

**Vorgeschlagene neuropsychologische Diagnostik
bei Kindern und Jugendlichen
mit Verdacht auf FASD**

Bereich	Testverfahren	Abkürzung	Altersbereich
Intelligenz / kognitive Leistungsfähigkeit	Snijders-Oomen Non- verbaler Intelligenztest	SON-R 2-8	2;0- 7;11 Jahre
		SON-R 6-40	6;0- 40;0 Jahre
	Wechsler Preschool and Primary Scale of Intelligence –Forth Edition – Deutsche Version	WPSSI- IV	2;6- 7;7 Jahre
		WISC- V	6;0 - 16;11 Jahre
		WAIS-IV	16;0 - 89;0 Jahre
Wechsler-Intelligenztest für Erwachsene			
Entwicklung	Klinisch- entwicklungsneurologische Beurteilung		
	Bayley Scales of Infant and Toddler Development III	BSID III	1 - 42 Monate

Sprache	Subtests „Wortschatz-Test“ und „Gemeinsamkeiten finden“ (WPSSI, WISC, WAIS)	WPSSI- IV WISC- V WAIS IV	2;6- 7;7 Jahre 6;0 - 16;11 Jahre 16;0 - 89;0 Jahre
	Skala „Sprachverständnis“ (WPSSI, WISC, WAIS)	WPSSI- IV WISC- V WAIS IV	2;6- 7;7 Jahre 6;0 - 16;11 Jahre 16;0 - 89;0 Jahre
	Sprachentwicklungstest für zweijährige Kinder	SETK-2	2;0 - 2;11 Jahre
	Sprachentwicklungstest für drei- bis fünfjährige Kinder	SETK 3-5	3;0 - 5;11 Jahre
	Sprachstandserhebungstest für Fünf- bis Zehnjährige SET 5-10	SET 5-10	5;0- 10;0 Jahre
Fein-/ Graphomotorik	Klinisch-neurologische Beurteilung		
und	Movement Assessment Battery for Children	M-ABC-2	3;0 - 16;11 Jahre
grobmotorische Koordination	Zürcher Neuromotorik		5;0 - 18;11 Jahre

Räumlich-visuelle Wahrnehmung und Räumlich-konstruktive Fähigkeiten	Frostigs Entwicklungstest der visuellen Wahrnehmung	FEW-3	4;0 - 10;11 Jahre
	Frostigs Entwicklungstest der visuellen Wahrnehmung Jugendliche und Erwachsene	FEW-JE	9;0- 90;0 Jahre
	Abzeichentest für Kinder	ATK	7;0 - 12;0 Jahre
	Rey Complex Figure Test and Recognition Trial	RCFT	6;0 - 89 Jahre
	Subtests „Mosaik-Test“ (SON-R, WPSSI, WISC, WAIS), „Figuren legen“ (WPSSI, WISC, WAIS)	SON-R WPSSI WISC WAIS	2;6 - 7;0 Jahre 2;6- 7;7 Jahre 6;0 - 16;11 Jahre 16;0 - 89;0 Jahre
Lern- und Merkfähigkeit	Verbaler Lern- und Merkfähigkeitstest	VLMT	6;0 – 79;0 Jahre
	Merk- und Lernfähigkeitstest für 6- bis 16-jährige	Basic MLT	6;0 – 16;11 Jahre
	Skala „Arbeitsgedächtnis“(WISC)	WISC	6;0 - 16;11 Jahre

Exekutive Funktionen	Testbatterie zur Aufmerksamkeitsprüfung (Untertests: GoNogo; Arbeitsgedächtnis; Flexibilität; Inkompatibilität)	TAP	6;0 - 90;0 Jahre
	Testbatterie zur Aufmerksamkeitsprüfung für Kinder (Untertests: GoNogo; Arbeitsgedächtnis; Flexibilität; Inkompatibilität)	KITAP	6;0 - 10;0 Jahre
	Verhaltensinventar zur Beurteilung exekutiver Funktionen	BRIEF	6;0- 16;0 Jahre
	Verhaltensinventar zur Beurteilung exekutiver Funktionen für das Kindergartenalter	BRIEF-P	2;0- 6;0 Jahre
	Regensburger Wortflüssigkeitstest	RWT	8;0 – 15;0 Jahre und ab 18;0 Jahre
	Turn von London-Deutsche Version	TL-D	6;0 - 15;0 Jahre und ab 18;0 Jahre
	Wisconsin Card Sorting Test	WCST	6;5 - 89;0 Jahre
	Behavioral Assessment of the Dysexecutive Syndrome	BADS	16;0 - 87;0 Jahre
Behavioral Assessment of the Dysexecutive Syndrome (in children)	BADS-C	8;0 – 15;11 Jahre	

Rechenfertigkeiten	Deutscher Mathematiktest	DEMAT 1+	Ende der 1. Klasse bis Anfang der 2.Klasse
		DEMAT 2+	Ende der 2. Klasse bis Anfang der 3.Klasse
		DEMAT 3+	Letzte 6 Wochen der 3. Klasse bis erste 6 Wochen der 4. Klasse
		DEMAT 4+	3 Wochen vor und nach Halbjahr der 4. Klasse bis 6 Wochen vor Ende der 4. Klasse
	Heidelberger Rechentest	HRT	1.-4.Klasse
	Testverfahren zur Dyskalkulie bei Kindern	ZAREKI-K ZAREKI-R	5;0 - 7,5 Jahre 6;6 - 13,5 Jahre

Aufmerksamkeit	d2-R Aufmerksamkeits-Belastungstest	d2-R	9;0 - 60;0 Jahre
	Testbatterie zur Aufmerksamkeitsprüfung (Untertests: Alertness, Daueraufmerksamkeit, geteilte Aufmerksamkeit)	TAP	6;0 - 90;0 Jahre
	Testbatterie zur Aufmerksamkeitsprüfung für Kinder (Untertests: Alertness, Daueraufmerksamkeit, geteilte Aufmerksamkeit)	KITAP	6;0 - 10;0 Jahre
	Fremd- und Selbst-beurteilungsbögen zum Störungsbereich ADHS aus dem „Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche II“ (DISYPS)	FBB-ADHS	3;0- 18;0 Jahre
		SBB-ADHS	11;0- 18;0 Jahre
	Intelligence and Development Scales (Untertest “Aufmerksamkeit selektiv”)	IDS-2	5;0- 20;0 Jahre
		IDS-P	3;0- 5;11 Jahre
Durchstreichtest vom Wechsler Intelligence Scale for Children –Fifth Edition – Deutsche Version	WISC- V	6;0 - 16;11 Jahre	

Soziale Fertigkeiten und Verhalten	Elternfragebogen über das Verhalten von Kindern und Jugendlichen = Child-Behavior-Checklist	CBCL 1,5-5 C-TRF 1,5-5	4;0 - 18;0 Jahre 1;5- 5;0 Jahre
	Deutsche Schulalter-Formen der Child Behavior Checklist	CBCL 6-18R,TRF 6-18R, YSR 11-18R	6;0- 18;0 Jahre
	Youth Self Report	YSR	11;0 - 18;0 Jahre
	Verhaltensfragebogen bei Entwicklungsstörungen	VFE-E	4;0 - 18;0 Jahre
	Strenghts and Difficulties Questionnaire	SDQ	6;0 - 16;0 Jahre
	Fremd- und Selbstbeurteilungsbögen zum Störungsbereich Störungen des Sozialverhaltens aus dem „Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche II“ (DISYPS)	FBB-SSV SBB-SSV FBB-ADHS SBB-ADHS	4;0- 18;0 Jahre 11;0- 18;0 Jahre 3;0- 18;0 Jahre 11;0- 18;0 Jahre
	Intelligence and Development Scales (Untertests “Emotionen erkennen”, “Emotionen regulieren”, Soziale Situationen verstehen”, “Sozial handeln”)	IDS-2 IDS-P	5;0- 20;0 Jahre 3;0- 5;11 Jahre

Güteparameter der vorgeschlagenen neuropsychologischen Testverfahren zur Diagnostik von Kindern und Jugendlichen mit Verdacht auf FAS

Intelligenz / kognitive Leistungsfähigkeit

Snijders-Oomen non verbaler Intelligenztest 2-8 (SON-R 2-8; 2018)

Kurzbeschreibung: Der SON-R 2-8 ist ein Intelligenztestverfahren für alle Kinder im Alter von 2;0 bis 8;0 Jahren. Durch seinen sprachfreien Charakter eignet sich das Verfahren auch besonders gut für Kinder mit Schwierigkeiten und Einschränkungen auf dem Gebiet der Sprach- und Sprechentwicklung und der Kommunikation, wie zum Beispiel Kinder mit Sprach- und Sprechstörungen, schwerhörige oder gehörlose Kinder, Kinder mit einer autistischen Störung, Kinder mit Problemen in der sozialen Entwicklung und Kinder, deren Muttersprache nicht Deutsch ist.

Normierung: Für die deutsche Normierung wurden in den Jahren 2016 bis 2017 insgesamt Daten von 762 Kindern in Deutschland erhoben, die mit einer niederländischen Stichprobe zusammengefasst wurden (Gesamt-N = 1.727). Im Gegensatz zu vielen anderen Intelligenztests basieren die Normwerte auf dem exakten Alter der Kinder. Neben dem Referenzalter für die sechs Untertests und dem Gesamtwert werden altersbezogene Unterteststandardwerte und Skalenwerte angegeben. Die Normen wurden für die Altersspanne von 2;0 bis einschließlich 7;11 Jahren berechnet.

Durchführungsdauer: Durchführung dauert etwa 50 Minuten (inkl. kurzer Pausen).

Reliabilität: Retest-Reliabilität gegeben: Messwiederholung nach durchschnittlich 3 Monaten bei n = 101 Kindern: mittlere bis hohe Retest-Reliabilität der Skalen (SON-IQ .81, Handlungsskala .76, Denkskala .66).

Interne Konsistenz gegeben: Berechnet mit Guttman-Modell (Items als Analyseeinheit) im Mittel für SON-IQ .91 (Range über die verschiedenen Altersgruppen .89-.93), für Handlungsskala .87 (Range .86-.89) und für Denkskala .82 (Range .79-.88).

Validität: Die durchschnittliche Korrelation des IQ-Wertes im SON-R 2-8 mit den IQ-Werten in anderen normierten Intelligenztests (SON-R 2½-7, SON-R 6-40, WNV und WPPSI-IV) beträgt .79 (N = 208). Diese hohe Korrelation bestätigt die kongruente Validität des SON-R 2-8. Für die Handlungsskala und die Denkskala des SON-R 2-8 beträgt die durchschnittliche Korrelation .77 bzw. .65. Somit hängt die Handlungsskala im Durchschnitt stärker mit den IQ-Werten der anderen normierten Intelligenztests zusammen als die Denkskala.

Snijders-Oomen non verbaler Intelligenztest 6-40 (SON-R 6-40;2012)

Kurzbeschreibung: Der SON-R 6-40 erfasst sprachfrei die allgemeine Intelligenz bei Kindern ab 6 Jahren, Jugendlichen und Erwachsenen bis 40 Jahren. Daneben eignet sich der Test

insbesondere für die Intelligenzdiagnostik bei sprachbeeinträchtigten, entwicklungsverzögerten und schwach begabten Personen. Der SON-R 6-40 lässt sich gegenüber seinem Vorgänger (SON-R 5,5-17) bei einem deutlich größeren Altersbereich einsetzen.

Durchführungsdauer: Insgesamt max. 60 Min.

Normierung: Der SON-R 6-40 wurde von 2009 bis 2011 in Deutschland und den Niederlanden anhand einer Gesamtstichprobe von N = 1.933 Kindern, Jugendlichen (N = 1441) und Erwachsenen (17 Altersgruppen von 6;0 bis 40;0 Jahren) normiert.

Reliabilität: Retest-Reliabilität ist gegeben. Für die Altersgruppen beträgt die Korrelation zwischen .90 und .93, für die Subtests .72-.84, für SON-IQ gesamt .92
Interne Konsistenz gegeben: interne Konsistenzen der Subtests gut (.85-.89), interne Konsistenz des Gesamtwerts sehr gut (.95).

Validität: Es wurden Korrelationen mit anderen Intelligenztests (WISC-IV= .77, WIE = .83, WNV= .77) berechnet, die die Validität des Verfahrens belegen. Zusätzlich wurden zahlreiche Untersuchungen mit verschiedenen klinischen Gruppen (Kinder mit LRS, ADHS, Hörbeeinträchtigung) durchgeführt. Vergleiche zeigen außerdem, dass Kinder mit Migrationshintergrund bei der Bearbeitung des SON-R 6-40 weniger benachteiligt sind als bei anderen Tests.

Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV;2018)

Kurzbeschreibung: Die WPPSI-IV ist ein umfassendes Intelligenztestverfahren zur Erfassung allgemeiner und spezifischer kognitiver Fähigkeiten bei Kindergarten- und Vorschulkindern im Alter zwischen 2;6 und 7;7 Jahren. Die Version beruht auf einer umfangreichen Überarbeitung des WPPSI-III.

Normierung: deutsche Adaption, Normierung (n= 895 Kindern) und Übersetzung von 2015 bis 2017 durchgeführt

Bearbeitungszeit: Die Bearbeitungszeit liegt bei 2;6-3;11 Jahren bei 25 bis 35 Minuten (Untertests zur Berechnung des Gesamt-IQs und der primären Indizes) und bei 4;0-7;7 Jahren bei 35 bis 55 Minuten (Untertests zur Berechnung des Gesamt-IQs und der primären Indizes).

Reliabilität: Die Testhalbierungsreliabilität für die Subtests ist gegeben. Gute bis sehr gute Testhalbierungsreliabilitäten der Subtests in allen Normgruppen (.77-.95).

Die Retest-Reliabilität ist gegeben. Messwiederholung nach durchschnittlich 26 Tagen bei 104 Kindern (n = 41 2;6-3;11-Jährige und n = 63 4;0-7;7-Jährige): Mittlere bis hohe Stabilitäten für Subtests und Prozesswerte (.55-.87), hohe Stabilitäten für Indizes (.67-.90) und Gesamt-IQ (.84-.88) in beiden Altersgruppen.

Interne Konsistenz ist gegeben. Gute bis sehr gute Werte für Indizes und Gesamt-IQ in allen Altersgruppen (.79-.96). Keine Angaben für Subtests.

