



Stand 29.09.2020

Paulina Cordero Donoso & Gerolf Renner

Paulina Cordero Donoso arbeitet seit 2011 als Psychologin innerhalb der sozial-psychiatrischen Versorgung von Kindern und Jugendlichen. Des Weiteren ist sie als Dozentin in der Ausbildung von Kinder- und Jugendlichenpsychotherapeut*innen und als Seminarleiterin im Bereich der Intelligenz- und Leistungsdiagnostik tätig.

Zitationsempfehlung:

Cordero Donoso, P. & Renner, G. (2020). *Testinformation zu den Raven's 2. Deutsche Fassung der Raven's 2 Progressive Matrices – Clinical Edition* (Dia-Inform Verfahrensinformationen 007-01). Ludwigsburg: Pädagogische Hochschule Ludwigsburg.

**Dieser Text wird veröffentlicht unter der Creative-Commons Lizenz
CC BY-NC-ND 4.0 DE**

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Diese Publikation darf dementsprechend nur unter Nennung der Urheber und ausschließlich zu nicht-kommerziellen Zwecken genutzt und weiterverbreitet werden, Modifikationen des Textes sind nicht zugelassen.

Potenzieller Interessenkonflikt:

Dia-Inform will unabhängige und nicht von kommerziellen Interessen beeinflusste Informationen zur Verfügung stellen. Wir legen deshalb Wert auf eine sorgfältige Dokumentation aller Tatsachen, die auf einen potenziellen Interessenkonflikt hinweisen könnten.

Es liegen keine Interessenskonflikte vor.

Das Projekt Dia-Inform

Projektleitung und Korrespondenz:

Prof. Dr. Gerolf Renner
Pädagogische Hochschule Ludwigsburg
Fakultät für Sonderpädagogik
Reuteallee 46
71634 Ludwigsburg
renner@ph-ludwigsburg.de

Prof. Dr. Markus Scholz
Pädagogische Hochschule Ludwigsburg
Fakultät für Sonderpädagogik
Reuteallee 46
71634 Ludwigsburg
markus.scholz@ph-ludwigsburg.de

Projektinformation:

Ziel des Projektes Dia-Inform an der Pädagogischen Hochschule Ludwigsburg ist es, ein umfangreiches Angebot an Informationsmaterialien zu diagnostischen Verfahren bereit zu stellen, die in der (sonder-)pädagogischen Diagnostik eingesetzt werden.

Informationen zu diagnostischen Verfahren sowie Informations-, Lehr- und Praxisvideos bilden ein abgestuftes System, das für unterschiedliche Informationsbedürfnisse von Lehrenden und Studierenden und verschiedene Einsatzzwecke (Veranstaltungen, Selbststudium, Vorbereitung auf diagnostische Gutachten) flexibel und passgenau verwendet werden kann und in dessen Weiterentwicklung Studierende aktiv mit eingebunden werden können.

Im Rahmen des Projektes soll ein umfassender, nicht von kommerziellen Interessen beeinflusster und anderweitig nicht verfügbarer Informationspool zu diagnostischen Verfahren entstehen.

Dia-Inform Verfahrensinformationen:

Dia-Inform Verfahrensinformationen geben einen Überblick über theoretische Grundlagen, Aufbau und psychometrische Eigenschaften eines diagnostischen Instrumentes. Sie weisen eine einheitliche Struktur auf, die Leser/innen die Orientierung erleichtert. Die Verfahrensinformationen diskutieren Stärken und Schwächen der Verfahren und weisen auf sinnvolle Einsatzmöglichkeiten in der sonderpädagogischen und klinisch-psychologischen Diagnostik hin. Dabei werden insbesondere die spezifischen Anforderungen einer Diagnostik bei Kindern und Jugendlichen mit Behinderungen berücksichtigt.

Dia-Inform Verfahrensinformationen werden von mindestens zwei Autor/innen gemeinsam verfasst, die unterschiedliche professionelle Perspektiven einbringen.

Dia-Inform Verfahrensinformationen stehen Ihnen kostenfrei auf OPUS-PHLB, dem Hochschulschriftenserver der PH Ludwigsburg, unter folgendem Link zur Verfügung:

<https://phbl-opus.phlb.de/solrsearch/index/search/searchtype/collection/id/16235>

Raven's 2

Deutsche Fassung der Raven's Progressive Matrices 2 – Clinical Edition

Vorbemerkung

Diese Testinformation bezieht sich in erster Linie auf den Einsatz der Raven's 2 bei Kindern und Jugendlichen.

Autor*innen

NCS Pearson (Hrsg.)

Erscheinungsjahr

2019

Altersbereich

4;0–69;11 Jahre

Verlag

Pearson, Frankfurt

Preis des Testverfahrens und der Verbrauchsmaterialien laut Verlagsseite

Test komplett 450,00 €; 50 Antwortbogen 150,00 €.

Lizenz zur digitalen Auswertung von Testungen der Papierform für 1 Jahr 79,00 €, für drei Jahre 199,00 €. Digitalisierte Durchführung, Auswertung und Berichterstellung der Kurzform 4,00 €, der Langform 5,00 €.

Bei Gruppentestungen muss für jede Testperson ein Stimulusbuch (Einzelpreis 120,00 €) bzw. bei digitaler Testung ein PC oder Tablet zur Verfügung stehen.

Alle Angaben zuzüglich Mehrwertsteuer.

Diagnostische Zielsetzung lt. Testmanual

Bei den Raven's 2 handelt es sich um eine Revision der drei Versionen der Raven's Progressive Matrices Testserie: Coloured Progressiv Matrices (CPM; J. C. Raven, J. Raven & Court, 2002) zur Erfassung der kognitiven Fähigkeiten im unteren Alters- und Leistungsbereich, Standard Progressive Matrices (SPM; J. C. Raven, 2009) zur Erfassung im mittleren Leistungsbereich und Advanced Progressive Matrices (APM; J. C. Raven, J. Raven & Court, 1998) zur Erfassung im oberen Leistungsbereich. Die Raven's 2 integrieren alle drei Verfahren in einem Test.

Die Raven's 2 sind laut Manual als „Verfahren mit nonverbalen Testinhalten zur Erfassung der allgemeinen kognitiven Fähigkeit“ konstruiert worden (Manual, S.13). Die Bearbeitung der Testaufgaben erfordere „kognitive Funktionen wie Wahrnehmung und Aufmerksamkeit für visuelle Details, induktives Schlussfolgern, fluide Intelligenz, breite visuelle Intelligenz, Klassifikationsfähigkeiten und räumliche Fähigkeiten, simultane Verarbeitung sowie Arbeitsgedächtnisleistungen“ (Manual, S. 14). Als Anwendungskontexte nennt das Manual psychiatrische und schulpsychologische Untersuchungen, Eignungsdiagnostik und Rehabilitationsmaßnahmen sowie Forschungsstudien. Besonders geeignet seien die Raven's 2 „für die Anwendung bei Personen mit Einschränkungen in der expressiven und rezeptiven Sprache, bei tauben oder schwerhörigen Personen oder bei

der Erfassung der kognitiven Fähigkeiten von (sprachbezogenen) Minderheiten im Bereich der Hochbegabung“ (Manual, S. 24).

