen und Buchstaben vorgelesen, die sie sich merken und anschließend wiedergeben muss.
- (14) Figuren legen:
Die Testperson muss Puzzleteile unter Zeitbeschränkung zu einer sinnvollen Figur zusammensetzen.

Altersgruppe: 16 bis 89 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt ca. 60–90 Minuten. Das Verfahren ist als **Einzeltest** durchzuführen.

Da das Verfahren komplexe Anforderungen an die Versuchsleiter/innen stellt, sollten diese über eine entsprechende psychodiagnostische Ausbildung in der Durchführung und Anwendung des WIE verfügen. Insbesondere müssen sie die Instruktionen und die Regeln zur Bewertung von Lösungen perfekt beherrschen und über hinreichende Sachkenntnis zur Interpretation der Ergebnisse verfügen (Aster et al., 2006, S. 15). In erster Linie kommen ausgebildete Psychologinnen/Psychologen und eingewiesene Ärztinnen/Ärzte als Testleiter/innen in Frage. Im Rahmen von Routineuntersuchungen oder Forschungsstudien können auch geschulte Studierende das Verfahren durchführen, diese sollten allerdings sachkundig supervidiert werden.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte im Zeitraum von 1999 bis 2005 an einer Stichprobe von $N = 1.790$ Probandinnen und Probanden. Für jede Altersgruppe wurden je 100 Versuchspersonen in Deutschland und je 20 Probandinnen/Probanden in Österreich und der deutschsprachigen Schweiz untersucht. Die Rohwerte wurden mit den Erwartungswerten der entsprechenden Altersgruppe in Wertpunkte transformiert und in weiterer Folge in IQ-Werte umgewandelt.



Testgüte

Reliabilität

Die Split-Half-Reliabilität der WIE-Subtests variiert zwischen $r = .70$ (Bilderordnen) und $r = .92$ (Matrizen).

Validität

Dem WIE wurden drei neue Subtests beigefügt und bestehende Subtests wurden mit neuen Items ergänzt. Die inhaltliche Validität bleibt gemäß Autorinnen und Autoren erhalten, denn die Neuerungen wurden mit Psychodiagnostikerinnen/-diagnostikern, Neuropsychologinnen/-psychologen und mit den Herausgeberinnen/Herausgebern von früheren Versionen diskutiert (zitiert nach Psyndex-Test Review).

Konfirmatorische und explorative Faktorenanalysen konnten die Vier-Faktorenstruktur des WIE bestätigen, jedoch nicht in allen Altersgruppen (Aster et al., 2006).

Autorinnenkommentar

Der WIE eignet sich hervorragend für die Diagnostik von Intelligenz im mittleren und höheren Bereich (bis zu 1,5 Standardabweichungen über dem Mittelwert).

Literatur

- Aster, M. v., Neubauer, A. & Horn, R. (Hrsg., 2006). *WIE. Wechsler Intelligenztest für Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler.* Frankfurt/M.: Harcourt.
- Baddeley, A. D. (2002). *Is working memory still working?* *European Psychologist*, 7, 85-97.
- Baddeley, A. D. & Hitch, G. (1994). *Working memory.* In G. H. Bower (Hrsg.), *The psychology of learning and motivation. Advances in research and theory* (Vol. 8, S. 47-90). San Diego, CA: Academic Press.
- Tewes, U. (1991). *Hamburg-Wechsler Intelligenztest für Erwachsene. Revision 1991. Handbuch und Testanweisung.* Stuttgart: Huber.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III.* San Antonio, TX: The Psychological Corporation.

3.45 WIT-2; Wilde-Intelligenz-Test 2

Kersting, M., Althoff, K. & Jäger, A. O. (2008). WIT-2. Wilde-Intelligenz-Test 2. Göttingen: Hogrefe.

Beschreibung

Mit dem WIT-2 kann die kognitive Leistungsfähigkeit insbesondere im Kontext der beruflichen Eignungsdiagnostik erhoben werden.

Der WIT-2 ist die überarbeitete Fassung des Wilde-Intelligenz-Tests (Jäger, 1963; Jäger & Althoff, 1983, 1994). Der WIT basiert auf der Intelligenzkonzeption von Thurstones Primary Mental Abilities (Thurstone, 1938). Das modifizierte Modell der Primary Mental Abilities (MMPMA) ergänzt Thurstones Modell um „(1) Facettenansatz, (2) Hierarchieannahme und (3) Kognitive Korrelate-Ansatz“ (Kersting et al., 2008, S. 23).

Gemäß dem MMPMA und den aktuellen Erkenntnissen der Intelligenzforschung wurden im WIT-2 acht Subtests entfernt und drei Subtests neu konstruiert.

Der WIT-2 erfasst nunmehr 11 Subtests, die folgenden Modulen bzw. Dimensionen zugeordnet werden können:

Modul 1: *Sprachliches Denken*

Die Fähigkeit, mit sprachlichen Konzepten umzugehen, wobei Wortschatz, Sprachverständnis und sprachlogisches Denken eine Rolle spielen.

Modul 2: *Rechnerisches Denken*

Die Fähigkeit, einfache Rechenoperationen der Addition, Subtraktion, Multiplikation und Division korrekt auszuführen.

Modul 3: *Räumliches Denken*

Die Fähigkeit zur Vorstellung räumlicher Relationen.

Modul 4: *Schlussfolgerndes Denken*

Die Fähigkeit, bestimmte logische Regeln und Gesetzmäßigkeiten zu erkennen und zweckentsprechend anwenden zu können.

Modul 5: *Merkfähigkeit*

Die Fähigkeit, sich kurz zuvor eingeprägte Informationen und Assoziationen zu merken und wieder zu erkennen.

Modul 6: *Arbeitseffizienz*

Die Leistung, komplexe Informationen eines vergleichsweise geringen Schwierigkeitsgrades mit andauernder Konzentration effizient zu verarbeiten und anzuwenden. „Effizienz“ wird dabei durch die gleichzeitige Optimierung von Arbeitsgeschwindigkeit und -genauigkeit bei der Selektion und Extraktion relevanter Informationen und deren anschließender Kombination erzielt.

Modul 7: *Wissen Wirtschaft*

Kenntnisse auf dem Gebiet Wirtschaft



Modul 8: *Wissen Informationstechnologie*
 Kenntnisse auf dem Gebiet Informationstechnologie

In den 11 Subtests (Kersting et al., 2008) sind von den Probandinnen und Probanden folgende Aufgaben zu lösen:

(1) *Analogien* (sprachliches Denken und schlussfolgerndes Denken):
 Auf der linken Seite eines Gleichheitszeichens sind zwei Wörter vorgegeben, die in einer bestimmten Beziehung zueinander stehen, auf der rechten Seite steht ein Wort. Von fünf Wahlwörtern ist dasjenige auszuwählen, das

auf der rechten Seite des Gleichheitszeichens eine analoge Beziehung herstellt.

- (2) *Abwicklungen* (räumliches Denken und schlussfolgerndes Denken):
 Zu einer aus mehreren Flächen mit verschiedenen Zeichen bestehenden Faltvorlage ist aus fünf vorgegebenen Körpern derjenige herauszufinden, der sich aus der Faltvorlage herstellen lässt.
- (3) *E-Mails Bearbeiten* (Arbeitseffizienz):
 Die Aufgabe simuliert die Bearbeitung des E-Mail-Posteingangs.
- (4) *Eingekleidete Rechenaufgaben* (rechnerisches Denken):
 Verbal eingekleidete Rechenaufgaben sind zu lösen.
- (5) *Grundrechnen* (Rechnerisches Denken):
 Rechenaufgaben zu den vier Grundrechenarten sind zu lösen.
- (6) *Gleiche Wortbedeutungen* (sprachliches Denken):
 Zu einem vorgegebenen Wort ist aus fünf anderen Wörtern das sinnähnlichste herauszusuchen.
- (7) *Information wiedererkennen* (Merkfähigkeit):
 In einem Text befinden sich verbal, numerisch und figural kodierte Informationen. Einer Einprägungszeit von 4 Minuten folgen für ca. 17 Minuten andere Subtests (Störphase). In dem anschließenden Reproduktionstest sollen Einzelheiten erinnert werden.
- (8) *Spiegelbilder* (räumliches Denken):
 Von fünf Strichfiguren lassen sich vier durch einfaches Verschieben in der Ebene, die fünfte dagegen nur durch Umklappen in der Vorstellung zur Deckung bringen; diese Figur ist herauszufinden.
- (9) *Wissen Informationstechnologie*
 Es werden Fragen zum Thema Informationstechnologie gestellt.
- (10) *Wissen Wirtschaft*
 Es werden Fragen zum Thema Wirtschaft gestellt.

(11) *Zahlenreihen* (schlussfolgerndes Denken):

Vorgegeben ist eine Folge von Zahlen, die nach einer Regel aufgebaut ist; auf dem Antwortbogen sind die Ziffern der Zahl durchzustreichen, die als nächstes Glied der Reihe folgen müsste.

Der WIT-2 wurde explizit für die Verwendung in der beruflichen Eignungsdiagnostik, Personalauswahl und Personalentwicklung konzipiert. Die acht Dimensionen decken ein breites Spektrum der interessierenden kognitiven Leistungsfähigkeiten ab, die bei der Beurteilung beruflicher Eignung relevant sein können. Einzelne Dimensionen können separat erhoben werden, so dass die Anwendung für gezielte Fragestellungen möglich ist.

Altersgruppe: Jugendliche ab 14 Jahren und Erwachsene

Hinweise zur Durchführung

Die Bearbeitungszeit beträgt 4 bzw. 3 Stunden (Lang- und Kurzform). Bei Verwendung nur einiger, ausgewählter Subtests für spezielle Fragestellungen ist das Verfahren entsprechend kürzer (z.B. Subtestauswahl für Hochschulabsolventinnen und -absolventen: ca. 50 Minuten). Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Die Testanweiser/innen sollten intellektuell und sozial kompetent, durchsetzungsfähig und verantwortungsbewusst sein (Kersting et al., 2008, S. 36).

Hinweise zu Auswertung und Interpretation

Die Größe der Normierungsstichprobe schwankt je nach Subtest zwischen $N = 2.234$ und $N = 1.024$ Personen. Der überwiegende Teil der Daten für den WIT-2 wurde im Kontext des Ernstfalls von beruflichen Bewerbungssituationen erhoben. Rohwerte werden in z-Werte transformiert, in den jeweiligen Bereichen gemittelt und in ein Profil eingetragen, ein Gesamtscore wird ermittelt.

Es bestehen bevölkerungsrepräsentative Altersnormen (Standardwerte) sowie Normen für spezielle Bildungsgruppen.

Testgüte

Reliabilität

Die Split-Half-Reliabilität beträgt $r = .98$; die Paralleltest-Reliabilität beträgt $r = .94$; die Retest-Reliabilität beträgt $r = .93$.

Validität

Konstruktvalidität: Strukturgleichungsmodelle belegten die Modellannahmen. Weiters bestätigte sich die Zuordnung der Subtests zu den Dimensionen und die Annahme übergeordneter

Faktoren in Form von fluider und kristalliner Intelligenz. Ein den Dimensionen übergeordneter Faktor Schlussfolgerndes Denken konnte ebenso bestätigt werden.

Die internationale Einordnung des WIT-2 erleichtert den Angaben der Autoren zufolge den Vergleich mit dem in den USA verbreiteten Wonderlic Intelligenztest (Wonderlic, 1992). Das schlussfolgernde Denken des WIT-2 korreliert mit diesem Verfahren mit $r = .66$.

Die Korrelationen der WIT-2-Dimensionen mit dem CFT 3 (Cattell & Weiss, 1971) betragen zwischen $r = .49$ (sprachliches Denken) und $r = .69$ (schlussfolgerndes Denken); die Korrelation des WIT-2 schlussfolgerndes Denken mit dem Leistungsprüfsystem (LPS; Horn, 1983) schlussfolgerndes Denken beträgt $r = .66$.

Kriteriumsvalidität: Fast alle Dimensionen des WIT-2 korrelieren mit den Abschlussnoten in Mathematik und Deutsch: Die Note in Mathematik korreliert zwischen $r = .14$ (Arbeitseffizienz) und $r = .44$ (räumliches Denken) mit dem WIT-2; die Note in Deutsch korreliert zwischen $r = .22$ (Merkfähigkeit) und $r = .46$ (sprachliches Denken) mit dem WIT-2.

Autorinnenkommentar

Der WIT-2 ist ein brauchbares Verfahren zur Messung intellektueller Fähigkeiten. Am besten misst er im Durchschnittsbereich bis zu 1,5 Standardabweichungen vom Mittelwert.

Literatur

Cattell, R. B. & Weiß, R. H. (1971). Grundintelligenztest Skala 3 (CFT-3). Göttingen: Hogrefe.

Horn, W. (1983). Leistungsprüfsystem (LPS). Handanweisung (2., erweit. und verbesserte Aufl.). Göttingen: Hogrefe.

Jäger, A. O. (1963). Der Wilde-Test. Ein neues Intelligenzdiagnostikum. Zeitschrift für experimentelle und angewandte Psychologie, 1. 260-278.

Jäger, A. O. & Althoff, K. (1983). Der Wilde-Intelligenz-Test (WIT). Ein Strukturdiagnostikum. Herausgegeben von der Deutschen Gesellschaft für Personalwesen e.V. Göttingen: Hogrefe.

Jäger, A. O. & Althoff, K. (1994). Der Wilde-Intelligenz-Test (WIT). Ein Strukturdiagnostikum. Herausgegeben von der Deutschen Gesellschaft für Personalwesen e.V. (2., revidierte Auflage). Göttingen: Hogrefe.

Kersting, M., Althoff, K. & Jäger, A. O. (2008). WIT-2. Wilde-Intelligenz-Test 2. Göttingen: Hogrefe.

Thurstone, L. L. (1938). Primary Mental Abilities. Chicago: University of Chicago Press.

Wonderlic, E. F. (1992). Wonderlic Personnel Test. Libertyville, IL: E.F. Wonderlic & Associates, Inc.

3.46 WMT; Wiener Matrizen-Test

Formann, A. K. & Piswanger, K. (1979). Wiener Matrizen-Test. Ein raschskaliertes sprachfreies Intelligenztest. Weinheim: Beltz.

Eine Computerversion wird im Rahmen des Wiener Testsystems von der Fa. Schuhfried angeboten. Im Rahmen des Hogrefe-Testsystems (HTS) liegt eine weitere Computerversion vor.

Beschreibung

Der Wiener Matrizen-Test WMT ist ein Verfahren zur sprachfreien Erfassung allgemeiner Intelligenz. Er ist einfach und ökonomisch anwendbar und geeignet für Schüler/innen und Erwachsene, bei denen die allgemeine intellektuelle Leistungsfähigkeit abgeklärt werden soll.

Die Testkonstruktion des WMT geschah in Anlehnung an die Standard Progressive Matrices von Raven (1938). Der WMT soll die Personeneigenschaft „Schlussfolgerndes Denken im Umgang mit abstrakten Symbolen“ messen. Da Matrizen tests allgemein hoch auf dem Spearman-Generalfaktor für Intelligenz laden, stellt den Autoren zufolge dieses Merkmal einen Indikator für das allgemeine Intelligenzniveau dar. Dem Test wird der Untersuchung von Piswanger (1977) zufolge Kulturunabhängigkeit bescheinigt.

Der WMT besteht aus 24 Aufgaben. Jedes Item besteht aus einer Reihe von abstrakten geometrischen Figuren, die nach bestimmten Prinzipien angeordnet sind. Aufgabe der Probandin/des Probanden ist es, aus jeweils acht Antwortvorgaben diejenige Lösung auszuwählen, die diese Figurenreihe logisch richtig fortsetzt.

Altersgruppe: 14 bis 40 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt ca. 30 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Er kann von jeder Person angewendet werden, die mit den Prinzipien psychologischer Diagnostik hinreichend vertraut ist.

Hinweise zu Auswertung und Interpretation

Die Normierung wurde an einer Stichprobe von $N = 2.248$ im Jahr 1979 an österreichischen Schülerinnen und Schülern im Alter zwischen 14 und 18 Jahren verschiedener Schultypen durchgeführt. Die Normen der Computerversion wurden 2002 publiziert. In den Normentabellen sind Prozentränge, Standardwerte und IQ-Werte ersichtlich.



Testgüte

Reliabilität

Durch die Verwendung des Rasch-Modells ist die innere Konsistenz des WMT gegeben. Die Split-Half-Reliabilität beträgt für die Papier-Bleistift-Version $r = .83$, für die Computerversion beträgt sie $r = .81$.

Validität

Die Ergebnisse bezüglich der Validität bestätigen die postulierte Eigenschaft des WMT, den Generalfaktor g von Spearman zu messen.

Die Korrelation der WMT mit dem I-S-T 70 beträgt $r = .85$, mit den Standard Progressive Matrices

(SPM) errechnet sich eine Korrelation von $r = .92$. Es ergibt sich eine geringe Korrelation des WMT mit der Mathematiknote ($r = .30$).

Autorinnenkommentar

Der Test besteht aus 24 figuralen Mehrfachwahl-Matrizenaufgaben mit jeweils acht Auswahlantworten. Die Distraktoren dieses Tests sind wenig effektiv: Bei mehr als 2/3 der Aufgaben – Mittring und Rost (2008) errechnen sogar 75% – kann mit einer recht einfachen Strategie (Abzählen von Matrixkomponenten, Teil- und Schnittmengenbildung) die jeweilige Lösung ohne Inspektion des Item-Stamms, also ohne das Finden der der Bezugs-Matrix zugrundeliegende(n) generellen Regel(n), gefunden werden. Es kann angenommen werden, dass die mangelnde Effizienz der Distraktoren die Konstruktvalidität dieses Tests besonders für die Gruppe der Hochbegabten deutlich beeinträchtigt.

Literatur

Formann, A. K. & Piswanger, K. (1979). *Wiener Matrizen-Test. Ein raschskaliertes sprachfreier Intelligenztest*. Weinheim: Beltz.

Formann, A. (1973). *Die Konstruktion eines neuen Matrizen-tests und die Untersuchung des Lösungsverhaltens mit Hilfe des linearen logistischen Testmodells*. Universität Wien: Phil. Dissertation.

Mittring, G. & Rost, D. (2008). *Die verflochtenen Distraktoren: Über den Nutzen einer theoretischen Distraktorenanalyse bei Matrizen-tests (für besser Begabte und Hochbegabte)*, *Diagnostica*, 54(4), 193–201.

Piswanger, K. (1977). *Interkulturelle Vergleiche mit dem Matrizen-test von FORMANN*. Universität Wien: Phil. Dissertation.

Raven, J. C. (1938). *Progressive Matrices*. London: Lewis.

3.47 WNV; Wechsler Nonverbal Scale of Ability

Wechsler, D. & Naglieri, J. A. (2006). Wechsler Nonverbal Scale of Ability (WNV). San Antonio, TX: Harcourt Assessment. (Anweisungen im Manual sind in deutscher Sprache vorhanden).

Beschreibung

Die WNV dient der nonverbalen Messung der allgemeinen kognitiven Fähigkeiten. Sie basiert auf einem schlüssigen und bewährten Testkonzept und stellt eine wertvolle Ergänzung der Serie von Wechsler-Tests um ein sprachfreies Verfahren dar.

Die Testbatterie gehört zur Serie der Wechsler-Intelligenztests (Wechsler Intelligenztest für Erwachsene [WAIS-III; Wechsler, 1997], Wechsler Intelligenztest für Kinder [WISC-IV; Wechsler, 2004] und Wechsler Intelligenztest für das Vorschulalter [WPPSI-III; Wechsler, 2002]). Im WNV werden jedoch, im Gegensatz zu den anderen Wechsler-Tests, nur nonverbale Aufgaben verwendet. Die WNV ist insbesondere zur Untersuchung von Personen geeignet, deren Muttersprache nicht Englisch bzw. nicht Deutsch ist oder die hohe Begabungen bei gleichzeitigen sprachlichen Defiziten aufweisen.

Intelligenz wird im Sinne von Wechsler als eine generelle kognitive Fähigkeit beschrieben, die sich in unterschiedlichen Formen manifestiert, weshalb sie mit Hilfe verschiedener Zusammenstellungen von Untertests abgebildet werden soll. Bei der Testbearbeitung der WNV werden lediglich minimale rezeptive Sprachfähigkeiten benötigt und der Einfluss expressiver Sprachkompetenz und rechnerischer Fertigkeiten auf die Testleistung und damit auf die Schätzung der Allgemeinen Intelligenz wird eliminiert (Wechsler & Naglieri, 2006, S. 1).

Die WNV besteht aus sechs Subtests mit teilweiser Zeitbegrenzung:

- (1) *Matrizen*
Die Testperson muss ein fehlendes Muster auf Grund von Gesetzmäßigkeiten ergänzen.
- (2) *Coding*
Die Testperson muss nach einer vorgegebenen Zahlen- oder Form-Symbol-Zuordnung aus einer Auswahl die dazu gehörigen Symbole möglichst schnell und fehlerfrei notieren.
- (3) *Object Assembly*
Die Testperson muss zwei bis acht Puzzleteile unter Zeitbegrenzung zu einer Figur zusammensetzen.
- (4) *Recognition*
Die Testperson muss ein geometrisches Muster wieder erkennen.
- (5) *Spatial Span*
Die Testperson muss farbige Blöcke in der gleichen Reihenfolge sowie in umgekehr-

ter Reihenfolge berühren, wie es unmittelbar zuvor die Testleiterin/der Testleiter getan hat.

(6) *Picture Arrangement*

Die Testperson muss Serien von Bildern in die logisch richtige Reihenfolge bringen.

Die meisten der in der WNV zusammengestellten Untertests haben sich bereits in anderen Wechsler-Tests als Maß des generellen Fähigkeitskonzepts bewährt und können daher zu einem entsprechenden Indikator zusammengefasst werden. Neben der allgemeinen Intelligenz erfassen sie jedoch auch spezifische Fähigkeitsaspekte, z.B. perzeptuelles Schlussfolgern, Wahrnehmungsorganisation, graphomotorische Geschwindigkeit oder visuell-räumliches Gedächtnis und können daher einzeln interpretiert werden. Neuartig an der WNV sind auch die bildlichen Instruktionen, die den Testpersonen vorgelegt und ggf. durch einfache verbale Erläuterungen unterstützt werden.

Altersgruppe: 4 bis ca. 22 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt abhängig vom Alter bei der Kurzversion ca. 20 Minuten, bei der Langversion ca. 45 Minuten.

Es handelt sich um einen **Einzeltest**. Der Startpunkt eines Subtests variiert abhängig vom Alter der Testperson. Abbruchregeln legen fest, dass ein Subtest nach einer bestimmten Zahl unbearbeiteter Aufgaben beendet wird. Das Verfahren sollte nur von Psychologinnen und Psychologen durchgeführt werden. Studierende oder Hilfskräfte können das Verfahren allenfalls unter Supervision durchführen, die Interpretation sollte aber fertigen ausgebildeten Fachleuten vorbehalten bleiben (Wechsler & Naglieri, 2006). Da das Testmanual bislang nur in englischer Sprache vorliegt, werden sehr gute englische Sprachkenntnisse vorausgesetzt, um eine adäquate Durchführung und Interpretation zu gewährleisten. Die Instruktionen für die Testpersonen sind in deutscher Sprache vorhanden.

