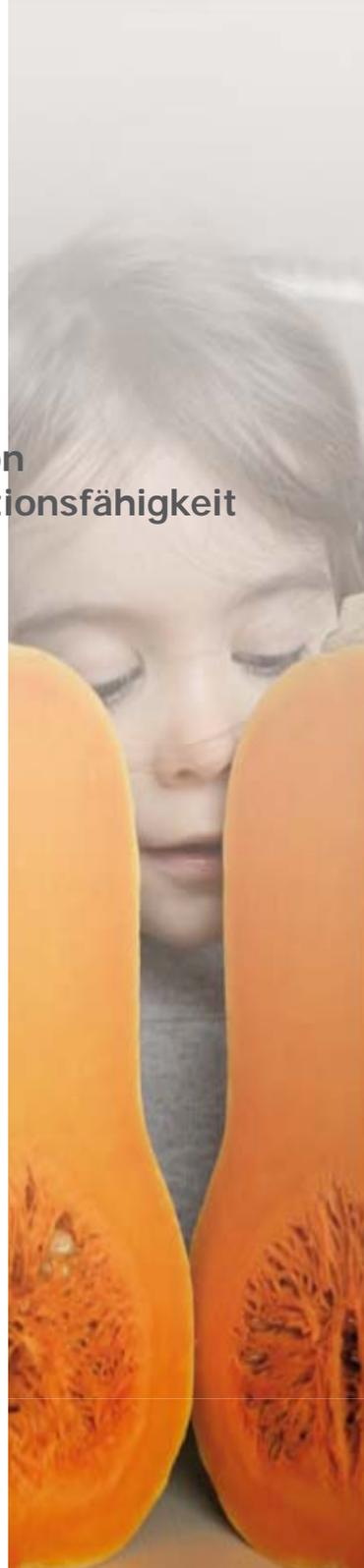


Psychologische Diagnostik von Aufmerksamkeits- und Konzentrationsfähigkeit im Kindergarten- und Schulalter

Ulrike Kipman | Astrid Fritz



Psychologische Diagnostik von Aufmerksamkeits- und Konzentrationsfähigkeit im Kindergarten- und Schulalter

MMag. DDr. Ulrike Kipman, B.Sc.
Mag. Dr. Astrid Fritz

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

Herausgeber
Österreichisches Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF)

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

Tel.: +43/662-439581
Fax: +43/662-439581-310
E-Mail: info@oezbf.at

© 2014

Lektorat/Korrektorat: MMag. Dr. Claudia Resch, Mag. Johanna Weber
Design/Fotografie: Mag. Christina Klaffinger
Grafik/Layout: Mag. Anna Klaffinger, BA
Druck: Print & Smile

Eigenverlag: Österreichisches Zentrum für Begabtenförderung und Begabungsforschung

ISBN: 978-3-9503401-5-0

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

Inhaltsverzeichnis

| | |
|---|----|
| Tabellenverzeichnis | 6 |
| Abbildungsverzeichnis | 7 |
| Vorwort des Herausgebers | 8 |
| Einleitung | 10 |
| 1. Aufmerksamkeit und Konzentration – Begriffsklärung und Modelle | 14 |
| 1.1. Begriffsklärung | 15 |
| 1.2. Theorien zu Aufmerksamkeit und Konzentration | 16 |
| 1.3. Theorien zur Verarbeitungskapazität | 20 |
| 2. Kriterien zur Auswahl wissenschaftlicher Tests | 24 |
| 2.1. Objektivität | 25 |
| 2.2. Reliabilität | 25 |
| 2.3. Validität | 26 |
| 2.4. Normierung | 27 |
| 3. Tests zur Messung von Aufmerksamkeits- und Konzentrationsfähigkeit im Kindergarten- und Schulalter | 30 |
| 3.1. BASIC-MLT; Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-Jährige | 33 |
| 3.2. COG; Cognitrone | 42 |
| 3.3. CompACT; Computerized Attention and Concentration Tests | 47 |
| 3.4. DL-KG; Differentieller Leistungstest – KG | 54 |
| 3.5. DL-KE; Differentieller Leistungstest – KE | 59 |
| 3.6. d2-R; Aufmerksamkeits- und Konzentrationstest | 63 |
| 3.7. FAKT-II; Frankfurter Adaptiver Konzentrationsleistungs-Test II | 68 |

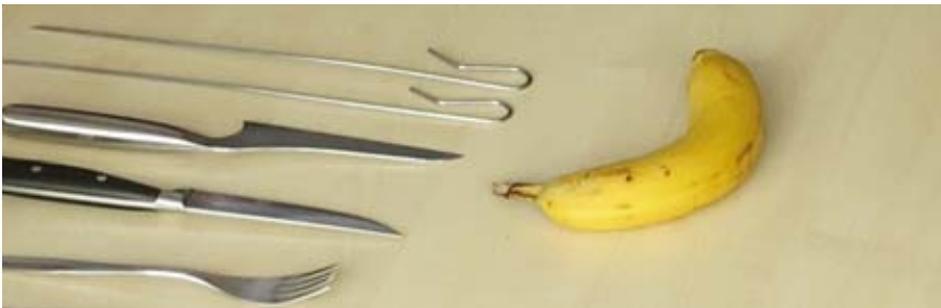
| | |
|---|-----|
| 3.8. FAIR-2; Frankfurter Aufmerksamkeits-Inventar 2 | 78 |
| 3.9. FWIT; Farbe-Wort-Interferenztest | 93 |
| 3.10. INKA; Inventar Komplexer Aufmerksamkeit | 101 |
| 3.11. KKA; Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige | 106 |
| 3.12. KHV-VK; Konzentrations-Handlungsverfahren für Vorschulkinder | 113 |
| 3.13. KLT-R; Konzentrations-Leistungs-Test – Revidierte Fassung | 116 |
| 3.14. KVT; Konzentrations-Verlaufs-Test | 124 |
| 3.15. KT 3–4 R; Konzentrationstest für 3. und 4. Klassen – Revidierte Fassung | 129 |
| 3.16. P-T; Pauli-Test | 136 |
| 3.17. Rev.T.; Revisions-Test | 144 |
| 3.18. SIGNAL; Signal-Detection | 149 |
| 3.19. TAP (inkl. KiTAP); Testbatterie zur Aufmerksamkeitsprüfung | 155 |
| 3.20. TPK; Testreihe zur Prüfung der Konzentrationsfähigkeit | 170 |
| 3.21. VIGIL; Vigilanz | 177 |
| 3.22. WAF; Wahrnehmungs- und Aufmerksamkeitsfunktionen | 185 |
| Gesamt-Literaturverzeichnis | 192 |
| Anhang 1 – | |
| Musterbefunde und -gutachten | 208 |
| Untersuchungsbericht | 210 |
| Klinisch-psychologisches Gutachten 1 | 212 |
| Klinisch-psychologisches Gutachten 2 | 217 |
| Psychologischer Befund | 221 |
| Anhang 2 – Testübersicht | 224 |

Tabellenverzeichnis

| | |
|--|-----|
| Tab. 1: Aufmerksamkeitsformen und Begrifflichkeiten | 16 |
| Tab. 2: Übersicht über die Module des CompACT (adaptiert übernommen aus Prieler, 2011) | 48 |
| Tab. 3: Itembeispiel für den KLT-R 4–6 | 117 |
| Tab. 4: Itembeispiele für den KLT-R 6–13 | 118 |
| Tab. 5: Die Verfahren der TAP und ihre Messbereiche (adaptiert nach Zimmermann & Fimm, 2004, S. 184) | 157 |
| Tab. 6: Untertests der Kinderversion der TAP „KiTAP“ (entnommen aus Zimmermann & Fimm, 2004, S. 199) | 166 |
| Tab. 7: Reliabilitäten für die Variablen R, F, und Mw.RT nach Schuhfried (1995, S. 10) | 182 |
| Tab. 8: Interkorrelationen der Variablen R, F, Mw.RT und St.RT für den Standardparameterblock S1 und die Stichproben Kinder, Studierende, Sportler/innen (Schuhfried, 2004, S. 12) | 183 |
| Tab. 9: Die Durchführungsdauer der einzelnen WAF-Tests | 186 |

Abbildungsverzeichnis

| | |
|--|----|
| Abb. 1: Aufmerksamkeit und Konzentration als unabhängige Konstrukte (adaptiert nach Schmidt-Atzert, Büttner & Bühner, 2004, S. 11) | 15 |
| Abb. 2: Theorien der Aufmerksamkeit (adaptiert nach Heubrock & Petermann, 2001, S. 12) | 17 |
| Abb. 3: Broadbents Filtertheorie (1958) (adaptiert nach Heubrock & Petermann, 2001, S. 12) | 17 |
| Abb. 4: Gängige Normskalen im Vergleich (entnommen aus Holling, Preckel & Vock, 2004, S. 83) | 28 |



Vorwort des Herausgebers (ÖZBF)

Liebe Leserin! Lieber Leser!

Aufmerksamkeit und Konzentration stellen im Kontext von Begabung wichtige Faktoren dar. In den Begabungsmodellen werden sie zum Teil als Voraussetzung und zum Teil als Moderatoren jenes Lernprozesses angesehen, in dem Begabungen zu Leistungen und schließlich zu Leistungsexzellenz transformiert werden.

Dennoch ist festzustellen, dass zunehmend häufiger bei Kindern und Jugendlichen auch das gleichzeitige Vorliegen einer Begabung und einer Aufmerksamkeitsstörung vermutet wird (vgl. Stapf, 2010). Tatsächlich können bei oberflächlicher Betrachtung manche Verhaltensweisen Begabter denen von Kindern ähneln, bei denen eine Aufmerksamkeitsstörung diagnostiziert wurde. Denn begabte Kinder sind in Kindergärten oder Volksschulen manchmal unterfordert, sodass sie mit Langeweile und in der Folge mit Unaufmerksamkeit bzw. Abschalten, Nebenbeschäftigungen oder Störungen reagieren.

In diesem Handbuch werden diagnostische Instrumente vorgestellt, die die verschiedenen Komponenten der Aufmerksamkeit und Konzentration erfassen und eine differentialdiagnostische Abklärung erlauben. Wir hoffen mit dieser Publikation Psychologinnen und Psychologen in der Auswahl geeigneter Testverfahren für entsprechende Fragestellungen unterstützen zu können und damit eine differenzierte Beurteilung der verschiedenen Konstrukte zu ermöglichen.

Die vorliegende Publikation stellt den dritten und letzten Band in einer Buchreihe dar, welche die Identifikation von Begabung behandelt. Eine sorgfältige und verantwortungsvolle Begabungsdiagnostik sollte stets mehr als die Mes-



sung von Intelligenz umfassen (Band 1) und verschiedene nicht-kognitive Persönlichkeitsmerkmale und Umweltvariablen mit einschließen. Zu den nicht-kognitiven Persönlichkeitsmerkmalen gehören neben (Leistungs-)Motivation, Schul- und Prüfungsangst, Selbstkonzept sowie Stressverarbeitung (Band 2) auch Aufmerksamkeit und Konzentrationsfähigkeit (vorliegender Band 3).

Wir danken den vielen Kindern und Jugendlichen, die an den Tests teilgenommen, und deren Eltern, welche den Testungen zugestimmt haben, sowie den Assistentinnen und Assistenten, die bei der Datenauswertung und -analyse mitgeholfen haben. Ein großer Dank geht auch an Lektorat/Korrektorat und Grafik/Layout, die durch ihre wertvolle Arbeit die Qualität dieses Handbuchs maßgeblich verbessert haben.

Auch beim Leibniz-Zentrum für psychologische Information und Dokumentation (ZPID) an der Universität Trier bedanken wir uns herzlich. Das ZPID betreibt die Referenzdatenbank PSYINDEX Tests, welche für unsere Recherchen von großem Wert war.

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

Einleitung



Zu den wichtigsten Voraussetzungen des Lernens gehört die Aufmerksamkeits- und Konzentrationsfähigkeit. Probleme in diesem Bereich stellen im Kindes- und Jugendalter die von Lehrerinnen und Lehrern am häufigsten genannten Beeinträchtigungen dar. Wichtig ist in diesem Kontext zu unterscheiden, ob es sich um eine klinische Störung handelt oder um eine nicht krankheitswertige reduzierte Aufmerksamkeits- und Konzentrationsfähigkeit¹, die im Schulalltag mitverantwortlich für Lernstörungen sein kann. Nur durch eine genaue entsprechende Abklärung können etwaige Interventionen und eine u.U. notwendige medikamentöse Therapie sinnvoll eingesetzt werden.

Die testdiagnostischen Instrumente, die in diesem Handbuch vorgestellt werden, erfassen die verschiedenen Komponenten der Aufmerksamkeit (geteilte Aufmerksamkeit, Alertness, selektive Aufmerksamkeit, Daueraufmerksamkeit, räumliche Aufmerksamkeit, fokussierte Aufmerksamkeit usw.) sowie Aspekte der Konzentration (Fehleranzahl im Verhältnis zur Bearbeitungsdauer usw.).

Vorangestellt werden muss, dass Instrumente zur Verhaltensbeobachtung und zur Exploration (Checklisten, standardisierte Interviews usw.) sowie Gesprächsmethoden, die selbstverständlich auch Teil der Diagnostik sein sollten, in diesem Band aus Platzgründen nicht behandelt werden. Hierzu sei auf Heubrock und Petermann (2001) sowie auf den Beitrag von Imhof im Buch von Büttner und Schmidt-Atzert (2004) verwiesen, die den mehrdimensionalen multimethodalen Ansatz umfassend beschreiben.

Weiters ist darauf hinzuweisen, dass die Überprüfung des intellektuellen Leistungsvermögens eines Kindes im Rahmen der Klärung einer möglichen Aufmerksamkeitsstörung unerlässlich ist, da die Intelligenzdiagnose² Aufschluss darüber geben kann, ob ein Mangel an Aufmerksamkeit u.U. mit einer Über- oder Unterforderung zusammenhängt. Zudem kann die Diagnostik von Persönlichkeitsmerkmalen³ im Rahmen der Aufmerksamkeitsdiagnostik hilfreich sein, um Fehldiagnosen zu vermeiden und in der Folge individuell angepasste Förderungsmaßnahmen zu ergreifen.

Aufgrund der Unüberschaubarkeit und Mannigfaltigkeit der Instrumente ist die Liste der vorgestellten Tests nicht vollständig. Es wurden solche Diagnoseinstrumente ausgewählt, die

- gängig sind,
- eine deutschsprachige Verfahrensversion haben,

¹ Zur Unterscheidung zwischen den Begriffen Konzentration und Aufmerksamkeit sei auf Schmidt-Atzert, Büttner & Bühner (2004) verwiesen.

² Siehe Band 1 „Psychologische Testverfahren zur Messung intellektueller Begabung“ (Kipman, Kohlböck & Weigluny, 2012)

³ Siehe Band 2 „Psychologische Diagnostik moderierender Persönlichkeitsmerkmale bei Kindern und Jugendlichen“ (Kipman, 2013)

- theoretisch verankert sind,
- empirisch validiert sind,
- gesicherte Reliabilitäten und
- Vergleichskennwerte besitzen.

In einem ersten kurzen Kapitel werden grundlegende Begriffsklärungen zu Aufmerksamkeit und Konzentration sowie zu Normierung und den Hauptgütekriterien wissenschaftlicher Tests (Objektivität, Reliabilität, Validität) vorgenommen.

Das zweite Kapitel behandelt die wichtigsten Theorien in Zusammenhang mit Aufmerksamkeit und Konzentration. Im dritten Kapitel werden die ausgewählten Verfahren zur Aufmerksamkeits- und Konzentrationsdiagnostik vorgestellt, die im Rahmen von schulpyschologischen Untersuchungen vorgegeben werden können.

Im Anhang findet sich zum einen Material aus der psychologischen Praxis, das als Muster im Rahmen der Aufmerksamkeits- und Konzentrationsdiagnostik dienen kann und zum anderen eine Kurzübersicht über die beschriebenen Testverfahren.

Literatur

- Büttner, G. & Schmidt-Atzert, L. (2004, Hrsg.). Diagnostik von Konzentration und Aufmerksamkeit. Göttingen: Hogrefe.*
- Heubrock, D. & Petermann, F. (2001). Aufmerksamkeitsdiagnostik. Göttingen: Hogrefe.*
- Kipman, U. (2013). Psychologische Diagnostik moderierender Persönlichkeitsmerkmale bei Kindern und Jugendlichen. Salzburg: ÖZBF.*
- Kipman, U., Kohlböck, G. & Weilguny, W. (2012). Psychologische Testverfahren zur Messung intellektueller Begabung. Salzburg: ÖZBF.*
- Schmidt-Atzert, L., Büttner, G. & Bühner, M. (2004). Theoretische Aspekte von Aufmerksamkeits-/Konzentrationsdiagnostik. In G. Büttner & L. Schmidt-Atzert (Hrsg.), Diagnostik von Konzentration und Aufmerksamkeit (S. 3–22). Göttingen: Hogrefe.*



1. Aufmerksamkeit und Konzentration – Begriffsklärung und Modelle



1.1. Begriffsklärung

Aufmerksamkeit und Konzentration werden alltagssprachlich als zwei unterschiedliche Konstrukte aufgefasst (vgl. Schwalbach, 2001). Während Aufmerksamkeit mehr mit Wahrnehmung assoziiert wird, verbindet man Konzentration eher mit der Art des Arbeitens.

Auch in der Wissenschaft werden die beiden Begriffe unterschieden, allerdings gibt es bislang keine konsensfähige Definition. Strittig ist, ob Konzentration in den Erklärungsrahmen der Aufmerksamkeitskomponenten eingegliedert werden kann, oder ob Konzentration konzeptuell von Aufmerksamkeit abgegrenzt werden muss (vgl. Goldhammer & Moosbrugger, 2006).

In der ersten Auffassung wird Konzentration als besondere Form der Aufmerksamkeit angesehen. Brickenkamp und Karl (1986, S. 195) etwa sehen Konzentration als „... eine spezifische Form von Aufmerksamkeit, eine Steigerungsform, die sich im Wesentlichen durch Willkürlichkeit und stärkere Fokussierung [...] auszeichnet“.

In der zweiten Auffassung werden Aufmerksamkeit und Konzentration als zwei unabhängige Konstrukte angesehen. Schmidt-Atzert, Büttner und Bühner (2004) beispielsweise sehen Aufmerksamkeit als in Verbindung mit Wahrnehmung stehend an (vgl. Abb. 1). Unter Aufmerksamkeit verstehen sie das selektive Beachten relevanter Informationen oder Reize. Dagegen bezieht sich in ihrer Definition der Begriff Konzentration auf alle Stufen der Verarbeitung von Informationen, von der selektiven Wahrnehmung (also Aufmerksamkeit) über die Kombination, Speicherung bis zur Handlungsplanung.

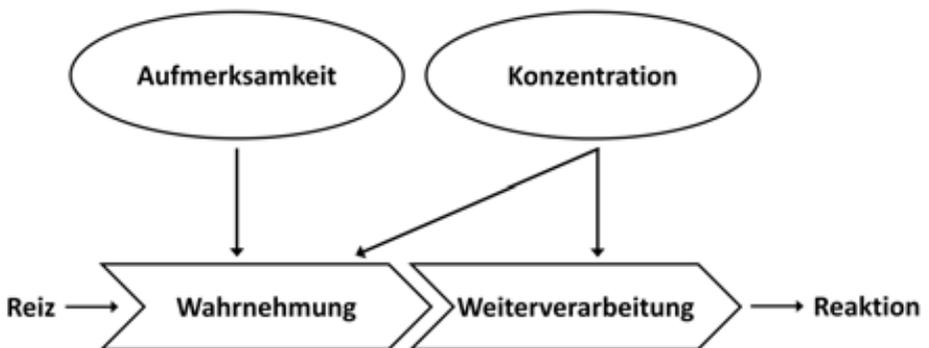


Abb. 1: Aufmerksamkeit und Konzentration als unabhängige Konstrukte (adaptiert nach Schmidt-Atzert, Büttner & Bühner, 2004, S. 11)

Tab. 1 liefert einen Überblick über gängige Aufmerksamkeitsformen, deren zentrale Aspekte und Bedeutungen.

Tab. 1: Aufmerksamkeitsformen und Begrifflichkeiten

| Aufmerksamkeitsbegriffe | zentraler Aspekt | Beschreibung |
|--|-------------------------|--|
| <i>Alertness oder Aufmerksamkeitsaktivierung</i> | Intensität | Zustand der gezielten und kontinuierlichen Aufmerksamkeit; kann phasisch durch Warnreize intensiviert werden |
| <i>Fokussierte, selektive oder gerichtete Aufmerksamkeit</i> | Selektion | Wahrnehmung wird selektiv auf einen bestimmten Bereich fokussiert bzw. auf ein Ziel gerichtet |
| <i>Geteilte Aufmerksamkeit</i> | Anzahl beachteter Reize | Beachtung von mehreren Zielreizen |
| <i>Daueraufmerksamkeit</i> | Zeitaspekt | Aufmerksamkeit über längere Zeit |
| <i>Vigilanz</i> | Zeitaspekt | Beobachtung seltener Reize über eine längere Zeit, mit gewisser Monotonie verbunden |

1.2. Theorien zu Aufmerksamkeit und Konzentration

Seit den 50er Jahren wurden unterschiedliche Theorien zur Aufmerksamkeit und Konzentration entwickelt. Die Theorien lassen sich grob in zwei Gruppen einteilen (vgl. Abb. 2; Heubrock & Petermann, 2001, S. 12):

1. Modelle, die von einer begrenzten Aufmerksamkeitskapazität ausgehen (Filtermodelle, begrenzter Kapazitätsvorrat)
2. Modelle, die die Selektions- und Integrationsfunktion der Aufmerksamkeit hervorheben

Die wichtigsten Theorien der beiden Gruppen werden im Folgenden kurz dargestellt.

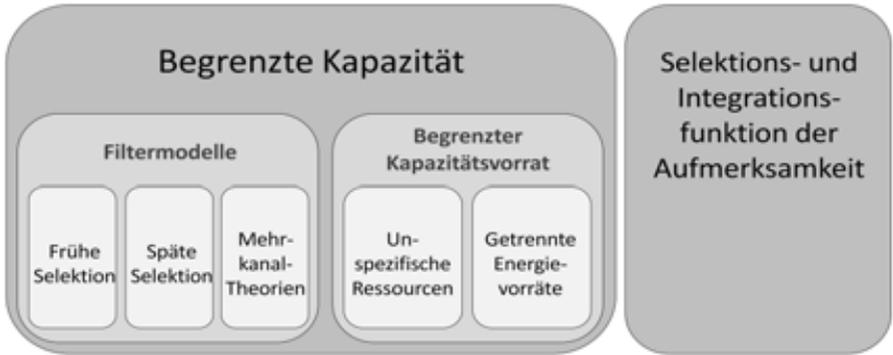


Abb. 2: Theorien der Aufmerksamkeit (adaptiert nach Heubrock & Petermann, 2001, S. 12)

Filtermodelle hatten einen großen und lang anhaltenden Einfluss auf die Erforschung von Aufmerksamkeitsprozessen: Sie gehen davon aus, dass es einen begrenzenden Faktor innerhalb der Informationsverarbeitung gibt, der eine Überlastung des zentralen Prozesses verhindern soll.

Die Filtertheorie von Broadbent (1958) geht von einem sensorischen Speicher aus, der alle Wahrnehmungen zunächst einmal aufnimmt (vgl. Abb. 3). Daraufhin werden die Informationen auf Basis ihrer physikalischen Charakteristika gefiltert und nur die verbliebenen Informationen semantisch verarbeitet. Das heißt, dass bei Gesprochenem auditive Reize wie etwa Stimmhöhe

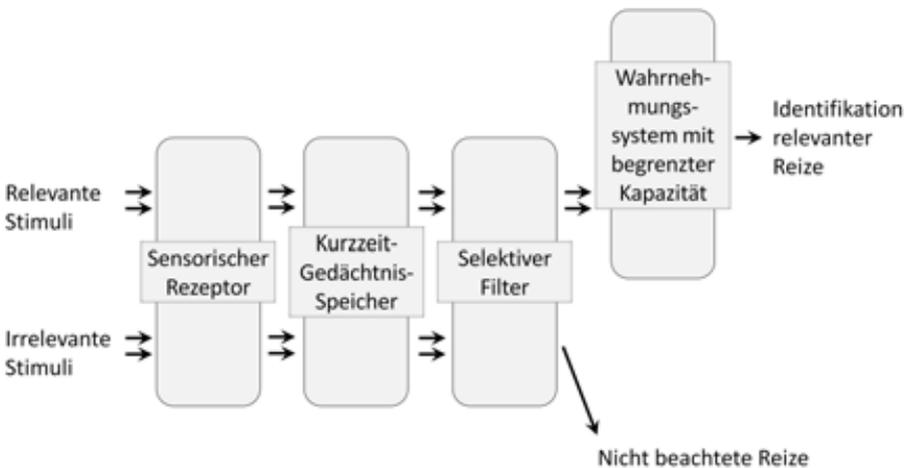


Abb. 3: Broadbents Filtertheorie (1958) (adaptiert nach Heubrock & Petermann, 2001, S. 12)

oder Tonlage und nicht etwa der tatsächliche Inhalt entscheidend dafür sind, ob der Stimulus verarbeitet wird. Nur die Inputs, die den selektiven Filter passieren, werden inhaltlich entschlüsselt und bewusst wahrgenommen. Die abgeblockten Reize verbleiben aber, für eventuelle spätere Zugriffe, kurze Zeit im Filter. Broadbent geht dabei von einer seriellen Verarbeitung aus, weshalb der Filter nötig ist, um das System vor Überlastung zu schützen. Nur die verarbeiteten Informationen können bewusst werden und in das Langzeitgedächtnis übernommen werden.

In späteren Untersuchungen wurde festgestellt, dass Probandinnen und Probanden bei dichotischen Hörtests⁴ auch irrelevante Informationen wahrnehmen und weiterverarbeiten können (Gray & Wedderburn, 1960; Moray, 1959; Treisman, 1960). Trotz konzentriertem Zuhören bei einem Konversationsstrang konnten die Probandinnen und Probanden relevante Information aus dem sogenannten „Störärauschen“ aus der Umgebung herausfiltern. So hörten sie zum Beispiel den eigenen Namen aus einer Unterhaltung heraus, obwohl sie ihr nicht bewusst zuhörten (Cocktail-Party-Effekt), was durch Broadbents Theorie nicht ausreichend erklärt werden konnte.

Um Schwächen von Broadbents Theorie zu beseitigen, entwickelte Anne Treisman im Jahr 1964 die Attenuations- oder Dämpfungstheorie. Diese besagt, dass der Filtermechanismus nicht nach dem „Alles-oder-Nichts-Prinzip“, sondern eher nach dem Prinzip eines Dämpfers funktioniert. Die Reizstärke auf dem unbeachteten Kanal wird reduziert statt gefiltert und so können Informationen aus diesem Kanal auch in abgeschwächter Form verarbeitet werden.

Bei Treismans Theorie ist die Position des „Flaschenhalses“ deutlich flexibler als bei Broadbent, wenn auch immer noch relativ früh im Prozess der Verarbeitung angesetzt. Die Selektion der zu beachtenden Reize erfolgt hierarchisch sowohl auf Grund der physikalischen Eigenschaften, der Struktur als auch auf Grund des Inhalts. Bei der Auswahl auditiver Reize würde dies bedeuten, dass zunächst auf der Basis von Reizmustern, dann Silben, die zu Wörtern werden, anschließend einer grammatischen Struktur und zuletzt der Satzbedeutung selektiert wird.

Im Gegensatz zu Treisman und Broadbent entwickelten Deutsch und Deutsch (1963) eine Theorie der späten Selektion. Der Theorie zufolge werden alle Nachrichten zunächst gleichermaßen (semantisch) verarbeitet. Die Auswahl der Nachrichten erfolgt erst auf einer späteren Verarbeitungsstufe je nach Relevanz für die Verhaltenssteuerung. Die Aufmerksamkeit fällt auf den inhaltlich relevantesten Reiz, der auch die Antwort auf die Stimuli bestimmt.

In Folge dieser unterschiedlichen Theorien kam es zu einer Kontroverse zwischen Treisman und Deutsch und Deutsch bezüglich des Zeitpunkts der Reizselektion. Moray versuchte schließlich, die Kontroverse zwischen früher und später Selektion zu lösen, indem er ein Filtermodell

⁴ Bei einem dichotischen Hörtest werden der Testperson zur selben Zeit auf jedem Ohr verschiedene Wörter (oder kurze Sätze) dargeboten, die anschließend nacheinander nachgesprochen werden sollen.

mit zwei Informationskanälen entwickelte. Später modifizierte auch Broadbent sein ursprüngliches Modell zu einem Mehrkanal-Modell, das sich empirisch absichern ließ (Johnston & Wilson, 1980).

Aufgrund der Entwicklung des Computers in den 70er Jahren des 20. Jahrhunderts entstanden einige Theorien, die einen begrenzten Energievorrat postulierten (Kahnemann, 1973; Posner & Boies, 1971). Diese Theorien ließen sich schließlich aber empirisch nicht absichern. In den 1980er Jahren wurden schließlich Modelle mit getrennten Energievorräten (kognitiv-energetische Modelle) diskutiert (Sanders, 1983; Wickens, 1983). Der Leitgedanke ist hier, dass sich mentale Vorgänge in der Reaktionszeit auf einen Stimulus abbilden und dass die gemessene Leistung von der Fähigkeit der Testperson sowie auch von ihrem Zustand, z.B. ihrer Motivation abhängig ist. In dem Modell werden einzelnen Verarbeitungsstufen spezifische energetische Niveaus der Informationsverarbeitung zugeordnet. Die erste Stufe ist für die Kodierung, Suche, Entscheidung und Organisation der motorischen Reaktion zuständig (strukturelle Ebene). Der zweiten Stufe werden Anstrengung, Arousal und Aktivierung zugeordnet (energetische Ebene). Eine übergeordnete Instanz agiert als exekutives Managementsystem und bringt auch eigene Ressourcen zur zielgerichteten Problembewältigung (z.B. Impulskontrolle, Verhaltensplanung und zeitliche Koordinierung) auf.

Manche Wissenschaftler/innen stellen jedoch die Vorstellung einer begrenzten Kapazität des Gehirnes generell in Frage. Neisser (1979) beispielsweise sieht die Ursachen für gegenseitige Störungen bei der gleichzeitigen Bearbeitung verschiedener Aufgaben in Koordinationsproblemen und Interferenz. Interferenz tritt auf, wenn man für zwei unvereinbare Vorhaben dieselben Wahrnehmungsschemata nutzt.

Neumann (1996) betont, dass es sich bei dieser Art von Interferenz nicht um einen Mangel, sondern um eine Fähigkeit handelt, ungewollte Handlungen zu unterdrücken und irrelevante Reize an einem Zugang zur Handlungssteuerung zu hindern. Anderenfalls würde ein Organismus unter den widersprechenden Handlungen in Chaos versinken.

Zusammenfassend lässt sich Aufmerksamkeit unabhängig von der Modellvorstellung als Prozess beschreiben, der eine große Menge an Eindrücken aufnimmt und sortiert und relevante Reize schließlich verarbeitet. Konzentration liegt dann vor, wenn bewusst Informationen, die gezielt ausgewählt wurden, verarbeitet werden. Wichtig erscheint in diesem Zusammenhang, dass irrelevante bzw. störende Informationen von der weiteren Verarbeitung ausgeschlossen werden müssen.

1.3. Theorien zur Verarbeitungskapazität

Da Aufmerksamkeits- und Konzentrationsfähigkeit eng mit Verarbeitungskapazität zusammenhängt, werden im Folgenden Theorien zur Verarbeitungskapazität skizziert.

In den meisten Gedächtnistheorien werden folgende Gedächtnisprozesse benannt: Kodierung, Speicherung und Abruf. Das bekannte Dreispeichermodell von Atkinson und Shiffrin (1971) geht von einem sensorischen Speicher, einem Kurzzeitspeicher und einem Langzeitspeicher aus. Umweltreize bewirken eine vorbewusste, extrem kurze Speicherung im sensorischen Speicher. Die meisten Informationen werden sogleich wieder vergessen, wohingegen wenige (7 ± 2) Informationseinheiten durch die auf sie gerichtete Aufmerksamkeit in den Kurzzeitspeicher und damit für etwa 30 Sekunden ins Bewusstsein kommen. Der Kurzzeitspeicher, bestehend aus dem auditiv-verbale und dem visuell-nonverbale Gedächtnis, kann die Information nicht nur ins Vergessen oder in eine Reaktion überführen, sondern steht auch in einer Wechselwirkung mit dem Langzeitspeicher, welcher zu einer dauerhaften und unbegrenzten Speicherung in der Lage ist. Nach älteren Ansichten ist für den Übergang in das Langzeitgedächtnis Memorieren nötig, während moderne Theorien von einem flexibleren System mit verschiedenen Einflussgrößen ausgehen. Beim Langzeitgedächtnis wird zwischen dem deklarativen oder expliziten System, zu dem das episodische und das semantische Gedächtnis zählen, und dem prozeduralen oder impliziten System, zu welchem beispielsweise automatisierte Fertigkeiten, Priming und Lernvorgänge gerechnet werden, unterschieden. Daneben gibt es moderne Gedächtnismodelle, wie das integrative Mehrspeichermodell von Hasselhorn (1996), das davon ausgeht, dass die nicht zerfallenen Informationen aus einem sensorischen Register entweder in ein semantisches Netzwerk übergehen oder im Arbeitsgedächtnis bewusst weiterverarbeitet werden, wobei die Intensität dieser Verarbeitung die Knotenaktivierung im Langzeitgedächtnis beeinflusst. Den Wissensabruf aus dem Langzeitgedächtnis steuern Echo- und Conversion-Prozesse der zentralen Exekutive. Das Arbeitsgedächtnis ist insofern vom Kurzzeitgedächtnis zu unterscheiden, als dass die Informationen dabei aktiv manipuliert und bearbeitet werden.

- Neben den generellen Gedächtnistheorien existieren viele klinische Studien und Theorien über die kindliche Gedächtnisentwicklung. Die Gedächtnisleistung von Kindern variiert positiv mit dem Alter.
- Die „Fuzzy-Trace-Theorie“ (Brainerd & Reyna, 1998) besagt, dass Interferenzen zunehmend besser ausgeblendet werden können.
- Das „Good Information Processing-Modell“ (Pressley, Borkowski, & Schneider, 1989) geht davon aus, dass neue und bessere Strategien zunächst noch viel kognitive Kapazität brauchen, wobei dies mit einer zunehmenden Automatisierung weniger wird.

Motoren der Gedächtnisentwicklung, welche einen altersabhängigen Einfluss ausüben, sind dabei die Gedächtniskapazität, Gedächtnisstrategien, das Metagedächtnis und das Vorwissen.

Beim Testen des Erinnerungsvermögens werden meist explizite und nicht implizite Gedächtnisleistungen gemessen. Es wird zwischen Aufgaben mit freier Erinnerungsleistung und bloßem Wiedererkennen unterschieden. Je nach Aufgabentyp können unterschiedliche Rückschlüsse auf Einspeicher- und Abrufstrategien gezogen werden.

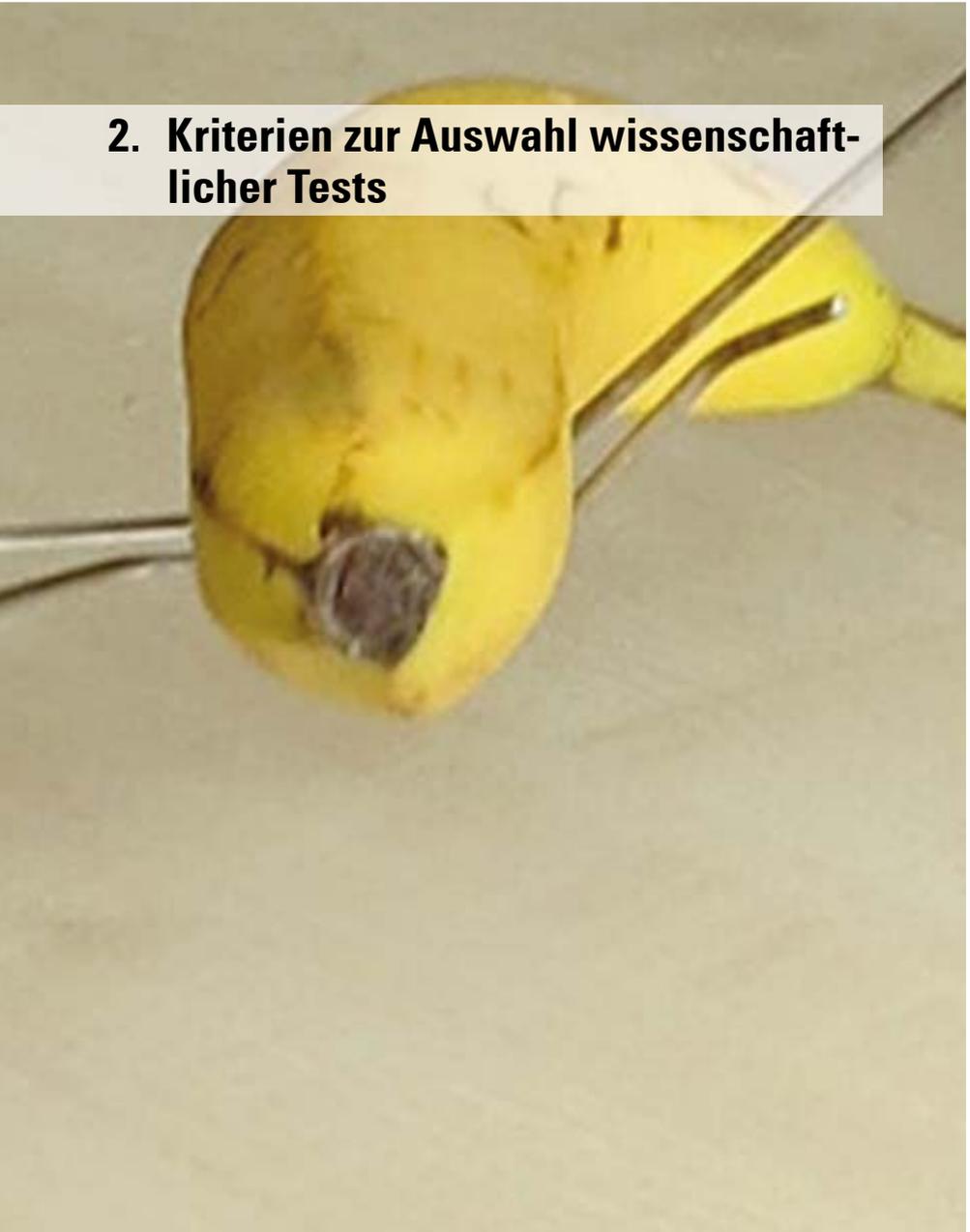
Literatur

- Atkinson, R. C. & Shiffrin, R. M. (1971). *The control of short-term memory*. *Scientific American*, 225, 82–90.
- Brainerd, C. J. & Reyna, V. F. (1998). *Fuzzy-Trace Theory and children's false memories*. *Journal of Experimental Child Psychology*, 71, 81–129.
- Brickenkamp, R. & Karl, G. A. (1986). *Geräte zur Messung von Aufmerksamkeit, Konzentration und Vigilanz*. In R. Brickenkamp (Hrsg.), *Handbuch apparativer Verfahren in der Psychologie* (S. 195–211). Göttingen: Hogrefe.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.
- Deutsch, J. A. & Deutsch, D. (1963). *Attention: some theoretical considerations*. *Psychological Review*, 70, 80–90.
- Goldhammer, F. & Moosbrugger, H. (2006). *Aufmerksamkeit*. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 16–33). Heidelberg: Springer.
- Gray, J. A. & Wedderburn, A. A. (1960). *Grouping strategies with simultaneous stimuli*. *Quarterly Journal of Experimental Psychology*, 12, 180–184.
- Hasselhorn, M. (1996). *Kategoriales Organisieren bei Kindern*. Göttingen: Hogrefe.
- Heubrock, D. & Petermann, F. (2001). *Aufmerksamkeitsdiagnostik*. Göttingen: Hogrefe.
- Johnston, W. A. & Wilson, J. (1980). *Perceptual processing of non-targets in an attention task*. *Memory & Cognition*, 8, 372–377.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Moray, N. (1970). *Attention: Selective Processes in vision and hearing*. New York: Academic Press.
- Neisser, U. (1979). *Kognition und Wirklichkeit. Prinzipien der Implikation der kognitiven Psychologie*. Stuttgart: Klett-Cotta.
- Neumann, O. (1996). *Theorien der Aufmerksamkeit*. In O. Neumann & A. F. Sanders (Hrsg.), *Aufmerksamkeit (Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung, Serie II, Kognition, Band 2, 559–643)*. Göttingen: Hogrefe.
- Posner, M. I. & Boies, S. J. (1971). *Components of attention*. *Psychological Review*, 78, 391–408.
- Pressley, M., Borkowski, J. G. & Schneider, W. (1989). *Good information processing: What it is and what education can do to promote it*. *International Journal of Educational Research*, 13, 857–867.
- Sanders, A. F. (1983). *Towards a model of stress and human performance*. *Acta Psychologica*, 53, 61–97.
- Schmidt-Atzert, L., Büttner, G. & Bühner, M. (2004). *Theoretische Aspekte von Aufmerksamkeits-/Konzentrationsdiagnostik*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 3–22). Göttingen: Hogrefe.

- Schwalbach, C. (2001). *Laienkonzepte über Aufmerksamkeit und Konzentrationsfähigkeit*. Marburg: Philipps-Universität Marburg, Fachbereich Psychologie.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York: Academic Press.



2. Kriterien zur Auswahl wissenschaftlicher Tests



Vorrangige Kriterien für die Auswahl der Testverfahren für die Aufmerksamkeits- und Konzentrationsdiagnostik sollten die so genannten Testgütekriterien Objektivität, Reliabilität und Validität sowie die Normierung der Testwerte sein. Es sollten z.B. nur Tests mit einer möglichst aktuellen Normierung verwendet werden. Als aktuell wird üblicherweise eine Normierung innerhalb der letzten zehn Jahre bezeichnet. Zudem ist darauf zu achten, dass die Normen differenziert für die entsprechende Altersgruppe und andere gegebenenfalls relevante Merkmale vorliegen.

Sofern die Normen nicht optimal sind, ist es für Anwender/innen unabdingbar zu wissen, in welchen Punkten die Normierung ungenügend ist, damit dies bei der Testinterpretation berücksichtigt werden kann. Selbstverständlich sollte das angewendete Testverfahren an Testperson und Zielformulierung angepasst sein, d.h. je nach Fragestellung sollten Testverfahren Anwendung finden, bei denen Boden- bzw. Deckeneffekte ausgeschlossen werden können.

Weiters sollten für eine möglichst umfassende Aufmerksamkeits- und Konzentrationsdiagnostik Tests verwendet werden, die das zu messende Konstrukt in mehreren Formen abbilden (siehe dazu Tab. 1). Sie sollten zudem nicht nur eine einzelne Aufgabenfacette (z.B. Rechnen) umfassen, da sonst Defizite in dieser Domäne fälschlicherweise der Aufmerksamkeits- und Konzentrationsleistung angerechnet werden könnten.

Im Folgenden werden die Hauptgütekriterien wissenschaftlicher Tests, die Objektivität, die Reliabilität und die Validität erklärt und das ebenfalls sehr wichtige Nebengütekriterium der Normierung beschrieben.

2.1. Objektivität

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

2.2. Reliabilität

Reliabilität ist ein Maß für die Zuverlässigkeit eines Testverfahrens, d.h. sie gibt an, wie konsistent das Verfahren das gleiche Konstrukt misst. Alle Items sollten die gleiche Eigenschaft oder Charakteristik messen (*interne Konsistenz*). Ein reliabler Test erbringt auch bei einer Wiederho-

lungsmessung (*Retest-Reliabilität*) bzw. der Messung mit jeweils nur der Hälfte der Aufgaben (*Split-Half-Reliabilität*) vergleichbare Ergebnisse. Der Kennwert für Reliabilität sollte nahe bei 1 liegen, auf jeden Fall über $r = .70$.

2.3. Validität

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

Die *Inhaltsvalidität* (*content validity*) beschreibt, inwieweit die Fragen in einem Test den Bereich, den man untersuchen will, abdecken.

Die *Konstruktvalidität* erfasst, wie gut ein Verfahren den relevanten Bereich abdeckt, aber auch, wie gut es das Konstrukt operationalisieren kann. Als Maß für die *Konstruktvalidität* werden die konvergente und die diskriminante (oder auch divergente) Validität herangezogen:

- Die *konvergente Validität* ist gegeben, wenn die Messdaten von Testverfahren, die dasselbe Konstrukt messen wollen, hoch miteinander korrelieren.
- Die *diskriminante Validität* ist gegeben, wenn die Messdaten von Testverfahren, die verschiedene Konstrukte messen, nur gering miteinander korrelieren (im besten Fall mit $r = .00$).

Die *Kriteriumsvalidität* beschreibt den Zusammenhang eines Tests im Vergleich zu Tests, die einen ähnlichen Bereich erfassen und auf ähnliche Weise messen (also mit einem „Außenmaß“). Die *Kriteriumsvalidität* zeigt demnach an, wie stark ein Test mit einem ähnlichen Test zusammenhängt bzw. ob der Test den Wert eines anderen Tests vorhersagen kann (*prädiktive Validität*). Untersuchungen zur *Kriteriumsvalidität* erfolgen im Rahmen der Aufmerksamkeitsdiagnostik häufig anhand von Korrelationen zwischen verschiedenen Aufmerksamkeits-tests.

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

2.4. Normierung

Um relative Aussagen über Leistungen, Fähigkeiten, Eigenschaften oder Merkmale von Personen machen zu können, werden die Testergebnisse einer Person mit denen anderer Personen verglichen. Will man beispielsweise die Frage beantworten, ob eine 12-jährige Testperson eine Aufmerksamkeitsstörung hat, vergleicht man ihre Testergebnisse mit einer sogenannten Normstichprobe. Damit lässt sich feststellen, wie stark die Aufmerksamkeitsleistung der getesteten Person von der Aufmerksamkeitsleistung anderer Zwölfjähriger (gleichen Geschlechts und vergleichbarer Bildung) abweicht. Wenn sie im Vergleich zur Norm besonders gering ausfällt, kann eine Aufmerksamkeitsstörung vermutet und mit weiteren Verfahren abgesichert werden.

Im optimalen Fall steht für die Normierung eines Tests eine repräsentative Stichprobe zur Verfügung. Repräsentativ bedeutet, dass die Stichprobe bezüglich der Merkmale (z.B. Alter, Geschlecht, Bildung) und der Merkmals- bzw. Fähigkeits- oder Eigenschaftsausprägung mit der Grundgesamtheit vergleichbar ist. Von dieser Normstichprobe liegen die Ergebnisse des Tests im Normalfall in Form von Summenwerten vor, diese Summenwerte („Rohwerte“) werden auf Normalverteilung⁵ geprüft und – sofern keine signifikante Abweichung von der Normalverteilung festgestellt werden kann – in z-Werte umgerechnet (standardisiert). Die z-Werte können wieder mittels linearer Transformation in verschiedene andere Normskalen transformiert werden (siehe S. 28, Abb. 4). Gängig sind bei Aufmerksamkeitstests die z-Metrik (Mittelwert 0, Standardabweichung 1) und die T-Metrik (Mittelwert 50, Standardabweichung 10). Anhand des erreichten Testergebnisses kann man den z- bzw. T-Wert der fraglichen Person ermitteln und somit feststellen, ob und inwieweit die Leistungen durchschnittlich, über- oder unterdurchschnittlich sind.

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

⁵ In der Normalverteilung gilt, dass im Bereich einer Standardabweichung rechts und links vom Mittelwert immer 68,28 % aller Messwerte liegen (siehe Abbildung zur Gauß'schen Glockenkurve). Im Bereich von zwei Standardabweichungen links und rechts vom Mittelwert befinden sich 95,46 % aller Messwerte und im Bereich von drei Standardabweichungen 99,74 %. Untersucht man eine genügend große Stichprobe, so liegt die Mehrzahl der Messwerte im mittleren Bereich und nur wenige Messwerte in den Extrembereichen.

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

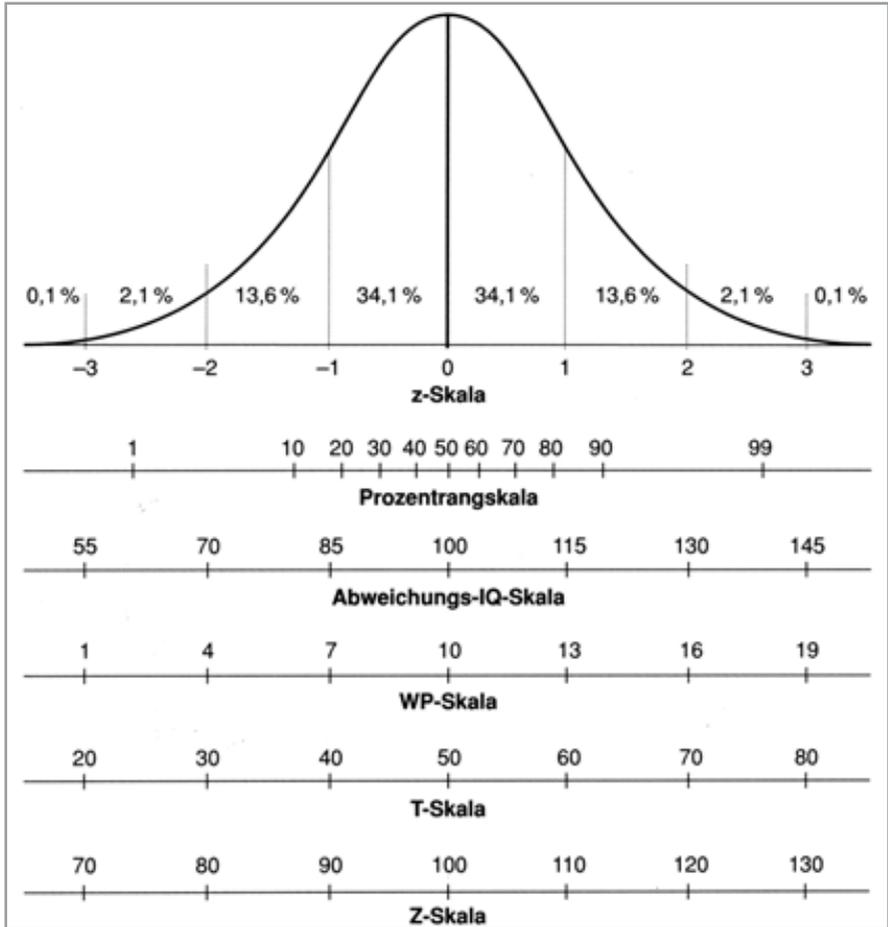


Abb. 4: Gängige Normskalen im Vergleich (entnommen aus Holling, Preckel & Vock, 2004, S. 83)

Literatur

Holling, H., Preckel, F. & Vock, M. (2004). Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6. Göttingen: Hogrefe.

3. Tests zur Messung von Aufmerksamkeits- und Konzentrationsfähigkeit im Kindergarten- und Schulalter



Im Folgenden sollen die im Bereich der Schulpsychologie gängigen und bewährten diagnostischen Instrumente zur Messung von Aufmerksamkeits- und Konzentrationsfähigkeit vorgestellt werden: Für jedes der Verfahren wurde der Einsatzbereich, eine Kurzbeschreibung des Verfahrens und Informationen zu Zuverlässigkeit, Gültigkeit, Normen und Bearbeitungsdauer zusammengestellt. Die ausführlichen Test-Informationen entsprechen – wenn nicht anders angegeben – modifizierten Beschreibungen aus der Psyndex-Datenbank⁶. Den Testbeschreibungen folgt ein Kommentar der Psyndex-Testreviewerinnen und -reviewer und darauf folgend ein kurzer Autorinnenkommentar.

Es muss vorangestellt werden, dass eine umfassende diagnostische Abklärung eines möglichen Aufmerksamkeitsdefizits im Idealfall eine Zusammenfassung von „Life-, Questionnaire- und Test-Daten“ (in Anlehnung an Cattell, 1973) sein sollte. Informationen über Verhaltensweisen in realen Lebenssituationen sollten also mit von der Probandin/dem Probanden selbst eingeschätzten Verhaltensweisen durch Fragebögen und Testdaten in einer Zusammenschau analysiert werden.

Weiters wird darauf hingewiesen, dass es keine Aufgaben gibt, die ausschließlich Aufmerksamkeit oder Konzentration verlangen. Die genannten Tests verwenden einfache Aufgaben, um den Einfluss von speziellen Fähigkeiten und Fertigkeiten auf die Testleistung zu minimieren. In manchen Fällen sind aber Kinder/Jugendliche mit Dyskalkulie oder Legasthenie oder Kinder/Jugendliche, die eine eingeschränkte Merkfähigkeit oder Wahrnehmungsgeschwindigkeit aufweisen, benachteiligt. Hierzu sei auf die Autorinnenkommentare zu den einzelnen Testverfahren verwiesen. Auch die Leistungsmotivation ist hier von großer Bedeutung. Es lässt sich insbesondere bei Jugendlichen nicht immer feststellen, ob die Testperson nicht besser (schneller) arbeiten *konnte* oder *wollte*.

⁶ PSYINDEX Tests: Ziel des Datenbanksegments PSYINDEX Tests ist es, umfassende Informationen über Testverfahren für den Bereich der Testdiagnostik zur Verfügung zu stellen. Dazu werden in den deutschsprachigen Ländern angewandte Tests, Skalen, Fragebögen, Interviewmethoden, Beobachtungsmethoden, apparative Testverfahren, Methoden der computergestützten Diagnostik und andere diagnostische Instrumente aus allen Bereichen der Psychologie und der Pädagogik nach einem standardisierten Raster beschrieben und bewertet. Gegenwärtig enthält PSYINDEX Tests mehr als 6000 Testnachweise. PSYINDEX Tests ist Teil der kostenpflichtigen Referenzdatenbank PSYINDEX, die aus den Segmenten PSYINDEX Literatur und AV-Medien sowie PSYINDEX Tests besteht. PSYINDEX, die keine Volltexte enthält, wird hergestellt und vertrieben vom Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). Das ZPID ist die überregionale Dokumentations- und Informationseinrichtung für das Fach Psychologie in den deutschsprachigen Ländern. Es ist das Forschungsdatenzentrum für die Psychologie. Das ZPID informiert Wissenschaft und Praxis aktuell und umfassend über psychologisch relevante Literatur, Testverfahren, audiovisuelle Medien, Forschungsdaten und Psychologie-Ressourcen im Internet. Kontaktdaten: Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID)

Universität Trier

D-54286 Trier

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

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

Die im Folgenden aufgeführten diagnostischen Instrumente erlauben es, verschiedene Aufmerksamkeitskomponenten oder Teile davon (Aufmerksamkeitsaktivierung, fokussierte Aufmerksamkeit, Daueraufmerksamkeit, selektive Aufmerksamkeit, geteilte Aufmerksamkeit, kognitive Flexibilität usw.⁷) differenziert zu erfassen.

⁷ Begriffsdefinitionen finden sich z.B. bei Schmidt-Atzert, Büttner & Bühner (2004).

3.1. BASIC-MLT; Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-Jährige

Lepach, A. C. & Petermann, F. (2008). Battery for Assessment in Children. BASIC-MLT. Merk- und Lernfähigkeitstest für 6- bis 16-Jährige. Bern: Huber.

Beschreibung

Der BASIC-MLT ist ein Merk- und Lernfähigkeitstest, der bei Kindern und Jugendlichen zwischen 6 und 16 Jahren auf Untertest-, Skalen- und Merquotientenniveau Einblicke in mögliche Defizite des Lern- und Merkvermögens liefert.⁸

Durch den Merk- und Lernfähigkeitstest für 6- bis 16-Jährige (BASIC-MLT; Lepach & Petermann, 2008) sollen Störungen beim Wiedererkennen und Wissenszuwachs bei Kindern und Jugendlichen erfasst werden. Gedächtnisstörungen werden dabei als „Einbußen des Lernens, Behaltens und des Abrufs gelernter Information“ (S. 11) definiert. Mögliche Gedächtnisstörungen werden in die bei Kindern selten auftretende globale Amnesie, in kongrade Amnesie nach einem Trauma und spezifische Gedächtnisstörungen unterschieden. Während die beiden erstgenannten meist durch eine bilaterale Schädigung hervorgerufen werden, entstehen letztere durch eine unilaterale Schädigung. Die spezifischen Störungen werden nach modalitätsspezifischen Gedächtnisdefiziten oder nach Gedächtnisprozessen charakterisiert.

Im BASIC-MLT sollen sowohl Enkodierung, kurz- und langfristiges Behalten als auch der Informationsabruf getestet werden. Auch der Lernverlauf und eine mögliche proaktive oder retroaktive Interferenz, also eine störende Auswirkung später oder früher gelernter Informationen, sollen damit überprüfbar sein. Sowohl verbales als auch visuelles Stimulusmaterial aus verschiedenen Komponenten wird verwendet.