Die Raven's 2 können nach Einschätzung der Testautoren bei diagnostischen Fragestellungen eingesetzt werden, bei denen die Erfassung der allgemeinen kognitiven Fähigkeiten durch eine mehrdimensionale Testbatterie entweder nicht möglich oder nicht erforderlich ist. Sie seien im Sinne eines Screening-Instruments geeignet den weiteren Bedarf einer ausführlicheren Leistungsdiagnostik festzustellen, „z. B. bei Lernstörungen, Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung, intellektueller Beeinträchtigung oder Hochbegabung“ (Manual, S. 23).

Einschränkend verweist das Manual darauf, dass die Raven's 2 „keine umfassende Evaluation der kognitiven Fähigkeit einer Testperson“ darstellen. Die Raven's 2 sollten daher „nicht als alleiniges Maß zur Diagnosestellung oder für Schulplatzierungsentscheidungen verwendet werden“ (Manual, S. 24).

Theoretischer Hintergrund

Die Raven's 2 stehen in der Tradition einer Testserie, die von dem britischen Psychologen John Carlyle Raven in den 30er Jahren des 20. Jahrhunderts begründet wurde. Die Raven-Tests wurden zur Erfassung einer Komponente der Allgemeinintelligenz entwickelt, die Spearman (1927) als „eduction“ bezeichnete. Dieser heute nicht mehr geläufige Begriff beschreibt die Fähigkeit von Menschen, auf Basis begrenzter Informationen Strukturen und Bedeutungen zu erkennen und neue Einsichten zu entwickeln („to construct meaning out of confusion“; J. Raven, 2008). Das Manual grenzt das eingeschränktere Konstrukt der eduktiven Fähigkeiten konzeptuell vom g-Faktor ab, der als gemeinsames Element aller kognitiven Tests verstanden wird, betont aber auch, dass die Raven-Tests eines der besten Maße von g darstellen (dazu kritisch Gignac, 2015). Diese begriffliche Abgrenzung wird an anderen Stellen im Manual und im Ergebnisbericht des Testsystems Q-Global wieder relativiert, und die Raven's 2 werden als Test zur Erfassung der allgemeinen kognitiven Fähigkeiten bezeichnet.

Setting, Durchführungsdauer und Auswertungsmöglichkeiten

Die Raven's 2 liegen in drei Formen vor: Papierform, digitale Langform und digitale Kurzform (s. Tabelle 1). Sie können als Einzel- oder Gruppentest durchgeführt werden.

Die maximalen Bearbeitungszeiten (reine Testzeit ohne Vorbereitung und Einführung) für die verschiedenen Testformen sind in Tabelle 1 aufgeführt. Angaben zur effektiven Bearbeitungszeit in der europäischen Normstichprobe liegen nicht vor (s. Tabelle 1 zu entsprechenden Daten aus der amerikanischen Normierung).

Digitale Testungen werden automatisch ausgewertet. Die Papierform kann manuell anhand einer Auswertungsschablone und der Normtabellen im Manual ausgewertet werden. Alternativ ist eine internetbasierte Auswertung und Berichterstellung über das Testsystem Q-global möglich. Der Bericht beinhaltet eine Beschreibung der erfassten Konstrukte und eine übersichtliche Darstellung der Ergebniswerte.

Informationen für Anwender*innen früherer Auflagen

Die wichtigsten Änderungen gegenüber früheren Verfahren auf einen Blick:

- CPM, SPM und APM wurden in einem Testverfahren integriert.
- Neuer Aufgabentyp (1 × 6 Reihe)
- Alle Items mehrfarbig gestaltet
- Möglichkeit einer digitalen Anwendung
- Einführung einer Abbruchregel bei digitalen Testungen
- Vereinfachte und verkürzte Testinstruktionen

Testaufbau, Testmaterial und Durchführung

Den Testpersonen werden Muster mit geometrischen Figuren vorgelegt, die nach einem bestimmten logischen Prinzip aufgebaut sind. Ein Teil des Musters fehlt. Die Aufgabe der Testpersonen besteht darin, das jeweilige Prinzip, das einer Aufgabe zugrunde liegt, zu erkennen, das fehlende Teil aus fünf vorgegebenen Antwortmöglichkeiten zu finden, und auf einem Antwortbogen die Nummer der richtigen Lösung zu markieren. Die Papierform besteht aus fünf Aufgabensets (A-E) mit jeweils 12 Aufgaben. Bei einer digitalen Testanwendung wird ein individuelles Set von Aufgaben automatisch aus einer Itembank von 329 Aufgaben generiert.

Tabelle 1

Aufbau der Raven's 2

Testform	Altersgruppe	Item-Zahl	Maximale Bearbeitungszeit	Bearbeitungszeit 50 % / 90 % ^a
Papierform	4;0–8;11	36 Aufgaben (Set A, B & C)	30 Min.	12 / 22 Min.
	9;0–69;11 ^b	48 Aufgaben (Set B, C, D & E)	45 Min.	23 / 43 Min.
Digitale Langform	4;0–8;11	36 Aufgaben	30 Min.	11 / 21 Min.
	9;0–69;11	48 Aufgaben	45 Min.	20 / 42 Min.
Digitale Kurzform	4;0–69;11	24 Aufgaben	20 Min.	10 / 19 Min.

Anmerkungen: ^a Daten aus der amerikanischen Normstichprobe (Manual, S. 31), nicht differenziert nach Altersgruppen. Angegeben sind die Zeiten, in denen 50 % bzw. 90 % der Testpersonen die Bearbeitung abgeschlossen hatten.

^b Bei Personen mit einer niedrigen Leistungsfähigkeit ist alternativ die Vorgabe der Sets A, B und C möglich.

Die durchgehend farbig gestalteten Aufgaben haben vier unterschiedliche Formate (vgl. Abbildung 1): Aufgaben mit einer 1 × 1-Struktur zeigen einen Rahmen, der mit einem lückenhaften Muster gefüllt ist; die Testpersonen müssen das fehlende Teil des Musters finden. Aufgaben mit einer 2 × 2-Struktur bestehen aus einem Muster mit drei Figuren, zu denen die passende vierte auszusuchen ist. Bei einer 3 × 3-Struktur enthalten die Aufgaben ein Muster mit acht Figuren, zu denen die neunte passende Figur gefunden

werden muss. Bei der neuen 1 × 6-Struktur zeigen die Aufgaben eine horizontal angeordnete Reihe aus sechs Figuren, bei der eine fehlende Figur gefunden werden muss.

Der Testkoffer für die Papierform enthält das Manual, ein Stimulusbuch, einen Satz Antwortbogen und eine Auswertungsschablone. Für die Durchführung einer digitalen Testung ist ein Zugang zur webbasierten Plattform Q-global erforderlich. Ein Computer oder Tablet mit Internetzugang wird für jede Testperson benötigt. Ein Übungslink für den/die Testleiter*in, um sich mit der digitalen Durchführung vertraut zu machen, wird zur Verfügung gestellt. Zusätzlich wird für alle Testformen eine Stoppuhr benötigt.