Hinweise zu Auswertung und Interpretation

Die Normierung der WNV erfolgte mit repräsentativen US-amerikanischen ($N = 1.323$) und kanadischen Stichproben ($N = 875$; Wechsler & Naglieri, 2006). Normtabellen mit altersnormierten T-Werten, Prozenträngen mit Konfidenzintervallen für Subtests und Gesamtwert sind vorhanden. Relative Stärken und Schwächen der Testperson können ermittelt werden.

Testgüte

Reliabilität

Es wurde eine hohe interne Konsistenz für den WNV-Gesamtscore über die Altersgruppen berechnet ($\alpha = .91$); die Retest-Reliabilität nach 10–52 Tagen beträgt bei der Langversion zwi-

schen $r = .74$ (4- bis 7-Jährige) und $r = .85$ (13- bis 21-Jährige), bei der Kurzversion zwischen $r = .64$ (4- bis 7-Jährige) und $r = .76$ (13- bis 21-Jährige).

Validität

Konfirmatorische Faktorenanalysen konnten das Einfaktorenmodell nur teilweise belegen, weitere Studien sind für den Nachweis der faktoriellen Validität der WNV erforderlich.

Es besteht eine Korrelation zwischen WNV-Langversion und dem WPPSI-III (Wechsler, 2002) von $r = .71$. Mit dem NNAT-I (Naglieri Nonverbal Ability Test; Naglieri, 2003) errechnet sich eine Korrelation von $r = .73$.

Autorinnenkommentar

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 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.

Literatur

- Naglieri, J. A. (2003). Naglieri Nonverbal Ability Test – individual administration. San Antonio, TX: Harcourt Assessment.*
- Wechsler, D. (1997). Wechsler Adult Intelligence Scale – III. San Antonio, TX: Harcourt Assessment.*
- Wechsler, D. (2002). Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI-III). San Antonio, TX: The Psychological Corporation.*
- Wechsler, D. (2004). Wechsler Intelligence Scale for Children – Fourth Edition: Canadian. Toronto, Ontario: Harcourt Assessment.*
- Wechsler, D. (2005). Wechsler Intelligence Scale for Children – Fourth Edition, Spanish. San Antonio, TX: Harcourt Assessment.*
- Wechsler, D. & Naglieri, J. A. (2006). Wechsler Nonverbal Scale of Ability (WNV). Technical and interpretative manual. San Antonio, TX: Harcourt Assessment.*

3.48 WPPSI-III; Wechsler Preschool and Primary Scale of Intelligence

Petermann, F. & Lipsius, M. (2009). WPPSI-III. Wechsler Preschool and Primary Scale of Intelligence – Third Edition. Deutsche Version. Frankfurt/M.: Pearson.

Beschreibung

Die WPPSI dient zur Erfassung allgemeiner und spezifischer intellektueller Fähigkeiten von Kindern im Vorschulalter.

Die WPPSI-III ist die deutschsprachige Adaptation der Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III; Wechsler, 2002). Die Altersspanne gegenüber der bisherigen Version wurde bei der WPPSI-III nach oben verschoben. Mit der Neuauflage der WPPSI-III steht jetzt ein anwenderfreundliches Manual zur Verfügung, in dem sowohl die Angaben zur Testentwicklung und Interpretation als auch die Informationen zur Durchführung und Auswertung enthalten sind. Mit der WPPSI-III kann ein Gesamt-IQ als Maß für den kognitiven Entwicklungsstand eines Kindes im Vorschulalter bestimmt werden. Zusätzlich können vier weitere übergeordnete Werte berechnet werden: Verbal- und Handlungsteil, Verarbeitungsgeschwindigkeit und Allgemeine Sprachskala.

Die Testbatterie besteht aus 14 Untertests:

Mosaik-Test

Das Kind soll innerhalb einer vorgegebenen Zeitspanne eine Serie von vorgegebenen geometrischen Mustern mit Hilfe von ein- oder zweifarbigen Würfeln nachbauen.

Allgemeines Wissen

Das Kind soll eine Serie mündlich gestellter Fragen über allgemein bekannte Ereignisse, Sachverhalte, Orte und Persönlichkeiten beantworten.

Matrizen-Test

Das Kind betrachtet eine unvollständige Matrix und soll das fehlende Teil aus vier bzw. fünf Antwortmöglichkeiten auswählen.

Wortschatz-Test

Das Kind soll eine Serie von mündlich vorgegebenen Wörtern definieren.

Bildkonzepte

Das Kind soll aus zwei Bildreihen jeweils ein Bild auswählen, um daraus eine Gruppe mit einer gemeinsamen Eigenschaft zu bilden.

Symbol-Suche

Das Kind vergleicht in einer begrenzten Zeit eine Gruppe von abstrakten Symbolen mit einem Zielsymbol und soll angeben, ob sich das Zielsymbol in der Suchgruppe befindet.

Begriffe erkennen

Das Kind entschlüsselt einen allgemeinen Begriff, der von der Testleiterin/vom Testleiter umschrieben wird.

Symbole kodieren

Einer Serie einfacher geometrischer Figuren soll das Kind abstrakte Symbole zuordnen. Es zeichnet in einer begrenzten Zeit die Symbole in die dazugehörigen Felder, indem es einen vorgegebenen Schlüssel verwendet.

Allgemeines Verständnis

Das Kind soll eine Serie von mündlichen Fragen beantworten, die das Verständnis von allgemeinen Prinzipien und sozialen Situationen oder Regeln erfordern.

Bilder ergänzen

Das Kind sieht eine Reihe von Bildern und soll das wichtigste Detail benennen (bzw. darauf zeigen), das auf dem jeweiligen Bild fehlt.

Gemeinsamkeiten finden

Das Kind soll das Gemeinsame von mündlich vorgegebenen Begriffspaaren benennen oder beschreiben. Die Begriffe beziehen sich auf Konzepte oder Gegenstände des Alltags.

Passiver Wortschatz

Das Kind sieht sich vier Bilder an und soll auf das Bild zeigen, das die Testleiterin/der Testleiter benennt.

Figuren legen

Das Kind soll eine sinnvolle Figur aus Puzzle-Teilen innerhalb einer festgelegten Zeit zusammensetzen.

Aktiver Wortschatz

Das Kind soll Bilder benennen, die ihm zuvor in einem Buch gezeigt wurden.

Altersgruppe: 3;0 bis 7;2 Jahre

Hinweise zur Durchführung

Die Bearbeitungsdauer beträgt für Kinder zwischen 3;0–3;11 Jahren 20 bis 40 Minuten, für Kinder zwischen 4;0–7;2 Jahren 40 bis 50 Minuten für die Kerntests.

Das Verfahren ist als **Einzeltest** konzipiert. Es werden keine Angaben zum beruflichen Hintergrund und der Qualifikation der Testleiter/innen gemacht.



Hinweise zu Auswertung und Interpretation

Die Normen wurden 2009 in Deutschland an N = 710 Kindern erhoben.

Testgüte

Reliabilität

Die Reliabilität der Untertests variiert zwischen $\alpha = .77$ und $\alpha = .88$ und auf Indexebene zwischen $\alpha = .87$ und $\alpha = .92$. Für den Gesamttest beträgt sie $\alpha = .95$.

Validität

Zum Nachweis der Validität liegen Interkorrelationsstudien, faktorenanalytische Studien, Korrelationsstudien mit anderen Messinstrumenten wie dem HAWIK-IV sowie klinische Validierungsstudien vor.

Autorinnenkommentar

Die WPPSI-III 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.

Literatur

- Petermann, F. & Lipsius, M. (2009). *WPPSI-III. Wechsler Preschool and Primary Scale of Intelligence – Third Edition. Deutsche Version. Frankfurt/M.: Pearson.*
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI-III). San Antonio, TX: The Psychological Corporation.*

3.49 WST A-P; Wortschatztest – aktiv und passiv

Ibrahimovic, N. & Bulheller, S. (2005). Wortschatztest – aktiv und passiv. Deutsche Version des Mill Hill Vocabulary Tests. Frankfurt/M.: Harcourt.

Beschreibung

Der WST A-P erfasst den aktiven und passiven Wortschatzumfang von Jugendlichen und Erwachsenen. Der Grundgedanke des WST A-P, den Wortschatz durch ein Verfahren zur aktiven Sprachproduktion und zum Verständnis von Wortbedeutungen zu erfassen, ist sehr attraktiv. Die praktische Umsetzung dieses Gedankens gilt es jedoch noch in vieler Hinsicht zu optimieren, bevor der Test für eine praktische Verwendung hilfreich ist.

Bei dem Verfahren handelt es sich um die deutsche Fassung der amerikanischen Mill Hill Vocabulary Scale (Raven, Raven & Court, 1998). Ausgangspunkt der Testkonstruktion war eine Stichprobe von 194 randomisiert aus dem Duden (1996) ausgewählten Begriffen, die nach mehreren Analysen zur vorliegenden Endform führten. Der Test erlaubt eine zuverlässige Erfassung des aktiven und passiven Wortschatzumfangs. Im Gegensatz zu vielen anderen Wortschatztests, die sich auf die Erfassung des passiven Wortschatzes beschränken, indem sie eine „Wiedererkennungsleistung“ (Wörter der deutschen Sprache von zufälligen Anordnungen von Buchstaben zu unterscheiden) oder eine Zuordnung von Begriffen zu Abbildungen verlangen, besteht mit diesem Test auch die Möglichkeit, den aktiven Wortschatz durch „Definieren“ von Begriffen zu erfassen.

Der Test besteht aus insgesamt 60 Aufgaben, die in zwei Testteile zu je 30 Items aufgeschlüsselt sind. Im ersten Testteil – aktiver Wortschatz – müssen 30 Begriffe schriftlich definiert werden. Im zweiten Testteil – passiver Wortschatz – besteht die Aufgabe darin, dass den 30 Items aus vorgegebenen Beschreibungen die richtigen Definitionen zugeordnet werden müssen. Der Test liegt in zwei Testformen, Form A und Form B, vor (keine Parallel-Testformen).

Altersgruppe: 14 bis 60+ Jahre

Hinweise zur Durchführung

Die Testdauer beträgt im Schnitt 23 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Es werden keine Angaben hinsichtlich der erforderlichen diagnostischen Kenntnisse der Testleiter/innen gemacht.

Hinweise zu Auswertung und Interpretation

Beide Testformen wurden in den Jahren 2000 bis 2004 deutschlandweit normiert. Für Form A (N = 2.053) existieren Prozentränge für Jugendliche (14 bis 19 Jahre) getrennt nach Realschule

und Gymnasium und Prozenträge für Erwachsene (20 bis 60+ Jahre). Für Form B (N = 873) werden Prozenträge für Jugendliche (14 bis 19 Jahre) getrennt nach Realschule und Gymnasium angeboten.

Testgüte

Reliabilität

Die Split-Half-Reliabilität beträgt für Testform A $r = .93$, für Testform B $r = .91$.

Validität

Die inhaltliche Validität ist laut den Autoren durch die repräsentative Zufallsauswahl von Wörtern aus dem Duden (1996) belegt. Moderate Zusammenhänge konnten mit der Intelligenz-Struktur-Analyse (ISA) nachgewiesen werden.

Autorinnenkommentar

Der WST ist für die Intelligenzdiagnostik im wissenschaftlichen Sinne ungeeignet (erfasst nur Wortschatz). Er kann lediglich im Rahmen eines Screenings, ob und inwieweit es Probleme mit der Sprache gibt, verwendet werden.

Literatur

Duden (1996). Die deutsche Rechtschreibung. Mannheim: Duden-Verlag.

Ibrahimovic, N. & Bulheller, S. (2005). Wortschatztest – aktiv und passiv. Deutsche Version des Mill Hill Vocabulary Tests. Frankfurt/M.: Harcourt.

Raven, J., Raven, J. C. & Court, J. H. (1998). Mill Hill Vocabulary Scale (1998 edition). Oxford: Psychologist Press.

3.50 ZVT; Zahlen-Verbindungs-Test

Oswald, W. D. & Roth, E. (1978). Der Zahlen-Verbindungs-Test (ZVT). Ein sprachfreier Intelligenz-Schnell-Test. Göttingen: Hogrefe.

Beschreibung

Der ZVT dient zur Erfassung der kognitiven Leistungsgeschwindigkeit. Er ist ein zeitloser Test, der für eine sehr große Altersgruppe anwendbar ist. Er ist praktisch und einfach in der Durchführung und bei den Probandinnen/Probanden beliebt.

Im ZVT geht es darum, die nach dem Zufallsprinzip angeordneten Zahlen 1 bis 90 so schnell wie möglich miteinander zu verbinden. Die jeweils aufzusuchende nächste Ziffer befindet sich in unmittelbarer Nachbarschaft der zuletzt bearbeiteten Ziffer und ist durch eine gerade Verbindungslinie erreichbar.

Der ZVT kann im Rahmen der Schulberatung eingesetzt werden, insbesondere eignet er sich für Vergleiche zwischen traditionellen Intelligenztests und „kognitiver“ Leistungsgeschwindigkeit.

Altersgruppe: 8 bis ca. 95 Jahre

Hinweise zur Durchführung

Die Testdauer beträgt lediglich 5–10 Minuten. Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Spezielle Durchführungsvoraussetzungen liegen nicht vor.

Hinweise zu Auswertung und Interpretation

Die deutsche Normierungsstichprobe besteht aus $N = 2.109$ Probandinnen und Probanden im Alter zwischen sechs und acht Jahren. Da die Daten in den 1970er Jahren erhoben wurden, sind die Normwerte aufgrund des Flynn-Effekts mit Vorsicht zu interpretieren. Es liegen T-Werte, Prozentränge und IQ-Werte vor.

Testgüte

Reliabilität

Für den Einzeltest liegt die Retest-Reliabilität über verschiedene Stichproben gemessen zwischen $r = .84$ und $r = .97$. Sie ist weitgehend unabhängig von Alter und Bildungsstand der Versuchsperson. Für den Gruppentest wurde eine Retest-Reliabilität nach 6 Wochen von $r = .81$ errechnet.

Validität

Faktorenanalysen ergaben hohe Ladungen des ZVT auf einem Faktor, der mit Bearbeitungsgeschwindigkeit umschrieben werden kann. Signifikante Restladungen sind zusätzlich auf dem Reasoning-Faktor vorhanden. Die Korrelationen zwischen dem ZVT und dem PSB betragen zwischen $r = .69$ und $r = .82$ (abhängig von der Schulart); mit dem IST ergibt sich eine Korrelation von $r = .56$ (Studierende) bzw. $r = .50$



(Hauptschüler/innen); mit den SPM lässt sich eine Korrelation von $r = .40$ errechnen, mit dem CFT 3 ergibt sich eine Korrelation von $r = .53$.

Ein geringer Zusammenhang besteht zwischen dem ZVT und Schulnoten: $r = .27$ in den sprachlichen Fächern (Hauptschüler/innen) und $r = .36$ mit mathematisch-naturwissenschaftlichen Fächern (Hauptschüler/innen). Extremgruppenvergleiche erbrachten signifikante Unterschiede zwischen Sonderschülerinnen/-schülern und Gymnasiastinnen/Gymnasiasten sowie zwischen in Heimen untergebrachten Kindern und Kindern, die zu Hause wohnen (Kontrollgruppen). Zwischen eineiigen Zwillingen ergaben sich keine signifikanten Unterschiede.

Autorinnenkommentar

Der ZVT kann als brauchbares Verfahren zur Abschätzung eines besonderen Aspekts der intellektuellen Leistungsfähigkeit (Verarbeitungsgeschwindigkeit) nur im Sinne eines Screenings empfohlen werden. Zur Begabungsdiagnostik scheint der ZVT ungeeignet. Abgesehen davon, dass der ZVT nur eine Intelligenzfacette – mit nur einem Aufgabentyp – erfasst, reicht er auch nicht aus, um eine verlässliche Intelligenzdiagnose stellen zu können. Es sollten unbedingt Tests mit unterschiedlichen Aufgabenfacetten dem ZVT vorgezogen werden bzw. der ZVT nur als Zusatzinstrument zur Erhebung der Verarbeitungsgeschwindigkeit eingesetzt werden. Die Normen bedürfen einer Überprüfung, obgleich der Verdacht einer Normveränderung weniger begründet ist als etwa bei wissens- oder problemlösebezogenen Leistungstests.

Literatur

Oswald, W. D. & Roth, E. (1978). Der Zahlen-Verbindungs-Test (ZVT). Göttingen: Hogrefe.

4 Spezielle Aspekte in der Hochbegabungsdiagnostik



Traditionellen Leistungs-, v.a. Intelligenztests mangelt es beim Einsatz in der Hochbegabungsdiagnostik oft an ausreichend schwierigen Items. In der entsprechenden Zielpopulation kommt es demnach oft zu Deckeneffekten und in der Folge zu wenig messgenauen Testwerten (vgl. z.B. Preckel, 2003). Der damit verbundene Umstand, dass ziemlich viele Items für die Testperson (zu) leicht sind, also für hochintelligente Personen wenig Herausforderung bedeuten, führt relativ häufig auch zu reduzierter Leistungsmotivation.

Deckeneffekte und auch verminderte Leistungsmotivation können verhindert werden, indem der richtige Test ausgewählt wird bzw. den Testpersonen nur Items vorgegeben werden, die ihrem Leistungsniveau entsprechen („adaptives Testen“). Alle anderen Items sind für die jeweils betroffene Personengruppe ohnehin nicht informativ, da der Ausgang ihrer Bearbeitung von vornherein bekannt, jedenfalls höchst wahrscheinlich ist. Ob sie vorgegeben werden oder nicht, ist weitgehend unerheblich. Es ist vorzuziehen, an ihrer Stelle andere, informativere Items einzusetzen, die in fein abgestufte Grade der gemessenen Fähigkeit zu differenzieren vermögen; jedes Testergebnis bedeutet so gesehen einen hohen Informationsgewinn.

Die genannten Möglichkeiten adaptiven Testens in der Intelligenzdiagnostik werden derzeit nicht annähernd praktisch genutzt, nur der AID 2 bietet die Möglichkeit standardisiert an (vgl. Kubinger, 2003, 2006, 2009; Kubinger & Holoche-Ertl, 2010). Der SON-R enthält diese Möglichkeit ebenfalls, eignet sich aber nicht für die Diagnostik von hoher Intelligenz.

Mittels automatischer Itemgenerierung (regelgeleitete, computergestützte Erzeugung von Items auf der Basis definierter Algorithmen) könnten entsprechende Items relativ einfach produziert und vorgegeben werden. Damit können im Prinzip beliebig viele Items mit einem bestimmten, vorab festgelegten Schwierigkeitsgrad konstruiert werden.

Weiters sprechen technologische Gründe, wie z.B. Kosten und Testsicherheit, als auch kognitionstheoretische Argumente für diese Form der Itemgenerierung. Ansätze eines computergestützten Itemgenerators finden sich zum Beispiel bei Arendasy & Gittler (2003) oder bei Holling, Bertling, Zeuch & Kuhn (2010). Auch für Wiederholungstestungen sind umfangreiche Itempools, wie sie durch die automatische Generierung erstellt werden könnten, von Vorteil.

Literatur

- Arendasy, M. & Gittler, G. (2003). IRT-basierter Vergleich zweier Varianten automationsgestützt erstellter Matrizenaufgaben. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 261-275.
- Holling, H., Bertling, J., Zeuch, N. & Kuhn, J. T. (2010). Automatische Itemgenerierung: Grundlagen und Darstellung automatisch generierter Items anhand Lateinischer Quadrate. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 211-232). Göttingen: Hogrefe.

- Kubinger, K. D. & Holoher-Ertl, S. (2010). Die Vorteile adaptiven Testens in der Hochbegabungsdiagnostik. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 197-210). Göttingen: Hogrefe.
- Kubinger, K. D. (2003). *Adaptives Testen*. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 1-9). Weinheim: Beltz/PVU.
- Kubinger, K. D. (2006). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Kubinger, K. D. (2009). *Adaptives Intelligenz Diagnostikum – Version 2.2 (AID 2) samt AID 2 Türkisch*. Göttingen: Beltz.
- Preckel, F. (2003). *Diagnostik intellektueller Hochbegabung. Testentwicklung zur Erfassung der fluiden Intelligenz*. Göttingen: Hogrefe.



Gesamt-Literaturverzeichnis



- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. Auflage). Stuttgart: Kohlhammer.
- Amthauer, R. (1955). *Intelligenz-Struktur-Test (I-S-T)*. Göttingen: Hogrefe.
- Amthauer, R. (1973). *Intelligenz-Struktur-Test (I-S-T 70)*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (1999). *Intelligenz-Struktur-Test 2000 (I-S-T 2000)*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D. & Beauducel, A. (2001). *I-S-T 2000 R. Intelligenz-Struktur-Test 2000 R*. Göttingen: Hogrefe.
- Arendasy, M. & Gittler, G. (2003). IRT-basierter Vergleich zweier Varianten automationsgestützt erstellter Matrizenaufgaben. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 261-275.
- Aster, M. v., Neubauer, A. & Horn, R. (Hrsg., 2006). *WIE. Wechsler Intelligenztest für Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler*. Frankfurt/M.: Harcourt.
- Baddeley, A. D. & Hitch, G. (1994). Working memory. In G. H. Bower (Hrsg.), *The psychology of learning and motivation. Advances in research and theory* (Vol. 8, S. 47-90). San Diego, CA: Academic Press.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7, 85-97.
- Barth, N. (1989). *Frankfurter Analogietest 6-9. Test des logisch-schlussfolgernden Denkens für 6. bis 9. Klassen [Testmappe mit Handanweisung, je einem Musteraufgabenheft Form A und B und einem Musterantwortbogen Form A und B]*. Weinheim: Beltz.
- Barth, N. (1996). Zensuren und logisches Denken – eine empirische Untersuchung zur Übereinstimmungsgültigkeit zwischen 17 Schulzensuren und den Leistungen im Frankfurter Analogietest. *Zeitschrift für internationale erziehungs- und sozialwissenschaftliche Forschung* 13(1), 139-155.
- Belser, H., Anger, H., Bargmann, R. (1965). *Frankfurter Analogietest. FAT 7-8. Begabungstest für 7. und 8. Klassen (2. Auflage) [Testmappe mit „Beiheft mit Anleitung und Normentabellen“, Testheften Form A und Form B sowie Klassenliste]*. Weinheim: Beltz.
- Belser, H., Anger, H., Bergmann, R. & Raatz, U., (1973). *Frankfurter Analogietest FAT 4-6. Begabungstest für 4. bis 6. Klassen (2. Auflage, Bearbeitung und Normierung 1972. 1. Auflage 1965. Testmappe mit „Beiheft mit Anleitung und Normentabellen“, Testheften Form A und Form B sowie Klassenliste)*. Weinheim: Beltz.
- Bennett, G. K. (1969). *Mechanical Comprehension Test*. New York: Psychological Corporation.
- Berg, C. A. (2000). Intellectual development in adulthood. In R. J. Sternberg (Hrsg.), *Handbook of intelligence* (S. 117-137). Cambridge: Cambridge University Press.
- Berg, M. & Schaarschmidt, U. (1989). *Diagnosticum für bildlich angeregte kognitive Leistung – BILKOG. Intelligenztest für das Kindergarten- und frühe Schulalter*. Berlin: Psychodiagnostisches Zentrum der Humboldt-Universität.
- Binet, A. & Simon, T. (1905). *Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux*. *L'Annee Psychologique*, 11, 191-244.
- Brunner, M. & Süß, H. M. (2007). Wie genau können kognitive Fähigkeiten gemessen werden? Die Unterscheidung von Gesamt- und Konstruktreliabilitäten in der Intelligenzdiagnostik für den Berliner Intelligenzstrukturtest. *Diagnostica*, 53, Heft 4, 184-193.
- Bulheller, S. & Häcker, H. O. (2003). *Peabody Picture Vocabulary Test. PPVT. Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene*. Frankfurt/M.: Swets.