Validität: Zum Nachweis der Validität liegen Interkorrelationsstudien, faktorenanalytische Studien, Korrelationsstudien mit anderen Messinstrumenten: WPPSI-III (Petermann et al., 2014): n = 68-197, mittlere bis hohe Korrelationen von Subtests, Indizes und Gesamt-IQ beider Testversionen (.60-.83) WISC-V (Petermann, 2017): n = 29-31, mittlere bis hohe Korrelationen von Subtests, Indizes und Gesamt-IQ beider Testversionen (.44-.75) ET-6-6-R (Petermann & Macha, 2013): n = 63-98, mittlere Korrelationen von WPPSI-IV Indizes (außer Verarbeitungsgeschwindigkeit und Kognitiver Leistungsindex) und Gesamt-IQ mit Skalen kognitiver und sprachlicher Entwicklung (.25-.68) sowie Validierungsstudien mit speziellen Gruppen (Hoch - /Minderbegabung) vor.

Wechsler Intelligenztest für Kinder (WISC-V; 2017)

Kurzbeschreibung: Die WISC-V ist die aktuelle Revision des weltweit am häufigsten eingesetzten Intelligenzdiagnostikums zur Erfassung kognitiver Fähigkeiten von Kindern und Jugendlichen der Altersgruppe von 6;0 bis 16;11 Jahren. Der Test stellt eine Überarbeitung des WISC-IV dar.

Normierung: Es liegen Normen aus dem Jahr 2016 vor, die auf einer Stichprobe mit ca. 1.100 deutschen Kindern und Jugendlichen basieren.

Bearbeitungsdauer: Die Bearbeitungsdauer liegt zwischen 60 und 90 Minuten.

Reliabilität: Die Reliabilität der Untertests variiert zwischen $r = .80$ und $r = .93$. Auf Ebene der Indexwerte variiert die Reliabilität zwischen $r = .89$ und $r = .96$. Für den Gesamttest beträgt sie $r = .96$.

Testhalbierungsreliabilität eingeschränkt gegeben: Keine Angaben für Indizes und Gesamt-IQ. Sehr gute Split-half-Reliabilitätskoeffizienten (.83-.93; minderungskorrigiert nach Spearman-Brown) für die Subtests (nicht berechnet für Geschwindigkeitstests ZST, SYS & DT).

Retest-Reliabilität gegeben: Hohe Stabilitäten der Indizes (.72-.88), des Gesamt-IQ (.90) und der Subtests (.73-.84) (Korrelationen jeweils korrigiert um Gesamtvarianz der Normierungsstichprobe).

Interne Konsistenz gegeben: Sehr gute Konsistenzwerte für die Indizes (.89-.95) und den Gesamt-IQ (.96).

Validität: Die faktorielle und kriteriumsbezogene Validität konnte umfassend bestätigt werden. Hierzu wurden Vergleiche mit anderen Intelligenzdiagnostischen Verfahren sowie Faktoranalysen auf Basis der Normstichprobe durchgeführt.

Zusammenhänge mit anderen Intelligenztests:
WISC-IV (Petermann & Petermann, 2011): n = 83; hohe Korrelation der Gesamt-IQs (.89), Indizes (.62-.81) und Subtests (.64-.77).
WPPSI-III (Petermann, Ricken, Fritz, Schuck & Preuß, 2014): n = 32; hohe Korrelation der Gesamt-IQs (.89), niedrige bis hohe Korrelation der Indizes (.15-.79) und mittlere bis hohe Korrelation der Subtests (.31-.79).
WAIS-V (Petermann, 2012): n = 31; hohe Korrelation der Gesamt-IQs (.78), Indizes (.69-.84) und mittlere bis sehr hohe Korrelationen der Subtests (.31-1.00).
KABC-II (Kaufmann & Kaufmann, 2015): n = 84; hohe Korrelation von Gesamt-IQ mit Fluid-Kristallin-Index der KABC II (.83) und mittlere bis hohe Korrelationen mit den

weiteren KABC-II Indizes (.53-.76); Korrelation der Indizes mit den verschiedenen KABC-II Indizes theoriekonform niedrig bis hoch (.10-.74). Validierungsstudien mit speziellen Gruppen (Hoch - /Minderbegabung). Signifikante Unterschiede zwischen durchschnittlich und hochbegabten Kindern (n = 13; Nulleffekte bis große Effekte) bzw. Kindern mit Intelligenzminderung (n = 9; sehr große Effekte).

Wechsler Intelligenztest für Erwachsene (WAIS-IV; 2012)

Kurzbeschreibung: Die WAIS-IV dient der Erfassung der allgemeinen Intelligenz und von spezifischeren kognitiven Fähigkeiten bei Jugendlichen ab 16 Jahren und bei Erwachsenen. Der WAIS-IV ist gegenüber seinem Vorgängerverfahren, dem WIE (Wechsler Intelligenztest für Erwachsene), deutlich verändert. Neu entwickelte Untertests (z. B. Formenwaage, Visuelle Puzzle) ermöglichen es, Facetten der Intelligenz zu erfassen, die sich in der aktuellen Forschung als bedeutsam herausgestellt haben.

Normen: Es liegen repräsentative Normen (1.425) für Deutschland aus dem Jahr 2012 vor.

Bearbeitungsdauer: Die Bearbeitungsdauer liegt bei ca. 90 Minuten (Kerntests) bzw. ca. 115 Minuten (inkl. optionaler Untertests).

Reliabilität: Testhalbierungsreliabilität gegeben: Sehr gute Split-half-Reliabilitätskoeffizienten (minderungskorrigiert nach Spearman-Brown) der Indizes (.90-.97) und des Gesamt-IQ (.98). Koeffizienten für die Subtests gut bis sehr gut (.78-.93).

Retest-Reliabilität gegeben: Zweimalige Testung von 166 Personen mit durchschnittlichem Abstand von 34 Tagen zwischen Testungen; hohe Stabilität der Indizes (.81-.91) und des Gesamt-IQ (.94); Koeffizienten für die Subtests mittel bis hoch (.61-.85).

Keine Angaben zur internen Konsistenz.

Validität: Inhaltsvalidität ist gegeben und die interne Struktur nachgewiesen, klinische Validierungsstudien mit Hochbegabten und Personen mit Intelligenzminderung liegen vor.

Entwicklung

Bayley Scales of Infant and Toddlers Development-III (BAYLEY III; 2015)

Kurzbeschreibung: Beurteilung des aktuellen kognitiven, sprachlichen und motorischen Funktionsniveaus.

Fünf Bereiche: Kognition, Sprache Rezeptiv und Sprache Expressiv, Feinmotorik und Grobmotorik.

Normierung: Gesamtstichprobe von N = 1.009 normiert. N = 878 deutschen Kindern ohne bekannte Beeinträchtigungen auf 17 Altersgruppen verteilt .

jüngeren Altersgruppen mit niederländischen Säuglingen (N = 131) ergänzt.

Bearbeitungsdauer: Bearbeitungsdauer Alters- und leistungsabhängig:

Bayley-III-Skalen: 50 Minuten (Säuglinge) bis 90 Minuten (Kindergartenalter).

Screening-Test: 20 bis 25 Minuten

Reliabilität: Interne Konsistenz: Reliabilitätskoeffizienten der Untertests:

Gesamtstichprobe zwischen $r = .77$ und $r = .89$; Skalen zwischen $r = .86$ und $r = .88$; mittlere Reliabilitätskoeffizienten für die Untertests des Screening-Tests zwischen $r = .68$ und $r = .83$.

Validität: Inhaltliche Validität und auch Konstruktvalidität konnten nachgewiesen werden.

Konstruktvalidität: Korrelation zwischen Skalen und Untertests zwischen $.66$ und $.69$.

Ebene Untertest höchste Korrelation $.41$, höchste Skaleninterkorrelation $.43$

Kriteriumsvalidität: Stichprobe Frühgeborene: Korrelationsmuster zwischen BALEY-III Skalen und Elternurteil zur Kognition und Sprachlichen Entwicklung

Klinische Validierung: Trisomie 21: deutlich unterdurchschnittliche Skalennwerte der Kognitiven Skala und der Sprach Skala

Für den Screening Test liegen für jeden Untertest Angaben zu Spezifität und Sensitivität (Häufungen knapp oberhalb von 80%) vor und fallen zufriedenstellend aus.

Sprache

Sprachentwicklungstest für zweijährige Kinder (SETK-2; 2016)

Kurzbeschreibung: Der SETK-2 eignet sich zur Untersuchung des sprachlichen Entwicklungsstandes und zur Abklärung von Entwicklungsstörungen im sprachlichen Bereich. Dieser erfasst mit vier Untertests, anhand von kindgerechtem Testmaterial, die rezeptive und produktive Sprachverarbeitungsfähigkeit. Mit Hilfe einer Kurzform des Verfahrens kann, im Sinne eines Screeninginstrumentes, eine Identifikation von Risikokindern erfolgen. Nach wie vor ist er im deutschsprachigen Raum der einzige standardisierte und normierte Individualtest, der für eine genaue und prognostisch valide Einschätzung des Sprachentwicklungsstandes bei zweijährigen Kindern geeignet ist.

Normierung: Es liegen aktuelle, deutsche Normen (T-Werte, Prozentränge) für die zwei Halbjahresschritte 2;0–2;5 und 2;6–2;11 Jahre vor ($N = 374$). Erhebungszeitraum zwischen 2013 und 2014.

Bearbeitungszeit: Die Bearbeitungsdauer beträgt etwa 15 bis 20 Minuten.

Reliabilität: Die erreichten Reliabilitätskennwerte können insbesondere für die beiden Sprachproduktions-Tests als sehr zufriedenstellend bewertet werden. Diese bewegen sich zwischen $.83$ und $.94$. Die Kennwerte für die Sprachverständnis-Tests liegen dagegen deutlich niedriger und gehen über $.61$ nicht hinaus. Dies kann darauf zurückgeführt werden, dass beide Sprachverständnis-Tests vergleichsweise wenige Antworten verlangen.

Validität: Aktuelle Untersuchungsbefunde bestätigen die bereits in der 1. Auflage belegte Differenzierungsfähigkeit und die prognostische Validität des SETK-2. Spätes Wortlernen sind dementsprechend Risikokinder ersten Ranges für persistierende Sprachschwierigkeiten.

Sprachentwicklungstest für drei- bis fünfjährige Kinder (SETK 3-5; 2015)

Kurzbeschreibung: Der SETK 3-5 erfasst mit vier Untertests bei den dreijährigen Kindern und mit fünf Untertests bei den vier- bis fünfjährigen Kindern rezeptive und produktive Sprachverarbeitungsfähigkeiten sowie auditive Gedächtnisleistungen.

Normierung: Für die aktuelle, deutsche Normierungsstudie wurden 934 Kinder im Alter von 3;0 bis 5;11 Jahren untersucht. Erhebungszeitraum Mitte 2012 bis März 2014. Es liegen Normen (T-Werte, Prozentränge) für die fünf Altersgruppen 3;0–3;5 Jahre, 3;6–3;11 Jahre, 4;0–4;5 Jahre, 4;6–4;11 Jahre und 5;0–5;11 Jahre vor.

Bearbeitungszeit: Die Bearbeitungsdauer beträgt 15 bis maximal 25 Minuten.

Reliabilität: Die erreichten Werte für die Reliabilität, die sich zwischen $\alpha = .65$ und $\alpha = .92$ bewegen, können als zufriedenstellend bis sehr gut bewertet werden.

Validität: Zahlreiche Untersuchungen belegen die Differenzierungsfähigkeit (Frühgeborene, spezifisch sprachgestörte Kinder, späte Wortlernen) und die prognostische Validität: Aufgabe des Sätze Nachsprechens (meist 1,5 Jahre vor Einschulung) erlaubt eine gute Vorhersage der Schriftsprachkompetenz zu Beginn der zweiten Klasse und kann Hinweise auf möglicherweise zu erwartende Schwierigkeiten beim Lesen- und Schreiben-Lernen liefern.

Sprachstandserhebungstest für Fünf- bis Zehnjährige (SET 5-10; 2018)

Kurzbeschreibung: Der SET 5-10 nimmt eine an den Entwicklungsstand angepasste, umfassende Beurteilung des Sprachstands vor. Er besteht aus 10 Untertests, die die Bereiche Wortschatz, Semantische Relationen, Verarbeitungsgeschwindigkeit, Sprachverständnis, Sprachproduktion, Grammatik/Morphologie und die auditive Merkfähigkeit überprüfen.

Normierung: Die Normen (T-Werte, Prozentrangwerte) basieren auf einer deutschlandweiten Erhebung an 1.052 Kindern. Es liegen Normen für sieben Altersgruppen vor.

Bearbeitungszeit: Die Durchführung erfolgt im Einzelsetting und nimmt etwa 45 Minuten in Anspruch.

Reliabilität: Die Mehrzahl der Untertests weist (nach Cronbachs Alpha) eine gute interne Konsistenz zwischen $\alpha = .71$ und $.91$ auf.

Validität: Die Validität des SET 5-10 ist in verschiedenen Studien ausführlich untersucht worden. Analysen zur Kriteriumsvalidität des Verfahrens ergaben mittlere bis hohe Korrelationen zwischen den Untertests des SET 5-10 und anderen Testverfahren, die vergleichbare Konstrukte erheben. Die Ergebnisse sprechen für die Validität des Verfahrens.

Fein-/Graphomotorik oder grobmotorische Koordination

Movement Assessment Battery for Children (M-ABC-2; 2015)

Kurzbeschreibung: Die Movement ABC-2 ist ein Testverfahren mit dem sich das kindliche Leistungsvermögen in verschiedenen motorischen Bereichen überprüfen lässt. Es stehen dabei drei verschiedene Testbatterien zur Verfügung: Altersgruppe 1 (3;0-6;11 Jahre), Altersgruppe 2 (7;0-10;11 Jahre) und Altersgruppe 3 (11;0-16;11 Jahre). Es werden drei wichtige Komponenten motorischer Funktionen mit insgesamt acht Aufgaben untersucht: Handgeschicklichkeit, Ballfertigkeit, statische und dynamische Balance.

Normierung: Die Normen basieren auf den Leistungen von 1.000 in Großbritannien untersuchten Kindern und zusätzlich auf den Ergebnissen von 634 Kindern aus verschiedenen deutschen Standorten.

Bearbeitungszeit: Die Bearbeitungsdauer liegt bei Ca. 20 bis 30 Minuten.

Reliabilität: Die Zuverlässigkeit für eine Testwiederholung nach zwei Wochen beträgt $r = .97$ ($N = 138$). Kinder mit motorischen Beeinträchtigungen erzielen niedrige Testwerte ($PR < 5$) ($N = 20$).