Das Verfahren wird ab 6 Jahren empfohlen, da Kinder in diesem Alter erstmalig gezielt strategische Speicher- und Erinnerungshilfen nutzen (Hasselhorn & Grube, 2006) und Defizite durch schulische Anforderungen deutlicher zutage treten. Anhand von klinischer Beobachtung wird von den folgenden Subtypen von Merk- und Lernstörungen ausgegangen: unaufmerksam, stagnerender und vergesslicher Typ.

Zudem gibt es häufig Komorbiditäten mit Aufmerksamkeitsstörungen und anderen Störungen der Hirnleistung. Es ist auch möglich, dass sich Merk- und Lernstörungen nachteilig auf andere Fähigkeiten auswirken, die z.B. Wissen, Sprache und soziale Anpassung betreffen. Das ICD-10

⁸ Konzentrationsfähigkeit und Verarbeitungskapazität sind hoch korreliert, weshalb dieser Test gerne für die Konzentrationsdiagnostik verwendet wird und daher an dieser Stelle angeführt ist.

sieht bislang keine eigenständige Kodierung für derartige Störungen von Teilfunktionen oder -leistungen vor. Mit dem BASIC-MLT wurde ein Verfahren für den deutschsprachigen Raum entwickelt, welches nicht nur eine differenzierte Diagnose, sondern auch Behandlungsempfehlungen auf diesem Gebiet ermöglichen soll.

Der BASIC-MLT kann inhaltlich in fünf Skalenwerte gegliedert werden:

1. Aufmerksamkeit und Konzentration (AK)
2. Visuelles Lernen (VL)
3. Auditives Lernen (AL)
4. Visuelles Merken (VM)
5. Auditives Merken (AM)

Diese fünf Skalenwerte setzen sich aus acht skalenrelevanten Untertests zusammen:

- Muster Lernen
- Wörter Lernen
- Räumliches Positionieren
- Zahlenfolgen
- Muster Lernen Delay
- Farbfolgen
- Wörter Lernen Delay
- Geräuschfolgen

Zusätzlich existieren folgende sechs optionale Untertests:

- Details Merken
- Muster Lernen Wiedererkennung
- Wörter Lernen Wiedererkennung
- Alltagssituationen
- Geschichten Merken
- Handlungsfolgen

Die optionalen Untertests können von der Testleiterin/vom Testleiter hinzugefügt oder weggelassen werden. Bei Kindern unter 9 Jahren gilt dies auch für die beiden skalenrelevanten Untertests „Wörter Lernen“ (WL) und „Wörter Lernen Delay“ (WLD). Es werden bei den optionalen Untertests großteils alltagsnahe Leistungen abgefragt, die von den Kindern gut angenommen werden und auch zur Testmotivation beitragen. Die optionalen Untertests sind jedoch zugleich kaum standardisierbar und dadurch eher für eine qualitative Auswertung und als Beobachtungsmöglichkeit geeignet. Aus diesem Grund fließen sie nicht in den Gesamtwert oder in die Skalenwerte mit ein.

Die Reihenfolge der Tests entspricht der nachfolgenden Nummerierung. Optionale, nicht skalenrelevante Verfahren sind mit einem * gekennzeichnet. Bei den übrigen Verfahren sind die entsprechenden Skalen in Klammern hinter dem Namen angegeben.

1. Details Merken*: Bei den 12 Items wird jeweils mit der Aufforderung, sich Objekte und Stellen einzuprägen, für 5 (bei den ersten sechs Items) bzw. 10 (bei den weiteren sechs Items) Sekunden eine Bildvorlage gezeigt. Anschließend wird ein verändertes Bild dargeboten und das Kind soll benennen, was und wo sich etwas verändert hat.

2. Muster Lernen (Skala VL): Dem Kind werden sieben geometrische Mustervorlagen hintereinander präsentiert. Die Präsentationszeit beträgt maximal 30 Sekunden pro Muster, kann aber auf Wunsch des Kindes auch früher enden. Nachdem alle Mustervorlagen gezeigt wurden, soll das Kind möglichst genau und möglichst viele der Muster in beliebiger Reihenfolge mit einem Quadrat, einem Rechteck, einem Dreieck und einem Kreis nachlegen. Dies wird insgesamt fünfmal durchgeführt. Es wird frühzeitig abgebrochen und alle eigentlich nachfolgenden Versuche werden als richtig bewertet, wenn das Kind zweimal richtig alle Muster gelegt hat. Anschließend werden die Muster noch einmal ohne Vorgabe reproduziert.

3. Wörter Lernen (Skala AL, erst ab 9 Jahren): Listen mit neun (unter 9 Jahren nur sieben) sinnfreien, aber assoziationsfähigen Pseudowörtern werden von einer CD vorgespielt, woraufhin das Kind sie in beliebiger Reihenfolge, aber exakt, reproduzieren soll. Dies wird insgesamt fünfmal wiederholt. Es wird früher abgebrochen, wenn das Kind zweimal richtig alle Pseudowörter benannt hat. Anschließend wird das Ganze noch einmal mit neun (bei unter Neunjährigen mit sieben) neuen Pseudowörtern aus einer Interferenzliste durchgeführt, welche ebenfalls reproduziert werden sollen. Letztendlich sollen im Anschluss die ersten Wörter noch einmal ohne Vorgabe reproduziert werden.

4. Räumliches Positionieren (Skala VM): Das Kind erhält je 5–10 Sekunden Zeit, um sich die Art und Position geometrischer Formen einzuprägen, die auf einem 3x3- bzw. 4x4-Feldraster abgebildet sind. Danach soll das Kind mit Formen auf einem entsprechenden Raster das Gesehene nachlegen. Es existieren 12 Aufgabenreihen mit jeweils zwei Darstellungen, die reihenweise hintereinander durchgeführt werden. Werden beide Darstellungen einer Aufgabenreihe mit 0 bewertet, wird vorzeitig abgebrochen.

5. Zahlenfolgen (Skala AK, AM): Das Kind hört zunehmend komplexer werdende Zahlenfolgen von einer CD, die es wiederholen soll. Es existieren 16 Items in acht Aufgaben mit jeweils zwei Darstellungen, die



reihenweise hintereinander durchgeführt werden. Werden beide Darstellungen einer Aufgabenreihe mit 0 bewertet, wird vorzeitig abgebrochen.

6a. Muster Lernen Delay (Skala VL): Das Kind soll die Muster aus der Aufgabe „Muster Lernen“ mit den Materialien noch einmal frei reproduzieren.

6b. Muster Lernen Wiedererkennung*: Das Kind soll bei 28 Items (unter 9 Jahren 21 Items) die Muster aus der Aufgabe „Muster Lernen“ wiedererkennen und von neuen Mustern unterscheiden.

7. Farbfolgen (Skala AK, VM): Nachdem das Kind alle Farben einmal richtig benannt hat, werden ihm zunehmend komplexer werdende Farbfolgen gezeigt, die es jeweils nachlegen soll. Es existieren 16 Items in acht Aufgabenreihen mit jeweils zwei Darstellungen, die reihenweise hintereinander durchgeführt werden. Werden beide Darstellungen einer Aufgabenreihe mit 0 bewertet, wird vorzeitig abgebrochen.

8a. Wörter Lernen Delay (Skala AL, erst ab 9 Jahren): Das Kind soll die Pseudowörter aus der Aufgabe „Wörter Lernen“ noch einmal frei reproduzieren.

8b. Wörter Lernen Wiedererkennung*: Das Kind soll bei 28 Items (ab 9 Jahren 36 Items) die Wörter aus der Aufgabe „Wörter Lernen“ wiedererkennen und von neuen Wörter unterscheiden.

9. Geräuschfolgen (Skala AM): Nachdem das Kind alle präsentierten Geräusche einmal richtig benannt hat, werden ihm die zunehmend länger und komplexer werdenden Geräuschsequenzen von einer CD vorgespielt, deren Reihenfolge das Kind anschließend benennen soll. Es existieren 16 Items in acht Aufgabenreihen mit jeweils zwei Geräuschsequenzen, die reihenweise hintereinander durchgeführt werden. Werden beide Geräuschsequenzen einer Aufgabenreihe mit 0 bewertet, wird vorzeitig abgebrochen.

10. Alltagssituationen*: Jede der drei Altersgruppen (bis zur Vollendung des 9., 13. und 17. Lebensjahres) hat zwei spezifische Fotovorlagen. Diese werden der Testperson je 20 Sekunden einzeln gezeigt. Im Anschluss wird die Testperson nach Details und Kontext befragt.

11. Geschichten Merken*: Für jede der drei Altersgruppen (bis zur Vollendung des 9., 13. und 17. Lebensjahres) sind zwei spezifische Geschichten vorbereitet. Diese werden der Testperson einzeln von einer CD vorgespielt und sie soll die Geschichte jeweils im Anschluss mit möglichst vielen Details korrekt nacherzählen.

12. Handlungsfolgen*: Nachdem überprüft wurde, dass die Testperson Formen und Farbkarten richtig zuordnen kann, erhält sie per CD Aufforderungen für Handlungen, die sie in genau der genannten Reihenfolge ausführen soll.

Für Untertests, bei denen ein Lerneffekt zu erwarten ist, gibt es auch Alternativversionen, welche ein erneutes Testen ermöglichen.

Altersgruppe: 6;0–16;11 Jahre

Hinweise zur Durchführung

Der BASIC-MLT ist als **Einzeltest** konzipiert. Durch die optionalen Untertests kann die Testleiterin/der Testleiter das Verfahren individuell auf die Testperson abstimmen, indem sie/er einzelne Untertests weglässt oder hinzufügt. Dadurch ist auch die Testdauer variabel.

Die standardisierte, kurze Instruktion erfolgt mündlich durch die Testleiterin/den Testleiter, die genauere Aufgabeninstruktion erfolgt dagegen in der Regel durch die Audio-CD. Die Testleiterin/der Testleiter ist angehalten, die Instruktion genau zu befolgen. Bei einigen Untertests kann die Testleiterin/der Testleiter wählen, ob sie/er die Items vorliest oder von der CD abspielt. Bei anderen wird aus Gründen der Objektivität dringend zu letzterem Vorgehen geraten. Es ist der Testleiterin/dem Testleiter erlaubt, während der ersten Aufgaben eines Untertests bestimmte Hinweise auszusprechen, falls der Testperson diesbezügliche Fehler unterlaufen sollten oder sie entsprechende Fragen stellt. Durch die Stimulusbücher und andere Materialien erhält das Kind häufig auch einen visuellen (und teilweise sogar taktilen) Bezug zu dem Gesagten.

Die CD sollte in angemessener und ausreichender Lautstärke vorgespielt werden. Zudem muss bei der Durchführung am Tisch eine genaue Sitzordnung eingehalten werden, wobei die Testleiterin/der Testleiter an einer Ecke und das Kind an der ihm schräg gegenübergestellten Breitseite des Tisches sitzt. Die Testleiterin/der Testleiter muss dabei beide Seiten des Stimulusbuches erkennen, dem Kind Materialien reichen und zugleich seine Lösungen überprüfen können. Eine weitere Voraussetzung ist, dass die Testleiterin/der Testleiter und die Testperson der deutschen Sprache mächtig sein müssen.

Es werden etwa 45 Minuten benötigt, um die Kernbatterie durchzuführen. Weitere 20 Minuten dauern die optionalen Untertests, sofern diese alle gewählt werden.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte zwischen Februar 2005 und Juni 2006 an 405 deutschen und Schweizer Kindern und Jugendlichen im Alter zwischen 6;0 und 16;11 Jahren. Die Geschlechterverteilung ist ausgeglichen (50,1 % Jungen). Die Probandinnen und Probanden stammten aus den deutschen Bundesländern Bremen, Niedersachsen und Nordrhein-Westfalen sowie dem Kanton Basel-Stadt in der Schweiz. Der Altersdurchschnitt lag bei 10;0 Jahren. Aufgrund der



Altersdifferenzen in der Entwicklung der mnestischen Funktionen wurden die Testpersonen für die Normierung in fünf Altersgruppen unterteilt: (1) 6;0 bis 7;11 Jahre, (2) 8;0 bis 8;11 Jahre, (3) 9;0 bis 9;11 Jahre, (4) 10;0 bis 12;11 Jahre und (5) 13;0 bis 16;11 Jahre. Von den Schülerinnen und Schülern besuchten 45,2 % die Grund-, 8,6 % die Haupt-, 10,7 % die Real-, 5,4 % die Gesamtschule und 25,3 % das Gymnasium. Es wurde bei der Normierung nicht nach Geschlecht unterschieden, weil sich bei einer Varianzanalyse hierbei keine signifikanten Unterschiede zeigten.

Für die Testauswertung werden pro Untertest Punkte vergeben. Die aufaddierten Punkte jedes

Untertests ergeben den entsprechenden Rohwert. Auf dem Deckblatt des Protokollbogens können die Werte eingetragen werden. Aus den Rohwerten können für jeden Untertest in den Normtabellen T-Werte und Prozentränge ermittelt werden. Aus diesen ist es wiederum möglich, T-Mittelwerte für die fünf Skalen zu berechnen, indem die zwei T-Werte der skalenrelevanten Untertests aufaddiert und durch zwei geteilt werden. Die aufsummierten und durch vier geteilten T-Mittelwerte der Skalen VL, AL, VM und AM (also alle außer AK) ergeben schließlich den Merkquotienten MQ. Bei „Geschichten merken“ ist zudem eine qualitative Auswertung möglich. Die Ergebnisse können sowohl als ganzheitlicher Merkquotient (MQ) als auch auf Skalen- oder auf Untertestebene interpretiert und bewertet werden.

Neben den detaillierten und übersichtlichen Angaben auf dem BASIC-MLT-Protokollbogen verfügt das Verfahren über Auswertungsvorgaben der richtigen Lösungen bei den Stimulusbüchern, Auswertungs- und Interpretationsvorgaben im Manual (Lepach & Petermann, 2008, S. 75), Auswertungsbeispiele (S. 76–79) und Normtabellen (S. 99–118) mit T-Werten und Prozenträngen zur besseren Vergleichbarkeit. Interpretationshinweise sind bei jedem einzelnen Untertest im Manual zu finden (S. 39–61).

Es werden keine Angaben zur Dauer der Auswertung gemacht. Die geschätzte Auswertungszeit beträgt 5–10 Minuten für die skalenrelevanten Untertests und 5 Minuten für die optionalen Untertests.

Testgüte

Objektivität

Durch die äußerst standardisierte Testsituation und die genauen Vorgaben, denen die Testleiterin/der Testleiter unterworfen ist, kann von einer sehr guten Durchführungsobjektivität ausge-

gangen werden. Auch die Auswertungs- und Interpretationsobjektivität kann aufgrund der genauen Bewertungskriterien und der Interpretationshinweise bei jedem Untertest weitgehend als gegeben angesehen werden. Zusätzlich wird in Beispielen näher darauf eingegangen, welche Störungsbilder bei bestimmten Ergebnissen ausgeschlossen werden müssen. Die weitere Interpretation der Ergebnisse in Bezug auf ihre Konsequenzen obliegt allerdings entsprechend qualifizierten Personen.

Reliabilität

Um die Zuverlässigkeit des BASIC-MLT zu bestimmen, wurde Cronbachs α errechnet. Dabei ergaben sich für die Subskalen interne Konsistenzen zwischen $\alpha = .78$ und $\alpha = .86$ und für die Untertests interne Konsistenzen in mittlerer Höhe zwischen $\alpha = .61$ und $\alpha = .88$. Bei den Untertests „Wörter lernen“ (5. Durchgang und nach Interferenz) sowie „Wörter Lernen Delay“, die zur Skala „Auditives Lernen“ gehören, wurde die Itemgüte nur für eine Gruppe von älteren Probandinnen und Probanden errechnet, da bezüglich der auditiven Merkspanne bei Jüngeren noch keine zuverlässigen Aussagen getroffen werden konnten. Bezüglich des gesamten Merkquotienten liegen die relativ hohen Werte je nach Altersgruppe zwischen $\alpha = .78$ und $\alpha = .86$.

Validität

Um nachzuprüfen, wie adäquat der BASIC-MLT die Merk- und Lernfähigkeit erfasst, wurden Untersuchungen über die Gültigkeit des Verfahrens durchgeführt.

Zunächst besitzt der BASIC-MLT inhaltlich Augenscheinvalidität, weil er die Merk- und Lernfähigkeit dadurch misst, dass Personen etwas erlernen bzw. sich etwas merken müssen.

a) Konstruktvalidität

Es wurden Berechnungen der Interkorrelationen von Untertests, Skalenwerten und dem Gesamttest für alle Altersgruppen erstellt, wobei sich bei allen Altersgruppen ein hoher bis sehr hoher Zusammenhang zwischen den Untertests und der dazugehörigen Skala ergab. Auch gab es positive Korrelationen innerhalb der einzelnen Altersgruppen. Der Merkquotient (MQ) korreliert sowohl mit den Skalenwerten aller Altersgruppen als auch mit den einzelnen Untertests überzufällig hoch. Insofern kann davon ausgegangen werden, dass die einzelnen Untertests bezüglich dessen, was sie erfassen, in einem Zusammenhang zu ihren jeweiligen Skalen und zu dem Gesamttest stehen und dass das Alter einen deutlichen Effekt hat.

b) Klinische Validität

Bei $n = 35$ Kindern aus einer klinischen Inanspruchnahmepopulation mit isolierten Gedächtnisstörungen wurde ein Vergleich zu $n = 35$ zufällig ausgewählten Kindern aus der Normstichprobe gezogen (Lepach & Petermann, 2008, S. 66). Hierbei ergaben sich signifikante Unterschiede ($p < .01$) zwischen den Gruppen und es konnte durch die Stichprobenzugehörigkeit 59 % der Varianz aufgeklärt werden. Die Summenwerte der klinischen Stichprobe waren dabei über alle Untertests hinweg niedriger als die der Vergleichsstichprobe. Varianzanalytisch waren die Differenzen der beiden Stichprobengruppen mit Ausnahme der Untertests „Geschichten Merken“ und „Handlungsfolgen“ alle signifikant und wiesen hohe Effektstärken auf. Aufgeteilt in die

zwei Altersgruppen (1) 6;00 bis 9;11 Jahre und (2) 10;00 bis 16;11 Jahre ergaben sich signifikante Unterschiede ($p = .04$) bezüglich der Altersgruppenzugehörigkeit, nicht jedoch aufgrund des Geschlechts. Auf dieser Grundlage wurde bei der Normierung nach Alter, nicht jedoch nach Geschlecht unterschieden.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Beim BASIC-MLT handelt es sich um ein objektives und reliables Testinstrument zur Messung der Lern- und Merkfähigkeit, das auf einer soliden und fundierten theoretischen Basis entwickelt wurde. Es ist sehr übersichtlich und praktikabel aufgebaut. So sind z.B. Abbruchkriterien bei einzelnen Untertests auf dem Protokollbogen vermerkt und damit gut anwendbar. Zudem wurde das Testinstrument für mehrere Interpretationsebenen der Ergebnisse ausgerichtet. So ist es möglich, die Ergebnisse auf Untertest- oder Skalenniveau zu betrachten oder den Merkquotienten als allgemeines Maß zur Beurteilung heranzuziehen. Auf diese Weise können spezifische Defizite in bestimmten Bereichen auffindig gemacht werden.

Ein Mangel besteht allerdings beim ersten Item des ersten Untertests „Details Merken“. Bei diesem Item könnten Kinder bemerken, dass bei einem der fotografierten Krapfen weniger Staubzucker aufgestreut ist als bei einem anderen. Bei Kindern, die somit nicht nur zwischen Zuckerguss und Staubzucker, sondern auch zwischen viel und wenig Staubzucker unterscheiden, ist die vorgegebene Lösung dann jedoch falsch. Es wurde nicht nur der rechte Zuckergusskrapfen mit dem linken Staubzuckerkrapfen vertauscht (wie angegeben), sondern anschließend wurden auch der stärker und der weniger stark mit Staubzucker bestreute Krapfen miteinander vertauscht. Dies ist jedoch in Anbetracht der Fülle von Aufgaben vernachlässigbar.

Auch auf die Konstruktion und Entstehung des Verfahrens hätte für interessierte Leser/innen noch detaillierter eingegangen werden können. Weiterführende Studien, z.B. zu Unterschieden der Ergebnisse zwischen den Schultypen, wurden vermisst. Eventuell wurde dies auch deshalb sehr kurz gefasst, da der Fokus auf der praktischen Anwendbarkeit des Verfahrens liegt. Es wird jedoch darauf hingewiesen, dass zum Zeitpunkt der Veröffentlichung noch weitergehende Untersuchungen bezüglich des BASIC-MLT durchgeführt wurden, so dass auch mit weiteren Veröffentlichungen oder Erweiterungen, in denen vielleicht auch detaillierter auf diese Aspekte eingegangen wird, zu rechnen ist. Hierbei muss berücksichtigt werden, dass es sich um ein umfangreiches Verfahren mit vielen Untertests handelt, das 2008 in Erstaufgabe erschien.

Das einzige wirkliche Manko besteht für die Anwenderin/den Anwender darin, dass der BASIC-MLT als Testbox mit 480 Euro in der Anschaffung relativ kostspielig ist, wobei allerdings Inhalt und Ausstattung der Testbox auch entsprechend umfangreich und hochwertig sind. Somit handelt es sich bei dem BASIC-MLT um ein gelungenes, ökonomisches Verfahren, das seiner Zielsetzung gerecht wird und für die Diagnostik von Merk- und Lernfähigkeit bei deutschsprachigen Kindern und Jugendlichen sehr empfehlenswert ist.

Autorinnenkommentar

Das BASIC MLT-Verfahren ist für die Diagnostik der Merk- und Lernfähigkeit insbesondere von Kindern der Volksschule und der Sekundarstufe I sehr gut geeignet. Er erfasst aber nicht die Konzentrationsfähigkeit, sondern primär die Verarbeitungskapazität, welche aber hoch mit der Konzentrationsfähigkeit korreliert und dafür ein guter Prädiktor ist. Will man Konzentrationsfähigkeit bzw. Aufmerksamkeit prüfen, sollte man zu diesem Test zusätzlich spezifischere aufmerksamkeitsprüfende Verfahren vorgeben.

Literatur

- Atkinson, R. C. & Shiffrin, R. M. (1971). *The control of short-term memory*. *Scientific American*, 225, 82–90.
- Brainerd, C. J. & Reyna, V. F. (1998). *Fuzzy-Trace Theory and children's false memories*. *Journal of Experimental Child Psychology*, 71, 81–129.
- Hasselhorn, M. (1996). *Kategoriales Organisieren bei Kindern*. Göttingen: Hogrefe.
- Hasselhorn, M. & Grube, D. (2006). *Gedächtnisentwicklung (Grundlagen)*. In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung (Enzyklopädie der Psychologie, Serie Entwicklungspsychologie, Band 2, S. 271–325)*. Göttingen: Hogrefe.
- Lepach, A. C., Gienger, C. & Petermann, F. (2008). *Neuropsychologische Befunde zu Merk- und Lernstörungen bei Kindern anhand des BASIC-MLT*. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 36 (6), 389–400.
- Lepach, A. C. & Petermann, F. (2008). *Battery for Assessment in Children. BASIC-MLT. Merk- und Lernfähigkeitstest für 6- bis 16-Jährige*. Bern: Huber.
- Lepach, A. C., Petermann, F. & Schmidt, S. (2007). *Neuropsychologische Diagnostik von Merk- und Lernstörungen mit der MLT-C. Kindheit und Entwicklung*, 16, 16–26.
- Lepach, A. C., Petermann, F. & Schmidt, S. (2008). *Comparisons of the BASIC-Memory and Learning Test with the WISC-IV under developmental aspects*. *Zeitschrift für Psychologie*, 126 (3), 180–186.
- Petermann, F. & Lepach, A. (2006). *Neuropsychologische Diagnostik und Therapie von Aufmerksamkeits- und Gedächtnisstörungen im Kindesalter*. *Verhaltenstherapie*, 16, 112–120.
- Pressley, M., Borkowski, J. G. & Schneider, W. (1989). *Good information processing: What it is and what education can do to promote it*. *International Journal of Educational Research*, 13, 857–867.

3.2. COG; Cognitrone

Schuhfried, G. (2012). Cognitrone (Version 43). Manual. Verfasst von M. Wagner & T. Karner. Mödling: Schuhfried.

Beschreibung

Bei dem Programm Cognitrone handelt es sich um einen computergestützten Aufmerksamkeits- und Konzentrationstest, dessen Hauptanwendungsbereiche in der Personalpsychologie, Klinischen & Gesundheitspsychologie, Neuropsychologie, Verkehrspsychologie, Flugpsychologie und Sportpsychologie liegen.

Aufmerksamkeit wird in der Literatur oft mit Selektion gleichgesetzt. Die Leistung in den einschlägigen Testverfahren basiert darauf, dass durch die Aufmerksamkeit ein relativ kleiner Ausschnitt der Umwelt fokussiert wird. „Dieser Ausschnitt wird in der Erfahrung akzentuiert und dadurch bewusster erlebt als nichtbeachtete Teile der Umwelt.“ (Schuhfried, 2007, S. 4) Im Test können, erst wenn die aufgabenrelevanten Merkmale akzentuiert werden, die geforderten intellektuellen oder motorischen Funktionen optimal eingesetzt werden.

Zur Beschreibung von Aufmerksamkeitschritten wird auf das Akkumulatormodell von Vickers (1970) Bezug genommen. Aus dem gesamten Informationsangebot werden demnach permanent Stichproben („inspections“) entnommen, die jeweils eine bestimmte Zeit („inspection time“) in Anspruch nehmen. Diese „inspection time“ kann als individuelle Konstante interpretiert werden. Das abschließende Urteil über eine Aufgabe wird gefällt, wenn die akkumulierte Information ein Kriterium erreicht, das Vickers (1970) als „degree of caution“, d.h. als „Vorsicht“ bezeichnet (vgl. Schuhfried, 2007, S. 5).

In Aufmerksamkeits-tests werden Aufgaben dargeboten, deren Lösung keine spezifischen Fähigkeiten erfordern. Das Reizmaterial ist bei solchen Testverfahren generell einfach strukturiert, kann jedoch im Programm Cognitrone an Komplexität erheblich zunehmen. Operationalisiert wird das Konstrukt durch die Variablen *Geschwindigkeit*, *Genauigkeit* und *Konstanz der Leistung*. Die Ausprägung dieser Variablen verweist auf das Ausmaß an Aufmerksamkeit, das investiert wurde.

Beim Cognitrone werden auf dem Bildschirm visuelle Inhalte in vier nebeneinander liegenden „Anzeigefeldern“ („Vorlagen“) und einem darunter liegenden „Aufgabenfeld“ dargeboten. Jede der Figuren kann aus bis zu 16 Linien bestehen. Die Aufgabe der Probandin/des Probanden ist es, zu entscheiden, ob eine der Figuren der Anzeigefelder mit der Figur des Aufgabenfeldes übereinstimmt. Nach jeder Antworteingabe wechselt die Figur im Aufgabenfeld, nach einer bestimmten Zahl von Aufgaben wechseln die Figuren in den Anzeigefeldern. Die Figuren können mit fester oder freier Bearbeitungszeit vorgegeben werden.

Mittlerweile stehen sechs Testformen mit freier Bearbeitungszeit (S1–S3, S8–S9, S11) und zwei Testformen mit fester Bearbeitungszeit (S4–S5) zur Verfügung (Schuhfried, 2012). Bei den Testformen mit freier Bearbeitungszeit wird das nächste Item nach der Antworteingabe dargeboten, bei den Testformen mit fester Bearbeitungszeit erfolgt die Signaldarbietung im 1,8-Sekunden-Rhythmus unabhängig davon, ob bereits eine Antwort erfolgt ist oder nicht. In einem weiteren Parameterblock S7 mit geringerer Belastung ist nur die Identität zwischen der Figur im Aufgabenfeld und der Figur in einem einzigen Anzeigenfeld zu prüfen. Die Muster der Anzeige- und Aufgabenfelder aller Parameterblöcke sind bei Schuhfried (2007) vollständig aufgeführt.

Das Urteil über die Identität bzw. Nichtidentität der Figuren ist in allen Parameterblöcken von der Testperson durch das Betätigen einer Taste anzugeben. Als Eingabemedium dient die Probandentastatur oder die Rechnertastatur. Der Testverlauf kann am Testleiter/innenmonitor (falls vorhanden) mitverfolgt werden. Es wird das aktuelle Antwortverhalten der Probandin/des Probanden dargestellt, falsche Urteile werden hervorgehoben.

Je nach verwendeter Testform werden die Ergebnisse zu diversen Variablen zusammengeführt. Interpretationshinweise zu den Variablen in den jeweiligen Testformen sind im Manual enthalten (Schuhfried, 2012). Nachfolgend sollen nur kurz die Hauptvariablen der unterschiedlichen Testformen angeführt werden:

- Hauptvariable der Testformen S1–S3, S11: Mittlere Zeit „Korrekte Zurückweisung“ (sec)
- Hauptvariablen der Testformen S4–S5: Summe „richtige Reaktionen“, Summe „falsche Reaktionen“
- Hauptvariablen der Testformen S8–S9: Summe „Reaktionen“ (richtige und falsche Reaktionen), Prozentsatz „falsche Reaktionen“

Altersgruppe: Je nach Testform ist die Vorgabe ab einem Alter von 4;0 Jahren möglich.

Hinweise zur Durchführung

Das Verfahren ist aufgrund der Darbietung am Monitor **einzel**n durchzuführen. Eine animierte Instrukionsphase und eine fehlersensitive



Übungsphase führen in die Aufgabenstellung ein. Nach dem letzten Übungssitem kann die Probandin/der Proband wiederum selbstständig durch Betätigen einer Panel-Taste die Testphase einleiten.

Bei Testformen mit freier Bearbeitungszeit ist es Aufgabe der Probandin/des Probanden, eine abstrakte Figur mit der Vorlage zu vergleichen und hinsichtlich Identität zu beurteilen. Nach der Antworteingabe wird automatisch zur nächsten Aufgabe weitergeschaltet. Bei Testformen mit fester Bearbeitungszeit ist eine Reaktion nur bei Identität der Figur mit der Vorlage gefordert. Nach Ablauf der Präsentationszeit wird automatisch zur nächsten Aufgabe weitergeschaltet. Das Überspringen einer Aufgabe oder Zurückblättern ist nicht möglich.

Die Bearbeitungsdauer liegt, je nach Testform, zwischen 5 und 20 Minuten (inkl. Instruktions- und Übungsphase).

Hinweise zu Auswertung und Interpretation

Die Normierung ist je nach Testform unterschiedlich. Für die Testformen S1–S5 und S8, S9 und S11 existieren Normstichproben in der Größe von $n = 165$ bis $n = 1.475$, die großteils auch getrennt nach Alter und Bildungsgrad vorliegen. Zusätzlich stehen für diese Formen auch manche der folgenden Spezialnormen zur Verfügung: Normen verkehrsauffälliger Personen, schwedischer Erwachsener, Arbeitssuchender und Stellenbewerber/innen sowie Schüler/innennormen.

Die Normen liegen als T-Werte und Prozentränge vor.

Die Auswertung erfolgt computergesteuert und ist daher nur mit einem minimalen Zeitaufwand verbunden. Der Resultatausdruck enthält neben dem Testprotokoll auch grafische Darstellungen des Testprofils und eine Verlaufsdarstellung. Das Profil erlaubt als grafische Darstellung der normierten Testwerte eine rasche Einordnung der erzielten Leistung der Probandin/des Probanden im Vergleich zur gewählten Referenzstichprobe. Der Messwert der Probandin/des Probanden wird dabei in Referenz zum Durchschnittsbereich dargestellt. Das Ergebnis der Testperson wird unter Einbeziehung der Reliabilität und der Irrtumswahrscheinlichkeit angezeigt. Die Verlaufsdarstellung ist eine grafische Ergebnisdarstellung, die überblicksartig darüber Auskunft gibt, wie sich die Leistungsfähigkeit der Probandin/des Probanden über die gesamte Testdauer hinweg verändert hat.

Testgüte

Objektivität

Durch die programmgesteuerte Vorgabe und den automatischen Ablauf ist die Objektivität hinsichtlich Durchführung und Auswertung gewährleistet. Da es sich um einen normierten

Leistungstest handelt, ist auch die Interpretationsobjektivität gegeben (Lienert & Raatz, 1994). Im Manual sind umfassende Interpretationshinweise angeführt.

Reliabilität

Die Reliabilitäten wurden für diverse Variablen unterschiedlicher Testformen geschätzt. Interne Konsistenzen, Split-half-Koeffizienten und Retest-Reliabilitäten sind durchgehend sehr zufriedenstellend.

Für die Testform S11 wurde außerdem in einer Längsschnittstudie an 82 Personen (48 % Männer, 52 % Frauen) im Altersbereich von 17 bis 78 Jahren über den Zeitraum von 5 Monaten eine Stabilität von $r=0.87$ berechnet.

Validität

Es liegt eine große Anzahl von Studien zu verschiedensten Validitätskonzepten (inhaltliche Gültigkeit, konvergente bzw. diskriminante Validität, Konstruktvalidität und Kriteriumsvalidität) vor, die durchgehend zeigen, dass das Verfahren valide misst. Zur inhaltlich-logischen Gültigkeit führt Schuhfried (2007, S. 11) an, dass im Cognitrone intellektuell anspruchslose Aufgaben gestellt werden, die über eine längere Zeit möglichst schnell und fehlerfrei zu lösen sind. Parallel zu anderen Leistungstests spielt die Versuchssituation eine Rolle. In einer schwach motivierenden Situation wird eher die Einstellung zum Test gemessen, in einer stark motivierenden die Aufmerksamkeitsleistung.

Zahlreiche Studien im Bereich der Verkehrspsychologie bestätigen ebenfalls die Gültigkeit des Verfahrens im Rahmen der Kriteriumsvalidität.

Beispielsweise konnte Karner (2000) in einer Studie signifikante Unterschiede zwischen alkoholauffälligen Kraftfahrern und der Normgruppe im Verfahren Cognitrone feststellen. Die Testergebnisse der alkoholauffälligen Kraftfahrer waren signifikant schlechter als die der Normpopulation. Neuwirth (2001) konnte an einer anfallenden Stichprobe von Probandinnen/Probanden der Fahreignungsdiagnostik einer Südtiroler Untersuchungsstelle zeigen, dass der Cognitrone zwischen allen untersuchten Zuweisungsgruppen (psychiatrische und neurologische Probandinnen/Probanden bzw. Probandinnen/Probanden nach Alkoholabusus) und der Normgruppe trennt.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Das dargestellte Testverfahren zur Messung der Aufmerksamkeitsleistung ist v.a. durch seine Ökonomie und Anwendungsbreite bedeutsam. Auch die Flexibilität in der Programmierung und Anwendung des Verfahrens ist hervorzuheben. Leider werden die vielfältigen Möglichkeiten im Testmanual nur angerissen und es fehlen Validitätsstudien zu den genannten Anwendungsbe-reichen. Die einfache Erstellung eigener Testformen stellt den Geltungsbereich der Normen in Frage, zu deren Repräsentativität allerdings Angaben fehlen.



Autorinnenkommentar

Das Cognitrone-Programm eignet sich gut, um Defizite in der Aufmerksamkeit zu lokalisieren. Probandinnen/Probanden haben mit dem Cognitrone meist keine Probleme und auch kleine Kinder können die Aufgaben des Cognitrone bereits lösen. Die Kennwerte sind verlässlich und zeitlich relativ stabil.

Literatur

- Karner, T. (2000). *Sind verkehrspsychologische Testverfahren geeignete Instrumente, um mögliche Leistungsminderungen alkoholauffälliger Kraftfahrer aufzuzeigen?* *Report Psychologie*, 25 (9), 578–583.
- Keller, I. (2009). *Cognitrone (COG). Untertest aus dem Wiener Testsystem (WTS)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 237–243). Göttingen: Hogrefe.
- Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testpraxis*. Weinheim: Beltz.
- Neuwirth, W. (2001). *Extremgruppenvalidierung verkehrspsychologischer Testverfahren anhand von Zuweisungsgruppen*. *Psychologie in Österreich*, 21, 206–212.
- Rodenhausen, T. (2000). *Das Cognitrone*. In E. Fay (Hrsg.), *Tests unter der Lupe III* (S. 45–60). Lengerich: Pabst.
- Schuhfried, G. (2001). *Cognitrone. COG (Version 29.00) (Computerprogramm)*. Mödling: Schuhfried.
- Schuhfried, G. (2004). *COG. Cognitrone*. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Auflage, S. 251–255). Lengerich: Pabst.
- Schuhfried, G. (2007). *Cognitrone (Version 37.01)*. Mödling: Schuhfried.
- Schuhfried, G. (2012). *Cognitrone (Version 43). Manual*. Verfasst von M. Wagner & T. Karner. Mödling: Schuhfried. Abrufbar unter www.schuhfried.at/fileadmin/content/1_Manuale_de/COG.pdf [17.07.2013].
- Schuhfried, G. (2013). *Wiener Testsystem. Computergestützte Verfahren zur Leistungs- und Persönlichkeitsdiagnostik (Katalog)*. Mödling: Schuhfried.
- Vickers, D. (1970). *Evidence for an accumulator model of psychophysical discrimination*. *Ergonomics*, 13, 37–58.

3.3. CompACT; Computerized Attention and Concentration Tests⁹

Prieler, J. (2011). Computerized Attention and Concentration Tests. Göttingen: Hogrefe.

Beschreibung

Die computerisierte Leistungstestbatterie Computerized Attention and Concentration Tests (CompACT) umfasst vier voneinander unabhängige Module, die ein breites Spektrum an Aufmerksamkeits- und Konzentrationsleistungsdimensionen (u.a. Alertness, selektive Aufmerksamkeit, geteilte Aufmerksamkeit, Vigilanz und Daueraufmerksamkeit) abdecken. Die Hauptanwendungsgebiete liegen in der Arbeits-, Betriebs- und Organisationspsychologie, der Verkehrspsychologie, der Flugpsychologie, der Sportpsychologie sowie in der Klinischen Psychologie und Neuropsychologie.

Die Konzeption der insgesamt 17 Testformen schließt an gängige und empirisch gesicherte Methoden der Aufmerksamkeitsforschung an (vgl. Cohen, 1993; Schmidt-Atzert & Bühner, 2000). Die Aufgaben wurden ursprünglich für die Bearbeitung mit dem Touchpanel des Hogrefe Test-Systems (HTS) konzipiert. Um Anwenderinnen und Anwendern entgegenkommen zu können, die die Verwendung einer Tastatur bevorzugen, wurden im Laufe der Testentwicklung zusätzlich Tastaturversionen von drei Modulen der CompACT-Testbatterie umgesetzt.

Die vier Module des CompACT sind in der nachfolgenden Übersicht (Tab. 2, S. 48) aufgeführt und werden im Folgenden einzeln beschrieben.

Das *Modul CompACT-SR* (Simple Reaction) umfasst insgesamt fünf Testformen:

- Die Testformen 1 („Optischer Stimulus“) und 2 („Akustischer Stimulus“) erfassen die Reaktionsleistung bei Vorgabe einfacher optischer bzw. akustischer Reize.
- Die Testformen 3 („Optischer Stimulus mit Vorankündigung“) und 4 („Akustischer Stimulus mit Vorankündigung“) widmen sich im Speziellen der Fähigkeit zur Leistungssteigerung bei Ankündigung des Reizes (Alertness).
- Testform 5 („Go/No-Go-Bedingung“) dient der Erfassung selektiver Aufmerksamkeitsprozesse bei Vorgabe einfacher Reizkonstellationen (optische Reize).

⁹ Die Testbeschreibung stammt aus Prieler (2011).

Tab. 2: Übersicht über die Module des CompACT (adaptiert übernommen aus Prieler, 2011)

| Modul | Abk. | Aufgaben | Vorgabemodus | Messbereich |
|---------------------------------|-------------|---|----------------------|--|
| <i>CompACT-Simple Reaction</i> | <i>SR</i> | Einfachreaktion und einfache Wahlreaktion (optische und akustische Reize) | Touchpanel, Tastatur | Alertness und selektive Aufmerksamkeit |
| <i>CompACT-Concentration</i> | <i>Co</i> | Mengenvergleich (optische Reize) | Touchpanel, Tastatur | Konzentration |
| <i>CompACT-Vigilance</i> | <i>Vi</i> | Seltene Ereignisse bei erhöhter Testdauer (optische Reize) | Touchpanel, Tastatur | Vigilanz und Dauer Aufmerksamkeit |
| <i>CompACT-Complex Reaction</i> | <i>CR</i> | Mehrfachwahlreaktion (optische und akustische Reize) | Touchpanel | Geteilte Aufmerksamkeit und reaktive Belastbarkeit |

Gemessen werden einfache Reaktionszeiten im Millisekundenbereich. Die dem Verfahren zugrundeliegende experimentelle Situation entspricht dem Prinzip der Einfachreaktion (dem Reaktionsreiz ist eine bestimmte Reizantwort zugeordnet) bzw. einfachen Wahlreaktion (mehrere Reize werden dargeboten, von denen nur einer mit einer bestimmten Reaktion zu beantworten ist). Die spezifische Aufgabenstellung der fünf Testformen wird nachfolgend detailliert beschrieben:

- Testform 1 („Optischer Stimulus“): Die Aufgabe der Probandin/des Probanden besteht darin, bei Einblenden des optischen Signals (rotes Licht) durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit, rasch und zuverlässig auf unmittelbar wahrgenommene optische Reize reagieren zu können.
- Testform 2 („Akustischer Stimulus“): Aufgabe der Probandin/des Probanden ist es, bei Ertönen des akustischen Signals (Telefonklingeln) durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit, rasch und zuverlässig auf unmittelbar wahrgenommene akustische Reize reagieren zu können.
- Testform 3 („Optischer Stimulus mit Vorankündigung“): Aufgabe der Probandin/des Probanden ist es, bei Einblenden des zuvor (akustisch) angekündigten optischen Reizes durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit zur Leistungssteigerung aufgrund extrinsisch kontrollierter Aufmerksamkeitsaktivierung.

- Testform 4 („Akustischer Stimulus mit Vorankündigung“): Die Aufgabe der Probandin/ des Probanden besteht darin, bei Ertönen des zuvor durch ein (optisches) Warnsignal angekündigten akustischen Reizes durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit zur Leistungssteigerung aufgrund extrinsisch kontrollierter Aufmerksamkeitsaktivierung.
- Testform 5 („Go/No-Go-Bedingung“): Aufgabe der Probandin/des Probanden ist es, bei farblicher Übereinstimmung der dargebotenen optischen Signale (Vorgabereiz und Reaktionsreiz) durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit, schnell und zuverlässig auf relevante Reize reagieren und irrelevante Reize ignorieren zu können (selektive Aufmerksamkeit).

Mit dem *Modul CompACT-Co* wird die Konzentrationsleistung bei einfachen Mengenvergleichsaufgaben erfasst. Aufgabe der Probandin/des Probanden ist es dabei, Kreise und Vierecke innerhalb eines von insgesamt acht Kreissegmenten zu vergleichen. Ein Segment steht für ein Item, das durch Betätigung des zugehörigen Reaktionsfeldes (bzw. der zugehörigen Taste) adäquat beantwortet werden soll: Sind mehr Kreise als Vierecke innerhalb des „aktiven“ Kreissegmentes sichtbar, ist die Kreistaste (bzw. die Taste F1) zu betätigen. Ist die Zahl der gezeigten Vierecke größer als die der Kreise, so muss die Taste, auf der ein Viereck abgebildet ist (bzw. die Taste F12), gedrückt werden.

Der Testleiterin/dem Testleiter stehen drei Testformen zur Auswahl:

1. Testform 1 („Konzentrationstest zur Feststellung des persönlichen Arbeitstempos“): In Testform 1 wird das darauffolgende Item erst eingeblendet, wenn das aktuelle (richtig oder falsch) beantwortet wurde (freie Bearbeitungszeit). Die Testdauer (Testphase) liegt zwischen 10 und 15 Minuten.
2. Testform 2 („Konzentrationsleistung unter Zeitdruck“): In Testform 2 beträgt die Reizdarbietungszeit 1500 ms, unabhängig davon, ob eine Reaktion erfolgt oder nicht (feste Darbietungszeit). Die Testdauer (Testphase) beträgt 15 Minuten.
3. Testform 3 („Test zur Feststellung der quantitativen/qualitativen Konzentrationsleistung“): Die Anzahl der Items in Testform 3 wird durch die Bearbeitungsgeschwindigkeit der Probandin/des Probanden bestimmt. Das darauffolgende Item wird – wie in Testform 1 – erst dann eingeblendet, wenn das aktuelle (richtig oder falsch) beantwortet wurde. Die Testphase wird jedoch nach genau sechs Minuten abgebrochen.

Mit dem *Modul CompACT-Vi* wird die Fähigkeit zur Aufrechterhaltung der Aufmerksamkeit bei Vorgabe von (seltenen) Zielreizen innerhalb einer gleichzeitig dargebotenen Menge von irrelevanten Reizen erfasst. Den Anwenderinnen und Anwendern stehen vier Testformen zur Auswahl:

- Testform 1 („Vigilanztest“) erfasst die Reaktionsleistung in monotonen Reizsituationen (die Zahl der Zielreize ist vergleichsweise gering). Die Aufgabe der Probandin/des Probanden besteht darin, bei Aufscheinen des relevanten Reaktionsreizes (Kreis) innerhalb

einer gleichzeitig dargebotenen Menge von Vierecken durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit zur Aufrechterhaltung der Aufmerksamkeit in einer monotonen Reizsituation.

- Testform 2 („Vigilanztest, Screening“) stellt eine Screeningvariante der Testform 1 dar. Die Testdauer für die 45 Items beträgt 15 Minuten. Lang- und Screeningform des Vigilanztests korrelieren mit $r = .97$.
- Testform 3 („Daueraufmerksamkeitstest“) prüft die Fähigkeit zur Aufrechterhaltung der Aufmerksamkeit über einen längeren Zeitraum hinweg (die Zahl der Zielreize ist vergleichsweise hoch). Die Aufgabe der Probandin/des Probanden ist es, bei Aufscheinen des relevanten Reaktionsreizes (Kreis) innerhalb einer gleichzeitig dargebotenen Menge von Vierecken durch Druck auf das Reaktionsfeld zu reagieren. Erfasst wird die Fähigkeit zur Aufrechterhaltung der Aufmerksamkeit über einen längeren Zeitraum hinweg. Zu beachten ist, dass die relevanten Reize (Kreise) in kürzeren Abständen dargeboten werden als im Vigilanztest.
- Testform 4 („Daueraufmerksamkeitstest, Screening“) stellt eine Screeningvariante der Testform 3 dar.

Mit dem *COMPACT-CR-Modul* werden die Fähigkeit zur Aufmerksamkeitssteilung und die Belastbarkeit der Probandin/des Probanden bei Reizüberflutung geprüft. Die fünf Testformen weisen folgende Besonderheiten auf:

- Der Einsatz einer „Ruhetaste“ in Testform 1 („Geteilte Aufmerksamkeit“) ermöglicht die getrennte Messung von Bewegungs- und Reaktionszeit bei gleichzeitiger und zeitversetzter Vorgabe von optischen und akustischen Reizen. Die Belastbarkeit der Probandin/des Probanden bei Reizüberflutung wird in den Testformen 2 bis 5 über die Vorgabe von Reizfolgen (optische und akustische Reize) mit hoher Darbietungsfrequenz erfasst.
- Die Testformen 2 und 3 („Komplexe Aufmerksamkeitsleistung“, „fremdgesteuertes Tempo“ und „Verkehrspsychologische Form“) sind durch Reize mit festen Darbietungszeiten gekennzeichnet (Reaktionsbedingung).
- In Testform 3 werden zudem vier (Belastungs-)Phasen unterschieden (Ausgangsleistung, Belastungsphase, Überbelastungsphase und Regenerierungsphase), die insbesondere der Erfassung des Reaktionsverhaltens der Probandin/des Probanden bei unterschiedlich hoher Belastung sowie der Fähigkeit zur Regeneration dienen.
- Die Itemvorgabe in Testform 4 („Komplexe Aufmerksamkeitsleistung, selbstgesteuertes Tempo“) erfolgt aktionsbedingt (der Folgereiz wird erst eingeblendet, wenn auf den aktuellen Reiz richtig reagiert wurde).
- Die Reizdarbietungsgeschwindigkeit in Testform 5 (Komplexe Aufmerksamkeitsleistung, Belastungsform) wird kontinuierlich an die Leistungsfähigkeit der Probandin/des Probanden angepasst (Adaptive Bedingung).

Altersgruppe: ab 14;0 Jahren

Hinweise zur Durchführung

Das CompACT-System wird als **Einzeltest** durchgeführt.

CompACT-SR: Die Testdauer (Testphase) beträgt drei Minuten in den Testformen 1 bis 4 und neun Minuten in Testform 5.

CompACT-Co: Die Testdauer (Testphase) beträgt 10 bis 15 Minuten in Testform 1, 15 Minuten in Testform 2 und 6 Minuten in Testform 3.

CompACT-Vi: Die Testdauer (Testphase) des Vigilanztests beträgt 30 Minuten. Für die Screeningform des Vigilanztests beträgt die Testdauer (Testphase) 15 Minuten. Die Testdauer (Testphase) des Daueraufmerksamkeitstests beträgt 20 Minuten in Testform 3 und 10 Minuten in Testform 4 (Screening).

CompACT-CR: Die Testdauer (Testphase) beträgt 10 bis 12 Minuten in Testform 1, 8 Minuten in Testform 2, 12 Minuten in Testform 3, 6 Minuten in Testform 4 und 6 Minuten in Testform 5.

Hinweise zu Auswertung und Interpretation

Reaktionszeit (in ms): Der Testkennwert „Reaktionszeit“ ist ein quantitatives Maß für die Aufmerksamkeitsleistung. Bei der Berechnung der Normwerte werden die zunächst negativ gepolten Zeitmaße in positiv gepolte Leistungsmaße umgewandelt. Hohe Normwerte entsprechen damit einer geringen Reaktionszeit bzw. einer hohen Reaktionsgeschwindigkeit.

Reaktionsstabilität: Die Variable „Reaktionsstabilität“ ist ein Maß für die Gleichmäßigkeit oder Homogenität der Konzentrationsleistung. Eine über die Zeit annähernd konstante Konzentrationsleistung geht mit einer entsprechend geringen Streuung der Reaktionszeiten bzw. hohen Reaktionsstabilität einher. Hohe Normwerte entsprechen damit einer geringen Reaktionszeit bzw. einer hohen Reaktionsgeschwindigkeit.

CompACT-SR: Die Normwerte basieren auf den Daten von 219 Probandinnen/Probanden im Alter von 14–79 Jahren. Neben einer Gesamtnorm stehen altersspezifische Normen zur Verfügung.

CompACT-Co: Die Normwerte basieren auf den Daten von 224 Probandinnen/Probanden im Alter von 14–91 Jahren. Neben einer Ge-



samtnorm stehen alters- und geschlechtsspezifische Normen zur Verfügung. Zusätzlich liegen tastaturspezifische Normen ($n = 96$) vor.

CompACT-Vi: Die Normwerte basieren auf den Daten von 215 Testpersonen im Alter von 14–91 Jahren. Neben einer Gesamtnorm stehen altersspezifische Normen zur Verfügung.

CompACT-CR: Die Normwerte zu den Testformen 2 bis 5 basieren auf den Daten von 163 Testpersonen im Alter von 14–78 Jahren. Die Normwerte zu Testform 1 basieren auf den Daten von 110 Probandinnen/Probanden im Alter von 15–79 Jahren. Es liegen alters- und geschlechtsspezifische Normen vor.

Gütekriterien

Objektivität

Da die Testvorgabe, die Registrierung und die Auswertung der Resultate standardisiert und computergesteuert erfolgen, ist von Objektivität in Durchführung und Auswertung auszugehen.

Reliabilität

CompACT-SR: Die internen Konsistenzen liegen zwischen $\alpha = .86$ und $\alpha = .98$ (Cronbachs α) bzw. $r = .80$ und $r = .97$ (Split half).

CompACT-Co: Die internen Konsistenzen liegen zwischen $\alpha = .80$ und $\alpha = .98$ (Cronbachs α) bzw. $r = .80$ und $r = .97$ (Split half).

CompACT-Vi: Die internen Konsistenzen liegen zwischen $\alpha = .94$ und $\alpha = .99$ (Cronbachs α) bzw. $r = .80$ und $r = .99$ (Split half) in den Langformen des Moduls. Für die Screeningform konnten Werte zwischen $\alpha = .95$ und $\alpha = .99$ (Cronbachs α) bzw. $r = .88$ und $r = .98$ (Split half) ermittelt werden.

CompACT-CR: Die internen Konsistenzen liegen zwischen $\alpha = .70$ und $\alpha = .99$ (Cronbachs α) bzw. $r = .76$ und $r = .99$ (Split half).

Validität

Die Validität des Moduls wird durch Studien zur konvergenten, diskriminanten und faktoriellen Validität sowie durch Interkorrelationen zwischen den einzelnen Testformen der Module gestützt.

Autorinnenkommentar

Das CompACT-Verfahren ähnelt dem WAF-Verfahren und auch der TAP in Aufbau und Durchführung. Die Erfahrungen mit Jugendlichen und Erwachsenen mit derartigen Verfahren sind gut, die Kennwerte sehr gut geeignet, um festzustellen, ob (bei ausreichend guter Motivation, ins-

besondere bei den Vigilanz- und Daueraufmerksamkeits-Modulen) ein krankheitswertiges Aufmerksamkeitsdefizit vorliegt. Bei Schülerinnen und Schülern sollte in jedem Fall ein Verfahren zur Überprüfung der Arbeitshaltung (Arbeitseinstellung, Leistungsmotivation, Anstrengungsvermeidung, Lern- und Arbeitsverhalten etc.) durchgeführt werden.

Literatur

Cohen, R. A. (1993). The neuropsychology of attention. New York: Plenum Press.

Prieler, J. (2011). Computerized Attention and Concentration Tests. Manuals. Göttingen: Hogrefe.

Schmidt-Atzert, L. & Bühner, M. (2000). Aufmerksamkeit und Intelligenz. In K. Schweitzer (Hrsg.), Intelligenz und Kognition (S. 125–152). Landau: Verlag Empirische Pädagogik.

3.4. DL-KG; Differentieller Leistungstest – KG

Kleber, E. W., Kleber, G. & Hans, O. (1999). Differentieller Leistungstest-KG (DL-KG). Test zur Erfassung des Leistungsverhaltens bei konzentrierter Tätigkeit im Grundschulalter (2., korrigierte Auflage). Göttingen: Hogrefe.

Beschreibung

Mit dem DL-KG kann bei Volksschülerinnen und -schülern von ca. 6–10 Jahren und Sonderschülerinnen und -schülern ab 9 Jahren die Konzentrationsleistung in quantitativer und qualitativer Hinsicht untersucht werden. Dadurch wird eine differentialdiagnostische Unterscheidung zwischen eindeutig konzentrationsbedingten und neurotischen bzw. organischen Lern- und Leistungsstörungen ermöglicht (Borchert, Knopf-Jerchow & Dahbashi, 1991).

Der Differentielle Leistungstest DL-KG schließt an die Tradition der Durchstreichtests an (vgl. Bartenwerfer, 1982; Bourdon, 1895). Er ist kein Konzentrationstest im herkömmlichen Sinne. Er verzichtet auf den Anspruch, Konzentrationsfähigkeit nur in einem Testwert zu erfassen, er liefert drei Messwerte für verschiedene Leistungsvariablen (Quantität, Qualität und Gleichmäßigkeit der Leistung).

Als Weiterentwicklung bestehender Konzeptionen, bei denen nur die „relevanten“ Stimuli durchgestrichen werden, sollen beim DL-KG auch die „irrelevanten“ Zeichen mit einem Punkt markiert werden. Damit soll der sogenannten „Unbedenklichkeitshaltung“ einer Probandin/eines Probanden vorgebeugt werden, indem die Beachtung jedes Zeichens erzwungen wird. Während der Testung werden akustische Signale gesetzt, bei denen die Probandinnen/Probanden das jeweilig bearbeitete Zeichen einkreisen.

Die Testblätter sind mit 20 Zeilen zu je 20 Zeichen bedruckt. Die Reihenfolge der Stimuli ist zufällig.

Die Zeichen bestehen aus stilisierten Abbildungen vertrauter Gegenstände, z.B. Baum, Blume, Stuhl, Tisch, Kamm, Bürste, Tretroller, Fahrrad usw.

Von den 7- und 8-jährigen Volksschülerinnen und -schülern und Kindern mit Lernproblemen sollen jeweils 3, von den 9- und 10-jährigen Volksschülerinnen und Volksschülern jeweils 4 Stimuli als „relevant“ erkannt und angestrichen werden.