- Burgemeister, B. B., Blum, C. H. & Lorge, I. (1954). *Columbia Mental Maturity Scale*. New York: Harcourt, Brace and World.
- Carlson, J., Goldman, R., Bollinger, J. & Wiedl, K. H. (1974). *Der Effekt von Problemverbalisation bei verschiedenen Aufgabengruppen und Darbietungsformen des „Raven Progressive Matrices Test“*. *Diagnostica*, 20, 133-141.
- Carpenter, P. A., Just, M. A. & Shell, P. (1990). *What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test*. *Psychological Review*, 97(3), 404-431.
- Carroll, J. B. (1991). *No demonstration that g is not unitary, but there is more to the story: Comment on Kranzler and Jensen*. *Intelligence*, 15, 423-436.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Cattell, R. B. & Weiß, R. H. (1971). *Grundintelligenztest Skala 3 (CFT-3)*. Göttingen: Hogrefe.
- Cattell, R. B. (1954). *Culture Fair Intelligence Test, Scales 1, 2 and 3*. Champaign, IL: IPAT.
- Cattell, R. B. (1960). *Culture Fair Intelligence Test, Scale 2*. Champaign, IL: IPAT.
- Cattell, R. B. (1966). *Culture Fair (or Free) Intelligence Test, Scale 1 (Manual)*. Champaign, Ill.: IPAT.
- Cattell, R. B. (1971). *Abilities: their structure, growth and action*. Boston: Houghton Mifflin.
- Cattell, R. B. (1987). *Intelligence: its structure, growth, and action*. Amsterdam: Elsevier.
- Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest CFT 1. Skala 1*. (5. revidierte Auflage). Göttingen: Hogrefe.
- Colom, R., Escorial, S. & Rebollo, I. (2004). *Sex differences on the progressive matrices are influenced by sex differences on spatial ability*. *Personality and Individual Differences*, 37(6), 1289-1293.
- Conrad, W., Büscher, P., Hornke, L., Jäger, R., Schweizer, H., Stünzner, W. v. & Wiencke, W. (1986). *Mannheimer Intelligenztest*. MIT. Weinheim: Beltz.
- Daniels, J. C. (1962). *Figure Reasoning Test. A nonverbal intelligence test*. London: Crosby Lockwood & Son Ltd.
- Daniels, J. C. (2004). *Figure Reasoning Test*. Frankfurt/M.: Swets.
- Das, J. P. & Jarman, R. (1991). *Cognitive integration. Alternative model for intelligence*. In H. A. H. Rowe (Hrsg.), *Intelligence: Reconceptualization and Measurement*. Hillsdale, NJ: Lawrence Erlbaum & Associates.
- Daseking, M., Petermann, F. & Waldmann, H.-C. (2008). *Der allgemeine Fähigkeitsindex (AFI) – eine Alternative zum Gesamt-Intelligenzquotienten (G-IQ) des HAWIK-IV?* *Diagnostica*, 54(4), 211-22.
- Daseking, M., Petermann, U. & Petermann, F. (2007). *Intelligenzdiagnostik mit dem HAWIK-IV. Kindheit und Entwicklung*, 16(4), 250-259.
- Davidson, J. E. & Sternberg, R. J. (1984). *The role of insight in intellectual giftedness*. *Gifted Child Quarterly*, 28(2), 58-64.
- DeShon, R. P., Chan, D. & Weissbein, D. A. (1995). *Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants*. *Intelligence*, 21(2), 135-155.
- Deutsches Institut für Internationale Pädagogische Forschung (Hrsg., 1994). *BT 1-2. Bildertest 1-2 (unveränderter Nachdruck der 2., überarbeiteten Auflage) [Testmappe mit Beiheft, Mustertestheft, Auswertungsschlüssel und Klassenliste]*. Weinheim: Beltz.

- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Dörner, D. (1986). *Diagnostik der operativen Intelligenz*. *Diagnostica*, 32, 209-308.
- Duden (1996). *Die deutsche Rechtschreibung*. Mannheim: Duden-Verlag.
- Dunn, L. M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised (PPVT-R)*. Manual for forms L and M. Bloomington, MN: Pearson Assessment.
- Dunn, L. M. & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III)*. Circle Pines, MN: American Guidance Service.
- Eggert, D. & Schuck, K. D. (1975). *Ein empirischer Beitrag zum Zusammenhang von Intelligenz, Motorik und Sozialstatus im Vorschulalter*. *Sonderpädagogik*, 1(5), 57-62 und 121-135.
- Emmet, W. G. (1956). *Deeside Picture Test for Seven-Years-Old*. London: Harrap.
- Eysenck, H. J. (1979). *The structure and measure of intelligence*. Berlin: Springer.
- Fay, E. (1997). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18(1/2), 36-38.
- Feldhusen, J. F. & Jarvan, F. A. (2000). *Identification of gifted and talented youth for educational programmes*. In K. A. Heller, F. J. Mönks, A. H. Passow & R. F. Subotnik (Hrsg.). *International handbook of giftedness and talent* (S. 233-252) (2nd edition). Kidlington: Elsevier Science.
- Feldhusen, J. F. (1986). *A conception of giftedness*. In R. J. Sternberg & J. E. Davidson (Hrsg.), *International handbook of giftedness* (S.112-127). Cambridge: Cambridge University Press.
- Formann, A. (1973). *Die Konstruktion eines neuen Matrizen- und die Untersuchung des Lösungsverhaltens mit Hilfe des linearen logistischen Testmodells*. Universität Wien: Phil. Dissertation.
- Formann, A. K. & Pischwanger, K. (1979). *Wiener Matrizen-Test. Ein raschskaliertes sprachfreies Intelligenztest*. Weinheim: Beltz.
- Frenzel, A. C. & Nett, U. (2008). *Berliner Intelligenzstrukturtest für Jugendliche: Begabungs- und Hochbegabungsdiagnostik (BIS-HB) von A. O. Jäger et al.* *Diagnostica*, 54, 221-225.
- Frohnholzer, F. (2001). *PSB-R 6-13. Untersuchung an der Hauptschule*. Unveröffentlichte Zulassungssarbeit, Universität Regensburg, Lehrstuhl Psychologie VI.
- Frye, D., Zelazo, P. D., Brooks, P. J. & Samuels, M. C. (1996). *Inference and action in early causal reasoning*. *Developmental Psychology*, 32 (1), 120-131.
- Funke, J. & Vaterrodt-Plünnecke, B. (1998). *Was ist Intelligenz?* München: Beck.
- Gardner, H. (1983). *Frames of Mind: The theory of multiple intelligences*. New York: Basic Books.
- 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.
- Guilford, J. P. (1964). *Persönlichkeit*. Weinheim: Beltz.
- Heller, K. & Geisler, H.-J. (1983). *Kognitiver Fähigkeits-Test (Grundschulform)*. Weinheim: Beltz.
- Heller, K. & Geisler, H.-J. (1983). *Kognitiver Fähigkeits-Test (Kindergarten)*. Weinheim: Beltz.
- Heller, K. A. & Perleth, C. (1998). *Kognitiver Fähigkeitstest für 5. bis 13. Klassen*. Göttingen: Hogrefe.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.
- Heller, K. A. & Perleth, C. (2007). *MHBT-P. Münchner Hochbegabungstestbatterie für die Primarstufe*. Göttingen: Hogrefe.
- Heller, K. A. & Perleth, C. (2007). *MHBT-S. Münchner Hochbegabungstestbatterie für die Sekundarstufe*. Göttingen: Hogrefe.

- Heller, K. A. (2001). *Hochbegabung im Kindes- und Jugendalter* (2. Auflage). Göttingen: Hogrefe.
- Heller, K. A., Gädike, A.-K. & Weinläder, H. (1985). *Kognitiver Fähigkeits-Test für 4. bis 13. Klassen (KFT 4-13+)* (2., verbesserte und erweiterte Auflage). Weinheim: Beltz.
- Heller, K. A., Kratzmeier, H. & Lengfelder, A. (1998). *Matrizen-Test-Manual, Band 1. Ein Handbuch mit deutschen Normen zu den Standard Progressive Matrices von J. C. Raven*. Weinheim: Beltz.
- Heller, K. A., Perleth, C. & Hany, E. A. (1994). *Hochbegabung – ein lange Zeit vernachlässigtes Forschungsthema. Einsichten – Forschung der Ludwig-Maximilians-Universität München*, Vol. 3, 1, 18-22.
- Hermann, T. (1976). *Lehrbuch der empirischen Persönlichkeitsforschung* (3. Auflage). Göttingen: Hogrefe.
- Holling, H. & Kanning, U. P. (1999). *Hochbegabung*. Göttingen: Hogrefe.
- Holling, H., Bertling, J., Zeuch, N. & Kuhn, J. T. (2010). *Automatische Itemgenerierung: Grundlagen und Darstellung automatisch generierter Items anhand Lateinischer Quadrate*. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 211-232). Göttingen: Hogrefe.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6*. Göttingen: Hogrefe.
- Holling, H., Wübbelmann, K. & Geldschläger, H. (1996). *Kriterien und Instrumente zur Auswahl von Begabten in der beruflichen Bildung*. In R. Manstetten (Hrsg.), *Begabtenförderung in der beruflichen Bildung* (S. 86-174). Göttingen: Hogrefe.
- Hörmann, H. (1964). *Aussagemöglichkeiten psychologischer Diagnostik*. Göttingen: Hogrefe.
- Horn, H. & Schwarz, E. unter Mitarbeit von G. Vieweger. (1977). *Bildertest BT 1-2. Deutsche Bearbeitung des Moray House Picture Intelligence Test 1. Neubearbeitung von R. Kühn und R. Heck-Möhling* (2. Auflage, 1. Auflage 1967). Weinheim: Beltz.
- Horn, J. L. (1968). *Organization of abilities and the development of intelligence*. *Psychological Review*, 75(3), 242-259.
- Horn, R. (2003). *Intelligenztests für Kinder, Heft 1/03. Keine kritische Anmerkung zum K-ABC*. *Report Psychologie*, 28, 189.
- Horn, W. (1962). *Leistungsprüfsystem (LPS)*. Göttingen: Hogrefe.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung (PSB)*. Göttingen: Hogrefe.
- Horn, W. (1983). *Leistungsprüfsystem (LPS). Handanweisung* (2., erweit. und verbesserte Aufl.). Göttingen: Hogrefe.
- Horn, W., Lukesch, H., Kormann, A. & Mayrhofer, S. (2002). *PSB-R 4-6. Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen. Revidierte Fassung*. Göttingen: Hogrefe.
- Horn, W., Lukesch, H., Mayrhofer, S. & Kornmann, A. (2004). *PSB-R 6-13. Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen. Revidierte Fassung*. Göttingen: Hogrefe.
- Huldi, M. (1992). *Replikation des Berliner Intelligenzstrukturmodells bei Schweizer Mittelschülern*. In L. Montada (Hrsg.), *Bericht über den 38. Kongress der Deutschen Gesellschaft für Psychologie in Trier 1992* (Band 1, S. 207-208). Göttingen: Hogrefe.
- Huldi, M. (1997). *Die Intelligenzstruktur von Mittelschülern: Zur Gültigkeit des Berliner Intelligenzstrukturmodells (BIS)*. Bern: Lang.
- Humphreys, L. G. (1985). *A conceptualization of intellectual giftedness*. In F. D. Horowitz & M. O'Brian

- (Hrsg.). *The gifted and talented: Development perspectives* (S. 331-360). Washington: APA.
- Hussy, W. (1987). Zur Steuerfunktion der Sprache beim Problemlösen. *Sprache & Kognition*, 1, 14-22.
- Hylla, E., Kraak, B., Horn, H., Schwarz, E. & Kühn, R. (1993). *Aufgaben zum Nachdenken 4+*. Göttingen: Hogrefe.
- Ibrahimovic, N. & Bulheller, S. (2005). *Wortschatztest – aktiv und passiv. Deutsche Version des Mill Hill Vocabulary Tests*. Frankfurt/M.: Harcourt.
- Ibrahimovic, N., Bulheller, S., Horn, R., Gittler, G., Institut für Test-und Begabungsforschung-GmbH (2006). *IBF. Intelligenz-Basis-Faktoren*. Frankfurt/M.: Harcourt.
- Ingenkamp, K. (1966). *Bildertest (BT 2-3)*. Weinheim: Beltz.
- Institut für Test und Begabungsforschung (ITB) & Gittler, G. (1998). *Intelligenz Struktur Analyse. Ein Test zur Messung der Intelligenz*. Frankfurt: Swets.
- 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(4), 505-524.
- Jacobs, C. & Petermann, F. (2007). *Grundintelligenztest (CFT-20R) von Rudolf Weiss (2006) [Testinformation]*. *Diagnostica*, 53, 109-113.
- Jäger, A. O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2006). *BIS-HB. Berliner Intelligenzstruktur-Test für Jugendliche: Begabungs- und Hochbegabungsdiagnostik*. Göttingen: Hogrefe.
- Jäger, A. O. & Althoff, K. (1983). *Der Wilde-Intelligenz-Test (WIT). Ein Strukturdiagnostikum*. Herausgegeben von der Deutschen Gesellschaft für Personalwesen e.V. Göttingen: Hogrefe.
- Jäger, A. O. & Althoff, K. (1994). *Der Wilde-Intelligenz-Test (WIT). Ein Strukturdiagnostikum*. Herausgegeben von der Deutschen Gesellschaft für Personalwesen e.V. (2., revidierte Auflage). Göttingen: Hogrefe.
- Jäger, A. O. (1963). *Der Wilde-Test. Ein neues Intelligenzdiagnostikum*. *Zeitschrift für experimentelle und angewandte Psychologie*, 1, 260-278.
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Jäger, A. O. (1982). *Mehrmodale Klassifikation von Intelligenzleistungen. Experimentell kontrollierte Weiterentwicklung eines deskriptiven Intelligenzstrukturmodells*. *Diagnostica*, 28, 195-225.
- Jäger, A. O. (1984). *Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven*. *Psychologische Rundschau*, 35, 21-35.
- Jäger, A. O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2004). *Berliner Intelligenzstrukturtest für Jugendliche: Begabungs- und Hochbegabungsdiagnostik (BIS-HB)*. Göttingen: Hogrefe.
- Jäger, A. O., Süß, H.-M & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test (BIS-4)*. Göttingen: Hogrefe.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger Publishers.
- Joswig, H. (2000). *Begabungen erkennen – Begabte fördern*. Rostock: Universität Rostock, Philosophische Fakultät.
- Kaufman, A. S. & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines, MN: American Guidance Service.
- Kersting, M., Althoff, K. & Jäger, A. O. (2008). *WIT-2. Wilde-Intelligenz-Test 2*. Göttingen: Hogrefe.
- Klauer, K. J. (2003). *Über internationale Entwicklungstendenzen sonderpädagogischer Diagnostik un-*

- ter förderdiagnostischem Aspekt. In G. Ricken, A. Fritz & C. Hofmann (Hrsg.), *Förderbedarf der sonderpädagogischen Diagnostik* (S. 127-142). Lengerich: Pabst Science Publisher.
- Kleine, D. & Jäger, A. O. (1987). Replikation des Berliner Intelligenzstrukturmodells (BIS) bei brasilianischen Schülern und Studenten. *Diagnostica*, 33, 14-29.
- Klix, F. (1983). Begabungsforschung – ein neuer Weg in der kognitiven Intelligenzdiagnostik? *Zeitschrift für Psychologie*, 191 (4), 360-387.
- Knaack, R. (1978). Zur Brauchbarkeit der Coloured Progressive Matrices (CPM) von Raven bei Kindern im Vorschulalter. *Psychologie in Erziehung und Unterricht*, 25, 159-167.
- Kramer, J. (1972). *Intelligenztest. Mit einer Einführung in Theorie und Praxis der Intelligenzprüfung* (4. revidierte Auflage). Solothurn: Antonius.
- Kratzmeier, H. & Horn, R. (1979). SPM. Raven-Matrizen-Test. *Standard Progressive Matrices*. Weinheim: Beltz.
- Kratzmeier, H. (1977). *Heidelberger Intelligenztest HIT 1-2*. Weinheim: Beltz.
- Kratzmeier, H. (1980). *APM. Manual RAVEN-Matrizen-Test*. Weinheim: Beltz.
- Kratzmeier, H. (1982). *Heidelberger Intelligenztest HIT 3-4*. Weinheim: Beltz.
- Kubinger, K. D & Wurst, E. (2000). *AID 2. Adaptives Intelligenz Diagnostikum 2*. Göttingen: Beltz.
- Kubinger, K. D. & Holoher-Ertl, S. (2010). Die Vorteile adaptiven Testens in der Hochbegabungsdiagnostik. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik. Band 8* (S. 197-210). Göttingen: Hogrefe.
- Kubinger, K. D. & Wurst, E. (1991). *AID. Adaptives Intelligenz Diagnostikum* (3., ergänzte Auflage). Weinheim: Beltz.
- Kubinger, K. D. (2003). Adaptives Testen. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 1-9). Weinheim: Beltz/PVU.
- Kubinger, K. D. (2006). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Kubinger, K. D. (2009). *Adaptives Intelligenz Diagnostikum – Version 2.2 (AID 2) samt AID 2 Türkisch*. Göttingen: Beltz.
- Kubinger, K. D., Formann, A. K. & Farkas, M. G. (1991). Psychometric shortcomings of Raven's Standard Progressive Matrices (SPM) in particular for computerized testing. *European Review of Applied Psychology*, 41, 295-300.
- Kuhn, J. T., Holling, H. & Freund, A. (2008). Begabungsdiagnostik mit dem Grundintelligenztest (CFT 20-R). *Psychometrische Eigenschaften und Messäquivalenz. Diagnostica*, 54, 184-192.
- Kühn, R. (1974). Untersuchungen zur Dimensionalität des BT 1-2. *Mitteilungen und Nachrichten Nr. 75/76, Deutsches Institut für international pädagogische Forschung*, 40-73.
- Lehrl, S. (1972). Intelligenzspurtests – psychopathologische Messinstrumente der Intelligenz. *Medizinische Welt*, 23, 167-168.
- Lehrl, S. (1977). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B*. Erlangen: Straube.
- Lehrl, S., Merz, J., Erzigkeit, H. & Galster, V. (1974). Der MWT-A – ein wiederholbarer Intelligenz-Kurztest, der weitgehend unabhängig von seelisch-geistigen Störungen ist. *Nervenarzt*, 45, 364-369.
- Lorge, I., Thorndike, R. L. & Hagen, E. (1964). *The Lorge-Thorndike Intelligence Tests (multilevel edition)*. Boston: Houghton-Mifflin.

- MacArthur, R. S. & Elley, W. B. (1963). *The reduction of socio-economic bias in intelligence testing. British Journal of Educational Psychology*, 33, 107-119.
- Marland, S. P. (1971). *Education of the gifted and talented. Report to the Congress of the United States by the Commissioner of Education. Washington DC: US Government Printing Office.*
- Matarazzo, J. D. (1982). *Die Messung und Bewertung der Intelligenz Erwachsener nach Wechsler. Bern: Huber.*
- McGee, M. G. (1979). *Human spatial abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences. Psychological Bulletin*, 86, 889-918.
- Melchers, P. & Preuss, U. (1991). *K-ABC. Kaufman-Assessment Battery for Children von Alan S. Kaufman und Nadeen L. Kaufman. Deutschsprachige Fassung. Amsterdam: Swets & Zeitlinger. (Anmerkung: 2006 erschien die 6. Auflage der K-ABC.)*
- Melchers, P., Schürmann, S. & Scholten, S. (2006). *K-TIM. Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene. Leiden: PITS.*
- Mellone, M. A. & Thomson, G. H. (1958). *Moray House Picture Intelligence Test 1. Edinburgh: Darien Press.*
- Metzler, P. & Schmidt, K.-H. (1992). *RASCH-Skalierung des Mehrfachwahl-Wortschatztests (MWT). Diagnostica*, 38, 31-51.
- Mills, C., Ablard, K. & Brody, L. (1993). *The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students. Roeper Review*, 15(3), 183-186.
- Mittring, G. & Rost, D. (2008). *Die verflixten Distraktoren: Über den Nutzen einer theoretischen Distraktorenanalyse bei Matrixentests (für besser Begabte und Hochbegabte), Diagnostica*, 54(4), 193–201.
- Mönks, F. J. (1995). *Hochbegabung – ein Mehrfaktorenmodell. Grundschule*, 28, 15-18.
- Moser, A. (2003). *Intelligenz und ihre Beziehung zur Schulleistung und Konzentration. Diplomarbeit: Universität Regensburg.*
- Naglieri, J. A. & Ford, D. (2003). *Addressing Underrepresentation of Gifted Minority Children Using the Naglieri Nonverbal Ability Test (NNAT). Gifted Child Quarterly*, 47(2), 155-160.
- Naglieri, J. A. (1985). *Matrix analogies test expanded form (MAT). San Antonio, TX: Psychological Corp.*
- Naglieri, J. A. (1997). *Naglieri Nonverbal Ability Test (NNAT). San Antonio, TX: Psychological Corp.*
- Naglieri, J. A. (2003). *Naglieri Nonverbal Ability Test – individual administration. San Antonio, TX: Harcourt Assessment.*
- Oswald, W. D. & Roth, E. (1978). *Der Zahlenverbindungstest (ZVT). Göttingen: Hogrefe.*
- Penrose, L. S. & Raven, J. C. (1936). *A new series of perceptual tests: Preliminary communication. British Journal of Medical Psychology*, 16, 97–104.
- Perleth, C. & Sierwald, W. (1996). *Stabilität und Veränderungen in einem zweidimensionalen Intelligenzstrukturmodell. Ein Beitrag zur psychometrischen Intelligenzforschung. Vortrag auf der 38. TEAP in Eichstätt.*
- Perleth, C. (2006). *Die Münchner Hochbegabungs-Testbatterie (MHBT). In H. Wagner (Hrsg.), Intellektuelle Hochbegabung. Aspekte der Diagnostik und Beratung. Tagungsbericht (in Zusammenarbeit mit der Thomas-Morus-Akademie Bensberg; S. 56-69). Bad Honnef: Bock.*
- Perleth, C., Hofmann, U., Schauer, S. & Wernberger, H. (1994). *Intelligence testing in a Bavarian comprehensive school. School Psychology International*, 15, 263-277.