Validität: Konstruktvalidität mittels WISC IV anhand einer Teilstichprobe

Korrelation IVI (Individueller Verarbeitungsindex) mit dem IQ Gesamtergebnis des WISC IV = .80

Korrelation FKI (Fluid Kristallin Index) mit dem IQ Gesamtergebnis des WISC IV = .88

Korrelation SFI (Sprachfrei-Index) mit dem IQ Gesamtergebnis des WISC IV = .75

Klinische Validierungsstudie (Umschrieben Entwicklungsstörung motorischer Funktionen, Sprachstörungen, ADHS, Autistische Störungen, Lernstörungen, Intelligenzminderung)

Zürcher Neuromotorik

Kurzbeschreibung: Die Zürcher Neuromotorik ist ein Testverfahren, das eine Untersuchung der motorischen Funktionen von Kindern und Jugendlichen zwischen 3 und 18 Jahren erlaubt. (Repetitive Bewegungen; Alternierende Bewegungen; Sequentielle Bewegungen; Adaptive Leistungen; Gleichgewicht; Haltung)

Normierung: Schweiz, 1997
Alter: 5;0 bis 18;11 J.

<u>Reliabilität:</u>	Zeitmessung:	Intra-Rater: $r = 0.56 - 1.0$
		Inter-Rater: $r = 0.32 - 0.99$
		Test-Retest: $r = 0.38 - 0.89$
Mitbewegungen:	Intra-Rater: $r = 0.53 - 0.90$	
	Inter-Rater: $r = 0.44 - 0.82$	
	Test-Retest: $r = 0.20 - 0.61$	
Indizes:	Intra-Rater: $r = 0.73 - 1.0$	
	Inter-Rater: $r = 0.62 - 0.98$	
	Test-Retest: $r = 0.40 - 0.91$	

Validität: Klinische Validität: motorische Ungeschicklichkeit, die pädagogischer oder medizinischer Intervention bedurfte, Frühgeborene, angeborener Herzfehler, Z.n. perinatale Asphyxie, spezifische Spracherwerbsstörung.

Sensitivität 93%

Spezifität 93%

Räumlich-visuelle Wahrnehmung und räumlich-konstruktive Fähigkeiten

Frostigs Entwicklungstest der visuellen Wahrnehmung (FEW-3; 2021)

Kurzbeschreibung: Der FEW-3 ist die deutsche Bearbeitung des Developmental Test of Visual Perception 3 (DTVP-3) und der Nachfolger des bewährten entwicklungsdiagnostischen Verfahrens FEW-2. Er ist darauf ausgerichtet, visuelle Wahrnehmungsfähigkeiten differenziert zu erfassen. Dabei wird explizit zwischen motorikreduzierten und visuo-motorischen Anteilen unterschieden.

Normierung: Die FEW-3-Normstichprobe setzt sich aus 1.708 Kindern zwischen 4;0 und 10;11 Jahren aus sechs deutschen Bundesländern zusammen. Es stehen altersspezifische Wertpunkte, Prozentränge und T-Werte für die Subtest-Rohwertsummen sowie Prozentränge und Index-Werte für die Indizes der visuellen Wahrnehmung zur Verfügung. Die Normierung wurde für 11 Altersgruppen vorgenommen, die zwischen 4;0 Jahren und 7;11 Jahren jeweils einen Bereich von sechs Monaten umfassen und ab 8;0 Jahren von 12 Monaten.

Bearbeitungszeit: Die FEW-3-Testdauer variiert je nach Kind und dessen Leistungsfähigkeit zwischen 20 und 40 Minuten. Normalerweise soll der FEW-3 in einer Sitzung durchgeführt werden. Für einzelne Kinder kann es aber auch sinnvoll sein, die Durchführung auf zwei Testsitzungen zu verteilen. Dies wird von der Testleiterin bzw. dem Testleiter individuell entschieden.

Reliabilität: Die drei Indizes der visuellen Wahrnehmung (VMI, MRVW und GVW) weisen mittlere interne Konsistenzwerte auf, die über dem Wert von Alpha = .80 liegen. Die Berechnung der internen Konsistenzen für die Indizes erfolgte nach dem von Guilford empfohlenen Verfahren und die durchschnittlichen internen Konsistenzen über alle Altersgruppen hinweg wurden mit Hilfe von Fishers z-Transformation ermittelt.

Validität: Die Kriteriumsvalidität des FEW-3 wird durch erwartungsgemäße Zusammenhänge zwischen Wahrnehmungsleistungen und feinmotorischen, kognitiven und schulischen Leistungen (erhoben mit den Verfahren MOVE 4-8, IDS-2, KLASSE 4) belegt. Das theoretische Modell der Wahrnehmungsdimensionen des FEW-3 wird durch konfirmatorische Faktorenanalysen bestätigt. Die Konstruktvalidität wird durch den erwarteten monotonen Anstieg der Wahrnehmungsleistungen über die Altersgruppen hinweg gestützt.

Frostigs Entwicklungstest der visuellen Wahrnehmung Jugendliche und Erwachsene (FEW- JE; 2012)

Kurzbeschreibung: Der FEW-JE ist die deutsche Bearbeitung des Developmental Test of Visual Perception – Adolescent and Adult (DTVP-A).

Normierung: Die Normstichprobe setzt sich aus insgesamt 1.440 Kindern, Jugendlichen (N= 624) und Erwachsenen im Alter von 9-90 Jahren aus allen Ländern der Bundesrepublik

Deutschland zusammen. Es stehen altersspezifische Wertpunkte, standardisierte Skalenwerte und Prozentränge für 15 Altersgruppen zur Verfügung.

Bearbeitungszeit: Die Durchführungsdauer liegt bei etwa 20 Minuten. Das Testen von Menschen mit Behinderungen kann aufgrund der (motorischen, visuellen) Einschränkungen etwas länger dauern. Die Auswertung der Testergebnisse dauert etwa 10 Minuten.

Reliabilität: Die Reliabilität der drei Skalenwerte (Batterie-Reliabilität nach Lienert und Raatz) beträgt über alle Altersgruppen hinweg .87 (Allgemeine visuelle Wahrnehmung), .83 (Visuo-motorische Integration) und .81 (Motorik-reduzierte visuelle Wahrnehmung).

Validität: Das Testkonzept wird durch konfirmatorische Faktorenanalysen bestätigt. Die klinische Validität wurde durch Untersuchungen an neurologisch erkrankten PatientInnen (u.a. Schlaganfälle, Schädel-Hirn-Traumata, Tumore) geprüft. In einer nach Geschlecht und Alter gematchten Kontrollgruppe ergeben sich in allen Skalen- und Untertestwerten signifikant höhere Mittelwerte als in den Patientengruppen.

Abzeichentest für Kinder (ATK; 2004)

Kurzbeschreibung: Der Abzeichentest für Kinder (ATK) ist besonders zur Diagnostik raumanalytischer und räumlich-konstruktiver Fähigkeiten und zum Nachweis entsprechender Funktionsstörungen bei 7- bis 12-jährigen Kindern geeignet.

Normierung: Deutschland, 2004 (Erhebung an 350 gesunden Kindern)
Alter 7;0 - 12;0 J.

Reliabilität: bisher keine Angaben dazu

Validität: Konstrukt: Klinik: neurologische Stichprobe von Kindern mit gesicherter räumlich-konstruktiver Störung und neurologische Stichprobe von Kindern ohne räumlich-konstruktive Störung.

Korrelation des GAT und Mosaiktest (MT) aus HAWIK: $r = 0.50$

Cut-off Wert (ATK) und Wertpunkt von ≤ 4 (MT): Übereinstimmung von 87,5 %; die Autoren gehen daher von einer guten klinischen Validität des ATK aus (S.30).

Rey Complex Figure Test and Recognition Trial (RCFT; 1995)

Kurzbeschreibung: Der Rey- Osterrieth Complex Figure Test (ROCF) ist ein neuropsychologischer Test zur Erfassung der Fähigkeit der räumlich visuellen Konstruktion und der visuellen Gedächtnisleistung. Auch wird der Test zur Erfassung exekutiver Funktionen eingesetzt.

Normierung: USA, 1993

Alter 6;0 - 89;0 J.

Test nur auf Englisch, jedoch Beurteilung nicht sprachgebundener Leistungen

Reliabilität: Interrater $r = 0.93 - 0.99$

Retest $r = 0.76 - 0.89$

Validität: Konstrukt: Klinik: Gehirnschäden

Korrelation mit Subtests anderer neuropsycholog. Testverfahren

Diskrimination Gesund / Psychiatrisch krank / Gehirnschaden

→ Copy: richtig positiv 57,8%

Recall: richtig positiv 61,1%

Recall and Recognition richtig positiv 77,8%

Exekutive Funktionen

Testbatterie zur Aufmerksamkeitsprüfung (TAP)

Version 2.3.1 (Version 2021)

Kurzbeschreibung: Die Testbatterie zur Aufmerksamkeitsprüfung (TAP) bietet eine Sammlung unterschiedlicher Testverfahren, mit denen die vielfältigen Teilaspekte der Aufmerksamkeit differenziert erfasst werden können. Auch damit verbundene Aspekte der visuellen Wahrnehmung werden von der Testbatterie abgedeckt.

Normierung: Europa Erhebung: Normierungsdaten aus unterschiedlichen Studien, aus denen jeweils unterschiedliche Teilmengen von Tests verwendet wurden; Tests sind nicht in gleichem Umfang für Erwachsene und Jugendliche normiert Alter: 6;0 bis 90;0 J. (abhängig vom jeweiligen Untertest)

Reliabilität: Testhalbierung (Odd-Even-Reliabilität): $r = 0.219 - 0.997$ (insgesamt für alle Tests; Kinder- und Jugendliche: 6-19 J.); insgesamt: die Reliabilitäten der Reaktionszeitmediane liegen überwiegend über 0.90, daher zufriedenstellend bis sehr gut; die Reliabilitäten der Fehlermaße sind hingegen häufig unzureichend; Test-Retest: siehe vorläufige Ergebnisse für Erwachsene S.100;

Es liegt eine Studie von Földényi et al. (2000) zur Reliabilität und Retest-Stabilität bei 95 Deutschschweizer Schulkindern vor (7-10 J.): $r_{tt} = 0.84$; insgesamt: mittlere bis hohe Reliabilitäten mit Ausnahme des Untertests: Geteilte Aufmerksamkeit: $r_{tt} = 0.32$; insgesamt: deutlich bessere Reliabilitäten bei Reaktionszeitmedianen im Vergleich zu Kriterien der Leistungsgüte (Auslassungen, Fehler).

Validität: Konstrukt: Studien zur faktoriellen Validität liegen vor.

Klinische Validität: Studien mit neurologischen Patienten, psychiatrischen

Patienten, pharmakologische Studien, funktionelle Bildgebungsstudien, Studien mit Kindern und Studien mit älteren Probanden;

Studien mit Kindern:

Földényi et al. (2000): klinische Validierungsstudie an Kindern mit ADHS : 90 % der Kinder konnten korrekt der ADHS-Gruppe oder Kontrollgruppe zugeordnet werden anhand Reaktionszeit-Schwankungen im Go/Nogo, Fehleranzahl bei Flexibilität (Wechsel: nonverbal) und Testalter; Kriteriumsvalidität: Korrelation TAP und Eltern-/Lehrerurteile (standardisierte Fragebögen): signifikante Korrelation Go/Nogo, Inkompatibilität, geteilte Aufmerksamkeit, Flexibilität

Koschak et al. (2003): ADHS-Kontrollgruppe

Tucha et al. (2005): ADHS-Gesunde Kinder
Heubrock et al. (2001): Inanspruchnahme neuropsychologischer Ambulanz
Kunert et al. (1996): faktorielle Validität

Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP;2002)

Kurzbeschreibung: Die Testbatterie zur Aufmerksamkeitsprüfung – Kinderversion (KiTAP) dient der Erfassung der Aufmerksamkeit von Kindern im Alter von 6;0 bis 10;11 Jahre.

Normierung: Deutschland, 2002

Erhebung: Normierungsdaten aus unterschiedlichen Studien, aus denen jeweils unterschiedliche Teilmengen von Tests verwendet wurden aus Deutschland, Frankreich, Belgien, Italien, Schweiz, Österreich; Alter: 6;0 bis 10;0 J.

Reliabilität: Testhalbierung (Split-Half-Reliabilität):

$r = 0.64 - 0.97$ (insgesamt für alle Tests; Kinder: 6-7 J.)

$r = 0.63 - 0.96$ (insgesamt für alle Tests; Kinder: 8-10 J.)

Insgesamt: die Reliabilitäten der Reaktionszeitmediane deutlich besser, z.T. sehr gut; die Reliabilitäten der Fehlermaße hingegen häufig im unteren Bereich, aufgrund insgesamt eher geringer Auslassungen als Artefakte zu betrachten.

Validität: Konstrukt: Studien zur faktoriellen Validität liegen vor.

Klinische Validität: Studien bei Kindern mit ADHS und nach Hirnschädigung.

Verhaltensinventar zur Beurteilung exekutiver Funktionen (BRIEF; 2013)

Kurzbeschreibung: Klinisches Fragebogenverfahren zur Erfassung exekutiver Funktionsbeeinträchtigungen. Es liegt in drei Versionen vor: zur Beurteilung durch Eltern, durch Lehrer und zur Selbstbeurteilung.

Fremdbeurteilung (86 Fragen). 2 Hauptindices gebildet:

Verhaltensregulations- Index aus den Skalen Hemmen, Umstellen und emotionale Kontrolle
Kognitiver Regulations-Index aus den Skalen Initiative, Arbeitsgedächtnis, Planen/Strukturieren, Ordnen/Organisieren und Überprüfen.

Beide Indices ergeben zusammen den Exekutiven Gesamtwert.

Die Selbstbeurteilung mit 80 Fragen (BRIEF-SB) ist weitgehend parallel aufgebaut.

Normierung: Für die Fremdbeurteilung (Eltern und Lehrer) existieren geschlechtsspezifische Normen (N = 921 [Eltern] bzw. N = 941 [Lehrer]) für die Altersstufen 6 bis 7, 8 bis 10, 11 bis 13 und 14 bis 16 Jahre. Für die Selbstbeurteilung existieren geschlechtsspezifische Normen (N = 531) für die Altersstufen: 11 bis 13 und 14 bis 16 Jahre.

Bearbeitungszeit: Die Bearbeitungszeit beträgt 10 bis 15 Minuten.