Neben der quantitativen Erfassung der Konzentrationsleistung (GZ = Gesamtzahl der angekreuzten relevanten Zeichen) werden 3 weitere Indikatoren für qualitative Aspekte der Aufmerksamkeitsleistung gebildet:

- F % T: Fehlergesamtzahl, relativiert auf Gesamtzahl der bearbeiteten Zeichen
- SB % / GZ: Schwankungsbreite der quantitativen Leistung
- SB / F %: Schwankungsbreite der qualitativen Leistung

Altersgruppe: Das Verfahren findet im Bereich der Volksschule seine Anwendung (frühestens ab Mitte des 1. Schuljahres). Das Alter der mit dem Verfahren testbaren Probandinnen/Probanden liegt zwischen dem 6. und dem 10. Lebensjahr. Außerdem liegen Hinweise für die Testung von Sonderschülerinnen und -schülern ab 9 Jahren vor.

Hinweise zur Durchführung

Das Verfahren kann als **Gruppen- und Einzeltest** eingesetzt werden.

Parallelformen können durch die Auswahl anderer „relevanter“ Stimuli gebildet werden.

Für die Durchführung des Verfahrens sollte eine Unterrichtsstunde veranschlagt werden.

Die Anweisungen sind wortwörtlich im Manual abgedruckt. Bei einer „erprobenden Instruktion“ werden auf einer Übungsvorlage 3 Zeilen durchgearbeitet.

Das Verfahren kann von Fachkräften, die mit den Prinzipien der Durchführung und Auswertung von Tests vertraut sind, angewendet werden.

Es empfiehlt sich, den Test unter Anwesenheit zweier Testleiter/innen durchzuführen, da durch Beobachtung des Verhaltens der Probandinnen/Probanden während der Testung wichtige zusätzliche Anhaltspunkte zur Erstellung von Diagnosen gewonnen werden können.

Hinweise zu Auswertung und Interpretation

Es liegen geschlechtsspezifische Normwerte für 7- bis 10-jährige Volksschüler/innen in Jahresschritten vor. In die Eichstichprobe waren 906 Schüler/innen aus 46 Volksschulklassen einbezogen.



Die Rohwerte für die einzelnen Leistungsindikatoren aus dem Test werden getrennt nach dem Geschlecht und nach Altersgruppen in 5 Prozentrangklassen umgerechnet (PR bis 5, bis 25, bis 75, bis 95, größer 95). Analoge Normierungen liegen für die 14 Zeitintervalle vor, wobei allerdings nur 3 Prozentrangklassen gebildet werden.

F %-Normen werden auf Zeitintervall-Ebene nur für die Prozentranggruppen bis 10 und 25 % angegeben, da dieser Indikator nur im unterdurchschnittlichen Bereich differenziert. Zusätzlich wurden Streuungen und kritische Differenzen (auf dem 5 %-Niveau) für alle Zeitintervalle und getrennt für Jungen und Mädchen und die Altersstufen von 7–10 Jahren berechnet.

Die Rohpunktermittlung erfolgt mit Hilfe transparenter Lösungsschablonen. Die ermittelten Leistungen können mit alters- und geschlechtsspezifischen Normwerten verglichen werden.

Die Auswertungszeit beträgt ca. 10 Minuten.

Testgüte

Objektivität

Aufgrund der standardisierten Durchführungs- und Auswertungsrichtlinien kann das Verfahren als weitgehend objektiv bezeichnet werden.

Reliabilität

Die Messgenauigkeit der quantitativen Leistungsvariable scheint hinreichend zu sein, wobei allerdings hinsichtlich der zeitlichen Stabilität dieses Merkmales z.T. deutliche Einschränkungen zu konstatieren sind. Für die qualitativen Leistungsindikatoren werden z.T. gar keine Reliabilitätsangaben berichtet (SB % / GZ), während die sehr pauschalierend mitgeteilten Angaben hinsichtlich des Fehlerprozentwertes deutlich machen, dass dieser Wert aufgrund von Reliabilitätsmängeln für individualdiagnostische Zwecke nicht interpretiert werden sollte.

Validität

Für den DL-KG wird inhaltlich-logische Gültigkeit beansprucht. Darüber hinaus wurden die Beziehungen zu verschiedenen Testverfahren und Lehrer/innen-Urteilen untersucht.

Zu Verfahren mit ähnlichem Validitätsanspruch (d_2) bestehen maximal mittelhohe Übereinstimmungen bis $r = .50$. Auch das Lehrer/innen-Urteil über denselben Verhaltensaspekt stimmt mit den in dem Verfahren erfassten Merkmalen nur in mittelhohem bis geringem Ausmaß überein. Der Vergleich zu Intelligenzverfahren erbringt nur in wenigen Fällen signifikante Beziehungen. Schulnoten sind von der Testleistung gar nicht bzw. nur in geringem Ausmaß abhängig.

Als weitere Validitätsbelege können Unterschiede aus Gruppenuntersuchungen herangezogen werden. Diese zeigten u.a., dass

- zwischen altersgleichen Volks- und Sonderschülerinnen und -schülern signifikante Unterschiede in der erwarteten Richtung bestehen,
- ein deutlicher Altersfortschritt bei der Gesamtleistung, jedoch nicht bei den qualitativen Leistungsindikatoren vorliegt,
- so gut wie fast keine geschlechtsspezifischen Unterschiede vorhanden sind.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Die Eichung des Verfahrens muss als vorläufig eingeschätzt werden. Es fehlen v.a. spezifische Angaben für Kinder mit Lernstörungen. Außerdem müsste die Gültigkeit der Normen für einen größeren geographischen Bereich abgesichert werden.

Aufgrund der Reliabilitätsüberprüfungen ist eine Interpretation der Indikatoren für die qualitative Seite der Testleistung ausgesprochen problematisch.

Der eigentliche differentialdiagnostische Anspruch, der mit dem Verfahren verbunden ist, wurde noch nicht systematisch evaluiert. Außerdem fehlen Nachweise über die Effizienz einer Verzahnung von Diagnose und Therapie bei den als behandlungsbedürftig eingeschätzten Schülerinnen und Schülern.

Autorinnenkommentar

Das Instrument ist zur Feststellung spezifischer Leistungsvariablen gut geeignet und ist in leicht abgewandelter Form auch bei jüngeren Probandinnen/Probanden anwendbar (anstatt 14 nur 10 Zeitintervalle). Es ist für Kinder leicht verständlich und die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht es sehr gut, aufgrund der Intervalltestung und des relativ langen Testzeitraums die fokussierte, selektive und die Daueraufmerksamkeit bei Kindern zu überprüfen und in Kombination mit Verfahren zur Überprüfung zur Vigilanz und zur räumlichen Aufmerksamkeit AD(H)S-Kinder mit hoher Wahrscheinlichkeit zu identifizieren. Allerdings ist aufgrund der teilweise fehlenden Normwerte und der Reliabilitätsmängel Vorsicht bei individualdiagnostischer Verwendung geboten.

Literatur

- Bartenwerfer, H. (1982). Allgemeine Leistungsdiagnostik. In K.-J. Groffmann & L. Michel (Hrsg.), Intelligenz- und Leistungsdiagnostik (Enzyklopädie der Psychologie, Themenbereich B, Serie II, Band 3. S. 482–512). Göttingen: Hogrefe.*
- Borchert, J., Knopf-Jerchow, H. & Dahbashi, A. (1991). Testdiagnostische Verfahren in Vor-, Sonder- und Regelschulen. Ein kritisches Handbuch für Praktiker (Testkurzdarstellung DL-KG – Differen-*



- tieller Leistungstest KG: S. 322–323). Heidelberg: Asanger.
- Bourdon, B. (1895). *Observations comparatives sur la reconnaissance, la discrimination et l'association*. *Revue philosophique*, 40, 153–185.
- Kleber, E. W., Kleber, G. & Hans, O. (1999). *Differentieller Leistungstest-KG (DL-KG)*. Test zur Erfassung des Leistungsverhaltens bei konzentrierter Tätigkeit im Grundschulalter (2., korrigierte Auflage). Göttingen: Hogrefe.

3.5. DL-KE; Differentieller Leistungstest – KE

Kleber, E. W. & Kleber, G. (1974). Differentieller Leistungstest – KE (DL-KE). Göttingen: Hogrefe.

Beschreibung

Mit dem DL-KE kann bei Volksschülerinnen und Volksschülern in der Eingangsstufe die Konzentrationsleistung in quantitativer und qualitativer Hinsicht erfasst werden. Damit kann eine differentialdiagnostische Unterscheidung zwischen eindeutig konzentrationsbedingten und neurotischen bzw. organischen Lern- und Leistungsstörungen ermöglicht werden.

Als Weiterentwicklung bestehender Konzeptionen, bei denen nur die „relevanten“ Stimuli durchgestrichen werden, sind beim DL-KE auch die „irrelevanten“ Zeichen von den Probandinnen/Probanden mit einem Punkt zu markieren. Damit soll der sogenannten „Unbedenklichkeitshaltung“ von Probandinnen/Probanden vorgebeugt werden, indem die Beachtung jedes Zeichens erzwungen wird. Während der Testung werden alle 90 Sekunden akustische Signale gesetzt, bei denen die Kinder um das gerade bearbeitete Zeichen einen Kreis machen sollen.

Die Testvorlage besteht aus 7 Bearbeitungsseiten mit 14 Zeilen zu je 10 Zeichen und einem Übungsblatt mit 11 Zeilen. Weiterhin liegen 2 Vorlagen mit je 10 Zeichen vor, auf denen geübt werden soll, welche der Zeichen als „relevant“ und als „irrelevant“ erkannt werden.

Die Zeichen bestehen aus stilisierten Abbildungen vertrauter Gegenstände wie z.B. Baum, Kamm, Kind, Auto, Ente etc. Es sollen bei dem Verfahren jeweils 2 Stimuli als relevant erkannt und durchgestrichen werden.

Die Gesamtbearbeitungszeit beträgt 15 Minuten (zehn 90-Sekunden-Intervalle).

Neben der quantitativen Erfassung der Konzentrationsleistung (GZT = Gesamtzahl der angekreuzten relevanten Zeichen) wurden 3 weitere Indikatoren zur Erfassung qualitativer Aspekte der Aufmerksamkeit konzipiert:

- F % T: Fehlergesamtzahl, relativiert auf Gesamtzahl der bearbeiteten Zeichen
- SB % / GZ: Schwankungsbreite der quantitativen Leistung
- SB / F %: Schwankungsbreite der qualitativen Leistung

Altersgruppe: Das Verfahren soll in der zweiten Hälfte der ersten Klasse der Grundschule Anwendung finden.

Hinweise zur Durchführung

Das Verfahren kann als **Gruppen- oder Einzeltestung** appliziert werden. Es wird empfohlen, die Gruppen auf 4–6 Kinder zu begrenzen. Eine Parallelversion (mit zwei anderen „relevanten“ Stimuli) ist vorgesehen.

Zu der Bearbeitungszeit von 15 Minuten muss mit 10 Minuten Instruktion und Einübung gerechnet werden.

Das Verfahren ist ein Papier-und-Bleistift-Test, zu dem die Unterlagen vollständig bezogen werden können. Die Testleiterin/der Testleiter benötigt zusätzlich eine Stoppuhr. Es empfiehlt sich, die Stoppuhr auf 90-Sekunden-Intervalle zu präparieren.

Die Anweisungen sind wortwörtlich im Manual abgedruckt. Eine Einübungsphase ist vorgesehen.

Das Verfahren kann von jeder Fachkraft, die mit den Prinzipien der Durchführung und Auswertung von Tests vertraut ist, angewendet werden.

Die Anwesenheit von 2 Testleiterinnen/-leitern bei der Durchführung ist von Vorteil, da dadurch eine genaue Beobachtung des Verhaltens der Testperson bzw. der Testpersonen während der Testung erleichtert wird.

Hinweise zu Auswertung und Interpretation

Es liegen Normwerte getrennt für Knaben und Mädchen vor. Die Eicherhebung fand im März 1972 statt; einbezogen waren 10 Schulen mit 18 ersten Klassen (160 Jungen, 170 Mädchen) in Schleswig-Holstein.

Für die vier Indikatoren der Leistung, die aus dem Test gewonnen werden, ist ein 5-stufiges Klassifikationsschema vorgesehen, wobei getrennt für Jungen und Mädchen die Rohwerte in Prozentrangwerte umgerechnet werden (weit überdurchschnittlich, überdurchschnittlich, durchschnittlich, unterdurchschnittlich, weit unterdurchschnittlich).

Für die quantitative und qualitative Leistung in den einzelnen Zeitintervallen sind ebenfalls Umwandlungen zu 5 Prozentranggruppen vorgesehen (unter 5, bis 25, bis 75, bis 95, über 95 %). Aus diesen für Jungen und Mädchen getrennt ermittelten Angaben lässt sich eine mittlere Leistungskurve berechnen.

Die Rohpunkteermittlung erfolgt mittels Lösungsschablonen. Die ermittelten Leistungen können mit vorliegenden Normwerten verglichen werden. Auf den für Jungen und Mädchen getrennt vorliegenden Auswertungsblättern sind die mittleren Leistungskurven für GZ und F % dargestellt.

Es ist mit ca. 6 Minuten Auswertungszeit zu rechnen.

Gütekriterien

Objektivität

Aufgrund der standardisierten Durchführungs- und Auswertungsrichtlinien kann das Verfahren als weitgehend objektiv bezeichnet werden. Eine quantitative Überprüfung der Objektivität liegt nicht vor.

Reliabilität

Die mittlere Korrelation der einzelnen Intervalle liegt zwischen $r = .80$ und $r = .95$, die Stabilität (Retest nach einer Woche) der einzelnen Auswertungsparameter zwischen $r_{tt} = .55$ und $r_{tt} = .86$.

Validität

Für das Verfahren wird inhaltlich-logische Gültigkeit beansprucht. Die 4 Subskalen korrelieren untereinander in mittelhohem Maße (GZT – F % T: $r = .42$; SB F % – GZT: $r = .40$; SB F % – F % T: $r = .76$).

Die Validierung anhand des Frankfurter Tests für Fünfjährige – Konzentration (FTF-K) erbrachte eine Korrelation von $r = .60$.

Aufgrund der in der Handanweisung mitgeteilten Korrelationen besteht eine relativ hohe Übereinstimmung zwischen dem Gesamleistungswert und Verfahren ähnlichen Validitätsanspruchs (FTF-K). Das Lehrer/innen-Urteil über Konzentration scheint hingegen zum Großteil aufgrund anderer Merkmale erschlossen zu werden als im DL-KE erfasst werden. Eine mittelhohe Beziehung ist auch noch zu Indikatoren der motorischen Entwicklung vorhanden, während Intelligenz-, Sprach- und Sozialstatuswerte nicht mit dem Gesamtwert korreliert sind.

Der Fehlerprozentsatz als Indikator der Leistungsgüte hängt hingegen in mittlerem Ausmaß mit Intelligenz-, Sprach- und Motorikleistungen zusammen. Außerdem zeigt sich für diesen Messwert auch eine Beziehung zur sozialen Schichtzugehörigkeit.

Weitere indirekte Hinweise über die Validität des Verfahrens ergeben sich aus verschiedenen Gruppenuntersuchungen. Dabei wurde u.a. gefunden, dass

- die Gesamtergebnisse und die Qualität der Leistung bei konkreten Reizvorlagen besser sind als bei abstrakten,
- die Testleistung bei Gruppenuntersuchungen ansteigt, die Leistungsgröße aber beeinträchtigt wird (besonders ab einer Gruppengröße von 7 Testpersonen),
- bei einer Durchführung mit störendem Lärm durch Andere die Leistung nicht beeinträchtigt wird,
- zurückgestellte Schüler/innen im Vergleich zu eingeschulten signifikant schlechtere

- quantitative und qualitative Ergebnisse erzielen,
- eindeutige Alterstrends bei der Gesamtleistung, jedoch nicht bei der qualitativen Leistungsvariable bestehen,
- eine deutliche Überlegenheit der Mädchen gegenüber den Jungen vorhanden ist.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Aufgrund der Eichstichprobe ist nur ein geringer Repräsentativitätsanspruch gesichert. Weitere Erhebungen zur Differenzierung der Normen nach dem Alter der Kinder und dem Zeitpunkt der Untersuchung (etwa vor dem Beginn der Grundschule, am Ende des 1. Schuljahres etc.) wären wünschenswert. Der diagnostische Anspruch des Verfahrens wird empirisch nicht systematisch untermauert. Die Beziehung zu Schulleistungsdaten ist noch ungeklärt. Aufgrund der geringen Reliabilität sollte nur der Gesamtleistungswert im Rahmen der Einzelfalldiagnostik interpretiert werden.

Autorinnenkommentar

Das Instrument ist – ebenso wie der DL-KG – zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Es ist für Kinder leicht verständlich und die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht es, die fokussierte, selektive und die Daueraufmerksamkeit bei Kindern zu überprüfen und ist insbesondere in Kombination mit anderen Verfahren wie der KiTAP geeignet, auch AD(H)S-Kinder zu identifizieren. Allerdings ist aufgrund der schon etwas älteren Normierung (aus dem Jahr 1972) Vorsicht geboten.

Literatur

- Barth, A. (2002). *Der Differentielle Leistungstest-KE (DL-KE)*. In E. Brähler, H. Holling, D. Leutner & F. Petermann (Hrsg.), *Brickenkamp Handbuch psychologischer und pädagogischer Tests* (3., vollständig überarbeitete und erweiterte Auflage, Band 1, S. 270–272). Göttingen: Hogrefe.
- Borchert, J., Knopf-Jerchow, H. & Dahbashi, A. (1991). *Testdiagnostische Verfahren in Vor-, Sonder- und Regelschulen. Ein kritisches Handbuch für Praktiker (Testkurzdarstellung DL-KE – Differentieller Leistungstest KE: S. 321–322)*. Heidelberg: Asanger.
- Ettrich, C. & Ettrich K.U. (2004). *Entwicklung der Konzentrationsfähigkeit im Vorschulalter und diagnostische Probleme ihrer Erfassung*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 205–231). Göttingen: Hogrefe.
- Kleber, E. W. & Kleber, G. (1974). *Differentieller Leistungstest – KE (DL-KE)*. Göttingen: Hogrefe.

3.6. d2-R; Aufmerksamkeits- und Konzentrations-test

Brickenkamp, R., Schmidt-Atzert, L. & Liepmann, D. (2010). Test d2 – Revision Aufmerksamkeits- und Konzentrationstest. Göttingen: Hogrefe.

Beschreibung

Der d2-R misst Konzentration bzw. konzentrierte Aufmerksamkeit mittels eines Durchstreichtests mit figuralem Material. Der Test erfordert die Selektion relevanter Reize unter Zeitdruck; er erfasst die Konzentrationsfähigkeit, die Schnelligkeit und Genauigkeit bei der Unterscheidung ähnlicher visueller Reize (Detail-Diskrimination).

Der Test d2-R stellt eine Weiterentwicklung des bewährten Tests d2 dar, der in seiner ersten Version bereits 1962 erschien und zuletzt in der 9. Auflage herausgegeben wurde (Brickenkamp, 2002). Die Testaufgabe und die zu bearbeitenden Zeichen blieben bei der Revision unverändert.

Beim d2-R handelt es sich um einen Durchstreichtest, der in Anlehnung an ähnlich konzipierte Verfahren wie z.B. den Durchstreichtest von Bourdon (1895) konzipiert wurde (Durchstreichen aller a, i, r, s in einem Text). Die pro Blatt benötigte Zeit sowie die Fehler (Auslassungen und Fehldurchstreichungen) wurden dabei errechnet. Sharp (1899) verwendete noch Texte, Whipple (1910) Buchstaben und Toulouse und Pieron (1911) acht geometrische Figuren. Schon 1925 verzeichnete Henning 29 Formen von Durchstreichtests von 28 Autorinnen und Autoren (Bartenwerfer, 1982).

Die Testvorlage besteht aus einer Bearbeitungsseite zu 14 Zeilen mit je 57 Zeichen. Damit wurde der d2-R im Vergleich zum d2 bei gleich bleibender Bearbeitungszeit von 20 Sekunden pro Zeile um 10 zusätzliche Zeichen am Zeilenende erweitert. Diese Testverlängerung stellt eine Maßnahme gegen Deckeneffekte dar. Die Zusammensetzung der Zeichen wurde ebenso wie das Verhältnis von Zielobjekten zu Distraktoren beibehalten.

Die Zeichen bestehen aus den Buchstaben d und p mit meist ein oder zwei kleinen Strichen unten oder oben. Wie Heinrich (1973) bemerkt,



wiederholen sich die Zeilen dem Aufbau nach, d.h. die Zeilen 1, 4, 7, 10 und 13 sind identisch, ebenso 2, 5, 8, 11 und 14 sowie 3, 6, 9 und 12.

Die „d“s mit 2 Strichen sollen als relevant erkannt und durchgestrichen werden. Nach 20 Sekunden beginnt die Testperson auf das Kommando der Testleiterin/des Testleiters „Halt, nächste Zeile“ eine neue Zeile zu bearbeiten.

Zur Erfassung quantitativer und qualitativer Aspekte der Konzentrationsleistung wurden im d2-R folgende Indikatoren konzipiert:

- BZO-Wert = Anzahl der bearbeiteten Zielobjekte (*Arbeitstempo*)
- KL-Wert = markierte Zielobjekte minus Verwechslungsfehler (*Konzentrationsleistung*)
- F % = Prozentsatz an Fehlern bezogen auf BZO (*Genauigkeit*)

Des Weiteren kann der Verlauf der Leistung dargestellt werden. Dazu wird der Test in 4 Blöcke zu je 3 Zeilen eingeteilt. Die erste und letzte Zeile werden dabei nicht ausgewertet. Durch die so entstandenen vier parallelen „Subtests“ sollen Rückschlüsse auf einen eventuellen Einstellungswechsel, den Grad der Ermüdung und möglicherweise auftretende Lerneffekte gewonnen werden.

Altersgruppe: 9;0–60;0 Jahre

Hinweise zur Durchführung

Das Verfahren kann als **Gruppen- und Einzeltest** appliziert werden. Zusätzlich zur reinen Testbearbeitungszeit von 4 Minuten und 40 Sekunden benötigt man für die Instruktion inklusive der Übungsaufgaben weitere 5 Minuten. Die Anweisungen sind wortwörtlich im Manual abgedruckt. Ausformulierte Instruktion sowie Kurzanleitung liegen zusätzlich in türkischer Sprache vor.

Das Verfahren kann von jeder Fachkraft, die mit den Prinzipien der Durchführung und Auswertung von Tests vertraut ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Die Normen für den d2-R wurden im Zeitraum von Ende 2007 bis Mitte 2008 erhoben. Für die Normierung wurde der Test mehr als 4.000 Personen ($n = 4.024$; 50,7 % männlich, 49,4 % weiblich) im Altersbereich von 9 bis 60 Jahren vorgelegt. Die getesteten Probandinnen und Probanden stammen aus 6 deutschen Bundesländern, die meisten aus Berlin (52,8 %). In der Normie-

rungsstichprobe wurden sämtliche Schulformen abgedeckt. Bei Schülerinnen und Schülern der weiterführenden Schulen wurden vorwiegend Gymnasiastinnen/Gymnasiasten, bei Schulabgängern vorwiegend Personen mit Realschulabschluss erhoben.

Im Manual sind Prozentrang- und Standardwertangaben für die Maße BZO, KL und F % zu finden. In der Gesamtstichprobe wurden für keines der drei Maße signifikante Geschlechtsunterschiede gefunden. Die Testleistungen sind erwartungsgemäß altersabhängig. Sie steigen von 9–19 Jahren zunächst steil, dann zunehmend schwächer mit dem Alter an, bis sie für die Altersgruppe von 20–39 ein Plateau erreichen. Ab der nächsten Altersgruppe fallen die Testleistungen deutlich ab.

Die Auswertung erfolgt manuell und dauert bedingt durch eine Vereinfachung der Kennwerte gegenüber Vorgängerversionen nur mehr 5 Minuten.

Testgüte

Objektivität

Aufgrund der hoch standardisierten Durchführungs- und Auswertungsrichtlinien kann das Verfahren als objektiv bezeichnet werden. Die Interpretationsobjektivität ist durch nach dem Alter gestufte und aktuelle Normen gegeben.

Reliabilität

Die interne Konsistenz liegt je nach Altersgruppe für den Kennwert F % bei $.80 \leq .91$; die Kennwerte KL und BZO variieren zwischen $.89 \leq .95$. Die Retestreliabilitäten nach 1 bzw. 10 Tagen erreichen für den KL-Wert $.94$ bzw. $.85$, für BZO $.91$ bzw. $.92$ und für F % $.84$ bzw. $.47$.

Validität

Für das Verfahren wird inhaltlich-logische Gültigkeit beansprucht. Zudem liegen zahlreiche Befunde zur Validität basierend auf den Vorgängerversionen vor. Darunter befinden sich Studien zur Konstruktvalidität, faktoriellen Validität und v.a. zur empirischen Validität, die aus den Gebieten der Klinischen, Pädagogischen und Experimentellen Psychologie sowie der Verkehrs-, ABO-, Sport-, Umwelt- und Pharmakopsychologie stammen.

Leider sind die Beziehungen zu Intelligenzdaten vorwiegend für die Mengenleistung des d2 (GZ-Wert) dokumentiert und nur selten zu den anderen Variablen. Insgesamt zeigt sich, dass die Beziehungen ein mittleres Ausmaß nicht übersteigen. Ausnahmen bilden dabei einzelne Subtests des HAWIE und IST sowie spezielle Stichproben (z.B. Kinder und psychiatrische Patientinnen/Patienten), bei denen die Zusammenhänge substantiell ausfallen. Interessant ist ferner die These von Hellwig (1975), nach der die Konzentrationsleistungsvariablen des d2 als notwendige Voraussetzungen für Intelligenzleistungen angesehen werden sollten. Letztendlich ist aber ungeklärt, in welche Richtung eine Kausalinterpretation sinnvoll ist (Jäger, 1973; Wiese & Kroj, 1972).



Eine weitere Frage betrifft die Beziehung zu anderen Verfahren desselben inhaltlichen Bereiches. Auch hierzu liegt eine Reihe von Befunden vor. Die mitgeteilten Zusammenhänge zeigen, dass gleichsinnige Leistungsindikatoren aus verschiedenen Tests in maximal mittlerem Maß korreliert sind. Die Ergebnisse sind dabei hinsichtlich der Maße für die Quantität der Leistung höher als für die der Qualität. Insgesamt zeigt sich eine deutliche Tendenz, dass diese Verfahren einen substantiellen spezifischen Anteil besitzen.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Alles in allem stellt der d2-R eine sehr gelungene Weiterentwicklung des langjährig bewährten, aber durchaus kritisch diskutierten Tests d2 dar. Mit der Revision und dem neu gestalteten Testmanual werden einige konzeptuelle und methodische Kritikpunkte an der Vorgängerversion d2 (z.B. unklare konzeptuelle Einordnung, Deckeneffekte bei leistungsstarken Testpersonen) entkräftet oder zumindest abgeschwächt. Die Einführung neuer Kennwerte und der Durchschreib-Testbogen vereinfachen die ohnehin schon ökonomische Auswertung weiter. Die ausführlichen und differenzierten Hinweise zur Interpretation der Testergebnisse sind vorbildlich; hervorzuheben ist insbesondere die empirisch fundierte Diskussion um die intentionale Verfälschbarkeit der Ergebnisse, die – obwohl eigentlich unabdingbar – bislang nur bei wenigen Leistungstests so systematisch geführt wird. Wünschenswert wären sicherlich – trotz der vergleichsweise hohen Äquivalenz von d2 und d2-R – weitere Befunde zur Validität des d2-R sowie zur Stabilität der Testwerte über längere Zeiträume. Auch lassen einige empirische Befunde die Reliabilität und Validität der Masse für die Arbeitsgenauigkeit weiterhin fraglich scheinen.

Autorinnenkommentar

Mit dem d2-R kann die fokussierte Aufmerksamkeit überprüft werden. Für Kinder erweist sich eine Testung mit dem DL-KG bzw. DL-KE als vorteilhafter, welche im Grunde dieselben Kompetenzen abprüfen wie der d2-R-Durchstreichtest, mit dem Unterschied, dass sich die Distraktoren beim DL-KG und DL-KE deutlicher unterscheiden. Für Schüler/innen der Sekundarstufen I und II eignet sich dieser Test insbesondere in Kombination mit computerbasierten Verfahren zur geteilten und räumlichen Aufmerksamkeit sehr gut, um Aufmerksamkeitsdefizite festzustellen.

Bei schwierigen Fällen hat es sich bewährt, zusätzlich zum d2-R – sofern ausreichend Zeit zur Verfügung steht – auch Verfahren zur Erfassung der Arbeitseinstellung und Motivation durchzuführen, um eine eventuelle Fehldiagnose im Sinne einer Aufmerksamkeitsstörung aufgrund mangelnder Motivation zu vermeiden.

Literatur

- Bartenwerfer, H. (1982). *Allgemeine Leistungsdiagnostik*. In K.-J. Groffmann & L. Michel (Hrsg.), *Intelligenz- und Leistungsdiagnostik (Enzyklopädie der Psychologie, Themenbereich B, Serie II, Band 3., S. 482–512)*. Göttingen: Hogrefe.
- Bourdon, B. (1895). *Observations comparatives sur la reconnaissance, la discrimination et l'association*. *Revue philosophique*, 40, 153–185.
- Brickenkamp, R. (2002). *Test d2. Aufmerksamkeits-Belastungs-Test (9., vollständig überarbeitete und neunormierte Auflage)*. Göttingen: Hogrefe.
- Brickenkamp, R., Schmidt-Atzert, L. & Liepmann, D. (2010). *Test d2 – Revision Aufmerksamkeits- und Konzentrationstest*. Göttingen: Hogrefe.
- Heinrich, H. C. (1973). *Einige Bemerkungen zum d2-Durchstreichtest nach Brickenkamp*. *Diagnostica*, 19, 118–124.
- Hellwig, H.-J. (1975). *Zur Differenzierung von Intelligenztest- und Konzentrationsleistungen*. Unveröffentlichte Dissertation, Technische Universität Berlin.
- Henning, H. (1925). *Die Untersuchung der Aufmerksamkeit*. In E. Abderhalden (Hrsg.), *Handbuch der biologischen Arbeitsmethoden (Abt. VI, Teil B. S. 598-802)*. Berlin: Urban & Schwarzenberg.
- Jäger, R. (1973). *Bemerkungen zur „Untersuchung über den Zusammenhang zwischen Intelligenz (Wechsler) und Konzentrationsfähigkeit (Test d2 nach Brickenkamp) von W. Wiese und G. Kroj“*. *Zeitschrift für experimentelle und angewandte Psychologie*, 20, 572–574.
- Sharp, S. E. (1899). *Individual psychology*. *American Journal of Psychology*, 10, 329–391.
- Toulouse, E. & Pieron, H. (1911). *Technique de psychologie experimentale I*. Paris.
- Whipple, G. M. (1910). *Manual of mental and physical tests*. Baltimore: Warwick and York.
- Wiese, W. & Kroj, G. (1972). *Untersuchung über den Zusammenhang zwischen Intelligenz (Wechsler) und Konzentrationsfähigkeit (Test d2 nach Brickenkamp)*. *Zeitschrift für experimentelle und angewandte Psychologie*, 19, 690–699.

3.7. FAKT-II; Frankfurter Adaptiver Konzentrationsleistungs-Test II

Moosbrugger, H. & Goldhammer, F. (2007). FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test II (grundlegend neu bearbeitete und neu normierte 2. Auflage des FAKT von H. Moosbrugger und M. Heyden). Bern: Huber.

Beschreibung

Das Ziel des FAKT-II ist die Messung der individuellen Konzentrationsleistung. Das Verfahren kann im Kontext der berufsbezogenen, pädagogischen und klinischen Diagnostik sowie der Forschung eingesetzt werden.

Der Frankfurter Adaptive Konzentrationsleistungs-Test II (FAKT-II; Moosbrugger & Goldhammer, 2007) erfasst das Merkmal Konzentration; dieses wurde hinsichtlich seiner konzeptuellen Beziehung zu Aufmerksamkeit von Moosbrugger, Goldhammer und Schweizer (2006) systematisch eingeordnet. Es lässt sich nach Goldhammer und Moosbrugger (2007) so definieren, dass hohe Konzentration im erfolgreichen Zusammenwirken von spezifischen Aufmerksamkeitskomponenten besteht, die unter Einsatz willentlicher Anstrengung eine andauernde Selektion, Koordination und Kontrolle von Handlungsschemata leisten. Wie das Testkonzept weiter in die allgemeine Theorienlandschaft und Befundlage zu Aufmerksamkeit und Konzentration eingebunden ist, geht aus dem Manual nicht hervor.

Der FAKT-II benutzt als psychodiagnostischen Zugang ein computerbasiertes adaptives Konzept, welches sich auf die Bereiche Testtempo, Länge der Übungsphase und Testlänge bezieht und mit dessen Hilfe man bei der psychometrischen Bestimmung der Konzentrationsleistung einen besonders hohen Qualitätsstandard erzielen kann. Im FAKT-II wird die Leistung $L(i)$ gemäß der physikalischen Definition: $L(i) = \text{Arbeitsmenge } A(i) / \text{Zeitintervall } t(i)$ als Verhältnis von Arbeitsmenge $A(i)$ und Zeitintervall $t(i)$ unmittelbar geschätzt. $A(i)$ bezeichnet hier jene Arbeitsmenge, welche erforderlich ist, um jeweils ein weiteres Item i in der Abfolge des Testitems zu beurteilen, wohingegen $t(i)$ den zur Bearbeitung des weiteren Items verfügbaren Zeitraum bezeichnet. Im Test wird die Itemschwierigkeit durch Variation der Itemdarbietungszeit individuell so oft erhöht bzw. verringert, bis ein Tempo gefunden ist, bei dem die Probandin/der Proband gerade noch in der Lage ist, mit der Wahrscheinlichkeit von $p = .50$ eine adäquate Antwort zu geben („liminales“ Testtempo). Das liminale Testtempo dient als Konzentrationsleistungsmaß, der prozentuale Anteil nicht falsch bearbeiteter Items als Maß für die Arbeitsgenauigkeit und die inverse intraindividuelle Variabilität der Bearbeitungszeiten als Maß für die Arbeitshomogenität der Probandin/des Probanden.

Bei dem FAKT-II handelt es sich um ein computergestütztes Testverfahren, das in drei unterschiedlich schwierigen Testformen und zwei Durchführungsarten vorliegt, so dass sechs un-

terschiedliche Testvarianten möglich sind. In allen Varianten werden als Stimulusmaterial die speziell für eine valide Konzentrations- und Aufmerksamkeitsdiagnostik konstruierten Frankfurter Diskriminationsitems verwendet, die bereits im Frankfurter Aufmerksamkeits-Inventar FAIR genutzt wurden (Moosbrugger & Öhlschlägel, 1996). Die leicht erfassbaren visuellen Zeichen, welche auf dem Computerbildschirm dargeboten werden, bestehen aus einem äußeren Kreis, welcher als Maskierung dient, und einem inneren Symbol. Die inneren Symbole unterscheiden sich auf zwei Stimulusdimensionen mit je zwei Stufen, nämlich der Dimension „Gestalt“ (Kreis oder Quadrat) und der Dimension „Punkte-Anzahl“ (zwei oder drei Punkte innerhalb des Symbols). Weiterhin unterscheiden sich die Items in einer dritten, jedoch von der Probandin/vom Probanden zu vernachlässigenden Dimension „Anordnung der Punkte“.

Nach dem Zufallsprinzip wurde eine Abfolge von 640 Items so angeordnet, dass die vier Itemarten gleich häufig (je 160 Mal) vertreten sind. Außerdem tritt jeder Itemtyp gleich häufig in beiden Stufen der irrelevanten Dimension „Anordnung der Punkte“ auf (je 80 Mal).

Aufgabe der Testperson ist es, auf zwei Zielzeichen (Kreis mit drei Punkten, Quadrat mit zwei Punkten, unabhängig von der Anordnung der Punkte) mit dem Drücken der Zieltaste zu reagieren („1“ in der Ziffernreihe der Tastatur), während auf die Distraktoren (Kreis mit zwei Punkten, Quadrat mit drei Punkten) mit dem Drücken der „0“ zu reagieren ist. Dabei soll möglichst schnell und fehlerfrei gearbeitet werden. Zur Erinnerung sind oberhalb des/jeweils zu beurteilenden Zielitems die zwei Zielitemarten und die zu drückende Taste genannt.

Das Testverfahren ist insofern adaptiv, als die Aufgabenschwierigkeit über eine leistungsabhängige Anpassung der Itemdarbietungsdauer variiert wird. Beginnend mit einer eher hohen Dauer wird die Darbietungszeit bei richtiger Reaktion zunächst proportional zu den vorangegangenen Itemdarbietungszeiten stufenweise so lange verkürzt, bis die Probandin/der Proband während der Itemdarbietungszeit keine oder eine falsche Antwort gibt. Dann wird die Darbietungszeit stufenweise wieder so lange verlängert, bis die/der Getestete eine richtige Antwort gibt. Dieses Prozedere wiederholt sich, bis sich die Vorgabegeschwindigkeit an der individuellen liminalen Konzentrationsleistungsschwelle einpendelt, an der die Probandin/der Proband mit einer Wahrscheinlichkeit von 50 Prozent in der Lage ist, die einzelnen Items adäquat zu beantworten.

Die drei Testformen des FAKT-II unterscheiden sich wie folgt:

- FAKT-E (Einzeldarbietung; hohe Belastung): Die zu beurteilenden Items werden einzeln auf dem Bildschirm dargeboten. Die Itemdarbietungszeit beträgt zunächst zwei Sekunden und wird bei richtiger Reaktion zunächst stufenweise verkürzt, bei ausbleibenden oder falschen Antworten stufenweise so lange verlängert, bis der individuelle liminale Schwellenwert (Wahrscheinlichkeit adäquater Bearbeitung von 50 Prozent) gefunden ist.
- FAKT-S (Simultandarbietung; mittlere Belastung): Es werden jeweils 10 zu beurteilende Zielitems in einer Zeile auf dem Bildschirm eingeblendet. Im ersten, nicht zeitbegrenzten Durchgang endet die Bearbeitungszeit, sobald die Testperson fünf Items bearbeitet hat.

Es wird die Richtigkeit der Lösungen ermittelt und die Darbietungszeit wird entsprechend stufenweise verlängert bzw. verkürzt, um über die Versuchsdurchgänge hinweg den Schwellenwert zu ermitteln, an dem die Testperson jeweils 50 Prozent der 10 Items korrekt beurteilt.

- FAKT-SR (Simultandarbietung mit Reaktionszeitmessung; deutlich geringste Belastung): Die Items werden wiederum in Zeilen zu je 10 Items dargeboten. Mit einem Pfeil ist das aktuell zu beurteilende Item markiert. Alle 10 Items sind zu bearbeiten, erhoben werden Reaktionszeit und Richtigkeit der Lösung.

Bei den Durchführungsarten werden drei Varianten unterschieden:

- adaptive Testzeit: Die Testung dauert mindestens 2 Minuten und endet mit der Unterschreitung eines vorab spezifizierten Genauigkeitskriteriums für die Schätzung des liminalen Schwellenwerts, spätestens jedoch nach 6 Minuten.
- Standardtestzeit: Die Testung dauert stets 6 Minuten (einschließlich 1-minütiger Übungsphase).
- Verlängerung der Testzeit in 6-Minuten-Schritten (12, 18, 24 und 30 Minuten): Diese Testvariante wird gewählt, wenn eine Beurteilung des Konzentrationsverlaufs von Interesse ist.

Altersgruppe: Grundsätzlich ist eine Anwendung des FAKT-II ab einem Alter von 10 Jahren möglich. Auch eine Testung von Seniorinnen/Senioren ist möglich, allerdings umfasst die Normierungsstichprobe lediglich den Altersbereich von 16 bis 55 Jahren.

Hinweise zur Durchführung

Der Test kann als **Einzel- oder Gruppentest** durchgeführt werden. Er liegt in sechs Varianten vor (drei Darbietungsformen, zwei Durchführungsarten). Ferner kann das Verfahren mit verlängerter Testzeit (maximal 30 Minuten) vorgegeben werden, um den Konzentrationsverlauf abzubilden.

Bei Wahl der adaptiven Testzeit endet die Testung, sobald die Messung ein Genauigkeitskriterium erreicht, spätestens aber nach 6 Minuten. Bei Wahl der Standardtestzeit beträgt die reine Testzeit konstant 6 Minuten. Insgesamt sind für diese Durchführungsvarianten ca. 10 Minuten zu veranschlagen. Eine dritte Durchführungsoption ermöglicht die Feststellung des Konzentrationsverlaufs; hierzu kann die Testzeit in 6-Minuten-Schritten auf bis zu 30 Minuten verlängert werden.

Für die Durchführung wird ein Computer (Betriebssystem Windows) benötigt, auf dem das Ho-greffe Testsystem, das separat zu erwerben ist, und die Programmsoftware installiert sind.

Die Instruktion wird durch das Starten des jeweils zu bearbeitenden Tests auf dem Computerbildschirm angezeigt. Die Probandin/der Proband muss sich diese selbstständig durchlesen.

Danach erfolgt eine Übungsphase. Wenn während dieser Übung von der Probandin/vom Probanden mehr als fünf Fehler gemacht werden, wird die Übung abgebrochen und es erscheint automatisch die Instruktion, die von der Probandin/vom Probanden erneut durchgelesen werden muss. Die Instruktion wird durch Beispiele unterstützt, wodurch sie gut verständlich ist. Trotzdem besteht die Gefahr, dass bei Probandinnen/Probanden mit Lesestörung das Verständnis der Instruktion nicht ausreichend gewährleistet ist und durch zu häufiges Wiederholen der Übungsphase die Testdurchführungsmotivation der Testperson sinkt.

Spezifische Durchführungsanforderungen an Testleiter/innen bzw. Testpersonen oder Untersuchungsbedingungen werden nicht genannt.

Hinweise zu Auswertung und Interpretation

Es liegen Prozentränge und Standard-Nine-Werte vor; zudem werden die 95-Prozent-Konfidenzintervalle der wichtigsten Testparameter mitgeteilt. Die für die Neunormierung verwendeten Daten stammen aus den Jahren 1997 bis 2005 und beziehen sich jeweils auf die Durchführungsart „Standardtestzeit 6 Minuten inklusive 1-minütiger Übungsphase“.

Sie wurden in folgenden Eichstichproben erhoben:

Normen für eine Ersttestung mit dem FAKT-II wurden an $n = 364$ (FAKT-E), an $n = 223$ (FAKT-S) bzw. $n = 224$ (FAKT-SR) Probandinnen/Probanden im Alter von 16 bis 55 Jahren gewonnen. Es handelte sich vorwiegend um Schüler/innen bzw. Studierende, ergänzend wurden (in allen drei Stichproben weniger als 15 Prozent) berufstätige Akademiker/innen und Nicht-Akademiker/innen einbezogen. Die Normen für die Zweittestung basieren auf Daten von $n = 202$ (FAKT-S; durchgeführt nach FAKT-S) bzw. $n = 331$ Probandinnen/Probanden (FAKT-SR; durchgeführt nach FAKT-E) im Alter von 16 bis 54 Jahren. Während das Geschlechterverhältnis in den Normierungsstichproben zum FAKT-S etwa ausgewogen war, bestanden die übrigen Stichproben zu zwei Dritteln bzw. zu nahezu 80 Prozent (FAKT-SR) aus Frauen.

Für die Testform FAKT-SR liegen zusätzlich Normen einer Bewerber/innenstichprobe für den gehobenen Polizeidienst vor (erhoben 2005–2006 an $n = 338$ Probandinnen/Probanden im Alter von 18–22 Jahren, davon 54 % männlich).



Die Auswertung erfolgt vollautomatisch durch die Software. Adäquate Antworten liegen vor, wenn die Testitems richtig beurteilt werden. Inadäquate Antworten hingegen umfassen fehlende Antworten, falsche Antworten und nur zufällig richtige Antworten.

Aus den registrierten Antwortmustern (innerhalb der 6-minütigen Standardtestzeit, der adaptiven Bearbeitungsdauer oder in jedem der 6-Minuten-Intervalle bei verlängerter Testzeit) werden – je nach Testform auf unterschiedlichen Formeln basierend – drei Testwerte ermittelt und in Form von Rohwerten und Prozenträngen (einschließlich Konfidenzintervall der Prozentränge auf dem 95 %-Niveau) mitgeteilt:

1. Konzentrations-Leistung (KL), welche ein Maß für die individuelle Arbeitsgeschwindigkeit darstellt.
2. Konzentrations-Genauigkeit (KG), welche den individuellen prozentualen Anteil der nicht falsch angegebenen Antworten an allen Antworten angibt.
3. Konzentrations-Homogenität (KH), welche über die individuelle Arbeitsgleichmäßigkeit (im Sinne mehr oder weniger schwankender Itemdarbeitungszeiten bzw. Reaktionszeiten über den Testverlauf) informiert.

Die Ergebnisse werden in Tabellen- und Profiform ausgegeben; für den wissenschaftlichen Einsatz können die Daten exportiert und mit Statistikprogrammen weiterverarbeitet werden. Das Manual enthält zwei Beispiel-Ergebnisbögen, die sehr knapp kommentiert werden (Moosbrugger & Goldhammer, 2007, S. 51 ff.).

Gütekriterien

Objektivität

Die Durchführungsobjektivität im Sinne von Testleiter/innen-Unabhängigkeit ist aufgrund der computergestützten Instruktionvorgabe und Testdurchführung gegeben. Die Verständlichkeit der FAKT-II-Instruktion wurde im Rahmen einer Untersuchung an einer Stichprobe von 249 Bewerberinnen/Bewerbern (95 Frauen, 154 Männer vorwiegend mit Mittlerer Reife bzw. Hauptschulabschluss; davon $n = 26$ Migrantinnen/Migranten) ermittelt. Auf einer fünfstufigen Skala (leicht verständlich bis schwer verständlich) wählten 96 % der Probandinnen/Probanden eine der drei ersten Antwortkategorien; die Instruktionverständlichkeit kann somit als gut bewertet werden.

Die Auswertungsobjektivität ist gegeben, da die automatische Testwertbestimmung Verrechnungssicherheit gewährleistet. Ebenso kann hohe Interpretationsobjektivität postuliert werden, da die Testwerte anhand von Testnormen eindeutig interpretiert werden können.

Reliabilität

Zur Beurteilung der internen Konsistenz wurde in den Normierungsstichproben auf der Basis der Minutenergebnisse Cronbachs α für alle drei Testwerte KL 1 (Konzentrations-Leistung),

KG 1 (Konzentrations-Genauigkeit) und KH 1 (Konzentrations-Homogenität) der Testformen E, S, SR bestimmt. Die Werte liegen für die Durchführungsart „Standardtestzeit von 6 Minuten inklusive 1-minütiger Übungsphase“ bei Ersttestung zwischen $\alpha = .77$ und $\alpha = .96$ und bei Zweittestung (nur FAKT-S und FAKT-SR) zwischen $\alpha = .79$ und $\alpha = .97$. Durchweg sehr hohe Konsistenzkoeffizienten wurden für den KL-Wert ermittelt ($\alpha \geq .93$); die Werte für KG und KH liegen etwas niedriger, jedoch meist über $\alpha = .80$ (Moosbrugger & Goldhammer, 2007, S. 34 f.). Alle Testformen weisen bereits zwischen Erst- und Zweittestung (Intervall beim FAKT-E; FAKT-S: 5 bis 60 Minuten; Fakt-SR: 3–24 Tage; $72 \leq n \leq 295$) eine gute bis sehr gute Retestreliabilität für den Testwert KL auf ($r_{tt} = .70$ für FAKT-E bis $r_{tt} = .91$ bei FAKT-S). Bei mehrmaliger Testung steigen die Retestreliabilitäten nochmals deutlich an ($r_{tt} = .89$ für FAKT-E; $r_{tt} = .96$ für FAKT-S), was darauf zurückgeführt werden kann, dass bei Zweit- und Dritttestung die durch die mangelnde Übung hervorgerufene Variabilität der Konzentrationsleistungswerte nur noch sehr gering ist. Bei FAKT-E kann durch mehrmalige Testung eine besonders große Stabilitätssteigerung erzielt werden.

Validität

Das Testmanual nimmt Stellung zur Inhalts- und Konstruktvalidität sowie zur konvergenten und diskriminanten Validität, zum sog. Geschwindigkeits-Genauigkeits-Dilemma und zur kriterienbezogenen Validität (Moosbrugger & Goldhammer, 2007, S. 37 ff.).

a) Inhaltsvalidität

Es kann von einer hohen Inhaltsvalidität ausgegangen werden, da der FAKT-II viele Merkmale abbildet, die bei Konzentrationsleistungen in der Realität benötigt werden. So müssen mehrere Informationsquellen beachtet, gewonnene Informationen integriert und auf dieser Basis Entscheidungen getroffen werden; dies geschieht unter experimentell erzeugtem Zeitdruck.

b) Konstruktvalidität

Die Homogenität der FAKT-II-Items kann als gegeben betrachtet werden, weil alle Items nach dem Prinzip der Kombination zweier unabhängiger Stimulusdimensionen konstruiert wurden. Eine empirische Überprüfung der Itemhomogenitäten im Sinne der Item-Response-Theorie ergab bei 72 Modellüberprüfungen (Analyse verschiedener Testformen unter verschiedenen Optionenkombinationen) durchweg Modellkonformität mit dem polytomen Rasch-Modell (richtige vs. falsche vs. fehlende Antworten). Goodness-of-Fit-Tests nach dem strengeren dichotomen Rasch-Modell (richtige vs. falsche/fehlende Antworten) dokumentierten für 68 der 72 Fälle Modellkonformität.

- Konvergente Validität

Zwischen den KL-Werten der drei Testformen des FAKT-II bestehen enge korrelative Zusammenhänge, was für die konvergente Validität des Verfahrens spricht.

Es ergaben sich Korrelationen von $r = .40$ bis $r = .61$ zwischen FAKT und dem Konzentrations-Leistungswert KL aus dem Aufmerksamkeits-Belastungstest d2 (Brickenkamp,



1994) sowie von $r = .44$ bis $r = .55$ zwischen FAKT und dem Leistungswert L aus dem Frankfurter Aufmerksamkeits-Inventar FAIR (Moosbrugger & Öhlschlägel, 1996). Es „wird aber auch deutlich, dass die im FAKT unter experimentell induzierten situativen Belastungen gemessene Konzentrations-Leistung von der mit den Paper-Pencil-Tests erfassten Aufmerksamkeit inhaltlich klar unterschieden werden muss, und zwar sowohl für die teiladaptive Testform FAKT-SR als auch in noch stärkerem Maß für die volladaptiven Testformen FAKT-E und FAKT-S“ (Moosbrugger & Goldhammer, 2007, S. 39).

Zudem wird auf eine Analyse mit Strukturgleichungsmodellen an Daten von $n = 232$ Studierenden verwiesen (Goldhammer, Moosbrugger & Schweizer, 2007), welche die Beziehungen zwischen dem FAKT-II und fünf Aufmerksamkeitsfähigkeiten nach Posner (z.B. Posner & Rafal, 1987) spezifizieren. Hier ließ sich der KL-Wert durch geteilte Aufmerksamkeit sowie die Fähigkeit zum Aufmerksamkeitswechsel vorhersagen; die multiple Korrelation zwischen Aufmerksamkeit und Konzentrationsleistung betrug $r = .51$ (FAKT-E) bzw. $r = .60$ (FAKT-SR). Allerdings lassen sich die im Manual dokumentierten Ergebnisse nur partiell nachvollziehen, da Angaben zur Operationalisierung der Aufmerksamkeitsfähigkeiten fehlen.

In einer weiteren konfirmatorischen Analyse an der gleichen Stichprobe (Moosbrugger et al., 2006) wurden verschiedene Aufmerksamkeits- und Konzentrationsmaße theoretisch den Faktoren „perceptual attention“ (wahrnehmungsgebundene Objekterkennung und -selektion) und „executive attention“ (intentionale Kontrolle dieser Kategorisierungen) zugeordnet. Der FAKT-II-Parameter KL 1 lud theoriekonform primär auf dem Faktor „executive attention“ ($a = .62$), wies aber ebenso erwartungsgemäß (da exekutive Prozesse die Verarbeitung eines perzeptuellen Inputs betreffen) eine Nebenladung auf dem Faktor „perceptual attention“ ($a = .49$) auf.

- Diskriminante Validität

Es kann von einer hinreichenden Unabhängigkeit zwischen FAKT und Intelligenz im Sinne diskriminanter Validität ausgegangen werden.

Der korrelative Zusammenhang zwischen der Konzentrations-Leistung KL 1 in der Testform FAKT-E und FAKT-SR (Durchführungsart „Standardtestzeit von 6 Minuten inklusive 1-minütiger Übungsphase“) und dem Intelligenzfaktor Reasoning beträgt beim Wiener Matrizen-Test (WMT; Formann & Piswanger, 1979) und den Advanced Progressive Matrices (APM Set II; Raven, 1962) $r = .18$ bis $r = .41$ ($n = 197$).

Belegt wurde ferner die Unabhängigkeit des KL 1-Parameters von den Big Five-Persönlichkeitsfaktoren (gemessen mit dem NEO-FFI; Borkenau & Ostendorf, 1993; $n = 55$). Die Korrelationen variierten zwischen $r = -.19$ (Neurotizismus) und $r = .17$ (Gewissenhaftigkeit).

Geschwindigkeits-Genauigkeits-Dilemma: Das Dilemma besteht darin, dass in herkömmlichen Konzentrationstests – trotz Gleichgewichtung von Geschwindigkeit und Genauigkeit in der Instruktion – die Arbeitsgeschwindigkeit bei der Testauswertung gegenüber der Genauigkeit überbewertet wird, d.h. die Konzentrationsleistung sorgfältig arbeitender Testpersonen wird eher unterschätzt. Die Testformen FAKT-E und FAKT-S verhindern eine Überbewertung der Arbeitsgeschwindigkeit, indem das Testtempo der individuellen Konzentrationsfähigkeit angepasst, d.h. die Itemdarbietungszeit nach falschen Antworten verlängert, nach richtigen hingegen verkürzt wird. Probandinnen/Probanden, die sorgfältig, aber langsam arbeiten, müssen somit schneller arbeiten, während Probandinnen/Probanden, die fehlerhaft antworten, ihr Tempo u.U. verlangsamen müssen, bis der adaptive FAKT-Algorithmus die individuelle liminale Leistungsfähigkeit erreicht. Auch in Testform FAKT-SR wird die Bevorzugung einer schnellen und ungenauen Bearbeitung dadurch reduziert, dass bei der Berechnung der Leistungsfähigkeit die Gesamtzahl bearbeiteter Items um die doppelte Fehlerzahl reduziert wird.

- Kriteriumsvalidität

Mit Blick auf den (vorläufigen) Studienerfolg im Fach Psychologie, gemessen an der Statistiknote, ließ sich eine substantielle Kriteriumsvalidität feststellen (Jonkisz, Moosbrugger & Mildner, 2006); der numerische Kennwert wird im Testmanual leider nicht genannt. In schrittweisen Regressionsanalysen wies der Testwert KL 1 nach der Abiturnote in Mathematik die höchste inkrementelle Validität auf. Weitere Studien zur Kriteriumsvalidität fehlen allerdings (Weis & Nuerk, 2011).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Der Frankfurter Adaptive Konzentrationsleistungs-Test FAKT-II ist ein rechnergestütztes Verfahren, mit dem eine rasche, ökonomische Feststellung der Konzentrationsfähigkeit von Personen im interindividuellen Vergleich möglich ist. Das Besondere an FAKT-II ist, dass er umso schwieriger wird, je besser die Testperson sich konzentriert, was ihn gerechter als viele andere Leistungstests macht, bei denen die Aufgabenschwierigkeit nicht dem Fähigkeitsniveau angepasst wird.

Darüber hinaus sind beim FAKT-II Durchführungs-, Auswertungs- und Interpretationsobjektivität gewährleistet. Die Instruktion wurde auf ihre Verständlichkeit überprüft. Die Reliabilitätswerte des Tests sind sehr gut und das Verfahren besitzt inhaltlich-logische Gültigkeit in allen untersuchten Bereichen. Durch das computerbasierte adaptive Testverfahren sowie eine ra-

sche und genaue Diagnose ist der FAKT-II sehr ökonomisch und gut zumutbar. Außerdem besteht Testfairness, da eine systematische Diskriminierung bestimmter Personengruppen vermieden wird. Die Robustheit gegen Übungseffekte ist im Vergleich zu anderen Verfahren als hoch einzuschätzen.