- Petermann, F. & Lipsius, M. (2009). WPPSI-III. Wechsler Preschool and Primary Scale of Intelligence – Third Edition. Deutsche Version. Frankfurt/M.: Pearson.
- Petermann, F. & Petermann, U. (2007). HAWIK-IV. Hamburg-Wechsler-Intelligenztest für Kinder – IV. Übersetzung und Adaptation der WISC-IV von David Wechsler. Bern: Huber.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piswanger, K. (1977). *Interkulturelle Vergleiche mit dem Matrizenest von FORMANN*. Universität Wien: Phil. Dissertation.
- Preckel, F. (2003). *Diagnostik intellektueller Hochbegabung. Testentwicklung zur Erfassung der fluiden Intelligenz*. Göttingen: Hogrefe.
- Preckel, F. (2010). *Intelligenztests in der Hochbegabungsdiagnostik*. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Tests und Trends, Jahrbuch der pädagogisch-psychologischen Diagnostik*. Band 8 (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.
- Raven, J. C. (1938). *Progressive Matrices*. London: Lewis.
- Raven, J. C. (1978). *Advanced Progressive Matrices: Set II (1962 Revision)*. London: H.K. Lewis & Co.
- Raven, J. C., Court, J. H. & Raven, J. (1994). *Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales. Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Raven, J. C., Raven, J. & Court, J. H. (2002). *Raven's Progressive Matrices und Vocabulary Scales. Coloured Progressive Matrices mit der Parallelfom des Tests und der Puzzle-Form. Deutsche Bearbeitung und Normierung Stephan Bulheller und Hartmut Häcker [Manual]*. Frankfurt/M.: Swets.
- Raven, J., Raven, J. C. & Court, J. H. (1998). *Mill Hill Vocabulary Scale (1998 edition)*. Oxford: Psychologist Press.
- Raven, J., Raven, J. C., & Court, J. H. (2001). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Oxford, England: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.
- Renzulli, J. S. (1978). *What makes giftedness: A re-examination of the definition of the gifted and talented*. Storrs, CT: University of Connecticut, Bureau of Educational Research
- Renzulli, J. S. (1986). *The three-ring-conception of giftedness*. In R. J. Sternberg & J. E. Davidson (Hrsg.), *Conceptions of giftedness* (S. 53-92). New York: Cambridge University Press.
- Renzulli, J. S. (2004). *Eine Erweiterung des Begabungsbegriffs unter Einbeziehung co-kognitiver Merkmale*. In C. Fischer, F. Mönks & E. Grindel (Hrsg.), *Curriculum und Didaktik der Begabtenförderung* (S. 54-82). Münster: LIT-Verlag.
- Ricken, G., Fritz, A., Schuck, K. D. & Preuss, U. (2007). *HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter – III*. Bern: Huber.
- Rost, D. H. (1991). *Identifizierung von Hochbegabung*. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 23(3), 197-231.
- Rost, D. H. (2000). *Hochbegabte und hochleistende Jugendliche*. Münster: Waxmann.
- Saccuzzo, D. P., Johnson, N. E. & Guertin, T. L. (1994). *The Raven's Progressive Matrices Test as an Equitable Measure of IQ in an Ethnically Diverse Population*. Manuscript submitted for publication.

- Sameroff, A. J. & Haith M. M. (1996). *The five to seven year shift: The age of reason and responsibility*. Chicago, IL: The University of Chicago Press.
- Schaarschmidt, U., Ricken, G., Kieschke, U. & Preuss, U. (2004). *BIVA. Bildbasierter Intelligenztest für das Vorschulalter*. Göttingen: Hogrefe.
- Schmidtke, A., Schaller, S. & Becker, P. (1980). *CPM. RAVEN-Matrizen-Test. Coloured Progressive Matrices. Manual (2. verbesserte Auflage)*. Weinheim: Beltz.
- Schröder, H. (2005). *KLI 4-5 R. Kombiniertes Lern- und Intelligenztest für 4. und 5. Klassen – Revidierte Form (6., revidierte Auflage)*. Göttingen: Hogrefe.
- Schuck, K. D. (1976). *Untersuchungen zur Erfassung von Umweltkorrelaten des verbalen und nichtverbalen kognitiven Leistungsniveaus vier- bis sechsjähriger Kinder*. Frankfurt: Universität, Dissertation.
- Schuck, K. D. (2001). *Diagnostik*. In J. Borchert (Hrsg.), *Handbuch der Sonderpädagogischen Psychologie* (S. 233-248). Göttingen: Hogrefe.
- Schuck, K. D., Eggert, D. & Raatz, U. (1975). *Columbia Mental Maturity Scale CMM 1-3. Sprachfreier Gruppenintelligenztest für die Grundschule*. Weinheim: Beltz.
- Schuck, K. D., Eggert, D. & Raatz, U. (1976). *Columbia Mental Maturity Scale. CMM 1-4. Sprachfreier Einzel- und Gruppenintelligenztest für die Grundschule und Allgemeine Sonderschule*. Wien: Ketterl.
- Schweizer, K., Goldhammer, F., Rauch, W. & Moosbrugger, H. (2007). *On the validity of Raven's Matrices Tests: Does spatial ability contribute to performance? Personality and Individual Differences*, 43, 198-201.
- Siegler, R. S. (1996). *Unidimensional thinking, multidimensional thinking, and characteristic tendencies of thought*. In A. J. Sameroff & M. M. Haith (Hrsg.), *The five to seven year shift: The age of reason and responsibility* (S. 63-84). Chicago, IL: The University of Chicago Press.
- Snijders, J. T. & Snijders-Oomen, N. (1970). *Snijders-Oomen nicht-verbale Intelligenztestreihe (S.O.N.)*. Groningen: Wolters-Nordhoff.
- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (1997). *Snijders-Oomen Non-verbaler Intelligenztest SON-R 5½-17 (2. korrigierte Auflage)*. Frankfurt/M.: Swets.
- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2005). *Non-Verbaler Intelligenztest SON-R 5½-17 (3. korrigierte Auflage)*. Göttingen: Hogrefe.
- Snijders, J. T., Tellegen, P. J. & Laros, J. A. (2007). *SON-R 2½-7. Non-verbaler Intelligenztest*. Göttingen: Hogrefe.
- Snijders-Oomen, N. (1943). *Intelligentieonderzoek van doofstomme kinderen (Intelligenzuntersuchung bei taubstummen Kindern)*. Nijmegen: Berkhout.
- Spearman, C. (1904). "General intelligence", objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spearman, C. (1938). *Measurement of intelligence*. *Scientia*, 64, 75-82.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (1986). *A triarchic theory of intellectual giftedness*. In R. J. Sternberg & J. E. Davidson (Hrsg.), *Conceptions of giftedness* (S. 223-246). Cambridge: Cambridge University Press.
- Sternberg, R. J. (1997). *The triarchic theory of intelligence*. In D. P. Flanagan, J. L. Genshaft & P. L. Har-

- risson (Hrsg.), *Contemporary intellectual assessment: Theories, tests, and issues* (S. 92-104). New York: Guilford Press.
- Stone, L. J., Smith, H. T. & Murphy, L. B. (Hrsg., 1973). *The competent infant – research and commentary*. New York: Basic Books.
- Stumpf, H. & Fay, E. (1981). *Entwicklung und Erprobung eines neuartigen Aufgabentyps zur Erfassung des räumlichen Vorstellungsvermögens*. *Diagnostica*, 27(2), 157-174.
- Stumpf, H. & Fay, E. (1983). *Schlauchfiguren – Ein Test zur Beurteilung des räumlichen Vorstellungsvermögens (Testmappe mit Handanweisung, Testheften A und B, 10 Antwortbögen und Schablonen A und B)*. Göttingen: Hogrefe.
- Süß, H.-M., Oberauer, K., Wittmann, W., Wilhelm, W. & Schulze, R. (2002). *Working-memory capacity explains reasoning ability – and a little bit more*. *Intelligence*, 30, 261-288.
- Süss-Burghart, H. (2006). *Der Intelligenztest „Naglieri Nonverbal Ability Test (NNAT)“ in der Diagnostik einer klinischen Stichprobe und im Vergleich mit der „Kaufman Assessment Battery for Children (K-ABC)“ und dem „Hamburg-Wechsler Intelligenztest für Kinder (HAWIK-III)“*. *Frühförderung interdisziplinär*, 25(3), 113-121.
- Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. & Laros, J. (1998). *Snijders-Oomen Nonverbal Intelligence Test, SON-R 2½-7. Manual and Research Report*. Lisse: Swets & Zeitlinger.
- Tewes, U. (1983). *Hamburg-Wechsler-Intelligenztest für Kinder – Revision 1983*. Bern: Huber.
- Tewes, U. (1991). *Hamburg-Wechsler Intelligenztest für Erwachsene. Revision 1991. Handbuch und Testanweisung*. Stuttgart: Huber.
- Tewes, U., Rossmann, P. & Schallberger, U. (Hrsg., 1999). *HAWIK-III. Hamburg-Wechsler-Intelligenztest für Kinder – Dritte Auflage. Übersetzung und Adaptation der WISC-III Wechsler Intelligence Scale for Children – Third Edition von David Wechsler*. Bern: Huber.
- Thorndike, R. L. & Hagen, E. P. (1977). *Measurement and Evaluation in Psychology and Education*. New York: Wiley.
- Thorndike, R. L., Hagen, E. & Lorge, I. (1968). *Cognitive Abilities Tests*. Boston: Houghton-Mifflin.
- Thorndike, Robert L. & Hagen, Elizabeth P. (1971). *Cognitive Abilities Test*. Boston.
- Thorndike, Robert L. & Hagen, Elizabeth P. (1993). *Form 5 Cog AT. Norms booklet*. Chicago.
- Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.
- Van der Meer, E. (1998). *Intelligenz als Informationsverarbeitung*. In E. Roth (Hrsg.), *Intelligenz* (S. 161-184). Stuttgart: Kohlhammer.
- Vock, M. (2008). *Non-verbaler Intelligenztest (SON R 2½-7)*. *Diagnostica*, 54, 112-115.
- Wagner, R. K. & Sternberg, R. J. (1985). *Practical intelligence in real world pursuits: The role of tacit knowledge*. *Journal of Personality and Social Psychology*, 49(2), 436-458.
- Wechsler, D. & Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability (WNV). Administration and scoring manual*. San Antonio, TX: Harcourt Assessment.
- Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.
- Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore: Williams & Wilkins.
- Wechsler, D. (1949). *Manual for the Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.
- Wechsler, D. (1956). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.
- Wechsler, D. (1956). *Intelligenztests für Kinder (HAWIK)*. Bern: Huber.

- Wechsler, D. (1967/2002). *Wechsler Primary and Preschool Scale of Intelligence™ -Third edition (WPPSI™-III)*. San Antonio, TX: Harcourt Assessment.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children – Revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised (WAIS-R)*. New York: Psychological Corporation.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children – Third Edition (WISC-III)*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale – III*. San Antonio, TX: Harcourt Assessment.
- Wechsler, D. (2002). *Wechsler Primary and Preschool Scale of Intelligence – Third edition (WPPSI-III)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2003). *The WISC-IV technical and interpretive manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2004). *Wechsler Intelligence Scale for Children – Fourth Edition: Canadian*. Toronto, Ontario: Harcourt Assessment.
- Wechsler, D. (2005). *Wechsler Intelligence Scale for Children – Fourth Edition, Spanish*. San Antonio, TX: Harcourt Assessment.
- Weiß, C. (2008). *Hochbegabtenberatung: Konzepte und Evaluation*. Unveröffentlichte Diplomarbeit. Institut für Pädagogische Psychologie, Julius-Maximilians-Universität Würzburg.
- Weiß, H. R. & Weiß, B. (2006). *CFT 20-R. Grundintelligenztest Skala 2 – Revision*. Göttingen: Hogrefe.
- Weiß, R. & Osterland, J. (1997). *Grundintelligenztest Skala 1. CFT 1 (Westermann Test)*. 5., revidierte Auflage. Göttingen: Hogrefe.
- Weiß, R. (1971). *Grundintelligenzskala 3 (CFT 3)*. Braunschweig: Westermann.
- Weiß, R. (1978). *Grundintelligenztest Skala 2 (CFT 20)*. Braunschweig: Westermann.
- Weiß, R. H. & Osterland, J. (1976). *Grundintelligenztest CFT 1. Skala 1*. Braunschweig: Westermann.
- Wenke, W. & Müller, U. (1966). *Möglichkeiten und Grenzen des Einsatzes einzelner diagnostischer Kurzverfahren bei der Schülersauslese*. *Zeitschrift für Psychologie*, 172, 82-116.
- Westmeyer, H. (1972). *Logik der Diagnostik*. Stuttgart: Kohlhammer.
- White, S. H. (1996). *The child's entry into the „Age of Reason“*. In A. J. Sameroff & M. M. Haith (Hrsg.), *The five to seven year shift: The age of reason and responsibility* (S. 17-32). Chicago, IL: The University of Chicago Press.
- Wilkes, J. & Weigel, A. (1998). *Vergleich von HAWIK-R und den Progressiven Matrizen tests (Raven) in einer klinischen Inanspruchnahmepopulation*. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 26, 261-265.
- Wolfram, H., Neumann, J. & Wiczorek, V. (1986). *Psychologische Leistungstests in der Neurologie und Psychiatrie. Methoden und Normwerte*. Leipzig: VEB Georg Thieme.
- Wonderlic, E. F. (1992). *Wonderlic Personnel Test*. Libertyville, IL: E.F. Wonderlic & Associates, Inc.
- Ziegler, A. (2008). *Hochbegabung*. München: Reinhardt.
- Ziegler, A. (2009). *„Ganzheitliche Förderung“ umfasst mehr als nur die Person: Aktiotop- und Soziotopförderung*. *Heilpädagogik online* 02/09, 5-34. Verfügbar unter: www.heilpaedagogik-online.com/2009/heilpaedagogik_online_0209.pdf [31.01.2011].

Anhang 1 – Mustergutachten



Dieser Abschnitt beinhaltet drei Mustergutachten, die eine Hilfe für die Erstellung von Gutachten im Rahmen der Begabungs- und Intelligenzdiagnostik bieten sollen.

Die Diagnostik von Begabung bzw. Intelligenz sollte stets von einer bestimmten Fragestellung abgeleitet werden, die auch die Verfahrensauswahl leitet. Die drei Mustergutachten behandeln folgende Fragestellungen:

- Vorzeitige Einschulung
- Überspringen einer Schulstufe
- Intelligenzdiagnostik im Rahmen einer Eignungsuntersuchung für eine berufliche Ausbildung

Vorzeitige Einschulung

An

Salzburg, am *****

Klinisch-psychologisches Gutachten

Aus den erhobenen Daten und den Ergebnissen der psychologischen Untersuchung wird im Auftrag der ***** nachstehend Befund und Gutachten über die mj. Susanne **** (geb. am *****, 5;2 Jahre) erstellt.

Grundlage der Begutachtung war die Verwertung der Untersuchungsdaten sowie die Angaben der Eltern und der Kindergartenpädagogin.

Die Fragestellung umfasst folgende Punkte:

- Vorzeitige Einschulung

Das psychologische Gutachten basiert auf:

- Explorationsgespräch mit der Kindergartenpädagogin und den Eltern
- Psychologische Untersuchung des Kindes mittels SON-R 5½-17, AID 2, KFT-K, K-ABC, *****

Explorationsgespräch

Die Exploration steht am Beginn jeder diagnostischen Untersuchung und beinhaltet die Erörterung der gesamten relevanten Vorgeschichte einschließlich der Beschreibung der aktuellen Situation sowie das Sammeln, Systematisieren und diagnostische Verarbeiten der Informationen. Sie gibt Hinweise für die Gestaltung der Untersuchung, der Auswahl der Tests und untersucht schon vorab Persönlichkeitseigenschaften, Interessen, Probleme und Denkweisen der zu testenden Person.

Psychologische Testverfahren

SON-R 5½-17 (Snijders Oomen Non-verbaler Intelligenztest)

Der SON-R 5½-17 ist ein Intelligenztest für Kinder im Alter zwischen 5;6 und 17 Jahren. Er besteht aus sieben Subtests: Kategorien, Analogien, Situationen, Bildgeschichten, Mosaik, Zeichenmuster und Suchbilder. Die ersten drei sind Mehrfachwahl-, die übrigen Handlungstests. Inhaltlich lassen sich die Subtests des SON-R 5½-17 in vier Gruppen einteilen: Tests für abstraktes Denken (Kategorien, Analogien), Tests für konkretes Denken (Situationen, Bildgeschichten), Tests für räumliches Vorstellungsvermögen (Mosaik, Zeichenmuster) und Perzeptionstests (Suchbilder).

AID 2 (Adaptives Intelligenzdiagnostikum – Version 2)

Das AID 2 dient der Erfassung komplexer und basaler Kognitionen bei Kindern und Jugendlichen. Zugrunde liegt die pragmatische intelligenztheoretische Position, möglichst viele Fähigkeiten zu erfassen. Gemessen werden die „verbal-akustischen“ Fähigkeiten Alltagswissen, angewandtes Rechnen, unmittelbares Reproduzieren-numerisch, Synonyme finden, Funktionen abstrahieren, soziales Erfassen und sachliches Reflektieren sowie die „manuell-visuellen“ Fähigkeiten Realitätssicherheit, soziale und sachliche Folge-richtigkeit, Kodieren und Assoziieren, Antizipieren und Kombinieren-figural, Analysieren und Synthetisieren-abstrakt; ergänzend gibt es die Zusatztests Unmittelbares Reproduzieren-figural/abstrakt, Merken und Einprägen und Strukturieren-visumotorisch. Alle elf

Untertests sowie die drei Zusatztests des AID 2 beruhen auf operationalen Definitionen, die die gemessenen Fähigkeiten jeweils festlegen. Neben der globalen Beurteilung der Intelligenz und der Profilinterpretation ermöglicht das Diagramm zur Diagnostik von Teilleistungsstörungen ein entsprechendes Screening in Bezug auf ausgewählte Teilleistungsfähigkeiten. Das Beiblatt für Beobachtungen der „Arbeitshaltungen“ dient der qualitativen Beurteilung des Arbeits- und Kontaktverhaltens bei Leistungsanforderung.

KFT-K (Kognitiver Fähigkeitstest – Kindergartenform)

Der Kognitive Fähigkeitstest (KFT-K) ist ein Niveautest zur Ermittlung kognitiver Lernfähigkeiten von 5- bis 6-jährigen Kindern. Er vermittelt Informationen über Sprachverständnis, Erkennen von Relationen, schlussfolgerndes sowie rechnerisches Denken und dient damit der Erfassung schulisch relevanter intellektueller Lern- und Leistungsvoraussetzungen.

K-ABC (Kaufmann Assessment Battery for Children)

Die K-ABC ist ein Testverfahren, das Intelligenz und Fertigkeiten misst. Die Grundlage der K-ABC ist die Definition der Intelligenz als Fähigkeit, Probleme durch geistiges Verarbeiten zu lösen, so dass bei der Diagnose der Prozess der Lösungsfindung und nicht der Inhalt der Aufgabe im Vordergrund steht. Die Messung intellektueller Fähigkeiten wird von der Messung des Standes erworbener Fertigkeiten (Lernen und Wissen) getrennt, um diese unterschiedlichen Bereiche mentaler Leistung einzeln und im Vergleich miteinander erfassen zu können. Deshalb ist die K-ABC in vier Skalen gegliedert: Skala einzelheitlichen Denkens, Skala ganzheitlichen Denkens (als Skalen intellektueller Fähigkeiten), Fertigkeitenskala und Sprachfreie Skala.

[...]

Zur Situation

[...]

Untersuchungsbericht

Intelligenztests

SON-R 5½-17 (Snijders Oomen Non-verbaler Intelligenztest)

Susanne erfasste die Aufgabenstellung jeweils sofort. Sie zeigte keine Ermüdungsercheinungen.

Sie erreichte folgende Ergebnisse:

(Angegeben werden die Rohwerte (RW) und Normwerte (N) für die einzelnen Subtests sowie der normierte Gesamtwert (Mittelwert), zusätzlich der Kennwert zur Überprüfung des Unterschieds zwischen den Subtestwerten (E2), die latenten Subtestwerte (L) und das zugehörige 80%-Intervall.)

<i>Subtest</i>	<i>RW</i>	<i>N*</i>	<i>E2**</i>	<i>L</i>	<i>(80%-Int.)</i>
Kategorien	24	157	139	146	136–157
Mosaik	20	151	62	148	140–156
Suchbilder	30	154	103	147	137–156
Zeichnmuster	13	151	60	148	140–156
Situationen	28	155	107	151	142–160
Analogien	30	151	66	148	140–156
Bildergeschichten	14	157	133	151	141–161
<i>Mittelwert</i>		<i>154</i>			

* Die Normwerte basieren auf den Kennwerten für 5;6 Jahre alte Kinder (MW=100, SD=15).

** Subtests unterscheiden sich nicht signifikant voneinander ($p > .05$).

Susanne erreicht bei allen Subtests Werte weit über ihrer Altersnorm. Vergleicht man die Leistungen zwischen den Subtests, ergibt sich kein signifikanter Unterschied. Auffallend ist, dass sie die Subtests für das räumliche Vorstellungsvermögen und den Perzeptionstest sehr schnell (weit unter der vorgegebenen Zeit) löst. Für die Situationen und Bildergeschichten muss sie hingegen länger nachdenken. Der Normmittelwert (154; MW=100,

SD=15) deutet auf eine weit überdurchschnittliche Intelligenz hin. Der spezifische IQ¹⁹ liegt bei 147, der generalisierte IQ²⁰ bei 140 (131–148). Es ergibt sich daher ein standardisierter IQ von 145, der einem Prozentrang von 99 entspricht, d.h. 99% der Vergleichspopulation sind in diesem Test schlechter als Susanne.

AID 2 (Adaptives Intelligenzdiagnostikum – Version 2)

Susanne erfasste die Aufgabenstellung jeweils sofort. Sie zeigte keine Ermüdungserscheinungen.

Sie erreichte folgende Ergebnisse²¹:
(Angegeben werden die T-Werte²² (T) für die einzelnen Subtests):

<i>Subtest</i>	<i>T</i>
Alltagswissen	55
Realitätssicherheit	62
Angewandtes Rechnen	65
Soziale und sachliche Folgerichtigkeit	67
Unmittelbares Reproduzieren	61
Synonyme finden	60
Kodieren und Assoziieren	65
Antizipieren und Kombinieren	72
Funktionen abstrahieren	62
Analysieren und Synthetisieren	68
Soziales Erfassen und Sachliches Reflektieren	62

¹⁹ Spezifische Fähigkeiten, die mit den 7 Subtests gemessen wurden.

²⁰ Intelligenzniveau unter Berücksichtigung aller vergleichbaren Subtests.

²¹ Referenzgruppe waren Kinder mit 6;0 Jahren.

²² MW=50, SD=10

Insgesamt ergibt die psychologische Untersuchung Susannes mit dem AID 2 eine überdurchschnittliche intellektuelle Leistungsfähigkeit, in keinem Bereich gibt es Schwächen. Die Leistungen in den Untertests streuen nicht merklich, es kann von ausgewogenen Ergebnissen gesprochen werden. Relative Schwierigkeiten hat sie lediglich bei den Fragen zum Alltagswissen, hier erreicht sie einen für 6-jährige Kinder durchschnittlichen Score (oberer Durchschnittsbereich). Am meisten Freude zeigt sie bei den Aufgaben zum angewandten Rechnen, beim Antizipieren und Kombinieren (figural) sowie beim Analysieren und Synthetisieren (abstrakt).

KFT-K (Kognitiver Fähigkeitstest – Kindergartenform)

Susanne erfasste die Aufgabenstellung jeweils sofort. Sie zeigte keine Ermüdungsercheinungen.