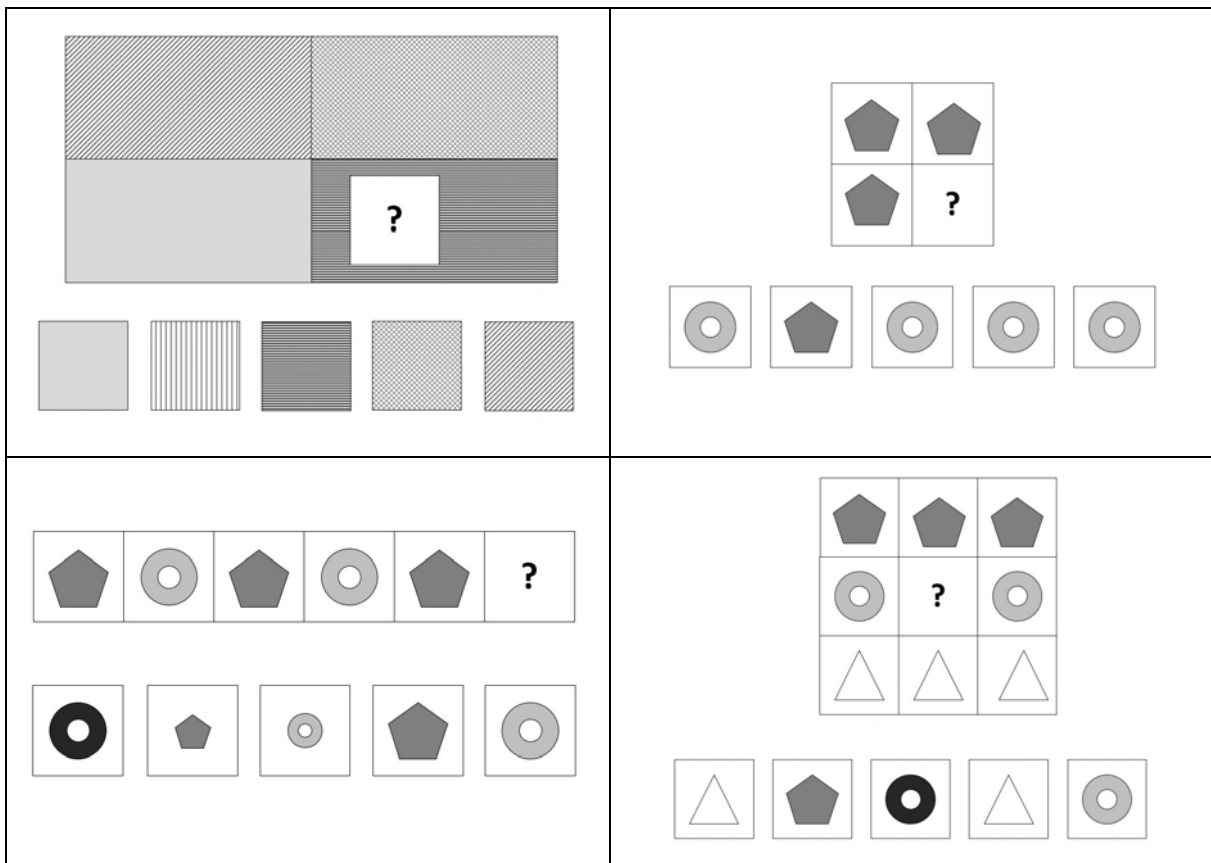


Abbildung 1. Aufgabenformate der Raven's 2 (keine Original-Items)

Die Anweisungen werden den Testpersonen bei allen Testformen vorgelesen und können mehrmals wiederholt und bei Bedarf in eigenen Worten formuliert werden (Manual, S. 44). Die digitale Testung beginnt mit sogenannten Demonstrationsaufgaben. Es handelt sich um eine animierte Anleitung am Monitor, die – begleitet von verbalen Hinweisen der Testleiter*innen – den Ablauf der digitalen Testung erläutert. Es folgen Übungsaufgaben, die in allen Testformen eingesetzt werden. Sie sollen sicherstellen, dass die Instruktionen verstanden wurden. Bei falscher Beantwortung einer Übungsaufgabe in der Papierform erhalten die Testpersonen „erweiterte Instruktionen“ mit einer korrigierenden Rückmeldung und ergänzenden Erläuterung der Aufgabenanforderungen. Bei einer digitalen Testung werden den Testpersonen ihre Fehler sofort automatisch angezeigt; sie erhalten eine animierte korrigierende Rückmeldung und werden dann aufgefordert eine neue Lösung zu wählen. Bei erneut falscher Lösung werden ihnen die erweiterten Instruktionen von den Testleiter*innen vorgelesen. Eine positive Rückmeldung bei einer richtigen Lösung erfolgt bei den Übungsaufgaben ausschließlich bei einer digitalen Testung.

Nur in den digitalen Testformen ist ein Abbruchkriterium definiert: Die Testung endet, wenn nach Item 10 sechs Items in Folge nicht gelöst wurden.

Auch in der digitalen Version wird die Bearbeitungszeit nicht automatisch überwacht und festgehalten. Bei Unterbrechungen der Testung, die sich z. B. durch Nachfragen der Testpersonen oder Pausen ergeben, muss die Zeitmessung gestoppt und anschließend wieder aufgenommen werden. Wie dies in einer Gruppentestung praktisch realisiert werden kann, bei der es bei einzelnen Teilnehmer*innen zu unterschiedlichen Zeitpunkten zu Unterbrechungen kommen kann, lässt das Manual offen.

Ergebnis(-werte)

Zentrale Ergebniswerte:

- IQ-Werte ($M = 100$, $SD = 15$) mit 90%- und 95%-Vertrauensintervall
- Prozenträge

Weitere Ergebniswerte:

- Leistungswerte. Dabei handelt es sich um auf der Item-Response-Theory basierende Werte, die aus den Itemparametern des Rasch-Modells abgeleitet werden. Leistungswerte erlauben u. a. den Vergleich der Leistungen von Personen, die verschiedene Testformen und Testsets durchgeführt haben.
- Testaltersäquivalente (Entwicklungsalter) für den Altersbereich von 4;0–19;11 Jahren

Objektivität

Die Testanweisungen sind größtenteils eindeutig formuliert und garantieren weitgehend eine *objektive Testdurchführung*. Es werden überwiegend – aber nicht durchgehend – wörtliche Instruktionen vorgegeben. Nicht hinreichend konkret geregelt ist die Option, dass Testleiter*innen die Protokollierung der Antworten übernehmen können, wenn dies den Testpersonen „nicht möglich ist“ (Manual, S. 27). Die Instruktion lässt offen, wann das Kriterium „nicht möglich“ erfüllt ist. Neben der offensichtlichen Möglichkeit, dass die Testperson nicht über die motorischen Fähigkeiten verfügt, die Lösung zu markieren oder die Maus zu bedienen, stellt sich die Frage, ob dies auch für Personen gilt, die Schwierigkeiten haben, sich auf dem Testformular zu orientieren, oder die aufgrund von Problemen in der Aufmerksamkeit und Selbststeuerung mit der selbstständigen Bearbeitung der Aufgaben überfordert sind. Besonders bei den jüngsten Altersgruppen und bei Testpersonen mit exekutiven Schwächen könnte das Ausmaß der Steuerung durch die Testleiter*innen einen Einfluss auf die Testleistungen haben.