Das Testprogramm des FAKT-II ist leicht am Computer zu installieren und auch die Durchführung bereitet keinerlei Probleme. Die Testauswertung erfolgt vollautomatisch, so dass die Testleiterin/der Testleiter keinerlei Arbeit hat und das Ergebnis sofort ausdrucken kann. Das Testmanual ist leicht zu handhaben, übersichtlich und – in Teilen – sehr ausführlich. Kritisch zu sehen ist – neben den hohen Anschaffungskosten – das teilweise sehr knapp gehaltene Manual, in dem bspw. die theoretische Einbettung des Verfahrens und mögliche Interpretationen nur angedeutet werden und die Darstellung einzelner Validitätsbelege – insbesondere zur konvergenten Validität – ohne Lektüre der zitierten Originalarbeiten unverständlich und intransparent bleibt, weil numerische Kennwerte nicht explizit genannt und Operationalisierungen nicht beschrieben werden. Zudem muss angesichts der hohen Korrelationen der drei Testwerte und des gleichzeitigen Fehlens von Validitätsbelegen für KG und KH die Frage aufgeworfen werden, ob nicht die Beschränkung auf KL als einzigen Testwert hinreichend wäre. Schließlich ist auf die mangelnde Repräsentativität der Normierungsstichproben hinzuweisen, die lediglich einen geringen Anteil Berufstätiger enthalten und in denen teilweise weibliche Teilnehmerinnen deutlich überwiegen.

Ungeachtet dieser Einwände handelt es sich bei dem FAKT-II um einen auf einem überzeugenden methodischen Konzept basierenden und empirisch gut validierten Test zur Erfassung von Konzentrationsleistungen, der dort, wo ein Einsatz rechnergestützter Testverfahren möglich ist, eine attraktive Alternative zu bewährten Paper-and-Pencil-Tests darstellt.

Autorinnenkommentar

Der FAKT II ist durch die Möglichkeit des adaptiven Testens eine attraktive Alternative zu den herkömmlichen Durchstreichtests, strengt aber die Testpersonen erfahrungsgemäß sehr an. Es dürfte oft sinnvoller sein, die TAP oder das WAF oder das CompACT-Verfahren durchzuführen, um ein vollständiges Bild der Aufmerksamkeitsleistung zu bekommen. Will man nur Teilaspekte genauer prüfen, so eignet sich dieser Test hervorragend.

Literatur

- Borkenau, P. & Ostendorf, F. (1993). NEO-FFI. NEO-Fünf-Faktoren Inventar nach Costa und McCrae. Göttingen: Hogrefe.*
- Brickenkamp, R. (1994). Test d2. Aufmerksamkeits-Belastungs-Test (8., erweiterte und neu gestaltete Auflage). Göttingen: Hogrefe.*

- Formann, A. K. & Piswanger, K. (1979). *Wiener Matrizen-Test*. Weinheim: Beltz.
- Goldhammer, F. & Moosbrugger, H. (2006). Aufmerksamkeit. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 16–33). Heidelberg: Springer.
- Goldhammer, F., Moosbrugger, H. & Schweizer, K. (2007). On the separability of cognitive abilities related to Posner's attention components and their contributions to conceptually distinct attention abilities related to working memory, action theory, and psychometric assessment. *European Psychologist*, 12, 103–118.
- Jonkisz, E., Moosbrugger, H. & Mildner, D. (2006). Die „Frankfurt Study“ zur Vorhersage des Studienerfolgs. In B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer & O. Vitouch (Hrsg.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (S. 342–348). Lengerich: Pabst.
- Moosbrugger, H. & Goldhammer, F. (2007). *FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test II (grundlegend neu bearbeitete und neu normierte 2. Auflage des FAKT von H. Moosbrugger und M. Heyden)*. Bern: Huber.
- Moosbrugger, H., Goldhammer, F. & Schweizer, K. (2006). Latent factors underlying individual differences in attention measures: Perceptual and executive attention. *European Journal of Psychological Assessment*, 22, 177–188.
- Moosbrugger, H. & Heyden, M. (1997). *Frankfurter Adaptiver Konzentrationsleistungs-Test (FAKT)*. Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (1996). *Frankfurter Aufmerksamkeits-Inventar (FAIR)*. Bern: Huber.
- Posner, M. I. & Rafal, R. D. (1987). Cognitive theories of attention and the rehabilitation of attentional deficits. In M. J. Meier, A. L. Benton & L. Diller (Hrsg.), *Neuropsychological rehabilitation* (S. 182–201). Edinburgh: Churchill Livingstone.
- Raven, J. C. (1962). *Advanced Progressive Matrices*. London: Lewis.
- Weis, S. & Nuerk, H.-C. (2011). TBS-TK Rezension: „FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test“. *Psychologische Rundschau*, 62, 139–141.

3.8. FAIR-2; Frankfurter Aufmerksamkeits-Inventar 2

Moosbrugger, H. & Öhlschlägel, J. (2011). FAIR-2. Frankfurter Aufmerksamkeits-Inventar 2 (2., überarbeitete, ergänzte und normenaktualisierte Auflage des FAIR von Moosbrugger & Öhlschlägel, 1996). Bern: Huber.

Beschreibung

Mit dem FAIR-2 sollen interindividuelle Unterschiede in der gerichteten Aufmerksamkeit abgebildet werden, definiert als Fähigkeit zur konzentrierten (genauen und schnellen) Unterscheidung visuell ähnlicher Zeichen unter Ausblendung aufgabenirrelevanter Information. Das Verfahren kann in der Forschung sowie in allen Praxisfeldern eingesetzt werden, in denen eine Diagnostik von Aufmerksamkeits- und Konzentrationsleistungen erforderlich ist.

Aufmerksamkeitsleistungen stellen die Grundvoraussetzung für die Bewältigung einer Vielzahl kognitiver und intellektueller Anforderungen des Alltagslebens dar. Das Ziel der Aufmerksamkeitsdiagnostik besteht darin, interindividuelle Unterschiede in dem Ausmaß abzubilden, in dem es Personen gelingt, ihre Aufmerksamkeit über einen längeren Zeitraum auf bestimmte Reize oder Tätigkeiten willentlich auszurichten (Moosbrugger & Öhlschlägel, 2011, S. 19). Die Testautorinnen/-autoren halten die vielfach vorgenommene Unterscheidung von Aufmerksamkeit und Konzentration für nicht notwendig (Moosbrugger & Öhlschlägel, 2011, S. 22); zur Begründung beziehen sie sich auf das mehrdimensionale Aufmerksamkeitsmodell von Sturm und Zimmermann (2000), das seinerseits auf dem neuropsychologisch orientierten Modell von Posner und Boies (1971) basiert, sowie auf eigene empirische Befunde. Als gemeinsame Definition von Aufmerksamkeit und Konzentration wird formuliert: „Hohe Konzentration besteht demnach im erfolgreichen Zusammenwirken jener Aufmerksamkeitskomponenten, die unter Einsatz willentlicher Anstrengung eine andauernde Selektion, Koordination und Kontrolle von Handlungsschemata leisten“ (Goldhammer & Moosbrugger, 2006, S. 31).

Nach Moosbrugger und Öhlschlägel (2011, S. 20 ff.) muss eine „konstruktgetreue Diagnostik“ von derart definierten Konzentrationsleistungen den folgenden 10 Anforderungen genügen:

- (1) Von den Testpersonen sind Diskriminationsurteile zu bestimmten Reizgegebenheiten unter zeitlicher Limitierung zu erbringen, d.h. „Konzentrationsleistung“ wird im Sinne der physikalischen Definition von Leistung als „Arbeit pro Zeit“ definiert.
- (2) Die Reizgegebenheiten sollen für alle Testpersonen leicht erfassbar und gleichermaßen geläufig sein; gleiches gilt für die geforderten Reaktionen.
- (3) Die Reizgegebenheiten sollen sich zumindest in zwei Dimensionen voneinander unterschei-

den, damit ihre Bearbeitung die mental koordinierte Beachtung mehrerer Dimensionen erfordert.

(4) Die Reizgegebenheiten sollen mindestens eine entscheidungsirrelevante Reizdimension aufweisen, die bei den Urteilen auszublenden ist.

(5) Die Verteilung der Reizgegebenheiten muss strengen Wahrscheinlichkeitstheoretischen Überlegungen genügen, um abschätzen zu können, zu welchem Anteil richtige Reaktionen der Testpersonen tatsächlich durch konzentriertes Verhalten bedingt (und nicht zufällig) sind.

(6) Die Bearbeitungsanforderung soll sicherstellen, dass unter Ausblendung der irrelevanten Information alle relevanten Reizdimensionen simultan betrachtet werden müssen.

(7) Das Arbeitsverhalten der Testperson muss während der gesamten Testbearbeitung vollständig und eindeutig protokolliert werden, damit klar entscheidbar ist, welche Abschnitte des Reizmaterials bearbeitet wurden.

(8) Die instruktionsgemäß vorgegebene Reihenfolge für die Bearbeitung der Reizgegebenheiten muss eingehalten und kontrolliert werden, damit gesichert ist, dass die Testpersonen tatsächlich mehrere Dimensionen des Reizmaterials simultan (und nicht sukzessive) beachtet haben.

(9) Die Verrechnung zu einem Testwert muss sicherstellen, dass der Testwert eine Mindestschätzung der Konzentrationsleistung darstellt; durch Korrekturformeln o.Ä. müssen Leistungsüberschätzungen durch unkonzentrierte, aber zufällig richtige Bearbeitung ausgeschlossen werden.

(10) Da die verfügbaren Aufmerksamkeitsressourcen teils auf die konkrete Arbeit selbst, teils auf die Qualitätskontrollfunktion für diese Arbeit aufgeteilt werden, soll jeder dieser Anteile durch geeignete Testwerte repräsentiert und zu einem gemeinsamen Testwert integriert werden.

Mit dem Frankfurter Aufmerksamkeits-Inventar (FAIR; Moosbrugger & Öhlschlägel, 1996) und seiner nunmehr vorliegenden aktualisierten Version FAIR-2 (Moosbrugger & Öhlschlägel, 2011) wird die Konzentrationsleistung unter Berücksichtigung dieser Anforderungen operationalisiert.



Bei Bearbeitung des FAIR sollen die folgenden Komponenten der Aufmerksamkeit aus dem Modell von Sturm und Zimmermann (2000) beansprucht werden (Moosbrugger & Öhlschlägel, 2011, S. 32):

- Alertness/Wachheit (Regulation der Reaktionsbereitschaft als Voraussetzung für die Testbearbeitung)
- Daueraufmerksamkeit/Vigilanz (Aufrechterhaltung von Aufmerksamkeit über die gesamte Testdauer durch mentale Anstrengung)
- selektive/fokussierte Aufmerksamkeit (bevorzugte Verarbeitung bestimmter Zielstimuli bei gleichzeitigem Ignorieren der Distraktoren)
- geteilte Aufmerksamkeit (Aufteilung der Aufmerksamkeitsressourcen zwischen verschiedenen Zielstimuli sowie den Bearbeitungsanforderungen Tempo und Qualität)

Das FAIR fand seit seiner Erstveröffentlichung 1996 weite Verbreitung und kann als einer der zwei „derzeit vermutlich bekanntesten Konzentrationstests“ (Wagner-Menghin, 2003, S. 253) bezeichnet werden. Das FAIR-2 ist mit dem Vorläufer konzeptuell identisch, Testmaterial, Instruktionen und zu errechnende Testkennwerte blieben ebenfalls unverändert. Veränderungen gegenüber der Erstaufgabe bestehen in der:

- Erweiterung und Aktualisierung der Normen,
- Ergänzung der Handanweisung um neuere Forschungsergebnisse und verbesserte Einordnung des Testkonzepts in die Forschung zu Aufmerksamkeit und Konzentration und
- Bezugnahme auf die Anforderungen der DIN 33430; die Erfüllung dieser Anforderungen soll durch eine Kurzdokumentation im Manual belegt werden, die aufzeigt, auf welchen Seiten des Testmanuals zu den einzelnen Kriterien der Norm Stellung genommen wird (Moosbrugger & Öhlschlägel, 2011, S. 135 ff.).

Das FAIR-2 liegt in zwei vergleichbar aufgebauten Formen A und B vor. Das jeweils achtseitige Testheft enthält auf der ersten Seite Raum für personenbezogene Angaben (Name, Geschlecht, Alter, Schulabschluss und Beruf) sowie Datum und Uhrzeit. Es folgen eine zweiseitige Instruktion mit einer Übungszeile sowie zwei Doppelseiten mit dem eigentlichen Test, auf denen sich jeweils links die relevanten Teststimuli und rechts Auswertungsfelder zum Eintragen von Zwischenergebnissen befinden. Auf der Hefrückseite sind zusammenfassende Auswertungen einzutragen.

Der Test umfasst insgesamt 640 Items, die auf zwei Seiten zu je 320 Items (16 Zeilen zu je 20 Items) verteilt sind. Jedes Item setzt sich zusammen aus einem Kreis (als Maskierungsreiz), in dessen Inneren Figuren abgebildet sind, welche sich auf drei Dimensionen unterscheiden:

1. Form (Kreis oder Quadrat),
2. Anzahl von Punkten innerhalb der Form (zwei oder drei Punkte) und
3. Anordnung der Punkte, wobei die letztgenannte Merkmalsdimension aufgabenirrelevant ist.

In jeder Testform sind zwei Arten von Items per Instruktion als Zielitems definiert (Form A: Kreis im Kreis mit drei Punkten oder Quadrat im Kreis mit zwei Punkten; Form B: Kreis im Kreis mit zwei Punkten, Quadrat im Kreis mit drei Punkten).

Aufgabe der Testperson ist es, innerhalb von 3 Minuten pro Testseite möglichst viele Zielitems zu markieren, indem sie einen Stift ohne abzusetzen in einer durchgehenden Linie unterhalb der Itemzeile entlang führt und jedes Zielitem durch einen Zacken der Linie in das Item hinein kennzeichnet und die Linie dann ohne Unterbrechung unter der Itemzeile weiter bis zum nächsten erkannten Zielitem führt. Aus den Ergebnissen werden Kennwerte für Aufmerksamkeitsquantität, -genauigkeit und -kontinuität errechnet.

Altersgruppe: 9;0–85;0 Jahre

Hinweise zur Durchführung

Bei dem FAIR-2-Verfahren handelt es sich um einen Paper-and-Pencil-Test, der als **Einzel- oder Gruppentest** durchgeführt werden kann.

Es liegen zwei Formen A und B vor, die aber nicht als echte Parallelformen zu bezeichnen sind, da sich ihre Mittelwerte signifikant voneinander unterscheiden: Form A differenziert besser im leistungsstarken, Form B besser im leistungsschwachen Bereich. Neben der deutschen Version des Verfahrens wurden Versionen des FAIR in englischer, französischer und koreanischer Sprache erstellt (Moosbrugger & Öhlschlägel, 2011, S. 11).

Die reine Bearbeitungszeit beträgt sechs Minuten (pro Testseite drei Minuten), die gesamte Bearbeitungsdauer einschließlich Instruktion und Übungszeile etwa 10 bis 12 Minuten.

Für die Durchführung werden pro Testperson ein Testbogen (Form A oder B), zuverlässiges, in dunkler Farbe schreibendes Schreibgerät und eine Stoppuhr mit Sekundenanzeige für die Testleiterin/den Testleiter benötigt. Die Testbögen können als Verbrauchsmaterial bezogen werden.

Auf dem Fragebogen ist eine zweiseitige schriftliche Instruktion abgedruckt, welche die Testperson selbstständig durcharbeiten soll. Die Instruktion beschreibt die Aufgabenstellung, veranschaulicht an Beispielen die korrekte und fehlerhafte Bearbeitung und lässt das Vorgehen an einer Übungszeile erproben. Moosbrugger und Öhlschlägel (2011, S. 38) empfehlen, die Instruktion vorzulesen, wenn Instruktionsverständnis oder Motivation zu selbstständigem Lesen bei den Testpersonen angezweifelt werden müssen. Fragen der Testpersonen sind „instruktionsnah“ zu beantworten. Es ist sicherzustellen, dass alle Testpersonen die Instruktion verstanden haben, bevor die Testung beginnt.



Die Durchführung des Verfahrens ist bei Probandinnen und Probanden mit motorischen Störungen oder Beeinträchtigungen der Sehfähigkeit nicht möglich. Die äußeren Durchführungsbedingungen entsprechen den für Leistungstests üblichen (d.h. Ausschaltung von Störfaktoren wie Lärm, beengte Arbeitsplätze, schlecht beleuchtete oder belüftete Räume, wechselseitige Ablenkung der Teilnehmer/innen bei Gruppentests).

Die Testleiterin/der Testleiter benötigt zur Durchführung des FAIR-2 keine besondere Vorbildung außer der Vertrautheit mit dem Verfahren und die Instruktion.

Hinweise zu Auswertung und Interpretation

Die Testformen A und B wurden getrennt normiert. Die Daten wurden in Gruppen- und Einzeltestungen erhoben, wobei die Stichproben sich aus Probandinnen und Probanden der Normierungsstichprobe der Erstaufgabe (Moosbrugger & Öhlschlägel, 1996) sowie neuen Probandinnen und Probanden aus jeweils mehreren Teilstichproben zusammensetzen (Moosbrugger & Öhlschlägel, 2011, S. 65 ff.).

Die Normierungsstichprobe der Form A umfasst $n = 1.990$ Probandinnen/Probanden, davon $n = 703$ aus der ersten Normierung (die Gewinnung dieser Stichprobe wird nicht beschrieben) sowie $n = 1.287$ neue Probandinnen/Probanden aus folgenden Teilstichproben:

- (1) Schüler/innen ($n = 744$, 9–17 Jahre; 62,3 % weiblich)
- (2) Studierende ($n = 237$, 18–30 Jahre, 54,9 % weiblich)
- (3) Seniorenstudierende ($n = 86$, 50–85 Jahre, 67,4 % weiblich)
- (4) Arbeitssuchende ($n = 116$ aus einer Gesamtstichprobe von $n = 1.839$, 16–61 Jahre, 31,5 % weiblich; Der Anteil dieser Teilstichprobe an der Normierungsstichprobe wurde auf 9,9 % beschränkt, da ihre Werte sich erheblich von denen der anderen Stichproben unterscheiden und angesichts des überproportionalen Anteils dieser Stichprobe zu deutlichen Verzerrungen der Normen geführt hätte.)
- (5) Jobbewerber/innen für ein Call-Center ($n = 30$, 20–67 Jahre, 50 % weiblich)
- (6) Besucher/innen einer Schule für Medizinisch-technische Assistenten ($n = 47$, 20–67 Jahre, 60 % weiblich)
- (7) Teilnehmer/innen einer eignungspsychologischen Untersuchung ($n = 29$, 14–54 Jahre, 69 % weiblich)

Form B wurde an $n = 1.003$ Probandinnen/Probanden normiert. Hier wurden $n = 850$ Probandinnen/Probanden aus der ersten Normierung (ebenfalls nicht beschrieben) um lediglich $n = 153$ neue Probandinnen/Probanden ergänzt. Letztere entstammten folgenden Teilstichproben:

- (1) Jobbewerber/innen für ein Call-Center ($n = 34$, 22–61 Jahre, 58,8 % weiblich)
- (2) Besucher/innen einer Schule für Medizinisch-technische Assistenten ($n = 69$, 16–61 Jahre, 75,4 % weiblich)
- (3) Schüler/innen eines Gymnasiums ($n = 50$, 14–17 Jahre, 50 % weiblich)

Die Gesamtstichproben für Form A und B werden jeweils hinsichtlich ihrer Zusammensetzung nach Alter, Geschlecht und Schulabschluss bzw. -besuch beschrieben; hier fällt der hohe Anteil fehlender Angaben zum Schulabschluss auf, der vermutlich zu einem erheblichen Teil darauf zurückgeht, dass Jugendliche, die noch keinen Schulabschluss besitzen, das Item nicht eindeutig beantworten konnten. Ferner werden die Verteilungen der Testwerte und ihre Beziehungen zu Geschlecht und Alter ausführlich dargestellt. Es zeigten sich für den Leistungswert, Qualitätswert und Kontinuitätswert jeweils leichte Geschlechtsunterschiede zu Gunsten von Frauen, doch wurde ein geschlechtsneutraler interindividueller Leistungsvergleich bevorzugt und daher keine getrennte Normierung vorgenommen. Angesichts der starken Leistungsunterschiede zwischen den Alterskategorien wurden dagegen altersdifferenzierte Normwerte ermittelt.

Für beide Testformen liegen Prozentrang- und Staninewerte für die Gesamtstichprobe sowie entsprechende altersspezifische Normwerte vor (Moosbrugger & Öhlschlägel, 2011, S. 119 ff.).

Form A wurde normiert für die fünf Altersgruppen 9–13 Jahre, 14–17 Jahre, 18–25 Jahre, 26–35 Jahre und 36 und mehr Jahre ($251 \leq n \leq 535$).

Form B weist Normen für die vier Altersgruppen 14–17 Jahre, 18–25 Jahre, 26–35 Jahre und 36 und mehr Jahre auf ($117 \leq n \leq 400$).

Zu beachten ist, dass die Normwerte nur für die Ersttestung Gültigkeit besitzen, da bei mehrfacher Darbietung des Verfahrens Übungseffekte in einer Größenordnung von 0,5 bis 1 Standardabweichungen (bei Studierenden) bzw. 0,27 bis 0,37 Standardabweichungen bei Seniorinnen/Senioren nachgewiesen wurden (Moosbrugger & Öhlschlägel, 2011, S. 87).

Die einzelnen Protokollierungs-, Auswertungs- und Interpretationsschritte werden bei Moosbrugger und Öhlschlägel (2011, S. 39–63) sehr detailliert erläutert und an Beispielen veranschaulicht.

Die Protokollierung umfasst fünf Schritte, die unter Verwendung von Auswertungsschablonen für die beiden Testseiten getrennt durchgeführt werden (S. 39); die resultierenden Summenwerte werden auf dem Protokollbogen notiert:

1. Bestimmen der Gesamtmenge G: Ermittlung des letzten bearbeiteten Zeichens auf der Seite.

2. Markieren und Zählen der Linienfehler FL: Auszählen gravierender Abweichungen vom vorgeschriebenen vollständigen Markierungsprinzip, d.h. Auslassungen (Unterbrechungen der Linie durch nicht markierte Zeichen), Mehrfachmarkierungen (nachträgliche Verbesserungen, doppelte Linien) und Einzelmarkierungen (Unterbrechungen der Linie links und rechts eines Zeichens). Die Fehler werden zeilenweise ermittelt und über alle Zeilen aufsummiert. Zeichen, für die ein Linienfehler vorliegt, werden bei der Auszählung der anderen Fehlerarten nicht mehr berücksichtigt, d.h. Fehler werden jeweils nur einmal kodiert. Liegen mehr als fünf Prozent Linienfehler vor, wird zu einer Testwiederholung geraten.
3. Bestimmen von Verpasser-Fehlern FV: Mit Hilfe von Schablonen wird zeilenweise die Anzahl nicht mit einer Zacke markierter Zielitems ermittelt und über alle Zeilen aufsummiert.
4. Bestimmen von Falscher-Alarm-Fehlern FA: Ebenfalls mit Hilfe der Schablonen wird zeilenweise die Anzahl mit einer Zacke markierter Nicht-Zielitems abgelesen und über alle Zeilen aufsummiert.
5. Bestimmen von Zusatzfehlern FZ: Wenn mehr als vier Verpasser-Fehler pro Testhälfte festgestellt werden, kann im Rahmen der Zusatzauswertung ermittelt werden, ob die Testperson den Bogen evtl. instruktionswidrig bearbeitet und systematisch nur eines der zwei Zielitems berücksichtigt hat (zur Bestimmung siehe Moosbrugger & Öhlschlägel, 2011, S. 107).

Für die weitere Auswertung werden die Summenwerte G, FL, FV, FA und FZ in den Auswertungsbogen auf der Testheftückseite übertragen und zu einem Gesamtsummenwert addiert. Sodann werden folgende Kennwerte zur Beurteilung des individuellen Aufmerksamkeitsverhaltens und des Instruktionsverständnisses ermittelt:

- Markierungswert M: Verständnis und Befolgen der Instruktion und des vollständigen Markierungsprinzips, berechnet als $(G-FL)/G$.
- Leistungswert L: Aufmerksamkeitsquantität; Menge der konzentriert bearbeiteten Items während der Testdauer, berechnet als $G-FL-2(FV+FA)$.
- Qualitätswert Q: Aufmerksamkeitsgenauigkeit, Anteil der konzentrierten bearbeiteten Items an der Gesamtzahl bearbeiteter Items, berechnet als L/G .
- Kontinuitätswert K: kontinuierlich erbrachte Aufmerksamkeitsleistung, berechnet als $L \times Q$.

Zur Beurteilung der Werte M, L, Q und K liegen Prozenträge und Staninewerte vor. Hinweise zur Interpretation der Testwerte geben Moosbrugger und Öhlschlägel (2011, S. 50 f.; S. 57 ff.): Liegt der Wert von M unter .95, so lässt dies auf Mängel bei der Befolgung der Instruktionen schließen, die durch unzureichendes Sprachverständnis bedingt sind; in diesem Fall sollte auf eine diagnostische Interpretation des Tests verzichtet werden. Werte > .95 deuten in Verbindung mit einem niedrigen Normwert auf eine geringe Selbstkontrolle bei der Testbearbeitung hin.

Die beiden Maße L und Q erfassen unterschiedliche Aspekte der Aufmerksamkeitsleistung; für sie erscheint eine multiplikative Verknüpfung gerechtfertigt. L bildet das Ausmaß kognitiver Ressourcen, welche zur Aufgabenbearbeitung eingesetzt wurden, Q hingegen die übergeordnete kognitive Selbstkontrolle. Eine isolierte Betrachtung beider Werte bringt die Aufmerksamkeitsleistung nicht umfassend zum Ausdruck, da beide Werte zugleich den Arbeitsstil kennzeichnen (tempo- bzw. sorgfaltsbetont). Der Wert K wiederum stellt eine gemeinsame Bewertung der Aspekte L und Q dar.

Ergänzend zu der Standardauswertung kann eine Zusatzauswertung durchgeführt werden, die aber lediglich bei höheren Fehlersummen (fünf oder mehr Verpasser-Fehler pro Testhälfte) angezeigt ist (Moosbrugger & Öhlschlägel, 2011, S. 101 ff.), um zu verhindern, dass durch die Standardauswertung die wahre Leistung der Testperson überschätzt wird. Sie beinhaltet die Bestimmung der unteren Leistungsschranke L_{\min} bei instruktionswidriger Testbearbeitung (berechnet als $L-2x[FV+FA]$) sowie eines adjustierten Testwertes L^* (berechnet als $L-2x[FZ]$), der eine genaue Lokalisation der Leistung ermöglicht: L^* liegt immer zwischen L und L_{\min} .

Die Testmappe enthält transparente Auswertungsschablonen. Die Ergebnisse werden auf dem Testbogen eingetragen, der auf zwei Seiten auch die Berechnungsvorschriften für die Standard- und Zusatzauswertung sowie eine Grafik zum Eintragen des Testprofils enthält. Zwei Ablesebeispiele für die Normen finden sich im Anhang des Manuals (Moosbrugger & Öhlschlägel, 2011, S. 119, S. 127). In der Handanweisung finden sich ferner ausführliche Hinweise zum Vorgehen bei der Auswertung und zur Interpretation, einschließlich Rechen- und Interpretationsbeispielen (S. 47–63). Ferner sind Standardmessfehler und Konfidenzintervalle der Testwerte angegeben (S. 89 f.).

Mit Hilfe eines zusätzlich zum Test zu erwerbenden Auswerteprogramms können auf Basis der manuell ermittelten Rohwerte die verschiedenen Testwerte der Standardauswertung automatisch ermittelt und in einer Form ausgegeben werden, die der letzten Seite des Testhefts entspricht.

Für die manuelle Auswertung sind nach Moosbrugger und Öhlschlägel (2011, S. 17) 5–10 Minuten pro Fall zu veranschlagen. Dieser Wert scheint aber unrealistisch knapp und dürfte eher für die rechnergestützte Auswertung angemessen sein, denn es muss auch bei vorhandener Routine angesichts der Vielzahl von Zähl- und Verrechnungsvorgängen von einem hohen Zeitaufwand ausgegangen werden.

Testgüte

Objektivität

Das Verfahren beansprucht Objektivität in Durchführung, Auswertung und Interpretation (Moosbrugger & Öhlschlägel, 2011, S. 85–87). Dieser Bewertung kann angesichts der klaren

und standardisierten Instruktionen und Auswertungshinweise, die zudem durch zahlreiche Beispiele illustriert werden, sowie der normbasierten Interpretation im Wesentlichen zugestimmt werden. Nach Fay (1996, S. 33) ist jedoch die Auswertungsobjektivität nicht unbedingt gegeben, z.B. könne es bei der Erfassung der Linienfehler zu Fehlentscheidungen kommen. Diese Kritik könnte möglicherweise durch eine einfache Untersuchung zur Auswerterübereinstimmung entkräftet werden.

Reliabilität

Analysen der Split-Half-Reliabilität erbrachten für die Testwerte L und K Reliabilitäten mit Werten von über $r = .90$. Der Qualitätswert Q weist Reliabilitäten mit Werten nahe $r = .80$ auf. Die Retest-Reliabilität beträgt bei einem Retest-Intervall von zwei Wochen $r = .81$ für die Testwerte L und K sowie $r = .73$ für den Testwert Q; bei einem Retest-Intervall von 6 Monaten betragen die Reliabilitäten für L und K ebenfalls $r = .81$, für Q $r = .40$. Analysen der Paralleltest-Reliabilität erbrachten für die Testwerte L, Q und K Werte von über $r = .80$.

Die Split-half-Reliabilität wurde an Teilen der aktualisierten Normierungsstichproben ermittelt. Es ergaben sich für L und K Werte von $r_{tt} > .90$ und für den Qualitätswert Q Werte nahe $r_{tt} = .80$; der Wert für M fiel dagegen mit $r_{tt} < .67$ deutlich ab. Auf Basis dieser Koeffizienten wurden auch Standardmessfehler und Konfidenzintervalle der Testwerte berechnet (Moosbrugger & Öhlschlägel, 2011, S. 89 f.).

Die weiteren Ergebnisse entstammen – soweit nicht anders mitgeteilt – der Erstaufgabe des FAIR. Eine Studie von Rothaar (1993) an $n = 38$ Testpersonen (5-, 10- oder 20-minütiges Intervall zwischen den Testformen) ergab eine Korrelation der L-Werte von $r_{tt} = .76$.

Die Retestreliabilität wurde von Rothaar (1993) für die Leistungswerte L über Intervalle von 5, 10 und 20 Minuten bestimmt. Sie liegt bei $r_{tt} = .85$ (Form A) und $r_{tt} = .91$ (Form B). Neuere Studien untersuchten die Retestreliabilität über längere Intervalle von zwei Wochen ($n = 53$ Studierende) bzw. sechs Monaten ($n = 68$ Seniorinnen/Senioren im Alter von 50–85 Jahren).

Insgesamt lassen sich die Reliabilitäten der Leistungs- und Kontinuitätswerte L und K unabhängig von der Bestimmungsmethode als befriedigend bis gut beurteilen und sprechen für die Stabilität des erhobenen Merkmals. Kritischer zu sehen sind die Zuverlässigkeiten des Qualitätswerts Q und v.a. des Markierungswerts M, die – sicherlich mitbedingt durch die schiefe Werteverteilung – deutlich niedriger ausfallen.

Validität

Die Validität des FAIR-2 ist durch zahlreiche empirische Belege und theoretische Argumente gestützt. Bezüglich der konvergenten Validität konnte gezeigt werden, dass hypothesenkonforme Zusammenhänge des FAIR-2 mit Tests bestehen, die ihrerseits das Konstrukt Aufmerksamkeit bzw. bestimmte Aufmerksamkeitsdimensionen erfassen (z.B. FAKT, d2, TAP, BIS-ZS, ZVT). Die diskriminante Validität wurde insbesondere im Hinblick auf Intelligenzmaße (z.B. CFT, HAWIE, LPS-4 und IST 70) untersucht, wobei sich – bis auf Probandinnen/Probanden mit geringer Intelli-

genz – hypothesenkonforme niedrige Zusammenhänge mit Intelligenz zeigten. Von Kriteriumsvalidität des FAIR kann ausgegangen werden.

Das Manual nimmt Stellung zur Inhalts- und Konstruktvalidität, zur konvergenten und divergenten Validität sowie zur Kriteriumsvalidität.

a) *Inhaltliche Validität*

Die inhaltliche Validität kann nach Moosbrugger und Öhlschlägel (2011, S. 92 f.) als gegeben angesehen werden, weil der inhaltliche Bezug zwischen den Anforderungen im Test und den Anforderungen in bestimmten Realsituationen (z.B. verschiedene Informationen beachten und integrieren, schnell und genau arbeiten müssen) unmittelbar gegeben erscheint.

b) *Konstruktvalidität*

Die Konstruktvalidität ist durch die vorhandene Itemhomogenität gegeben, da alle Items nach dem gleichen Konstruktionsprinzip entwickelt wurden und das Verfahren den 10 Postulaten der Testautorinnen/-autoren entsprechend als Aufmerksamkeitstest einzustufen ist (Moosbrugger & Öhlschlägel, 2011, S. 93).

Im Zuge der *konvergenten Validierung* wurden Zusammenhänge mit anderen Aufmerksamkeitstests ermittelt (Moosbrugger & Öhlschlägel, 2011, S. 93 ff.). Die Befunde finden sich wiederum bereits – wenn nicht anders genannt – im Manual der Vorgängerversion:

- Test d2 (Brickenkamp, 1981): Schäfer und Moosbrugger (1993) ermittelten an $n = 601$ Testpersonen für den Kennwert GZ-F (fehlerbereinigte Gesamtleistung) Korrelationen von $r = .49$ mit FAIR-L und FAIR-K. Bei Hallwachs (1994) betrug die Korrelation zwischen GZ-F und FAIR-L ebenfalls $r = .49$ ($n = 88$), bei Jung (1992) $r = .76$ ($n = 78$). Für eine französische Version beider Verfahren betrug die Korrelation $r = .51$ ($n = 268$; Küchemann, 2002).
- Frankfurter Adaptiver Konzentrations-Leistungs-Test (FAKT; Moosbrugger & Heyden, 1997): In einer neuen Untersuchung ($n = 64$) wurden Korrelationen von $.44 < r < .55$ zwischen FAIR-L und drei Versionen des computergestützten FAKT ermittelt, die im Rahmen einer Testbatterie vorgegeben wurden.
- Testbatterie zur Aufmerksamkeitsprüfung (TAP; Zimmermann & Fimm, 1993): An $n = 117$ Versuchspersonen fand sich eine Korrelation von $r = .38$ zwischen dem Subtest „Sustained attention“ und FAIR.



Die *diskriminante Validität* wurde über Korrelationen zwischen FAIR und verschiedenen Intelligenztests geprüft; die Stichproben(-größen) lassen sich nicht immer erschließen:

- Test d'Intelligence (CFT; Les Edition du Centre de Psychologie Appliquée, 1986): Küchemann (2002) fand in einer franko-kanadischen Studie eine Korrelation zwischen der mit dem CFT gemessenen Grundintelligenz und dem FAIR-L von $r = .30$. Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE-R): Mit dem Gesamtscore korrelierte FAIR-L zu $r = .36$.
- Leistungsprüfsystem (LPS; Horn, 1983): Mit Skala 4 (Problemlösefähigkeit) war FAIR-L zu $r = .12$ korreliert.
- Intelligenz-Struktur-Test IST-70 (Amthauer, 1970): Es fanden sich für alle Aufmerksamkeitskennwerte signifikante Korrelationen mit dem Gesamtwert (FAIR-L: $r = .43$; FAIR-Q: $r = .29$; FAIR-K: $r = .44$); die Korrelationen fielen bei Testpersonen mit einem Intelligenzquotienten ≥ 90 niedriger aus ($r < .20$) als bei Testpersonen mit einem IQ < 90 ($r > .26$); vor diesem Hintergrund wird die Korrelation zwischen Intelligenz und FAIR auf interindividuelle Unterschiede im Instruktionsverständnis zurückgeführt. Fay (1996) hingegen weist kritisch auf die durch die Stichprobenaufteilung bedingte mögliche Varianzeinschränkung und so künstlich niedrige Korrelation hin. Zudem stellt er die Repräsentativität der Stichprobe in Frage, wenn 52 % der Testpersonen einen IST-70-Standardwert ≤ 90 aufweisen, was einem IQ von 85 entspricht.

Eine gemeinsame Hauptkomponentenanalyse von IST-70, d2 und FAIR erbrachte u.a., dass Test d2 und FAIR jeweils separate Faktoren markierten, was trotz hoher konvergenter Validität darauf hinweist, dass die Tests teils Unterschiedliches messen (Moosbrugger & Öhlschlägel, 2011, S. 97). Schmidt-Atzert, Bühner und Enders (2006) fanden dagegen in einer gemeinsamen Faktorenanalyse verschiedener Intelligenz- und Konzentrationstests, dass FAIR und d2 gemeinsam hoch auf den beiden Faktoren „Konzentration“ und „Figurale Fähigkeiten“ luden. Insgesamt belegen die Befunde nach Ansicht von Moosbrugger und Öhlschlägel (2011, S. 98) die hinreichende Unabhängigkeit der mit FAIR gemessenen Aufmerksamkeitsleistungen und Intelligenz.

c) Kriteriumsvalidität

Zur Kriteriumsvalidität des FAIR liegen Befunde vor, welche die differentialdiagnostische Validität bei klinisch auffälligen Kindern und Jugendlichen belegen sollen. Oh (1999) sowie Oh, Moosbrugger und Poustka (2005) konnten bei jugendlichen Schizophrenie-Patientinnen und -Patienten in der akuten Krankheitsphase deutlich niedrigere Aufmerksamkeitsleistungen auf allen drei FAIR-Maßen (insbesondere der Selbstkontrollfunktion FAIR-Q) nachweisen als bei Patientinnen/Patienten mit anderen Diagnosen (z.B. affektive Störungen, neurotische Störungen, Essstörungen, hyperkinetische Störungen, Störungen des Sozialverhaltens). Diskriminanzanalytisch konnten 77,8 % der jugendlichen Schizophrenen korrekt ihrer Diagnosegruppe zugeordnet werden. Siebert (2002) fand, dass $n = 104$ Kinder mit Aufmerksamkeits-/Hyperaktivitätsstörungen höhere Werte auf den Scores FAIR-L und FAIR-K erzielten als eine klinisch unauffällige Kinderstichprobe, jedoch tendenziell schlechtere kognitive Selbstkontrolle (FAIR-Q) zeigten. Im

Manual wird ferner darauf hingewiesen, dass sich in einer Untersuchung an Seniorinnen/Senioren innerhalb einer umfangreichen Testbatterie der FAIR-L als bester Prädiktor zur Früherkennung von Demenz erwies (Moosbrugger & Öhlschlägel, 2011, S. 99).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Beim FAIR-2 handelt es sich um ein sorgfältig erarbeitetes Verfahren zur Untersuchung des individuellen Aufmerksamkeitsverhaltens, das sich auf ein klar expliziertes Testkonzept und transparente, wohldurchdachte Konstruktionsprinzipien stützt; hier hat das Manual gegenüber der Erstauflage deutlich profitiert (siehe auch Turß, 2002). Insbesondere die Einführung des vollständigen Markierungsprinzips wird von mehreren Rezensenten als Fortschritt gegenüber herkömmlichen Durchstreichaufgaben bewertet (Hagemeister & Westhoff, 1998; Thimm, 2009; Wagner, 2003). Fay (1996, S. 37) merkt allerdings zum Konzept kritisch an, dass z.B. die Forderung, in Konzentrationstests seien unbedingt Diskriminationsurteile zu erbringen, ebenso wie die vorgelegte Definition von Konzentration keinen Absolutheitsanspruch besitze. Dass gegenüber der Erstauflage FAIR (Moosbrugger & Öhlschlägel, 1996) ein neuer Titel gewählt wurde, wirkt irritierend; suggeriert wird ein neues, revidiertes Verfahren, obschon Testkonzept, Auswertung und Interpretation identisch blieben und sogar weite Teile der Normierungsstichproben in die Neuauflage übernommen, d.h. insgesamt keine fundamentalen Änderungen vorgenommen wurden.

Die zu den Gütekriterien mitgeteilten Ergebnisse weisen das Verfahren als brauchbar für den angestrebten Anwendungsbereich aus. In der Neuauflage wurden u.a. die bislang fehlenden Befunde zur Retestreliabilität über längere Zeiträume ergänzt, wenngleich die Stichproben klein und nur begrenzt aussagefähig erscheinen (insbesondere im Fall einer Seniorinnen/Senioren-Stichprobe, in der über ein 6-monatiges Intervall durchaus Merkmalsveränderungen zu erwarten sind). Die empirische Validierung erscheint trotz einiger wertvoller Erweiterungen (z.B. zur Differenzierungsfähigkeit im klinischen Bereich) mit Blick auf die kriterienbezogene und prädiktive Validität weiterhin lückenhaft. So fehlen nach wie vor Korrelationen zu Außenkriterien wie Schulnoten, Studienergebnissen oder beruflichen Leistungen, welche die Nützlichkeit des Tests im begabungs- und eignungsdiagnostischen Kontext belegen könnten (siehe auch Wagner, 2003). Ferner kritisiert Fay (1996, S. 38) die Interpretation der unerwünscht hohen Korrelation zum IST-70 und die Angaben zur faktoriellen Eigenständigkeit des FAIR v.a. gegenüber dem d2. Ferner sei die Frage nach der Trainierbarkeit, ein Grundproblem von Konzentrationstests, noch nicht beantwortet; allerdings tragen die Autorinnen/Autoren möglichen diesbezüglichen Einwänden Rechnung, indem sie angesichts festgestellter Übungseffekte deutlich darauf hinweisen, die Normen seien lediglich bei einer Ersttestung mit dem Verfahren heranzuziehen. Wünschenswert wäre indes eine Auseinandersetzung mit dem Problem des „faking bad“ von Konzentrationsleistungen; der Test erscheint in dieser Hinsicht nicht weniger verfälschungsanfällig als andere Leistungstests auch und die Gewinnung diagnostischer Hinweise auf solche Verfälschungstendenzen wären für einen Einsatz des FAIR-2 in der Begutachtungspraxis von zentraler Bedeutung.



Das Testmanual gibt über die wesentlichen Aspekte des Verfahrens zufriedenstellende Auskunft. Eine Ausnahme bildet die lückenhafte Beschreibung der Normierungsstichproben; hier fehlen Angaben zum Vorgehen bei der Gewinnung der Stichproben aus der Erstaufgabe, zu den Zeitpunkten der Normierung und zur Repräsentativität (Hagemeister & Westhoff, 1998). Ebenso bleiben nach Fay (1996, S. 39) die Schwierigkeitsunterschiede zwischen den beiden Testformen unerklärt.

Alles in allem betrachtet handelt es sich jedoch bei dem FAIR-2 um ein ökonomisches Testverfahren mit akzeptablen Gütekriterien, das eine

interessante Alternative zu anderen verbreiteten Konzentrationstests (insbesondere dem Test d2) darstellt.

Autorinnenkommentar

Das FAIR-2 eignet sich als Alternative zum d2 und zum DL-KG. Kinder im Grundschulalter haben erfahrungsgemäß mehr Spaß am DL-KG. Interessant ist das Profil von Kindern mit AD(H)S im Vergleich zu normalen Kindern. Es zeigt sich, dass Kinder mit AD(H)S im Ausmaß der verfügbaren kognitiven Ressourcen (Testwert L) sowie im Ausmaß der kontinuierlich aufrechterhaltenen Konzentration (Testwert K) signifikant über einer klinisch unauffälligen Kinderstichprobe lagen, im Ausmaß der Selbstkontrollfunktion (Testwert Q) aber tendenziell niedrigere Werte erzielen, das Profil von Kindern mit AD(H)S ist also V-förmig im Vergleich zu der klinisch unauffälligen Stichprobe, welches A-förmig ist (Siebert, 2002).

Um ein umfassendes Bild der Konzentrationsleistung zu bekommen, bietet es sich an, zusätzlich oder stattdessen computergestützte Verfahren zur räumlichen und geteilten Aufmerksamkeit durchzuführen.

Literatur

- Amthauer, R. (1970). *Intelligenz-Struktur-Test (I-S-T 70, 3. erw. Auflage)*. Göttingen: Hogrefe.
- Brickenkamp, R. (1981). *Test d2. Aufmerksamkeits-Belastungs-Test (6. Auflage)*. Göttingen: Hogrefe.
- Fay, E. (1996). *Frankfurter Aufmerksamkeits-Inventar. FAIR (Testrezension)*. In E. Fay, *Tests unter der Lupe. Aktuelle Leistungstests – kritisch betrachtet (Band 1, S. 29–40)*. Heidelberg: Asanger.
- Goldhammer, F. & Moosbrugger, H. (2006). *Aufmerksamkeit*. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik (S. 16–33)*. Heidelberg: Springer.

- Hagemeister, C. & Westhoff, K. (1998). FAIR Frankfurter Aufmerksamkeits-Inventar. H. Moosbrugger & J. Öhlschlägel. *Diagnostica*, 44 (4), 225–229.
- Hallwachs, M. (1994). Messen FAIR und d2 dasselbe? Unveröffentlichtes Manuskript. Hamburg: Psychologischer Dienst des Berufsbildungswerks.
- Horn, W. (1983). Leistungsprüfsystem (LPS) (2. Auflage). Göttingen: Hogrefe.
- Jung, T. (1992). Untersuchungen zu Aspekten der Validität psychologischer Konzentrationstests: Ein Vergleich zwischen dem Aufmerksamkeits-Belastungs-Test d2 und dem Frankfurter Aufmerksamkeits-Inventars (FAIR). Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Küchemann, D. (2002). Erprobung und Validierung einer französischen Version des Frankfurter Aufmerksamkeits-Inventar: Eine interkulturelle Äquivalenzuntersuchung. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Les Edition du Centre de Psychologie Appliquée. (1986). Manuel d'Application du Test d'Intelligence de R.B. Cattell (Culture Free Test). Echelle 3. Paris: Les Edition du Centre de Psychologie Appliquée.
- Moosbrugger, H. & Heyden, M. (1997). Frankfurter Adaptiver Konzentrations-Leistungs-Test (FAKT). Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (1996). FAIR. Frankfurter Aufmerksamkeits-Inventar. Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (2011). FAIR-2. Frankfurter Aufmerksamkeits-Inventar 2. Bern: Huber.
- Moosbrugger, H. & Reiß, S. (2004). Das Frankfurter Aufmerksamkeits-Inventar FAIR. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 103–118). Göttingen: Hogrefe.
- Oh, H. (1999). *Aufmerksamkeitsdefizite und jugendpsychiatrische Störungen*. Frankfurt am Main: Peter Lang.
- Oh, H., Moosbrugger, H. & Poustka, F. (2005). Kann eine spezifische Aufmerksamkeitsdiagnostik zur Differentialdiagnostik psychischer Störungen beitragen? *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 33 (3), 181–189.
- Petermann, F. (2011). Frankfurter Aufmerksamkeits-Inventar 2 (FAIR-2) [Testbesprechung]. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 59 (4), 325–326.
- Posner, M. I. & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78, 391–408.
- Rothaar, P. (1993). Untersuchung von Interferenz und Übungseffekten bei serieller Anwendung von Parallelformen des Frankfurter Aufmerksamkeits-Inventars (FAIR). Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Schäfer, N. & Moosbrugger, H. (1993). Einige Validitätsaspekte des Frankfurter Aufmerksamkeits-Inventars (FAIR) (Arbeiten aus dem Institut für Psychologie der Johann Wolfgang Goethe-Universität, Heft 10/1993). Frankfurt am Main: Johann Wolfgang Goethe-Universität Frankfurt am Main, Institut für Psychologie.
- Schmidt-Atzert, L., Bühner, M. & Enders, P. (2006). Messen Konzentrationstests Konzentration? Eine Analyse von Konzentrationstestleistungen. *Diagnostica*, 52 (1), 33–44.
- Siebert, U. (2002). Frankfurter Aufmerksamkeits-Inventar (FAIR, Moosbrugger & Öhlschlägel, 1996) als differentialdiagnostisches Instrument bei Kindern mit Aufmerksamkeitsdefizitsyndrom mit oder ohne Hyperaktivität (ADD/ADHD) im Alter von 8–12 Jahren. Unveröffentlichte Diplomarbeit,

Johann Wolfgang Goethe-Universität Frankfurt am Main.

- Sturm, W. & Zimmermann, P. (2000). Aufmerksamkeitsstörungen. In W. Zimmermann, M. Herrmann & C. Wallesch (Hrsg.), *Lehrbuch der klinischen Neuropsychologie* (S. 345–365). Lisse, NL: Swets & Zeitlinger.
- Thimm, M. (2009). Frankfurter Aufmerksamkeits-Inventar (FAIR). In D. Schellig, R. Drechsler, D. Heine-
mann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 181–186). Göttingen: Hogrefe.
- Turß, M. (2002). Frankfurter Aufmerksamkeits-Inventar (FAIR). In U. P. Kanning & H. Holling (Hrsg.),
Handbuch personaldiagnostischer Instrumente (S. 149–154). Göttingen: Hogrefe.
- Wagner, M. M. (2003). H. Moosbrugger & J. Öhlschlägel (1996). Frankfurter Aufmerksamkeits-In-
ventar (FAIR) (Testrezension). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24 (1),
81–83.
- Wagner-Menghin, M. M. (2003). Konzentrationstests. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüs-
selbegriffe der psychologischen Diagnostik* (S. 249–253). Weinheim: Beltz.
- Zimmermann, P. & Fimm, B. (1993). *Testbatterie zur Aufmerksamkeitsprüfung (TAP)*. Freiburg: Psytest.

3.9. FWIT; Farbe-Wort-Interferenztest

Bäumler, G. (1985). Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop. Göttingen: Hogrefe.

Beschreibung

Als objektiver und zuverlässiger, mehrdimensionaler Leistungstest misst der Farbe-Wort-Interferenztest elementare Fähigkeiten der Informationsverarbeitung (Auswahl, Codierung und Decodierung) im optisch-verbale Funktionsbereich. Mit dem Test werden Verarbeitungsgeschwindigkeiten erfasst (Wahrnehmung, begriffliche Umsetzung und verbale Wiedergabe von Reizen durch das Lesen bzw. Benennen). Mittels des Farbe-Wort-Inkongruenzprinzips nach J. R. Stroop wird zudem die Stressbelastung und damit die konzentrierte Beanspruchung variiert. Die mit dem Test messbaren kognitiven Leistungsfunktionen sind: Nomination (Geschwindigkeit der Namenfindung, Benennung), Selektivität (konzentrativer Widerstand gegenüber dominierenden Reaktionstendenzen oder Interferenzneigung), Alertness (Grundgeschwindigkeit der Informationsverarbeitung) sowie – unter bestimmten Voraussetzungen – die Lesegeschwindigkeit. Die Auswertung liefert ein dreidimensionales Leistungsprofil.

Der FWIT geht zurück auf eine auf dem Stroopschen Interferenzprinzip (Stroop, 1935) beruhende Testtafel, die bereits in den 40er Jahren in den USA eingesetzt wurde. In Deutschland wurde die differentiell-psychologische Verwertbarkeit der Interferenzneigung (Rigidität) v.a. von Thomae (1951), Hörmann (1956) und Merz (1961) erkannt und untersucht.

Der FWIT ist ein sensomotorischer Speed-Leistungstest und soll die kognitiven Grundfunktionen des Lesens, Benennens und der Selektivität (Farbe-Wort-Interferenz) erfassen und einen Einblick in die allgemeine psychische Leistungsfähigkeit der Probandin/des Probanden vermitteln. „Die im Test gemessenen kognitiven Leistungsfunktionen sind: Lesegeschwindigkeit, Benennungsgeschwindigkeit (Nomination), konzentrativer Widerstand gegenüber dominierenden Reaktionstendenzen (sog. Selektivität bzw. Interferenzneigung) und sensomotorische Aktionsgeschwindigkeit (allgemeiner Speedfaktor oder Grundgeschwindigkeit der Informationsverarbeitung)“ (Bäumler, 1985, S. 7).

Der FWIT besteht aus einer Übungstafel und 9 Testtafeln mit je 72 in drei Spalten angeordneten Items.

Im Test sind die folgenden drei Aufgabenarten (Subtests) enthalten:

- Farbwörter-Lesen (FWL): Lesen von in schwarzer Farbe gedruckten Farbwörtern.
- Farbstriche-Benennen (FSB): Benennen der Farben von Farbstrichen.

- Interferenzversuch (INT): Benennen der Farbe von farbig gedruckten Farbwörtern bei Farbe-Wort-Inkongruenz (z.B. das Wort „grün“ ist in gelber Farbe gedruckt).

Jede Aufgabe wird mit drei Testtafeln, die mit a, b und c gekennzeichnet sind, geprüft, wobei die drei Subtests in gleichbleibendem Schema aufeinanderfolgen:

Farbwörter-Lesen (FWL-a)
 Farbstriche-Benennen (FSB-a)
 Interferenzversuch (INT-a)

Farbwörter-Lesen (FWL-b)
 Farbstriche-Benennen (FSB-b)
 Interferenzversuch (INT-b)

Farbwörter-Lesen (FWL-c)
 Farbstriche-Benennen (FSB-c)
 Interferenzversuch (INT-c)

Die Sequenz der einzelnen Items ist innerhalb der Testtafeln zufallsverteilt, allerdings folgen nie gleiche Items unmittelbar aufeinander oder mit nur einem Item dazwischen. Bei der Interferenztafel treten ausschließlich inkongruente Farbe-Wort-Kombinationen auf und die verschiedenen Farben und Wörter treten in jeder Tafel und in jeder Spalte gleich häufig auf. Die von der Probandin/vom Probanden benötigte Bearbeitungszeit ist das wichtigste Leistungskriterium des Tests. Bei den Interferenztafeln werden außerdem Benennungsfehler und Verbesserungen registriert, jedoch nicht in die Testscores eingearbeitet.

Relevante Hinweise sollen zudem die im Test auftretenden stressbezogenen Verhaltensweisen liefern. Zur Erfassung der Leistungen im FWIT wurden folgende Indikatoren konzipiert:

- FWL: „Lesewert“, errechnet aus dem Median der Lesezeiten (Rohwerte) der drei FWL-Tafeln – in T-Werte transformiert.
- FSB: „Benennungswert“, errechnet aus dem Median der Benennungszeiten (Rohwerte) der drei FSB-Tafeln – in T-Werte transformiert.
- INT: „Konfliktwert“, errechnet aus dem Median der Benennungszeiten (Rohwerte) der drei INT-Tafeln – in T-Werte transformiert.

Diese drei Variablen liefern das „Subtestprofil“. Außerdem werden noch zwei regressionsbereinigte Residualwerte als Kenngrößen berechnet:

NON: Bereinigte Benennungsgeschwindigkeit oder „Nomination“: Das Farbwörter-Lesen (FWL) wird aus dem Farbstriche-Benennen (FSB) herauspartialisiert (bereinigtes Farbstriche-Benennen). Als Synonyme für NON werden FSB, FWL, Delta FSB oder dFSB verwendet.

SEL: Bereinigte oder spezifische Interferenzvariable, konzentrativer Widerstand oder „Selektivität“: Leistung im Interferenzversuch (INT) nach Herauspartialisieren des Farben-Benennens. Die Variable SEL soll die individuelle Störanfälligkeit des Benennens bei interferierendem Reizmaterial ausdrücken. Als Synonyme für SEL werden INT, FSB, Delta INT oder dINT verwendet.

Die Variablen FWL, NON und SEL ergeben das Drei-Faktoren-Profil. „Das Drei-Faktoren-Profil hat allerdings nur dann Gültigkeit, wenn angenommen werden kann, dass die jeweilige Stichprobe bzw. Population keine nennenswerte Variation im vierten Faktor (der spezifischen Lesegeschwindigkeit) enthält, bzw. wenn im konkreten Einzelfall der Proband keine spezifische Lesestörung (z.B. Legasthenie) – oder auch keine außerordentlichen Fähigkeiten im Schnellesen – aufweist. Andernfalls würde die bereinigte Variable NOM auch den Faktor der spezifischen Lesefähigkeit (abgek. LEF) enthalten und NOM wäre nicht mehr eindeutig interpretierbar“ (Bäumler, 1985, S. 9).

Altersgruppe: 10;0–84;0 Jahre

Hinweise zur Durchführung

Das Verfahren soll als **Einzeltest** durchgeführt werden. Parallelformen liegen nicht vor.

Die Dauer der gesamten Testdurchführung wird mit 10–15 Minuten angegeben.

Die auf der Übungstafel des Testheftes vorliegende Instruktion kann entweder wörtlich von der Testleiterin/vom Testleiter vorgelesen oder der Probandin/dem Probanden zur selbstständigen Kenntnisnahme vorgelegt werden. Im letzteren Fall braucht die Probandin/der Proband nur noch „von einigen wenigen Hinweisen des Testleiters“ (Bäumler, 1985, S. 34) unterstützt zu werden. Jede der drei Aufgabentypen wird auf der ersten Tafel des Testheftes anhand von 15 Items geübt.

Genauere Anweisungen zum Verhalten der Testleiterin/des Testleiters und der Testperson sowie zum Testablauf sind in der Handanweisung abgedruckt.