Sie erreichte folgende Ergebnisse:

(Angegeben werden die Rohwerte (RW), Prozenträge (PR) und T-Werte²³ (T) für die einzelnen Subtests sowie die Werte für die Gesamtleistung):

<i>Subtest</i>	<i>RW</i>	<i>PR</i>	<i>T</i>
Sprachverständnis	11	86	61
Beziehungserkennen	12	92	64
Schlussfolgerndes Denken	13	96	67
Rechnerisches Denken	14	93	65
<i>Gesamtleistung</i>	<i>50</i>	<i>96</i>	<i>67</i>

Das kognitive Fähigkeitsniveau von Susanne ist sehr hoch. Der GL-Wert von T=67 (PR=96) besagt, dass in der Vergleichsgruppe 96% der Kinder ein schlechteres Ergebnis als Susanne erzielen würden. Das Profil insgesamt ist ausgeglichen: In allen Subtests erreicht Susanne überdurchschnittliche Scores (T>60).

Vergleicht man die Subtests untereinander, so fällt auf, dass Susanne im Verhältnis den geringsten Score (nur leicht über dem Durchschnitt) beim Sprachverständnis erreicht. Dies ist aber wohl nicht auf ihre intellektuellen Fähigkeiten zurückzuführen, sondern viel

²³ MW=50, SD=10

eher darauf, dass die Begriffe mit ihr nicht eingeübt wurden. Zudem gilt der KFT-K als ein-dimensionales Verfahren, was eine Interpretation von Subtests nur bedingt ermöglicht.

K-ABC (Kaufmann Assessment Battery for Children)

Susanne erfasste die Aufgabenstellung jeweils sofort. Sie zeigte keine Ermüdungserscheinungen.

Sie erreichte folgende Ergebnisse:

(Angegeben werden die Standardwerte²⁴ für die Gesamtskalen und für die Untertests der Fertigkeitenskala):

<i>Gesamtskalen</i>	<i>Standardwerte*</i>
Skala einheitlichen Denkens	140
Skala ganzheitlichen Denkens	144
Skala intellektueller Fähigkeiten	146
Fertigkeitenskala	142
Sprachfreie Skala	138

* keine signifikanten Unterschiede zwischen den Skalen

<i>Fertigkeitenskalen</i>	<i>Standardwerte</i>
Wortschatz	134
Gesichter und Orte	140
Rechnen	145
Rätsel	147
Lesen/Verstehen	136

²⁴ MW=100, SD=15

Susanne erreicht bei der K-ABC bei den Gesamtskalen und bei den Fertigkeitenskalen durchwegs überdurchschnittliche Ergebnisse mit Standardwerten über 134. Zwischen den Standardwerten für die Gesamtskala zeigen sich keine signifikanten Unterschiede.

Konzentrationstests

[...]

Persönlichkeitstests

[...]

Zusammenfassung

Das Gutachten soll Aufschluss über eine eventuelle vorzeitige Einschulung der mj. Susanne **** geben. Es wurden Intelligenz, Konzentrationsfähigkeit und diverse maßgebliche Persönlichkeitsmerkmale erfasst.

Die Intelligenz wurde mit dem SON-R 5½-17, dem AID 2, dem KFT-K und der K-ABC untersucht, die Konzentrationsfähigkeit mit ****, die relevanten Persönlichkeitseigenschaften mit ****. Zusätzlich wurde ein Explorationsgespräch mit den Eltern und der Kindergartenpädagogin durchgeführt.

[...]

Anhand der Testergebnisse kann man feststellen, dass einer vorzeitigen Einschulung der mj. Susanne *** aus psychologischer Sicht nichts entgegensteht. Mit großer Wahrscheinlichkeit wird sie weder kognitiv noch aufgrund ihrer Persönlichkeit Schwierigkeiten – für den Fall einer vorzeitigen Einschulung – haben.

[...]

Hochachtungsvoll,

Überspringen einer Klasse

An

Salzburg, am *****

Klinisch-psychologisches Gutachten

Aus den erhobenen Daten und den Ergebnissen der psychologischen Untersuchung wird im Auftrag der ***** nachstehend Befund und Gutachten über den mj. Andreas***** (geb. am *****; 15;5 Jahre) erstellt.

Grundlage der Begutachtung war die Verwertung der Untersuchungsdaten sowie die Angaben des Untersuchten, der Eltern und der Klassenlehrer/innen.

Die Fragestellung umfasst folgende Punkte:

- Überspringen einer Klasse

Das psychologische Gutachten basiert auf:

- Exploration
- Psychologische Untersuchung mit dem BIS-HB, der MHBT-S, dem AID 2, *****

Explorationsgespräch

Die Exploration steht am Beginn jeder diagnostischen Untersuchung und beinhaltet die Erörterung der gesamten relevanten Vorgeschichte einschließlich der Beschreibung der aktuellen Situation sowie das Sammeln, Systematisieren und diagnostische Verarbeiten der Informationen. Sie gibt Hinweise für die Gestaltung der Untersuchung, die Auswahl der Tests und untersucht schon vorab Persönlichkeitseigenschaften, Interessen, Probleme und Denkweisen der zu testenden Person.

Psychologische Testverfahren

BIS-HB (Berliner Intelligenzstrukturtest für Jugendliche – (Hoch-)Begabungsdiagnostik)

Der BIS-HB eignet sich für die allgemeine und die strukturelle Intelligenzdiagnostik bei durchschnittlich und hoch intelligenten Kindern und Jugendlichen von 12 bis 16 Jahren. Das Verfahren basiert auf dem Berliner Intelligenzstrukturmodell (BIS) und erfasst die in diesem Modell spezifizierten Intelligenzdimensionen: Einfallsreichtum, Bearbeitungsgeschwindigkeit, Merkfähigkeit, Verarbeitungskapazität, die Fähigkeit zum Umgang mit verbalem, figuralen und numerischem Material sowie als Zusammenfassung aus diesen Fähigkeiten die allgemeine Intelligenz. Die verschiedenen, insgesamt 45 Aufgabentypen des Tests machen diesen in der Durchführung sehr abwechslungsreich. Die Vielfalt der Anforderungen unterstützt zudem die Akzeptanz des Verfahrens bei den getesteten Personen. Besonders zeichnet sich der BIS-HB dadurch aus, dass er eine große Vielfalt intellektueller Fähigkeiten erfasst. Auch wird über die Dimension Einfallsreichtum die in den meisten Intelligenztests vernachlässigte kognitive Leistungsdimension der Kreativität berücksichtigt.

MHBT-S (Münchner Hochbegabungstestbatterie für die Sekundarstufe)

Die MHBT-S enthält Test- und Fragebogenskalen zur Erfassung unterschiedlicher Begabungsformen auf hohem Niveau für Sekundarstufenschüler/innen der Klassen 4 bis 12+. Einsatz findet der Test dabei u.a. bei der Talentsuche für Förderprogramme oder auch in der Einzelfallberatung zur Abklärung von Hochbegabung.

AID 2 (Adaptives Intelligenzdiagnostikum – Version 2)

Das AID 2 dient der Erfassung komplexer und basaler Kognitionen bei Kindern und Jugendlichen. Zugrunde liegt die pragmatische Intelligenztheoretische Position, möglichst viele Fähigkeiten zu erfassen. Gemessen werden die „verbal-akustischen“ Fähigkeiten Alltagswissen, angewandtes Rechnen, unmittelbares Reproduzieren-numerisch, Synonyme finden, Funktionen abstrahieren, soziales Erfassen und sachliches Reflektieren sowie die „manuell-visuellen“ Fähigkeiten Realitätssicherheit, soziale und sachliche Folgerichtigkeit, Kodieren und Assoziieren, Antizipieren und Kombinieren-figural, Analysieren und Synthetisieren-abstrakt; ergänzend gibt es die Zusatztests unmittelbares Reproduzieren-figural/abstrakt, Merken und Einprägen und Strukturieren-visumotorisch. Alle elf Untertests sowie die drei Zusatztests des AID 2 beruhen auf operationalen Definitionen, die die gemessenen Fähigkeiten jeweils festlegen. Neben der globalen Beurteilung der Intelligenz und der Profilinterpretation ermöglicht das Diagramm zur Diagnostik von Teilleistungsstörungen ein entsprechendes Screening in Bezug auf ausgewählte Teilleistungsfähigkeiten. Das Beiblatt für Beobachtungen der „Arbeitshaltungen“ dient der qualitativen Beurteilung des Arbeits- und Kontaktverhaltens bei Leistungsanforderung.

[...]

Zur Situation

[...]

Untersuchungsbericht

Intelligenztests

BIS-HB (Berliner Intelligenzstrukturtest für Jugendliche – (Hoch-)Begabungsdiagnostik)

Andreas erfasste die Aufgabenstellung jeweils sofort. Er zeigte keine Ermüdungserscheinungen.

Er erreichte folgende Ergebnisse:
(Angabe sind die Normwerte²⁵ (N) und Konfidenzintervalle²⁶ (95%) zu den einzelnen Subskalen):

<i>Subskala</i>	<i>N</i>	<i>KI (95%)</i>
Verarbeitungskapazität	136	127–144
Bearbeitungsgeschwindigkeit	124	115–133
Merkfähigkeit	130	117–143
Einfallsreichtum	116	103–127
Leistungen bei sprachlichen Aufgaben	131	124–138
Leistungen bei Aufgaben mit Zahlen	136	126–146
Leistungen bei Aufgaben mit Figuren und Abbildungen	134	126–142
Allgemeine Intelligenz	135	124–146

Andreas erreicht auf allen Skalen Werte, die über dem Durchschnitt der Vergleichspopulation liegen ($N > 115$). Die Leistungen in den Subtests unterscheiden sich nicht deutlich. Lediglich bei der Skala Einfallsreichtum kann eine relative Schwäche beobachtet werden.

²⁵ MW=100, SD=15

²⁶ Psychologische Messungen weisen immer einen bestimmten Messfehler auf, ein einzelner Messwert entspricht daher nie 100%-ig dem tatsächlichen Wert. Diese Unsicherheit lässt sich mit dem Konfidenzintervall abschätzen. Dieses wird so bestimmt, dass es mit 95%-iger Wahrscheinlichkeit den Normwert enthält.

MHBT-S (Münchner Hochbegabungstestbatterie für die Sekundarstufe)

Andreas erfasste die Aufgabenstellung jeweils sofort. Er zeigte keine Ermüdungsercheinungen.

Er erreichte folgende Ergebnisse:
(Angegeben werden die Normwerte²⁷ (N) zu den einzelnen Leistungs-Skalen):

<i>Skala</i>	<i>N</i>
Abwicklungen	56
Spiegelbilder	59
Aufgaben aus Physik und Technik	65
Verbale Denkfähigkeiten	62
Mathematische Denkfähigkeiten	64
Technische Denkfähigkeiten	66
Kognitives Fähigkeitsniveau	65

In den kognitiven Fähigkeiten und im technischen Bereich liegt Andreas im überdurchschnittlichen Bereich, hier erreicht er überall Prozenträge über 88,5, d.h. 88,5% der Vergleichspopulation sind in diesen Bereichen schlechter als er. Beim räumlichen Denken (Abwicklungen, Spiegelbilder) sind die Scores im oberen Durchschnittsbereich, was jedoch für den gymnasialen Bildungserfolg ohnehin weniger relevant sein dürfte.

AID 2 (Adaptives Intelligenzdiagnostikum – Version 2)

Andreas erfasste die Aufgabenstellung jeweils sofort. Er zeigte keine Ermüdungsercheinungen.

Er erreichte folgende Ergebnisse:

²⁷ MW=50, SD=10

(Angegeben werden die T-Werte²⁸ (T) für die einzelnen Subtests):

<i>Subtest</i>	<i>T</i>
Alltagswissen	68
Realitätssicherheit	63
Angewandtes Rechnen	63
Soziale und sachliche Folgerichtigkeit	67
Unmittelbares Reproduzieren	58
Synonyme finden	63
Kodieren und Assoziieren	61
Antizipieren und Kombinieren	66
Funktionen abstrahieren	56
Analysieren und Synthetisieren	69
Soziales Erfassen und Sachliches Reflektieren	57

Insgesamt ergibt die psychologische Untersuchung von Andreas mit dem AID 2 eine überdurchschnittliche intellektuelle Leistungsfähigkeit, in keinem Bereich gibt es Schwächen. Die Leistungen in den Untertests streuen nicht merklich, es kann von ausgewogenen Ergebnissen gesprochen werden. Die besten Leistungen im Vergleich zur Altersgruppe zeigt er beim Alltagswissen (T=68), die relativ schlechtesten beim sozialen Erfassen und sachlichen Reflektieren (hier liegt er mit einem T-Wert von 57 im oberen Durchschnittsbereich).

Konzentrationstests

[...]

Persönlichkeitstests

[...]

²⁸ MW=50, SD=10

Zusammenfassung

Das Gutachten soll Aufschluss über ein eventuelles Überspringen des mj. Andreas **** geben. Es wurden Intelligenz sowie ***** [bestimmte relevante Persönlichkeitsmerkmale und Konzentrationsfähigkeit] erfasst.

Die Intelligenz wurde mit dem BIS-HB, der MHBT-S und dem AID 2 untersucht, die Konzentrationsfähigkeit mit ****, die relevanten Persönlichkeitseigenschaften mit ****. Zusätzlich wurde ein Explorationsgespräch mit den Eltern und den Lehrpersonen durchgeführt [...].

Anhand der Testergebnisse kann man feststellen, dass dem Überspringen der nächsten Klasse aus psychologischer Sicht nichts entgegensteht. Mit großer Wahrscheinlichkeit wird er keine Schwierigkeiten haben, dem Unterricht in der nächsthöheren Klasse zu folgen.

[...]

Hochachtungsvoll,

Intelligenz (Eignungsuntersuchung)

An

Salzburg, am *****

Klinisch-psychologisches Gutachten

Aus den erhobenen Daten und den Ergebnissen der psychologischen Untersuchung wird im Auftrag der ***** nachstehend Befund und Gutachten über Herrn ***** (geb. am *****, 18;6 Jahre) erstellt.

Grundlage der Begutachtung war die Verwertung der Untersuchungsdaten sowie die Angaben des Untersuchten.

Die Fragestellung umfasst folgende Punkte:

- Eignungsuntersuchung für die Ausbildung zum *****

Das psychologische Gutachten basiert auf:

- Exploration
- Psychologische Untersuchung mit dem I-S-T 2000-R und dem CFT 3, ***, ***,
*** und ****

Explorationsgespräch

Die Exploration steht am Beginn jeder diagnostischen Untersuchung und beinhaltet die Erörterung der gesamten relevanten Vorgeschichte einschließlich der Beschreibung der aktuellen Situation sowie das Sammeln, Systematisieren und diagnostische Verarbeiten der Informationen. Sie gibt Hinweise für die Gestaltung der Untersuchung, die Auswahl der Tests und untersucht schon vorab Persönlichkeitseigenschaften, Interessen, Probleme und Denkweisen der zu testenden Person.

Psychologische Testverfahren

I-S-T 2000 R (Intelligenz-Struktur-Test 2000 R)

Der I-S-T 2000 R ist ein vielseitig einsetzbarer, ökonomischer Intelligenztest. Im Rahmen des neuen, theoretisch fundierten und empirisch begründeten Strukturkonzepts können mit dem I-S-T 2000 R folgende 11 Fähigkeiten erfasst werden: verbale Intelligenz, figural-räumliche Intelligenz, rechnerische Intelligenz, Merkfähigkeit, schlussfolgerndes Denken, verbales Wissen, figural-bildhaftes Wissen, numerisches Wissen und Wissen (Gesamt) sowie fluide und kristallisierte Intelligenz.

CFT 3 (Culture Fair Test, Grundintelligenztest)

Der CFT 3 ermöglicht eine Ermittlung des Maßes für die „allgemeine intellektuelle Leistungsfähigkeit“ durch Vergleich mit der Altersgruppe und verschiedenen Bildungsgruppen.

[...]

Zur Situation

[...]

Untersuchungsbericht

Intelligenztests

I-S-T 2000 R (Intelligenz-Struktur-Test 2000 R)

Herr ***** erfasste die Aufgabenstellung jeweils sofort. Er zeigte keine Ermüdungserscheinungen.

Er erreichte folgende Ergebnisse:

(Angegeben als Rohwerte (RW) für die einzelnen Untertests (UT) und als Standardwerte (SW) für die einzelnen Skalen: verbale Intelligenz, numerische Intelligenz, figurale Intelligenz, Merkfähigkeit, verbales Wissen, numerisches Wissen und figurales Wissen):

UT	SW	Konstrukt
Grundmodul		
SE (Satzergänzung)	16	Urteilsbildung
AN (Analogien)	20	Kombinationsfähigkeit
GE (Gemeinsamkeiten)	17	Sprachl. Abstraktionsfähigkeit
∑ verbale Intelligenz	53	128 (=> IQ: 142)
RE (Rechenaufgaben)	19	Praktisch-rechnerisches Denken
ZR (Zahlenreihen)	20	Theoretisch-rechnerisches Denken
RZ (Rechenzeichen)	19	logische Beziehungen herstellen
∑ numerische Intelligenz	58	126 (=> IQ: 139)

UT	SW	Konstrukt
FA (Figurenauswahl)	18	Figurales Vorstellen
WÜ (Würfelaufgaben)	19	Räumliches Vorstellen
MA (Matrizenaufgaben)	19	logische Relationen herstellen
Σ figurale Intelligenz	56	130 (=> IQ: 145)
ME (Merkaufgaben figural)	11	figurale Merkfähigkeit
ME (Merkaufgaben verbal)	10	verbale Merkfähigkeit
Σ Merkfähigkeit	21	120 (=> IQ: 130)
Erweiterungsmodul		
VW (verbales Wissen)	27	125 (=> IQ: 137)
NW (numerisches Wissen)	19	102 (=> IQ: 103)
FW (figurales Wissen)	23	111 (=> IQ: 116)

Prozentränge (PR) und Vertrauensintervalle (CI) (95%):

	PR	CI
Verbale Intelligenz	100	$142 \pm 1,96 * 15 * \sqrt{1-0,89} = [151,75; 132,25]$
Numerische Intelligenz	100	$139 \pm 1,96 * 15 * \sqrt{1-0,96} = [144,88; 133,12]$
Figurale Intelligenz	100	$145 \pm 1,96 * 15 * \sqrt{1-0,88} = [155,18; 134,82]$
Merkfähigkeit	98	$130 \pm 1,96 * 15 * \sqrt{1-0,93} = [137,77; 122,22]$
Verbales Wissen	99	–
Numerisches Wissen	58	–
Figurales Wissen	86	–

Der Gesamt-Standardwert für die fluide Intelligenz betrug 131, ihm entspricht ein IQ von 145 und ein Prozentrang von 100, welcher besagt, dass die Leistung höher liegt als bei etwa 99 Prozent der Altersgruppe mit derselben Schulbildung.

Der Gesamt-Standardwert für die kristallisierte Intelligenz betrug 117, ihm entspricht ein IQ von 125 und ein Prozentrang von 96, welcher besagt, dass die Leistung höher liegt als bei etwa 95 Prozent der Altersgruppe mit einem vergleichbaren Schulabschluss.

Die Ergebnisse sprechen für eine überdurchschnittliche Gesamtintelligenz, wenn man sie auf die Gruppe der gleichaltrigen Maturantinnen und Maturanten bezieht.

Was die einzelnen Teilfunktionen des Grundmoduls betrifft, so sind im Hinblick auf die Altersgruppe die sprachlichen Leistungen, das zahlengebundene Denken, die Raumvorstellung und die Merkfähigkeit überdurchschnittlich ausgeprägt, d.h. die Fähigkeit zum Umgang mit sprachlichem Material, die Rechenfertigkeit und die Fähigkeit, logische Beziehungen zwischen Zahlen herzustellen, die Fähigkeit zum Umgang mit figural-bildlichem Material (mit zwei- und dreidimensionalen Figuren) und die Fähigkeit zum aktiven Einprägen und kurzfristigen Wiedererkennen von Informationen sind überdurchschnittlich ausgeprägt.

Das schlussfolgernde Denken, das sich aus den Skalen verbales Wissen, numerisches Wissen und figurales Wissen zusammensetzt, ergibt einen Prozentrang von 100, was auf eine weit überdurchschnittliche Fähigkeit im induktiven und deduktiven Denken schließen lässt.

Die hohe fluide Intelligenz, die durch diesen Begriff bezeichnet wird, weist darauf hin, dass Herr ***** gut Beziehungen zwischen Stimuli und Informationen herstellen kann und eine ausgeprägte Fähigkeit zum formal-logischen Denken hat.

Im Vergleich zum individuellen Leistungsniveau findet sich ein Schwerpunkt beim theoretisch-rechnerischen Denken, dagegen ein Leistungstiefpunkt bei den sprachlichen Funktionen im Sinne der Urteilsbildung.

Im Erweiterungsmodul Wissen erreichte Herr ***** einen Gesamtstandardwert von 117, was einem Prozentrang von 96 entspricht.

Die Kodierungsart des Wissens bei Herrn ***** kann, gemessen an den Standardwerten, als eher verbal bezeichnet werden.

In der numerischen Codierung erreicht er den niedrigsten Standardwert.

Die Wissensskala entspricht der kristallisierten Intelligenz, die besagt, dass eine Person in der Lage war, ein umfangreiches Wissen in unserer Kultur zu erwerben.

Auch auf diesem Gebiet hat Herr ***** im Vergleich zur Alters- und Bildungsgruppe überdurchschnittliche Werte erreicht, verglichen mit seinen anderen Leistungen im Test ist er aber auf diesem Gebiet als unterdurchschnittlich einzustufen.

[...]

CFT 3 (Culture Fair Test, Grundintelligenztest)

Herr ***** erfasste die Aufgabenstellung jeweils sofort. Er zeigte keine Ermüdungserscheinungen.

Herr ***** erreichte folgende Ergebnisse:
(Angabe als Standardwerte (SW) für die einzelnen Untertests (UT)):

UT	SW	Konstrukt
Test I:		
Reihenfortsetzung	8	Reasoning
Klassifikationen	12	Beziehungsstiftendes Denken
Matrizen	7	Reasoning
Topologisches Folgern	6	Interferenz
=> IQ: 145		
Test II:		
Reihenfortsetzung	9	Reasoning
Klassifikationen	11	Beziehungsstiftendes Denken
Matrizen	6	Reasoning
Topologisches Folgern	10	Interferenz
=> IQ: 142		

Der Gesamt-Standardwert betrug im ersten Testteil 33, ihm entspricht ein IQ von 145 (alterskorrigiert).

Im zweiten Test betrug der Gesamt-SW 36, was einem IQ von 142 entspricht (alterskorrigiert).

Die Ergebnisse sprechen für eine überdurchschnittliche Gesamtintelligenz.

Der Gesamtintelligenzwert ergibt einen IQ von 150 (laut Tabelle, alterskorrigiert), was in Einbezugnahme des Vertrauensintervalls einer Intelligenz im Bereich zwischen 143,2 und 156,2 Punkten entspricht.

Was die einzelnen Untertests betrifft, so sind die Teilbereiche „topologisches Schlussfolgern“ und „Klassifikationen“ (Subtests 2 und 4) in beiden Untertests besser als die Teilbereiche „Reihen und Matrizen“ (Subtests 1 und 3).