Reliabilität: Die interne Konsistenz der Skalen und Indices der BRIEF-Eltern- und -Lehrerversionen liegt zwischen .79 und .98, die des BRIEF-SB zwischen .73 und .96. Retest-Reliabilitäten liegen beim BRIEF-Eltern für die meisten Skalen über .80, bei der Lehrerversion meist über .90. Retest-Reliabilitäten des BRIEF-SB liegen zwischen .78

und .86. Die Hauptindices der Lehrer- und Elternskalen lassen sich faktorenanalytisch reproduzieren.

Validität: Die klinische Validität wird anhand von Profildarstellungen verschiedener Störungsbilder (u.a. ADHS, Autismus, Frühgeborene, Epilepsie, FASD) aufgezeigt. Das BRIEF ist in der Lage, zuverlässig auf Gruppennittelwert-Ebene zwischen Kindern und Jugendlichen mit und ohne ADHS, mit und ohne Autismus-Spektrum-Störungen und mit und ohne fetales Alkoholsyndrom zu unterscheiden. Bei Frühgeborenen und Tumoren des ZNS gibt es keine bedeutsamen Gruppenunterschiede. Internationale Studien mit der Original-BRIEF-Version fanden ebenfalls bei Kindern mit Epilepsie oder Kindern mit Zustand nach Schädelhirntrauma nur geringe Unterschiede zu neurologisch unbeeinträchtigten Kindern.

Verhaltensinventar zur Beurteilung exekutiver Funktionen für das Kindergartenalter (BRIEF-P; 2013)

Kurzbeschreibung: Das BRIEF-P ist ein Fremdbeurteilungsfragebogen zur Einschätzung von exekutiven Funktionen bei Kindern im Kindergartenalter. Es kann durch die Eltern und/oder durch andere Bezugspersonen beispielsweise die Erzieher/-innen im Kindergarten ausgefüllt werden.

Normierung: Es liegen alters- und geschlechtsgetrennte T-Werte und Prozentränge für Kinder im Alter von 2;0 bis 6;11 vor (N = 460; Deutschland; 2012-2013).

Bearbeitungszeit: Die Bearbeitungsdauer liegt zwischen 10 bis 15 Minuten.

Reliabilität: Die interne Konsistenz der Skalen liegt für das Elternurteil zwischen .75 und .89, für das Erzieherurteil werden Koeffizienten zwischen .82 und .94 erreicht. Für die Indizes liegen die Werte der Elternversion zwischen .86 und .91 und die Werte der Erzieherversion zwischen .89 und .95. Die interne Konsistenz für die Gesamtskala beträgt in der Elternversion .95, in der Erzieherversion .96. Die Interrater-Reliabilität erreicht für den Gesamtwert .56.

Validität: Es bestehen signifikante Korrelationen von $r = .40$ bis $.58$ zwischen den BRIEF-P-Skalen Arbeitsgedächtnis sowie Planen/Organisieren und den Skalen Aufmerksamkeitsprobleme, Sozialer Rückzug und Emotionale Reaktivität der CBCL 1½-5. Weiterhin korrelieren die Skalen Inhibition sowie Planen/Organisieren vor allem mit der Skala Hyperaktivität ($r > .70$) des Behavior Assessment System for Children (BASC). Aufmerksamkeitswechsel und Emotionale Kontrolle des BRIEF-P korrelieren mit den Skalen Rückzug, Depression und Anpassungsfähigkeit des BASC ($r = .51$ bis $.76$). Faktorenanalytische Auswertungen des BRIEF-P weisen darauf hin, dass der Test Informationen erfasst, die nicht von anderen Beurteilungsskalen erfasst werden (Sherman & Brooks, 2010).

Regensburger Wortflüssigkeitstest (RWT;2001)

Kurzbeschreibung: Der RWT ist ein diagnostisches Verfahren zur Erfassung der Wortflüssigkeit, bei dem über einen Zeitraum von ein oder zwei Minuten Lösungen verbal generiert werden müssen.

Normierung: Die Normen stammen aus dem Jahr 2000 (N (Kinder) = 184).

Reliabilität: Keine Angaben zur internen Konsistenz. Retest-Reliabilitätskoeffizienten zwischen .72 und .89 (N = 80). Interrater-Reliabilität = .99.

Validität: Die Autoren verweisen auf die umfangreiche Literatur zu diesem gut eingeführten Verfahren. Für eine Unterstichprobe (N = 94) wurden außerdem Korrelationen mit Leistungen in IQ- und anderen neuropsychologischen Tests berechnet. Am höchsten korrelieren die Ergebnisse mit einem Test zur figuralen Flüssigkeit (5 Punkte Test, Regard, 1982; Korrelation mit S-Wörtern $r = .512$). Bezüglich der klinischen Validität verweisen die Autoren auf die Sensibilität dieses etablierten Verfahrens für verschiedene klinische Gruppen und auf die Ergebnisse in denen von ihnen erhobenen Patientengruppen.

Turm von London - Deutsche Version (TL-D; 2004)

Kurzbeschreibung: Beim Turm von London handelt es sich um eine Transformationsaufgabe zur Erfassung des konvergenten problemlösenden Denkens.

Normierung: Der TL-D in dieser Version ist standardisiert und normiert. Es liegen Normierungen für Kinder und Jugendliche (Normstichprobe N = 299) im Alter von 6 – 15 Jahren vor sowie für Erwachsene (Normstichprobe N = 1263). Rohwerte sowie altersspezifische und bildungsspezifische Prozentrangwerte (PR) können somit ermittelt werden. Die Personen können Rohwerte von 0 bis 20 erzielen. Dies entspricht der Anzahl der gelösten Aufgaben. Außerdem wird die Planungszeit protokollarisch erfasst. Für die Jugendlichen im Alter von 16 und 17 Jahren, für die keine gesonderten Normen vorliegen, wurden in dieser Untersuchung im Falle der 16-jährigen Jugendlichen die Normierung für 15-Jährige herangezogen und im Falle der 17-Jährigen sowie 18-Jährigen die bildungsspezifische Normierung der Erwachsenenstichprobe.

Reliabilität: Zusammenfassend ergeben sich gute bis zufriedenstellende Resultate bei der Überprüfung dieses Verfahrens auf Reliabilität (Tucha, O. und Lange, W., 2004).

a) Interne Konsistenz: Cronbachs Alpha = 0.785

b) Retest-Reliabilität: $r = 0.861$ (Halbtest-Reliabilitätskoeffizient nach Spearman-Brown)

Validität: Inhaltsvalidität: In zahlreichen Studien wurde der „Turm von London“ verwendet und dessen Sensitivität nachgewiesen. Eine Auflistung dieser klinischen Studien findet sich im Testhandbuch (Tucha, O. und Lange, W., 2004).

Konstruktvalidität: Der Zusammenhang zwischen dem Gesamtscore des TL-D und anderen kognitiven Leistungen wie allgemeine Intelligenz (Korrelation = 0.306), verbale Merkspanne (Korrelation = 0.297), verbales Arbeitsgedächtnis (Korrelation = 0.483), kognitive Verarbeitungsgeschwindigkeit (Korrelation = -0.509) und kognitive Flexibilität (Korrelation = -0.535) wurden anhand einer Teilstichprobe der Normstichprobe (n = 248) berechnet. Alle Korrelationen werden im Testhandbuch als statistisch bedeutsam ($p < 0.01$) berichtet. Es zeigte sich ein engerer Zusammenhang zur kognitiven Flexibilität (Korrelation = -0,535) und kognitiven Verarbeitungsgeschwindigkeit, die als exekutive Leistungen gesehen werden können. Auf der anderen Seite wird das im TL-D erfasste Merkmal nicht von anderen Verfahren erfasst.

Wisconsin Card Sorting Test (WCST; 1993)

Kurzbeschreibung: Der WCST wurde ursprünglich entwickelt, um die Fähigkeit zu abstraktem Denken und zur Änderung kognitiver Strategien in Reaktion auf veränderte Umweltkontingenzen zu erfassen und kann als Messinstrument der exekutiven Funktionen betrachtet werden, da die Fähigkeit zur Entwicklung und Beibehaltung einer angemessenen Problemlösungsstrategie und über verschiedene Reizbedingungen hinweg gefordert wird, um ein zukünftiges Ziel zu erreichen.

Normierung: Die Normierung erfolgte altersspezifisch von 6,5 Jahre bis 89 Jahre. Die Normierungsstichprobe bestand aus 899 Personen, wovon 453 Kinder und Jugendliche waren. Die Auswertung erfolgt nicht hinsichtlich einer Gesamtleistung; es werden 16 verschiedene Variablen erfasst, wie Anzahl der erkannten Kategorien, die gesamte Anzahl richtiger und falscher Antworten, ineffizientes Lernen, Perseverationen, Perseverationsfehler, Aufrechterhalten einer adäquaten Strategie, ineffiziente Konzeptualisierung.

Bearbeitungszeit: Die Kurzversion besteht aus der Sortierung des ersten Kartenstapels mit 64 Karten. Dieser Test dauert laut Testhandbuch 20 – 30 Minuten.

Objektivität: Die vorliegende Version ist bzgl. der Durchführung des Verfahrens und der Auswertung standardisiert. Zur Übereinstimmung der Auswertung zwischen verschiedenen Anwendern liegen einige Studien vor (Axelrod et al., 1992; Huettner et al., 1989, zitiert nach Heaton et al., 1993). Die Inter- und Intrascorer-Übereinstimmung wurden als exzellent beschrieben.

Reliabilität: Die Autoren berichten eine moderate bis gute Reliabilität bei Kindern und Jugendlichen. Der „generalizability coefficient“ lag zwischen 0.39 und 0.72, mit einem Mittelwert von 0.57 und Median von 0.60.

Validität: Der WCST wurde in zahlreichen Studien und bei sehr unterschiedlichen Gruppen eingesetzt. Viele der Studien werden im Testhandbuch aufgelistet und näher beschrieben. Insgesamt betrachten die Autoren die Validität des WCST zur Erfassung der Exekutivfunktionen bei neurologischen Patienten als nachgewiesen.

Behavioral Assessment of the Dysexecutive Syndrome (BADS; 1996)

Kurzbeschreibung: Das Behavioral Assessment of the Dysexecutive Syndrome (BADS) beinhaltet sechs Untertests, mittels derer verschiedene Aspekte exekutiver Funktionen erfasst werden können. Es liegt eine deutsche Übersetzung vor (Karen Ufer, 2000).

Normierung: Die Beurteilung der Testergebnisse erfolgt anhand der englischen Normierungsstudie. Deutsche Normen liegen bislang nicht vor. Andere Länder haben bisher ebenfalls auf die englischen Normierungsdaten zurückgegriffen (Karen Ufer, 2000). Die englische Normierungsstudie bestand aus 216 Kontrollpersonen im Alter von 16 bis 87 Jahren (M = 46,6; SD = 19,8), der NART-IQ lag durchschnittlich bei 102,7 (SD = 16,2, Bereich 69 – 129).

Objektivität: Aufgrund der Standardisierung der Testdurchführung und Auswertung sollte die Objektivität des Tests gewährleistet sein.

Reliabilität:

Inter-Rater-Reliabilität: Die Inter-Rater-Reliabilität wird für alle 6 Untertests als ausreichend hoch angegeben (zwischen 0,88 und 1,00). Für 8 der 18 Parameter wurde eine vollständige Übereinstimmung gefunden. Eine Auflistung der Werte für alle 18 Parameter findet sich im Testhandbuch.

Test-Retest-Reliabilität: Die Autoren der Testbatterie fanden einen klaren Trend zur Verbesserung der Testwerte in einer zweiten Testung, auch wenn der Unterschied der Mittelwerte nicht signifikant war ($t > 0.05$). Dies konnte in einer neueren Studie bestätigt werden (Jelicic, Henquet, Derix & Jolles, 2001).

Validität: Zur Untersuchung der Validität des BADS wurden 78 neurologische Patienten mit dem BADS und einer Reihe von anderen Verfahren (dem WAIS-R; Wechsler, 1981, dem Cognitive Estimates Test; Shallice und Evans, 1978 und der Kurzform des Wisconsin Card Sorting Tests untersucht. Im Vergleich der Patientenstichprobe mit der Kontrollstichprobe ergaben sich durch paarweise t-Tests signifikante Unterschiede der Gesamtprofilwerte der Gruppen ($p < 0,0001$). Die Kontrollgruppe zeigte nicht nur im Gesamtprofilwert deutlich bessere Punktwerte, sondern sie waren auch in jedem einzelnen Untertest signifikant besser als die Patienten. Der geringste Unterschied zwischen den Gruppen ergab sich für die Schlüsselsuche, der größte für den Modifizierten Sechs-Elemente-Test. Weitere Studien unterstützten die Anwendbarkeit und Nützlichkeit des BADS bei Patienten mit Schizophrenie (Ihara, Berrios & McKenna, 2000; Krabbendam, de Vugt, Derrix & Jolles, 1999), bei chronisch alkoholkranken Patienten (Moriyama et al., 2000) und bei depressiven Patienten (Paelecke-Haberman, Pohl & Lelow, 2005).

Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C; 2003)

Kurzbeschreibung: Behavioral Assessment of the Dysexecutive Syndrome in Children (BADS-C) ist eine Reihe von Tests zur Erfassung der Exekutivfunktionen. BADS-C ist sowohl für Kinder als auch für Jugendliche anwendbar. Vom BADS-C existiert eine britische, keine deutsche, Version.

Normierung: Die Normen (N = 259) sind von 2003.

Numerische Angaben zu Reliabilität und Validität werden nicht gemacht.

Validität: Zur Untersuchung der Validität wurden jedoch Vergleiche mit zwei anderen internationalen Testverfahren (Strengths and Difficulties Questionnaire (SDQ) und Dysexecutive Syndrome Questionnaire for Children (DEX-C)) angestellt: "BADS-C would therefore seem to be a valid test of executive functioning in everyday life." (Manual) „BADS-C is a scientifically valid and reliable battery of tests of executive functioning for children and adolescents...“ (Pearson Assessment). Durch Erfahrungen in der klinisch-psychologischen Anwendung des BADS-C schätzen die Leitlinien-Autoren, dass das Verfahren trotz mangelnder Überprüfung der Gütekriterien eine hohe klinische Validität besitzt.

Rechenfertigkeiten

Deutscher Mathematiktest für die 1.Klasse (DEMAT 1+; 2021)

Kurzbeschreibung: Zur Überprüfung der Mathematikleistungen von Grundschulern basierend auf den Inhalten der Mathematiklehrpläne der 1. Klassen (aller deutschen Bundesländer) sowie zur frühzeitigen Diagnose einer Rechenschwäche anhand curricularer Inhalte vom Ende der 1. Klasse und Anfang der 2. Klasse.