Das Verfahren kann von jeder Fachkraft, die mit den Prinzipien der Anwendung und Auswertung von Tests vertraut ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Die Leistungen im FWIT können mit sechs verschiedenen T-Wert-Normen verglichen werden, denen jeweils verschiedene Eichstichproben zugrunde lagen:

- Allgemeine („absolute“) Leistungsnorm:
Die Eichstichprobe umfasste 480 männliche und weibliche Probandinnen/Probanden im Alter von 16–84 Jahren, bestehend aus Angestellten, Akademikerinnen/Akademikern, Klinikpatientinnen/-patienten, Seniorinnen/Senioren, Soldatinnen/Soldaten und Berufskraftfahrerinnen/-fahrern und wird vom Autor als bevölkerungsrepräsentativ bezeichnet. Normentabellen ohne Altersdifferenzierung liegen für die Grundvariablen FWL, FSB und INT sowie für die regressionsbereinigten Faktoren NOM (dFSB) und SEL (dINT) vor.
- Altersnorm für Jugendliche (selegierte Stichprobe):
Die Eichstichprobe umfasste 250 männliche und weibliche Jugendliche im Alter von 10–20 Jahren und bestand ausschließlich aus Gymnasialschülerinnen/-schülern. Sie ist deshalb für die Gesamtpopulation der Jugendlichen nicht repräsentativ. Es liegen altersspezifische Normen (Ein-Jahres-Schritte) für die drei Grundvariablen FWL, FSB und INT vor.
- Altersnorm für Erwachsene:
Die Eichstichprobe umfasste 850 männliche und weibliche Probandinnen/Probanden im Alter von 16–84 Jahren und bestand ähnlich wie die allgemeine Leistungsnorm aus Personen „aus sehr unterschiedlichen Lebensbereichen“ (Bäumler, 1985, S. 38). Altersspezifische Normentabellen (14 Altersklassen zu je 5 Jahren) liegen für die Grundvariablen FWL, FSB und INT vor.
- Angestellten-Norm:
Die Eichstichprobe umfasste 170 nichtakademische Angestellte verschiedener öffentlicher Einrichtungen im Alter von 16–45 Jahren (Haupt- und Realschulabschluss bzw. mittlerer Dienst) beiderlei Geschlechts. Normentabellen existieren für die Grundvariablen PWL, FSB und INT sowie für die Faktoren NON (dFSB) und SEL (dINT).
- Studierenden- und Akademiker/innen-Norm:
Die Eichstichprobe umfasste 150 Studierende und Akademiker/innen im Alter von 19–50 Jahren. Normentabellen liegen für die drei Grundvariablen FSB, FWL und INT sowie für die Faktoren NON (dFSB) und SBL (dINT) vor.

- Übungsnormen (Studierende):

Die Eichstichprobe umfasste 70 Studierende im Alter von 19–30 Jahren, die den Test insgesamt dreimal im Abstand von zwei bis drei Tagen durchführten.

Außerdem sind in der Handanweisung eine allgemeine Transformationstabelle zur Ermittlung einer logarithmischen Messwerttransformation (log FWL, log FSB und log INT) sowie Normentabellen zur Berechnung der Erwartungswerte zu log FSB und log INT aus log FWL bzw. aus log FSB abgedruckt. Die Tabellen zur Berechnung der regressionsbereinigten Faktoren NON und SEL sind nur für die allgemeine Leistungsnorm, die Angestelltenorm und die Akademiker/innen-Norm relevant. Der Standardmessfehler und der Vertrauensbereich der drei Grundvariablen FWL, FSB und INT sind ebenso wie die kritischen Differenzen des Dreier- (FWL, FSB, INT) und des Zweierprofils (NON und SEL) ausgewiesen.

Zuerst werden die Bearbeitungszeiten der neun Testtafeln bzw. die Medianwerte der Subtests auf den Protokollbögen eingetragen. Dann folgt die Eintragung des graphischen Rohwertprofils auf der Rückseite des Bogens. Die Medianwerte der drei Subtests werden anhand der Tabellen in T-Werte überführt. Dann erfolgt die logarithmische Transformation der Medianwerte, die Ermittlung der Erwartungswerte für log FSB und log INT und schließlich die Ermittlung der T-Normen für die Differenz zwischen den empirischen log-Werten (log FSB bzw. log INT) und den aus log FWL bzw. log FSB erwarteten Werten mittels der dafür vorgesehenen Tabellen in der Handanweisung. Die T-Werte der Differenzen dFSB und dINT kennzeichnen die individuelle Ausprägung des bereinigten Benennungsfaktors NON (Nomination) und des bereinigten Interferenzfaktors SEL (Selektivität).

Die Normweltergebnisse können auch graphisch im Protokollbogen als Normen-Profile dargestellt werden.

Anweisungen zur inhaltlichen Interpretation des Testergebnisses, zu der auch Fehler und Verbesserungen (bei der INT-Tafel) und das Testverhalten der Probandin/des Probanden herangezogen werden können, sind in der Handanweisung enthalten.

Auf der Vorderseite der Protokollbögen ist zudem die Itemfolge der Interferenztafeln abgedruckt.

Testgüte

Objektivität

Die Objektivität der Testdurchführung ist dadurch beeinträchtigt, dass unerfahrene Testleiter/innen Unterschiede im Motivierungsverhalten, bei der Anerkennung von Benennungsfehlern (Subtest INT) und bei Entscheidungen über sogenannte Repetitionen aufweisen, die sich aber nach Meinung des Autors nur geringfügig auf die Testergebnisse auswirken. Die Auswertung des Verfahrens hingegen kann als weitgehend objektiv gelten.



Was die empirische Überprüfung der Objektivität betrifft, so wird vom Autor eine Gesamtobjektivität von .98 – .99 berichtet. Leider sind weitere Angaben bezüglich der Ermittlung dieser Werte nicht dokumentiert.

Reliabilität

Die Konsistenzen und Wiederholungszuverlässigkeiten der Subtests des FWIT sowie die Profilzuverlässigkeiten des FWIT liegen auf relativ hohem Niveau und genügen damit durchaus den herkömmlichen Ansprüchen an die Reliabilität eines Tests. Allerdings sind die verwendeten Stichproben nicht hinreichend spezifiziert, d.h. es fehlen sowohl Angaben über ihren Umfang als auch über

ihre genaue Zusammensetzung. Außerdem sagt der Autor auch nichts über die bei der Berechnung der Retest-Reliabilität benutzten zeitlichen Zwischenräume.

Ebenfalls berechnet wurden die Subtestinterkorrelationen für homogene und heterogene Stichproben. Dabei ergaben sich Korrelationskoeffizienten von $r = .48$ bis $r = .79$.

Validität

Für das Verfahren wird inhaltlich-logische Validität beansprucht. Zum Nachweis der mit dem Test erfassten Leistungsfaktoren wurde eine Dimensionsanalyse (innere Validität) durchgeführt.

Eine nach zehnmaliger Wiederholung durchgeführte Faktorenanalyse der drei Subtests zeigte keine Veränderung der Faktorenstruktur aufgrund von Übungseinflüssen. Hierbei wurden vier Faktoren extrahiert: ein Generalfaktor (allgemeine Aktionsgeschwindigkeit), ein allgemeiner Benennungsfaktor, ein Interferenzfaktor und ein spezifischer Lesefaktor.

Vom Autor werden hinsichtlich der empirischen Validität zahlreiche Untersuchungen aufgeführt. Das Zustandekommen der angegebenen Korrelationen, die meist nur als bestimmte Bereiche dargestellt sind, ist zudem bezüglich der verwendeten Stichproben nur ungenügend dokumentiert.

Für den Zusammenhang zwischen den FWIT-Subtests und Intelligenztests bzw. kognitiven Tests werden Korrelationen von $r = .00$ bis ca. $r = .55$ genannt. Bei Merkfähigkeits- und Wortflüssigkeitstests reichen die Werte bis $r = .44$.

Die Korrelationen der FWIT-Subtests mit Konzentrations- und Aufmerksamkeitstests liegen im Mittel bei $r = .45$.

Der Zusammenhang zwischen Sensumotorik-Tests und FWIT-Subtests erreicht höchstens $r = .42$.

„Selbstvertrauen“, „soziale Unbekümmertheit“ und „emotionale Stabilität“ (aus einem mehrdimensionalen Persönlichkeitsfragebogen von Bäumler, unveröffentlicht) korrelieren zwischen $r = .20$ und $r = .30$ mit allen drei Subtests. Auch zwischen Schulnoten und dem Generalfaktor des FWIT bzw. dem Subtest FWL zeigen sich positive Beziehungen von $r = .20$ bis $r = .40$ bzw. $r = .30$ bis $r = .50$.

Lehrer/innen- und Vorgesetztenurteile (über Schüler/innen bzw. Firmenangestellte) zeigen Zusammenhänge mit dem FWIT-Subtest, die zwischen ca. $r = .25$ und $r = .60$ liegen.

Vom Autor werden außerdem verschiedene Untersuchungen zu Gruppendifferenzen hinsichtlich Alter, Geschlecht, Begabung, verschiedenen psychopathologischen Gruppen und bestimmten Berufs- bzw. Leistungsgruppen sowie Legasthenie und Farbenschwäche in der Handanweisung aufgeführt.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Aufgrund der angeführten Ergebnisse kann der FWIT als durchaus objektives und zuverlässiges Verfahren bezeichnet werden. Allerdings – und das gilt für sämtliche Gütekriterien – sind die angeführten Untersuchungen hinsichtlich der verwendeten Stichproben und der ermittelten Ergebnisse meist unzureichend und ungenau dokumentiert. Für die mittels des Verfahrens erhaltenen Nebeninformationen (Fehler und Verbesserungen, Testverhalten der Probandin/des Probanden) kann Auswertungsobjektivität nicht beansprucht werden und auch ihre Bedeutung und diagnostische Valenz ist nicht geklärt. Die Auswertung des Tests erweist sich als etwas kompliziert und die verschiedenen Abkürzungen für ein und dieselbe Variable (z.B. SEL = INT, FSB = Delta INT = dINT) tragen nicht unbedingt zur Vereinfachung der Testdurchführung bei. Außerdem wäre bei einigen Eichstichproben eine zahlenmäßige Aufstockung und bei der Altersnorm für Jugendliche die Ausweitung mittels einer unselektierten Stichprobe wünschenswert.

Von den für die Eichung verwendeten Stichproben her macht der FWIT ohnehin eher den Eindruck, in erster Linie für Probandinnen/Probanden aus der gehobenen Mittelschicht bzw. aus dem höheren Bildungsbereich entworfen worden zu sein. Eine erneute Eichung an repräsentativeren Stichproben bzw. eine Erweiterung der Normen für andere Populationen (z.B. Jugendliche aus Haupt- und Realschulen) ist daher angeraten.

Autorinnenkommentar

Der FWIT ist sehr gut geeignet, um die Interferenz-Fähigkeit von Personen zu überprüfen. Es handelt sich um ein schnell erklärtes und kurzweiliges Verfahren, mit dem sehr schnell die Interferenzfähigkeit festgestellt werden kann. Will man allerdings eine umfassende Aufmerksamkeitsdiagnostik betreiben, sollte unbedingt ein weiteres Verfahren hinzugezogen werden.

Literatur

- Bäumler, G. (1985). *Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop. Handanweisung*. Göttingen: Hogrefe.
- Hänze, M. (1997). Mood and the Stroop interference effect. *Psychologische Beiträge*, 39 (3), 229–235.
- Heinemann, D. (2009). *Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 971–978). Göttingen: Hogrefe.
- Hörmann, H. (1956). Zum Problem der psychischen Starrheit (Rigidität). *Zeitschrift für experimentelle und angewandte Psychologie*, 3, 662–683.
- Jost, K. (1992). Die Interferenzneigung Schizophrener in der Version des Stroop-Tests nach G. Bäumler. *Psychologische Beiträge*, 34, 3–15.
- Lamberti, G. (1983). Untersuchungen zur Farbe-Wort-Interferenz bei schizophrenen Erkrankungen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 4, 141–149.
- Merz, F. (1961). Über die individuelle Interferenzneigung. *Zeitschrift für experimentelle und angewandte Psychologie*, 8, 381–392.
- Puhr, U. & Wagner, M. (2001). *Interferenztest nach Stroop. STROOP (Version 22.00. Computerprogramm mit Manual)*. Mödling: Schuhfried.
- Schwenkmezger, P. (1986). *Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop. & G. Bäumler. Diagnostica*, 32 (2), 171–173.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Thomae, H. (1951). Experimentelle psychologische Diagnostik. *Fortschritte der Neurologie, Psychiatrie und ihrer Grenzgebiete*, 19, 1–22.

3.10. INKA; Inventar Komplexer Aufmerksamkeit

Heyde, G. (1995). Inventar Komplexer Aufmerksamkeit (INKA). Frankfurt am Main: Swets Test Services.

Beschreibung

Das INKA misst die Konzentrationsfähigkeit über die Kapazität der Informationsverarbeitung. In 18 Buchstabenreihen sollen bestimmte Buchstaben und/oder -Pärchen gesucht werden. Die Buchstaben müssen zuerst über eine Umwandlungstabelle gefunden werden.

Beim Inventar Komplexer Aufmerksamkeit handelt es sich um einen Konzentrationstest, bei dessen Bearbeitung nicht nur die Schnelligkeit der Testperson ausschlaggebend ist. Erforderlich ist vielmehr eine komplexe Konzentrationsleistung, die von Heyde (1995, S. 7) als „Güte der Informationsverarbeitung“ definiert wird. Ausgehend von der Kritik an gängigen Konzentrationstests, dass hohe Leistung auch ohne Konzentration möglich sei (Öhlschlägel & Moosbrugger, 1991), und dem dadurch ausgelösten wissenschaftlichen Disput (z.B. Brickenkamp, 1991; Öhlschlägel & Moosbrugger, 1991), wurde mit dem INKA ein bis dahin noch nicht realisiertes Aufgabenprinzip entwickelt. Vor der Diskrimination visueller Reize ist eine Umkodierung der vorgegebenen Reize und deren Memorierung zu leisten. Die Aufgaben können nur insgesamt fehlerfrei, fehlerhaft oder gar nicht bearbeitet werden. Der Schwierigkeitsgrad der Aufgaben nimmt in dem raschskalierten Verfahren im Laufe der Testbearbeitung zu, so dass statt der reinen Speedkomponente auch eine Messung des Fähigkeitsniveaus erfolgt. Zugleich soll durch die Aufgabenkonstruktion gewährleistet sein, dass die unterschiedliche Ausprägung individueller Fertigkeiten oder intellektueller Fähigkeiten sich nicht auf die Testleistung auswirken kann (Heyde, 1995).

Das raschskalierte Verfahren besteht aus insgesamt 18 Aufgaben. Eine Aufgabe wird durch eine „Suchreihe“ von zufällig kombinierten Konsonanten gebildet, die von Aufgabe zu Aufgabe nach verschiedenen Konsonanten oder Konsonantenpärchen durchsucht werden soll (RFLPHZR usw.). Die zu suchenden Konsonanten sind über eine Vorgabe (z.B. B, TH) und eine Umwandlungstabelle (B --> Z, C --> R, H --> T, T --> D usw.) zu erschließen. Die Vorgaben bestehen aus einem, zwei oder drei Elementen, jedes Element ist entweder ein Konsonant oder ein Konsonantenpärchen.

Die Suchreihen sind von links nach rechts abzarbeiten. Zur Aufgabenlösung sind diejenigen Buchstaben am rechten Rand der Suchreihe zu notieren, die links neben den gesuchten Konsonanten stehen, so dass sich nur eine einzige richtige Lösungsreihe aus 5, 6 oder 7 Konsonanten ergeben kann. Die Wahrscheinlichkeit, durch zufälliges Notieren von Buchstaben die richtige Lösung zu erhalten, geht gegen Null. Während der Aufgabenbearbeitung darf nichts notiert

oder in der Suchreihe markiert werden. Für die Bearbeitung der Aufgaben stehen 15 Minuten zur Verfügung. Die Anzahl der richtig gelösten Aufgaben bildet den Gesamtstandardwert des Tests, der innerhalb des Raschmodells als Personenparameter interpretiert wird.

Altersgruppe: 14;0–39;0 Jahre

Hinweise zur Durchführung

Der Test ist **einzel**n oder **in der Gruppe** durchzuführen.

Es liegen zwei Parallelformen für Rechts- und Linkshänder vor, die sich lediglich in der Anordnung der Suchreihen, der Vorgaben und dem freien Platz für die Lösungsbuchstaben unterscheiden. Statistisch sind die beiden Testformen äquivalent, Mittelwertsunterschiede sind nicht signifikant (t-Tests; Heyde, 1995, S. 27).

Für die Durchführung des Verfahrens sind ca. 25 Minuten zu veranschlagen (inklusive Instruktion und Übungsphase mit zehnminütiger Dauer).

Dem Verfahren ist eine ausführliche mündliche Instruktion voranzustellen, deren standardisierter Wortlaut bei Heyde (1995, S. 17 f.) abgedruckt ist.

Zur Durchführung des Verfahrens sind keine besonderen Qualifikationen erforderlich.

Hinweise zu Auswertung und Interpretation

Die Höhe der Korrelation des INKA-Gesamtwertes mit der Schulbildung ($r = .35$, Redundanz = 12 %) gab den Ausschlag für die Erstellung schulbildungsspezifischer Normen in Form von Z-Werten ($M = 100$, $SD = 10$; Schulbildung: Hauptschule, mittlerer Bildungsweg, Abitur bzw. Fachhochschulreife, Hochschule; Heyde, 1995, S. 23). Es stehen aber auch Normen der Gesamtstichprobe ($n = 615$ im Rahmen der Berufseignungsdiagnostik anfallende Personen) zur Verfügung. Mittelwerte und Standardabweichungen der Gesamtstichprobe und der nach Schulbildung differenzierten Teilstichproben werden ebenfalls bei Heyde (1995, S. 25) aufgeführt.

Die Auswertung erfolgt mit Hilfe einer Schablone, die eine schnelle Kontrolle der Lösungen erlaubt. Je ein Punkt wird für eine richtig gelöste Aufgabe vergeben, anschließend wird durch Addition der Gesamtröhrewert ermittelt.

Die Auswertung nimmt pro Fall nur wenige Minuten in Anspruch.

Testgüte

Objektivität

Aufgrund der standardisierten Instruktion und der Eindeutigkeit der Aufgabenlösung kann beim Befolgen der Anwendungsvorschriften die Durchführungs- und Auswertungsobjektivität des Verfahrens als gewährleistet angesehen werden.

Reliabilität

Das Verfahren ist im Sinne des Raschmodells homogen. Eine Überprüfung der Retestreliabilität im Sinne der Klassischen Testtheorie fand nicht statt.

Validität

Zur Validierung des Verfahrens wurde einer Stichprobe von 615 Personen neben dem INKA der Intelligenz-Struktur-Test IST-70 (Amthauer, 1973) und der Aufmerksamkeits-Belastungs-Test d2 (Brickenkamp, 1987) vorgelegt. Die gemeinsame Faktorenanalyse (Hauptkomponentenanalyse mit Varimaxrotation) erbrachte eine Dreifaktorenlösung, bei der auf einem der Faktoren nur der INKA-Gesamtwert, der d2 und der IST-Untertest „Merkaufgaben“ (IST-ME) luden (Faktorladungen $> .60$). Dies ist nicht überraschend, da beim INKA ausdrücklich auch eine Memorierungsleistung gefordert ist. Der Faktor wird von Heyde (1995) als „Konzentrationsfähigkeit“ interpretiert. Soziodemographische Daten der Stichprobe sowie die Ergebnisse der Faktorenanalyse werden bei Heyde (1995, S. 20f) im Einzelnen aufgeführt.

Mit den standardisierten Testergebnissen derselben Stichprobe (INKA-Standardgesamtwert, doppelt fehlerkorrigierter d2-Gesamtwert, IST-70-Standardwert sowie die Variablen Alter, Geschlecht und Schulbildung) wurden Produkt-Moment-Korrelationen nach Pearson berechnet, die Matrix der Interkorrelationen und der Redundanzen der Variablen ist ebenfalls bei Heyde (1995, S. 22) aufgeführt. Demnach liegen die Korrelationen zwischen INKA und d2 bei $r = .46$, zwischen INKA und IST-ME bei $r = .33$, wobei andere IST-Untertests zum Teil höher mit INKA korrelieren (INKA/IST-Rechenaufgaben: $r = .40$).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Mit dem Inventar Komplexer Aufmerksamkeit liegt ein Konzentrationsstest vor, dessen Aufgaben nach einem neuartigen Prinzip entwi-





ckelt worden sind. Er verspricht, bisherige Mängel von Konzentrationstests zu beheben. Es wird nicht die reine Speedkomponente, sondern das Fähigkeitsniveau überprüft. Dabei sind neben dem Absuchen einer Buchstabenreihe auch ein Dekodierungs- und ein Memorierungsprozess zu leisten. Es handelt sich um ein raschskaliertes Verfahren, das im Großen und Ganzen als ein sehr brauchbares Instrument zur Erfassung von Konzentrationsleistungen erscheint. Weitere Untersuchungen sollten den Überlegenheitsanspruch des Verfahrens gegenüber im Moment noch gebräuchlichen Konzentrationstests untermauern. Fay (1996) nennt u.a. als Schwächen des seiner Ansicht nach „verfrüht veröffentlichten Tests“ (S. 4)

die teilweise nicht einfach zu verstehende Instruktion, den zu geringen Stichprobenumfang für Hochschulabsolventinnen und -absolventen ($n = 41$), sowie die unzureichende Explizierung des zu erfassenden Konstrukts: „Es fehlt an einer Beschreibung, einer näheren Bestimmung, einer Einbettung des Konstrukts: Das Aufzeigen von Ähnlichkeit zu einem Verfahren, dem Test d2, von dem sich INKA gerade abheben soll, ist da nicht die *via regia*“ (S. 65).

Autorinnenkommentar

Das INKA hat den Vorteil, dass sich unterschiedliche Ausprägungen individueller Fertigkeiten nicht auf die Bearbeitungsleistung auswirken und dass sich das Inventar in der Praxis gut bewährt hat. AD(H)S kann mit diesem Test nicht festgestellt werden, vielmehr lässt sich eine Aussage über die Konzentrationsfähigkeit und die Arbeitsspeicherkapazität der Testperson treffen. Bei Anwendung des Verfahrens ist es wichtig, die Motivation der Testperson in die Diagnose miteinzubeziehen, da diese beim INKA eine große Rolle spielen kann.

Literatur

- Amthauer, R. (1973). *Intelligenz-Struktur-Test (I-S-T 70, 4. Auflage)*. Göttingen: Hogrefe.
- Brickenkamp, R. (1987). *Test d2. Aufmerksamkeits-Belastungs-Test (7. Auflage)*. Göttingen: Hogrefe.
- Brickenkamp, R. (1991). *Fehlinterpretationen von Testleistungen? Anmerkungen zum Beitrag „Konzentrationsleistung ohne Konzentration?“*. *Diagnostica*, 37, 52–57.
- Fay, E. (1996). *Tests unter der Lupe. Aktuelle Leistungstests – kritisch betrachtet (Band 1. Testrezension INKA, S. 57–65)*. Heidelberg: Asanger.
- Heyde, G. (1995). *Inventar Komplexer Aufmerksamkeit (INKA). Handanweisung*. Frankfurt am Main: Swets Test Services.

- Heyde, G. (2004). INKA – Inventar Komplexer Aufmerksamkeit. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 133–142). Göttingen: Hogrefe.
- Heyde, G. (2004). INKA. Inventar Komplexer Aufmerksamkeit. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Auflage, S. 409–412). Lengerich: Pabst.
- Öhlschlägel, J. & Moosbrugger, H. (1991). Konzentrationsleistung ohne Konzentration? Zur Schätzung wahrer Leistungswerte im Aufmerksamkeits-Belastungs-Test d2. *Diagnostica*, 37, 42–51.
- Öhlschlägel, J. & Moosbrugger, H. (1991). Überraschende Validitätsprobleme im Aufmerksamkeits-Belastungs-Test d2. *Report Psychologie*, 16 (9), 16–25.
- Petermann, F. (2009). Inventar Komplexer Aufmerksamkeit (INKA). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 217–223). Göttingen: Hogrefe.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Testkurzdarstellung INKA Inventar Komplexer Aufmerksamkeit: S. 297–299)*. Lengerich: Pabst.

3.11. KKA; Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige

Krampen, G. (2007). KKA. Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige. Göttingen: Hogrefe.

Beschreibung

Die KKA erfasst die kurzzeitige selektive Aufmerksamkeits- und Konzentrationsleistung (Leistungsmenge in vorgegebener Zeit) von Vorschul- und Primarschulkindern im Alter von 3 bis 8 Jahren. Es handelt sich bei der KKA um einen Durchstreichtest, bei dem die Kinder die Aufgabe haben, in mehreren Reihen mit verschiedenen Abbildungen die jeweiligen Zielobjekte anzuzeichnen, wobei die Bearbeitungszeit pro Testreihe auf 10 Sekunden festgelegt ist. Zeitnahe oder zeitferne Wiederholungsmessungen sind durch 6 Parallelformen der KKA möglich; ergänzend können Leistungsverläufe bei der Anwendung von KKA-Testserien zu einem Testzeitpunkt (maximal sechs KKA-Parallelformen) und die Anzahl von Verwechslungsfehlern zur Interpretation herangezogen werden.

Mit der Kaseler-Konzentrations-Aufgabe (KKA) für 3- bis 8-Jährige liegt ein Leistungs- und Entwicklungstest zur Erfassung der kurzzeitigen selektiven Aufmerksamkeits- und Konzentrationsleistung bei Kindern der Vorschule und der Primarschule (ersten zwei Klassen) vor. Der Test erfasst ausschließlich die Leistungsmenge kurzzeitiger selektiver Aufmerksamkeitsleistung und eignet sich auch für einen Einsatz im Rahmen einer multimethodalen Psychodiagnostik. Die Anwendungsschwerpunkte liegen abgesehen von der Grundlagen- und anwendungsbezogenen Forschung v.a. in der Klinischen Kinderpsychologie und Pädagogischen Psychologie (z.B. im Kontext der remedialen Entwicklungsdiagnostik, der Frühförderung, der Schulreife diagnostik oder der Evaluation von Fördermaßnahmen). Das Verfahren in Form eines Durchstreichtests eignet sich auch für eine Anwendung in der Erziehungsberatung, der Heimpfychologie und der Pädiatrie (Krampen, 2007, S. 38).

Die Entstehung der KKA geht auf die Ad-hoc-Entwicklung eines möglichst kurzen, ökonomischen Forschungsinstruments zur Abschätzung der kurzzeitigen Konzentrationsleistung bei Vorschulkindern im Rahmen einer Studie zurück, in der die Aufmerksamkeitsbelastung untersucht wurde, die durch die Testung von Kindern mit dem Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 1998, 2002) entsteht. Aus einer ersten Vorversion entstand im Zuge vielfältiger Anwendung, Erprobung (v.a. im universitären Kontext sowie in der Praxis der Klinischen Kinderpsychologie, der Pädiatrie und der Pädagogischen Psychologie), Überarbeitung und gütebezogener Untersuchungen das vorliegende Verfahren. In Bezug auf die Form eines Durchstreichtests diente der Aufmerksamkeits-Belastungs-Test d2 (Brickenkamp, 1962, 2002) als Vorbild. Im Unterschied zum d2 sind allerdings keine Buchstaben, sondern einfache Symbole zur Bearbeitung vorgegeben. Der Test erlaubt die Erfassung der Leistungsmenge, nicht jedoch der Leistungsqualität.

Die Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige besteht aus einem Übungsblatt und dem eigentlichen Testblatt. Auf dem querformatigen Testbogen sind neun Reihen mit Symbolen in Schwarz-weiß-Druck untereinander angeordnet. Die Symbole wurden den „Wingdings“ (Microsoft Word (c) 2000) entnommen.

Sie umfassen:

- die Vorderseite eines Briefumschlages,
- die Rückseite eines Briefumschlages,
- ein aufgeschlagenes Buch,
- eine (Wand-)Uhr,
- eine Glocke und
- eine geöffnete Hand.

Die Anordnung der Symbole in den neun Reihen ist zufällig. Zur optischen Auflockerung und visuellen Strukturierung enthält jede der neun Reihen an zufälliger Stelle zusätzlich das Symbol einer schwarzen Bombe mit Zündschnur. Alle Symbole mit Ausnahme der Bombe sind potentielle Zielsymbole für die Aufgabe des Durchstreichens. Jede Reihe umfasst insgesamt jeweils vier der sechs möglichen Symbole und einmal den visuellen Distraktor Bombe, so dass jede Reihe 25 Symbole enthält und der gesamte Testbogen aus 225 Abbildungen besteht.

Altersgruppe: 3;0–8;0 Jahre

Hinweise zur Durchführung

INKA kann sowohl als **Einzel-** als auch als **Gruppentest** eingesetzt werden.

Mit den insgesamt sechs potenziellen Ziel-Symbolen für die Durchstreichaufgabe bietet das Verfahren sechs Parallelförmigkeiten für Testwiederholungen. Eine Besonderheit liegt darin, dass die Testinstruktionen außer in Deutsch auch in türkischer, russischer, englischer, portugiesischer und luxemburgischer Sprache vorliegen. Somit kann die KKA auch bei Kindern mit einer anderen Primärsprache als der deutschen angewendet werden. Alle Erläuterungen und Anweisungen, die die Testleiterin/der Testleiter in der Übungsphase und während der Durchführung der KKA geben soll, sind im Manual wortwörtlich vorgegeben. Die Instruktionen sind dabei sehr übersichtlich (z.B. durch Hervorhebungen) dargestellt.

Für die Bearbeitung des Testbogens sind feste Zeitumfänge vorgesehen. Jeweils 10 Sekunden stehen dem Kind für die Bearbeitung jeder der neun Symbolzeilen zur Verfügung. Somit ergibt sich eine Phase von 1:30 Minuten für die reine Bearbeitungszeit eines Testbogens. Die Durchführungszeit der KKA richtet sich nach der gewählten Anwendungsvariante: Bei der einmaligen

Durchführung nimmt die KKA bei einer Zeitdauer von ca. 5 Minuten für Instruktion, Beispielübung und 1:30 für die Bearbeitung des Testbogens insgesamt ca. 6:30 Minuten in Anspruch. Bei Durchführung von Testserien verlängert sich die benötigte Zeit entsprechend der Anzahl der durchgeführten Testungen bis zu einer maximalen Dauer von ca. 16:30 Minuten bei der Bearbeitung von sechs Testbögen bzw. aller möglichen sechs Zielsymbole.

Im Testmanual finden sich einige Angaben in Bezug auf Rahmenbedingungen der Testanwendung: Zum einen weist Krampen (2007, S. 38 f.) darauf hin, dass die KKA in einem ruhigen, störungsfreien Raum durchgeführt werden sollte, der dem Kind möglichst wenige Ablenkungsmöglichkeiten bietet. Zudem sind der Größe des Kindes entsprechende Möbel (Tisch und Stuhl) erforderlich. Weiterhin empfiehlt der Autor bis zur zweiten Klassenstufe die Durchführung der KKA im Einzelsetting. Die Testleiterin/der Testleiter sollte möglichst über Eck mit dem Kind am Tisch – bei Rechtshändern auf deren linker Seite und bei Linkshändern auf deren rechter Seite – sitzen.

Hinweise zu Auswertung und Interpretation

Für das Verfahren (KKA-Rohwerte) liegen Normen vor (T-Werte und Prozentrangwerte), die anhand einer umfassenden Stichprobe von $n = 5.314$ Kindern aus dem südwest- und südostdeutschen Raum sowie aus Luxemburg gewonnen wurden. Die Geschlechterverteilung in der Normstichprobe ist nahezu ausgewogen. Für die Gruppen der 3-, 4-, 5- sowie 6- bis 8-Jährigen liegen getrennte Normen vor, wobei jede der Kohorten über 800 Kinder umfasst und die Verteilung der Kinder nach ihrem Alter relativ ausgewogen ist. Auf die Erstellung spezifischer Normen für die Geschlechtergruppen, für Kinder mit verschiedenen Primärsprachen sowie monolinguale und bilinguale Kinder wurde auf Grund der Befunde zur differenziellen Validität verzichtet. Die Normwerte beziehen sich auf Einzeltestungen. Auf Grund des enormen Aufwands wurde auf die Erstellung von Normen für Testserien verzichtet. Der Autor empfiehlt, im Falle der Durchführung von Testserien auf die altersgruppenspezifischen Normwerte für die Einzeldurchführung zurückzugreifen (Krampen, 2007, S. 72). Angaben zum Zeitpunkt der Datenerhebung für die Normierung sind im Manual nicht zu finden.

Die Auswertung eines Testergebnisses erfolgt durch einfaches Auszählen der richtigen Treffer (korrekt durchgestrichene Symbole) pro Reihe und deren Addition zu einem Gesamtwert, dem Rohwert für die Leistungsmenge. Auf einem zusätzlich zur Verfügung stehenden Auswertungsbogen kann dieser Wert notiert und anhand einer Normtabelle in Prozentrangwerte und T-Werte transformiert werden. Zusätzlich schlägt der Autor vor, die Anzahl der Verwechslungsfehler (falsch durchgestrichene Symbole) auszuzählen und auf dem Auswertungsbogen zu vermerken. Neben der Einzelanwendung kann die KKA auch in Form einer Testserie zum Einsatz kommen. Hier werden nacheinander verschiedene Symbole als Durchstreichziele vorgegeben und es werden somit bis zu sechs Durchgänge ermöglicht. Für die Auswertung von Testserien mit der KKA bietet der Auswertungsbogen eine Tabelle, in der die Ergebnisse jedes Durchgangs (Rohwert

für die Leistungsmenge, Prozentrang und T-Werte sowie die Anzahl der Verwechslungsfehler) eingetragen werden können und auch eine graphische Darstellung des Verlaufs der T-Werte und der Menge an Verwechslungsfehlern möglich ist.

Die Auswertung nimmt etwa 5 bis 10 Minuten in Anspruch.

Testgüte

Objektivität

Die Durchführungsobjektivität ist durch die wörtlich im Manual vorgegebene Instruktionsanweisung für die Testleiterin/den Testleiter gegeben. Dies wird auch anhand einer Untersuchung mit $n = 916$ Kindern belegt. Untersuchungen zur Auswertungsobjektivität während der Phase der Testentwicklung ergaben eine hundertprozentige Übereinstimmung zwischen Erst- und Zweitauswerterinnen/-auswertern, was im Sinne einer hohen Auswertungsobjektivität der KKA gewertet werden kann. Im Fall des Verdachts von Flüchtigkeitsfehlern beim Auszählen der richtig angekreuzten Symbole empfiehlt Krampen (2007, S. 56), eine zweite Person zur Auswertung hinzuzuziehen. Auch die Interpretationsobjektivität kann – bei engem Bezug der Diagnostikerin/des Diagnostikers zu den Normwerten, den Vertrauensintervallen und kritischen Differenzen – als gegeben angesehen werden.

Reliabilität

Für die KKA liegen Angaben zur internen Konsistenz, zur Retestreliabilität, zur Paralleltestreliabilität, zur Reliabilität von Testserien sowie Itemkennwerte (Schwierigkeitsindizes und Trennschärfekoeffizienten) vor, die auf unterschiedlichen, überwiegend umfangreichen Stichproben der Erprobungsphase basieren.

Die Kennwerte für die interne Konsistenz (Cronbachs α) und die Paralleltestreliabilität sind als sehr hoch einzustufen (bis $\alpha = .98$) und entsprechen dem Befundmuster zur Reliabilität ähnlicher Skalen für (ältere) Kinder, Jugendliche und Erwachsene (Krampen, 2007, S. 56). Weder in Bezug auf das Alter noch auf die Primärsprache zeigten sich signifikante Unterschiede in den Koeffizienten der internen Konsistenz und der Testwiederholungsreliabilität. Besonders der Befund zum Alter ist positiv im Sinne des Anspruchs eines Entwicklungstests zu werten. Die Paralleltestreliabilität wurde dreifach überprüft: Die summarische Bestimmung über alle sechs Paralleltestvarianten



bestätigte die Paralleltestreliabilität für die gesamte Altersgruppe der 3- bis 8-Jährigen und auch getrennt für die spezifischen Altersgruppen. Eine getrennte Ermittlung der Reliabilität der sechs verschiedenen KKA-Parallelformen erbrachte hohe interne Konsistenzen. Die Testung aller möglichen 15 Kombinationen der sechs Parallelformen in kurzen Zweierreihen an kleineren Stichproben ($79 < n < 95$) mit jeweils zwei verschiedenen Zielsymbolen kann ebenfalls als Bestätigung der Paralleltestreliabilität aller sechs KKA-Varianten gewertet werden (Krampen, 2007, S. 57 f.). Für KKA-Testserien liegen Befunde zur internen Konsistenz vor, in die die Skalenwerte der einzelnen Parallelformen als Items eingehen. Die Höhe der internen Konsistenzen geht mit zunehmender Anzahl der durchgeführten Parallelformen leicht zurück, bleibt aber in einem für Testserien akzeptablen Bereich ($\alpha = .78$ bei einer Sechser-Testserie).

Validität

Im Testmanual werden Ergebnisse in Bezug auf die konvergente und diskriminante Validität, zur differenziellen Validität sowie zur Änderungssensitivität der KKA dargestellt.

Für die diskriminante und konvergente Validität liegen umfangreiche Befunde über Zusammenhänge mit verschiedenen Maßen vor, die anschaulich im Manual dargestellt werden. Im Folgenden seien einige Ergebnisse exemplarisch wiedergegeben:

- (1) Erwartungsgemäß hohe Zusammenhänge zeigten sich mit anderen Konzentrationstests (FTF, d2, KLT und Pauli-Test), insbesondere mit anderen Konzentrations-Durchstreichtests.
- (2) Mit Maßen der Fremdeinschätzung von Konzentrationsleistungen finden sich statistisch bedeutsame Beziehungen bei Einschätzungen durch Kindergarten-Erziehungskräfte und Primarschul-Lehrkräfte, nicht jedoch für Eltern, was der Autor im Sinne geringerer entwicklungsdiagnostischer Fähigkeiten von Eltern für die eigenen Kinder wertet (Krampen, 2007, S. 59 ff.).
- (3) Schwächere Zusammenhänge zeigen sich in Bezug auf Indikatoren der Sprachentwicklung und der allgemeinen Intelligenz. Das Befundmuster bestätigt laut Krampen (2007, S. 61) den Anspruch der KKA, „kurzzeitige selektive Aufmerksamkeits- und Konzentrationsleistungen bei Vorschulkindern und Primarschulkindern der ersten und zweiten Klasse hinreichend unabhängig von ihrem Sprachentwicklungsstand und ihrer Intelligenz zu erfassen“.

In Bezug auf die differenzielle Validität des Verfahrens liegen Befunde aus einer Reihe von Subgruppenvergleichen vor (Geschlechtervergleich, klinische Gruppen und Vergleiche in Bezug auf die Primärsprache und den familiären Hintergrund). Hier sollen die Befunde zur Differenzierungsfähigkeit der KKA zwischen den verschiedenen Altersgruppen referiert werden, da diese für das Verfahren als bereichsspezifischer Entwicklungstest von entscheidender Bedeutung sind. Im Kontext zahlreicher Querschnittsbefunde konnte die differenzielle Validität für die Altersgruppen der 3-, 4-, 5-, und der 6- bis 8-Jährigen bestätigt werden. Eine Differenzierung der KKA zwischen den Gruppen der 6-, 7-, und 8-Jährigen konnte nicht bestätigt werden, daher wurden diese Grup-

pen bei der Normierung zusammengefasst. Darüber hinaus liegen auch Längsschnittdaten vor, die durch die viermalige (im Abstand von jeweils einem Jahr) durchgeführte Testung von $n = 156$ Dreijährigen, $n = 161$ Vierjährigen und $n = 154$ Fünfjährigen gewonnen wurden. Für die ersten beiden Gruppen führten Varianzanalysen zu statistisch bedeutsamen Befunden zur Altersdifferenzierung. Für die dritte Kohorte ergab sich lediglich ein Unterschied zwischen dem Alter von fünf und dem von sechs bis acht Jahren. Zudem weisen die Längsschnittdaten auf eine relativ hohe positionale Stabilität der KKA-Ergebnisse hin (Krampen, 2007, S. 63).

Die Änderungssensitivität der KKA in Bezug auf die Abbildung externer Einflüsse wurde einerseits in Bezug auf die Effekte möglicher Belastungen (z.B. Anwesenheit anderer Kinder, Lärm, Tageszeit) und zum anderen in Hinblick auf die Effekte systematischer Fördermaßnahmen untersucht. Dabei zeigte sich, dass die Durchführung im Einzelsetting – auf Grund statistisch signifikant schlechterer Leistungen bei Anwesenheit eines weiteren zuschauenden Kindes als sozialer Stressor – vorzuziehen ist. Weiterhin fanden sich signifikante Unterschiede der Testleistungen in Bezug auf die Tageszeit zwischen dem Vormittag und Testung unmittelbar nach dem Mittagessen bei $n = 120$ Vorschulkindern. Auch konnten leichte, aber statistisch signifikante Effekte (Reduktion der Testleistung im KKA) durch die Teilnahme an einer zeitlich aufwändigeren psychodiagnostischen Testung (z.B. WET oder HAWIE) nachgewiesen werden. Die Änderungssensitivität für Effekte systematischer Fördermaßnahmen ist bislang nur exemplarisch untersucht worden. Erste Ergebnisse weisen auf eine Änderungssensitivität der KKA für Interventionen, wie regelmäßiges Bilderbuchlesen mit dem Kind, regelmäßige Spiele (Regelspiele wie Mikado oder Memory) mit erwachsenem Spielpartner und die Vermittlung der Grundübungen des Autogenen Trainings hin. Erste positive Befunde liegen auch für das Marburger Konzentrations-training für Kindergarten- und Vorschulkind vor (MKT; Krowatschek, Albrecht & Krowatschek, 2004), die jedoch auf Grund der sehr geringen Stichprobengröße von $n = 12$ äußerst vorsichtig interpretiert werden sollten.

Kommentar der Psyindex-Testreviewerinnen und -reviewer

Mit der Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige (KKA) liegt ein ökonomisches Verfahren zur Erfassung der kurzzeitigen, selektiven Konzentrations- und Aufmerksamkeitsleistung von Vorschulkindern und Kindern der ersten beiden Klassen der Primarschule vor. Die KKA zeichnet sich insbesondere durch eine moderne Gestaltung aus, indem aktuelle, zeitgemäße Symbole als Ziele für die Durchstreichaufgabe gewählt wurden. Eine weitere positiv zu wertende Besonderheit besteht darin, dass die Durchführungsanweisung für die Testleiterin/den Testleiter und die Instruktion neben dem Deutschen in weiteren Sprachen vorliegen. Hierdurch eröffnet sich die Möglichkeit für einen breiten Einsatz des Verfahrens in vielfältigen Kontexten. Wird der Empfehlung des Autors Folge geleistet und die Diagnostik auf die Erhebung der Leistungsmenge beschränkt, kann die KKA im Rahmen einer multidimensionalen Diagnostik, in der zusätzlich Q-, L- und L'-Daten erhoben werden, als ein valider, reliabler und v.a. ökonomischer und fairer diagnostischer Baustein eingestuft werden.



Autorinnenkommentar

Die KKA ähnelt der DL-KG und der DL-KE. Sie ist zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht es sehr gut, die fokussierte und die selektive Aufmerksamkeit bei Kindern zu überprüfen und sollte in Kombination mit weiteren Verfahren zur Überprüfung der Vigilanz und der räumlichen Aufmerksamkeit angewendet werden, sofern ein umfassendes Bild der Aufmerksamkeitsleistung gewünscht wird.

Literatur

- Brickenkamp, R. (1962). *Test d2. Aufmerksamkeits-Belastungs-Test*. Göttingen: Hogrefe.
- Brickenkamp, R. (2002). *Test d2. Aufmerksamkeits-Belastungs-Test (9., vollständig überarbeitete und neu normierte Auflage)*. Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (1998). *Der Wiener Entwicklungstest (WET). Ein allgemeines Entwicklungstestverfahren für Kinder von 3 bis 6 Jahren*. Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (2002). *WET. Wiener Entwicklungstest. Ein Verfahren zur Erfassung des allgemeinen Entwicklungsstandes bei Kindern von 3 bis 6 Jahren (2., überarbeitete und neu normierte Auflage)*. Göttingen: Hogrefe.
- Krampen, G. (2007). *KKA. Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige*. Göttingen: Hogrefe.
- Krampen, G. (2008). *Kognitive Entwicklung bei 3- bis 8-Jährigen. Konzentrationsleistung und Übergang vom vor-operatorischen zum konkret-operatorischen Denken*. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40 (2), 79–86.
- Krampen, G. (2009). *Entwicklungsförderung bei Kindergartenkindern mit Konzentrationsschwächen: Vergleichende Untersuchungen unter Einbezug von Entspannungsverfahren*. *Entspannungsverfahren*, 26 (26), 8–25.
- Krowatschek, D., Albrecht, S. & Krowatschek, G. (2004). *Marburger Konzentrationstraining für Kindergarten- und Vorschul Kinder (MKT)*. Göttingen: Hogrefe.

3.12. KHV-VK; Konzentrations-Handlungsverfahren für Vorschulkinder

Ettrich, K.U. & Ettrich, C. (2005). KHV-VK. Konzentrations-Handlungsverfahren für Vorschulkinder. Göttingen: Hogrefe.

Beschreibung

Die Konzentrationsfähigkeit eines Kindes stellt eine Grundbedingung für das Lernen und das Erbringen von schulischen Leistungen dar. Mögliche Defizite sollten daher schon im Vorschulalter festgestellt werden. Zu diesem Zweck wurde das Konzentrations-Handlungsverfahren für Vorschulkinder (KHV-VK) entwickelt. Als kindgemäßes Sortierverfahren, das die Parameter Fehler (Sorgfaltsleistung) und Zeit (Arbeitstempo) erfasst, ermöglicht es Aussagen zum Niveau konzentrativer Fähigkeiten und stellt auch bei der Diagnostik von AD(H)S im Vorschulalter eine wichtige Hilfe dar. Die Aufgabe der Kinder besteht im Sortieren von 44 Karten mit gegenständlichen Abbildungen nach bestimmten Merkmalen. Für altersgerecht entwickelte Kinder wurde ein „Vierer-Sort“ (für das Sortieren nach vier Merkmalen) und für entwicklungsbeeinträchtigte Kinder ein „Zweier-Sort“ (zwei Merkmale) gestaltet.

Das KHV-VK basiert auf der Unterscheidung zwischen unwillkürlicher und willkürlicher Aufmerksamkeit und Konzentrationsfähigkeit. Diese Differenzierung ist gut geeignet, um die Entwicklung von Kindern zu markieren: Bei sehr jungen Kindern dominiert die unwillkürliche Aufmerksamkeit, die an die Attraktivität des Reizes und seine individuelle Nützlichkeit gebunden ist. Mit zunehmendem Alter gelingt der Wechsel zur willkürlichen Aufmerksamkeit immer besser. Kinder können sich interessanten Reizen bewusst zuwenden. Konzentrationsfähigkeit setzt ferner die Bereitschaft des Kindes voraus, Fremdziele zu übernehmen und Aufgaben auch zu Ende zu führen. Als Material, um diesen Entwicklungsstand zu testen, verwendet das KHV-VK 44 Karten, auf denen jeweils 12 Figuren dargestellt sind. Es handelt sich um einen Sortieretest in der Tradition von Abels (1954): Die Karten müssen von den Kindern nach bestimmten Merkmalen in eine Box mit vier Fächern einsortiert werden. Die kritischen Merkmale sind in Form A Kamm und Baum und in Form B Blume und Bürste. Es kann die „Zweiersortierung“ oder die „Vierersortierung“ gewählt werden. Bei der „Vierersortierung“ sind die Karten der Form A nach den Merkmalen Baum, Kamm, Baum und Kamm, weder Baum noch Kamm zu sortieren, in der Form B sinngemäß nach Blume und Bürste. Im Gegensatz zum KHV gibt es hier kein Täuschungsfach, d.h. alle Sortiermerkmale sind auch tatsächlich im Kartensatz vorhanden.

Die Instruktion lautet: „Die Karten sollst Du jetzt in diese vier Kästchen einsortieren. Auf den Karten siehst Du viele Bilder (Karte zeigen und Bilder benennen lassen). Auf einige Bilder musst Du beim Einsortieren besonders gut achten. Es sind der Baum (zeigen) und der Kamm (zeigen). Findest Du auf der Karte einen Baum, dann legst Du sie in dieses Fach (zeigen)...“.

Die ersten vier Karten werden mit dem Kind gemeinsam abgelegt, Fehler werden korrigiert und ggf. das Zuordnungsprinzip neu erklärt. Die Arbeitszeit beträgt 10 Minuten.

Altersgruppe: 3;0–6;11 Jahre

Hinweise zur Durchführung

Es handelt sich um einen **Einzeltest**.

Die gesamte Durchführung (Instruktion und Arbeitszeit) dauert zwischen 15 und 20 Minuten und kann von jeder geschulten Helferson durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Das Verfahren wurde an $n = 1.887$ Kindern normiert. Es liegen Stanine-Werte für vier verschiedene Altersgruppen vor.

Testgüte¹⁰

Objektivität

Die Durchführungsobjektivität ist durch die standardisierte Anweisung gegeben.

Reliabilität

Die Reliabilität des Verfahrens wurde mit Hilfe der Odd-Even-, der Retest- und der Paralleltest-Methode nachgewiesen und liegt für den Zeit-Wert im Bereich von $r = .83$ bis $r = .89$ und für den Fehler-Wert zwischen $r = .73$ und $r = .83$.

Validität

Die Gültigkeit des KHV-VK wurde durch den Vergleich von Kindern mit unterschiedlichem Entwicklungsniveau (normal entwickelt bis deutlich entwicklungsgestört) sowie aus der Analyse der altersbezogenen Leistungen hinsichtlich Zeit- und Fehlerwerten erschlossen. Zusätzlich wurde die eigenständige Indikationsleistung des Tests durch Korrelationen mit anderen kognitiven Leistungen ermittelt.

¹⁰ vgl. www.testzentrale.de/programm/konzentrations-handlungsverfahren-fur-vorschulkinder.html (Abfragedatum: 27.12.2012)

Autorinnenkommentar

Das Verfahren ist eines der wenigen, das sich für Kinder im Vorschulalter eignet. Die Kinder haben erfahrungsgemäß Freude an der Durchführung und verstehen die Instruktion gut. Defizite in der Aufmerksamkeitsleistung können somit schon im Vorschulalter festgestellt werden wodurch frühzeitig Fördermaßnahmen gesetzt werden können. Für eine umfassende AD(H)S-Diagnostik eignet sich der Test nicht. Hierzu braucht man unbedingt Informationen von Kindergartenpädagoginnen und -pädagogen, Eltern und eine diagnostische Verhaltensbeobachtung.

Literatur

Abels, D. (1954). *Konzentrations-Verlaufs-Test K-V-T*. Göttingen: Hogrefe.
Ettrich, K. U. & Ettrich, C. (2005). *KHV-VK. Konzentrations-Handlungsverfahren für Vorschulkinder*. Göttingen: Hogrefe.



3.13. KLT-R; Konzentrations-Leistungs-Test – Revidierte Fassung

Düker, H., Lienert, G. A., Lukesch, H. & Mayrhofer, S. (2001). KLT-R. Konzentrations-Leistungs-Test – Revidierte Fassung. Göttingen: Hogrefe.

Beschreibung

Der KLT-R misst die individuelle Konzentrationsfähigkeit. Im Unterschied zu Tests der Kurzzeitanspannung wird hier die Langzeitanspannung gemessen. Erfasst werden sowohl die Quantität als auch die Qualität der Dauerbeanspruchung sowie der Leistungsverlauf einer Testperson. Die leichtere Form des KLT-R (KLT-R 4–6) ist für die Anwendung bei Schülerinnen und Schülern der 4.–6. Schulstufe geeignet. Zusätzliche Einsatzmöglichkeiten dieser Form liegen im Bereich der Förderschule. Die schwierigere Form (KLT-R 6–13) ist in der 6.–13. Schulstufe und bei Erwachsenen anwendbar.

Der Revidierte Konzentrations-Leistungs-Test (KLT-R) ist die revidierte und neu normierte Version des Konzentrations-Leistungs-Tests (KLT; Düker & Lienert, 1965). Der KLT zählt zu den objektiven Leistungstests; er geht auf ein zuerst von Düker (1943, 1949) beschriebenes Verfahren zur Feststellung der allgemeinen psychischen Leistungsfähigkeit zurück. Mit dem Test wird der Anspruch erhoben, Konzentrationsfähigkeit im Sinne von Belastbarkeit, Ausdauer und Ermüdungsresistenz zu messen, wobei die Testergebnisse nur wenig von intellektuellen Voraussetzungen abhängen sollen. Düker (1949, S. 10) selbst baute das Verfahren auf dem Konzept der Koordination auf, die er verstand als das „zu einer Gesamttätigkeit geordnete Zusammenwirken der Einzeltätigkeiten, die zur Erreichung eines bestimmten Zweckes erforderlich sind“. Koordination erfordere „eine psychische Anpassung, deren Intensität von der Schwierigkeit der auszuführenden Leistung abhängig“ sei (S. 11), wobei hier die Fähigkeit zur Anpassung zum Zwecke der Koordination als Konzentrationsfähigkeit bezeichnet wird. Auf dieser Sichtweise von Koordination basieren auch die Anforderungen, die der KLT bzw. KLT-R an die Testprobandinnen/-probanden stellt: Das Lösen der KLT-Aufgaben macht das geordnete Zusammenwirken der Einzeltätigkeiten „Auffassen“, „Rechnen“, „Merken“, „Regelabruf“ und „Entscheiden“ zu einer Gesamttätigkeit erforderlich. In dem ursprünglichen Verfahren ist sowohl die Messung der Leistungsmenge als auch der Leistungsgüte vorgesehen. „Beide sind gleichermaßen Ausdruck der Konzentrationsfähigkeit mit dem Unterschied, dass die Leistungsmenge mehr der Antriebsfunktion, die Leistungsgüte mehr der Kontrollfunktion entspricht“ (Düker & Lienert, 1965, S. 4). Im Unterschied zu Tests für Kurzzeitanspannung (z.B. der Test d2 von Brickenkamp, 1994) soll der KLT Langzeitanspannung erfassen. Mit der revidierten Form KLT-R werden ebenfalls wie beim KLT Quantität und Qualität dieser Art von Dauerbeanspruchung diagnostisch erfasst, aber zusätzlich Indikatoren für den Leistungsverlauf gebildet.

Der KLT-R liegt in zwei unterschiedlichen Schwierigkeitsstufen vor, welche aus je neun Blöcken mit jeweils 20 Rechenaufgaben bestehen. Pro Aufgabenblock sind 2 Minuten Bearbeitungszeit

vorgesehen. Das Verfahren besteht aus komplexen Rechenaufgaben, die auch Kurzzeitspeicherleistungen einbeziehen. Jedes Item besteht aus der Addition und Subtraktion von zwei mal drei einstelligen Zahlen; die Zwischenergebnisse müssen gemerkt werden; sodann wird ein Gesamtergebnis nach unterschiedlichen Regeln errechnet. Für die Bestimmung des Wertes für jedes Item ist entweder die Anwendung einer Regel (leichtere Form: KLT-R 4–6) oder von zwei Regeln (schwierigere Form: KLT-R 6–13) vorgesehen.

Bei der leichteren Form (KLT-R 4–6) wird nach Berechnung der beiden Zwischenergebnisse das Endergebnis durch die Subtraktion des niedrigeren von dem größeren Zwischenergebnis berechnet.

Bei der schwierigeren Form (KLT-R 6–13) ist das Endergebnis jeder Rechenaufgabe aufgrund der Beachtung von zwei Regeln zu bilden: Falls das erste Zwischenergebnis größer ist als das zweite, ist die Differenz aus beiden zu bilden, wenn jedoch das erste Ergebnis kleiner als das zweite ist, sind beide zu addieren.

Für den KLT-R werden folgende Leistungsindikatoren gebildet:

- Gesamtleistungsmenge, definiert als die Anzahl der richtig und falsch bearbeiteten Aufgaben ($GL = RWR + RWF$)
- Anzahl richtig bearbeiteter Items ($RWR = \text{Rohwert richtig gelöster Aufgaben}$)
- Anzahl falsch bearbeiteter Items ($RWF = \text{Rohwert falsch gelöster Aufgaben}$)
- Fehlerprozent ($F \% = [RWF] \times 100 / [RWR + RWF]$)
- Schwankungsbreite ($SBR = \text{Differenz aus den richtigen Lösungen aus dem Aufgabenblock mit der größten und der geringsten Leistung}$); dieser Wert wird nur für die richtig gelösten Aufgaben berechnet, da die Differenzen in Bezug auf die Fehlerwerte wesentlich geringer sind und die Schwankungsbreite bezüglich der Gesamtleistung im Wesentlichen durch die Anzahl der richtig gelösten Aufgaben bestimmt wird.
- Für die Interpretation der Testergebnisse ist es zusätzlich noch hilfreich, eine Verteilung der gemachten Fehler auf die drei Testdrittel vorzunehmen. Diese Verteilung kann der qualitativen Ergänzung der Angaben zur Leistungskonstanz dienen.