Damit ist gemeint, dass die Fähigkeit des „Reasoning“ (Subtests 1 und 3) im Vergleich zu den Fähigkeiten des beziehungsstiftenden Denkens (Subtest 2) und der Interferenz (Subtest 4) weniger ausgeprägt ist.

Bei Herrn ***** findet sich ein Schwerpunkt beim beziehungsstiftenden Denken (Subtest 4), bei dem er in beiden Testteilen sehr gute Werte erreichte.

Ein Leistungstiefpunkt ist bei den Matrizen zu sehen, bei denen meist nur die Hälfte der gesamtöglichen richtigen Lösungen erreicht wurde.

Zwischen den Bereichen mathematischen Denkens und dem Komplex der Grundintelligenz besteht nachweislich eine hohe Gemeinsamkeit. Da die grundlegenden getesteten Faktoren alle dem mathematisch-naturwissenschaftlichen Denken entspringen, weisen die hohen erreichten Intelligenzwerte auf jeden Fall auf eine mathematische Begabung hin.

[...]

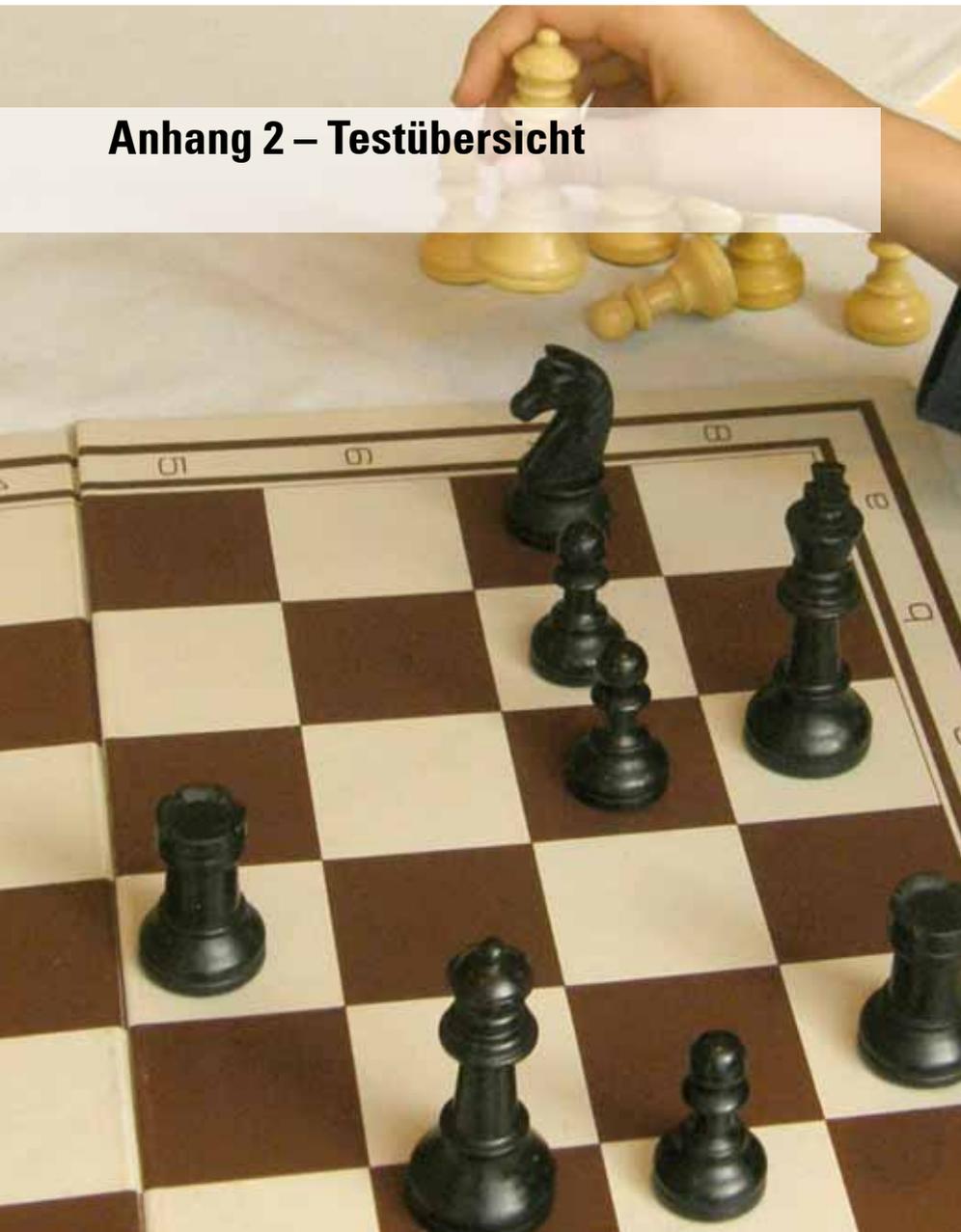
Zusammenfassung

Die psychologische Untersuchung ergab, dass Herr ***** über eine überdurchschnittliche Intelligenz verfügt, die in den Bereichen der mathematischen und figuralen Fähigkeiten am besten ausgeprägt ist. Im I-S-T 2000 R zeigt sich eine hohe fluide Intelligenz, wobei auch die kristallisierte Intelligenz im Vergleich zur Altersgruppe als überdurchschnittlich zu werten ist.

Der CFT misst hauptsächlich mathematische Fähigkeiten. Herr ***** erzielte dort sehr gute Ergebnisse, vor allem beim beziehungsstiftenden Denken. Grundsätzlich sind also vor allem die überdurchschnittlichen Fähigkeiten im logischen und abstrakten Denken sowie die räumliche Vorstellungskraft als besonders zu nennen.

Hochachtungsvoll,

Anhang 2 – Testübersicht



Der folgende Abschnitt beinhaltet eine Kurzbeschreibung aller vorgestellten Verfahren (alphabetische Reihung nach Testkürzel). Für die Hochbegabungsdiagnostik geeignete Tests sind grafisch mit der Dame-Figur gekennzeichnet.



AID 2 – Adaptives Intelligenzdiagnostikum – Version 2

<i>Altersbereich</i>	6;0–15;11 Jahre
<i>Struktur</i>	Basiert pragmatisch auf intelligenztheoretischer Position, möglichst viele Fähigkeiten zu erfassen; Inhalte der Untertests sind an Wechsler-Test orientiert.
<i>Kurzbeschreibung</i>	Testaufbau: 11 Untertests (+ 3 zusätzliche Tests). Ermittelt werden können neben Untertestkennwerten die Intelligenzquantität bzw. kognitive Mindestfähigkeit über den (zweit-)niedrigsten Untertestwert. Einzeltest
<i>Normen</i>	Normierung: 1995–1997 Reliabilität: Für die 9 Untertests und alle Zusatztests gegeben (Gültigkeit des Rasch-Modells).
<i>Testgüte</i>	Konstruktvalidität über explorative und konfirmatorische Faktorenanalysen belegt. Diskriminante Validität über Vorgänger belegt, z.T. nur geringe Korrelationen mit konstrukt nahen Tests.
<p>Das AID 2 bietet die Möglichkeit des adaptiven Testens, das gerade im Bereich der Hochbegabungsdiagnostik sehr viele Vorteile (Erhöhung der Messgenauigkeit, Verringerung der Testlänge, Optimierung der Entscheidungsstrategie) mit sich bringt. Das im AID 2 umgesetzte Konzept bietet abweichend von der standardmäßigen Vorgabe eine attraktive Testadministration und eignet sich hervorragend gerade für die Testung (hoch)begabter Kinder.</p> <p><i>Ausführliche Beschreibung ab S. 36.</i></p>	

APM – Advanced Progressive Matrices

<i>Altersbereich</i>	12–60 Jahre
<i>Struktur</i>	Ähnlich wie der CFT 20 sollen die APM die allgemeine (fluide, angeborene) Intelligenz erfassen. Schwierigste Version der Raven's Progressive Matrices; für die Testung von Personen mit überdurchschnittlichen Fähigkeiten geeignet.
<i>Kurzbeschreibung</i>	Testaufbau: 12 Übungsaufgaben (Set 1) und 26 Testaufgaben, die in aufsteigender Schwierigkeit angeordnet sind (Set 2); Matrizenaufgaben (3x3-Felder) mit 8 Antwortalternativen Einzel- und Gruppentest
<i>Normen</i>	Normierung: 1996/1997 Reliabilität: ausreichend
<i>Testgüte</i>	Konstrukt- und Kriteriumsvalidität durch Korrelationen mit Tests und Schulnoten belegt. Bei Schulnoten höchste Zusammenhänge mit Note in Mathematik.

Bezogen auf die Intelligenzdiagnostik ist zu beachten, dass die APM nur eine Intelligenzfacette – und dies zudem mit nur einem Aufgabentyp – erfassen. Beim Einsatz mit hochintelligenten 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 Schulstufen 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 Personen 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“).

Um die testspezifische Fehlervarianz gering zu halten, sollen Tests mit unterschiedlichen Aufgabenfacetten den APM vorgezogen werden. Der alleinige Einsatz figuraler Reasoning-Tests wie der APM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe auch Preckel, 2010).

Ausführliche Beschreibung ab S. 40.

AzN 4+ – Aufgaben zum Nachdenken 4+

<i>Altersbereich</i>	4. und 5. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Kombinierter Lern- und Intelligenztest</p> <p>Testaufbau: Insgesamt 76 Aufgaben und 17 Übungsbeispiele, gegliedert in 5 Subtests, die jeweils 12 bis 20 Einzelaufgaben beinhalten.</p> <p>Subtests 1 + 2: Rechnen und Zahlenreihen erfassen, rechnerisches und logisches Denkvermögen</p> <p>Subtests 3 + 4: Analogien und Satzergänzung erfassen, verbale und sprachlogische Fähigkeiten</p> <p>Subtest 5: Instruktionsverständnis prüft die Anpassungsfähigkeit der Kinder an neuartige Aufgabenstellungen</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1976</p> <p>Reliabilität:</p> <ol style="list-style-type: none"> (1) Anweisungsverständnis (2) Unterscheidungsfähigkeit (3) Verständnis von Folgen (4) Raumorientierung (5) Analogiebildung <p>Split-Half-Reliabilität sehr gut</p> <p>Validität gegeben</p> <p>Faktorenanalytische Überprüfung ergab zweidimensionale Struktur: Rechnerisch-logisches Denken und sprachlogische Fähigkeiten</p> <p>Konvergente Validität geprüft mit LPS und PSB sowie mit Schulnoten (mittlere Korrelationen).</p>
<p>Die AzN eignen sich sehr gut zur Prognose von Schulerfolg/Schulversagen, können aber eine intellektuelle Hochbegabung nicht hinreichend gut prüfen, da oft nur gelernte Inhalte abgefragt werden. Da diese nicht zwischen Intelligenz und Fleiß differenzieren, erscheint eine zusätzliche Individualtestung mit einem weiteren Test (Reasoning) sinnvoll.</p> <p><i>Ausführliche Beschreibung ab S. 44.</i></p>	

BIS-4 – Berliner-Intelligenzstruktur-Test

<i>Altersbereich</i>	16–19 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Test zur Intelligenz-Differenzialdiagnostik</p> <p>Testaufbau: 45 sehr verschiedene repräsentativ ausgewählte Aufgabentypen zu folgenden sieben Fähigkeiten:</p> <ul style="list-style-type: none"> - Bearbeitungsgeschwindigkeit - Merkfähigkeit, Einfallsreichtum - Verarbeitungskapazität - anschauungsgebundenes Denken - figural-bildhaftes Denken - sprachgebundenes Denken - zahlengebundenes Denken <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: keine Angaben zum Jahr; Normierungsstichprobe klein</p> <p>Reliabilität: hinreichend ($\alpha = .75$ bis $\alpha = .79$)</p> <p>Faktorielle Validität anhand des BIS-Modells geprüft.</p> <p>Kriteriumsvalidität geprüft (mittlere Korrelationen)</p>
<p>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 zufriedenstellend sind (die Werte liegen zwischen $r = .77$ für die Merkfähigkeit und $r = .93$ für g), die Konstruktreliabilitäten der spezifischen kognitiven Fähigkeiten unabhängig vom verwendeten Koeffizienten aber nicht zufriedenstellend sind (die Werte liegen zwischen $r = .17$ für die numerische Fähigkeit und $r = .67$ 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.</p> <p><i>Ausführliche Beschreibung ab S. 46.</i></p>	

BIS-HB – Berliner Intelligenzstruktur-Test für Jugendliche:
Begabungs- und Hochbegabendiagnostik



<i>Altersbereich</i>	12;6–16;5 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Basiert theoretisch auf dem BIS-Modell von Jäger</p> <p>Testaufbau: 3 Testhefte, insgesamt 45 Aufgabentypen</p> <p>4 operative Fähigkeiten:</p> <ul style="list-style-type: none"> - Verarbeitungskapazität - Einfallsreichtum - Merkfähigkeit - Bearbeitungsgeschwindigkeit <p>3 inhaltsgebundene Fähigkeiten:</p> <ul style="list-style-type: none"> - sprachgebundenes Denken - zahlengebundenes Denken - figural-bildhaftes Denken <p>Allgemeine Intelligenz BIS-g</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2002/2003</p> <p>Reliabilität: gut bis sehr gut ($\alpha = .80$ bis $\alpha = .95$)</p> <p>Validität mit Begabtenstichproben durch konfirmatorische Faktorenanalysen belegt.</p> <p>Konstrukt- und Kriteriumsvalidität durch Korrelationen mit Tests und Schulnoten belegt.</p>
<p>In einer Rezension folgern Frenzel & Nett (2008, S. 224): „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 großteils beigepflichtet werden, zur Unterschätzung ist jedoch zu sagen, dass der Test zwar eher konservative Schätzungen liefert, es sich jedoch um keine „erhebliche Unterschätzung“ handelt, für die Hochbegabungsdiagnostik ist er durchaus empfehlenswert.</p> <p><i>Ausführliche Beschreibung ab S. 49.</i></p>	

BIVA – Bildbasierter Intelligenztest für das Vorschulalter

<i>Altersbereich</i>	3;6–7;6 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testaufbau: 8 Subtests:</p> <ul style="list-style-type: none"> - Merkmale herauslösen - Vergleich von Merkmalen - Objekte herauslösen - Wort-Bild-Vergleich - Geschichten-Folgen - Reihen fortsetzen - Geschichten-Analogien - Reihen-Analogien <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: erfolgt; Zeitraum nicht definiert</p> <p>Reliabilität: mäßig bis gut ($\alpha=.65$ bis $\alpha=.88$)</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (geringe bis mittlere Korrelationen)</p>
<p>Der BIVA eignet sich für die Begabungsdiagnostik über dem Durchschnittsbereich nicht. Es treten schon bei IQs, die mehr als eine Standardabweichung über dem Mittelwert liegen, Deckeneffekte auf.</p> <p><i>Ausführliche Beschreibung ab S. 52.</i></p>	

BT 1-2 – Bildertest für 1. und 2. Klassen

<i>Altersbereich</i>	6–8 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testkonstruktion auf Grundlage der „general ability“</p> <p>Testaufbau: 8 Subtests mit jeweils 10–13 Aufgaben. Insgesamt 91 Aufgaben zu den Bereichen</p> <ul style="list-style-type: none"> - Instruktionen - Nichtpassendes - Ergänzungen - Unsinniges - Spiegelbilder - Folgen - Wesentliches - Reihen <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1975/1976</p> <p>Neunormierung liegt laut Angaben der Autorinnen und Autoren vor.</p> <p>Reliabilität: gering ($\alpha = .34$)</p> <p>Konstruktvalidität $r = .62$ (CMM 1-3)</p>
<p>Der BT 1-2 ist ein sehr ansprechender Test für Kinder im Grundschulalter. Erfahrungsgemäß machen ihn die Kinder sehr gerne und haben Spaß an den Aufgaben. Intellektuelle Hochbegabung kann mit ihm allerdings nicht abgeprüft werden, da die Testaufgaben zumeist sehr einfach sind („Deckeneffekte“) und auch die Trennschärfen der Items in vielen Fällen nicht hinreichend groß sind. Zudem ist die Reliabilität (interne Konsistenz) des Tests gering, weshalb eine „allgemeine Intelligenz“ im Sinne eines Gesamt-IQs ohnehin nicht festgestellt werden kann. Spezifische Neigungen und auch Begabungen können aufgrund der Testergebnisse angenommen werden. Bevor eine intellektuelle Hochbegabung diagnostiziert wird, sollte aber unbedingt ein zweiter unabhängiger Test durchgeführt werden.</p> <p><i>Ausführliche Beschreibung ab S. 56.</i></p>	

BT 2-3 – Bildertest für 2. und 3. Klassen

<i>Altersbereich</i>	7;0–9;3 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testkonstruktion auf Grundlage der „general ability“</p> <p>Testaufbau: 5 Subtests:</p> <ul style="list-style-type: none"> - Anweisungsverständnis - Unterscheidungsfähigkeit - Raumorientierung - Verständnis von Folgen - Analogiebildung <p>Insgesamt 100 Aufgaben</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1975</p> <p>Reliabilität: interne Konsistenz nicht angegeben, mittelhohe Iteminterkorrelationen zwischen $r = .41$ und $r = .52$</p> <p>Faktorielle Validität: g-Faktor klärt 89% der Varianz auf.</p>
<p>Der BT 2-3 ist ebenso wie der BT 1-2 ein sehr ansprechender Test für Kinder im Grundschulalter. Positiv sind die Verschiedenartigkeit der Aufgaben und die hohe prognostische Validität für den Schulerfolg. Intellektuelle Hochbegabung kann mit ihm allerdings nicht abgeprüft werden, da die Testaufgaben zumeist sehr einfach sind („Deckeneffekte“).</p> <p><i>Ausführliche Beschreibung ab S. 58.</i></p>	

CFT 1; Skala 1 – Grundintelligenztest

<i>Altersbereich</i>	5;3–9;5 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Culture Fair Intelligence Test</p> <p>Basiert auf der Intelligenztheorie von Cattell</p> <p>Testaufbau: 5 Subtests:</p> <ul style="list-style-type: none"> - Substitutionen - Labyrinth - Klassifikationen - Ähnlichkeiten - Matrizen <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1976 Überprüfung der Normen 1995</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizienten zwischen $r = .65$ und $r = .86$</p> <p>Faktorielle Validität abgesichert</p> <p>Konstrukt- und Kriteriumsvalidität abgesichert durch Korrelationen mit anderen Tests und Schulnoten (mittlere Zusammenhänge)</p>
<p>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).</p> <p><i>Ausführliche Beschreibung ab S. 60.</i></p>	

CFT 20-R; Skala 2 – Grundintelligenztest

<i>Altersbereich</i>	8;5-60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Culture Fair Intelligence Test</p> <p>Test erfasst über figurale Aufgaben fluide Intelligenz (nach Cattell, 1971)</p> <p>Testaufbau: 2 gleichartig aufgebaute Testteile, je 4 Untertests:</p> <ul style="list-style-type: none"> - Reihen fortsetzen - Klassifikationen - Matrizen - topologische Schlussfolgerungen <p>Ergänzungstests: Wortschatz und Zahlenfolgen zur Erfassung der verbalen und numerischen Verarbeitungskapazität (nach Jäger)</p> <p>Versionen: Testteil 1 mit kurzer oder verlängerter Testzeit, Kurz- und Langform, Pseudoparallelforn</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2003/2004</p> <p>Reliabilität: gut ($\alpha = .96$)</p> <p>Faktorielle Validität: explorative Analysen ergeben 3 Faktoren, konfirmatorische Analysen belegen einen latenten Faktor g.</p> <p>Konstrukt- und Kriteriumsvalidität durch Korrelationen mit Tests und Schulnoten belegt. Bei Schulnoten höchste Zusammenhänge mit Note in Mathematik.</p>
<p>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) Kuhn, Holling & Freund (2008) dokumentieren, dass etliche Testitems für begabtere Schüler/innen (IQ>120) einen geringeren Informationsgehalt aufweisen als für weniger begabte Gruppen (niedrige Trennschärfe bei Begabten) und damit verbunden die Messgenauigkeit der Subtests für Begabte verringert ist.</p>	
<p><i>Ausführliche Beschreibung ab S. 63.</i></p>	

CMM 1-3 – Columbia Mental Maturity Scale 1-3

<i>Altersbereich</i>	6–10 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Ursprünglich zur Diagnose geistiger Behinderung eingesetzt. Testaufbau: Nur ein Aufgabentyp - Kinder wählen ein Bild aus, das nicht zu den übrigen passt. Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1971/1972 Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizienten zwischen $r = .87$ und $r = .88$ Validität: Konstruktvalidität geprüft durch Korrelationen mit anderen Tests (mittlere Korrelationen).
<p>Die CMM 1-3 eignet sich nur bedingt zur Testung von intellektueller Hochbegabung; 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 Aufgaben sind zudem wenig abwechslungsreich, weshalb Ermüdungseffekte nicht auszuschließen sind.</p> <p><i>Ausführliche Beschreibung ab S. 66.</i></p>	

CMM 1-4 – Columbia Mental Maturity Scale 1-4

<i>Altersbereich</i>	3–10 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Ursprünglich zur Diagnose geistiger Behinderung eingesetzt. Testaufbau: Nur ein Aufgabentyp - Eines aus 5 Bildern passt nicht zu den anderen, dieses soll identifiziert werden. Einzeltest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1968–1972 Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .88$ Konstrukt- und Kriteriumsvalidität abgesichert durch Korrelationen mit anderen Tests und Schulnoten (mittlere Zusammenhänge).
<p>Die CMM 1-4 eignet sich wie auch die CMM 1-3 nur bedingt zur Testung von intellektueller Hochbegabung; 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 Aufgaben sind zudem wenig abwechslungsreich, weshalb Ermüdungseffekte nicht auszuschließen sind.</p> <p><i>Ausführliche Beschreibung ab S. 67.</i></p>	

CPM – Coloured Progressive Matrices

<i>Altersbereich</i>	3;9–11;8 Jahre
<i>Struktur</i>	Beurteilt wird die kognitive Entwicklung bis max. zum Stadium, in dem das Kind beginnt, in Analogien zu denken.
<i>Kurzbeschreibung</i>	Testaufbau: 3 Sets mit je 12 Aufgaben. Kinder sollen zu einer unvollständigen farbigen geometrischen Figur die richtige Ergänzung aus 6 möglichen Vorgaben auswählen. Einzel- und Gruppentest
<i>Normen</i>	Neunormierung: 2002 Reliabilität: mäßig bis sehr gut
<i>Testgüte</i>	($\alpha = .67$ bis $\alpha = .93$) Validität: gegeben; Faktor simultanes Verarbeiten wird mit Ladungen zwischen $r = .75$ und $r = .85$ erfasst.
<p>Die CPM sind populäre Verfahren zur Erfassung der fluiden Intelligenz. Bezogen auf die Hochbegabungsdiagnostik ist zu beachten, dass die CPM wie auch die APM nur eine Intelligenzfazette – und dies mit zudem nur einem Aufgabentyp – erfassen. Vielfach werden Geschlechtereffekte zugunsten männlicher Testpersonen dokumentiert. Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den CPM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der CPM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe auch Preckel, 2010).</p> <p><i>Ausführliche Beschreibung ab S. 70.</i></p>	