Generell kommen in den Instruktionen Besonderheiten bei der Testung von jungen Kindern und Testpersonen mit Behinderungen zu kurz (z. B. angemessene Sitzposition bei Kindern im Vorschulalter, angemessene Beleuchtung bei Personen mit Sehbeeinträchtigungen usw.). Auf Besonderheiten der Testung bei stark sprachbeeinträchtigten Kindern sowie bei jüngeren Kindern, die nach Erhalten der *erweiterten Instruktionen* weiterhin Verständnisprobleme aufweisen, wird nicht näher eingegangen. Hinweise zum Umgang mit schwierigen Testsituationen (z. B. bei Verweigerung der Mitarbeit) bietet das Manual nicht.

Bei der digitalen Testung ist die erforderliche Mindestgröße des Bildschirms nicht geregelt. Der Übungslink, der die Testleiter*innen mit der Durchführung der digitalen Testung vertraut machen soll, ist auf Q-global nicht intuitiv zu finden. Dieser ist erst, nachdem eine Testung einer neuen Testperson zugewiesen wurde, unter dem Link „Test-Hilfe“ zu finden.

Die Länge der möglichen Testpausen wird wenig exakt als „kurz“ beschrieben.

Der altersabhängige Testanstieg und die maximale Testzeit sind festgelegt, wobei den Testleiter*innen auch ein gewisser Spielraum eingeräumt wird, eine längere Bearbeitungszeit zuzulassen.

Empfehlungen zur Beziehungsgestaltung, zur Testvorbereitung und zur Gruppengröße bei Gruppentestungen liegen vor.

Die Auswertungsobjektivität ist durch festgelegte Lösungsvorgaben gesichert. Die Interpretationsobjektivität ist durch den Bezug auf Normwerte gewährleistet.

Reliabilität

Die Reliabilitätskoeffizienten der Raven's 2 wurden für die Papierform und für beide digitale Formen mittels der Split-half-Methode berechnet. Für die Papierform basieren die Werte auf der europäischen Normierungsstichprobe. Die Reliabilitäten der digitalen Testformen gehen nur auf die US-amerikanische Normierungsstichprobe zurück. Für die verschiedenen Altersgruppen liegen die Reliabilitätsschätzungen vorwiegend im befriedigenden Bereich (Tabelle 2).

Die niedrigsten Reliabilitätswerte im Rahmen der europäischen Normstichprobe der Papierform finden sich in den Altersgruppen von 4 und 8 Jahren (.75 und .76), die höchsten in den Altersgruppen von 15 und 16 Jahren (.92 und .90) sowie 20 bis 24 Jahren (.93). Reliabilitäten für Personen mit Intelligenzminderung werden für alle Testvarianten im Manual aufgeführt. Ein Koeffizient von .85 wird bei der Papierform der europäischen Version genannt.

Für die Papierform und für die beiden digitalen Formen wurden Retest-Reliabilitäten bei einer US-amerikanischen Stichprobe von 239 Personen ermittelt. Die Werte liegen zwischen .80 und .89; Übungseffekte zeigen einen Zuwachs von 0.9 bis 5.5 IQ-Punkten. Für die Papierform liegen Retest-Reliabilitäten bei altersgemischten Stichproben aus den Niederlanden (29 Personen) und Spanien (101 Personen) bei .92 bzw. .80, mit einem mittleren Zuwachs von 4.5 bzw. 4.2 IQ-Punkten. Aus der europäischen Normierung liegen keine Daten zur Paralleltestreliabilität (Papier – Digital) vor.

Tabelle 2

Ausgewählte Reliabilitätskennwerte der Raven's 2 Progressive Matrices

Testform (Datenherkunft)	Spannbreite der Split-half-Werte (4-19 Jahre)	Spannbreite der Split-half-Werte (20-69 Jahre)	Retest- Stabilität (4-16 Jahre)
Papierform (Europa; Retest USA)	.75 - .92	.85 - .93	.85 - .88
Digitale Langform (USA)	.85 - .89	.89 - .90	.84 - .87
Digitale Kurzform (USA)	.79 - .81	.80 - .81	.82

Anmerkung: Angaben laut Testmanual.

Validität

Inhaltliche Validität: Die in den Raven's 2 verwendeten Matrizenaufgaben sind fraglos inhaltlich valide für den Bereich der fluiden Intelligenz. Bei der Entwicklung der amerikanischen Version der Raven's 2 seien Items von Experten des Testverlags und externen Experten im Hinblick auf die „inhaltliche Abdeckung und Relevanz“ (Manual, S.

81) geprüft worden. Es sei bei der Itemkonstruktion auch darauf geachtet worden, konstruktirrelevante Einflüsse (z. B. irreführende Distraktoren, farbliche Gestaltung) auszuschließen. Zu den dabei angewandten Kriterien finden sich keine näheren Angaben.

Faktorielle Validität: Die Raven's 2 sind ein eindimensionales Verfahren. Daten zur Frage, ob diese Faktorenstruktur angemessen ist, finden sich im Manual nicht, obwohl die Dimensionalität im Rahmen des verwendeten Rasch-Modells überprüft werden kann.

Konvergente Validität: Zum Zusammenhang der europäischen Raven's 2 mit anderen intelligenzdiagnostischen Verfahren werden keine Daten vorgelegt. Die Autoren berichten lediglich von entsprechenden US-amerikanischen Befunden zu den früheren Ausgaben der Raven's Progressive Matrices (CPM, SPM und APM).

Zum Zusammenhang mit Schulleistungen, zur *divergenten Validität* und zur *prognostischen Validität* liegen keine Daten vor.

Diskriminative Validität: Die Autoren berichten über Ergebnisse einer US-amerikanischen Studie (digitale Testform) und einer Studie im Rahmen der europäischen Normierung (Papierform), bei denen jeweils Personen mit Hochbegabung und Personen mit intellektuellen Beeinträchtigungen mit den Raven's 2 untersucht wurden. Als Kriterium wurden bereits vorliegende IQ-Werte verwendet, die mindestens zwei Standardabweichungen über bzw. unter dem Mittelwert liegen mussten. Eine gesicherte klinische Diagnose einer Intelligenzminderung lag wohl nicht vor. In der Validierungsstudie der Extremgruppenvergleiche für die europäische Fassung (Papierform) fällt der mittlere IQ mit einem Wert von 123.3 in der Gruppe der Personen mit Hochbegabung signifikant höher als die Werte der gematchten Kontrollstichprobe. Personen mit einer intellektuellen Beeinträchtigung (N=173) im Alter zwischen 4 und 48 Jahren zeigen signifikant niedrigere Werte (mittlerer IQ 64.8) im Vergleich zu der gematchten Kontrollstichprobe aus der Normierungsstudie.

Normierung

Die Ermittlung europäischer Normen der Papierform der Raven's 2 erfolgte von Mai 2018 bis Juli 2019. Die Normstichprobe umfasst insgesamt 1.200 Personen mit jeweils gleich großen Teilstichproben aus sechs europäischen Ländern bzw. Regionen (Deutschland, Frankreich, Niederlande, Spanien, Skandinavien und Großbritannien). Die Untersuchung der deutschen Teilstichprobe wurde „durch ein renommiertes Sozialforschungsinstitut durchgeführt“ (Manual, S. 62). Die Testleiter*innen wurden lt. Manual für die Testdurchführung der Raven's 2 sowie für die Rekrutierung der Stichprobe geschult. Ausführliche Auskünfte zur Qualifikation, Rekrutierung und Qualitätskontrolle der Testdurchführungen geben die Autoren nicht.