Zu Beginn werden einige persönliche Daten erfasst (Geburtsdatum, Schulart, Beruf/Klasse und Geschlecht). Zudem werden auf dem Deckblatt die für die Aufgabenbearbeitung anzuwendenden Regeln an Beispielen erklärt (die Anweisungen sind für die Testformen A und B identisch, jedoch unterscheiden sich die Instruktionen der beiden Schwierigkeitsstufen).

Tab. 3: Itembeispiel für den KLT-R 4–6

| | |
|---------------|-------|
| obere Zeile: | 5-2+6 |
| untere Zeile: | 4+7+5 |

Instruktion:

Zuerst die obere Zeile ausrechnen (Tab. 3) und sich das Ergebnis merken (= 9). Genauso wird mit der unteren Zeile verfahren (= 16). Beide Zahlen im Kopf behalten! Regel: Stets die kleinere Zahl von der größeren abziehen (16-9) und das Ergebnis eintragen (= 7).

Tab. 4: Itembeispiele für den KLT-R 6–13

| | Beispiel A | Beispiel B |
|---------------|------------|------------|
| obere Zeile: | $8-3+7$ | $8+5-7$ |
| untere Zeile: | $5+6-3$ | $4+9-5$ |

Instruktion:

Zuerst die obere Zeile ausrechnen und sich das Ergebnis merken (Beispiel A = 12). Genauso wird mit der unteren Zeile verfahren (Beispiel A = 8). Beide Zahlen im Kopf behalten!

1. Regel: Ist die obere Zahl größer (wie im Beispiel A), dann die untere von der oberen abziehen (12-8) und das Ergebnis (= 4) eintragen.
2. Regel: Ist die obere Zahl kleiner (wie im Beispiel B), dann die untere zu der oberen hinzuzählen (6+8) und das Ergebnis (= 14) eintragen.

Altersgruppe: 4.–6. Schulstufe (KLT-R 4–6) bzw. 6.–13. Schulstufe (KLT-R 6–13)

Hinweise zur Durchführung

Der KLT-R kann als **Einzel- oder Gruppentest** durchgeführt werden.

Für jede Schwierigkeitsstufe sind zwei echte Parallelversionen vorhanden.

Die reine Bearbeitungszeit beträgt 18 Minuten (2 Minuten pro Aufgabenblock). Zusätzlich müssen ca. 10 Minuten für Einführung und Instruktion veranschlagt werden.

Die Instruktionen zum KLT-R sind wortwörtlich auf den Testformularen auf der ersten Seite abgedruckt. Das Manual enthält zudem weitere Testanweisungen für die Testleiterin/den Testleiter.

Aufgrund der maximierten Objektivität kann das Verfahren nicht nur von Psychologinnen/Psychologen, sondern von jeder Fachkraft (z.B. Beratungslehrer/innen), die mit den Prinzipien der

Anwendung und Auswertung von Tests vertraut ist, durchgeführt werden. Vor der Durchführung zu diagnostischen Zwecken sind ein Selbstversuch sowie Probedurchführungen zu empfehlen; die genaue Einhaltung der Testanweisungen wird vorausgesetzt.

Hinweise zu Auswertung und Interpretation

Zur Ausarbeitung von Normen für den KLT-R 4–6 konnte eine Normierungsstichprobe von $n = 860$ bayerischen Schülerinnen und Schülern, für den KLT-R 6–13 eine Stichprobe von $n = 2.600$ erreicht werden. Bei keinem der quantitativen Testindikatoren des KLT-R 4–6 oder des KLT-R 6–13 ist ein Mittelwertsunterschied zwischen den beiden Testformen nachzuweisen. Für die Normierung bedeutet dies, dass sie für beide Parallelversionen gemeinsam vorgenommen werden kann. Ebenso sind in keinem einzigen Fall geschlechtsspezifische Unterschiede nachweisbar. Eine geschlechtsbezogene Normierung ist daher nicht notwendig. Es zeigen sich jedoch deutliche schulstufenspezifische und schulartspezifische Leistungsunterschiede.

Für den KLT-R 4–6 wurden Normtabellen nach folgenden Kriterien ausgearbeitet:

1. Schulstufenbezogene Normen für die 4., 5. und 6. Schulstufe. Bei der schulstufenbezogenen Normberechnung für die 5. und 6. Schulstufe wurden die Anteile nach Schularten entsprechend der Aufteilung der Schüler/innen in bayerischen Schulen gewichtet.
2. Für Schüler/innen der 5. und 6. Schulstufe stehen auch schulartspezifische Normwerte zur Verfügung.

Für den KLT-R 6–13 stehen folgende Normtabellen zur Verfügung:

1. Schulstufenbezogene Normen für die 6.–7., 8.–9., 10. sowie 11.–13. Schulstufe. Bei der schulstufenbezogenen Normberechnung wurden die Anteile nach Schularten wieder entsprechend der Aufteilung der Schüler/innen in den bayerischen Schulen gewichtet.
2. Für Schüler/innen der 6.–7., 8.–9., 10. sowie 11.–13. Schulstufe können auch schulartspezifische Normwerte ermittelt werden.
3. Zudem sind noch vorläufige Normen für Erwachsene (Personen älter als 18 Jahre) vorhanden.



Es ist ökonomisch, mit der Auszählung der richtigen Lösungen pro Aufgabenblock zu beginnen. Diese Werte werden in die dafür vorgesehenen Kästchen unter die Aufgabenblöcke eingetragen. Sodann wird darüber die Summe gebildet; dieser Wert ist in das Auswertungsblatt unter RWR (= Rohwert richtige) einzutragen. Analog verfährt man mit den gemachten Fehlern. Der Fehlerwert ist in das Auswertungsblatt unter RWF (= Rohwert falsche) einzutragen. Aufgabenauslassungen werden nicht mitgezählt. Die Summe der falsch bearbeiteten Aufgaben aus jeweils den ersten drei, den nachfolgenden drei und den letzten drei Blöcken wird noch unter Fehlerverteilung in dem Auswertungsblatt vermerkt. Die Gesamtleistungsmenge (GL) wird als die Summe der richtigen und falschen Lösungen bestimmt und ebenfalls in das Auswertungsblatt eingetragen. Der Fehler-Prozentwert (F %) wird nach der angegebenen Formel ($F \% = RWF \times 100 / GL$) berechnet. Die Schwankungsbreite wird als Differenz der höchsten Anzahl der richtigen Lösungen (RWR) abzüglich der niedrigsten Anzahl der richtigen Lösungen pro Block bestimmt. Der Leistungsverlauf kann zudem auf dem Auswertungsblatt visualisiert werden. In einem nächsten Schritt werden für die einzelnen Indikatoren (mit Ausnahme der Fehlerverteilung) Normwerte bestimmt.

Die Auszählung der richtigen und falschen Lösungen erfolgt mit Hilfe der Lösungsschablonen. Die Normwerte können in den entsprechenden Tabellen des Tabellenanhangs abgelesen werden. Es stehen Prozentrangnormen und Standardnormen zur Verfügung.

Die Auswertungszeit beträgt ca. 10 Minuten.

Testgüte

Objektivität

Die Testdurchführung ist voll standardisiert, damit ist in Bezug auf die Person der Testleiterin/des Testleiters Durchführungsobjektivität gesichert. Aufgrund der Auswertung mit Schablonen ist die Auswertungsobjektivität gewährleistet und durch das Vorliegen von Normtabellen ist auch die Interpretationsobjektivität gegeben.

Reliabilität

Die neue Testform des KLT-R erlaubt es, die Reliabilität des Verfahrens mittels Koeffizienten der internen Konsistenz zu schätzen. Sowohl für den KLT-R 4–6 als auch für den KLT-R 6–13 weisen die Homogenitätskoeffizienten der Eichstichprobe das Verfahren bezüglich der richtig gelösten Aufgaben pro Block als hoch reliabel aus ($.94 \leq r_{tt} \leq .97$). Etwas niedrigere, aber immer noch hinreichende Homogenitäten sind hinsichtlich der falschen Lösungen pro Block gegeben ($.79 \leq r_{tt} \leq .93$).

Validität

Es liegen verschiedene Untersuchungen zur konvergenten Validität vor. Zwischen den Indikatoren der Leistungsmenge aus dem d2 (GZ bzw. GZ-F) und dem KLT-R bestehen relativ hohe signifikante Korrelationen von $r = .46$ bis $r = .57$ (Marsch, 2000).

An einer Stichprobe von 127 Realschülerinnen/-schülern (Hofgärtner, 1998, S. 59) wurde zwischen der Leistungsmenge des KLT-R 6–13 und dem GZ-F-Wert aus dem d2 eine signifikante Korrelation von $r = .27$ gefunden; zwischen den Fehlermaßen betrug der Zusammenhang sogar $r = .37$. Mit den Subtests 9 und 10 aus dem Prüfsystem für Schul- und Bildungsberatung (PSB; Horn, 1969) wird ebenfalls versucht, die Konzentrationsleistung zu erfassen. Die korrelativen Beziehungen zwischen der Mengenleistung aus dem KLT-R 6–13 und dem PSB 9 variierten zwischen $r = .51$ und $r = .58$, die Korrelationen zum PSB 10 lagen zwischen $r = .25$ und $r = .79$ (Bonkowsky, 1999, S. 117; Gottwald, Knan & Marsch, 1997, S. 19 f.).

Die divergente Validität des KLT-R wurde anhand des Rechentests 4–6 (RT 4–6; Lobeck, Frei & Blöching, 1990) und des PSB überprüft. Beispielsweise wurden in zwei Studien (Bonkowsky, 1999; Hofgärtner, 1998) die Reasoningaufgaben des PSB als Intelligenzindikator verwendet. Bezüglich der Leistungsmenge und der Leistungsgüte des KLT-R ergaben sich signifikante Korrelationen. Von einer Unabhängigkeit der Intelligenz- und Konzentrationsleistungen kann also nicht ausgegangen werden. Die Fähigkeit zu längerer Konzentration scheint vielmehr eine Vorbedingung für die erfolgreiche Bewältigung von Intelligenzaufgaben zu sein. Zudem resultieren aus Gemeinsamkeiten bei der Testdurchführung (v.a. die Durchführung unter Zeitvorgaben) weitere überlappende Varianzanteile.

Als externe Validierungskriterien wurden verschiedene Schulleistungsindikatoren herangezogen. Beispielsweise wurden von Burgstaller (2001) an einer Stichprobe von Gymnasiastinnen/Gymnasiasten der 9. und 10. Schulstufe ($n = 86$) Korrelationen zwischen den Werten aus dem KLT-R 6–13 und Schulnoten aus dem Jahreszeugnis berechnet. Obwohl sich erwartungsgemäße Beziehungen andeuten, ließ sich hier allerdings nur ein Zusammenhang zwischen einem hohen F %-Wert und schlechten Mathematiknoten statistisch absichern.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Mit dem KLT-R liegt ein vielversprechendes Verfahren zur Messung der individuellen Konzentrationsfähigkeit vor. Neben Indikatoren der Leistungsmenge und der Leistungsqualität werden mit dem KLT-R auch Indikatoren des Leistungsverlaufs erhoben. Der Test kann in vielen Bereichen der Psychologie eingesetzt werden (z.B. in der Schulpsychologie, Betriebspsychologie, Verkehrspsychologie oder der Pharmakopsychologie). Die Ergebnisse zur Reliabilität (Cronbachs α) des KLT-R sind als hoch bis hinreichend einzustufen. Brauchbare Ergebnisse für die Retestreliabilität liefert der alte KLT, für den KLT-R wurden noch keine Testwiederholungsstudien durchgeführt. Es liegen erste Untersuchungen zur konvergenten, divergenten und kriteriumsbezogenen Validität vor, jedoch sind noch weitere Validierungsstudien nötig. Für die Anwendung sind nach Schulstufen und Schularten gut differenzierte und übersichtliche Normtabellen entwickelt worden. Dennoch ist die Normierungsfrage noch nicht vollständig geklärt. Die Normierungsstichprobe stammt aus Bayern, daher wären Vergleichsstudien mit Stichproben aus anderen Bundesländern anzustreben. Da sich zur Zeit der Normierung die 5. und 6.



Klassenstufe der Realschule noch in einer Übergangsphase befand, sind die Normierungsstichproben für diese Klassen zu klein. Zudem sind für den Bereich der Erwachsenen bisher nur vorläufige Normen vorhanden. Es wurde hierbei auf die untersuchten Schüler/innen-Stichproben zurückgegriffen und damit auf Probandinnen/Probanden mit einem eingeschränkten Altersspektrum. Für Erwachsene sind noch umfassende Stichproben zu erheben. Da Konzentrationsfähigkeit immer anhand eines bestimmten Materials erfasst werden muss, geht der gewandte Umgang mit diesem Material auch in die Leistung eines Konzentrationstests ein. Es wäre also denkbar, dass Probandinnen/Probanden mit einer Rechen-

schwäche im KLT-R eventuell benachteiligt sind. Diese Frage müsste jedoch noch genauer überprüft werden.

Autorinnenkommentar

Der KLT eignet sich hervorragend zur Überprüfung der Konzentrationsfähigkeit bei Kindern und Jugendlichen, die mit dem Rechnen keine Probleme haben, nach Möglichkeit gute Kopfrechner/innen sind und bei welchen auch kein Motivationsdefizit bzw. keine unzureichende Arbeitseinstellung vorliegt. Weiters ist die Testleistung nicht unabhängig von der Arbeitsspeicherkapazität. Zur Feststellung von AD(H)S ist dieser Test eher ungeeignet.

Literatur

- Bonkowsky, D. (1999). *Intelligenz – familiäre und schulische Aspekte. Unveröffentlichte Zulassungsarbeit, Universität Regensburg.*
- Brickenkamp, R. (1994). *Test d2. Aufmerksamkeits-Belastungs-Test (8., erweiterte und neu gestaltete Auflage). Göttingen: Hogrefe.*
- Burgstaller, S. (2001). *Validierungsstudien zum PSB-R 6–13. Unveröffentlichte Diplomarbeit, Universität Regensburg.*
- Düker, H. (1943). *Psychopharmakologische Untersuchungen über die Wirkung von Keimdrüsenhormonen auf die geistige Leistungsfähigkeit. Archiv für experimentelle Pathologie und Pharmakologie, 202, 262–313.*
- Düker, H. (1949). *Über ein Verfahren zur Bestimmung der geistigen Leistungsfähigkeit. Psychologische Forschung, 23, 10–24.*
- Düker, H. & Lienert, G.A. (1965). *Konzentrations-Leistungs-Test. K-L-T. Göttingen: Hogrefe.*

- Düker, H., Lienert, G. A., Lukesch, H. & Mayrhofer, S. (2001). *KLT-R. Konzentrations-Leistungs-Test – Revidierte Fassung*. Göttingen: Hogrefe.
- Gottwald, D., Knan, A. & Marsch, E. (1997). *Entwicklung einer neuen Variante eines Konzentrations-tests*. Unveröffentlichte Forschungspraktikumsarbeit, Universität Regensburg.
- Hofgärtner, S. (1998). *Zusammenhang zwischen Konzentration und Intelligenz*. Unveröffentlichte Zulassungsarbeit, Universität Regensburg.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung. PSB*. Göttingen: Hogrefe.
- Lobeck, A., Frei, M. & Blöching, R. (1990). *Schweizer Rechentest 4.–6. Klasse*. Basel: Beltz.
- Lukesch, H. (2004). *KLT-R. Revidierter Konzentrations-Leistungs-Test*. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Auflage, S. 453–457). Lengerich: Pabst.
- Marsch, E. (2000). *Konvergente und divergente Validität des KLT-R bei Schüler/innen der 4. bis 6. Schulstufe*. Unveröffentlichte Diplomarbeit, Universität Regensburg.
- Mayrhofer, S. (2004). *Der Revidierte Konzentrations-Leistungs-Test (KLT-R)*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 119–131). Göttingen: Hogrefe.
- Nell, V. (2003). *Konzentrations-Leistungs-Test, revidierte Fassung (KLT-R)*. In E. Fay (Hrsg.), *Tests unter der Lupe 4. Aktuelle psychologische Testverfahren – kritisch betrachtet* (S. 59–75). Göttingen: Vandenhoeck & Ruprecht.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Testkurzdarstellung KLT-R Revidierter Konzentrations-Leistungs-Test: S. 331–334)*. Lengerich: Pabst.

3.14. KVT; Konzentrations-Verlaufs-Test

Abels, D. (1974). Konzentrations-Verlaufs-Test KVT (2., verbesserte Auflage). Göttingen: Hogrefe.

Beschreibung

Der Konzentrations-Verlaufs-Test erlaubt eine genaue Analyse des Arbeitsverhaltens unter Daueraufmerksamkeitsbedingungen. Die Probandin/der Proband soll einen Kartenstoß nach vier Kriterien durchsehen und dahingehend sortieren, ob die Karten die eine oder andere von zwei vorgegebenen Zahlen enthalten und ob beide Ziel-Zahlen bzw. keine der beiden Ziel-Zahlen enthalten ist. Gemessen werden Arbeitszeit und Fehlerzahl; es ergibt sich eine Arbeits-Verlaufskurve.

Der im Jahre 1954 von Abels veröffentlichte KVT geht auf den Kartensortierversuch von Münsterberg (1912) und einen in den zwanziger Jahren im Leipziger Psychologischen Institut verwendeten Zahlensortiertest zurück. Diese Verfahren lieferten zwar Werte für Arbeitszeit und Fehlerzahl, eine Analyse des Arbeitsvorgangs war mit ihnen jedoch nicht möglich. Erst die vom Autor eingeführten Neuerungen, wie die Nummerierung der Karten auf der Rückseite, die Standardisierung der Darbietungsreihenfolge und die besondere Art der Fixierung der Sorgfaltsleistung, erlauben eine genaue Analyse des Arbeitsvorgangs, der in einer Verlaufsskala festgehalten werden kann.

Anhand der Verlaufskurve erhebt Abels den Anspruch, dass der KVT zusätzlich zu quantitativen Leistungswerten einen Einblick in den Arbeitsprozess liefert. Hoch- und Tiefpunkte der Konzentrationsfähigkeit sollen z.B. anhand von Fehlerhäufungen an bestimmten Stellen festgestellt werden. Zudem schreibt der Autor dem Verfahren eine größere Berufsnähe zu als den meisten anderen Konzentrationstests.

Die Probandin/der Proband ist angehalten, aus einem Kartenpäckchen mit 60 Zahlenkarten jeweils das oberste Kärtchen abzuheben und nachzusehen, ob die Zahlen „43“, „63“, „43 und 63“ unter den 36 aufgedruckten Zahlen sind oder ob keine der beiden Zahlen vorkommt. Dann soll das Kärtchen auf das mit der entsprechenden Zahl gekennzeichnete Feld auf den Arbeitsplatz gelegt werden. Das Arbeitsblatt ist mit vier quadratischen Zahlenfeldern aus je 60 Zahlen (1–60) bedruckt, die entsprechend der geforderten Aufgabe mit „43“, „63“, „43+63“ bzw. einem waagrechten Strich gekennzeichnet sind.

Eine besondere Erschwernis des Tests besteht darin, dass Karten, die nur eine „63“ enthalten, im Testmaterial überhaupt nicht vorkommen. Jede in das mit „63“ gekennzeichnete Feld abgelegte Karte ist also ein Fehler. „Die Aufmerksamkeit wird durch die Erwartung eines noch nicht eingetretenen Falles in besonderer Weise in Anspruch genommen“ (Abels, 1961, S. 7).

Nach Beendigung des Tests wird auf dem Auswertungsblatt von der Testleiterin/vom Testleiter die Zeit notiert.

Altersgruppe: 14;0–60;0 Jahre

Hinweise zur Durchführung

Das Verfahren kann als **Einzeltest und bei Gruppen** bis zu 25 Personen angewendet werden. Eine Parallelförmigkeit liegt nicht vor.

Eine Zeitbegrenzung ist nicht vorgesehen. Zur reinen Bearbeitungszeit von ca. 7 bis 16 Minuten ist noch die Zeit für die Instruktion und das Rücksortieren der Karten zu rechnen, wofür jedoch keine Angaben vorliegen.

Die Instruktion ist wortwörtlich im Manual vorgegeben. Übungsbeispiele sind nicht vorgesehen.

Das Verfahren kann von jeder Fachkraft, die mit den Prinzipien der Anwendung und Auswertung von Tests vertraut ist, durchgeführt werden. Der Autor schlägt bei Gruppenuntersuchungen zusätzlich die Mitarbeit von Testhelfer/innen vor, v.a. wegen der von ihm empfohlenen Interventionen und diagnostischen Beobachtungen während der Testdurchführung.

Hinweise zu Auswertung und Interpretation

Es liegen für Fehler- und Zeitwerte ($n = 1.171$) sowie für kombinierte Fehler-Zeit-Werte Normen in Standardwerten vor ($n = 1.201$). Die Stichproben sind zwar sehr groß, ihre Charakterisierung ist jedoch äußerst mangelhaft. Der Autor weist nur darauf hin, dass sich „diese Probanden [...] berufsmäßig etwa zu gleichen Teilen aus Angestellten und Anwärtern des öffentlichen Dienstes und aus Angehörigen verschiedener technischer und kaufmännischer Berufe“ (Abels, 1961, S. 3) zusammensetzen. Selbst wenn „vom Akademiker bis zum Volksschulabsolventen“ (Abels, 1961, S. 3) alle Bildungsbereiche abgedeckt sind, ist eine ausreichende Repräsentativität der Stichproben nicht unbedingt gewährleistet.

Eine weitere Ergänzung und Aufgliederung der Normen in Bezug auf Lebensalter, Bildungsklassen und Berufsgruppen erscheint durchführbar und notwendig.

In diese Richtung weisen auch die Ergebnisse von Seyfried (1962; zit. nach Bartenwerfer, 1964), die eine hohe Differenzierungsfähigkeit des KVT bei verschiedenen Altersklassen von Kindern und bei gleichaltrigen Gymnasiastinnen/Gymnasiasten, Hauptschüler/innen, Nicht-Hauptschulreife und Hilfsschüler/innen erkennen ließen.

Die Auswertung des KVT erfolgt hinsichtlich Arbeitszeit und Fehlerzahl. Die Nummern der aufgelegten Kärtchen werden auf dem jeweiligen Feld des Arbeitsblattes angestrichen.

Die Ermittlung der Rohpunktwerte für die Fehler geschieht dann mit Hilfe einer transparenten Schablone.

Die Zeit- und Fehlerwerte werden schließlich an Hand von Tabellen in Standardwerte für Tempoleistung (Zeit), Konzentrationsleistung (Fehler) und kombinierte Zeit-Fehler-Werte umgerechnet. Die Fehler sind zudem in eine Verlaufsskala einzutragen.

Bei den Fehlern wird zwischen drei Arten unterschieden, die eine differenzierte Auswertung und eine qualitative Deutung ermöglichen sollen:

- Fehler, bei denen entweder die „43“ oder die „63“ übersehen worden ist (einfacher Fehler).
- Fehler, bei denen beide Zahlen („43“ und „63“) übersehen worden sind (Doppel-Fehler).
- Fehler, bei denen eine Zahl gesehen wurde, die auf dem Kärtchen gar nicht vorhanden war (Illusions-Fehler).

Es ist mit ca. 1 bis 2 Minuten Auswertungszeit zu rechnen.

Testgüte

Objektivität

Durchführung und Auswertung des Verfahrens sind weitgehend standardisiert. Ein wesentlicher Kritikpunkt ergibt sich allerdings bei der Durchführung des Tests in der Gruppe. Hier soll die Testleiterin/der Testleiter den bereits fertigen Testpersonen leise auftragen, die Ergebnisse ihrer Sortierarbeit auf dem Arbeitsblatt anzustreichen und die Karten wieder richtig (von 1 bis 60) zu ordnen, um selbst Zeit und Mühe zu sparen. Diese Interventionen dürften erhebliche Auswirkungen auf die noch arbeitenden Probandinnen/Probanden haben.

Reliabilität

Die Split-half-Reliabilität des KVT wird mit $r = .67$ ($n = 157$) angegeben. Bartenwerfer (1962; zit. nach Bartenwerfer, 1964) ermittelt die Paralleltest-Reliabilität durch die Kommunalitäten zweier Faktorenanalysen ($n = 198$, $n = 152$) und kommt auf Werte von .40 und .59. Die Zuverlässigkeit des Verfahrens kann damit nur als unzureichend bewertet werden.

Validität

In der Handanweisung findet sich kein empirischer Nachweis der Gültigkeit des Tests. Zwischen Zeit- und Fehlerbewertung ermittelt der Autor eine leicht negative Korrelation von $r = -.19$ ($n = 503$). Korrelationen der KVT-Leistungen mit nicht weiter definierten Intelligenzleistungen

und der Tempoleistung beim Formensortieren sind mit $r = .03$ bzw. $r = .04$ nicht signifikant ($n = 163$ bzw. $n = 353$).

Der KVT lässt zwar die Aufeinanderfolge der Fehler erkennen, ihre genaue zeitliche Lokalisation ist jedoch nicht möglich. Der Anspruch des Verfahrens, den Leistungsverlauf sichtbar zu machen, kann somit nur eingeschränkt bestätigt werden.

Legt man den vom Autor angegebenen ziemlich niedrigen Reliabilitätskoeffizienten von $r = .67$ zugrunde, so können hohe Validitäten ohnehin nicht erwartet werden, da die tatsächliche Zuverlässigkeit der Gültigkeitskriterien erheblich unter dem idealen Wert von $r = 1$ liegen dürften (vgl. Guilford, 1954; Brickenkamp, 1963).

Kommentar der Psyn dex-Testreviewerinnen und -reviewer

Beim KVT handelt es sich um ein sehr ökonomisches und vielseitig einsetzbares Verfahren.

Allerdings sind auch eine Reihe schwerwiegender Kritikpunkte aufzuführen. Das Verfahren wird beispielsweise dem Anspruch, den Verlauf der Konzentrationsleistung zu messen, nicht gerecht, da eine genaue zeitliche Lokalisation der Fehler nicht möglich ist.

Weitere Belege zur Reliabilität und Validität sind notwendig, die Höhe der bisher ermittelten Gütekriterien kann nur als gering bewertet werden. Außerdem ist die durchgeführte Normierung problematisch, da die Stichprobe nicht eindeutig definiert und aufgeschlüsselt ist und aufgrund der vorhandenen Angaben nicht ohne Weiteres als repräsentativ gelten kann. Eine weitere Differenzierung der Normen hinsichtlich Lebensalter, Berufsgruppen und Bildungsklassen ist wünschenswert.

Die Objektivität der Durchführung des Tests in der Gruppe ist aufgrund der vom Autor empfohlenen Intervention stark beeinträchtigt. Zudem hat das Freibleiben des 4. Zahlenquadrats auf dem Arbeitsblatt zur Folge, dass die Probandinnen/Probanden verschieden lang unter „Erwartungsspannung“ arbeiten, abhängig vom Zeitpunkt, an dem ihnen diese Täuschung klar wird. Die Anwendung des KVT bei der Erfassung verschiedener Persönlichkeitsmerkmale (vgl. Freytag, 1962; Dirks, 1957) ist ebenso äußerst fragwürdig. Birke und Köhne (1967) kamen



durch die Untersuchung der Testkonstruktion des KVT anhand einer Stichprobe von $n = 400$ 16- bis 20-jährigen männlichen Untersuchungshäftlingen zu dem Schluss, dass die Mängel des Verfahrens eine völlige Neukonstruktion und Neustandardisierung notwendig machen.

Autorinnenkommentar

Der KVT eignet sich nur bedingt für den Einsatz in der Schulpsychologie. Es ist (ähnlich wie auch bei anderen Verfahren wie dem FAIR-2 oder dem d2) möglich, Arbeitszeit und Fehleranzahl zu überprüfen, also die Schülerin/den Schüler innerhalb der 4-Felder-Tafel „schnell/langsam x genau/ungenau“ einzustufen. Eine AD(H)S-Abklärung kann mit diesem Test nicht gemacht werden. Die Aktualität der Normen (aus den Jahren 1960 und 1974) ist fraglich.

Literatur

- Abels, D. (1954). *Konzentrations-Verlaufs-Test K-V-T*. Göttingen: Hogrefe.
- Abels, D. (1961). *Konzentrations-Verlaufs-Test K-V-T (überarbeitete Version)*. Göttingen: Hogrefe.
- Abels, D. (1974). *Konzentrations-Verlaufs-Test KVT (2., verbesserte Auflage)*. Göttingen: Hogrefe.
- Bartenwerfer, H. (1964). *Allgemeine Leistungstests*. In R. Heiss (Hrsg.), *Handbuch der Psychologie*. Band 6, *Psychologische Diagnostik* (S. 385–410). Göttingen: Hogrefe.
- Birke, W. & Köhne, H. (1967). *Zur Kritik des KVT (nach Abels)*. *Diagnostica*, 13, 130–133.
- Brickenkamp, R. (1963). *Einige theoretische Grundzüge und praktische Methoden der Zuverlässigkeitsschätzung psychodiagnostischer Verfahren*. *Psychologische Rundschau*, 14, 13–25.
- Dirks, H. (1957). *Über die Reichweite des KVT (nach Abels) zur Diagnose der Willensstruktur (nach Herwig-Dirks)*. *Diagnostica*, 3, 57–62.
- Freytag, G. (1962). *Experimentalpsychologie und mathematische Typenanalyse bei Depressionszuständen*. *Zeitschrift für Psychologie*, 166, 241–297.
- Guilford, J. P. (1954). *Psychometric Methods*. New York: McGraw-Hill.
- Kühne, W. (2009). *Konzentrations-Verlaufs-Test (K-V-T)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: *Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 92–98). Göttingen: Hogrefe.
- Münsterberg, H. (1912). *Experimentalpsychologie und Berufswahl*. *Zeitschrift für experimentelle Pädagogik und pädagogische Psychologie*, 13, 1–7.

3.15. KT 3–4 R; Konzentrationstest für 3. und 4. Klassen – Revidierte Fassung

Nell, V., Bretz, J. & Sniehotta, F. F. (2004). KT 3–4 R. Konzentrationstest für 3. und 4. Klassen. Revidierte Fassung. Göttingen: Beltz.

Beschreibung

Der KT 3–4 R dient zur Erfassung der Aufmerksamkeit und Konzentrationsleistung von Kindern der dritten und vierten Schulstufe. Der Test kann in der Schulpsychologie, in der klinisch-psychologischen Praxis und der pädagogisch-psychologischen Forschung eingesetzt werden.

Mit dem KT 3–4 R (Nell, Bretz & Sniehotta, 2004) sollen Aufmerksamkeit und Konzentrationsleistung von Kindern der dritten und vierten Schulstufe über einen schulrelevanten Zeitraum untersucht werden. „Aufmerksamkeit“ wird aufgefasst als eine allgemeine Leistungskomponente, welche die Vorbedingung weiterer, spezifischerer Leistungen darstellt (Nell et al., 2004, S. 7). Die „Konzentrationsleistung“ wird im Einklang mit gängigen Auffassungen (z.B. Brickenkamp, 2002; Westhoff, 1995) definiert als „zügige und korrekte Bearbeitung einer Anzahl gleicher Aufgaben über einen längeren, pädagogisch relevanten Zeitraum“ (Heck-Möhling, 1993, S. 3).

Bei dem KT 3–4 R handelt es sich um die revidierte Fassung des KT 3–4 von Heck-Möhling (1993). Ebenso wie das Vorgängerverfahren ist der KT 3–4 R als Speed-Test konzeptualisiert und der Gruppe der sog. Durchstreichtests zuzurechnen. Er stellt keine großen Anforderungen an schriftsprachliche, rechnerische oder sonstige schulische Fertigkeiten, so dass Kinder, die Schwierigkeiten beim Erwerb dieser Kulturtechniken haben, bei der Erfassung ihrer Konzentrationsleistung nicht benachteiligt werden.

Die Überarbeitung des KT 3–4 war durch das Bestreben motiviert, Richtlinien für die Diagnostik bei Konzentrationsproblemen zu gewinnen, welche in der schulpsychologischen Praxis verwendet werden können.

Die Veränderungen im KT 3–4 R werden in fünf Punkten zusammengefasst (Nell et al., 2004, S. 11):

1. Einführung des vollständigen Markierungsprinzips: Jedes Item ist vor seiner Bearbeitung zu markieren, damit bei der Auswertung erkennbar wird, ob die Testperson Items, Zeilen oder Seiten im Testheft übersprungen hat.
2. Einführung von Zeitmarken: Alle fünf Minuten ist auf Ansage ein waagerechter Strich unter das gerade bearbeitete Item zu setzen. Mit Hilfe dieser Markierungen soll einerseits der Leistungsverlauf diagnostisch beurteilt werden. Andererseits eröffnet die Einführung von Zeitintervallen die Möglichkeit, die interne Konsistenz des Verfahrens (Cronbachs α) als Reliabilitätsmaß zu bestimmen.



3. Einführung standardisierter Bild-Distraktoren: Zwischen den relevanten Testitems werden comic-artige Bilder präsentiert, die für Kinder attraktiv sind und daher von der Bildbearbeitung ablenken können; sie dienen der Erhöhung der ökologischen Validität des Tests.
4. Einführung neuer Testkennwerte: Neue Leistungsmaße beziehen sich auf die Gesamtleistung sowie die Leistungsveränderung über die vier abgegrenzten Zeitintervalle.
5. Berechnung neuer Normwerte

Das Testheft zum KT 3–4 R umfasst 13 Seiten mit jeweils 29 Aufgaben. Oben auf jeder Seite finden sich vier so genannte „Musterwürfel“, die

aus einem Würfel mit unterschiedlicher Augenverteilung bestehen, über und unter dem je eine Ziffer abgedruckt ist. Darunter sind fünf Zeilen mit jeweils sechs Feldern gezeigt; in einem ist eine comic-ähnliche Zeichnung enthalten, in den 29 weiteren sind Würfel (ebenfalls oben und unten versehen mit einer Ziffer) abgebildet. Aufgabe der Kinder ist es, diese Teststimuli in der vorgegebenen Reihenfolge mit den Musterwürfeln zu vergleichen. Diejenigen Teststimuli, die in allen Aspekten (Augenverteilung auf dem Würfel und Ziffern) mit einem der Musterwürfel übereinstimmen, sind durchzustreichen. Da sich die Musterwürfel von Seite zu Seite unterscheiden, können sie nicht gelernt werden. Die Aufgabe soll so schnell, aber auch so sorgfältig wie möglich bearbeitet werden. Um die instruktionsgemäße Bearbeitung zu kontrollieren, wurden das vollständige Markierungsprinzip und Zeitmarken eingeführt (s.o.).

Altersgruppe: 3. und 4. Schulstufe

Hinweise zur Durchführung

Der KT 3–4 R kann als **Einzel- oder Gruppentest** vorgegeben werden. Es liegen keine Parallelformen vor.

Die reine Testzeit beträgt 20 Minuten. Einschließlich Instruktionen ist in Einzeluntersuchungen mit einer Durchführungszeit von 25 Minuten zu rechnen. In Gruppenuntersuchungen sollte eine Schulstunde (ca. 40 Minuten) veranschlagt werden.

Eine sehr ausführliche, standardisierte Instruktion für die Durchführung in Gruppen findet sich im Testmanual (Nell et al., 2004, S. 13 ff.). Sie ist auch im Testheft abgedruckt und sollte dort von den Kindern mitgelesen werden. In einer Einzeltestung muss die Instruktion von der Testleiterin/vom Testleiter in den Singular übertragen werden; das Kind sollte in diesem Fall nicht mitlesen. Die Testpersonen sollten über hinreichende motorische Fertigkeiten sowie deutsche Sprach-

kenntnisse verfügen, die das Verständnis der Instruktionen erlauben. Als wichtigste Voraussetzungen auf Seiten der Testleiterin/des Testleiters werden die gründliche Vertrautheit mit der Instruktion und pädagogische Kompetenz genannt; sofern eine Gruppendurchführung geplant ist, sollte die Testleiterin/der Testleiter über Erfahrungen mit Kindergruppen entsprechenden Alters verfügen. Das Verfahren kann auch von Lehrerinnen und Lehrern durchgeführt werden, wenn diese in das angemessene Testleiter/innen-Verhalten eingewiesen wurden.

Hinweise zu Auswertung und Interpretation

Die Normierung erfolgte in zwei Erhebungswellen (März bis Mai 2001 und März bis Mai 2002) in einer für Berlin repräsentativen Gebietsstichprobe sowie einer Klumpenstichprobe aus einer ländlichen Region Mecklenburg-Vorpommerns. Die Normstichprobe setzt sich aus $n = 2.928$ Kindern (48,5 % Mädchen, 51,5 % Jungen; $M = 10;1$ Jahre, $SD = 11;2$ Monate) aus 150 Schulklassen (72 dritte und 78 vierte Klassen) zusammen.

Mitgeteilt werden Prozentränge und T-Werte für die Leistungskennwerte GZ, KL und GZxd, Prozentränge für den Sorgfaltskennwert R % sowie Quartilwerte für DGZ und SR %. Die Werte wurden getrennt für die dritte und vierte Klassenstufe sowie für sechs Altersgruppen (8;6–8;11 Jahre, 9;0–9;5 Jahre, 9;6–9;11 Jahre, 10;0–10;5 Jahre, 10;6–10;11 Jahre, 11;0–11;11 Jahre; $231 \leq n \leq 644$) ermittelt. Eine Normierung nach Geschlecht oder regionaler Herkunft war nicht erforderlich.

Ferner werden (ebenfalls für zwei Klassenstufen und sechs Altersgruppen) die 95 %-Konfidenzintervalle und kritische Differenzen für die Rohwerte der Indizes GZ, R %, dxGZ und KL mitgeteilt (Nell et al., 2004, S. 79).

Die Daten wurden ausschließlich in Gruppentestungen gewonnen, können jedoch auch zur Bewertung von Leistungen in Einzeltestungen herangezogen werden, da sich in einer Kontrolluntersuchung ($n = 82$) keine bedeutsamen Unterschiede zwischen Leistungen unter Einzel- vs. Gruppentestbedingungen ergeben hatten. Alle Normwerte gelten nur für eine Erstuntersuchung, da sich in der Reteststudie für die Mengenleistung GZ ein durchschnittlicher Leistungszuwachs von 28,5 Items gezeigt hatte.

Als Richtwerte zur Beurteilung der Klassenleistungen sind bei Nell et al. (2004, S. 76) Quartilwerte für Klassenmittelwerte und -standardabweichungen der KL-Werte aus $n = 72$ dritten und $n = 78$ vierten Klassen aufgeführt.

Die Auswertung des Verfahrens erfolgt manuell. Zunächst wird geprüft, ob das Kind mindestens 75 Prozent der Würfel vor der Markierung, welche den zuletzt bearbeiteten Würfel kennzeichnet, mit einem Punkt als „bearbeitet“ gekennzeichnet hat, da der Test nur dann ausgewertet werden darf.

Sodann werden mit Hilfe von Auswertungsschablonen folgende Werte ermittelt (jeweils separat für die vier Zeitintervalle 1 bis 4 sowie aufsummiert für den Gesamttest):

- RL: Anzahl richtig als mit einem Musterwürfel übereinstimmend gekennzeichnete Items
- RN: Anzahl richtig als nicht mit einem Musterwürfel übereinstimmend und daher nicht gekennzeichnete Items
- FA: Anzahl mit einem Musterwürfel übereinstimmender, aber nicht entsprechend gekennzeichnete Items
- FB: Anzahl nicht mit einem Musterwürfel übereinstimmender, aber dennoch als übereinstimmend gekennzeichnete Items

Aus diesen Werten werden im Rahmen der Standardauswertung die folgenden Kennwerte für die Leistung und Leistungsveränderungen berechnet:

- Mengenleistung GZ: Summe aller bearbeiteter Items, d.h. $RL+RN+FA+FB$
- Sorgfaltsleistung R %: Anteil richtig bearbeiteter Items an der Gesamtzahl bearbeiteter Items, d.h. $([RL+RN]/GZ) \times 100$
- Konzentrationsleistung KL: Differenz zwischen der Anzahl richtig durchgestrichener und der Anzahl falsch durchgestrichener Items, d.h. $RL-FB$
- Tempoveränderung DGZ: Differenz zwischen der Anzahl der markierten Items in Zeitintervall 4 und Zeitintervall 1, d.h. $GZ4-GZ1$

Im Rahmen der Zusatzauswertung können ferner berechnet werden:

- Diskriminations-Mengen-Leistung $dxGZ$: Multiplikation der Mengenleistung GZ mit einem analog zur Signalentdeckungstheorie entwickelten Diskriminationsindex d
- Sorgfaltsstreuung SR %: Streuung der Sorgfaltsleistung über die vier Intervalle R % 1 bis R % 4

Es liegen farbig gestaltete transparente Auswertungsschablonen sowie ein Auswertungs- und Ergebnisbogen vor, der auch eine Profildarstellung erlaubt. Ausführliche Interpretationshinweise (einschließlich zweier Fallbeispiele) finden sich bei Nell et al. (2004, S. 16 ff.). Dort wird darüber hinaus auf ein optional erhältliches Auswertungsprogramm verwiesen, jedoch keine Bezugsquelle spezifiziert.

Explizite Angaben zur Auswertungszeit fehlen. Die manuelle Standardauswertung dürfte ca. 10–15 Minuten beanspruchen. Hingegen scheint die Zusatzauswertung recht zeitaufwändig zu sein.

Testgüte

Objektivität

Aufgrund der standardisierten Materialien und Instruktionen, der standardisierten Auswertung

und der normorientierten Interpretation kann das Verfahren Objektivität beanspruchen. Allerdings ergaben sich in den Normierungsstudien, die von neun trainierten Testleiterinnen und Testleitern angewiesen worden waren, bedeutsame Testleiter/innen-Effekte (Effektstärken von $.01 \leq \text{Eta-Quadrat} \leq .07$).

Reliabilität

Die interne Konsistenz liegt für verschiedene Kennwerte zwischen $87 \leq \alpha \leq .92$. Die Retest-reliabilität nach 2–3 Monaten beträgt für das Arbeitstempo $r_{tt} = .75$, für die Konzentrationsleistung $r_{tt} = .74$. Für die – stärker situationsbedingten Schwankungen unterliegende – Arbeitsorg-falt ergibt sich eine Retestreliabilität von $r_{tt} = .47$.

Validität

Umfangreiche Befunde zur Validität des Verfahrens werden von Nell et al. (2004, S. 34 ff.) berichtet. Die inhaltlich-logische Gültigkeit des KT 3–4 R wird durch Lehrer/innen-Urteile über die Vorgängerversion KT 3–4 sowie Probandinnen/Probanden- und Expertinnen/Experten-Ein-schätzungen der Revision belegt.

Im Zuge der kriterienbezogenen Validierung des KT 3–4 R wurden in verschiedenen Stichproben Partialkorrelationen seiner Indizes (unter Kontrolle des Alters) mit folgenden Variablen berechnet:

- Lehrer/innen-Urteile: Geschwindigkeit der Aufgabenbearbeitung, Fehlerhaftigkeit der Aufgabenbearbeitung, Konzentriertheit im Unterricht, Dauerhaftigkeit der Konzentrati-onsleistung; Symptome von AD(H)S
- Schulleistungen: Deutsch und Mathematik
- Elternurteile: Skalen zu Konzentriertheit bei den Hausaufgaben ($k = 13$) und Konzentriert-heit beim Spiel ($k = 13$); Skalen zu Symptomen von AD(H)S: Unaufmerksamkeit ($k = 10$), Impulsivität ($k = 5$), Hyperaktivität ($k = 5$)
- Selbsteinschätzungen der schulischen Konzentriertheit

Insgesamt fielen die Zusammenhänge niedrig bis allenfalls moderat aus ($r < .40$). Die höchsten Validitätskoeffizienten fanden sich für die Indikatoren der Mengenleistung (GZ) und der Gesamtleistung (KL, dxGZ), während sich v.a. die Tempoveränderung DGZ zur Vorhersage der Kri-terien meist als untauglich erwies. Mit Blick auf die differenzielle Validität wurden zusätzlich die Zusammenhangsmuster zwischen KT 3–4 R und den Kriteriumsvariablen in den Stichproben der Mädchen vs. Jungen resp. der 3. vs. 4. Klassenstufe verglichen; es ergaben sich lediglich geringe Unterschiede in den Validitätskoeffizienten.

Die Annahme der konvergenten Validität des KT 3–4 R wird – allerdings ebenfalls nur bedingt – gestützt durch Korrelationen mit dem Differentiellen Leistungstest-Konzentration (DL-KG; Kleber, Kleber & Hans, 1999): Die DL-KG-Mengenleistung korrelierte bedeutsam mit den Kennwerten GZ ($r = .48$), KL ($r = .41$), dxGZ ($r = .32$) und R % ($r = .17$). Niedrige, überwiegend nicht-signifikante Zusam-menhänge fanden sich hingegen zwischen der DL-KG-Variable Fehleranteil und den KT 3–4 R-Kenn-

werten. Mit dem ebenfalls als Speed-Test konzeptualisierten Zahlenverbindungstest (ZVT; Oswald & Roth, 1987) waren alle Variablen des KT 3–4 R moderat korreliert ($.19 \leq r \leq .35$).

Als Beleg für die diskriminante Validität des KT 3–4 R werden geringe bis moderate Korrelationen ($r \leq .32$) mit mehreren Subtests („Wortklassifikation“, „Figurenanalogien“, Gesamtwert für „Schlussfolgerndes Denken“) aus dem KFT 4-12+R (Heller & Perleth, 2000) aufgeführt. Die KT 3–4 R-Leistungen erwiesen sich ferner als relativ unabhängig von der Motivation der Kinder. Es fanden sich Korrelationen von $r \leq .25$ mit Selbsteinschätzungen im „Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen“. Zudem waren keine Unterschiede zwischen den Testleistungen von Gruppen mit vs. ohne angekündigte Belohnung für gute Testleistungen zu verzeichnen (Nell, Bretz & Sniehotta, 2001). Ebenfalls gering ($r < .22$) fielen Korrelationen der Testwerte mit dem durch Lehrer/innen beurteilten Sprachverständnis der Kinder aus.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Mit dem KT 3–4 R liegt ein Test zur Erfassung der Konzentrationsleistungen von Schülerinnen und Schülern vor, der sich auf ein bewährtes Testkonzept stützt und für den eine umfangreiche Datenbasis vorliegt. Die psychometrischen Befunde, insbesondere die hohen Reliabilitätskoeffizienten, weisen das Verfahren als brauchbar für den angestrebten Einsatzkontext aus. Weniger überzeugend erscheinen die Korrelationsmuster mit validitätskonvergenten und -divergenten Verfahren, was jedoch teils auf die Auswahl der Kriterien zurückgehen könnte. Anstelle kombinierter Speed-Power-Tests zur Intelligenzmessung könnten hier z.B. reine Power-Tests (z.B. die Raven-Matrizen-Tests) herangezogen werden, um die Frage der Unabhängigkeit von Konzentrationsleistung und Intelligenz zu prüfen. Völlig ungeeignet für die Leistungsbeurteilung scheint jedoch der Index für die Tempoveränderung DGZ, der weder mit den übrigen Testwerten noch mit Validitätskriterien konsistente Zusammenhänge aufweist. Schließlich bleibt zu hoffen, dass die Autorinnen/Autoren baldmöglichst das im Manual erwähnte Auswertungsprogramm publizieren, um die Handhabung des Verfahrens zu erleichtern.

Autorinnenkommentar

Der Test ähnelt vom Prinzip der DL-KG, dem FAIR-2 und der KKA, hat jedoch andere Stimuli. Er ist zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Besonders Kinder mit einer Begabung im räumlichen Denken haben Spaß bei der Durchführung. Andere Kinder – insbesondere AD(H)S-Kinder – verlieren, wenn sie das Material sehen, oft das Interesse und die Lust, den Test durchzuführen. Der Test ermöglicht es sehr gut, die fokussierte und die selektive Aufmerksamkeit bei Kindern zu überprüfen und sollte in Kombination mit weiteren Verfahren zur Überprüfung der Vigilanz und der räumlichen Aufmerksamkeit angewendet werden, sofern ein umfassendes Bild der Aufmerksamkeitsleistung gewünscht wird. Eine Ausdehnung des Anwendungsbereichs auf Jugendliche könnte sinnvoll sein.

Literatur

- Brickenkamp, R. (2002). *Test d2. Aufmerksamkeits-Belastungs-Test (9., vollständig überarbeitete und neu normierte Auflage)*. Göttingen: Hogrefe.
- Heck-Möhling, R. (1993). *KT 3–4. Konzentrationstest für 3. und 4. Klassen (2. Auflage)*. Weinheim: Beltz.
- Heller, K. A. & Perleth, C. (2000). *KFT 4-12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.
- Kleber, E. W., Kleber, G. & Hans, O. (1999). *Differentieller Leistungstest – KG (DL-KG). Test zur Erfassung des Leistungsverhaltens bei konzentrierter Tätigkeit im Grundschulalter (2., korrigierte Auflage)*. Göttingen: Hogrefe.
- Nell, V., Bretz, J. & Sniehotta, F. F. (2001). *Bedingungsvariirtes Testen: Eine neue Konzeption zur Sicherung der ökologischen und kriterienbezogenen Validität in der Konzentrations- und Aufmerksamkeitsdiagnostik*. In Deutsche Psychologen Akademie (Hrsg.), *Psychologie am Puls der Zeit* (S. 348–349). Bonn: Deutscher Psychologen Verlag.
- Nell, V., Bretz, J. & Sniehotta, F. F. (2004). *KT 3–4 R. Konzentrationstest für 3. und 4. Klassen. Revidierte Fassung*. Göttingen: Beltz.
- Oswald, W. D. & Roth, E. (1987). *Der Zahlenverbindungstest (ZVT). Ein sprachfreier Intelligenz-Test zur Messung der „kognitiven Leistungsgeschwindigkeit“ (2., überarbeitete und erweiterte Auflage)*. Göttingen: Hogrefe.
- Westhoff, K. (1995). *Aufmerksamkeit und Konzentration*. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie, Serie VIII: Differentielle Psychologie und Persönlichkeitsforschung, Band 2: Verhaltens- und Leistungsunterschiede* (S. 375–402). Göttingen: Hogrefe.



3.16. P-T; Pauli-Test

Arnold, W. (1975). Der Pauli-Test. Anweisung zur sachgemässen Durchführung, Auswertung und Anwendung des Kräpelinischen Arbeitsversuchs (5., korrigierte Auflage). Berlin: Springer.

Beschreibung

Beim Pauli-Test handelt es sich um einen Test zur Erfassung der Konzentrationsleistung. Die Probandin/der Proband soll auf einem standardisierten Rechenbogen fortlaufend und möglichst schnell je zwei einstellige Zahlen addieren und zwar derart, dass in jede Lücke rechts die Summe der darüber und darunter stehenden Zahl zu schreiben ist, unter Weglassung der ersten Ziffer, wenn das Ergebnis zweistellig ist. Alle drei Minuten muss die Probandin/der Proband eine Markierung (Querstrich) in der Zahlenreihe vornehmen, um eine Verlaufsauswertung zu ermöglichen. Das Ergebnis des Arbeitsversuchs ist dann unter den Gesichtspunkten Menge, Güte und Verlauf der Leistung zu betrachten.

Der Pauli-Test geht auf die Versuche des Kräpelin-Schülers Oehr (1896) zurück, der erstmals die Methode fortlaufenden Addierens einstelliger Zahlen zur Bestimmung individueller Differenzen im Leistungsverlauf verwendete. Auf die Veröffentlichungen von Kräpelin und seinen Schülerinnen/Schülern folgten zahlreiche Untersuchungen mit Aufgaben dieser Art, bis Pauli (1936) die vorliegenden Erfahrungen zu einem standardisierten Test verarbeitete. Kräpelins Leitgedanke bei der Begründung der experimentellen Arbeitspsychologie besteht darin, dass die „seelischen Tatbestände“, die mit der Arbeit gegeben sind, mit Hilfe einer geeigneten Tätigkeit planmäßig herbeigeführt und untersucht werden sollen. Beim Pauli-Test handelt es sich nun um eine Tätigkeit, die nicht dem eigenen Antrieb entspringt, sondern aufgrund einer fremden Forderung zustande kommt. Maßgebend ist die Eintönigkeit des Geschehens, man kann die Leistung als fortlaufende Reihe einfacher Reaktionen auffassen. Die Dauerbelastung, die mit dem Test gegeben ist, hängt demnach v.a. von der persönlichen Anstrengung ab. Das im Test erzielte Ergebnis drückt die willentliche Selbstbelastung bei einem Mangel fast aller Anreize aus. Arnold (1975) geht davon aus, dass diese innere Teilnahmslosigkeit ein wesentliches Kennzeichen der heutigen industriellen Berufsarbeit ist und dass jeder Beruf die Anforderung stellt, auch das zu tun, was nicht den eigenen Wünschen entspricht. Da auch der Erfolg beim Pauli-Test nicht einfach von der besonderen Fähigkeit zur fortlaufenden Addition abhängt, soll dieser Arbeitsversuch auch Hinweise auf Arbeitsfähigkeit, Willensstärke, Zähigkeit, Ausdauer, Selbstbeherrschung usw. liefern, also auf Persönlichkeitseigenschaften, die in unmittelbarer Beziehung zur Leistung stehen.

Beim Pauli-Test handelt es sich um ein homogenes Verfahren mit zeitlicher Untergliederung in 20 Teilzeiten (3 Minuten-Intervalle). Die Probandin/der Proband soll auf einem standardisierten

Rechenbogen mit insgesamt 4.000 Items fortlaufend und möglichst schnell je zwei einstellige Zahlen (Ziffern 2 bis 9) addieren. Die Markierungen ermöglichen nicht nur Menge und Güte, sondern auch den Verlauf der Leistung zu betrachten. Dazu werden folgende Merkmale der Arbeitskurve bestimmt: die Gesamtmenge der Additionen, die Prozentsätze von Fehlern und Verbesserungen, die durchschnittliche Schwankung der Arbeitskurve gegenüber der doppelt ausgeglichenen Kurve bezogen auf die mittlere Teilzeitleistung, die Steighöhe der Kurve und die Lage des Kurvengipfelpunktes.

Altersgruppe: ab 7;0 Jahren

Hinweise zur Durchführung

Das Verfahren kann als **Einzel- und Gruppentest** angewendet werden.

Es existieren sowohl eine Kurzform (Bochow, 1967) als auch die Möglichkeit, den Pauli-Test als apparatives Verfahren mit einem elektronischen Gerät (LPS von Schuhfried) durchzuführen (Kallenbach, 1978). Außerdem gibt es Sonderformen für Blinde in Blindenschrift (Weiss, 1950) sowie für Analphabeten und Schreibunkundige (Reuning, 1960).

Die reine Testdauer beträgt 60 Minuten. Zur Dauer der Vorbereitung (Instruktion, Probeversuch) werden keine Angaben gemacht.

Eine Instruktion ist nur in Form inhaltlicher Anweisungen zur Testdurchführung für die Testleiterin/den Testleiter vorhanden, die/der diese in ihren/seinen eigenen Worten an die Probandin/den Probanden weitergeben muss. Vor Testbeginn erhält die Testperson anhand eines Probestreifens mit zwei Reihen zu je 10 Ziffern Gelegenheit, mit der Tätigkeit vertraut zu werden.

Das Verfahren kann von jeder Fachkraft, die mit der Anwendung und Auswertung von Tests vertraut ist, durchgeführt werden.

Hinweise zu Auswertung und Interpretation

Das Textbuch zum Pauli-Test (Arnold, 1975) enthält eine Reihe von Normentabellen für verschiedene Stichproben, wie z.B. männliche und weibliche Jugendliche und Erwachsene mit Volksschulabschluss, schwererziehbare und normale Jugendliche, Personen mit Hirnverletzungen, gehörlose Jugendliche, Lehrlinge, Jugendliche und Erwachsene von höheren Schulen und Hochschulen usw., wobei die untersuchten Altersbereiche insgesamt zwischen 7 und 35 Jahren liegen. Die unterschiedlichen Untersuchungsbedingungen machen die Normentabellen für die Praxis ziemlich unbrauchbar, so dass „der praktisch arbeitende Psychologe Normen für



seine speziellen Bedingungen selbst aufstellen muss“ (Bartenwerfer, 1964).

Als Auswertungshilfen sind Auswertungsbögen und ein Auswertungsgerät vorhanden. Zuerst wird aufgrund der Summe der Additionen für alle 20 Teilzeiten der Gesamtleistungswert ermittelt. Die Leistungen in den einzelnen Teilzeiten werden für die drei Arbeitskurven gebraucht. Aus der Verbindung der einzelnen Teilleistungswerte entsteht die erste Arbeitskurve, die in das Auswertungsformular eingetragen wird.