FAT 6-9 – Frankfurter Analogietest 6-9

<i>Altersbereich</i>	6.–9. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Basiert auf der Theorie von Spearman</p> <p>Ermittelt Fähigkeit zum logisch-schlussfolgernden Denken</p> <p>Testaufbau: 50 Aufgaben vier verschiedener Typen:</p> <ul style="list-style-type: none"> - Wortanalogien - gegenständliche Bildanalogien - abstrakte Bildanalogien - gemischte Analogien <p>Aus 5 Wahlmöglichkeiten soll die passende gefunden werden.</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1982</p> <p>Reliabilität: mäßig bis gut (interne Konsistenzen zwischen $\alpha = .67$ und $\alpha = .94$)</p> <p>Konstruktvalidität gegeben: Es ergeben sich 4 Faktoren:</p> <ul style="list-style-type: none"> - verbal-logische Fähigkeiten - Raumvorstellung - allgemeine Schulleistung und - naturwissenschaftliches Verständnis <p>Kriteriumsvalidität gering</p>
<p>Der FAT ist für die Testung von intellektueller Hochbegabung ungeeignet, da die Aufgaben im oberen Leistungsbereich zu leicht sind und Deckeneffekte schon im oberen Durchschnittsbereich auftreten können. Die Validität ist zudem eher gering. Eine Kombination mit anderen Verfahren wird – um eine Diagnose stellen zu können – unerlässlich sein.</p> <p><i>Ausführliche Beschreibung ab S. 74.</i></p>	

FRT bzw. FRT-J – Figure Reasoning Test

<i>Altersbereich</i>	FRT-J: 10–14 Jahre FRT: 14–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Basiert auf der Theorie von Spearman Testaufbau: Test prüft das Erkennen von logischen Zusammenhängen: Es soll die passende Figur zu einem unvollständigen Muster aus 8 Figuren gewählt werden. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Neunormierung: 2004 Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .96$ Faktorielle Validität gegeben, g-Faktor klärt 80% der Leistung auf. Konstrukt- und Kriteriumsvalidität mit anderen Tests und Schulnoten geprüft (mittlere Korrelationen).
<p>Der FRT enthält nur einen Aufgabentyp, was die Diagnostik einer allgemeinen Intelligenz bzw. einer intellektuellen Begabung erschwert. Deckeneffekte im oberen Begabungsbereich können vielfach beobachtet werden. Oft werden auch Geschlechtereffekte zugunsten männlicher Testpersonen beobachtet. Aus diesem Grund sollten Tests mit unterschiedlichen Aufgabenfacetten dem FRT vorgezogen werden. Will man lediglich das Reasoning überprüfen, so eignen sich andere figurale Matrizentests besser als der FRT.</p> <p><i>Ausführliche Beschreibung ab S. 76.</i></p>	

HAWIE-R – Hamburg-Wechsler-Intelligenztest für Erwachsene

<i>Altersbereich</i>	16–74 Jahre
<p><i>Struktur</i></p> <p><i>Kurz- beschreibung</i></p>	<p>Erfassung der allgemeinen Intelligenz</p> <p>Testaufbau: 11 Untertests (6 Verbaltests und 5 Handlungstests)</p> <p>Sprachliche Intelligenz wird über folgende Untertests geprüft:</p> <ul style="list-style-type: none"> - Allgemeines Wissen - Zahlen nachsprechen - Wortschatz-Test - rechnerisches Denken - allgemeines Verständnis - Gemeinsamkeiten finden <p>Praktische Intelligenz wird über folgende Untertests geprüft:</p> <ul style="list-style-type: none"> - Bilder ergänzen - Bilder ordnen - Mosaik-Test - Figuren legen - Zahlen-Symbol-Test <p>Tests haben durch Zeitbegrenzungen eine Speed-Test-Komponente.</p> <p>Einzeltest</p>
<p><i>Normen</i></p> <p><i>Testgüte</i></p>	<p>Normierung: vorhanden (ohne Datum)</p> <p>Reliabilität hinreichend hoch ($\alpha = .71$ bis $\alpha = .96$)</p> <p>Faktorielle Validität kann angenommen werden.</p>
<p>Dem HAWIE-R sind andere Verfahren der Intelligenzmessung vorzuziehen. Zwar sind die Testaufgaben gut normiert und abwechslungsreich, dennoch kommen Testpersonen mit anderen Testmaterialien oft besser zurecht.</p> <p><i>Ausführliche Beschreibung ab S. 78.</i></p>	

HAWIK-III – Hamburg-Wechsler-Intelligenztest für Kinder III

<i>Altersbereich</i>	6;0–16;11 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testung im Sinne des Globalkonzepts von Wechsler</p> <p>Testaufbau: Erfasst werden über 13 spezifische Untertests die praktische, verbale und allgemeine Intelligenz.</p> <p>Weiters werden 4 Indizes ermittelt:</p> <ul style="list-style-type: none"> - sprachliches Verständnis - Wahrnehmungsorganisation - Unablenkbarkeit - Arbeitsgeschwindigkeit <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1995–1998</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient zwischen $r = .91$ und $r = .96$</p> <p>Faktorielle Validität: 2-Faktoren-Lösung</p> <p>Kriteriumsvalidität durch Korrelationen mit Schulnoten geprüft (mittlere Korrelationen).</p>
<p>Der HAWIK-III ist ein Standardinstrument der allgemeinen Intelligenzdiagnostik. Es erfährt einen weiten Anwendungsbereich, seine Differenzierungsstärke liegt allerdings im mittleren Leistungsbereich. Somit ergeben sich in den Extrembereichen der kognitiven Leistungsfähigkeit Boden- bzw. Deckeneffekte. Das heißt, die Items sind größtenteils zu schwer oder zu leicht, es können daher keine präzisen Aussagen 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.</p> <p><i>Ausführliche Beschreibung ab S. 80.</i></p>	

HAWIK-IV – Hamburg-Wechsler-Intelligenztest für Kinder IV

<i>Altersbereich</i>	6–16;11 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Basiert theoretisch auf dem Cattell-Horn-Caroll-Modell</p> <p>Testaufbau: 15 Untertests (davon 5 optional): Ermittelt werden allgemeine Intelligenz und 4 Subindizes:</p> <ul style="list-style-type: none"> - Sprachverständnis - wahrnehmungsgebundenes, logisches Denken - Arbeitsgedächtnis - Verarbeitungsgeschwindigkeit <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2005/2006</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .97$</p> <p>Faktorielle Validität belegt</p> <p>Konstrukt- und Kriteriumsvalidität durch Korrelationen mit anderen Testverfahren und Schulnoten überprüft.</p>
<p>Relevant für die Hochbegabungsdiagnostik ist, dass im Bereich hoher Begabung nicht selten extreme Leistungsunterschiede zwischen den vier Indexwerten bzw. den beiden Indexpaaren (Sprachverständnis – logisches Denken vs. Arbeitsgedächtnis – Verarbeitungsgeschwindigkeit) vorkommen, was die Interpretierbarkeit des Gesamt-IQs erschwert (Daseking, Petermann & Petermann, 2007). Zudem zeigt sich in Validierungsstudien zur Hochbegabung, dass Kinder, welche in anderen Verfahren einen IQ>130 erzielten, im HAWIK-IV einen niedrigeren IQ erreichten. Bei Hochbegabten empfiehlt sich daher die Verwendung des Allgemeinen Fähigkeitsindex alternativ zum Gesamt-IQ, da für 2/3 der Begabten (IQ>120) der AFI höher ausfällt als der Gesamt-IQ. Der Zusammenhang mit der fluiden Intelligenz 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.</p> <p><i>Ausführliche Beschreibung ab S. 83.</i></p>	

HAWIVA III – Hannover-Wechsler-Intelligenztest für das Vorschulalter III

<i>Altersbereich</i>	2;6–6;11 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Auf Basis der Wechsler Preschool and Primary Scale of Intelligence</p> <p>Testaufbau: Neben den allgemeinen intellektuellen Fähigkeiten eines Kindes werden Werte für folgende Skalen ermittelt:</p> <ul style="list-style-type: none"> - Verbalteil - Handlungsteil - Verarbeitungsgeschwindigkeit - Allgemeine Sprachskala <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2004/2005</p> <p>Reliabilität: gut bis sehr gut ($\alpha = .89$ bis $\alpha = .95$)</p> <p>Es ergeben sich 2 (Altersbereich 2;6–3;11 Jahre) bzw. 3 Faktoren (Altersbereich 4;0–7;3 Jahre).</p> <p>Konstrukt- und Kriteriumsvalidität durch Korrelationen mit anderen Testverfahren und Lehrer/innenurteilen überprüft (mittlere Korrelationen).</p>
<p>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 CFT 1-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)).</p> <p><i>Ausführliche Beschreibung ab S. 87.</i></p>	

HIT 1-2 und HIT 3-4 – Heidelberger Intelligenztest für 1. und 2. bzw. 3. und 4. Klassen

<i>Altersbereich</i>	HIT 1-2: 7–8 Jahre HIT 3-4: 9–10 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Testaufbau: 5 Teiltests mit insgesamt 71 Aufgaben: <ul style="list-style-type: none"> - Relationen - Logik - Reproduktion - Differenzierung - Kreativität <p>Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	Normierung: 1976 (HIT 1-2) 1980 (HIT 3-4) Reliabilität: gut ($\alpha = .85$ bei HIT 1-2; $\alpha = .92$ bei HIT 3-4) Faktorielle Validität bis auf die Items zur Kreativität gegeben. Kriteriumsvalidität bis auf Skala Kreativität geprüft (geringe bis mittlere Zusammenhänge)
<p>Der HIT 1-2 und auch der HIT 3-4 eignen sich gut für ein Intelligenzscreening in Form einer Gruppentestung im Schulkontext, lassen aber Individualdiagnosen nur bedingt zu. Um eine verlässliche Intelligenzdiagnostik machen zu können, ist es sowohl für den oberen als auch für den unteren Leistungsbereich unabdingbar, weitere Tests einzusetzen.</p> <p><i>Ausführliche Beschreibung ab S. 91 bzw. S. 95.</i></p>	

IBF – Intelligenz-Basis-Faktoren

<i>Altersbereich</i>	14–60+ Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Erfassung intellektueller Leistungsfähigkeit Testaufbau: 2 Versionen (leicht und schwer) Jeweils 7 Untertests (drei sprachliche, zwei mathematisch-numerische, ein Test für das räumliche Vorstellungsvermögen und ein Gedächtnistest) mit 12 Aufgaben Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 2004 Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient zwischen $r = .73$ und $r = .91$ Faktorielle Validität gegeben
<p>Der IBF liefert Informationen zu individuellen Stärken und Schwächen. Als alleiniges Intelligenzdiagnostikum eignet er sich nur bedingt.</p> <p><i>Ausführliche Beschreibung ab S. 95.</i></p>	

IDS – Intelligence and Development Scales

<i>Altersbereich</i>	5–10 Jahre
<i>Struktur</i> <i>Kurz-</i> <i>beschreibung</i>	<p>Testaufbau: 7 kognitive Untertests, die den vier Bereichen</p> <ul style="list-style-type: none"> - Wahrnehmung - Aufmerksamkeit - Gedächtnis - Denken <p>zugeordnet werden und sich zu einem Intelligenzwert verrechnen lassen:</p> <ul style="list-style-type: none"> - Wahrnehmung visuell - Aufmerksamkeit selektiv - Gedächtnis phonologisch - Gedächtnis räumlich-visuell - Gedächtnis auditiv - Denken bildlich - Denken konzeptuell <p>Allgemeine Entwicklung wird in den IDS mit den folgenden fünf Funktionsbereichen erhoben:</p> <ul style="list-style-type: none"> - Psychomotorik - sozial-emotionale Kompetenz - Mathematik - Sprache - Leistungsmotivation <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2007/2008</p> <p>Reliabilität mäßig bis sehr gut (bei den kognitiven Untertests zwischen $\alpha = .68$ und $\alpha = .96$, bei den Tests zur allgemeinen Entwicklung zwischen $\alpha = .57$ und $\alpha = .88$; für den Intelligenzwert sehr hoch mit $\alpha = .92$)</p> <p>Konstruktvalidität nachgewiesen</p> <p>Kriteriumsvalidität geprüft (mittlere Korrelationen)</p>
<p>Die IDS können für die Leistungsdiagnostik im Alter von 5 bis 8 Jahren durchaus empfohlen werden (obwohl die Altersspanne bis 10 Jahre angegeben ist). Für die Diagnostik im oberen Extrembereich eignen sie sich eher nicht.</p> <p><i>Ausführliche Beschreibung ab S. 97.</i></p>	

ISA – Intelligenz Struktur Analyse

<i>Altersbereich</i>	14–55 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Basiert auf dem Berliner Intelligenzstrukturmodell von Jäger et al.</p> <p>Erfasst allgemeine Intelligenz sowie die Fähigkeitsbereiche „verbale Intelligenz“, „numerische Intelligenz“, „figural-räumliche Intelligenz“ und „Gedächtnisleistung“</p> <p>Testaufbau: Mit neun Subtests sollen die vier oben genannten Fähigkeitsbereiche gemessen werden. Insgesamt 177 Aufgaben, die den vier Fähigkeitsbereichen zugeordnet werden:</p> <ul style="list-style-type: none"> - Verbale Intelligenz oder Verarbeitungskapazität - Numerische Intelligenz oder Verarbeitungskapazität - Figurale Intelligenz oder Verarbeitungskapazität - Merkfähigkeit oder Gedächtnis <p>Zwei Kurzformen: ISA-L (leicht) ISA-S (schwer) Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1997–1999</p> <p>Reliabilität: gut/sehr gut ($\alpha = .87$ bis $\alpha = .89$ für die Subtests, $\alpha = .97$ für den Gesamtwert)</p> <p>Faktorielle Validität nachgewiesen</p> <p>Konstruktvalidität geprüft (mittlere Korrelationen)</p>
<p>Die ISA eignet sich sehr gut für die Messung intellektueller Begabung gesamt und in Teilgebieten auch im oberen Bereich (verbale numerische und figurale Intelligenz). Deckeneffekte können nur in den Extrembereichen beobachtet werden.</p> <p><i>Ausführliche Beschreibung ab S. 101.</i></p>	

I-S-T 2000 R – Intelligenz-Struktur-Test 2000 R



<i>Altersbereich</i>	15–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Erfasst verschiedene Intelligenzbereiche sowie Wissen und schlussfolgerndes Denken</p> <p>Testaufbau: Grundmodul besteht aus 11 Untertests mit je 20 Aufgaben zu</p> <ul style="list-style-type: none"> - verbaler Intelligenz - numerischer Intelligenz - figuraler Intelligenz - schlussfolgerndem Denken - Merkfähigkeit <p>Ein Erweiterungsmodul besteht aus einem Test zur Erfassung des erworbenen Wissens und soll fluide und kristallisierte Intelligenz erfassen.</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: vorhanden (keine Jahresangabe)</p> <p>Reliabilität: sehr gut ($\alpha = .87$ bis $\alpha = .97$)</p> <p>Faktorielle Validität nachgewiesen</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (mittlere Korrelationen)</p>
<p>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.</p> <p><i>Ausführliche Beschreibung ab S. 103.</i></p>	

K-ABC – Kaufman-Assessment Battery for Children (Deutsche Version)

<i>Altersbereich</i>	2;6–12;5 Jahre
<i>Struktur</i> <i>Kurzbeschreibung</i>	<p>Basiert theoretisch auf dem Modell von Horn & Cattell</p> <p>Testaufbau: 16 Untertests, die 4 Skalen zugeordnet sind: Die Skala „einzelheitliches Denken“ und die Skala „ganzheitliches Denken“ werden zur Skala „intellektuelle Fähigkeiten“ aggregiert (fluide Intelligenz).</p> <p>Eine Fertigkeitsskala erfasst die kristalline Intelligenz in Form von Faktenwissen, Sprachkonzepten und schulbezogenen Fertigkeiten.</p> <p>Ab 4 Jahren kann eine sprachfreie Skala für hörgeschädigte, sprach- und sprechgestörte Kinder sowie fremdsprachige Kinder zusammengestellt werden.</p> <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1986–1989</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient zwischen $r = .83$ und $r = .98$</p> <p>Faktorielle Validität durch explorative und konfirmatorische Analysen nachgewiesen.</p> <p>Konstruktvalidität durch Korrelationen mit Tests überprüft (mittlere Zusammenhänge).</p>
<p>Die K-ABC wird nach einer Umfrage aus dem Jahr 2009 an verschiedenen begabungspsychologischen Stellen in Österreich und Deutschland häufig für die Intelligenzdiagnostik 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 für die Diagnostik hoher Intelligenz 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).</p> <p><i>Ausführliche Beschreibung ab S. 105.</i></p>	

KFT 1-3 – Kognitiver Fähigkeitstest für 1.–3. Klasse (Grundschulform)

<i>Altersbereich</i>	1.–3. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Es sollen in differenzieller Weise schulisch relevante intellektuelle Lern- und Leistungsvoraussetzungen erfasst werden.</p> <p>Testaufbau: Insgesamt 60 Aufgaben zu</p> <ul style="list-style-type: none"> - Sprachverständnis - Beziehungserkennen - Schlussfolgerndes Denken - Rechnerisches Denken <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1978/1979</p> <p>Reliabilität: mäßig bis gut ($\alpha=.56$ bis $\alpha=.77$ bei den Subtests, $\alpha= .85$ für den Gesamttest)</p> <p>Faktorielle Validität mäßig</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (geringe bis mittlere Korrelationen)</p>
<p>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.</p> <p><i>Ausführliche Beschreibung ab S. 109.</i></p>	

KFT 4-12+R – Kognitiver Fähigkeitstest für 4.–12. Klasse

<i>Altersbereich</i>	4.–12. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	Erfassung der Verarbeitungskapazität Testaufbau: <ul style="list-style-type: none"> - insgesamt neun Subtests und 507 Aufgaben, die sich drei Testteilen (verbal, quantitativ, nonverbal) zuordnen lassen - zudem ab der 5. Schulstufe Inventar zur Erfassung des räumlichen Vorstellungsvermögens Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1996/1997 und aus 2002/2003. Reliabilität: gut ($\alpha = .79$) Faktorielle Validität durch explorative Analysen belegt. Konstrukt- und Kriteriumsvalidität geprüft (keine bis mittlere Korrelationen mit Schulnoten)
<p>Die nicht speziell für die Hochbegabungsdiagnostik vorgesehenen KFT-Versionen, insbesondere der KFT 4-12+R, werden zwar häufig auch für die Hochbegabungsdiagnostik eingesetzt (Weiß, 2008). In diesem Bereich sollte aber jedenfalls den KFT-Versionen aus dem MHBT aufgrund der höheren Aufgabenschwierigkeit und den aktuelleren Normdaten der Vorzug gegeben werden. Die MHBT enthält Versionen des KFT zur Hochbegabungsdiagnostik.</p> <p><i>Ausführliche Beschreibung ab S. 111.</i></p>	

KFT-K – Kognitiver Fähigkeitstest (Kindergarten)

<i>Altersbereich</i>	5–6 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Verfahren erfasst kognitive Fähigkeiten im Vorschulalter. Testaufbau: 4 Subtests mit je 15 Mehrfachwahl-Aufgaben zu <ul style="list-style-type: none"> - Sprachverständnis - Beziehungserkennen - Schlussfolgerndem Denken - Rechnerischem Denken Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1979–1981 Reliabilität: mäßig bis gut (Gesamttest $\alpha=.90$) Konstruktvalidität überprüft (mittlere Korrelationen)
<p>Der KFT-K ist ein geeignetes Instrument zur Feststellung der kognitiven Fähigkeiten von Kindergartenkindern und eignet sich auch gut zur Beantwortung der <i>Frage nach einer früheren Einschulung</i>. 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.</p> <p><i>Ausführliche Beschreibung ab S. 115.</i></p>	

KLI 4-5 / KLI 4-5 R – Kombinerter Lern- und Intelligenztest für 4. und 5. Klassen

<i>Altersbereich</i>	10–11 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testaufbau: Der Test teilt sich in zwei Teile: Der <i>Lernteil</i> bezieht sich auf das Erlernen und Anwenden von Regeln in der verbalen Kommunikation und im numerischen Bereich. Der „Behaltentest“ ermittelt das Behalten des in Geheimschrift und Zahlenumwandeln Gelernten. Der <i>Intelligenzteil</i> ermittelt im Sprachbereich Abstraktionsfähigkeit und Leseverständnis, im Zahlenbereich logisches Schließen und Kritikfähigkeit.</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2004</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .95$</p> <p>Validität: hinreichend hoch</p>
<p>Der KLI 4-5 ist für eine verlässliche Intelligenzdiagnostik nur bedingt geeignet. Er erlaubt einen guten Überblick über die Fähigkeiten der Kinder, lässt aber nur wenige Aussagen über die Gesamtleistungsfähigkeit zu.</p> <p><i>Ausführliche Beschreibung ab S. 117.</i></p>	

K-TIM – Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene

<i>Altersbereich</i>	11–80+ Jahre
<p><i>Struktur</i></p> <p><i>Kurz- beschreibung</i></p>	<p>Umfassender Intelligenztest, der auf der Theorie fluider gegenüber kristalliner Intelligenzanteile beruht.</p> <p>Testaufbau: 8 Untertests zur Intelligenzmessung sowie zwei weitere, fakultativ durchzuführende Untertests zur Beurteilung verzögerter gegenüber unmittelbarer Gedächtnisleistungen.</p> <p>Die 8 Untertests sind:</p> <ul style="list-style-type: none"> - Worträtsel - Auditives Verständnis - Doppelte Bedeutungen - Persönlichkeiten - Symbole lernen - Logische Denkschritte - Zeichen entschlüsseln - Figurales Gedächtnis <p>Die fakultativen Untertests sind:</p> <ul style="list-style-type: none"> - Symbole – Abruf nach Intervall - Auditives Verständnis – Abruf nach Intervall <p>Einzeltest</p>
<p><i>Normen</i></p> <p><i>Testgüte</i></p>	<p>Normiert</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .95$</p> <p>Faktorielle Validität nachgewiesen</p> <p>Konstruktvalidität überprüft (mittlere bis hohe Korrelationen)</p>
<p>Der K-TIM ist zeitaufwändig, erlaubt aber eine verlässliche, umfassende und objektive Aussage über die intellektuellen Fähigkeiten der getesteten Personen und ist zur <i>Diagnostik auch im oberen Intelligenzbereich</i> zu empfehlen (jedoch nicht im oberen Extrembereich).</p> <p><i>Ausführliche Beschreibung ab S. 119.</i></p>	

MHBT-P inkl. KFT-HB 3 und KFT HB-4

– Münchner Hochbegabungstestbatterie für die Primarstufe



<i>Altersbereich</i>	1.–4. Schulstufe (KFT-HB 3+4 für 3. und 4. Schulstufe)
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Beinhaltet folgende Skalen:</p> <ul style="list-style-type: none"> - Kognitive Fähigkeiten (KFT-HB 3+4 für Schüler/innen sowie Lehrer/innen-Checkliste) - Kreativität (Schüler/innen-Test und Fragebogen sowie Lehrer/innen-Checkliste) - Soziale Kompetenzen (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste) - Motivation (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste) - Arbeitsverhalten (Schüler/innen-Fragebogen und Lehrer/innen-Checkliste) - Psychomotorische Fähigkeiten (Lehrer/innen-Checkliste) - Musikalische Fähigkeiten (Lehrer/innen-Checkliste) - Einzeltest <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2002/2003</p> <p>Reliabilität: mäßig bis sehr gut ($\alpha < .60$ bis $\alpha = .95$)</p> <p>Faktorielle Validität nachgewiesen</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (mittlere Zusammenhänge)</p>
<p>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.</p> <p>Das Material ist ansprechend und abwechslungsreich. Die Kinder haben erfahrungsgemäß Spaß an den Aufgaben, lösen diese gerne und langweilen sich nicht</p> <p><i>Ausführliche Beschreibung ab S. 121.</i></p>	