Zur Gewährleistung der Repräsentativität der europäischen Teilstichproben erfolgte die Rekrutierung nach Quotenplänen. In der Stichprobenbeschreibung wird über die Stratifizierung nach Alter, Geschlecht, Bildungshintergrund (höchster Schulabschluss; bis 19;11 Jahre höchster Schulabschluss der Hauptbezugsperson), teilweise hinsichtlich der regionalen Herkunft (Spanien) und der ethnischen Zugehörigkeit (Großbritannien) berichtet. Eine Gleichverteilung nach Geschlecht ist in einigen Ländern nicht in allen Altersgruppen gegeben. Die Einteilung des Bildungshintergrundes entspricht laut Manual weitgehend der Verteilung in den beteiligten Ländern. Nicht in die Normierungsstichprobe aufgenommen wurden Personen mit verminderten Hör- oder Sehleistungen, mit beeinträchtigter Fein- und Grobmotorik sowie mit neurologischen und/oder psychiatrischen Störungen.

Weitere detaillierte Informationen zur Stichprobenrekrutierung finden sich im Manual nicht. Somit kann die Repräsentativität der Normen nicht abschließend beurteilt werden.

Für die Normierung wurden alle Teilstichproben zu einer europäischen Gesamtstichprobe zusammengefasst. Laut Manual habe es keine signifikanten Unterschiede zwischen den Mittelwerten der einzelnen Regionen gegeben. Eine eigene Nachberechnung (einfaktorielle Varianzanalyse auf Basis der im Manual angegebenen Mittelwerte und Standardabweichungen) zeigt dagegen, dass durchaus ein signifikanter Effekt der Datenherkunft vorliegt, allerdings bei kleiner Effektstärke.

Die Normwerte wurden auf Basis eines kontinuierlichen Normierungsmodells berechnet und sind im Manual für 36 Altersgruppen tabelliert: für die 4- bis 8-Jährigen im 4-Monatsabstand; für die 9- bis 13-Jährigen im 6-Monatsabstand; für die 15- und 16-Jährigen im Jahresabstand; danach in Abständen von 3 bis 10 Jahren.

In keiner Altersgruppe fanden sich relevante Boden- und Deckeneffekte. Einem Rohwert von 1 wird maximal ein IQ-Wert von 42 zugeordnet.

Problematische Itemgradienten sind in den Normtabellen schwer zu erkennen, da dort zahlreiche Leistungswerte eingetragen sind, die in der Normstichprobe gar nicht vorkamen. Stichprobenartige Prüfungen zeigen, dass es vereinzelt zu deutlichen Unterschieden in den Standardwerten kommen kann, wenn sich der Rohwert um 1 verändert. Dies betrifft vor allem den oberen ($IQ > 140$) und den unteren ($IQ < 80$) Normwertbereich und konnte in den fünf überprüften Altersgruppen im Extremfall einen Unterschied von 13 IQ-Punkten ausmachen (z. B. zwischen Rohwert 5 und Rohwert 6 bei den 4;0-Jährigen).

Abrupte Normwertveränderungen beim Übergang zwischen den normierten Altersgruppen konnten nicht festgestellt werden.

Bei einem Test mit Items im Multiple-Choice-Format ist auch zu bedenken, welche Rohwerte bei rein zufälliger Aufgabenbeantwortung erzielt werden können. In der Papierform für 4- bis 8-Jährige (36 Items mit jeweils 20% Lösungswahrscheinlichkeit) führt der entsprechende Rohwert im Alter von 4;0 Jahren zu einem IQ von 66 und im Alter von 5;0 Jahren zu einem IQ von 60. Es ist also nicht zu erwarten, dass eine zufällige Bearbeitung unterdurchschnittliche Fähigkeiten maskieren kann.

Weitere Gütekriterien

Ökonomie: Für relativ hohe Anschaffungskosten bieten die Raven's 2 die Abdeckung eines sehr engen Spektrums kognitiver Leistungen. Ihre Zeitökonomie, charakterisiert durch den kurzen Vorbereitungs-, Durchführungs-, und Auswertungsaufwand, geht zu Lasten einer deutlich eingeschränkten Konstruktrepräsentanz.

Zielgruppengerechte Gestaltung

Bei der folgenden Bewertung beziehen wir uns in erster Linie auf Kinder- und Jugendliche, die in klinisch-psychologischen und sonderpädagogischen Anwendungskontexten untersucht werden.

Kinder mit grafomotorischen Schwierigkeiten und Kinder mit Sehbeeinträchtigungen können mit der selbstständigen Protokollierung ihrer Lösungen auf dem eng bedruckt Antwortbogen überfordert sein. Da es nur einen Antwortbogen für alle Altersgruppen gibt, müssen Kinder auch beachten, dass die Spalten für Set D und Set E für sie irrelevant sind.

Die Standarddurchführung der Raven's 2 erwartet von den Testpersonen, dass sie nach Vorgabe der Übungitems 36 Items (4 bis 8 Jahre) bzw. 48 Items (ab 9 Jahre) eigenständig bearbeiten. Das ist für die jüngsten Kinder generell eine hohe Anforderung und wird auch für ältere Kinder mit Beeinträchtigungen in Aufmerksamkeit und Selbststeuerung eine erhebliche Herausforderung oder Überforderung darstellen. Die Option, auch ab dem Alter von 9 Jahren die Sets A, B und C vorzugeben, sollte daher bei Testpersonen mit vermuteten kognitiven Einschränkungen stets bedacht werden. Zudem werden anfangs einige relevante Anweisungen gegeben, die erst später im Testablauf zu beachten sind (z. B. „Instruieren Sie die Testpersonen ggf. außerdem, die grüne Zwischenseite vor Set B zu ignorieren und einfach weiterzuarbeiten“; Manual, S. 49) und die daher eine entsprechende Merkfähigkeit voraussetzen. Mit dem Hinweis, dass 30 Minuten Bearbeitungszeit zur Verfügung stehen, werden 4-Jährige wenig anfangen können. Anweisungen wie „Wenn Du z. B. die Übungsaufgabe A (...) bearbeitet hast, fülle bitte deine Antwort auf dem Antwortbogen in der Zeile ÜA aus“ sind keine Musterbeispiele für eine entwicklungsangemessene Sprache im Vorschulalter.

Bei den Testinstruktionen der Papierform findet sich eine ungünstige Formulierung, die bei den Probanden zu Missverständnissen führen kann: „Wenn du glaubst, dass du das richtige Teil gefunden hast, male den entsprechenden Kreis auf dem Antwortbogen aus“ (Manual, S.45). Nach eigenen Erfahrungen (PC) verstehen einige Kinder dies – verständlicherweise – als Aufforderung, den für die Markierung der richtigen Lösung vorgesehenen Kreis sorgfältig auszumalen.