Die anderen beiden Kurven ergeben sich aus Umrechnungsprozeduren dieses ersten Kurvenverlaufs. Der Autor geht davon aus, dass diese drei Linien weitreichende Rückschlüsse auf die Persönlichkeit der Probandin/des Probanden ermöglichen. Zusätzlich werden auch noch Gipfelflage, Steighöhe und Schwankungsprozent der Arbeitskurve berechnet. Im Textbuch finden sich ausführliche Hinweise zur psychologischen Deutung der Hauptmerkmale der Arbeitsleistung.

Testgüte

Objektivität

Die Durchführungsobjektivität kann nicht als völlig gesichert gelten, da die Testleiter/innen die Probandinnen/Probanden nach Pauli auf unterschiedliche Weise zu Höchstleistungen motivieren sollen. Die Objektivität der Auswertung ist gegeben, wobei die Diskussion über die Art und Anzahl der brauchbaren Aspekte der Arbeitskurve noch nicht abgeschlossen ist.

Reliabilität

Bisher sind nur wenig verwertbare Angaben über die verschiedenen Arten der Reliabilität vorhanden. Besonders bei Gruppen, die sich nach Alter, Geschlecht und Bildung unterscheiden, fehlen Koeffizienten für die verschiedenen Reliabilitätsaspekte. Anhaltspunkte für die Reliabilitätsverhältnisse geben die im Folgenden aufgeführten Untersuchungen:

- Plössl (1941) fand für einen Zeitabstand von einem Monat einen Retest-Koeffizienten (Rangkorrelation) von $r_{tt} = .96$ bezüglich einer Indexzahl aus Menge und Güte und von $r_{tt} = .68$ bezüglich der Fehler ($n = 58$).
- Bartenwerfer (1963) teilt für 152 Gymnasiastinnen/Gymnasiasten eine Split-half-Reliabilität (1. und 2. Testhäfte) von $r_{tt} = .97$ für das Merkmal „Gesamtmenge“ mit. Für die Merkmale „Fehler im Gesamttest“ bzw. „Verbesserungen“ ergaben sich hier Split-half-

Koeffizienten von $r_{tt} = .80$ bzw. $r_{tt} = .88$.

- Reuning (1959) ermittelte für das Merkmal „Gesamtmenge“ durch eine fragwürdige Halbierung nach gerad- und ungeradzahligem Additionen einen Split-Half-Koeffizienten von $r_{tt} = .99$. Für einen Test mit ausgesprochenem Speed-Charakter führt eine solche Halbierung jedoch zu einer Überschätzung des Zusammenhangs (vgl. Bartenwerfer, 1963).
- Bochow (1967) fand bei einer Kurzform des Pauli-Tests (20 mal 1 Min.) für eine Stichprobe von $n = 1.060$ eine Split-half-Reliabilität von $r_{tt} = .97$ für die Gesamtadditionsmenge.

Für die Merkmale „Schwankung“, „Steighöhe“ und „Gipfellage“ wurden in den bisher genannten Untersuchungen keine Reliabilitätsschätzungen vorgenommen.

Wulsten und Ulich (1971) führten an Schülerinnen und Schülern einer höheren Schule vier Testwiederholungen im Abstand von jeweils ca. acht Wochen durch und errechneten für sieben Merkmale des Verfahrens die Retest-Reliabilitäten mit Hilfe des Korrelationskoeffizienten von Bravais-Pearson. Die Ergebnisse beziehen sich auf die erste, zweite und vierte Testaufnahme. Eine Wiederholung der Berechnungen mittels des Spearman-Rangkorrelationskoeffizienten erbrachte nur unerhebliche Unterschiede zur ersten Berechnung. Für die Merkmale „Gesamtmenge“, „Fehler (v.H.)“, „Verbesserung (v.H.)“ und „Schwankung (v.H.)“ ergaben sich signifikante Retest-Koeffizienten in allen verglichenen Testaufnahmen. Den Merkmalen „Steighöhe“, „Gipfellage“ und „Schwankung (Mittel)“ sprechen Wulsten und Ulich (1971) eine Wiederholungszuverlässigkeit ab.

Insgesamt scheint die Gesamtmenge im Pauli-Test hoch reliabel zu sein, während die Zuverlässigkeit der übrigen Testwerte eher fragwürdig ist oder zumindest nicht als gesichert gelten kann.

Validität

Zahlreiche Untersuchungen haben den Pauli-Test als diagnostisches Hilfsmittel bei verschiedenen Problemen miteinbezogen (z.B. Arnold, 1958; Fischer & Ulich, 1961; Gebauer, 1965; Gruhnwald & Ulich, 1959; Gutewa, 1957; Hitpass, 1961, 1963; Jäger, 1960; Kohlmann, 1954; Ulich, 1958).

Außerdem sind mehrere Faktorenanalysen mit dem Pauli-Test durchgeführt worden (z.B. Arnold, 1961; Bäumlner, 1964; Bäumlner & Breitenbach, 1970; Bäumlner & Weiss, 1967; Bartenwerfer, 1964; Bernstein & Kuenzel, 1976; Seitz, 1971).

Eine von Bäumlner (1964) durchgeführte Faktorenanalyse ergab drei Faktoren, den numerischen Faktor „Grundfähigkeit des rechnerischen Umgangs mit Zahlen“, der hier auch als Interferenzfaktor interpretiert werden kann, den Faktor der Arbeitsgeschwindigkeit und einen Faktor „sinngemäßes Operieren mit Begriffen und Symbolen, speziell aber mit sprachlichen Begriffen“. Die Ladungen der Pauli-Testmenge auf den einzelnen Faktoren betragen $r = .40$, $r = .70$ und $r = .36$, der Faktor „Arbeitsgeschwindigkeit“ ist als der charakteristische Pauli-Test-Faktor anzusehen.

Eine Untersuchung von Bernstein und Kuenzel (1976) zum Vergleich von Struktur und Niveau der Intelligenz einer Gruppe hirngeschädigter, körperbehinderter ($n = 87$) und einer Gruppe nichthirngeschädigter, körperbehinderter Jugendlicher ($n = 72$) ergab eine weitgehende Übereinstimmung der Faktorenstrukturen der beiden Stichproben, während sie sich im Niveau des intellektuellen Verhaltens deutlich unterschieden. Der Pauli-Test zeigte auch hier eine hohe Ladung auf einem Faktor „allgemeine Aufmerksamkeit bei Tempomotivation“. Außer der Konzentrationskraft scheinen für diesen Faktor aber noch andere psychische Prozesse ausschlaggebend zu sein.

Insgesamt dürfte ein Problem der Verwendung von Rechenaufgaben zur Messung der allgemeinen Leistungsfähigkeit darin liegen, dass damit immer auch elementare Rechenfähigkeiten erfasst werden.

Eine Untersuchung von Gymnasialschülerinnen und -schülern im Alter von 15 und 17 Jahren ($n = 51$) ergab korrelative und faktoranalytische Zusammenhänge zwischen diversen Merkmalen des Pauli-Tests und verschiedenen Persönlichkeitsdimensionen (Seitz, 1971). Eine persönlichkeitsdiagnostische Relevanz zeigten v.a. drei Pauli-Test-Syndrome:

- Additionssumme und Schwankung kovariieren positiv mit „ich-schwacher Feldabhängigkeit“.
- Hohe Additionssumme, hohe Steighöhe und geringe Fehlerzahl stehen in Zusammenhang mit „selbstsicher-realistischer Robustheit“.
- Hohe Verbesserungszahl und geringe Gipfellage haben mit „Nachlässigkeit“ zu tun.
- Eine Generalisierung der Ergebnisse ist jedoch aufgrund der verwendeten kleinen Stichprobe keinesfalls möglich. Außerdem erreichen die Zusammenhänge höchstens ein mittleres Niveau.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Die Angaben zu wichtigen teststatistischen Kriterien sind z.T. lückenhaft, fragwürdig in ihrer Aussagekraft oder nicht genügend untersucht. Die Objektivität der Durchführung ist nicht gegeben, es sind nur wenig verwertbare Angaben zu Reliabilitätskoeffizienten in der Literatur zu finden und die unterschiedlichen Untersuchungen zur praktischen Anwendbarkeit des Pauli-Tests tragen nur wenig zur Abklärung der testtheoretischen Gültigkeit bei. Auch die verschiedenen Normentabellen können nur für die speziellen Bedingungen der einzelnen Untersuchungen, in denen sie erstellt wurden, verbindlich sein. Allgemeingültige Normen, die für eine breite Anwendung in der Praxis brauchbar sind, sind nicht vorhanden.

Was den Einsatz des Pauli-Tests in der charakterologischen Diagnostik betrifft, so sind kaum gesicherte Aussagen hinsichtlich der Ausprägung verschiedener Persönlichkeitsmerkmale zu machen.

Auch die Annahme des Autors, die im Pauli-Test geforderte Tätigkeit entspräche dem Wesen der Berufsarbeit und würde deshalb nicht nur den Leistungsverlauf, sondern auch die Einstellung zur Arbeit und andere Persönlichkeitseigenschaften diagnostizieren, ist äußerst problematisch. Bei der beruflichen Arbeit dürfte es sich höchstwahrscheinlich nicht nur um eine rein zweckgebundene Tätigkeit handeln, „deren Endziel nicht – wie beim Spiel – in der Tätigkeit als solcher gelegen ist, sondern in der Schaffung eines Werkes jenseits der Tätigkeit“ (Arnold, 1975, S. 20).

Autorinnenkommentar

Der Pauli-Test bevorzugt Probandinnen/Probanden mit guten Kopfrechenfähigkeiten. Zudem spielt hier die Kapazität des Arbeitsgedächtnisses eine Rolle im Hinblick auf die Anzahl der bearbeiteten Aufgaben. Es kann daher meist nicht differenziert werden, ob mögliche schlechte Leistungen auf mangelnde Konzentrationsfähigkeit oder schlichtweg auf die mangelnde Fähigkeit, im Kopf schnell zu rechnen, zurückzuführen sind. Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind/der Jugendliche mit dem Kopfrechnen Probleme hat oder an einer möglichen Rechenschwäche leidet.

Literatur

- Arnold, W. (1958). *Neue Erfahrungen mit dem Pauli-Test. Eine kritische Erwiderung. Zeitschrift für experimentelle und angewandte Psychologie*, 5, 534–541.
- Arnold, W. (1961). *Beiträge zur Faktorenanalyse des Pauli-Tests. Psychologische Beiträge*, 5, 312–327.
- Arnold, W. (1975). *Der Pauli-Test. Anweisung zur sachgemässen Durchführung, Auswertung und Anwendung des Kräpelinischen Arbeitsversuchs (5., korrigierte Auflage)*. Berlin: Springer.
- Bartenwerfer, H. (1963). *Mitteilungen zur Frage der Reliabilität dreier Merkmale des Pauli-Tests. Diagnostica*, 9, 77–79.
- Bartenwerfer, H. (1964). *Allgemeine Leistungstests. In R. Heiss (Hrsg.), Handbuch der Psychologie. Band 6, Psychologische Diagnostik (S. 385–410)*. Göttingen: Hogrefe.
- Bäumler, G. (1964). *Zur Faktorenstruktur der Pauli-Testleistung unter besonderer Berücksichtigung des sogenannten numerischen Faktors. Diagnostica*, 10, 107–119.
- Bäumler, G. & Breitenbach, W. (1970). *Zusammenhänge zwischen Intelligenz, Konzentration, Angst und Leistungsmotivation bei einer studentischen Stichprobe. Psychologie und Praxis*, 14, 37–40.
- Bäumler, G. & Weiss, R. (1967). *Eine Zweifaktorentheorie der nach der TAT-Methode gemessenen Leistungsmotivation (Heckhausen). Psychologie und Praxis*, 11, 23–45.
- Bernstein, J. & Kuenzel, R. (1976). *Vergleich der Intelligenz jugendlicher Cerebralparetiker und jugendlicher Körperbehinderter ohne Hirnschädigung. Sonderpädagogik*, 6, 119–128.
- Bochow, R. (1967). *Erfahrungen mit einer Abwandlungsform des Pauli-Tests (20 x 1 Min.)*. *Zeitschrift für experimentelle und angewandte Psychologie*, 14, 570–599.
- Fischer, M. & Ulich, H. (1961). *Über die Abhängigkeit einer kurzzeitigen Konzentrationsleistung von*

- der Tageszeit bei Kindern und Jugendlichen verschiedenen Alters. *Zeitschrift für experimentelle und angewandte Psychologie*, 8, 282–296.
- Gebauer, T. (1965). Vergleichende Untersuchungen über den Voraussagewert von Aufnahmeprüfung und Testuntersuchung für den Erfolg auf weiterführenden Schulen. In *Schülerkonflikt und Schülerhilfe*. Weinheim: Beltz.
- Gruhnwald, E. & Ulich, H. (1959). Über die Anwendung des Arbeitsversuchs bei psychisch Beeinträchtigten unter besonderer Berücksichtigung einer Untersuchung an taubstummen Jugendlichen. *Zeitschrift für experimentelle und angewandte Psychologie*, 6, 274–292.
- Gutewa, J. (1957). Der Wert psychologischer Tests für die Untersuchung der Wirkung von Medikamenten auf psychische Funktionen. *Psychiatria et Neurologia*, 134, 224–235.
- Heller, K. A. & Perleth, C. (1991). Der Pauli-Test. Anweisung zur sachgemäßen Durchführung, Auswertung und Anwendung des Kräpelinischen Arbeitsversuches nach R. Pauli, herausgegeben von W. Arnold. In K. A. Heller (Hrsg.), *Begabungsdagnostik in der Schul- und Erziehungsberatung* (S. 184–185). Bern: Huber.
- Hitpass, J. (1961). Vergleichende Untersuchung über den Voraussagewert von Aufnahmeprüfung und Testprüfung zur Erfassung der Eignung für die weiterführenden Schulen. *Schule und Psychologie*, 8, 65–71.
- Hitpass, J. (1963). Bericht über eine sechsjährige Bewährungskontrolle von Aufnahmeprüfung und Testprüfung. *Schule und Psychologie*, 10, 211–218.
- Jäger, A. O. (1960). Zum prognostischen Wert psychologischer Eignungsuntersuchungen. *Psychologische Rundschau*, 11, 160–178.
- Kallenbach, K. (1978). Auswirkungen unterschiedlicher Verfahren zum Pauli-Arbeitsversuch. Ein Vergleich von Abwandlungsformen: Kurzform eines Papier-Bleistift-Verfahrens – apparatives Verfahren. *Psychologie und Praxis*, 22, 157–163.
- Kohlmann, T. (1954). Psychologische Untersuchungen mit Rorschach- und Kräpelin-Versuch an vegetativen Neurosen. *Zeitschrift für diagnostische Psychologie und Persönlichkeitsforschung*, 2, 101–126.
- Oehr, A. (1896). Experimentelle Studien zur Individualpsychologie. *Psychologische Arbeiten*, 1, 92–151.
- Pauli, R. (1936). Beiträge zur Kenntnis der Arbeitskurve. *Archiv für die gesamte Psychologie*, 97, 465–532.
- Plössl, P. (1941). Die Arbeitskurve als diagnostisches Hilfsmittel bei Schwererziehbarkeit. *Zeitschrift für angewandte Psychologie*, 61, 2–69.
- Reuning, H. (1959). Why not artifactor analysis? *Journal of the National Institute for Personnel Research*, 7, 192–200.
- Reuning, H. (1960). *Test administrator's manual for continuous letter checking*. Johannesburg: National Institute for Personnel Research.
- Seitz, W. (1971). Über den Zusammenhang von Leistungen im Pauli-Test und diversen Persönlichkeitsmerkmalen. *Psychologische Beiträge*, 13, 221–238.
- Ulich, E. (1958). Neue Erfahrungen mit dem Pauli-Test. *Zeitschrift für experimentelle und angewandte Psychologie*, 5, 108–126.
- Weise, G. (1975). *Psychologische Leistungstests. Ein Handbuch für Studium und Praxis. Band 1: Intel-*

*ligenz – Konzentration – Spezielle Fähigkeiten (Pauli-Test: S. 121–130).
Göttingen: Hogrefe.*

*Weiss, G. (1950). Der charakterologische Arbeitsversuch (Pauli-Test) in
seiner Anwendung bei Blinden. Unveröffentlichte Dissertation, Uni-
versität München.*

*Wulsten, A.-R. & Ulich, E. (1971). Zur Frage der Wiederholungszuverlässig-
keit im Pauli-Test. Psychologische Rundschau, 22, 203–212.*





3.17. Rev.T.; Revisions-Test

Marschner, G. (1972). Revisions-Test (Rev.T.) nach Dr. Berthold Stender. Ein allgemeiner Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit. Göttingen: Hogrefe.

Beschreibung

Der Revisions-Test dient als allgemeiner Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit. Mittels einfacher Kopfrechenaufgaben wird ermittelt, wie die Konzentrationsfähigkeit anhält. Der Test umfasst 15 Testzeilen mit jeweils 44 Items. Die Items setzen sich aus zufällig auf den Aufgabenbögen verteilten Additions- und Subtraktionsaufgaben zusammen. Die im Test verwendeten Zahlen sind einstellig, ebenso das Ergebnis. Die Probandinnen und Probanden haben die Aufgabe, die Ergebnisse auf ihre Richtigkeit zu überprüfen. Die Bewältigung dieser Aufgabe erfolgt unter einem zeitlich vorgegebenen Rahmen. Die Bewertung erfolgt sowohl quantitativ als auch qualitativ.

Der Rev.T. ist ein allgemeiner Leistungstest (i.S. von Bartenwerfer, 1964) zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit sowie zur Ermittlung des Arbeitsstils und Arbeitsverhaltens. Es handelt sich bei dem Verfahren um einen Test mit ausgesprochenem Speed-Charakter, bei dem geistige Routinearbeit unter Zeitdruck verlangt wird. Der Test verlangt die Überprüfung der Ergebnisse einfacher Rechenaufgaben.

Der Autor geht davon aus, dass ein Teil der bei der Aufgabenkontrolle geleisteten neurophysiologischen Aktivierung, als deren Resultat „Aufmerksamkeit“, „Zentrierung“ oder „Vigilanz“ zu beobachten sind, als eine persönlichkeitsgebundene Reaktion angesehen werden muss, also etwas vom Individuum willentlich Beeinflussbares und somit Verantwortbares ist.

Die rein psychometrische Ermittlung der Konzentrationsfähigkeit wird deshalb durch die qualitative Analyse der Arbeitskurve und des Arbeitsverhaltens auch zur Prüfung der „Anstrengungsbereitschaft“ aufgewertet (i.S. der „Strukturtypen des praktischen Handelns und Verhaltens“ nach Herwig & Dirks, 1951). Im Vergleich zum Pauli-Test brauchen beim Rev.T. die Rechengänge nicht an sich durchgeführt werden, sondern es wird nur eine Nachkontrolle verlangt. Die Art der Aufgaben im Rev.T. ist Kindern vom 10. Lebensjahr an sehr geläufig. Bei Jugendlichen und Erwachsenen werden derartige Prüfungen einfacher Additionen bzw. Subtraktionen dann in steigendem Maße automatisiert durchgeführt.

„Damit sind kognitive Anteile bei der Aufgabenbewältigung weitgehend herausgenommen worden. Es ist nun zu vermuten, dass auf diese Weise die kortikalen Anteile an den Erregungsprozessen, die zur Aufmerksamkeitsleistung führen, gegenüber den subkortikalen zurücktreten. Diese Hypothese könnte man auch so formulieren, dass durch Testverfahren mit geringem kognitivem Anspruch der Vigilanzanteil des Aufmerksamkeitsverhaltens stärker in Erscheinung tritt, vergleichbare Intelligenz vorausgesetzt“ (Royl, 1977).

Das Verfahren besteht aus einstelligen Rechenaufgaben, deren Ergebnisse kontrolliert werden sollen. In der Form A handelt es sich dabei um Additionen, in der Form B um Subtraktionen. Der Test umfasst 15 waagrechte Zeilen mit je 44 Aufgaben auf einem DIN A3-Bogen. Die Aufgaben mit falschem Ergebnis sind durch einen senkrechten Strich durch Aufgabe und Lösung, die mit richtigem Ergebnis durch einen kleinen Haken kenntlich zu machen. Für jede zu bearbeitende Zeile wird der Probandin/dem Probanden eine Zeit von 30 Sekunden zugestanden, nach Ablauf dieser Zeit muss sie/er die nächste Aufgabenzeile bearbeiten bzw. den Test beenden.

Richtige und falsche Ergebnisse sind ohne System gemischt, wobei die Verteilung der falschen Ergebnisse in beiden Formen identisch ist.

Vor Beginn des eigentlichen Tests wird das Lösungsverhalten auf der Rückseite des Aufgabensystems an 34 Aufgaben geübt. Sowohl für die Form A als auch für die Form S existiert eine eigene Handanweisung (Form A: Handanweisung 1972; Form S: Handanweisung Teil II, 1980).

Altersgruppe: 9;6–75;0 Jahre

Hinweise zur Durchführung

Jede der beiden Testformen kann als **Einzel- und als Gruppentest** appliziert werden. Jedoch ist es aus untersuchungstechnischen Gründen (verschiedene Instruktionen) nicht möglich, die beiden Formen in einer Gruppe alternierend einzusetzen.

Die Parallelform S ist hauptsächlich für Kontroll- und Wiederholungsuntersuchungen vorgesehen.

Zusätzlich zur Bearbeitungszeit von ca. 7 ½ Minuten muss mit ca. 5 bis 10 Minuten Einführung und Instruktion gerechnet werden. Der genaue Wortlaut der Instruktion ist in den Handanweisungen vorgeschrieben.

Das Verfahren kann von jeder Fachkraft, die mit der Durchführung und Auswertung von Tests vertraut ist, durchgeführt werden. Lediglich die Analyse der Arbeitskurve ist nach Meinung des Autors nur einer Fachpsychologin/einem Fachpsychologen nach längerem Training möglich. Eine Möglichkeit zur Selbstanwendung besteht nicht.

Hinweise zu Auswertung und Interpretation

Normierungen erfolgten für die Form A erstmals 1970/71. Sie wurden 1976 korrigiert und 1980 für höhere Lebensalter erweitert. Für die Form S wurde die Normierung 1980 abgeschlossen. Die Eichung der Form A basiert auf insgesamt 12.000 Fällen, die der Form S auf 5.000 Fällen. Altersnormen liegen für die Form A von 9;6 bis 75 Jahren, für die Form S von 9;6 bis 39;6 Jahren und älter vor. Die schulische und berufliche Vorbildung der dafür verwendeten Eichstichproben umfasst alle Schulsysteme von der Sonderschule bis zur Universität, von Hilfsarbeiterinnen/Hilfsarbeitern bis zu Führungskräften in Betrieben unterschiedlicher Größe.

Für einen besonders qualifizierten Teil der Population (Schüler/innen von Fachoberschulen, technischen und kaufmännischen Fachschulen; Studierende von Fachhochschulen und technischen Hochschulen; technische und kaufmännische Führungs-Nachwuchs-Kräfte und besonders qualifizierte technische Führungskräfte aus mittleren und größeren Betrieben) wurden Qualifikationsgruppen-Normen erstellt, die für beide Testformen für einen Altersbereich von 17;6 bis 39;6 Jahren und älter vorliegen.

Sowohl die Alters- als auch die Qualifikationsgruppennormen wurden ausschließlich an rein männlichen Stichproben gewonnen.

Für die Form A des Rev.T. liegen noch Berufsgruppen-Normen für männliche und weibliche Jugendliche in kaufmännischer Berufsausbildung vor, und zwar von 13;6 bis 15;6 Jahren und älter. Die verwendeten Eichstichproben werden als repräsentativ für die Bundesrepublik Deutschland sowie für Betriebe verschiedener Größe angenommen, solange keine anderen Ergebnisse vorliegen.

Die Ermittlung der Rohwerte erfolgt mit Hilfe einer transparenten Auswertungsschablone, die für beide Testformen gleich ist. Über die Normentafeln in den Handanweisungen können dann Prozentrang und Standardwert-Normen ermittelt werden. Außerdem finden sich Interpretationsbeispiele in den beiden Handanweisungen.

Die Auswertungszeit beträgt je nach Routine der Testleiterin/des Testleiters zwischen 2 und 5 Minuten.

Testgüte

Objektivität

Die Durchführungsobjektivität kann durchaus etwas beeinträchtigt werden, da die Länge der Pausen zwischen den Teilzeiten nicht genau festgelegt ist und so abhängig von der Testleiterin/dem Testleiter unterschiedlich sein dürfte. Die Auswertung von Mengenleistung und Fehlern (quantitative Leistungswerte) ist als objektiv zu betrachten. Die Interpretation der Arbeitskurve (qualitative Auswertung) jedoch erfordert längere Erfahrung einer Fachpsychologin/eines Fach-

psychologen mit dem Rev.T., d.h. dieser Teil der Testauswertung hat mehr „ideographischen“ Charakter und beabsichtigt mehr die „typical performance“ zu ermitteln. Die Objektivität ist hier eher anzuzweifeln.

Reliabilität

Die interne Konsistenz liegt zwischen $\alpha = .90$ und $\alpha = .97$. Die Retest-Reliabilität liegt nach zwei Jahren bei $r_{tt} = .86$, nach drei Jahren bei $r_{tt} = .75$.

Validität

Für das Verfahren wird logische Validität beansprucht. Die Übereinstimmungsvalidität mit anderen allgemeinen Leistungstests beträgt zwischen $42 \leq r \leq .84$. Bei 20 Untersuchungen zur Vorhersagevalidität wurde eine mittlere Korrelation von $r = .64$ ermittelt (Gesamt $n = 1.900$).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Aufgrund der kurzen Durchführungs- und Auswertungszeit und der Anwendungsmöglichkeit als Gruppentest stellt der Rev.T. ein ökonomisches Verfahren dar.

Er soll nicht nur Auskunft über anhaltende Konzentration bei routinemäßiger, geistiger Tempoarbeit („Belastbarkeit“) geben, sondern darüber hinaus auch Erkenntnisse über Arbeitsstil und Arbeitsverhalten („Arbeitskurve“) liefern, d.h. über das Zusammenwirken von Antriebs- und Steuerungskräften, über Leistungsbereitschaft und Leistungsreserven und auch über Veränderungen des Verhaltens im Verlauf der Ausbildung oder einer Therapie.

Dazu ist anzumerken, dass sämtliche Untersuchungen zur Reliabilität und Validität ausschließlich die Leistungsmenge des Rev.T. betreffen, zu Fehlerwerten und Arbeitskurven liegen keine Angaben vor. Die Fehlerwerte besitzen nur eine geringe Aussagekraft und die Interpretation der Arbeitskurve erweist sich als nicht sehr objektiv und differenziert. Sowohl die Reliabilitäts- als auch die Validitätskoeffizienten sind befriedigend. Allerdings sind die Angaben über die verwendeten Stichproben teilweise nicht genügend transparent. Dem Problem von Extremgruppen wurde bei den Berechnungen zu wenig Aufmerksamkeit geschenkt. Für weibliche Probandinnen werden bis auf die Berufsgruppen-Normen keine gesonderten Normen mitgeteilt.

Außerdem wäre bei der Größe der Eichstichproben eine weitere Aufgliederung im unteren Altersbereich wünschenswert und möglich.

Autorinnenkommentar

Dieser Test führt insbesondere bei Kindern im Alter zwischen 9 und 12 Jahren oft zu einer Verzerrung, da Probandinnen/Probanden mit guten Kopfrechenfähigkeiten hier signifikant besser

abschneiden als Kinder/Jugendliche mit Problemen beim Kopfrechnen oder mit Dyskalkulie. Zudem spielt hier die Kapazität des Arbeitsgedächtnisses eine Rolle im Hinblick auf die Anzahl der bearbeiteten Aufgaben. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit oder schlichtweg an der mangelnden Fähigkeit, im Kopf schnell zu rechnen bzw. einer geringen Kapazität des Arbeitsgedächtnisses liegen. Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind/der Jugendliche mit dem Kopfrechnen Probleme hat oder an einer möglichen Rechenschwäche leidet.

Literatur

- Arndt, K., Stender, B. & Marschner, G. (1976). *Revisions-Test (Rev.T.)*. *Diagnostica*, 22, 138–139.
- Bartenwerfer, H. (1964). *Allgemeine Leistungstests*. In R. Heiss (Hrsg.), *Handbuch der Psychologie*. Band 6, *Psychologische Diagnostik* (S. 385–410). Göttingen: Hogrefe.
- Heber, I. A. (2009). *Revisions-Test (Rev. T.)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: *Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 187–196). Göttingen: Hogrefe.
- Herwig, B. & Dirks, H. (1951). *Strukturtypen des praktischen Handelns*. *Psychologische Rundschau*, 2, 20–28.
- Krumm, S., Schmidt-Atzert, L. & Eschert, S. (2008). *Investigating the structure of attention. How do test characteristics of paper-pencil sustained attention tests influence their relationship with other attention tests?* *European Journal of Psychological Assessment*, 24 (2), 108–116.
- Marschner, G. (1960). *Zur Diagnostik eines höheren F % im KLT. Eine vergleichende Untersuchung von Konzentrationsleistungstest und Revisions-Test in der betriebspsychologischen Praxis*. *Diagnostica*, 6, 66–72.
- Marschner, G. (1971). *Der Revisionstest als allgemeiner Leistungstest. Beiträge zur Standardisierung und Validierung des Revisions-Tests nach B. Stender*. *Diagnostica*, 17, 14–26.
- Marschner, G. (1972). *Revisions-Test (Rev.T.) nach Dr. Berthold Stender. Ein allgemeiner Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit*. Göttingen: Hogrefe.
- Royl, W. (1977). *Aus der wissenschaftlichen Begleitung der Gesamtschulversuche in Schleswig-Holstein. Die Rolle der Konzentrationsfähigkeit bei der Bearbeitung von Lernerfolgstests im Fach Deutsch*. In W. Arnold (Hrsg.), *Texte zur Schulpsychologie und Bildungsberatung*. Band 2. Braunschweig: Westermann.
- Schmidt-Atzert, L. & Ising, M. (1997). *Ein Beitrag zur Konstruktvalidität von d2 und Revisionstest*. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18 (4), 241–250.

3.18. SIGNAL; Signal-Detection

Schuhfried, G. (2007). Signal-Detection. Test zur Erfassung der Signalentdeckung (Version 26.05). Mödling: Schuhfried.

Beschreibung

SIGNAL ist ein Test zur Quantifizierung der Aufmerksamkeitsleistungen und der visuellen Differenzierungsleistung. Basierend auf der Signalentdeckungstheorie wird die visuelle Detailerfassung von komplexen Reizanordnungen unter zeitkritischer Bedingung über längere Zeit hinweg erfasst.¹¹

Das vorliegende Verfahren ist vor dem Hintergrund der Signalentdeckungstheorie von Green und Swets (1966) zu sehen, die Phänomene der Wahrnehmung und Entdeckung schwacher Signale vor einem rauschenden, sich verändernden Hintergrund beschreibt. Der Test Signal-Detection gibt jedoch nicht Reize vor, die aufgrund ihrer Schwellennähe eine visuelle Differenzierungsleistung erfordern, sondern Reize, die als Basisleistungen Aufmerksamkeit und Konzentration voraussetzen.

Die allgemeinere Frage der Signaldetektionstheorie kann nach Dorsch (1987, S. 617) auch lauten: „Unter welchen Bedingungen wird das Vorhandensein eines schwachen Signals vor einem Hintergrund von Störsignalen oder gegenüber anderen Signalen, welche gelegentlich mit dem relevanten Signal verwechselt werden können, gemeldet?“. Nach Schuhfried (2007) wird hier die enge Verbindung der Signalentdeckungstheorie mit der Entscheidungstheorie deutlich. Der Prozess der Reaktion „kritisches Signal vorhanden“ oder „kritisches Signal nicht vorhanden“ ist eher ein Problem der Entscheidung als der sensorischen Unterschiedsempfindlichkeit. Die Messung der Leistungsgüte bei dieser Art von Test kann über die Genauigkeit der Bearbeitung erfolgen (Schuhfried, 2007). Dabei sind die beiden falschen Reaktionen „Auslassungsfehler“ (kritisches Signal wird übersehen) und „Falscher

¹¹ vgl. www.schuhfried.at/wiener-testsystem-wts/alle-tests-von-a-bis-z/test/signal-signal-detection/ (Abfragedatum: 27.12.2012)

Alarm“ (Reaktionen ohne Vorhandensein des kritischen Reizes) möglich, die die Testleistung negativ beeinflussen.

Beim computergestützten Test Signal-Detection werden der Versuchsperson über den gesamten Bildschirm verteilt Punkte dargeboten. Ein sogenannter „Schritt“ im Testverlauf lässt einige dieser Punkte verschwinden und fügt andere hinzu (Störreize). Der Bildschirmhintergrund ist nicht „verrauscht“, die Punkte heben sich deutlich vom Hintergrund ab. Als „kritische Reize“ sind von der Probandin/vom Probanden Quadrate bestimmter Größe zu entdecken, die von vier Punkten gebildet werden. Beim Erkennen einer solchen kritischen Punktconstellation ist die schwarze Taste des Panels zu betätigen, als Bestätigung erfolgt ein leiser Piepston. Es wird jedoch kein Feedback über die Richtigkeit des Erkennens oder über ausgelassene kritische Reize gegeben. Der Testverlauf ist in 20 Teilzeiten eingeteilt, die aus jeweils 50 Schritten bestehen.

In die Auswertung gehen als Variablen der Median der Detektionszeit (MDDT), die Anzahl der Richtigen und Verspäteten (R+V), die Anzahl der Richtigen (R), die Anzahl der Verspäteten (V), die Anzahl der Ausgelassenen (A) und die Anzahl der Falschen (F) ein. Die Auswertung kann für den gesamten Bildschirm, für die rechte und linke Bildschirmhälfte oder für die einzelnen Quadranten getrennt erfolgen. Bei getrennter Auswertung wird zusätzlich die Anzahl der Geforderten, d.h. die Anzahl der kritischen Reize im jeweiligen Ausschnitt des Monitors, mit ausgewertet. Die Betrachtung der Ergebnisse der einzelnen Teilzeiten gibt zusätzlich Aufschluss über den Leistungsverlauf.

Altersgruppe: ab 6;0 Jahren

Hinweise zur Durchführung

Aufgrund der computerbasierten Durchführung ist das Verfahren **einzel**n durchzuführen.

Es liegen vier Standardtestformen vor. Die Version Standard S1 bietet weiße Punkte auf schwarzem Grund an, die inverse Version Standard S2 schwarze Punkte auf weißem Grund. Die Standardversion S3 ist durch eine sehr kurze Signaldauer von nur 1,5 Sekunden (statt 3,75 Sekunden) gekennzeichnet, und Version Standard S4 bietet die gleiche Anzahl von kritischen Signalen in beiden Bildschirmhälften an. Letztere eignet sich somit besonders zur Neglect-Diagnostik. Durch das Parameterprogramm sind beliebige weitere Testformen von der Anwenderin/dem Anwender zu erstellen. Variiert werden können u.a. die Parameter „Anzahl der Teilzeiten“, „Anzahl der kritischen Reize pro Teilzeit“, „Länge der kritischen Reize“, „Maximale Punktänderung pro Schritt“ sowie einige Parameter der Übungsphase. Der Test ist nach Schuhfried (2007) einfach in andere Sprachen zu übertragen. Während der Durchführung des Programms kann im Hintergrund das Zusatzprogramm „Physio-Unit“ des Wiener Testsystems eingesetzt werden, das physiologische Marker zu bestimmten Testzeitpunkten erhebt.

Für die Durchführung des Verfahrens sind mit Übungsphase ca. 13 Minuten zu veranschlagen.

Die Instruktion ist vollständig durch die Software vorgegeben. Nach der Erläuterung der Aufgabenstellung folgt ein Übungsprogramm, das mit weniger als vier Fehlern absolviert werden muss. Durch Tastendruck startet die Versuchsperson selbst die Testphase. Zur Durchführung des Verfahrens sind keine besonderen Qualifikationen erforderlich. Die Testleiterin/der Testleiter sollte sich zu Beginn des Tests lediglich vergewissern, dass bei der Probandin/beim Probanden ausreichendes Instruktionsverständnis vorliegt.

Hinweise zu Auswertung und Interpretation

Für die drei Standardtestformen liegen Normen unterschiedlicher Normierungsstichproben vor. Für die Testform S1 gibt es Vergleichswerte von Minenarbeiterinnen und -arbeitern in Südafrika ($n = 105$) sowie von Pilotenbewerberinnen und -bewerbern des österreichischen Bundesheeres ($n = 141$), für S2 liegen Vergleichswerte von Sportlerinnen und Sportlern ($n = 151$) vor. Für die Testform S3 liegen ebenfalls Vergleichswerte von Bewerberinnen/Bewerbern einer Pilotenausbildung beim österreichischen Bundesheer ($n = 482$) vor. Personen- und Leistungsmerkmale dieser Stichproben sind bei Schuhfried (2007, S. 14) aufgeführt, die Normtabellen mit T-Werten und Prozenträngen bei Schuhfried (2007, Anhang).

Die Auswertung erfolgt computergesteuert. Der Ergebnisausdruck liefert die unter Testaufbau genannten Variablen, eine Darstellung der Ergebnisse nach Schnelligkeit und Genauigkeit, Statistiken des Testverlaufs (Teilzeiten), eine graphische Darstellung des Testverlaufs, Quadranten-Graphiken und das Testprotokoll. Ein Beispielausdruck befindet sich bei Schuhfried (2007, Anhang). Darüber hinaus können die Testergebnisse mit Hilfe des Datenkonvertierprogramms im ASCII-Format gespeichert und von Statistikprogrammen weiterverarbeitet werden.

Die computerisierte Auswertung bedeutet einen minimalen Zeitaufwand.

Testgüte

Objektivität

Durch die programmgesteuerte Vorgabe und den automatischen Ablauf ist die Objektivität hinsichtlich Durchführung und Auswertung gewährleistet.

Reliabilität

Die Split-half-Reliabilitäten des Verfahrens wurden anhand dreier Stichproben berechnet. Es handelte sich um eine Stichprobe von 105 Arbeiterinnen/Arbeitern in Südafrika (S1), eine Stichprobe von 151 Sportlerinnen/Sportlern (S2) sowie eine Stichprobe von 482 Bewerberinnen/Bewerbern einer Pilotenausbildung (S3).



Anhand der drei Stichproben berechnete Schuhfried (1993) auch die Interkorrelationen zwischen den Auswertungsvariablen des SIGNAL. Als wesentliches Ergebnis ist festzuhalten, dass zwischen der Detektionszeit (Mittelwert) und „Anzahl der Richtigen“ nur eine sehr geringe Korrelation besteht (Standard 1: $r = -.25$, Standard 2: $r = -.19$), die etwas ansteigt, wenn nur sehr wenig Zeit für das Erkennen des Signals bleibt (Standard 3: $r = -.53$; Schuhfried, 1993, S. 13).

Validität

Dahms und Meissner (1995) prüften die konvergente und diskriminante Validität des Signal-Detection-Tests zusammen mit dem Daueraufmerksamkeitstest, der ebenfalls als Verfahren zur Signalentdeckung zu verstehen ist und die Aufmerksamkeit und ihre Schwankungen über einen längeren Zeitraum misst. Als Stichprobe wurden 44 männliche Wehrpflichtige von 18–23 Jahren herangezogen. Zur Prüfung der konvergenten Validität wurden die Testergebnisse (Reaktionszeit, Treffer, Falscher Alarm) mit dem Wiener Determinationsgerät (WDG; Schuhfried, ohne genaue Angaben) korreliert. Die Ergebnisse sind bei Dahms und Meissner (1995, S. 100) aufgeführt. Sie zeigen eine signifikant positive Korrelation ($r = .44$ bis $r = .62$) zwischen der Anzahl der Treffer aller drei Verfahren. Positiv signifikant korreliert waren auch die Reaktionszeiten von DAUF und SDT (bzw. SIGNAL). Die Zusammenhänge zwischen Treffern und Reaktionszeiten erwiesen sich nur zwischen dem Wiener Determinationsgerät und dem DAUF als positiv signifikant. Für die falschen Alarme aller drei Verfahren ließ sich kein systematischer Zusammenhang nachweisen. Um den faktoriellen Zusammenhang zwischen den Reaktionszeiten, den Treffern und den falschen Alarmen näher zu bestimmen, wurde eine Faktorenanalyse (Hauptachsenmethode, Varimaxrotation) durchgeführt. Auf dem zweiten der drei gefundenen Faktoren laden nur die Reaktionszeit und die falschen Alarme des SIGNAL hoch. Dies weist darauf hin, dass sich der Test Signal-Detection erheblich von den anderen beiden Verfahren unterscheidet (Dahms & Meissner, 1995, S. 101 f.).

Die diskriminante Validität des Verfahrens wurde durch Korrelation mit einigen Untertests der Eignungs- und Verwendungsprüfung (EVP) für Wehrpflichtige der Bundeswehr (Wotho & Rausch, 1988) sowie der Motorischen Leistungsserie (MLS; Schoppe, 1974) überprüft. Zwischen den Signalentdeckungstests (DAUF, SDT bzw. SIGNAL) und den Referenztests fanden sich erwartungsgemäß keine signifikanten Zusammenhänge. Eine anschließende Faktorenanalyse (Hauptachsenmethode, Varimaxrotation) wies ebenfalls deutlich auf die unterschiedlichen Konstrukte hin, die den Signalentdeckungstests einerseits, der Testbatterie EVP und dem MLS andererseits zugrunde liegen. Der Versuch einer Extremgruppenvalidierung wurde von Günther (1986) geleistet. Als Personen mit Berufen, die ein hohes Ausmaß an visueller Differenzierungsleistung erfordern, wurden Steuermänner für Regattaboote getestet. Die Variablen „Anzahl

falscher Reaktionen“ und „Anzahl übersehener kritischer Reize“ wurden mit dem Wettkampfalter (WA) und der Regattaleistung (RL) sowie der Regattaleistung bei auspartialisierendem Wettkampfalter (RL/WA) in Beziehung gesetzt. Die Ergebnisse bestätigen erwartungsgemäß die gute Leistung der Steuermänner.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Mit dem Signal-Detection-Test aus dem Wiener Testsystem liegt ein computergestütztes Verfahren zur Messung von Konzentration und Aufmerksamkeit sowie der visuellen Differenzierungsleistung vor. Positiv hervorzuheben sind die vom Programm vorgegebenen äußerst differenzierten Möglichkeiten der Auswertung, die neben zahlreichen Variablen auch mehrere graphische Darstellungen umfasst. Die Ergebnisse der Validierungsstudie von Dahms und Meissner (1995) weisen den SIGNAL im Vergleich mit dem Wiener Determinationsgerät und dem Daueraufmerksamkeitstest als dasjenige Verfahren aus, bei dessen Bearbeitung die Geschwindigkeitskomponente die geringste Bedeutung hat. Auch Schuhfried (2007) weist darauf hin, dass das Verfahren eher einen Power-Test darstellt. Dahms und Meissner (1995) empfehlen den Einsatz des Verfahrens nur in Längsschnittuntersuchungen, da sie die Standardisierung (der Version für das Wiener Testsystem PC/S) für unzureichend halten. Mit der inzwischen verbreiteten Datenbasis stellt das Programm SIGNAL jedoch ein ökonomisches und brauchbares Instrument dar, das zudem durch das Parameterprogramm recht vielseitig einsetzbar ist. Schuhfried (2007) verweist u.a. auf die Möglichkeit, SIGNAL als Screeningverfahren der Klinischen Psychologie zu nutzen und zu diesem Zweck extrem kurze Testformen zu erstellen. Für eine breitere Anwendung sollten die verschiedenen Testformen des Verfahrens allerdings besser teststatistisch überprüft werden.

Autorinnenkommentar

SIGNAL eignet sich hervorragend als Zusatzinstrument für die Diagnostik von Aufmerksamkeitsdefiziten und kann als Screeningverfahren gut eingesetzt werden. Für eine umfassende AD(H)S-Diagnostik im Schulkontext ist dieses Programm nicht geeignet, liefert aber wertvolle Hinweise auch zu Verarbeitungsgeschwindigkeit und Reaktion.

Literatur

- Dahms, P. & Meissner, A. (1995). *Korrelationsstatistische Ergebnisse zur Validität von zwei computergestützten Signal-Entdeckungs-Tests*. *Diagnostica*, 41 (2), 95–107.
- Dorsch, F. (1987). *Psychologisches Wörterbuch*. (11., ergänzte Auflage). Bern: Huber.
- Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Günther, R. (1986). *Anwendung der statistischen Entscheidungstheorie (SDT) bei der visuellen Sig-*



- nalerkennung von Fahrzeugführern. *Verkehrsmedizin*, 33, 237–242.
- Günther, T. (2009). *Signal-Detection (SIGNAL). Untertest aus dem Wiener Testsystem (WTS)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 130–137). Göttingen: Hogrefe.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Testkurzdarstellung Signal Detection: S. 495–497)*. Lengerich: Pabst.
- Schoppe, K. J. (1974). *Das MLS-Gerät: Ein neuer Testapparat zur Messung feinmotorischer Leistungen*. *Diagnostica*, 20, 43–47.
- Schuhfried, G. (1993). *Signal Detection (Version 4.00)*. Mödling: Schuhfried.
- Schuhfried, G. (2007). *Signal-Detection (Version 26.05)*. Mödling: Schuhfried.
- Schuhfried, G. (2013). *Wiener Testsystem. Computergestützte Verfahren zur Leistungs- und Persönlichkeitsdiagnostik (Katalog)*. Mödling: Schuhfried.
- Schuhfried, G. (2004). *Signal Detection*. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Auflage, S. 723–726). Lengerich: Pabst.
- Wothe, K. & Rausch, K. (1988). *Testhandbuch der Eignungs- und Verwendungsprüfung Wehrpflichtiger (TestHB-EVP). Testaufbau und Testanalyse*. In Bundesminister für Verteidigung (Hrsg.), *Fachliche Arbeitsanweisung für den Psychologischen Dienst der Bundeswehr im Wehrersatzwesen*. Bonn.

3.19. TAP (inkl. KiTAP); Testbatterie zur Aufmerksamkeitsprüfung¹²

Zimmermann, P. & Fimm, B. (1993). Testbatterie zur Aufmerksamkeitsprüfung (TAP). Freiburg: Psytest.

Beschreibung

Die Testbatterie zur Aufmerksamkeitsprüfung erfasst mit computergestützten Testverfahren die unterschiedlichen Komponenten von Aufmerksamkeit, u.a. Alertness, selektive Aufmerksamkeit, geteilte Aufmerksamkeit und Vigilanz. Die Verfahren sind von geringer Komplexität, um die Testleistungen nicht durch sensorische oder motorische Ausfälle, Gedächtnisstörungen, Sprachstörungen oder andere Defizite zu beeinflussen. Das Verfahren wird vorzugsweise in der neuropsychologischen Diagnostik eingesetzt, da Beeinträchtigungen der Aufmerksamkeits- und Konzentrationsleistung eine der häufigsten Schädigungsfolgen bei traumatischen Hirn-schädigungen darstellen. Es findet aber auch in der Diagnostik von Aufmerksamkeitsstörungen bei Kindern und Jugendlichen Einsatz.

Aufmerksamkeit ist im TAP nicht als einheitliche Funktion aufzufassen: Die Kontrolle des Informationsflusses im kognitiven System erfolgt durch eine große Zahl von Teilfunktionen. Sturm, Hartje, Orgass und Willmes (1994, S. 15) unterscheiden folgende vier Komponenten von Aufmerksamkeit:

1. Alertness oder allgemeine Reaktionsbereitschaft: die Fähigkeit, in Erwartung eines Reizes mit hoher Priorität das Aufmerksamkeitsniveau zu steigern und auf diesem erhöhten Niveau zu halten
2. selektive Aufmerksamkeit: die Fähigkeit, sich relevanten Reizen zuzuwenden und sie selektiv zu erfassen
3. geteilte Aufmerksamkeit: die Fähigkeit, Aufmerksamkeit auf zwei oder mehr Reize zugleich zu richten
4. Vigilanz: die Fähigkeit, Aufmerksamkeitsleistungen über längere Zeit aufrecht zu erhalten und nur selten auftretende relevante Reize in unregelmäßigen Abständen zwischen einer Vielzahl irrelevanter Reize zu entdecken.

In der vorliegenden Testbatterie zur Aufmerksamkeitsprüfung werden zu jedem dieser Bereiche Tests vorgegeben. Die Einteilung wird noch weiter differenziert, es werden insgesamt 13, zum Teil hoch spezifische Teilfunktionen erhoben. Die Verfahren sind dabei von geringer

¹² Informationen zur KiTAP und Demoversionen zu TAP und KiTAP finden sich unter www.psytest.net/index.php?page=downloads&hl=de_DE (Abfragedatum: 30.12.2012)

Komplexität, um die Beeinflussung der Testleistungen durch sensorische oder motorische Ausfälle, Gedächtnisstörungen, Sprachstörungen oder andere Defizite so gering wie möglich zu halten.

Die klinische Relevanz des Instruments beruht darauf, dass eine intakte Aufmerksamkeitsfunktion grundlegende Voraussetzung für die allgemeine Leistungsfähigkeit des kognitiven Systems ist. Jede Form von praktischer oder intellektueller Tätigkeit kann bei rascher Ermüdbarkeit und Ablenkbarkeit, reduzierter Bewusstseinselligkeit oder verminderter Konzentrationsfähigkeit erheblich eingeschränkt sein. Das Verfahren bildet bereits einen festen Bestandteil der neuropsychologischen Diagnostik. Beeinträchtigungen der Aufmerksamkeits- und Konzentrationsleistung stellen eine der häufigsten Schädigungsfolgen bei traumatischen Hirnschädigungen dar (Fimm, 1997, S. 25). Ursprünglich zur Erfassung von Aufmerksamkeitsstörungen bei traumatischen Hirnschädigungen oder anderen schädigungsbedingten, degenerativen oder psychopathologischen Funktionsstörungen entwickelt, hat die TAP inzwischen auch in die Diagnostik von Aufmerksamkeitsstörungen bei Kindern und Jugendlichen Einzug gehalten. Den theoretischen Hintergrund des Verfahrens bildet die Multi-Component-Theory von Posner und Boies (1971). Deren Schwachpunkt, die unzureichende Differenzierung des Konstrukts der selektiven Aufmerksamkeit, wurde durch den Einbezug mehrerer Teilfunktionen ausgeglichen (Zimmermann, North & Fimm, 1994, S. 7).

Die Testbatterie setzt sich aus 13 Tests zusammen, die im Folgenden detailliert beschrieben werden. Tab. 5 gibt einen Überblick über die Tests und ihren jeweiligen Messgegenstand. Bei allen Verfahren gilt es, nach dem Erscheinen optischer und/oder akustischer Reize auf dem Computerbildschirm, so schnell wie möglich eine von zwei Reaktionstasten zu betätigen. Bei einigen Testverfahren erscheint in der Mitte des Monitors ein kleines Quadrat, das als Fixationspunkt dient.

(1) Alertness:

Gemessen wird die Reaktionszeit auf einen einfachen, im Zentrum des Monitors dargebotenen visuellen Reiz (Kreuz). Durchgang A wird ohne, Durchgang B mit Vorgabe eines auditiven Warnreizes kurz vor dem Erscheinen des visuellen Reizes durchgeführt. Es ist eine Reaktionsbeschleunigung auf den durch einen Warnreiz angekündigten kritischen Reiz zu erwarten. Die Fähigkeit, das Aufmerksamkeitsniveau in Erwartung eines Reizes zu steigern, d.h. die Differenz zwischen Durchgang A und B, wird als phasische Alertness interpretiert (Becker, Sturm, Willmes & Zimmermann, 1996, S. 3).

(2) Arbeitsgedächtnis:

Das Arbeitsgedächtnis wird als zentrale Instanz für die Steuerung des Informationsflusses angesehen. Die Probandin/der Proband muss kontinuierlich den Informationsfluss durch den Kurzzeitspeicher kontrollieren, da ein Vergleich der auf dem Bildschirm gezeigten Zahl mit einer vorher dargebotenen Zahl gefordert wird. Die Reaktionstaste ist zu bedienen, wenn diese Zahlen identisch sind. Der Test kann in drei Schwierigkeitsabstufungen vorgegeben werden:

1. zweistellige Zahlen, letztgezeigte Zahl als Vergleich; 2. zweistellige Zahlen, vorletztgezeigte Zahl als Vergleich; 3. einstellige Zahlen, vorletztgezeigte Zahl als Vergleich.

Tab. 5: Die Verfahren der TAP und ihre Messbereiche (adaptiert nach Zimmermann & Fimm, 2004, S. 184)

| Test | Messgegenstand |
|---|--|
| <i>Basisleistungen</i> | |
| 1. Alertness | Intrinsische und phasische Alertness |
| 2. Arbeitsgedächtnis | Konzentration |
| 3. Geteilte Aufmerksamkeit | Aufmerksamkeitskapazität |
| 4. Go/No-Go-Test | Selektive Aufmerksamkeit |
| 5. Reaktionswechsel | Flexibilität, Umstellfähigkeit |
| 6. Vigilanztests | Vigilanz, Daueraufmerksamkeit |
| 7. Daueraufmerksamkeit | Daueraufmerksamkeit |
| 8. Visuelles Scanning | Selektive Aufmerksamkeit, Konzentration |
| <i>Spezifische Einzelleistungen</i> | |
| 9. Augenbewegung | Prüfung der Sakkadenlatenz nach links und rechts |
| 10. Gesichtsfeld-/Neglectprüfung | Einschränkungen des Gesichtsfeld, Neglect |
| 11. Inkompatibilität | Interferenzneigung |
| 12. Intermodaler Vergleich | Transmodale Integrationsfähigkeit |
| 13. Verdeckte Aufmerksamkeitsverschiebung | Fähigkeit zur verdeckten Aufmerksamkeit (Posner-Paradigma) |

(3) Geteilte Aufmerksamkeit:

Der Test kann in einer visuellen, einer auditiven oder einer gemischten Variante vorgegeben werden. Bei der visuellen Aufgabe erscheinen in einer Matrix von Punkten an unterschiedlichen Stellen mehrere Kreuze. Die Reaktionstaste ist zu betätigen, wenn vier der Kreuze ein Quadrat bilden. Bei der auditiven Variante ertönt abwechselnd ein hoher und ein tiefer Ton, die Reaktionstaste ist zu betätigen, wenn die regelmäßig alternierende Abfolge unterbrochen wird, d.h. zweimal derselbe Ton dargeboten wird. Bei der gemischten Variante werden die Aufgaben gemeinsam vorgegeben, die Reaktionstaste ist bei beiden Zielreizen zu drücken. Ausgewertet werden die Reaktionszeiten.

(4) Go/No-Go-Test:

Es wird die spezifische Fähigkeit zur Unterdrückung einer nicht-adäquaten Reaktion erhoben. Diese Fähigkeit zur Zurückweisung irrelevanter Reize kann nach Schädigungen des präfronta-

len Kortex defizitär sein. Der Test wird in zwei Varianten durchgeführt, es ist nur auf zuvor als kritisch definierte Reize zu reagieren. Bei der einfacheren Version ist auf einen kritischen Zielreiz (diagonales Kreuz) bei zwei Reizen insgesamt (diagonales u. gerades Kreuz) zu reagieren, bei der schwereren Version auf zwei Zielreize bei fünf vorgegebenen Reizen (zwei Muster aus fünf möglichen).

(5) Reaktionswechsel (Flexibilität):

Neben der Fokussierung auf relevante Reize setzt selektive Aufmerksamkeit die Fähigkeit zum Wechsel des Aufmerksamkeitsfokus voraus. Es werden simultan rechts und links vom Fixationspunkt konkurrierende Reize dargeboten. Es ist jeweils die Taste auf der Seite zu betätigen, auf welcher sich der Zielreiz befindet. Der Test kann in zwei Varianten vorgegeben werden, in der verbalen Variante werden ein Buchstabe und eine Zahl, in der nonverbalen Variante eine runde und eine eckige Form vorgegeben. Es ist entweder durchgängig auf einen der Reize oder abwechselnd auf beide zu reagieren. Es wird ein Differenzmaß gebildet (Differenz der Reaktionszeiten im Durchgang mit und ohne Handwechsel), das als Maß der Flexibilität gilt.

(6) Vigilanztest:

Die Daueraufmerksamkeit wird dadurch erhoben, dass in einer alternierenden Folge von Reizen ein unregelmäßig auftauchender kritischer Reiz entdeckt werden muss. Der Test liegt in vier Varianten vor (eine akustische, zwei optische und eine optisch-akustische). Bei den auditiven Tests sind in einer alternierenden Abfolge von hohen und tiefen Tönen Unregelmäßigkeiten zu entdecken, bei den visuellen Varianten große Ausschläge eines vertikal oszillierenden Lichtbalkens oder das zweimalige Auftauchen eines Musters in derselben Hälfte eines Quadrats. In der visuell-auditiven Variante ist auf zwei kritische Reize zu reagieren, ein hoher und ein tiefer Ton, jeweils gefolgt von einem bestimmten Buchstaben. Die maximale Länge des Tests beträgt 50 Minuten. Der Unterschied zwischen den Reaktionszeiten der ersten und der zweiten Testhälfte wird als Maß der Daueraufmerksamkeit herangezogen.