MHBT-S – Münchner Hochbegabungstestbatterie für die Sekundarstufe



<p><i>Altersbereich</i></p>	<p>4.–12. Schulstufe</p>
<p><i>Struktur</i> <i>Kurz- beschreibung</i></p>	<p>Enthält folgende Verfahren: Kognitiver Fähigkeitstest für Hochbegabte (KFT-HB 4-12+)</p> <p>Testaufbau: Fragebogen zur Erfassung wichtiger Begabungskomponenten wie</p> <ul style="list-style-type: none"> - Abwicklungen - Spiegelbilder - Erfassung physikalisch-technischer Problemlösekompetenzen - Soziale Kompetenzen - Kreativität - Interessen - Erkenntnisstreben - Leistungsmotivation - Arbeitsverhalten - Schulklima - Familienklima <p>Zusätzlich fünf Checklisten für Lehrkräfte zu den wichtigsten Hochbegabungsbereichen:</p> <ul style="list-style-type: none"> - Intelligenz - Kreativität - Musikalität - Sozialbegabung - Psychomotorik <p>Einzeltest</p>
<p><i>Normen</i> <i>Testgüte</i></p>	<p>Normierung: 1996/1997</p> <p>Reliabilität: hoch ($\alpha = .92$ bis $\alpha = .94$)</p> <p>Faktorielle Validität gut</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (geringe bis mittlere Korrelationen)</p>
<p>Gerade im Bereich über dem Durchschnitt und im oberen Extrembereich eignet sich die MHBT sehr gut für die Erstellung einer verlässlichen Hochbegabungsdiagnostik. 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 und Jugendlichen haben erfahrungsgemäß Spaß an den Aufgaben, lösen diese gerne und langweilen sich nicht.</p> <p><i>Ausführliche Beschreibung ab S. 124.</i></p>	

MIT – Mannheimer Intelligenztest

<i>Altersbereich</i>	12–45 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Sowohl auf der Basis des globalen Intelligenzmodells als auch nach dem Modell einer Kombination von teilweise unabhängigen Teilfähigkeiten entwickelt. Definition von Intelligenz orientiert sich an den Konzepten von Thurstone (1938), Guilford (1964) und Jäger (1967).</p> <p>Testaufbau: 10 Untertests mit sprachfreien und sprachgebundenen Aspekten der Intelligenz:</p> <ul style="list-style-type: none"> - Figurenreihen - Wortbedeutungen - Dominos - Buchstabengruppen - Zahlenreihen - Wortverhältnisse - Mosaik - Sprichwörter - Zahlensymbole - Unmöglichkeiten <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1969/1970</p> <p>Reliabilität: sehr gut ($\alpha = .96$ bis $\alpha = .97$)</p> <p>Faktorielle Validität nachgewiesen</p> <p>Konstrukt- und Kriteriumsvalidität geprüft (geringe bis mittlere Korrelationen)</p>
<p>Der MIT ist ein brauchbares Verfahren zur Messung von intellektuellen Fähigkeiten. Am besten misst er im Durchschnittsbereich bis zu 1,5 Standardabweichungen vom Mittelwert.</p> <p><i>Ausführliche Beschreibung ab S. 127.</i></p>	

MWT-A; MWT-B – Mehrfachwahl-Wortschatz-Test

<i>Altersbereich</i>	18–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Kurztest Testaufbau: Umgangssprachliches bzw. wissenschaftliches Wort soll aus Auswahl von verschiedenen sinnlosen Wörtern identifiziert werden. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1992 Reliabilität: interne Konsistenz nicht angegeben, Retest-Reliabilität $r = .96$ Validität durch Korrelation mit LPS geprüft ($r = .74$)
<p>Der MWT eignet sich nicht zur Diagnostik allgemeiner Intelligenz, da er keine zuverlässige und differenzierte Aussage über die Intelligenz der Testperson leisten kann. Als Screening ist er durchaus empfehlenswert.</p> <p><i>Ausführliche Beschreibung ab S. 129.</i></p>	

NNAT – Naglieri Nonverbal Ability Test

<i>Altersbereich</i>	5–17 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Messung schlussfolgernden Denkens Testaufbau: 4 Itemtypen: - Muster komplettieren - Analogien - serielles Schlussfolgern - räumliche Visualisierung Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung für deutschsprachige Länder nicht vorhanden. Reliabilität: mäßig bis sehr gut ($\alpha = .64$ bis $\alpha = .96$) Validität durch Vergleich mit K-ABC, HAWIK III geprüft (US-Stichprobe).
<p>Der NNAT erlaubt eine relativ schnelle und dennoch gute Aussage zur allgemeinen Intelligenz, welche aber oft überschätzt wird. Für die Diagnostik von Hochbegabung ist daher unbedingt eine Testung mit einem anderen Intelligenztest (MHBT, AID 2 u.a.) zu empfehlen.</p> <p><i>Ausführliche Beschreibung ab S. 131.</i></p>	

PPVT – Peabody Picture Vocabulary Test

<i>Altersbereich</i>	14–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Überprüft (passiven) Wortschatz und Sprachverständnis Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung vorhanden Reliabilität: sehr gut ($\alpha = .93$) Konstruktvalidität durch Korrelationen mit anderen Testverfahren geprüft (mittlere bis hohe Korrelationen).
<p>Der PPVT ist für die Diagnostik von intellektueller Begabung ungeeignet und kann für die psychologische Begabungsdagnostik nicht empfohlen werden.</p> <p><i>Ausführliche Beschreibung ab S. 133.</i></p>	

PSB-R 4-6 – Prüfsystem für Schul- und Bildungsberatung für 4. bis 6. Klassen

<i>Altersbereich</i>	4.–6. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>In der Tradition des Thurstone'schen Intelligenzkonzepts entwickelt.</p> <p>Testaufbau: 10 Subtests zu</p> <ul style="list-style-type: none"> - Allgemeinwissen - Zahlenreihen - Buchstabenreihen - Figurale Reihen - Wortflüssigkeit - Gliederungsfähigkeit - Raumvorstellung - Gemeinsamkeiten finden - Zahlenaddition - Zahlenvergleich <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1999/2000</p> <p>Reliabilität: mäßig bis sehr gut ($\alpha = .72$ bis $\alpha = .94$)</p> <p>Faktorielle Validität: nur bedingt nachgewiesen</p> <p>Konstrukt- und Kriteriumsvalidität: geprüft (geringe bis mittlere Korrelationen)</p>
<p>Der PSB-R 4-6 ist stark an der Schulleistung orientiert. Für eine differenzierte Begabungsdiagnostik ist dieser Test weniger geeignet.</p> <p><i>Ausführliche Beschreibung ab S. 135.</i></p>	

PSB-R 6-13 – Prüfsystem für Schul- und Bildungsberatung für 6. bis 13. Klassen

<i>Altersbereich</i>	6.–13. Schulstufe
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>In der Tradition des Thurstone’schen Intelligenzkonzepts entwickelt.</p> <p>Testaufbau: 9 Untertests zu</p> <ul style="list-style-type: none"> - Allgemeinwissen - Zahlreihen (Reasoning) - Buchstabenreihen (Reasoning) - Figurale Reihen (Reasoning) - Wortflüssigkeit - Raumvorstellung - Gemeinsamkeiten finden - Zahlenaddition – Konzentration - Zahlenvergleich – Konzentration <p>Die neun Untertests werden drei Faktoren (Verbal-, Reasoning- und Konzentrationsfaktor) zugeordnet.</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 2001</p> <p>Reliabilität: interne Konsistenz gut ($\alpha = .79$ bis $\alpha = .91$)</p> <p>Faktorielle Validität: bestätigt (3-Faktoren-Lösung)</p> <p>Konstrukt- und Kriteriumsvalidität: geprüft (mittlere bis hohe Korrelationen)</p>
<p>Das PSB-R 6-13 eignet sich als Screeninginstrument zur Erfassung intellektueller Begabung vorrangig im mittleren Bereich. Eine verlässliche Intelligenzdiagnostik, besonders im oberen Bereich, ist aufgrund der gewonnenen Ergebnisse des PSB-R 6-13 nicht möglich.</p> <p><i>Ausführliche Beschreibung ab S. 138.</i></p>	

SCH-F – Schlauchfiguren-Test

<i>Altersbereich</i>	15–20 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Überprüft wird die Raumvorstellung als Primärfaktor der Intelligenz. Testaufbau: Es werden Fotos von Würfeln mit verdrehten Schläuchen vorgelegt. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1982 Reliabilität: gut ($\alpha = .80$) Konvergente und diskriminante Validität geprüft (geringe bis mittlere Korrelationen).
<p>Bezogen auf die Hochbegabungsdiagnostik ist zu beachten, dass die Schlauchfiguren nur eine Intelligenzfacette (räumliches Vorstellungsvermögen) – mit nur einem Aufgabentyp – erfassen. Geschlechtereffekte zugunsten männlicher Testpersonen gibt es eindeutig, dies lässt sich u.U. durch den Vorteil von männlichen Personen im räumlichen Denken erklären, welches positiv mit der Testleistung der Schlauchfiguren korreliert. Dadurch, dass ausschließlich das räumliche Denken abgeprüft wird, sind weibliche Probandinnen stark benachteiligt. Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den Schlauchfiguren vorgezogen werden bzw. dieser Test nur als Zusatzinstrument zur Erhebung der räumlichen Vorstellungskraft eingesetzt werden.</p> <p><i>Ausführliche Beschreibung ab S. 141.</i></p>	

SON-R 2½-7 – Snijders-Oomen Non-verbaler Intelligenztest

<i>Altersbereich</i>	2;6–7;12 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Sprachfreier Intelligenztest zur Erfassung der allgemeinen Intelligenz (kein explizites Intelligenzmodell als Basis); nonverbale Instruktion möglich. Testaufbau: 6 Untertests (jeweils zwischen 14 und 17 Items), die zur Denkskala (fluide Intelligenz ohne räumliche Komponente) und zur Handlungsskala (fluide Intelligenz mit räumlicher Komponente) zusammengefasst werden. Einzeltest
<i>Normen</i> <i>Testgüte</i>	Normierung: 2005 Reliabilität: gut ($\alpha = .90$) Faktorielle und Konstruktvalidität sind belegt.
<p>Der SON-R 2½-7 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. Der SON-R 2½-7 ist insbesondere für die Anwendung mit sprachbeeinträchtigten, entwicklungsverzögerten oder schwach begabten Kindern konstruiert. Daher sind Deckeneffekte bei der Testung – v.a. älterer – hochintelligenter Kinder zu erwarten. Beim Einsatz mit jüngeren Kindergartenkindern ist zudem zu beachten, dass in diesem Alter erhobene Intelligenzschätzungen eher instabil sind und eine Intelligenzdiagnostik 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 SON-R 2½-7 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.</p> <p><i>Ausführliche Beschreibung ab S. 143.</i></p>	

SON-R 5½-17 – Snijders-Oomen Non-verbaler Intelligenztest

<i>Altersbereich</i>	5;6–17 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Sprachfreier Intelligenztest zur Erfassung der allgemeinen Intelligenz (kein explizites Intelligenzmodell als Basis); nonverbale Instruktion möglich.</p> <p>Testaufbau: 7 Subtests:</p> <ul style="list-style-type: none"> - Kategorien - Analogien - Situationen - Bildgeschichten - Mosaik - Zeichenmuster - Suchbilder <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1984/1985</p> <p>Reliabilität: gut ($\alpha = .85$)</p> <p>Faktorielle Validität bestätigt (Einfaktorenstruktur)</p> <p>Kriteriumsvalidität gering</p>
<p>Auch in der Altersvariante für das Schulalter liegt mit dem SON-R 5½-17 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 von den Anwender/innen 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), sodass 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 5½ -7 alleine eher nicht zu empfehlen.</p> <p><i>Ausführliche Beschreibung ab S. 145.</i></p>	

SPM – Standard Progressive Matrices

<i>Altersbereich</i>	6–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Erfasst Intelligenz über figurale Aufgaben, in denen jeweils ein Muster vervollständigt werden muss, sowie über klassische Matrizenaufgaben; mittelschwere Version der Raven Progressive Matrices. Testaufbau: 60 Items, die in 5 Sets A-E zusammengefasst sind. Einzel- und Gruppentest.
<i>Normen</i> <i>Testgüte</i>	Normierung: 1996/1997 Reliabilität mäßig bis sehr gut ($\alpha = .75$ bis $\alpha = .97$) Konstrukt- und Kriteriumsvalidität durch Korrelationen mit Tests und Schulnoten belegt. Bei Schulnoten höchste Zusammenhänge mit Note in Mathematik.
<p>Bezogen auf die Intelligenzdiagnostik ist zu beachten, dass die SPM (wie auch die APM) nur eine Intelligenzfacette mit nur einem Aufgabentyp erfassen. Beim Einsatz mit hochintelligenten Schülerinnen und Schülern findet sich bei den SPM noch mehr als bei den APM ein deutlicher Deckeneffekt. Geschlechtereffekte zugunsten männlicher Testpersonen sind dokumentiert. Kubinger (1991) stellte anhand von Analysen nach dem Rasch-Modell zudem fest, dass von insgesamt 60 Aufgaben 43 nicht eindimensional zu den anderen skaliert sind. Um die testspezifische Fehlervarianz gering zu halten, sollten Tests mit unterschiedlichen Aufgabenfacetten den SPM vorgezogen werden. Der alleinige Einsatz figuraler Reasoningtests wie der SPM ist für die Intelligenzdiagnostik im Kontext schulischer Begabungsförderung nicht optimal (siehe Preckel, 2010).</p> <p><i>Ausführliche Beschreibung ab S. 148.</i></p>	

WIE – Wechsler Intelligenztest für Erwachsene

<i>Altersbereich</i>	16–89 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Testaufbau: Ermittelt werden mittels 14 Subtests vier Indizes zur Abbildung kognitiver Teilfunktionen:</p> <ul style="list-style-type: none"> - Sprachliches Verständnis - Wahrnehmungsorganisation - Arbeitsgedächtnis - Arbeitsgeschwindigkeit <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: 1999–2005</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient zwischen $r = .75$ und $r = .92$</p> <p>Faktorielle Validität nicht für alle Altersstufen gegeben.</p>
<p>Der WIE eignet sich hervorragend für die Diagnostik von Intelligenz im mittleren und höheren Bereich (bis zu 1,5 Standardabweichungen über dem Mittelwert).</p> <p><i>Ausführliche Beschreibung ab S. 151.</i></p>	

WIT-2 – Wilde-Intelligenz-Test 2

<i>Altersbereich</i>	14–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Erhebt kognitive Leistungsfähigkeit. Basiert auf der Intelligenzkonzeption von Thurstone.</p> <p>Testaufbau: 11 Subtests, die den Modulen</p> <ul style="list-style-type: none"> - sprachliches Denken - rechnerisches Denken - Merkfähigkeit - Arbeitseffizienz - Wissen - Wissen zu Informationstechnologien <p>zugeordnet werden können.</p> <p>Einzel- und Gruppentest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: vorhanden (ohne Datum)</p> <p>Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .98$</p> <p>Konstruktvalidität gesichert</p> <p>Kriteriumsvalidität: geprüft (mittlere Korrelationen)</p>
<p>Der WIT-2 ist ein brauchbares Verfahren zur Messung von intellektuellen Fähigkeiten. Am besten misst er im Durchschnittsbereich bis 1,5 Standardabweichungen über dem Mittelwert.</p> <p><i>Ausführliche Beschreibung ab S. 155.</i></p>	

WMT – Wiener Matrizen-Test

<i>Altersbereich</i>	14–40 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Basiert auf dem Intelligenzmodell von Spearman. Testaufbau: Misst anhand von 24 Aufgaben schlussfolgerndes Denken im Umgang mit abstrakten Symbolen. Eine Reihe von abstrakten Symbolen soll fortgesetzt werden. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1979 Reliabilität: durch Gültigkeit des Rasch-Modells nachgewiesen Konstrukt- und Kriteriumsvalidität geprüft Geringe bis mittlere Korrelationen mit anderen Tests und Schulnoten
<p>Der Wiener Matrizen-Test besteht aus 24 figuralen Mehrfachwahl-Matrizenaufgaben mit jeweils acht Auswahlantworten. Die Distraktoren dieses Tests sind nicht sonderlich effektiv: Bei mehr als 2/3 der Aufgaben – Mittring & Rost (2008) errechnen sogar 75% – kann mit einer recht einfachen Strategie (Abzählen von Matrixkomponenten, Teil- und Schnittmengenbildung) die jeweilige Lösung ohne Inspektion des Item-Stamms, also ohne das Finden der der Bezugs-Matrix zugrunde liegende(n) generellen Regel(n), gefunden werden. Es kann angenommen werden, dass die mangelnde Effizienz der Distraktoren die Konstruktvalidität dieses Tests besonders im Bereich Hochbegabung deutlich beeinträchtigt.</p> <p><i>Ausführliche Beschreibung ab S. 159.</i></p>	

WNV – Wechsler Nonverbal Scale of Ability

<i>Altersbereich</i>	4–22 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	<p>Basiert auf dem Konzept von Wechsler und misst allgemeine kognitive Fähigkeiten.</p> <p>Testaufbau: 6 Subtests:</p> <ul style="list-style-type: none"> - Matrizen - Coding - Object Assembly - Recognition - Spatial Span - Picture Arrangement <p>Einzeltest</p>
<i>Normen</i> <i>Testgüte</i>	<p>Normierung: vorhanden (ohne Datum)</p> <p>Reliabilität: hoch ($\alpha = .91$)</p> <p>Faktorielle Validität nicht abgesichert</p> <p>Konstruktvalidität überprüft (mittlere Korrelationen)</p>
<p>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 zur Testung der spezifischen Intelligenz empfohlen werden. Für den Extrembereich ist es ratsam, noch weitere Verfahren einzusetzen, da hier die Messfehler größer werden.</p> <p><i>Ausführliche Beschreibung ab S. 161.</i></p>	

WPPSI-III – Wechsler Preschool and Primary Scale of Intelligence.
Deutsche Version



<i>Altersbereich</i>	3;0–7;2 Jahre
<i>Struktur</i>	Erfasst allgemeine und spezifische intellektuelle Fähigkeiten. Es kann der Gesamt-IQ als Maß für den kognitiven Entwicklungsstand eines Kindes im Vorschulalter bestimmt werden. Zusätzlich können vier weitere übergeordnete Werte berechnet werden: Verbal- und Handlungsteil, Verarbeitungsgeschwindigkeit und Allgemeine Sprachskala.
<i>Kurzbeschreibung</i>	Testaufbau: Testbatterie besteht aus 14 Untertests, die sich in drei Gruppen aufteilen lassen: - Kerntests - optionale Untertests - zusätzliche Untertests Einzeltest
<i>Normen</i>	Normierung: 2009 (kleine Stichprobe)
<i>Testgüte</i>	Reliabilität: sehr gut ($\alpha = .95$ – Gesamttest) Validität wurde überprüft
<p>Die WPPSI-III ist ein geeignetes Instrument zur Feststellung der kognitiven Fähigkeiten von Kindergartenkindern und eignet sich auch gut zur Beantwortung der <i>Frage nach einer früheren Einschulung</i>. Jedoch 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.</p> <p><i>Ausführliche Beschreibung ab S. 164.</i></p>	

WST A-P – Wortschatztest – aktiv und passiv

<i>Altersbereich</i>	14–60 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Erfasst aktiven und passiven Wortschatzumfang Testaufbau: 60 Aufgaben in 2 Testteilen: Teil 1 erfasst den aktiven Wortschatz. Teil 2 erfasst den passiven Wortschatz. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 2000–2004 Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient $r = .91$ (Form B) bzw. $r = .93$ (Form A) Inhaltliche Validität durch repräsentative Auswahl aus dem Duden belegt (Angabe der Autoren).
<p>Der WST A-P ist für die Intelligenzdiagnostik im wissenschaftlichen Sinne ungeeignet (erfasst nur Wortschatz). Er kann lediglich im Rahmen eines Screenings, ob und inwieweit es Probleme mit der Sprache gibt, verwendet werden.</p> <p><i>Ausführliche Beschreibung ab S. 167.</i></p>	

ZVT – Zahlen-Verbindungs-Test

<i>Altersbereich</i>	8–95 Jahre
<i>Struktur</i> <i>Kurz- beschreibung</i>	Erfasst die Informationsverarbeitungsgeschwindigkeit Testaufbau: Nach dem Zufallsprinzip angeordnete Zahlen sollen so schnell wie möglich verbunden werden. Einzel- und Gruppentest
<i>Normen</i> <i>Testgüte</i>	Normierung: 1970er Jahre Reliabilität: interne Konsistenz nicht angegeben, Split-Half-Koeffizient zwischen $r = .84$ und $r = .97$ Faktorielle Validität nachgewiesen. Konstrukt- und Kriteriumsvalidität geprüft (geringe bis mittlere Korrelationen)
<p>Der ZVT kann als brauchbares Verfahren zur Abschätzung eines besonderen Aspekts der intellektuellen Leistungsfähigkeit (Verarbeitungsgeschwindigkeit) nur im Sinne eines Screenings empfohlen werden. Zur Begabungsdiagnostik scheint der ZVT ungeeignet. Abgesehen davon, dass der ZVT nur eine Intelligenzfacette mit nur einem Aufgabentyp erfasst, reicht er auch nicht aus, um eine verlässliche Intelligenzdiagnose stellen zu können. Es sollten unbedingt Tests mit unterschiedlichen Aufgabenfacetten dem ZVT vorgezogen werden bzw. der ZVT nur als Zusatzinstrument zur Erhebung der Verarbeitungsgeschwindigkeit eingesetzt werden. Die Normen bedürfen einer Überprüfung, obgleich der Verdacht einer Normveränderung weniger begründet ist als etwa bei wissens- oder problemlösebezogenen Leistungstests.</p> <p><i>Ausführliche Beschreibung ab S. 168.</i></p>	



Ein herzliches Dankeschön an die Salzburger Schachjugend, bei deren Mannschaftsmeisterschaft die Fotos für diese Publikation entstanden!

In der vorliegenden Broschüre werden nach einem Überblick über verschiedene Intelligenztheorien und -modelle und einer Klärung der wichtigsten Begriffe im Zusammenhang mit dem Testen von Intelligenz (u.a. Normierung, Testgüte) die derzeit gängigen Intelligenztestverfahren ausführlich beschrieben.

Diese Beschreibungen beinhalten u.a. Angaben zur Durchführung, Auswertung und Interpretation (inklusive Testgütekriterien) sowie zur Eignung der Verfahren für die Diagnostik von Hochbegabung. Ein kurzes Kapitel zu speziellen Aspekten in der Hochbegabungsdagnostik (automatische Itemgenerierung, adaptives Testen) schließt die Broschüre ab.

Im Anhang finden sich neben einem Überblickskatalog mehrere Mustergutachten zu den Themenbereichen vorzeitige Einschulung, Überspringen einer Klasse und allgemeine Intelligenzdiagnostik.