Normierung: Normierung nach Klassenstufen und Geschlecht getrennt, geschlechtsübergreifende Normen (T-Werte und Prozentränge) für den letzten Monat des 1. Schuljahres und die ersten drei Monate des 2. Schuljahres
Bundesweit normiert (N = 2.839 Neunormierung; 2.936 Erstnormierung)

Bearbeitungsdauer: Gruppentest 40 Minuten; Einzeltest 20–35 Minuten.

Reliabilität: Die interne Konsistenz der Subtests liegt zwischen $r = .56$ und $r = .85$; die des Gesamttests beträgt $r = .90$ bzw. $r = .88$ für die 1. bzw. 2. Klasse. Zwischen den Leistungen des ersten und zweiten Schuljahres ergab sich eine Retest-Reliabilität von $r = .65$ (N = 52).

Validität: Es existiert Lehrplanvalidität, da der Testkonstruktion die Mathematiklehrpläne aller deutschen Bundesländer zu Grunde lagen. Die Übereinstimmungsvalidität mit dem Lehrerurteil beträgt $r = .66$. Korrelationen mit anderen Verfahren zur Erfassung der mathematischen Kompetenz liegen bei $r = .67$ (DEMAT 2+), $r = .57$ (DEMAT 4), $r = .64$ (MBK 0), $r = .68$ (MBK 1+).

Deutscher Mathematiktest für die 2.Klasse (DEMAT 2+; 2020)

Kurzbeschreibung: Zur Überprüfung der Mathematikleistungen von Grundschulern in Bezug auf die Inhalte der Mathematiklehrpläne der 2. Klassen (aller deutschen Bundesländer) sowie zur frühen Diagnose einer Rechenschwäche anhand curricularer Inhalte Ende der 2.Klasse bis Anfang der 3.Klasse.

Für die 2. Auflage wurde das Manual aktualisiert und neue Normen aus dreizehn deutschen Bundesländern erhoben.

Normierung: Normen nach Klassenstufe getrennt (T-Werte und Prozentränge) von 6.344 Kindern für das Ende des 2. Schuljahres und den Beginn des 3. Schuljahres. Es werden geschlechtsübergreifende sowie geschlechtsspezifische Normen angegeben.

Bearbeitungsdauer: Gruppentest 45 Minuten. Einzeltest 20 bis 40 Minuten.

Reliabilität: Interne Konsistenz (Cronbachs Alpha): 2. Klasse: $\alpha = .93$, 3. Klasse: $\alpha = .92$; Testhalbierungsreliabilität (korrigiert nach Spearman Brown): 2. Klasse: $r = .95$, 3. Klasse: $r = .94$.

Validität: Es existiert Lehrplanvalidität, da der Testkonstruktion die Mathematiklehrpläne aller deutschen Bundesländer zu Grunde lagen. Die Übereinstimmungsvalidität mit dem Lehrerurteil beträgt $r = .66$. Korrelationen mit anderen Verfahren zur Erfassung der

mathematischen Kompetenz liegen bei $r = .66$ (DIRG), $r = .68$ (DEMAT 1+), $r = .68$ (MBK 1+).

Deutscher Mathematiktest für die (3.Klasse DEMAT 3+; 2018)

Kurzbeschreibung: Einsatzbereich sind die letzten sechs Wochen der dritten Klasse und die ersten sechs Wochen der vierten Klasse. Es ist ein Gruppentest zur Erfassung von besonderen mathematischen Leistungsstärken und -schwächen mit den Zielen.

Normierung: Es existieren nach Klassenstufe getrennte, aktualisierte und erweiterte Normen an 6.185 Grundschulkindern aus allen deutschen Bundesländern.. Dies sind Klassennormen zur Beurteilung der mittleren Leistungen und der Leistungsstreuung einer ganzen Schulklasse.

Bearbeitungsdauer: Nettobearbeitungstestzeit 28 Minuten. Gesamttestzeit ca. 45 Minuten im Gruppentest.

Reliabilität: Der DEMAT 3+ zeigt voll befriedigende Reliabilitätskoeffizienten (Paralleltestreliabilität von .83, Testhalbierungsreliabilität von .85 und Cronbachs Alpha von .83).

Validität: Es besteht curriculare Validität und inhaltliche Validität. Itemzusammenstellung orientiert auf der Basis einer Analyse der Lehrpläne aller deutschen Bundesländer an allen drei großen Inhaltskomplexen der Mathematik für die dritte Klassenstufe.

Empirische konkurrente Validität: Zusammenhang mit Schulnoten im Fach Mathematik in Höhe von .63 und Zusammenhang mit landesweiter Vergleichsarbeit im Fach Mathematik in Höhe von .66. Empirische prognostische Validität: Zusammenhang mit Mathematikleistung nach ca. 10 Monaten .68.

Deutscher Mathematiktest für die 4.Klasse (DEMAT 4; 2006)

Kurzbeschreibung: Der Einsatzbereich liegt in den drei Wochen vor und nach dem Halbjahreswechsel des vierten Schuljahres bzw. sechs Wochen vor Ende des vierten Schuljahres und dient zur Überprüfung der mathematischen Kompetenz in Bezug auf die Inhalte der Mathematiklehrpläne der 4. Klassen sowie zur frühen Diagnose einer Rechenschwäche bzw. besonderer Mathematikstärken.

Normierung: Es existieren Normen aus allen deutschen Bundesländern ($N = 5.266$). Klassen- und geschlechtsspezifische Normen: Neben den Individualnormen stehen Klassennormen zur Beurteilung ganzer Schulklassen zur Verfügung, welche die differenzierte Analyse der Stärken und Schwächen sowie der Unterschiedlichkeit innerhalb einer ganzen Schulklasse erlauben ($N = 287$).

Bearbeitungsdauer: Die Nettotestzeit beträgt 29,5 Minuten. Gesamttestdauer ca. 45 Minuten.

Reliabilität: Konsistenzschätzung über Cronbachs Alpha (40 Items): Mitte 4. Klasse ($N = 2.250$): .84. Ende 4. Klasse ($N = 3.016$): .85. Paralleltest-Reliabilität (innerhalb einer Woche, $N = 143$): .82.

Validität: Lehrplanvalidität: Der Testkonstruktion liegen die Mathematiklehrpläne aller 16 deutschen Bundesländer zu Grunde. Kriterienbezogene externe konkurrente Validität: Korrelation zu Schulnoten im Fach Mathematik, Ende 4. Klasse (N = 563): $r = -.70$; Deutsch, Ende 4. Klasse (N = 567): $r = -.58$; Korrelation mit der Empfehlung für die weitere Beschulung, Ende 4. Klasse (N = 617): $r = -.64$. Kriterienbezogene interne konkurrente Validität: Korrelation zu Testverfahren zur Erfassung mathematischer Grundlagenkenntnisse, Heidelberger Rechentest HRT (N = 81): $r = .72$.

Heidelberger Rechentest (HRT; 2005)

Kurzbeschreibung: Der HRT 1-4 ist zur Erfassung mathematischer Grundlagenkenntnisse als Gruppen- oder Einzeltest im Grundschulalter ab Ende der 1. Klasse bis Anfang der 5. Klassenstufe zu jedem Zeitpunkt des Schuljahres anwendbar.

Normierung: Die Eichstichprobe umfasst N = 3.075 Grundschul Kinder aus 4 Bundesländern sowie zusätzlich Kinder aus Sprachheilschulen und Förderschulen. Für die Normenberechnungen wurde der Anteil der Sonderschüler entsprechend der Verteilung im Primarbereich in Deutschland gewichtet. Für alle Untertests liegen Prozentrang- und T-Wert-Normen für 4 Quartale pro Schuljahr ab Ende der 1. Klasse bis Ende der 4. Klasse vor.

Bearbeitungsdauer: In Schulklassen beträgt die Bearbeitungszeit ca. 50 bis 60 Minuten, als Einzeltest ca. 45 Minuten

Reliabilität: Die Ergebnisse von Wiederholungsmessungen bei N = 246 Schülern im Abstand von 1–2 Wochen belegen eine hohe Messzuverlässigkeit der Skalenwerte ($r_{tt} = .87-.93$) und ausreichende Reliabilität der Untertests ($r_{tt} = .69-.89$) für eine zuverlässige Profilinterpretation.

Validität: Für die Aufgaben zur Beherrschung der Rechenoperationen kann inhaltliche Validität angenommen werden, d.h. die Aufgaben stellen eine typische Auswahl aus dem Spektrum der Grundrechenarten und Gleichungsaufgaben dar. Im Bereich kriteriumsbezogener Validität zeigt sich eine hohe und spezifische Übereinstimmung mit der Schulnote im Fach Mathematik ($r = -.67$) und mit dem DEMAT 4 ($r = .72$).

Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern – Kindergartenversion (Zareki-K; 2009)

Kurzbeschreibung: Der Erwerb rechnerischer Fertigkeiten beginnt nicht mit dem Schuleintritt, sondern baut auf Vorkenntnissen und Vorläuferfertigkeiten auf, die im Kleinkind- und Kindergartenalter erworben werden. Mit der ZAREKI-K werden bei Kinder im Alter von 5-7 Jahren die relevanten Aspekte des Zahlenverständnisses, der Zahlenverarbeitung und des rechnerischen Operierens möglichst separat in verschiedenen Aufgabengruppen geprüft.

Normierung: Es existiert eine Repräsentative Stichprobe von Kindergartenkindern des Kantons Zürich sowie High-Risk-Stichprobe von Kindern mit Entwicklungsauffälligkeiten

Bearbeitungsdauer: Es gibt keine Zeitbegrenzung, durchschnittlich werden ca. 30 bis 40 Minuten benötigt.

Reliabilität: Die Interne Konsistenz beträgt (Cronbachs Alpha) bei der Eichstichprobe .93 und bei der High-Risk-Stichprobe .94.

Validität: Die Homogenität der Anforderungen wurde nach dem dichotomen logistischen Modell von Rasch geprüft. Die prognostische Validität wurde über ein 2-Jahres-Intervall untersucht. 61,5 % der Kinder, die in der 2. Klasse laut ZAREKI-R-Gesamtwert eine Dyskalkulie entwickelten, waren bereits mittels ZAREKI-K als „auffällig“ klassifiziert worden. Die Rate falsch positiver Diagnosen mit ZAREKI-K betrug lediglich 4,5 %.

Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (Zareki-R; 2005)

Kurzbeschreibung: Zareki-R ist ein Testverfahren zur Erfassung einer Dyskalkulie bei Grundschulkindern (1. bis 4. Klasse).

Normierung: Es liegen Prozentrangnormen (Gesamt-N = 764) für die Klassenstufen 1 bis 4 vor, die an n = 421 deutschen Kindern aller Klassenstufen sowie n = 343 Schweizer Zweitklässlern gewonnen wurden.

Bearbeitungsdauer: Es gibt kein Zeitlimit, durchschnittlich werden zwischen 15 und 30 Minuten benötigt.

Reliabilität: Die Interne Konsistenz (Cronbachs Alpha) beträgt bei der Eichstichprobe $r = .89$ und bei der klinischen Gruppe $r = .90$.

Validität: Eine zur Konstruktvalidierung berechnete exploratorische Faktorenanalyse (n = 343) ergab eine interpretierbare Vierfaktorenlösung (Varianzaufklärung: 58,4 %) mit den Faktoren „Zahlen- und Faktenwissen“, „Analog-semantische und arithmetische Fähigkeiten“, „Zählfertigkeiten“ und „Numerisches Arbeitsgedächtnis“.

Lern- oder Merkfähigkeit

Verbaler Lern- und Merkfähigkeitstest (VLMT; 2001)

Kurzbeschreibung: Der VLMT ist ein Test zum seriellen Listenlernen mit nachfolgender Distraction, Abruf nach Distraction und halbstündiger Verzögerung sowie einem Wiedererkennungsdurchgang

Normierung: Die Normen stammen aus dem Jahr 2001 (N = 515).

Bearbeitungszeit: Die Bearbeitungszeit beträgt 20-25 Minuten effektive Testzeit- inklusive halbstündiger Verzögerung ca.50-55 Minuten.

Reliabilität: Zur internen Konsistenz werden keine Angaben gemacht. Zur Untersuchung der Retest-Reliabilität wurde bei einer nicht näher beschriebenen Gruppe von 149 Patienten eine zweite Messung nach 8 bis 12 Monaten mit einer der beiden Parallelformen durchgeführt. „Retest-Paralleltestkorrelationen“ zwischen den beiden Messungen wurden berechnet und darauf aufbauend kritische Differenzen für Leistungsveränderungen.

Validität: Summarisch werden faktorenanalytische Studien mit den VLMT-Variablen angeführt, auf die sich die Herausarbeitung der „Hauptvariablen“ des Tests stützt und die die geringe klinische Relevanz der Fehlervariablen unterstreichen. Faktorenanalysen im Kontext von Testbatterien würden nahe legen, dass der Test sowohl „Kurz- als auch Langzeitaspekte“ des Verbalgedächtnisses abzubilden vermag. Weiterhin wurde die Vergleichbarkeit von VLMT und ähnlichen international verwendeten Verfahren überprüft und bestätigt. In einer kleinen Vergleichsstudie mit 21 gesunden Kontrollprobanden zeigte sich, dass die zentralen Parameter des VLMT mit bildhaften Gedächtnisleistungen korrelieren, aber dass lediglich der erste Lerndurchgang (Dg1) mit Verfahren zum Arbeitsgedächtnis bedeutsam korrelieren, nicht aber die „Hauptparameter“.

Klinische Validität: Schwerpunkt bilden Studien zur funktionellen Neuroanatomie. Der VLMT erwies sich als sensitiv für linksseitige mesio-temporale Funktionsstörungen, insbesondere eine verzögerte Abrufleistung. Temporo-kortikale Läsionen wirken sich stärker auf das Lernen bzw. die Aufnahme ins Langzeitgedächtnis aus.

Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-jährige (Basic MLT; 2008)

Kurzbeschreibung: Der BASIC-MLT ist ein Verfahren zur Erfassung globaler sowie modalitäts- und funktionspezifischer Störungen der Merk- und Lernfähigkeit, das speziell für Kinder und Jugendliche konzipiert wurde. Überprüft werden die unmittelbare Merkspanne, die Fähigkeit zum Lernzuwachs inkl. Interferenzanfälligkeit, der verzögerte Abruf sowie Wiedererkennungsleistungen.

Normierung: Die Datenerhebung zur Normierung fand zwischen Februar 2005 Juli 2006 statt (N=405).

Bearbeitungszeit: Die Bearbeitungszeit der Kernbatterie beträgt 60 Minuten. Optionale Untertests dauern ca. 25 Minuten

Reliabilität: Die Reliabilität der einzelnen Skalen liegt zwischen $r = 0.78$ und $r = 0.86$.