Einzelne Begriffe der erweiterten Instruktionen wie z. B. „Quadrat“ (Manual, S. 46) oder „Reihenfolge“ (Manual, S. 47) erschweren Vorschulkindern und Kindern mit kognitiver Einschränkung das Aufgabenverständnis. Somit ist der/die Testleiter*in gefordert, diese Begriffe in andere Worte zu fassen oder zu erklären.

Das Testmaterial ist wenig abwechslungsreich und bietet keine aktiven Handlungsmöglichkeiten. Das kommt den entwicklungstypischen Bedürfnissen der jüngsten Altersgruppen nicht entgegen.

Bei Durchführung der Papierform gibt es keine Abbruchregel. Das bedeutet für jüngere Kinder im unteren Leistungsbereich, dass sie überwiegend Aufgaben bearbeiten müssen, die sie nicht bewältigen können. Beispielsweise löst ein 5;0-jähriges Kind mit einem IQ von 85 nur 13 von 36 Aufgaben, bei einem IQ von 60 sogar nur 7 von 36.

Zugangsfertigkeiten / Einsatz bei behinderten Kindern / Testfairness

Da ausschließlich visuelles Material verwendet wird, sind die Raven's 2 generell für Personen mit Sehbehinderungen nicht geeignet. Bei direkter künstlicher Beleuchtung kann es auf den Vorlagen zu Reflexblendungen kommen. Laut Manual wurden mögliche Auswirkungen von Störungen der Farbwahrnehmung bei der Itemgestaltung berücksichtigt. In einigen ausgewählten Simulationen ließen sich keine gravierenden Erschwernisse für verschiedene Formen der Farbenblindheit feststellen; lediglich bei Vorliegen der extrem seltenen Achromatopsie wird die Testfairness erheblich beeinträchtigt.

Die motorischen Anforderungen im Rahmen einer Einzeltestung beschränken sich auf eine einfache Zeigereaktion, wenn die Protokollierung und das Umblättern im Stimulusbuch von den Testleiter*innen übernommen werden. Es liegen keine speziellen Durchführungshinweise für gehörlose Kinder vor.

CHC-theoretische Einordnung

In der Cattell-Horn-Carroll Intelligenztheorie (Schneider & McGrew, 2018; deutschsprachige Darstellung Mickley & Renner, 2019) können die Raven's 2 dem Schicht-II-Faktor Fluide Intelligenz (Schicht I: Induktion) zugeordnet werden. Fluide Intelligenz wird hier als Fähigkeit beschrieben, neuartige Probleme ohne Rückgriff auf vorhandenes Wissen lösen zu können.

Stärken

- Aktuelle Normierung
- Einsatz über einen breiten Altersbereich möglich
- Einsatz über einen sehr breiten Leistungsbereich möglich (± 4 SD)
- Frei von Boden- und Deckeneffekten
- Als Einzeltest weitgehend unkomplizierte Handhabung
- Einsatz als Gruppentest möglich (wenn auch aus unserer Sicht für die klinische Anwendung nicht zu empfehlen)
- Bei den meisten Testpersonen geringe Durchführungszeit
- Reliabilität wurde auch an einer Stichprobe von Menschen mit Intelligenzminderungen überprüft
- Leichtere Handhabung durch separate Instruktionen für Kinder (Du-Form) und Erwachsene (Sie-Form)
- Mehrere Übungssitems zur Sicherung des Instruktionsverständnisses, bei digitaler Durchführung zusätzliche Demonstrationsitems
- Sowohl digitale Testung als auch Testung in Papierform möglich
- Im Vergleich zu den früheren Versionen erspart die Abbruchregel bei digitaler Durchführung den Testpersonen im (weit) unterdurchschnittlichen Leistungsbereich die frustrierende Bearbeitung von vielen zu schwierigen Items.
- Potenziell sehr große Anzahl von Parallelformen bei digitaler Testung.

Einschränkungen / Schwächen

- Eingeschränkte Konstruktrepräsentanz
- Stichprobenumfang angesichts des breiten Altersbereichs eher gering
- Daten zur Stichprobenrekrutierung wenig detailliert aufgeführt
- Zur Qualifikation der Testleiter*innen und zur Qualitätskontrolle bei der Normierung werden nur vage Informationen gegeben
- Wenig abwechslungsreiche Aufgabenstellungen
- Antwortbogen nicht kindgerecht gestaltet
- Im Fall von Unterbrechungen der Testung Zeitmessung potenziell fehleranfällig
- Keine Daten zur Angemessenheit der eindimensionalen Teststruktur
- Keine Daten zum Zusammenhang mit Schulleistungen
- Keine Daten zur konvergenten Validität der europäischen Version mit anderen Intelligenztests
- Keine Daten zur Übereinstimmung der verschiedenen Durchführungsvarianten
- Nur spärliche Daten zur Validität bei klinischen Gruppen
- Das Verfahren wird als Screening-Instrument bezeichnet, es liegen aber keinerlei Daten zu Sensitivität und Spezifität vor.
- Keine empirischen Belege für die behauptete Kulturfairness (vgl. zu dieser Problematik Brachet, 2003)

- Die Gestaltung der Testsituation und die Instruktionen sind weniger geeignet für die Intelligenzdiagnostik im Vorschulalter.
- Bei Standarddurchführung müssen schon die jüngsten Kinder über eine längere Zeit selbstgesteuert arbeiten.
- Teilweise ungünstige Itemgradienten

Sinnvolle Ergänzungen

Da es sich bei der Raven's 2 um ein eindimensionales Verfahren handelt, das nur den Intelligenzfaktor fluide Intelligenz anspricht, sind je nach Ausgangslage und Fragestellung unterschiedliche Ergänzungen erforderlich, um eine breitere Erfassung kognitiver Fähigkeiten zu erzielen. In Frage kommen z. B.:

- Mehrdimensionale Testbatterien zur Erfassung der kognitiven Leistungsfähigkeit, z. B. WISC-V (Wechsler, 2017), KABC-II (P. Melchers & M. Melchers, 2015), WPPSI-IV (Wechsler, 2018), AID 3 (Kubinger & Holocher-Ertl, 2014)
- Spezifische Leistungstests, z. B.
 - Sprachentwicklungstests
 - Testverfahren zur auditiven Verarbeitung
 - Gedächtnistests (Arbeitsgedächtnis, Langzeitgedächtnis)
 - Tests zur Erfassung der Verarbeitungsgeschwindigkeit
 - Testverfahren zur visuellen Verarbeitung

Anforderungen an die Testleiterinnen & Testleiter

Das Manual formuliert sehr eindeutig, dass die Anwendung in erster Linie Psycholog*innen (MA, Diplom) vorbehalten ist, und nur erfolgen darf, wenn die Interpretation von Personen vorgenommen wird, die über eine angemessene Ausbildung in Testpsychologie verfügen.

Zu ergänzen ist, dass Intelligenzdiagnostik bei Kindern im Vorschulalter und bei Testpersonen mit deutlichen kognitiven Einschränkungen sowie motorischen oder sensorischen Beeinträchtigungen generell nur von Personen vorgenommen werden sollte, die über besondere Erfahrung in diesem Bereich verfügen.