Validität: Hinsichtlich der Konstruktvalidität erfolgten Validierungsstudien auf der Subtest-/Skalenebene über die Interkorrelationen der Subtestwerte und den Testskalen einerseits sowie über einen Stichprobenvergleich zwischen der Normstichprobe und einer klinischen Population andererseits.

Hinsichtlich der klinischen Validität wurden die diskriminanten diagnostischen Eigenschaften über den Vergleich der Normstichprobe mit einer klinischen Inanspruchnahmepopulation ermittelt. Dabei wurde deutlich, dass sich klinische und Normstichprobe signifikant voneinander unterscheiden.

Aufmerksamkeit

Aufmerksamkeits-Belastungstest (d2-R; 2010)

Kurzbeschreibung: Der d2-R ist ein Einzel- und Gruppentest für den Altersbereich von 9;0 bis 60;0 Jahren zur Untersuchung der individuellen Aufmerksamkeit und Konzentrationsfähigkeit. Er stellt eine Weiterentwicklung des Test d2 von 2002 dar, der zuletzt in der 9. Auflage erschienen war.

Normierung: Der Test wurde an rund 4.000 Personen für den Altersbereich von 9 bis 60 Jahren neu normiert (altersspezifische Normen).

Bearbeitungszeit: Die reine Testbearbeitungszeit beträgt 4 Minuten und 40 Sekunden. Die Instruktion nimmt einschließlich der Übungsaufgaben knapp 5 Minuten in Anspruch. Die Auswertung erfolgt manuell und dauert etwa 5 Minuten.

Reliabilität: Die Konzentrationsfähigkeit und die Schnelligkeit (Anzahl bearbeiteter Zielobjekte) werden mit hoher Messgenauigkeit erfasst (Cronbachs Alpha je nach Altersgruppe zwischen .89 und .95, Retest-Reliabilität nach 1 bzw. 10 Tagen = .94 und .91 bzw. .85 und .92). Der Kennwert für die Genauigkeit (Fehlerrate) weist befriedigende Reliabilitäten auf (Cronbachs Alpha: .80 bis .91, Retest-Reliabilität: .84 bzw. .47).

Validität: Zahlreiche Befunde zur Validität belegen, dass der Test Konzentrationsfähigkeit erfasst. Die Testleistung steht erwartungsgemäß nur in einer schwachen Beziehung zu Intelligenz, Leistungsmotivation, motorischer Schnelligkeit und Belastbarkeit. Zur Kriteriumsvalidität liegen aus mehreren Anwendungsbereichen (Studien aus dem Bereich: Verkehrs-, Pädagogische-, Umwelt-, Klinische und Medizinische Psychologie) unterstützende Befunde vor. Der Test kann nicht durch bestimmte Bearbeitungsstrategien „nach oben“ verfälscht werden. Eine Verfälschung „nach unten“ (Vortäuschen einer schlechten Konzentrationsfähigkeit) kann in vielen Fällen anhand unplausibler Buchstabenfehler- zwei oder mehr Buchstabenfehler- als Indikator für Täuschung -erkannt werden.

Testbatterie zur Aufmerksamkeitsprüfung TAP

Untertests: Alertness, Daueraufmerksamkeit, geteilte Aufmerksamkeit

Version 2.3.1 (Version 2021)

Kurzbeschreibung: Die Testbatterie zur Aufmerksamkeitsprüfung (TAP) bietet eine Sammlung unterschiedlicher Testverfahren, mit denen die vielfältigen Teilaspekte der Aufmerksamkeit differenziert erfasst werden können. Auch damit verbundene Aspekte der visuellen Wahrnehmung werden von der Testbatterie abgedeckt.

Normierung: Europa

Erhebung: Normierungsdaten aus unterschiedlichen Studien, aus denen jeweils unterschiedliche Teilmengen von Tests verwendet wurden;

Tests sind nicht in gleichem Umfang für Erwachsene und Jugendliche normiert

Alter: 6;0 bis 90;0 J. (abhängig vom jeweiligen Untertest)

Reliabilität: Testhalbierung (Odd-Even-Reliabilität): $r = 0.440 - 0.997$ (insgesamt für alle

Tests; Kinder- und Jugendliche: 6-19 J.); insgesamt: die Reliabilitäten der Reaktionszeitmediane liegen überwiegend über 0.90, daher zufriedenstellend bis sehr gut; die Reliabilitäten der Fehlermaße sind hingegen häufig unzureichend; Test-Retest: siehe vorläufige Ergebnisse für Erwachsene S.100; Es liegt eine Studie von Földényi et al. (2000) zur Reliabilität und Retest-Stabilität bei 95 Deutschschweizer Schulkindern vor (7-10 J.): $r_{tt} = 0.84$; insgesamt: mittlere bis hohe Reliabilitäten mit Ausnahme des Untertests: Geteilte Aufmerksamkeit: $r_{tt} = 0.32$; insgesamt: deutlich bessere Reliabilitäten bei Reaktionszeitmedianen im Vergleich zu Kriterien der Leistungsgüte (Auslassungen, Fehler).

Validität: Konstrukt: Studien zur faktoriellen Validität liegen vor.

Klinische Validität: Studien mit neurologischen Patienten, psychiatrischen Patienten, pharmakologische Studien, funktionelle Bildgebungsstudien, Studien mit Kindern und Studien mit älteren Probanden;

Studien mit Kindern:

Földényi et al. (2000): klinische Validierungsstudie an Kindern mit ADHS : 90 % der Kinder konnten korrekt der ADHS-Gruppe oder Kontrollgruppe zugeordnet werden anhand Reaktionszeit-Schwankungen im Go/Nogo, Fehleranzahl bei Flexibilität (Wechsel: nonverbal) und Testalter; Kriteriumsvalidität: Korrelation TAP und Eltern-/Lehrerurteile (standardisierte Fragebögen): signifikante Korrelation Go/Nogo, Inkompatibilität, geteilte Aufmerksamkeit, Flexibilität

Koschak et al. (2003): ADHS-Kontrollgruppe

Tucha et al. (2005): ADHS-Gesunde Kinder

Heubrock et al. (2001): Inanspruchnahme neuropsychologischer Ambulanz

Kunert et al. (1996): faktorielle Validität

Testbatterie zur Aufmerksamkeitsprüfung für Kinder KITAP (Untertests: Alertness, Daueraufmerksamkeit, geteilte Aufmerksamkeit)

Kurzbeschreibung: Die Testbatterie zur Aufmerksamkeitsprüfung – Kinderversion (KiTAP) dient der Erfassung der Aufmerksamkeit von Kindern im Alter von 6;0 bis 10;11 Jahre.

Normierung: Deutschland, 2002

Erhebung: Normierungsdaten aus unterschiedlichen Studien, aus denen jeweils unterschiedliche Teilmengen von Tests verwendet wurden aus Deutschland, Frankreich, Belgien, Italien, Schweiz, Österreich; Alter: 6;0 bis 10;0 J.

Reliabilität: Testhalbierung (Split-Half-Reliabilität):

$r = 0.64 - 0.97$ (insgesamt für alle Tests; Kinder: 6-7 J.)

$r = 0.63 - 0.96$ (insgesamt für alle Tests; Kinder: 8-10 J.)

Insgesamt: die Reliabilitäten der Reaktionszeitmediane deutlich besser, z.T. sehr gut; die Reliabilitäten der Fehlermaße sind hingegen häufig im unteren Bereich, aufgrund insgesamt eher geringer Auslassungen als Artefakte zu betrachten.

Validität: Konstrukt: Studien zur faktoriellen Validität liegen vor. Klinische Validität: Studien bei Kindern mit ADHS und nach Hirnschädigung.

Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche – (DISYPS-III; 2017)

Fremdbeurteilungsbogen / Selbstbeurteilungsbogen für den Störungsbereich

ADHS

(FBB-ADHS /SBB-ADHS)

Kurzbeschreibung: DISYPS-III ist die überarbeitete Version von DISYPS-II und dient der Erfassung von psychischen Störungen bei Kindern und Jugendlichen.

Die vorhandenen Materialien ermöglichen sowohl eine dimensionale Erfassung als auch eine kategoriale Diagnostik psychischer Störungen nach ICD-10 und DSM-5. Zudem liegt neu ein Screening vor, mit welchem Hinweise auf ein breites Spektrum an psychischen Störungen in relativ kurzer Zeit erhoben werden können, um danach gezielt störungsspezifische Instrumente des DISYPS-Systems einzusetzen. Auch das Screening umfasst Selbst- und Fremdeinschätzungen (Eltern, Lehrer/Erzieher) sowie eine Diagnose-Checkliste.

Hier sind die Skalen für den Störungsbereich ADHS interessant.

Normierung: Es liegen Repräsentativnormen für die Fremdbeurteilungsbogen (Elternurteil) und die Selbstbeurteilungsbogen zur Erfassung von ADHS, Störungen des Sozialverhaltens, Depressiven Störungen, Angststörungen, Zwangs-Spektrum-Störungen sowie für den Screeningbogen vor. Zusätzlich liegen klinische Normen für die Fremdbeurteilungsbogen (Elternurteil; Erzieher- und Lehrerurteil) und die Selbstbeurteilungsbogen für ADHS, Störungen des Sozialverhaltens, Depressive Störungen, Angststörungen und Tic-Störungen vor.

Bearbeitungszeit: Die Selbst- und Fremdbeurteilungsbogen dauern jeweils etwa 10 Minuten. Die Auswertung von Fragebogen bzw. Checklisten nimmt etwa 5 Minuten in Anspruch.

Reliabilität: Für die meisten Skalen der Fragebogen konnten mindestens zufriedenstellende interne Konsistenzen mit Werten von α gleich .70 bis α gleich .90 ermittelt werden.

Validität: Es liegen psychometrische Analysen für Repräsentativ-Stichproben und/oder für Klinik-Stichproben für alle DISYPS-II- bzw. DISYPS-III-Verfahren vor, mit Ausnahme der Verfahren zu Trauma- und Belastungsbezogenen Störungen sowie Bindungs- und Beziehungsstörungen. Die Konstruktvalidität der wichtigsten Verfahren wurde anhand von Faktorenanalysen überprüft.

Intelligence and Development Scales (IDS- 2 – Untertest „Aufmerksamkeit selektiv“; 2018)

Kurzbeschreibung: Die Intelligence and Development Scales - 2 (IDS-2) sind eine vollständige Überarbeitung und konzeptuelle Erweiterung der IDS. Die IDS-2 umfassen die Funktionsbereiche Intelligenz, Exekutive Funktionen, Psychomotorik, Sozial-Emotionale Kompetenz, Schulische Kompetenzen und Arbeitshaltung. Hier ist der Untertest „Aufmerksamkeit selektiv“ interessant.

Normierung: Die IDS-2 wurden von 2015 bis 2017 an 1.672 Kindern und Jugendlichen im Alter von 5;0 bis 20;11 Jahren aus Deutschland, Österreich und der Schweiz normiert.

Bearbeitungszeit: Die IDS-2 können modular eingesetzt und so an individuelle Fragestellungen angepasst werden. Der Untertest „Aufmerksamkeit selektiv“ insg. 9 Zeilen – pro Zeile max. 15 Sekunden

Reliabilität: Interne Konsistenz: Cronbach's Alpha für IQ-Werte = .96; Retest-Reliabilität nach 2 Wochen: Mittlere bis hohe Stabilitäten für Intelligenzfaktoren und Intelligenzwerte (.65-.89) sowie für Subtests exekutiver Funktionen (.72-.75) und sozial-emotionaler Kompetenz (.71-.85). Interne Konsistenz gegeben: Gute bis sehr gute Werte für Intelligenzfaktoren und Intelligenzwerte in einzelnen Altersgruppen (.77-.97) und für die Gesamtgruppe (.91-.98). Gute bis sehr gute Werte für Exekutive Funktionen (.88; Gesamtgruppe), Psychomotorik (.82-.87; 2 Altersgruppen), Sozial-emotionale Kompetenz (.82-.86; 2 Altersgruppen), Schulische Kompetenzen (.93-.97; 6 Altersgruppen bzw. 4 Klassenstufengruppen) und Arbeitshaltung (.90-.92; 3 Altersgruppen).

Validität: Es liegen Studien zur Konstruktvalidität (z. B. IDS, WISC-IV, M-ABC-2), Kriteriumsvalidität (z. B. Schulleistungen) und differenziellen Validität (z. B. ADHS, Autismus-Spektrum-Störung, Intelligenzminderung, Hochbegabung) vor.

Signifikante Unterschiede in kognitiven und allgemeinen Entwicklungsfunktionen zwischen durchschnittlich und überdurchschnittlich intelligenten oder mathematisch talentierten Kindern (mittlere bis große Effekte) bzw. Kindern mit Intelligenzminderung (große bis sehr große Effekte). Keine Unterschiede zwischen Kindern mit AD(H)S mit Medikation und nach Alter, Geschlecht und mütterlicher Ausbildung parallelisierter Kontrollgruppe. Erwartungsgemäße Unterschiede in Einzelskalen und/oder Gesamtttest beim Vergleich mit motorisch auffälligen Kindern oder Kindern mit Autismus-Spektrum-Störung. Nicht-erwartungsgemäß mittlere bis große Unterschiede in allen Skalen (außer Psychomotorik) für Jugendliche mit LRS, aber größte Effekte für Lesen und Rechtschreiben.

Konstruktvalidität gegeben: Überwiegend niedrige bis mittlere positive Korrelationen zwischen den kognitiven und allgemeinen Entwicklungsfunktionen. Intelligenzwerte korrelieren hoch untereinander (.77-.95) und mittelhoch mit Gesamtwert Exekutive Funktionen (.47-.64). Kriteriumsvalidität gegeben:

Zusammenhänge mit anderen Test und Fragebögen: WISC-IV (Petermann & Petermann, 2011): n = 114, mittlere Korrelationen von WISC-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.52-.69) und mit Exekutive Funktionen (.47); mittlere Korrelationen der WISC-Indizes mit den entsprechenden Intelligenzfaktoren (.43-.71). RIAS (Hagmann-von Arx & Grob, 2014): n = 180, mittlere Korrelationen von RIAS-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.47-.55). SON-R 6-40 (Tellegen et al., 2012) n = 139, mittlere Korrelationen von SON-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.43-.67).