Aufgrund ihrer eindimensionalen Struktur und des klaren Testablaufs sind die Anforderungen an die praktische Testdurchführung im Vergleich zu komplexeren Testbatterien deutlich niedriger. Dennoch sind Probetestungen vor einer ersten Anwendung zu empfehlen.

Zusammenfassende Bewertung von Paulina Cordero Donoso

In der praktischen Anwendung innerhalb einer sozialpsychiatrischen Praxis habe ich die Raven's 2 bei Kindern und Jugendlichen zwischen 4 und 16 Jahren mehrmals einsetzen können. Was die praktische Mitarbeit angeht, ist mir bei der Anwendung mit Vorschulkindern ein wenig motivierender Testeinstieg aufgefallen: Testinstruktionen, die nicht dem Entwicklungsstand angepasst sind und eine geringe Interaktion zwischen Kind und Testleiter ermöglichen, sowie Übungsaufgaben, bei denen die erbrachte Leistung der Kinder nicht angemessen gewürdigt werden darf, können zum schnellen Motivationsabfall führen. Auf eine kindgerechte Gestaltung der Untersuchungssituation, die auf ein freundliches, zugewandtes und validierendes Vorgehen achtet, wird unzureichend eingegangen.

Die Raven's 2 sind in ihrer Durchführung und Auswertung unkompliziert. Dennoch ergeben sich einige Unklarheiten wie z. B. die Entscheidung zwischen Einzel- oder

Gruppentestung und zulässige Modifikationen der Testanweisungen bei Sprachverständnisproblemen. Auch wenn die sprachlichen Anforderungen der Raven's 2 als reduziert zu bewerten sind, erfordert ihre Anwendung bei Kindern und Jugendlichen mit mangelnden Kenntnissen der deutschen Sprache eine kompetente Einschätzung und Abdeckung ihres Unterstützungsbedarfs. Hierbei ist m. E. auf die Relevanz einer Einzeltestung zu achten.

Hinsichtlich des theoretischen Hintergrunds ist der Versuch der Autoren, relevante Fachbegrifflichkeiten zum gemessenen Intelligenzkonstrukt der Raven's 2 zu beschreiben und ihre Zusammenhänge zu erklären, m. E. nicht zufriedenstellend gelungen. Es entsteht dadurch die Gefahr, dass sowohl die Interpretation als auch die Ergebnisrückmeldung ohne klaren Theoriebezug stattfindet und demzufolge Testergebnisse falsch verstanden werden. In diesem Zusammenhang bietet beispielweise die Textbeschreibung der automatischen Berichterstellung eine unübersichtliche Darstellung des gemessenen Intelligenzkonstruktes und kann eine klare Vermittlung der Testergebnisse erschweren. Eine gelungene Ergebnisvermittlung ist ein wichtiges Fundament für die Therapiemotivation sowie eine Chance für Kinder und Jugendlichen, Expert*innen im Umgang mit ihren Schwierigkeiten zu werden und somit ihre Kompetenzen zu erweitern.

Die Bezeichnung der Raven's 2 als Verfahren zur Erfassung der allgemeinen kognitiven Fähigkeit könnte den falschen Eindruck erwecken, es handele sich um ein Testverfahren, das ein umfassendes Bild der kognitiven Leistung der Testpersonen bietet. Die Entscheidung für den Einsatz der Raven's 2 sollte berücksichtigen, dass der Test primär fluide Intelligenz erfasst und damit weitere wichtige Bereiche der Intelligenz nicht einbezieht. Raven's 2 ist somit nicht das Verfahren der Wahl, um wichtige diagnostische Entscheidungen im Bereich der kognitiven Leistungen zu treffen.

Im Rahmen der sozialpsychiatrischen Praxis wende ich die Raven's 2 als Ergänzung anderer Testverfahren im Bereich der fluiden Intelligenz und bei Patient*innen, die ausschließlich wegen einer emotionalen oder verhaltensauffälligen Symptomatik untersucht werden und keine Hinweise auf eine Intelligenzbeeinträchtigung zeigen.

Zusammenfassende Bewertung von Gerolf Renner

In der eigenen klinisch-sozialpädiatrischen Praxis hatte ich gelegentlich eine der Vorläuferversionen, die CPM, eingesetzt, wenn die Erfassung kognitiver Leistungen für die klinische Fragestellung nicht zentral war, aber dennoch eine grobe Abschätzung des Intelligenzniveaus sinnvoll erschien. Ein zweiter Anlass für die Verwendung der CPM war die Ergänzung einer Intelligenzdiagnostik, wenn das eingesetzte Basisverfahren keine Beurteilung der fluiden Intelligenz erlaubte. CPM und auch SPM erschienen mir jedoch nicht ausreichend, um die typischen Fragestellungen der Sozialpädiatrie abzuklären, da bedeutsame Intelligenzfaktoren wie Arbeitsgedächtnis, kristalline Intelligenz, auditive Verarbeitung, Verarbeitungsgeschwindigkeit, visuelle Verarbeitung und Langzeitgedächtnis nicht gezielt untersucht werden konnten. Zudem ließen die Manuale im Hinblick auf die Gütekriterien und die Normierung viele Fragen offen.

An dieser Einschätzung hat sich auch mit dem Erscheinen der Raven's 2 nichts grundsätzlich geändert. Wie auch im Manual dargestellt, können die Raven's 2 eine umfassende Intelligenztestbatterie nicht ersetzen. Ebenso wichtig ist der Hinweis, dass die Raven's 2 keine „Schulplatzierungen“ (Manual, S. 24) begründen können und auf keinen Fall dazu benutzt werden dürfen, um Diagnosen, z. B. Intelligenzminderungen, zu stellen. In der Verlagswerbung werden diese Einschränkungen allerdings nicht benannt.

Das Testformat mit wenig aktiven Handlungsmöglichkeiten und hohen Anforderungen an die Selbststeuerung erscheint mir für den Einsatz in klinisch-psychologischen und sonderpädagogischen Kontexten nur bedingt geeignet. Die Protokollierung der Antworten sollte bei jüngeren Kindern und bei Personen mit kognitiven Beeinträchtigungen von den Testleiter*innen übernommen werden. Zumindest bei den einfacheren Items wäre eine alternative Variante, bei der z. B. Bildkarten abgelegt werden, kindgerechter und hätte zudem den Vorteil, dass die Instruktionen sprachlich noch einfacher gestaltet werden könnten (vgl. das Vorgehen beim Nonverbalen Intelligenztest SON-R 2-8; Tellegen, Laros & Petermann, 2018).

In der Papierform wurde die Anordnung der Items in Sets beibehalten, die Itemschwierigkeit steigt dadurch nicht kontinuierlich an, was die Etablierung eines Abbruchkriteriums verhindert. Leistungsschwache Testpersonen – wieder sind die jüngsten Kinder am stärksten betroffen – werden daher relativ viele Misserfolge erleben. Generell habe ich den Eindruck, dass Testpersonen im unteren Leistungsbereich bei der Testentwicklung und in den Darstellungen des Manuals wenig berücksichtigt werden.