Maße der Aufmerksamkeit und Exekutive Funktionen: n = 38, erwartungsgemäße mittlere Korrelationen mit TMT (Reitan, 1992; -.31/-.36), kleine bis mittlere mit TEA-Ch (Horn & Jäger, 2008; z.B. selektive Aufmerksamkeit: -.31). M-ABC-2 (Petermann, 2015): n = 54, mittlere Korrelationen von M-ABC-2-Gesamtwert und Psychomotorik (.49). ISK-K (Kanning, 2009): n = 182, kleine Korrelation von ISK-K-Gesamtwert mit Sozial-emotionale Kompetenz (.17).

Zusammenhänge mit Schulleistungen: Elterneinschätzungen: n = 294-728, Normierungsstichprobe; kleine bis mittlere Korrelationen von Einschätzungen in einzelnen Fächern mit IQ und IQ-Profil (.23-.41) und mit Schulische Kompetenzen (.27-.43). Noten: n = 194-478, Normierungsstichprobe; kleine bis mittlere Korrelationen von fachspezifischen Noten mit IQ und IQ-Profil (.17-.35), mit Schulische Kompetenzen (.16-.41) und mit

Arbeitshaltung (.16-.23). Lese- und Rechtschreibtests: $n = 60$, erwartungskonforme mittlere bis hohe Zusammenhänge der IDS-2-Subtests Lesen und Rechtschreiben mit verschiedenen sprachlichen Kompetenzen erfasst mittels DEUTSCH 9-10 (Segerer et al., in Vorb.; r zwischen .23-.73) und R-FIT 9-10 (Lenhart et al., in Vorb.; r zwischen .34-.76).

Persönlichkeitsmaße: mittlere Korrelationen von Arbeitshaltung mit Elterneinschätzung der Gewissenhaftigkeit mittels FFFK (Asendorpf, 1998; $n = 785$, $r = .28$), Selbsteinschätzung der Gewissenhaftigkeit mittels NEO-PI-R (Ostendorf & Angleitner, 2004; $n = 185$, $r = .61$) und verschiedenen Fragebögen zur Erfassung der Leistungsmotivation (.51-.65; n zwischen 184 bis 358).

Intelligence and Development Scales (IDS- P 2013- Untertest „Aufmerksamkeit selektiv“; 2018)

Kurzbeschreibung: Die IDS-P eignen sich für das gesamte Spektrum der Leistungs- und Entwicklungsdiagnostik, für die Kindergarten- und Schuleingangsdiagnostik sowie für klinische Fragestellungen. Die IDS-P erfassen mit insgesamt 15 Untertests fünf Funktionsbereiche: Kognition (7 Untertests), Psychomotorik (3 Untertests), Sozial-Emotionale Kompetenz (1 Untertest), Denken Logisch-Mathematisch (1 Untertest) und Sprache (3 Untertests). Hier ist der Untertest „Aufmerksamkeit selektiv“ interessant.

Normierung: Die IDS-P wurden an 700 Kindern aus Deutschland, der deutschsprachigen Schweiz und Österreich normiert. Der Gesamttest dauert circa 60 bis 90 Minuten.

Bearbeitungszeit: Die einzelnen Funktionsbereiche dauern: Kognition: ca. 40 Minuten (darunter „Aufmerksamkeit selektiv“ Zeitbegrenzung 90 Sekunden) ; Psychomotorik: ca. 15 Minuten; Sozial-Emotionale Kompetenz: ca. 5 Minuten; Denken Logisch-Mathematisch: ca. 5 Minuten; Sprache: ca. 10 Minuten.

Reliabilität: Die Interne Konsistenz der Funktionsbereiche beträgt: Kognition (Intelligenz) $\alpha = .91$, Psychomotorik $\alpha = .92$, Sozial-Emotionale Kompetenz $\alpha = .72$, Denken Logisch-Mathematisch $\alpha = .84$, Sprache $\alpha = .85$

Retest-Reliabilität der Funktionsbereiche nach 5 Monaten: Kognition (Intelligenz) $r_{tt} = .90$, Psychomotorik $r_{tt} = .85$, Sozial-Emotionale Kompetenz $r_{tt} = .53$, Denken Logisch-Mathematisch $r_{tt} = .80$, Sprache $r_{tt} = .69$.

Validität: Eine konfirmatorische Faktorenanalyse bestätigt die erwartete Struktur der Funktionsbereiche.

Kriteriumsvalidität erfolgte durch Vergleiche mit bewährten Verfahren (K-ABC, SETK 3-5, RIAS) sowie standardisierten Fragebogen für Eltern und Lehrkräfte (CBQ, SDQ). Kriteriumsvalidität weitgehend gegeben: Reynolds Intellectual Assessment Scale (RIAS; Hagemann von Arx & Grob, 2014): $n = 54$ Schweizer Kinder aus Normstichprobe (28 Jungen), Durchschnittsalter 4;6 Jahre; zeitlicher Abstand zwischen IDS-P und RIAS im Durchschnitt 4 Tage; mittlere Korrelation des RIAS-Gesamtwerts mit IDS-P Funktionsbereich Kognition (.63); Zusammenhänge mit anderen Funktionsbereichen werden

nicht berichtet. K-ABC (Melchers & Preuß, 2009): n = 26 Kinder (9 Jungen), Durchschnittsalter 4;5 Jahre; zeitlicher Abstand zwischen IDS-P und K-ABC wird nicht genannt; hohe Korrelation des Untertests Rechnen der K-ABC mit Funktionsbereich Denken logisch-mathematisch (.86); Zusammenhänge mit anderen Funktionsbereichen werden nicht berichtet. Sprachentwicklungstest für drei- bis fünfjährige Kinder (SETK 3-5; Grimm, 2001): n = 26 Schweizer Kinder aus Normstichprobe (12 Jungen), Durchschnittsalter 4;8 Jahre; zeitlicher Abstand zwischen IDS-P und SETK 3-5 wird nicht genannt; mittlere Korrelationen der Skalen des SETK 3-5 und des Funktionsbereiches Sprache (.49 - .61). Strength and Difficulties Questionnaire (SDQ; Goodman, 1997): Befragung von 534 Eltern und 451 extrafamiliären Bezugspersonen (keine weiteren Angaben) von 254 Mädchen und 280 Jungen, Durchschnittsalter 4;8 Jahre; niedrige Korrelationen der Skalen des SDQ mit dem Funktionsbereich sozial-emotionale Kompetenz (.11 - .18) und den drei Subtests zum Umgang mit der Testsituation (.09 - .28). Die Leistungen von fremdsprachigen Kindern, Kindern mit Trisomie 21, sprachentwicklungsauffälligen Kindern, Kindern mit allgemeiner Entwicklungsauffälligkeit und frühgeborenen Kindern unterscheiden sich von den Leistungen unauffälliger Kinder, was ein Beleg für die differentielle Validität ist.

Soziale Fertigkeiten oder Verhalten

Child Behavior Checklist Elternfragebogen (CBCL 1,5-5;2021)

Child Behavior Checklist für Erzieher (C-TRF 1.5-5; 2021)

Kurzbeschreibung: Die deutschen Formen der Child Behavior Checklist für Klein- und Vorschulkinder dienen der Erfassung von Verhaltensauffälligkeiten, emotionalen Auffälligkeiten und somatischen Beschwerden.

Die CBCL 1½-5 wird von Eltern bzw. Erziehungsberechtigten oder anderen Personen ausgefüllt, die das Kind in einer familiären Umgebung erleben. Die Anwendung der C-TRF 1½-5 wird für Erzieher/innen oder andere Personen empfohlen, die das Verhalten des Kindes in einer Gruppe von mindestens vier Kindern beurteilen können. Der Beurteilungszeitraum bezieht sich in beiden Fragebögen auf die letzten zwei Monate.

Normierung: Es stehen Normdaten (T-Werte) aus Feldstichproben für den Altersbereich 2;0 bis 5;11 Jahre zur Verfügung (CBCL 1½-5: N = 850; C-TRF 1½-5: N = 815 Jungen, N = 764 Mädchen). Zusätzlich können im Altersbereich 2;0 bis 5;11 Jahre für die Problemskalen beider Versionen klinische Vergleichswerte (Stanine, Prozentrangbänder) aus einer gemischten klinischen Stichprobe bestimmt werden (CBCL 1½-5: N = 155; C-TRF 1½-5: N = 169).

Bearbeitungszeit: Die Bearbeitung dauert pro Fragebogen etwa 10 bis 15 Minuten. Die Auswertung nimmt zusätzlich ca. 10 -15 Minuten in Anspruch.

Reliabilität: In beiden Perspektiven (CBCL/1½-5, C-TRF/1½-5) ist über alle Stichproben hinweg die interne Konsistenz der Skalen zweiter Ordnung (Internale Probleme, Externale Probleme) mit Werten für Cronbachs Alpha > .82 als durchgehend mindestens sehr gut zu bewerten. Die interne Konsistenz der Gesamtauffälligkeit fällt mit einem Alpha von mindestens .92 in beiden Perspektiven über alle Stichproben hinweg exzellent aus.

Validität: Die Problemskalen wurden im amerikanischen Original empirisch auf der Basis faktorenanalytischer Methoden abgeleitet. Die Faktorenmodelle beider Instrumente konnten mittels konfirmatorischer Faktorenanalysen an deutschen Datensätzen zufriedenstellend repliziert werden. Die Faktorenmodelle können zudem nach Analysen kulturübergreifender internationaler Datensätze durch die Autoren der amerikanischen Originalversion als robust angesehen werden.

Child Behavior Checklist Elternfragebogen (CBCL 6-18R,2014)

Child Behavior Checklist Lehrerfragebogen (TRF 6-18R;2014)

Child Behavior Checklist für Kinder und Jugendliche (YSR 11-18R; 2014)

Kurzbeschreibung: Die CBCL 6-18R richtet sich an Eltern von Kindern und Jugendlichen von 6 bis 18 Jahren, die TRF 6-18R an Lehrer von Kindern und Jugendlichen von 6 bis 18 Jahren und die YSR 11-18R an Kinder und Jugendliche im Alter von 11 bis 18 Jahren. Die Child Behavior Checklist (CBCL 6-18R) dient der Erfassung von Verhaltensauffälligkeiten, emotionalen Auffälligkeiten, somatischen Beschwerden sowie sozialen Kompetenzen von Kindern und Jugendlichen im Schulalter aus Sicht der Eltern. Die beiden Fragebögen TRF 6-18R und YSR 11-18R wurden direkt aus der CBCL abgeleitet und erlauben eine Beurteilung der weitgehend identischen Merkmale aus Sicht von Lehrern und Jugendlichen.

Normierung: Bundesweit repräsentative Normdaten (T-Werte) liegen für die CBCL 6-18R (N = 2.471) und die YSR 11-18R (N = 1.798) vor. Für die TRF 6-18R stehen Vergleichsdaten einer Großstadtstichprobe von Grundschulkindern (6-11 Jahre) (N = 397) zur Verfügung. Für alle drei Fragebogenfassungen liegen zudem Normdaten (Stanine-Werte und Prozenträge) einer Klinikstichprobe (Aufteilung Patienten aus einer Kinder- und Jugendpsychiatrischen Klinik und einer psychotherapeutischen Ausbildungsambulanz anhand der Achse I-Diagnose des MAS in die Gruppe „Internale Störung“, „Externale Störung“ und „Andere gemischte Störung“) vor (CBCL/6-18R: N = 1.217; YSR/11-18R: N = 718; TRF/6-18R: N = 793).

Bearbeitungszeit: Die Bearbeitung dauert pro Fragebogen etwa 15 bis 20 Minuten.

Reliabilität: Sowohl in Klinik- als auch in Feldstichproben liegt bei allen drei Fragebögen die interne Konsistenz der Skalen zweiter Ordnung (Internale Probleme, Externale Probleme) mit Werten für Cronbachs Alpha $> .80$ im guten Bereich. Die interne Konsistenz der Gesamtauffälligkeit ist mit einem Cronbachs Alpha von mindestens $.93$ in allen drei Bögen sehr gut.

Validität: Es liegen internationale Studien an repräsentativen Feldstichproben vor, welche die (im Original US-amerikanischen) Faktorenmodelle der drei Fragebogenversionen mittels konfirmatorischer Faktorenanalyse kulturübergreifend überzeugend replizieren, und damit die Skalierung 2001 bestätigen.

Verhaltensfragebogen bei Entwicklungsstörungen (VFE; 2007)

Kurzbeschreibung: Deutschsprachige Bearbeitung der Developmental Behavior Checklist (DBC) von Einfeld und Tongue; die an der deutschen Stichprobe durchgeführten Analysen findet sich separat (Steinhausen und Winkler, 2005); die deutsche Normierung wurde an 721 Kindern und Jugendlichen ermittelt aus 75 Institutionen

Normierung: Deutschland, 2007; Alter: 4;0 – 18;0 J.

Reliabilität: Re-test: ICC = 0.83 – 0.89 (deutsche Normierung)
Interrater (Eltern): ICC = 0.80 (DBC: australische Normierung),
Interrater (Veränderungssensitivität: DBC und klinische Beurteilungswerte):
r = 0.86 (DBC: australische Normierung).

Validität: Inhalt: Krankengeschichten, Interviews, erfahrene Kliniker.
Konstrukt: Faktorenanalyse: interne Konsistenz am niedrigsten bei ‚Angst‘, Zuverlässigkeit dieser Subskala unbefriedigend, ansonsten für VFE-E gegeben: $\alpha = 0.58 - 0.90$
Diskriminante (Interkorrelationen der Subskalen sowie der 3 Grade der GB für VFE-E):
überwiegend unter $r = 0.60$ (hinreichende Unabhängigkeit der Skalen), $r = 0.17 - 0.69$
Differenzierungsfähigkeit: fetales Alkoholsyndrom, Prader-Willi-Syndrom, fragiles X-Syndrom, tuberöse Sklerose sowie Differenzierung zwischen Autismus und geistiger Behinderung (Steinhausen & Winkler Metzke, 2004).
Klinische Validität des DBC: überprüft an Autismus, Depression, Psychose, Hyperaktivität, Angst.

Sensitivität für den DBC-Gesamtwert: 83%
Spezifität für den DBC-Gesamtwert: 85%

Strengths and Difficulties Questionnaire (SDQ-DEU ;2002)

Kurzbeschreibung: Deutschsprachige Bearbeitung der Strengths and Difficulties Questionnaire (SDQ) von Goodman (die deutsche Übersetzung erfolgte 1997). Es existiert eine Elternversion, Lehrerversion und Selbstversion.

Die deutsche Normierung (Elternversion) wurde an 930 Eltern ermittelt.

Anmerkung: frei verfügbarer Fragebogen: www.sdqinfo.org, hier finden sich Übersetzungen und eine Auflistung und Zusammenfassung bisheriger Publikationen

Normierung: Deutschland, 2002 (Elternversion); Alter 6;0 – 16;0 J.