Im Manual werden noch sehr wenige Daten zur Validität berichtet. Solange dieser Mangel nicht behoben ist, kann ich mir einen Einsatz der Raven's 2 bei wichtigen diagnostischen Entscheidungen kaum vorstellen. Hier besteht noch beträchtlicher Bedarf an weiterer Forschung, z. B. zur konvergenten Validität mit gebräuchlichen Intelligenzdiagnostischen Verfahren.

Im Manual habe ich an einigen Stellen die nötige kritische Distanz zum eigenen Produkt vermisst. So wird etwa die Kulturunabhängigkeit der Raven's 2 betont, ohne dass dies durch aktuelle Studien untermauert wird. Im Manual der CPM (J. C. Raven et al., 2002) fanden sich durchaus Hinweise, die gegen die Annahme einer völlig kulturunabhängigen Testung sprachen. Zwischen den verschiedenen europäischen Stichproben gab es zwar nur geringfügige Unterschiede, aber dies belegt keineswegs, dass die Raven's 2 die Intelligenz von Testpersonen aus anderen Kulturkreisen (z. B. Kinder mit Fluchterfahrungen) fair erfassen.

Reliabilitäten von IQ-Werten schon ab .90 als „exzellent“ und ab .80 als „gut“ zu beschreiben, erscheint mir nicht angemessen (diese Bewertung wäre m. E. nur für Subtests von Testbatterien passend; vgl. Bracken, 1987). Bei einer Reliabilität von .85, die nicht in allen Altersgruppen erreicht wurde, umfasst das 90%-Konfidenzintervall bei einem IQ-Wert von 85 fast 20 IQ-Punkte (75.4 – 94.6) – eine durchaus beträchtliche Spanne.

Die mit der Verwendung der Bezeichnung *Allgemeinintelligenz* verbundene Problematik (s. o.) wird zwar an einer Stelle des Manuals treffend benannt, ich hätte mir aber eine konsequente Vermeidung dieses Begriffs gewünscht. Testergebnisse der Raven's 2 sollten in Beratung und Befunddokumentation nicht als Allgemeinintelligenz oder allgemeine kognitive Fähigkeiten beschrieben werden, da dies den Eindruck erwecken könnte, es hätte eine umfassende Abklärung der Intelligenz stattgefunden.

Durchaus denkbar erscheint mir der Einsatz der Raven's 2, wenn im Rahmen einer Intelligenzdiagnostik eine ergänzende Erfassung fluider Intelligenzleistungen angestrebt wird oder bereits vorliegende Befunde abgesichert werden sollen.

Der Verzicht auf expressivsprachliche Anforderungen kommt Personen entgegen, die in einer Testsituation nicht verbal kommunizieren können oder wollen. Der breite

Altersbereich erleichtert langfristige Verlaufsmessungen. Die Option von Gruppentestungen wird in der klinisch-psychologischen und sonderpädagogischen Anwendung weniger ins Gewicht fallen, zumal diese im Hinblick auf die praktische Umsetzung (Zeitmessung, s. o.) kaum praktikabel sind, wenn mit Rückfragen oder anderen Unterbrechungen durch die Testpersonen zu rechnen ist. Eine Kombination von ausführlicher Einzeltestung und Gruppentestung könnte u. U. auch einen Eindruck davon verschaffen, ob und wie sich das Arbeitsverhalten von Testpersonen verändert, wenn die Anforderungen an die Selbststeuerung steigen und ein erhöhtes Ablenkungspotenzial gegeben ist. In anderen Anwendungskontexten werden die Vorteile einer digitalen Testung und die Option von Gruppenuntersuchungen möglicherweise bei der Entscheidung über den Einsatz der Raven's 2 stärker gewichtet werden.

Aktuelle Entwicklungen

Angesichts der gerade erfolgten Revision sind baldige Neuerungen vermutlich nicht zu erwarten.

Testrezensionen

Cordero Donoso, P. & Mickley, M. (2020). Neuere Testverfahren. Raven's 2. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 69(4), 378–388.
<https://doi.org/10.13109/prkk.2020.69.4.378>

Zitierte Literatur

- Brachet, I. (2003). Ist kulturunabhängiges Testen der Intelligenz möglich? *Sonderpädagogik*, 33(3), 150–161.
- Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, 4, 313–326.
- Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence*, 52, 71–79.
<https://doi.org/10.1016/j.intell.2015.07.006>
- Kubinger, K. D. & Holocher-Ertl, S. (2014). *AID 3. Adaptives Intelligenz Diagnostikum 3*. Göttingen: Beltz Test.
- Melchers, P. & Melchers, M. (2015). *KABC-II. Kaufman Assessment Battery for Children – II von Alan S. Kaufman & Nadeen L. Kaufman. Deutschsprachige Fassung*. Frankfurt a. M.: Pearson.
- Mickley, M. & Renner, G. (2019). Auswahl, Anwendung und Interpretation Auswahl, Anwendung und Interpretation deutschsprachiger Intelligenztests für Kinder und Jugendliche auf Grundlage der CHC-Theorie: Update, Erweiterung und kritische Bewertung. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 68(4), 323–343.
<https://doi.org/10.13109/prkk.2019.68.4.323>
- Raven, J. (2008). The Raven Progressive Matrices tests: Their theoretical basis and measurement model. In J. Raven & J. Raven (Hrsg.), *Uses and abuses of intelligence: Studies advancing Spearman and Raven's quest for non-arbitrary metrics* (S. 17–67). New York: Royal Fireworks Press.
- Raven, J. C., Raven, J. & Court, J. H. (1998). *APM. Raven's Progressive Matrices and Vocabulary Scales. Advanced Progressive Matrices. Deutsche Bearbeitung und Normierung Stephan Bulheller und Hartmut Häcker*. Frankfurt a. M.: Swets.
- Raven, J. C., Raven, J. & Court, J. H. (2002). *CPM - Coloured Progressive Matrices. Deutsche Bearbeitung und Normierung Stephan Bulheller und Hartmut Häcker*. Frankfurt a. M.: Swets.

- Schneider, W. J. & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment. Theories, tests, and issues* (4th ed., pp. 73–163). New York: The Guilford Press.
- Spearman, C. (1927). *The abilities of man*. New York: Macmillan.
- Tellegen, P. J., Laros, J. A. & Petermann, F. (2018). *SON-R 2-8. Non-verbaler Intelligenztest*. Göttingen: Hogrefe.
- Wechsler, D. (2017). *WISC-V. Wechsler Intelligence Scale for Children - Fifth Edition. Deutsche Bearbeitung Franz Petermann*. Frankfurt: Pearson Assessment.
- Wechsler, D. (2018). *Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition (WPPSI-IV) (dt. Bearbeitung durch F. Petermann u. M. Daseking)*. Frankfurt a. M.: Pearson Assessment.

Internet

- NCS Pearson (Hrsg.) (2019). *Raven's 2. Manual. Deutsche Fassung der Raven's Progressive Matrices 2 – Clinical Edition*. Frankfurt Pearson. Zugriff am 19.08.2020 <https://www.pearsonclinical.de/pub/media/wysiwyg/Pearson%5FDE/Pearson-Ravens-2-Manual-Leseprobe.pdf> (frei zugängliche Leseprobe mit der Einleitung des Manuals).
- Englischsprachige Einführung des Testverlags:
<https://www.youtube.com/watch?v=Z1dJKBG4QAO>