Reliabilität: Cronbachs Alpha: $\alpha = 0.58 - 0.76$ (einzelne Skalen).
Cronbachs Alpha: $\alpha = 0.82$ (Gesamtproblemwert).
Es liegt eine zufriedenstellende interne Konsistenz vor (Klasen et al., 2003).

Validität: Inhalt: orientiert an älteren Fragebögen (Rutter) bzw. Faktorenanalysen längerer Fragebögen und nosologischen Konzepten.

Konstrukt: Faktorenanalysen: Bestätigung der englischen Skalen

Kriteriumsbezogen: $r = 0.68 - 0.82$ (Skalen des SDQ und CBCL; Klasen et al., 2000)

Klinische Validität: überprüft an einer klinischen Stichprobe (Becker et al., 2004)

Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche – (DISYPS-III; 2017)
Fremdbeurteilungsbogen / Selbstbeurteilungsbogen für den Störungsbereich Störungen des Sozialverhaltens (FBB-SSV / SBB-SSV)

Kurzbeschreibung: DISYPS-III ist die überarbeitete Version von DISYPS-II und dient der Erfassung von psychischen Störungen bei Kindern und Jugendlichen.

Die vorhandenen Materialien ermöglichen sowohl eine dimensionale Erfassung als auch eine kategoriale Diagnostik psychischer Störungen nach ICD-10 und DSM-5. Zudem liegt neu ein Screening vor, mit welchem Hinweise auf ein breites Spektrum an psychischen Störungen in relativ kurzer Zeit erhoben werden können, um danach gezielt störungsspezifische Instrumente des DISYPS-Systems einzusetzen. Auch das Screening umfasst Selbst- und Fremdeinschätzungen (Eltern, Lehrer/Erzieher) sowie eine Diagnose-Checkliste.

Hier sind die Skalen für den Störungsbereich des Sozialverhalten interessant.

Normierung: Es liegen Repräsentativnormen für die Fremdbeurteilungsbogen (Elternurteil) und die Selbstbeurteilungsbogen zur Erfassung von ADHS, Störungen des Sozialverhaltens, Depressiven Störungen, Angststörungen, Zwangs-Spektrum-Störungen sowie für den Screeningbogen vor. Zusätzlich liegen klinische Normen für die Fremdbeurteilungsbogen (Elternurteil; Erzieher- und Lehrerurteil) und die Selbstbeurteilungsbogen für ADHS, Störungen des Sozialverhaltens, Depressive Störungen, Angststörungen und Tic-Störungen vor.

Bearbeitungszeit: Die Selbst- und Fremdbeurteilungsbogen dauern jeweils etwa 10 Minuten. Die Auswertung von Fragebogen bzw. Checklisten nimmt etwa 5 Minuten in Anspruch.

Reliabilität: Für die meisten Skalen der Fragebogen konnten mindestens zufriedenstellende interne Konsistenzen mit Werten von α gleich .70 bis α gleich .90 ermittelt werden.

Validität: Es liegen psychometrische Analysen für Repräsentativ-Stichproben und/oder für Klinik-Stichproben für alle DISYPS-II- bzw. DISYPS-III-Verfahren vor, mit Ausnahme der Verfahren zu Trauma- und Belastungsbezogenen Störungen sowie Bindungs- und Beziehungsstörungen. Die Konstruktvalidität der wichtigsten Verfahren wurde anhand von Faktorenanalysen überprüft.

Intelligence and Development Scales (IDS- 2 – Untertests “Emotionen erkennen”, “Emotionen regulieren”, Soziale Situationen verstehen”, “Sozial kompetent handeln”; 2018)

Kurzbeschreibung: Die Intelligence and Development Scales - 2 (IDS-2) sind eine vollständige Überarbeitung und konzeptuelle Erweiterung der IDS. Die IDS-2 umfassen die Funktionsbereiche Intelligenz, Exekutive Funktionen, Psychomotorik, Sozial-Emotionale Kompetenz, Schulische Kompetenzen und Arbeitshaltung.

Hier sind die Untertests im Bereich Emotionaler Kompetenz interessant wie: „Emotionen erkennen“, „Emotionen regulieren“, „Soziale Situationen verstehen“ oder „Sozial kompetent“ handeln. Das Modul Sozial-Emotionale Kompetenz dauert ca. 15 Minuten.

Normierung: Die IDS-2 wurden von 2015 bis 2017 an 1.672 Kindern und Jugendlichen im Alter von 5;0 bis 20;11 Jahren aus Deutschland, Österreich und der Schweiz normiert.

Bearbeitungszeit: Die IDS-2 können modular eingesetzt und so an individuelle Fragestellungen angepasst werden. Das Modul Sozial-Emotionale Kompetenz dauert ca. 15 Minuten.

Reliabilität: Interne Konsistenz: Cronbach's Alpha für IQ-Werte = .96; Retest-Reliabilität nach 2 Wochen: Mittlere bis hohe Stabilitäten für Intelligenzfaktoren und Intelligenzwerte (.65-.89) sowie für Subtests exekutiver Funktionen (.72-.75) und sozial-emotionaler Kompetenz (.71-.85). Interne Konsistenz gegeben: Gute bis sehr gute Werte für Intelligenzfaktoren und Intelligenzwerte in einzelnen Altersgruppen (.77-.97) und für die Gesamtgruppe (.91-.98). Gute bis sehr gute Werte für Exekutive Funktionen (.88; Gesamtgruppe), Psychomotorik (.82-.87; 2 Altersgruppen), Sozial-emotionale Kompetenz (.82-.86; 2 Altersgruppen), Schulische Kompetenzen (.93-.97; 6 Altersgruppen bzw. 4 Klassenstufengruppen) und Arbeitshaltung (.90-.92; 3 Altersgruppen).

Validität: Es liegen Studien zur Konstruktvalidität (z. B. IDS, WISC-IV, M-ABC-2), Kriteriumsvalidität (z. B. Schulleistungen) und differenziellen Validität (z. B. ADHS, Autismus-Spektrum-Störung, Intelligenzminderung, Hochbegabung) vor.

Signifikante Unterschiede in kognitiven und allgemeinen Entwicklungsfunktionen zwischen durchschnittlich und überdurchschnittlich intelligenten oder mathematisch talentierten Kindern (mittlere bis große Effekte) bzw. Kindern mit Intelligenzminderung (große bis sehr große Effekte). Keine Unterschiede zwischen Kindern mit AD(H)S mit Medikation und nach Alter, Geschlecht und mütterlicher Ausbildung parallelisierter Kontrollgruppe. Erwartungsgemäße Unterschiede in Einzelskalen und/oder Gesamtttest beim Vergleich mit motorisch auffälligen Kindern oder Kindern mit Autismus-Spektrum-Störung. Nicht-erwartungsgemäß mittlere bis große Unterschiede in allen Skalen (außer Psychomotorik) für Jugendliche mit LRS, aber größte Effekte für Lesen und Rechtschreiben.

Konstruktvalidität gegeben: Überwiegend niedrige bis mittlere positive Korrelationen zwischen den kognitiven und allgemeinen Entwicklungsfunktionen. Intelligenzwerte korrelieren hoch untereinander (.77-.95) und mittelhoch mit Gesamtwert Exekutive Funktionen (.47-.64).

Kriteriumsvalidität gegeben: Zusammenhänge mit anderen Test und Fragebögen: WISC-IV (Petermann & Petermann, 2011): n = 114, mittlere Korrelationen von WISC-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.52-.69) und mit Exekutive Funktionen (.47); mittlere Korrelationen der WISC-Indizes mit den entsprechenden Intelligenzfaktoren (.43-.71). RIAS (Hagmann-von Arx & Grob, 2014): n = 180, mittlere Korrelationen von RIAS-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.47-.55). SON-R 6-40 (Tellegen et al., 2012) n = 139, mittlere Korrelationen von SON-Gesamtwert mit Intelligenzwerten und Schulischen Kompetenzen (.43-.67).

Maße der Aufmerksamkeit und Exekutive Funktionen: n = 38, erwartungsgemäße mittlere Korrelationen mit TMT (Reitan, 1992; -.31/-.36), kleine bis mittlere mit TEA-Ch (Horn & Jäger, 2008; z.B. selektive Aufmerksamkeit: -.31). M-ABC-2 (Petermann, 2015): n = 54, mittlere Korrelationen von M-ABC-2-Gesamtwert und Psychomotorik (.49). ISK-K (Kanning, 2009): n = 182, kleine Korrelation von ISK-K-Gesamtwert mit Sozial-emotionale Kompetenz (.17).

Zusammenhänge mit Schulleistungen: Elterneinschätzungen: n = 294-728, Normierungsstichprobe; kleine bis mittlere Korrelationen von Einschätzungen in einzelnen

Fächern mit IQ und IQ-Profil (.23-.41) und mit Schulische Kompetenzen (.27-.43). Noten: $n = 194-478$, Normierungsstichprobe; kleine bis mittlere Korrelationen von fachspezifischen Noten mit IQ und IQ-Profil (.17-.35), mit Schulische Kompetenzen (.16-.41) und mit Arbeitshaltung (.16-.23). Lese- und Rechtschreibtests: $n = 60$, erwartungskonforme mittlere bis hohe Zusammenhänge der IDS-2-Subtests Lesen und Rechtschreiben mit verschiedenen sprachlichen Kompetenzen erfasst mittels DEUTSCH 9-10 (Segerer et al., in Vorb.; r zwischen .23-.73) und R-FIT 9-10 (Lenhart et al., in Vorb.; r zwischen .34-.76). Persönlichkeitsmaße: mittlere Korrelationen von Arbeitshaltung mit Elterneinschätzung der Gewissenhaftigkeit mittels FFFK (Asendorpf, 1998; $n = 785$, $r = .28$), Selbsteinschätzung der Gewissenhaftigkeit mittels NEO-PI-R (Ostendorf & Angleitner, 2004; $n = 185$, $r = .61$) und verschiedenen Fragebögen zur Erfassung der Leistungsmotivation (.51-.65; n zwischen 184 bis 358).

Intelligence and Development Scales (IDS- P; 2013)

Kurzbeschreibung: Die IDS-P eignen sich für das gesamte Spektrum der Leistungs- und Entwicklungsdiagnostik, für die Kindergarten- und Schuleingangsdiagnostik sowie für klinische Fragestellungen. Die IDS-P erfassen mit insgesamt 15 Untertests fünf Funktionsbereiche: Kognition (7 Untertests), Psychomotorik (3 Untertests), Sozial-Emotionale Kompetenz (1 Untertest), Denken Logisch-Mathematisch (1 Untertest) und Sprache (3 Untertests). Hier ist der Funktionsbereich Sozial-Emotionale Kompetenz interessant.

Normierung: Die IDS-P wurden an 700 Kindern aus Deutschland, der deutschsprachigen Schweiz und Österreich normiert. Der Gesamttest dauert circa 60 bis 90 Minuten.

Bearbeitungszeit: Die einzelnen Funktionsbereiche dauern: Kognition: ca. 40 Minuten; Psychomotorik: ca. 15 Minuten; Sozial-Emotionale Kompetenz: ca. 5 Minuten; Denken Logisch-Mathematisch: ca. 5 Minuten; Sprache: ca. 10 Minuten.

Reliabilität: Die Interne Konsistenz der Funktionsbereiche beträgt: Kognition (Intelligenz) $\alpha = .91$, Psychomotorik $\alpha = .92$, Sozial-Emotionale Kompetenz $\alpha = .72$, Denken Logisch-Mathematisch $\alpha = .84$, Sprache $\alpha = .85$

Retest-Reliabilität der Funktionsbereiche nach 5 Monaten: Kognition (Intelligenz) $r_{tt} = .90$, Psychomotorik $r_{tt} = .85$, Sozial-Emotionale Kompetenz $r_{tt} = .53$, Denken Logisch-Mathematisch $r_{tt} = .80$, Sprache $r_{tt} = .69$.

Validität: Eine konfirmatorische Faktorenanalyse bestätigt die erwartete Struktur der Funktionsbereiche.

Kriteriumsvalidität erfolgte durch Vergleiche mit bewährten Verfahren (K-ABC, SETK 3-5, RIAS) sowie standardisierten Fragebogen für Eltern und Lehrkräfte (CBQ, SDQ). Kriteriumsvalidität weitgehend gegeben: Reynolds Intellectual Assessment Scale (RIAS; Hagemann von Arx & Grob, 2014): $n = 54$ Schweizer Kinder aus Normstichprobe (28 Jungen), Durchschnittsalter 4;6 Jahre; zeitlicher Abstand zwischen IDS-P und RIAS im Durchschnitt 4 Tage; mittlere Korrelation des RIAS-Gesamtwerts mit IDS-P Funktionsbereich Kognition (.63); Zusammenhänge mit anderen Funktionsbereichen werden nicht berichtet. K-ABC (Melchers & Preuß, 2009): $n = 26$ Kinder (9 Jungen), Durchschnittsalter 4;5 Jahre; zeitlicher Abstand zwischen IDS-P und K-ABC wird nicht

genannt; hohe Korrelation des Untertests Rechnen der K-ABC mit Funktionsbereich Denken logisch-mathematisch (.86); Zusammenhänge mit anderen Funktionsbereichen werden nicht berichtet. Sprachentwicklungstest für drei- bis fünfjährige Kinder (SETK 3-5; Grimm, 2001): n = 26 Schweizer Kinder aus Normstichprobe (12 Jungen), Durchschnittsalter 4;8 Jahre; zeitlicher Abstand zwischen IDS-P und SETK 3-5 wird nicht genannt; mittlere Korrelationen der Skalen des SETK 3-5 und des Funktionsbereiches Sprache (.49 - .61). Strength and Difficulties Questionnaire (SDQ; Goodman, 1997): Befragung von 534 Eltern und 451 extrafamiliären Bezugspersonen (keine weiteren Angaben) von 254 Mädchen und 280 Jungen, Durchschnittsalter 4;8 Jahre; niedrige Korrelationen der Skalen des SDQ mit dem Funktionsbereich sozial-emotionale Kompetenz (.11 - .18) und den drei Subtests zum Umgang mit der Testsituation (.09 - .28). Die Leistungen von fremdsprachigen Kindern, Kindern mit Trisomie 21, sprachentwicklungsauffälligen Kindern, Kindern mit allgemeiner Entwicklungsauffälligkeit und frühgeborenen Kindern unterscheiden sich von den Leistungen unauffälliger Kinder, was ein Beleg für die differentielle Validität ist.