(7) Daueraufmerksamkeit:

Bei diesem Test erscheinen auf dem Bildschirm nacheinander Reize, die in verschiedenen Merkmalsdimensionen (Farbe, Form, Größe, Füllung) variieren. Ein kritischer Reiz liegt dann vor, wenn ein gezeigter Reiz in einer oder in einer von zwei vorher festgelegten Reizdimensionen mit dem vorangehenden Reiz übereinstimmt (z.B. die gleiche Form bei unterschiedlicher Farbe, Größe und Füllmuster aufweist). Um die Schwierigkeit der Aufgaben an das Leistungsniveau einer Testperson anzupassen, kann zwischen den Schwierigkeitsstufen „Form“ und „Farbe oder Form“ gewählt werden.

(8) Visuelles Scanning:

Dieser Test misst die Fähigkeit zum visuellen Abtasten des Gesichtsfeldes. Es wird eine Anordnung von Reizen (Rechtecke, die an einer der vier Seiten geöffnet sind) in einem Matrixmuster (5 x 5 = 25 Rechtecke) vorgegeben. Bei Entdeckung eines kritischen Reizes (an der oberen

Seite geöffnetes Rechteck) ist die linke Reaktionstaste zu bedienen, wenn kein kritischer Reiz in der Matrix enthalten ist, die rechte Taste.

(9) Augenbewegung:

Eine wichtige Funktion der selektiven Informationsaufnahme ist die Ausrichtung der Augen auf einen relevanten Gesichtsfeldausschnitt. Die Probandin/der Proband wird aufgefordert, nur nach rechts oder links zu sehen, wenn dort ein Reiz erscheint. Zentral, links und rechts vom Fixationspunkt werden unregelmäßig Rechtecke vorgegeben. Die Reaktionstaste ist zu betätigen, wenn die obere Seite eines Rechtecks offen ist.

(10) Gesichtsfeld- bzw. Neglectprüfung:

Der Test prüft in einem groben Scanning auf Gesichtsfeldausfälle. Während der Gesichtsfeldprüfung wird in der Mitte des Monitors ein Rechteck, darin ein Buchstabe, vorgegeben. Dieser Buchstabe wechselt und ist von der Probandin/vom Probanden jeweils laut zu benennen. Die Reaktionstaste ist zu betätigen, wenn auf dem Monitor zusätzlich rasch wechselnde Zahlen erscheinen, die als flackernder Reiz wahrgenommen werden. Diese Reize erscheinen in Zufallsposition in zufällig variierenden Intervallen. Während der Neglectprüfung ist der Bildschirm zusätzlich permanent mit Zahlen angefüllt; die schnell wechselnden Zahlen erscheinen in den Lücken dieses Zahlenteppichs. Die Auswertung erfolgt getrennt für die vier Gesichtsfeldquadranten sowie die beiden Gesichtsfeldhälften.

(11) Inkompatibilität:

Diese Aufgabe erzeugt durch die Nicht-Übereinstimmung von räumlicher Position des kritischen Reizes und der reagierenden Hand eine Reiz-Reaktionsinkompatibilität (S-R-Inkompatibilität nach Fitts und Seeger, 1953) und prüft die Interferenzneigung bei konkurrierenden Informationen. Es werden Kreuze links oder rechts vom Fixationspunkt dargeboten. Kurz zuvor erscheinen nach links bzw. rechts gerichtete Pfeile. Wenn das Kreuz auftritt, soll je nach Pfeilrichtung mit der linken oder rechten Hand reagiert werden, und zwar unabhängig von der Seite der Kreuzpräsentation. Die Reaktionszeiten in den inkompatiblen Bedingungen werden mit denjenigen in den kompatiblen Bedingungen verglichen.

(12) Intermodaler Vergleich (Crossmodale Integration):

Dieser frühe Prozess gerichteter Aufmerksamkeit stellt die Kontrolle des Inputs aus verschiedenen sensorischen Kanälen sicher. Eine Be-



einrächtigung der supramodalen Kontrolle wird erhoben, indem kritische Reize vorgegeben werden, bei denen simultan ein optischer und ein akustischer Reizaspekt gleichzeitig beachtet werden muss. Durch die Struktur der Aufgaben ist sichergestellt, dass es zu keiner strukturellen Interferenz zwischen den Informationskanälen kommt. Die Reaktionstaste ist zu drücken, wenn Tonhöhe (hoher vs. tiefer Ton, akustischer Reiz) und Pfeilrichtung (hinauf vs. hinunter, visueller Reiz) übereinstimmen. Andernfalls ist keine Taste zu bedienen.

(13) Verdeckte visuelle Aufmerksamkeitsverschiebung:

Der Fokus der visuellen Aufmerksamkeit kann ohne Augenbewegung verlagert werden. Rechts oder links vom Fixationspunkt werden einfache Reize dargeboten, auf deren Erscheinen so schnell wie möglich durch Tastendruck reagiert werden soll. Vor dem kritischen Reiz erscheint in der Mitte des Bildschirms ein Hinweisreiz (Pfeil), der mit hoher Wahrscheinlichkeit (80 %) nach der Seite zeigt, auf welcher der Reiz erscheint, (valider Hinweisreiz), in 20 % der Darbietungen jedoch nicht („invalid“ Hinweisreiz). Zeitliche Differenzen zwischen Reaktionszeiten auf valide und invalide Hinweisreize können als Zeitbedarf für die verdeckte Aufmerksamkeitsverschiebung gelten.

Bei allen Tests kann die Anzahl der Reizvorgaben bzw. die Länge des Tests variiert werden. Die Testleiterin/der Testleiter startet den Vor- und Hauptversuch. Die Auswertung erfolgt computergesteuert, es werden Reaktionszeiten, Fehler, richtige Reaktionen, Auslassungen und Differenzmaße als Kennwerte einbezogen.

Altersgruppe: Je nach Test liegen Normen für Kinder/Jugendliche und/oder Erwachsene vor (teilweise Normen für 6;0–9;0 Jahre).

Hinweise zur Durchführung

Das Verfahren ist als **Einzeltestung** durchzuführen.

Im TAP-Programm können je nach Test die Anzahl der Durchgänge oder die Testdauer, die Art und Anzahl der Zielreize sowie der Schwierigkeitsgrad gewählt werden. Parallelformen gibt es nicht.

Zur Dauer der Testdurchführung liegen keine Angaben vor.

Jedem der Tests ist eine computergesteuert vorgegebene Instruktion vorangestellt. Die Testleiterin/der Testleiter wählt zwischen wiederholbarem Vor- und Hauptversuch. Die Testleiterin/der Testleiter sollte sich vor dem Hauptversuch vergewissern, dass die Testperson die Instruktion verstanden hat. Földényi, Giovanoli, Tagwerker-Neuenschwander, Schallberger und Steinhäuser (1999, S. 90) zufolge sind die Instruktionen für einen Einsatz bei Kindern zu modifizieren, da sie sich für diese Zielgruppe als wenig geeignet erwiesen haben.

Auf Seiten der Probandinnen/Probanden ist ein im zentralen Bereich intaktes Gesichtsfeld für viele der Tests Voraussetzung. Auf Seiten der Testleiterin/des Testleiters sind keine besonderen Qualifikationen erforderlich.

Hinweise zu Auswertung und Interpretation

Für die 13 Tests liegen Normwerte aus unterschiedlichen Stichproben vor. Die Möglichkeit zum Einfügen aktualisierter Normen in die Computerversion ist vorgesehen, die Benutzerin/der Benutzer wird über das Vorliegen aktualisierter Normen informiert.

Eine Normierungsstudie für sieben Untertests wurde 1988/89 an einer Stichprobe von $n = 200$ gesunden Probandinnen/Probanden (Alter von 20 bis 60 Jahren) durchgeführt. Die Stichprobe ist bei Becker et al. (1996, S. 8) beschrieben, über die Studie (mit $n = 210$) berichten auch Sturm und Willmes (1994). Es waren die Tests Alertness, Gesichtsfeldprüfung, Neglect-Prüfung, Geteilte Aufmerksamkeit, Go/No-Go-Test, Vigilanz akustisch, Vigilanz optisch miteinbezogen.

Als Normwertskalen für die nicht-normalverteilten (rechtsschiefen) Rohwerte wurden Prozentränge und T-Werte ermittelt. Die Normen liegen für die intraindividuellen Reaktionszeitmediane, die relevanten Differenzbeträge und die intraindividuellen Standardabweichungen vor. Da bei allen Untertests Fehler nur sehr selten auftraten, konnten als Normangaben für die Fehlerhäufigkeiten lediglich Cut-off-Werte oder grobe Prozentrangklassen ermittelt werden. Hinsichtlich der Reaktionszeiten zeigten sich bei allen sieben Untertests signifikante ($p < .01$) Altersunterschiede. Anhand einer linearen Regressionsanalyse wurden Alterskorrekturwerte ermittelt. Signifikante Geschlechtsunterschiede traten nicht auf. Die Schulbildung führte lediglich im Go/No-Go-Test zu einem signifikanten Unterschied (Becker et al., 1996, S. 9).

Für einige Subtests liegen Normen für folgende Altersbereiche vor:

- Anhand einer Stichprobe von $n = 150$ Kindern zwischen 6 und 10 Jahren ermittelten Földényi et al. (1999, S. 96–98) Prozentränge für drei Altersgruppen bezüglich der Untertests Alertness, Geteilte Aufmerksamkeit, Inkompatibilität, Visuelles Scanning, Go/No-Go-Test, Reaktionswechsel und Vigilanztest. Földényi et al. (1999) halten geschlechtsspezifische Normen für erforderlich, da sich in dieser Altersgruppe im Gegensatz zur Erwachsenenstichprobe bei Sturm und Willmes (1994) auch Geschlechtsunterschiede zeigten. Die Jungen erreichten schnellere Reaktionszeiten, während die Mädchen deutlich weniger Fehler machten. Computervorerfahrungen wirkten sich nicht signifikant auf die Testleistungen aus.
- Kunert, Derichs und Irlle (1996, S. 108–113) führen Prozentränge für die Untertests Geteilte Aufmerksamkeit, Go/No-Go, Inkompatibilität, Reaktionswechsel und Visuelles Scanning auf, die auf einer Stichprobe von $n = 187$ Kindern basieren (9–12 Jahre; drei Altersgruppen; Erhebung 1992–1994). Auch für die Kinder-Stichproben von Földényi et al. (1999) und von Kunert et al. (1996) wurden signifikante Alterseffekte auf die Testleis-



tungen nachgewiesen.

- Für den Untertest „Alertness“ erstellten Bodenburger, Popp und Kawski (2001) ergänzende Normdaten für die Altersgruppe über 60 Jahren ($n = 50$ Probandinnen/Probanden; Quartilnormen und Prozentränge für Mittelwerte, Streuungen, Mediane der Durchgänge mit und ohne Warnton), mit zunehmendem Alter stiegen die Reaktionszeiten an (Datenerhebung 1993–2000).

Die Auswertung erfolgt computergesteuert und ist daher nur mit einem minimalen Zeitaufwand verbunden. Die Ergebnisse können graphisch oder

als Statistiken auf dem Monitor wiedergegeben, ausgedruckt oder auf Datenträgern gespeichert werden.

Als Parameter werden ausgewertet:

- Reaktionszeiten
- Anzahl der Fehler
- Anzahl der richtigen Reaktionen
- Verzögerte Reaktionen, d.h. Ausreißer (Reaktionszeiten länger als die mittlere Reaktionszeit plus 2,3-mal die Standardabweichung)

Zusätzlich werden bei einigen Tests Differenzmaße gebildet.

Ein ausführliches Fallbeispiel, an dem die differenzierte Aufmerksamkeitsprüfung und deren Interpretation anschaulich dargestellt wird, findet sich bei Fimm (1997).

Testgüte

Objektivität

Die Objektivität des Verfahrens ist durch die computergesteuerte Instruktion, Durchführung und Auswertung gewährleistet.

Reliabilität

Sturm und Willmes (1994) ermittelten die Split-half-Reliabilität (Odd-even-Methode) von sieben TAP-Tests an einer Stichprobe von $n = 210$ Erwachsenen (105 Frauen, 95 Männer) im Alter von 20–60 Jahren. Mit Ausnahme des Untertests „Geteilte Aufmerksamkeit“ wiesen die Verfahren gute bis sehr gute Reliabilitätskoeffizienten auf. Földényi, Giovanoli, Tagwerker-Neuenschwander, Schallberger und Steinhausen (2000) ermittelten anhand einer Stichprobe von $n = 95$ Kin-

dern (Alter 7–10 Jahre, 52 Mädchen, 43 Jungen) die Retestreliabilität einiger TAP-Tests, wobei das Testintervall zwischen 16 und 81 Tagen variierte. Zum Vergleich wurden die Daten einer Erwachsenenstichprobe ($n = 36$, Testintervall zwischen 18 und 43 Tagen) von Zimmermann und Fimm (1993) herangezogen.

Validität

a) Konstruktvalidität

Becker et al. (1996, S. 10) führten anhand der Normierungsstichprobe eine hierarchische Clusteranalyse (Complete-Linkage-Kriterium) auf der Basis der Produkt-Moment-Korrelationen aller Paare von Testvariablen durch, um zu ermitteln, inwieweit das Reaktionsverhalten der Probandinnen/Probanden sich bei den unterschiedlichen Tests ähnelt. Die höchsten Koeffizienten finden sich zwischen den einzelnen Variablen eines Testverfahrens, während die Korrelationen zwischen den Variablen verschiedener Verfahren deutlich geringer ausfallen. Entsprechend der Testkonzeption unterscheiden sich die Vigilanzaufgaben stark von den anderen Testaufgaben, aber auch die übrigen Untertests zeigen jeweils eigenständige Reaktionsmuster. Auch ein non-metrisches multidimensionales Skalierungsverfahren, die Smallest-Space-Analyse nach Guttman und Lingoes (Lingoes, 1973), bestätigte die Konzeption der TAP. Das zweidimensionale Ergebnis (Alienationskoeffizient = .06) zeigt, dass die zu einem Untertest oder einem dem Konzept nach ähnlichen Aufmerksamkeitsbereich zählenden Verfahren räumlich nahe beieinander liegen (Becker et al., 1996, S. 11).

b) Faktorielle Validität

Kunert et al. (1996) führten eine Faktorenanalyse (Hauptkomponentenanalyse mit Varimaxrotation) mit den Reaktionszeiten der Untertests Geteilte Aufmerksamkeit, Go/No-Go-Test, Inkompatibilität, Reaktionswechsel und Visuelles Scanning durch ($n = 61$ Kinder, 9–12 Jahre). Dabei wurden folgende drei Faktoren extrahiert (Eigenwertkriterium), die 67,5 % Varianz aufklärten:

- Faktor 1: Alertness, Go/No-Go, Intermodaler Vergleich, Reaktionswechsel, Arbeitsgedächtnis
- Faktor 2: Vigilanz, Geteilte Aufmerksamkeit, Inkompatibilität
- Faktor 3: Reaktionswechsel und Visuelles Scanning

Eine zweite Faktorenanalyse ohne den Subtest Alertness ($n = 99$) führte zu einer ähnlichen Dreifaktorenlösung mit 58,4 % Varianzaufklärung. Beide Ergebnisse entziehen sich Kunert et al. (1996, S. 101) zufolge einer einfachen Interpretation. Die Autorinnen/Autoren schlussfolgern, dass eine Differenzierung der Aufmerksamkeitskomponenten, die sich auch in einer differenzierten Faktorenstruktur niederschlägt, erst im Erwachsenenalter stattfindet.

Im Gegensatz dazu stehen die Ergebnisse von Földényi et al. (1999, S. 95), die eine Differenzierung der Aufmerksamkeitsfunktionen in einer Stichprobe mit $n = 67$ sechs- bis zehnjährigen Kindern durchaus nachweisen konnten (Hauptkomponentenanalyse mit Varimaxrotation). Dabei wurden jedoch neben den Reaktionszeiten auch die Fehler und Auslassungen als Kennwerte

der Subtests Alertness, Geteilte Aufmerksamkeit und Go/No-Go-Test einbezogen. Die gefundene Dreifaktorenlösung mit folgenden Faktoren klärte 72 % der Gesamtvarianz auf:

- Faktor 1: Fehler/Geteilte Aufmerksamkeit; Fehler und Auslassungen/Go/No-Go-Test; eine hohe negative Ladung weisen auf diesem Faktor die Reaktionszeitmediane im Go/No-Go-Test auf.
- Faktor 2: Reaktionszeiten des Alertness-Tests
- Faktor 3: Reaktionszeiten und Auslassungen im Subtest Geteilte Aufmerksamkeit

Földényi et al. (1999, S. 100) interpretieren die drei Faktoren als Ausdruck der Aufmerksamkeitskomponenten Wachheit, Selektive Aufmerksamkeit und Geteilte Aufmerksamkeit. Ergebnisse zu Faktorenanalysen der Daten einer Erwachsenenstichprobe liegen nicht vor.

c) Externe Validität

Földényi, Imhof und Steinhausen (2000) gaben einer Stichprobe von $n = 20$ Kindern mit Aufmerksamkeits-/Hyperaktivitätsstörungen und $n = 20$ männlichen Kontrollkindern (Alter 7,5–13,5 Jahre) die TAP vor. Die Kinder mit Aufmerksamkeits-/Hyperaktivitätsstörungen zeigten verglichen mit der Kontrollgruppe mehr Fehler und Auslassungen sowie erhöhte intraindividuelle Streuungen der Reaktionszeiten in den Subtests. Allerdings bearbeiteten die hyperaktiven Kinder die Subtests nicht generell langsamer. Durch die Reaktionszeit im Go/No-Go-Test und die Fehlerzahl im Test Reaktionswechsel konnten unter Berücksichtigung des Lebensalters 90 % der Kinder beider Gruppen richtig klassifiziert werden. Die einzelnen Testmaße klassifizierten zwischen 70 % und 75 % der Kinder mit Aufmerksamkeits-/Hyperaktivitätsstörungen und 72 % bis 85 % der Kontrollkinder richtig (Földényi, Imhof & Steinhausen, 2000, S. 160). Zwischen den TAP-Testwerten und den Aufmerksamkeitseinschätzungen der Eltern und der Lehrer/innen ergaben sich einige signifikante Zusammenhänge, die jedoch nicht sehr hoch ausfielen. Hohe Korrelationen wurden lediglich zu den Reaktionszeitschwankungen ($r = .61$) und den Fehlern ($r = .57$) im Go/No-Go-Test nachgewiesen (S. 164). Die Ergebnisse sprechen für die klinische Validität des Verfahrens zur Untersuchung von Kindern mit Aufmerksamkeits-/Hyperaktivitätsstörungen.

- Noterdäme, Sitter, Mildenberger und Amorosa (2000) untersuchten $n = 52$ sprachbeeinträchtigte und $n = 52$ normalentwickelte Kinder mit der TAP. Aufmerksamkeitsstörungen gehören zu den häufigsten begleitenden Verhaltensauffälligkeiten bei Kindern mit gestörter Sprachentwicklung. Die Kinder mit gestörter Sprachentwicklung zeigten erwartungsgemäß deutlich schlechtere Leistungen, d.h. längere Reaktionszeiten, größere Streuung der individuellen Reaktionszeiten und signifikant mehr Fehler. Es konnten vier Cluster von unterschiedlichen Aufmerksamkeitsprofilen identifiziert werden.
- Sprengelmeyer, Zimmermann, Lange und Hömberg (1994) verglichen $n = 20$ Patientinnen/Patienten mit Chorea Huntington mit $n = 27$ gesunden Kontrollpersonen. Die Patientinnen/Patienten zeigten signifikant längere Reaktionszeiten in den Untertests Vigilanz, Go/No-Go-Test, Geteilte Aufmerksamkeit, Alertness-Test und Reaktionswechsel, aber

nicht signifikant mehr Fehler. Im Test Intermodaler Vergleich machte die Patientengruppe darüber hinaus mehr Fehler. Im Alertness-Test konnten die Huntington-Patientinnen/Patienten ihr Aufmerksamkeitslevel im Durchgang mit Warnton gegenüber dem ohne Warnton ebenso wie die Kontrollgruppe steigern.

- Sturm et al. (1994) evaluierten Aufmerksamkeitstrainings mit Hilfe der TAP. Es wurden $n = 22$ links- und $n = 16$ rechtshemisphärisch geschädigte Patientinnen/Patienten mit vaskulärer Ätiologie einbezogen, die die TAP-Tests Alertness, Selektive Aufmerksamkeit, Geteilte Aufmerksamkeit und Optische Vigilanz bearbeiteten. Das Training (14 Stunden computergestütztes Training) wurde sukzessiv in den beiden jeweils am stärksten beeinträchtigten Bereichen durchgeführt. Signifikante spezifische, d.h. durch das jeweilige Training bedingte Trainingseffekte konnten für alle vier Untertests nachgewiesen werden. In den komplexen Aufmerksamkeitsfunktionen „Geteilte Aufmerksamkeit“ und „Selektive Aufmerksamkeit“ wurde die Reaktionsschnelligkeit jedoch auch durch das Training elementarer Funktionen (Alertness und Selektive Aufmerksamkeit bei „Geteilter Aufmerksamkeit“; Alertness bei „Selektiver Aufmerksamkeit“) verbessert (S. 26).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Die Testbatterie zur Aufmerksamkeitsprüfung kann als sehr gut evaluiertes Verfahren zur differenzierten Erfassung von Aufmerksamkeitsstörungen gelten. Die computergestützten Testverfahren geben Reaktionsparadigmen von geringer Komplexität vor, in denen die Testperson selektiv auf optische oder akustische Reize durch Tastendruck reagieren soll. Abgesehen vom Untertest „Geteilte Aufmerksamkeit“ fallen die Reliabilitätskoeffizienten der TAP sehr gut aus. Auch zur Validität des Verfahrens liegen hinreichende empirische Belege vor. Normdaten für die Altersgruppe über 60 Jahren wurden bislang durch die Studie von Bodenbun et al. (2001) lediglich für den Untertest Alertness erstellt, weitere Normen für diese Altersgruppe wären wünschenswert. Insgesamt eignet sich die TAP für die neuropsychologische Diagnostik von Aufmerksamkeitsstörungen sehr gut. Was allerdings fehlt, ist ein Manual, das die verstreut publizierten Befunde und Normen in sich vereinigt und damit leichter zugänglich machen würde.

Exkurs: Die KiTAP

Für die KiTAP wurden die Aufgaben der TAP in kindergerechte Situationen eingebaut. Beispielsweise soll bei der Go/No-Go-Aufgabe eine Vampirfledermaus so schnell wie möglich verjagt werden, bevor sie einer Katze wehtut.¹³ Die Untertests der Kinderversion „KiTAP“ sind in Tab. 6 dargestellt.

¹³ Demoversionen zu KiTAP unter www.psytest.net/index.php?page=downloads&hl=de_DE (Abfragedatum: 30.12.2012)

Tab. 6: Untertests der Kinderversion der TAP „KiTAP“ (entnommen aus Zimmermann & Fimm, 2004, S. 199)

| | Test | Kindgerechte Bezeichnung |
|----|-------------------------|---|
| 1. | Ablenkbareit | „Das fröhliche und das traurige Gespenst“ |
| 2. | Alertness | „Die Hexe“ |
| 3. | Daueraufmerksamkeit | „Der Tanz der Geister“ |
| 4. | Flexibilität | „Das Haus der Drachen“ |
| 5. | Geteilte Aufmerksamkeit | „Die Eulen“ |
| 6. | Go/No-Go | „Die Fledermaus“ |
| 7. | Vigilanz | „Der Spiegel“ |
| 8. | Scanning | „Der Ausflug der Hexen“ |

Autorinnenkommentar

Die TAP eignet sich sehr gut für eine umfassende Abklärung eines möglichen Aufmerksamkeitsdefizits bei Jugendlichen und mit Einschränkungen auch bei Kindern. Bei diesem Testverfahren kann bei Jugendlichen eine mangelnde Arbeitseinstellung oder fehlende Motivation zu Fehleinschätzungen führen. Für kleinere Kinder bietet sich alternativ die KiTAP an. Es empfiehlt sich, die KiTAP in mehreren Sitzungen durchzuführen, um Müdigkeitseffekte oder ein Absinken der Motivation zu vermeiden.

Literatur

- Beck, L., Heusinger, A., Boecker, M., Niemann, H. & Gauggel, S. (2008). Convergent and predictive validity of two computerized attention tests in brain-damaged patients. *Zeitschrift für Neurologie*, 19 (4), 213–222.
- Becker, M., Sturm, W., Willmes, K. & Zimmermann, P. (1996). Normierungsstudie zur Aufmerksamkeitstestbatterie (TAP) von Zimmermann und Fimm. *Zeitschrift für Neuropsychologie*, 7 (1), 3–15.
- Bodenburg, S., Popp, B. & Kawski, S. (2001). Ergänzende Normdaten zu dem Untertest Alertness aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP) in der Altersgruppe 60+. *Zeitschrift für Neuropsychologie*, 12 (2), 125–130.
- Bühner, M. (2009). Alertness. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 59–66). Göttingen: Hogrefe.

- Bühner, M. (2009). Arbeitsgedächtnis. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 797–803). Göttingen: Hogrefe.
- Bühner, M. (2009). Gesichtsfeld-/Neglectprüfung. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 138–143). Göttingen: Hogrefe.
- Bühner, M. (2009). Geteilte Aufmerksamkeit. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Aufmerksamkeit, Gedächtnis und exekutive Funktionen* (S. 265–272). Göttingen: Hogrefe.
- Bühner, M. (2009). Go/No-Go-Test. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 224–229). Göttingen: Hogrefe.
- Bühner, M. (2009). Inkompatibilität. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 197–203). Göttingen: Hogrefe.
- Bühner, M. (2009). Intermodaler Vergleich. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 259–264). Göttingen: Hogrefe.
- Bühner, M. (2009). Reaktionswechsel (Flexibilität). Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 289–295). Göttingen: Hogrefe.
- Bühner, M. (2009). Verdeckte Aufmerksamkeitsverschiebung. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 124-129). Göttingen: Hogrefe.
- Bühner, M. (2009). Vigilanztest. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 110–116). Göttingen: Hogrefe.
- Bühner, M. (2009). Visuelles Scanning. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 117–123). Göttingen: Hogrefe.
- Bühner, M., Ziegler, M., Bohnes, B. & Lauterbach, K. (2006). Übungseffekte in den TAP Untertests Test Go/No-Go und Geteilte Aufmerksamkeit sowie dem Aufmerksamkeitsbelastungstest (d2).

- Zeitschrift für Neuropsychologie*, 17 (3), 191–199.
- Drechsler, R. (2009). Testbatterie zur Aufmerksamkeitsprüfung TAP (Version 2.1). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 296–310). Göttingen: Hogrefe.
- Drechsler, R. (2009). Testbatterie zur Aufmerksamkeitsprüfung TAP-M (Version Mobilität 1.2). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 311–319). Göttingen: Hogrefe.
- Dreisörner, T. & Georgiadis, J. (2011). Sensitivität und Spezifität computergestützter Verfahren zur Diagnostik von Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) im Kindes- und Jugendalter □ Die Testbatterie zur Aufmerksamkeitsprüfung (TAP) und Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP). *Empirische Sonderpädagogik*, 3 (1), 3–19.
- Echterhoff, J., Golzarandi, A.G., Morsch, D., Lehmkühl, G. & Sinzig, J. (2009). Ein Vergleich computergestützter Testverfahren zur neuropsychologischen Diagnostik bei Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung. *Zeitschrift für Neuropsychologie*, 20 (4), 313–325.
- Eikermann, A., Petermann, F. & Daseking, M. (2009). Aufmerksamkeitsstörungen nach Schlaganfällen im Kindesalter. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 36 (6), 419–426.
- Fimm, B. (1997). Mikroanalyse von Aufmerksamkeitsprozessen. In S. Gauggel & G. Kerkhoff (Hrsg.), *Handbuch der Klinischen Neuropsychologie. Praxis der Neurorehabilitation* (S. 25–38). Göttingen: Hogrefe.
- Fitts, P. M. & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46, 199–210.
- Földényi, M., Giovanoli, A., Tagwerker-Neuenschwander, F., Schallberger, U. & Steinhausen, H.-C. (1999). Die Aufmerksamkeitsleistungen von 6-10jährigen Kindern in der TAP. *Zeitschrift für Neuropsychologie*, 10 (2), 87–102.
- Földényi, M., Giovanoli, A., Tagwerker-Neuenschwander, F., Schallberger, U. & Steinhausen, H.-C. (2000). Reliabilität und Retest-Stabilität der Testleistungen von 7-10jährigen Kindern in der computerunterstützten TAP. *Zeitschrift für Neuropsychologie*, 11 (1), 1–11.
- Földényi, M., Imhof, K. & Steinhausen, H.-C. (2000). Klinische Validität der computerunterstützten TAP bei Kindern mit Aufmerksamkeits-/Hyperaktivitätsstörungen. *Zeitschrift für Neuropsychologie*, 11 (3), 154–167.
- Kunert, H. J., Derichs, G. & Irle, E. (1996). Entwicklung von Aufmerksamkeitsfunktionen im Kindesalter: Ergebnisse einer vorläufigen Normierung der computergestützten Testbatterie zur Aufmerksamkeitsprüfung (TAP) an 9- bis 12jährigen Kindern. *Zeitschrift für Neuropsychologie*, 7 (2), 92–113.
- Lingoes, J. C. (1973). *The Guttman-Lingoes Nonmetric Program Series*. Ann Arbor: Mathesis Press.
- Noterdäme, M., Sitter, S., Mildenerger, K. & Amorosa, H. (2000). Differenzierung der Aufmerksamkeitsleistung bei sprachentwicklungsgestörten Kindern anhand der Testbatterie zur Aufmerksamkeitsprüfung. In W. Maier, R. R. Engel & H.-J. Möller (Hrsg.), *Methodik von Verlaufs- und Therapiestudien in Psychiatrie und Psychotherapie* (S. 204–207). Göttingen: Hogrefe.
- Posner, M. I. & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78, 391–408.
- Sprengelmeyer, R., Zimmermann, P., Lange, H. & Hömberg, V. (1994). Attention disorders in extrapy-

- ramidal disease: A preliminary report. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 31–40). Tübingen: Gunter Narr.
- Sturm, W. & Willmes, K. (1994). A normative study on the European Attention Test Battery. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 17–20). Tübingen: Gunter Narr.
- Sturm, W., Hartje, W., Orgass, B. & Willmes, K. (1994). Effektivität eines computergestützten Trainings von vier Aufmerksamkeitsfunktionen. *Zeitschrift für Neuropsychologie*, 5 (1), 15–28.
- Zimmermann, P. & Fimm, B. (1993). Testbatterie zur Aufmerksamkeitsprüfung (TAP). Freiburg: Psytest.
- Zimmermann, P. & Fimm, B. (2001). *Ergänzungsmanual zur Testbatterie zur Aufmerksamkeitsprüfung (TAP) Version 1.6 – Kinder- und Jugendlichennormen*. Herzogenrath: Psytest.
- Zimmermann, P. & Fimm, B. (2004). Die Testbatterie zur Aufmerksamkeitsprüfung (TAP). In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 177–202). Göttingen: Hogrefe.
- Zimmermann, P., North, P. & Fimm, B. (1994). Diagnosis of attentional deficits: theoretical considerations and presentations of a test battery. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 3–15). Tübingen: Gunter Narr.



3.20.TPK; Testreihe zur Prüfung der Konzentrationsfähigkeit

Kurth, E. & Büttner, G. (1999). Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK) (2., neu bearbeitete Auflage). Göttingen: Hogrefe.

Beschreibung

Die TPK ermöglicht bei Schülerinnen und Schülern der Klassen 2–6 eine differenzierte Diagnostik von Konzentrationsstörungen mit dem Ziel, sonderpädagogischen Förderbedarf zu ermitteln bzw. geeignete Trainingsmaßnahmen einzuleiten.

Die Testreihe zur Prüfung der Konzentrationsfähigkeit (Kurth & Büttner, 1999) ist die zweite überarbeitete Auflage (Neunormierung) des 1984 von Kurth veröffentlichten Verfahrens. Nach Brickenkamp und Karl (1986) ermöglicht die Konzentrationsfähigkeit dem Individuum, sich aufgabenrelevanten internen oder externen Reizen selektiv und ununterbrochen zuzuwenden und diese schnell und korrekt zu analysieren. Insofern stellt die Konzentrationsfähigkeit eine allgemeine Voraussetzung u.a. für schulische Leistungen dar (Kurth & Büttner, 1999, S. 7). So kommt Linderkamp (1996) zu dem Ergebnis, dass 80–90 % aller Kinder mit Aufmerksamkeitsstörungen Lernschwierigkeiten aufweisen. In einer Studie zeigte Orthmann (1993) anhand von drei hinsichtlich ihrer Leistungsstärke unterschiedlichen Untergruppen von Kindern der 5. Klasse (leistungsstarke, leistungsschwache und lernbehinderte Kinder), dass die beste Differenzierung über die Konzentrationswerte des TPK gelang (Kurth & Büttner, 1999, S. 10).

In Anlehnung an Neumann (1992, 1996) gehen die Autorinnen/Autoren davon aus, dass „den Sachverhalten, die deskriptiv unter dem Begriff der Aufmerksamkeit zusammengefasst werden, unterschiedliche Mechanismen zugrunde liegen, die von der Orientierung auf neue Anforderungen bis zu automatisierten Handlungen reichen“ (Kurth & Büttner, 1999, S. 7). Sie schlussfolgern, dass es nicht möglich ist, mit diagnostischen Verfahren unabhängig von der Aufgabenanforderung Aussagen über „die Aufmerksamkeit“ oder „die Konzentration“ zu machen und dass der Handlung eine zentrale Rolle zukommt. Als Konzentrationsstörung wird von Kurth und Büttner (1999, S. 8) in Anlehnung an DSM-IV und ICD 10 eine lang andauernde und chronische Herabsetzung von Aufmerksamkeit und Konzentration verstanden und von vorübergehenden Beeinträchtigungen der Konzentrationsfähigkeit durch fehlende Motivation, Ermüdung, Krankheit etc. abgegrenzt.

Die TPK besteht aus einer dreiteiligen Aufgabenserie:

1. Abschreibtext: 1–2 von Bildern umrahmte Geschichten sind in 10 Minuten abzuschreiben. Alle 2 Minuten erfolgt das Kommando: „Wort zu Ende schreiben, eine Zeile freilassen, weiterschreiben“. Kinder, die vor Ablauf der vorgesehenen Zeit fertig sind, bearbeiten

anschließend eine weitere Geschichte.

2. Tiergeschichte: Dieser Untertest dient als rezeptive Aufmerksamkeitsprobe: Es wird eine Geschichte vorgelesen. Die darin vorkommenden Tierarten sind anschließend aufzuschreiben.
3. Rechentest: Der Rechentest besteht aus Additions- und Subtraktionsaufgaben. Innerhalb von 10 Minuten sollen so viele Aufgaben wie möglich bearbeitet werden. Jeweils nach 2 Minuten wird das Kind angewiesen, einen Strich zu machen, um später die Anzahl der Rechenaufgaben pro 2-Minuten-Einheit zu errechnen.

Die Auswertung erfolgt nach folgenden Variablen:

- TSL: Silbenzahl in 10 Minuten (Abschreibetext)
- TF %: Prozentualer Fehleranteil (Abschreibetext)
- TS %: Stabilitätswerte (Abschreibetext)
- Ti: Anzahl der reproduzierten Tiere (Tiergeschichte)
- RL: Anzahl in 10 Minuten (Rechenaufgaben)
- RF %: Prozentualer Fehleranteil (Rechenaufgaben)
- RS %: Stabilitätswerte (Rechenaufgaben)

Demnach handelt es sich bei den Variablen TSL, Ti und RL um Mengenvariablen (Quantität), bei TF % und RF % um Genauigkeitswerte (Qualität) und bei den Variablen TS % und RS % handelt es sich um Werte, die den Konzentrationsverlauf abbilden.

Beim Abschreibetext wird die Zahl der jeweils in 2 Minuten abgeschriebenen Silben, die Gesamtzahl nach 10 Minuten und die Fehlerzahl berechnet. Entsprechend wird beim Rechentest vorgegangen. Bei der Tiergeschichte wird die Anzahl der genannten Tierarten summiert (sofern nicht mehr als drei falsche Tiere aufgeschrieben wurden und die Aufgabenbewertung beendet wird). Die Gesamtleistungen der drei Untertests (TSL, Ti und RL) werden im Auswertungsbogen notiert. Zur Berechnung der Genauigkeitswerte (TF % und RF %) werden auf der Rückseite des Protokollbogens Angaben gemacht: Die mit 100 multiplizierte Fehlerzahl ist durch die Gesamtzahl der Silben/Rechenaufgaben zu dividieren. Zur Berechnung der Stabilitätswerte (TS % und RS %) wird die Differenz zwischen den in 2 Minuten maximal geschriebenen Silben/gelösten Rechenaufgaben und den minimal erreichten Werten gebildet, mit 100 multipliziert und durch die Gesamtleistung dividiert. Die berechneten Werte werden anschließend auf der Vorderseite des Auswertungsbogens eingetragen. Die Rohwerte werden anhand der Normentabellen in C-Werte umgerechnet und können in das Profil der Konzentrationsleistungen eingetragen werden. Aus der Summe der C-Werte von TSL, TF %, Ti, RL und RF % kann anhand einer Umrechnungstabelle der Gesamtkonzentrationswert KW abgelesen werden. Dies ist allerdings nur dann sinnvoll, wenn es sich um ein ausgeglichenes Leistungsprofil handelt. Die Stabilitätswerte wurden nicht in den Gesamtwert einbezogen, da sie die niedrigsten Validitätskoeffizienten aufwiesen (Kurth & Büttner, 1999, S. 15). Sie werden neben dem KW als zusätzliche Information angegeben.



Altersgruppe: 2.–6. Schulstufe

Hinweise zur Durchführung

Das Verfahren sollte als **Gruppentest** durchgeführt werden, um die normale Klassensituation nachzubilden. Die TPK kann aber **auch als Einzeltest** vorgegeben werden.

Der gesamte Test erfordert etwa 40 Minuten. Die Instruktion ist im Handbuch wortwörtlich vorgegeben.

Die Testung kann durch pädagogisch geschultes Personal erfolgen. Der Test sollte in einer ruhigen Atmosphäre durchgeführt werden. Wegen der Nutzung schulischer Grundfertigkeiten ist der Test nicht bei Kindern mit Lese-Rechtschreib-Schwäche oder Rechenschwäche anwendbar. Ebenso können erhebliche feinmotorische Störungen die Tempowerte negativ beeinflussen.

Hinweise zu Auswertung und Interpretation

Normen (C-Werte von 1–10) werden von einer $n = 1.774$ umfassenden Stichprobe nach Schulklassen getrennt angegeben. Die Erhebung fand 1996/1997 statt. Auf eine geschlechtsgetrennte Normierung wurde verzichtet, weil sich nur hinsichtlich einer Variablen signifikant bessere Leistungen der Mädchen ergaben. In der Stichprobe sind alte und neue Bundesländer der Bundesrepublik Deutschland vertreten (Bayern, Schleswig-Holstein, Mecklenburg-Vorpommern und Sachsen) und aus diesen jeweils ganze Klassen nach den Aspekten Stadt ($n = 903$) und Land ($n = 871$) ausgesucht worden. In diesem Zusammenhang wurde ein leichtes Übergewicht der Mädchen ($n = 934$) gegenüber den Jungen ($n = 840$) akzeptiert. Die KW-Mittelwerte nach Länderstichproben ergaben ein leichtes Süd-Nord-Gefälle (um 3 KW-Punkte) und einen leichten Vorteil zugunsten der Landkreisstichproben. Zusätzlich werden im Manual die KW-Mittelwerte von früher gezogenen Stichproben mitgeteilt.

Im Manual liegen eine detaillierte Anleitung zur Auswertung sowie Hinweise zur Fehlerbewertung beim Abschreibtext und Angaben zur Berechnung der Testvariablen auf dem Auswertungsbogen vor. Für den Rechentest werden außerdem Schablonen zur schnellen Identifizierung richtiger Antworten geliefert. Die Rohwerte können mit Hilfe der Normen interpretiert und die entsprechenden C-Werte in ein Testprofil eingetragen werden.

Die Auswertung erfordert ca. 10 Minuten.

Testgüte

Objektivität

Aufgrund der wörtlichen Instruktion im Manual und der eindeutigen Auswertungsrichtlinien können Durchführung und Auswertung als objektiv gelten. Ebenso ist Interpretationsobjektivität gegeben (Normen).

Reliabilität

Die Retestreliabilität nach 2 Monaten bei Schülerinnen/Schülern der 3. und 5. Schulstufe ($n = 50$, Teilstichprobe von 1979) ergab für den Gesamt-Konzentrations-Wert einen Koeffizienten von $r = .88$. Dieses Ergebnis konnte 13 Jahre später an einer Stichprobe von $n = 231$ Schülerinnen/Schülern der 2.–6. Schulstufe aus Würzburg bestätigt werden. Für die einzelnen Testvariablen ergaben sich Koeffizienten zwischen $r = .55$ und $r = .95$, für den Gesamttest von $r = .80$. Es zeigte sich, dass die Mengenleistungswerte zuverlässiger sind als die Fehlerwerte und die den Verlauf widerspiegelnden Schwankungswerte.

Die Halbierungsreliabilitäten für den Abschreibtext und den Rechentest wurden für eine Teilstichprobe von $n = 830$ Kinder der Klassen 2–6 berechnet. Es wurden jeweils die Mengenleistungen in der 2. und 3. mit denen in der 4. und 5. Teilzeit korreliert. Für den Abschreibtext ergab sich (nach Spearman-Brown) ein Koeffizient von $r = .96$, für den Rechentest von $r = .94$.

Validität

Zur Überprüfung der kriterienbezogenen Validität wurden die TPK-Ergebnisse mit verschiedenen anderen Verfahren korreliert. Die TPK-Ergebnisse (Abschreibtext und Rechentest) von 68 Kindern mit Konzentrationsstörungen aus der 2. und 3. Klasse (durchschnittlicher IQ = 106) wurden mit den Ergebnissen aus dem Pauli-Test und dem Test d2 korreliert. Die höchste Korrelation von $r = .63$ ergab sich zwischen dem TPK-Rechentest und dem Pauli-Rechentest.

Außerdem wurden die TPK-Werte mit dem Lehrer/innen-Urteil korreliert (Kurth & Büttner, 1999, S. 18 f.). Die Lehrer/innen wurden gebeten, ihre Schüler/innen hinsichtlich deren Konzentrationsfähigkeit in eine Rangreihe zu bringen. In die Auswertung gingen die 47 Klassen ein, bei denen dies von der Lehrerin/vom Lehrer verwirklicht wurde. Für verschiedene Testvariablen (Gesamtsummen der drei Untertests und Genauigkeitswerte) ergaben sich Korrelationen zwischen $r = .36$ und $r = .40$, für den Gesamt-Konzentrations-Wert KW ergab sich eine Korrelation von $r = .59$. Das Ergebnis wird von Kurth und Büttner (1999, S. 19) als Bestätigung des Testkonzeptes gewertet, welches verschiedene Testanforderungen beinhaltet.

Der Annahme, dass zumindest eine gewisse Korrelation zwischen TPK-Ergebnissen und Intelligenz zu erwarten sei, wurde durch verschiedene Untersuchungen nachgegangen. Zu keinem der 1980 in Rostock erhobenen Intelligenzmaße (CPM, $n = 19$; HAWIK, $n = 31$; CFT 1, $n = 39$) ergaben sich jedoch Korrelationen in nennenswerter Höhe ($r = .17$ bis $r = .30$), die zudem statistisch nicht signifikant waren. Entsprechend erbrachte eine Untersuchung von Büttner

(1994) an 94 Haupt- und Grundschülerinnen und -schülern und 113 Schülerinnen und Schülern der Schulen für Erziehungshilfe zwischen der TPK-Variablen Ti (Gesamtmenge reproduzierter Tiere) und dem Zahlennachsprechen vorwärts des HAWIK nur sehr niedrige Korrelationen von $r = .22$ bzw. $r = .05$. Demnach handelt es sich bei dieser TPK-Aufgabe nicht um eine reine Gedächtnisanforderung.

Um sicherzugehen, dass die TPK unabhängig von der schulischen Lernmotivation ist, wurde der KW-Wert der TPK mit der Skala „Schulunlust“ des Angstfragebogens für Schüler/innen (AFS; Wieczerkowski, Nickel, Janowski, Fittkau & Rauer, 1975) korreliert (Kurth & Büttner, 1999, S. 19 f.; Büttner & Kurth, 1996). Es ergaben sich an zwei 1996 in Würzburg und Rostock untersuchten Stichproben von $n = 329$ bzw. $n = 393$ Schüler/innen geringe, wenn auch statistisch signifikante Korrelationen von $r = -.13$ und $r = -.19$. Somit ist der Einfluss von Schulunlust auf die TPK-Ergebnisse als recht gering anzusehen.

Ein Extremgruppenvergleich zeigte, dass sich verhaltensauffällige Kinder (heute Schule für Erziehungshilfe), Kinder einer Kliniksonderschule, die aufgrund von Konzentrationsschwächen ein Training durchlaufen sollten, und Schüler/innen der Schule für Lernbehinderte hinsichtlich der KW-Werte signifikant von den Normwerten der Vergleichsstichproben (Normierungsstichprobe von 1980) unterschieden (Kurth & Büttner, 1999, S. 20). Zusätzlich wurden die TPK-Werte von 40 Schüler/innen der 2.–4. Klasse der Schule für Erziehungshilfe mit denjenigen aus der Normierungsstichprobe von 1996 verglichen. Für die meisten Testvariablen ergaben sich signifikante Unterschiede. Die Testergebnisse zum Extremgruppenvergleich deuten darauf hin, dass die TPK eine Differenzierung zwischen verschiedenen schulischen Anforderungen erlaubt.

In einer Studie von Barchmann, Ettrich und Kinze (1987) konnte gezeigt werden, dass das Konzept der TPK (schulrelevante Aufgaben, Gruppensituation) Vorteile gegenüber anderen Konzentrationstests aufweist: Untersucht wurden 250 Normalschulkinder der 1.–4. Klasse, die kinderneuropsychiatrisch stationär behandelt wurden. Ihnen wurden vier verschiedene Konzentrationstests dreifach vorgegeben (zusätzlich zum ersten Mal nach vier Wochen und 3 Monate nach stationärer Entlassung). Der Vergleich der Leistungsveränderungen über die verschiedenen Messzeitpunkte bei KHV und TPK ergab, dass sich die TPK-Ergebnisse weniger verändert hatten als diejenigen des KHV (KHV; Koch & Pleissner, 1985). Die Autorinnen/Autoren schlussfolgern, dass die Prüfung schulrelevanter Aufgaben in der Gruppensituation geringere Veränderungen ergibt als spielerisch orientierte Anforderungen, die in der Einzelsituation zu bearbeiten sind (Barchmann et al., 1987, S. 158).

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Die Testreihe zur Prüfung der Konzentrationsfähigkeit unterscheidet sich von zahlreichen anderen Konzentrationstests durch die Vielgestaltigkeit der Testanforderungen, die die im Schulalltag üblichen Anforderungen möglichst genau abprüfen sollen. Auch die Überlegung, ein mög-

lichst als Gruppentest anzuwendendes Verfahren zu konzipieren, um eine in einer Schulstunde übliche Situation abzubilden, erscheint logisch. Das Manual ist weitestgehend übersichtlich aufgebaut. Zu kritisieren sind jedoch z.T. fehlende Literaturangaben, unzureichende Legenden bei Tabellen sowie die recht undurchschaubare Aufführung von Stichproben.

Die Untersuchungen zur Gültigkeit der TPK weisen darauf hin, dass es sich um ein valides Verfahren handelt, das gegenüber anderen Verfahren Vorteile aufweist, wie z.B. eine geringere spontane Schwankungsbreite der Parameter Qualität und Quantität. Die Reliabilität des Gesamtwerts KW ist zufriedenstellend. Die Normierung fand an einer ausreichend großen Stichprobe statt, deren Repräsentativität unterstellt werden kann. Allerdings kritisieren Barchmann et al. (1987, S. 159) zu Recht, dass Klassennormen eine zu geringe Gliederung der Leistungen ermöglichen. Der Anstieg der Norm von einer Klasse zur anderen ist so erheblich, dass es einen eklatanten Unterschied im C-Wert ergibt, je nachdem ob ein Kind am Ende einer Klasse oder am Anfang der nächsthöheren Klasse getestet wurde. Es wird deshalb empfohlen, mittels einer größeren Eichstichprobe z.B. Halbjahresnormen zu errechnen. Als ökonomisches und nützliches Instrument kann die Anwendung der TPK für den Schulbereich empfohlen werden.

Autorinnenkommentar

Dieser Test eignet sich hervorragend zur Überprüfung von schulnahen Aspekten der Konzentration. Will man ein Aufmerksamkeitsdefizit feststellen, eignet er sich nur bedingt, da Teil 1, wie auch der Revisions-Test, insbesondere bei Kindern im Alter zwischen 9 und 12 Jahren oft zu einer Verzerrung führt, da Probandinnen/Probanden mit guten Kopfrechenfähigkeiten hier signifikant besser abschneiden als Kinder/Jugendliche mit Problemen beim Kopfrechnen oder mit Dyskalkulie. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit liegen oder schlichtweg an der mangelnden Fähigkeit, im Kopf schnell zu rechnen. Beim zweiten Teil spielt die Rechtschreibfähigkeit zusätzlich zur Konzentrationsfähigkeit eine nicht unbedeutende Rolle, wenn es um die Fehleranzahl und Schnelligkeit beim Text-Abschreiben geht. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit liegen oder an mangelnden Rechtschreibkenntnissen bzw. Legasthenie. Der dritte Teil des TPK eignet sich hervorragend, um Konzentrationsfähigkeit (und Gedächtnis) zu überprüfen und macht den Kindern erfahrungsgemäß großen Spaß. Dieser kann z.B. in Kombination mit anderen Verfahren vorgegeben werden.

Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind bzw. die/der Jugendliche mit dem Kopfrechnen und/oder mit der Rechtschreibung Probleme hat bzw. an einer möglichen Rechenschwäche bzw. Legasthenie leidet. Unter Umständen kann man auch nur mit dem dritten Teil arbeiten, da dieser sehr aufschlussreiche Zusatzergebnisse im Rahmen einer Aufmerksamkeitsdiagnostik liefern kann.

Literatur

- Barchmann, H., Ettrich, K. U. & Kinze, W. (1987). *Methodische Aspekte zur Diagnostik der Konzentrationsfähigkeit im Kindesalter*. *Psychologie für die Praxis*, 2, 156-160.
- Brickenkamp, R. & Karl, G. A. (1986). *Geräte zur Messung von Aufmerksamkeit, Konzentration und Vigilanz*. In R. Brickenkamp (Hrsg.), *Handbuch apparativer Verfahren in der Psychologie* (S. 195–211). Göttingen: Hogrefe.
- Büttner, G. (1994). *Aspekte zur Validität eines schulnahen Konzentrationstests*. Poster auf dem 39. Kongress der Deutschen Gesellschaft für Psychologie. Hamburg: Universität.
- Büttner, G. & Kurth, E. (1996). *Konzentrationsleistungen in der schulnahen Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK) – eine regionale Vergleichsstudie*. *Zeitschrift für Pädagogische Psychologie*, 10 (3/4), 187–197.
- Dunkel, B. (1999). *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)*. In S. Grubitzsch, *Testtheorie – Testpraxis. Psychologische Tests und Prüfverfahren im kritischen Überblick* (2. unveränderte Auflage der vollständig überarbeiteten und erweiterten Neuauflage 1991. Kurzanalyse S. 435–437). Eschborn: Klotz.
- Koch, J. & Pleissner, S. (1985). *Konzentrations-Handlungs-Verfahren (KHV)*. Berlin: Psychodiagnostisches Zentrum.
- Kurth, E. & Büttner, G. (1999). *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)* (2., neu bearbeitete Auflage). Göttingen: Hogrefe.
- Kurth, E. & Büttner, G. (2004). *Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK)*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 143-159). Göttingen: Hogrefe.
- Kurth, E. (1984). *Testreihe zur Prüfung der Konzentrationsfähigkeit TPK (Testmappe mit Manual, je einem Musterrechenblatt und Musterauswertungsbogen, zwei Abschreibtexten und zwei Schablonen)*. Berlin: Psychodiagnostisches Zentrum.
- Linderkamp, F. (1996). *Zur Homogenität des Störungsbildes und die Notwendigkeit zur Subgruppendiffektion*. *Kindheit und Entwicklung*, 5, 89–92.
- Neumann, O. (1992). *Theorien der Aufmerksamkeit: Von Metaphern zu Mechanismen*. *Psychologische Rundschau*, 43, 83–101.
- Neumann, O. (1996). *Theorien der Aufmerksamkeit*. In O. Neumann & A. F. Sanders (Hrsg.), *Aufmerksamkeit (Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung, Serie II, Kognition, Band 2, 559–643)*. Göttingen: Hogrefe.
- Neumann, O. (1996). *Komponenten der Aufmerksamkeit und ihre Störungen*. *Kindheit und Entwicklung*, 5 (2), 75–79.
- Orthmann, D. (1993). *Vergleichende Untersuchungen zu den biosozialen Entwicklungsbedingungen bei leistungsstarken, leistungsschwachen und lernbehinderten Schülern*. Frankfurt: Peter Lang.
- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1975). *Angstfragebogen für Schüler. AFS. Westermann Test* (2. veränderte Auflage). Göttingen: Hogrefe.

3.21.VIGIL; Vigilanz

Schuhfried, G. (2004). Vigilanz. VIGIL (Version 25.00) (Computerprogramm mit Manual). Mödling: Schuhfried.

Beschreibung

VIGIL ist ein Test zur Erfassung der Aufrechterhaltung der Aufmerksamkeit in einer monotonen Reizsituation. Die Erfassung der Vigilanz auf Basis von monotonen Beobachtungsaufgaben ist sehr realitätsnah, valide und hoch reliabel.¹⁴

Unter Vigilanz wird in der Psychologie die Aufrechterhaltung der Aufmerksamkeit in monotonen Reizsituationen verstanden. Sie kennzeichnet einen Zustand der Funktionsbereitschaft des Organismus bei zufälligen, schwelennahen und selten auftretenden Ereignissen (Roth, 1989). Diese Fähigkeit ist beispielsweise bei Überwachungstätigkeiten gefordert. Der Ursprung der Vigilanzforschung liegt in der Radarschirmüberwachung (Buckner & McGrath, 1963; Mackworth, 1950), aus deren Kontext sich die Vigilanzmessung mit apparativen Verfahren entwickelte. Studien zur Vigilanz stellten in den darauf folgenden Jahren fest, dass die Leistungseffizienz der Probandinnen/Probanden über längere Zeit umso besser war, je stärker sich der kritische, zu beachtende Reiz in Intensität und Größe vom Hintergrund abhob und je größer die Erwartung war, die für das Auftreten des kritischen Reizes erzeugt wurde (Fraser, 1957).

Umgekehrt erwies sich die Anzahl der falschen Reaktionen als umso größer, je höher die Gesamt-Signalfrequenz und je geringer die Wahrscheinlichkeit des Auftretens eines kritischen Reizes ist. Als Indikatoren der Vigilanz gelten die Genauigkeit der Bearbeitung der Aufgabe sowie die Geschwindigkeit, mit der über längere Zeit hinweg die Aufgabe beantwortet wird. Der über die Zeit eintretende Leistungsabfall bei Vigilanzexperimenten wird mit der Abnahme des Aktivierungsniveaus und der entsprechenden Zunahme des Latenzniveaus der Probandinnen/Probanden erklärt; von Graf (zit. in Arnold, 1980)

¹⁴ vgl. www.schuhfried.at/wiener-testsystem-wts/alle-tests-von-a-bis-z/test/vigil-vigilanz/ (Abfragedatum: 30.12.2012).



bezeichnet dies als „Überforderung durch Unterforderung“. Das Programm „Vigilanz“ im Wiener Testsystem (Schuhfried, 1995) stellt die Möglichkeit bereit, mittels dreier unterschiedlicher Standardparameterblöcke computergestützt die Vigilanz der Probandinnen/Probanden im Zeitverlauf zu erfassen.

Das Testprogramm „Vigilanz“ beinhaltet drei Standardparameterblöcke (S1, S2 und S3) zur Erfassung der Vigilanz, die zwei voneinander verschiedenen Testvorgabeverionen zugeordnet sind (Schuhfried, 1995, S. 6):

1. *Version Quatember-Maly* (Standardparameterblock S1):

Bei dieser Testversion zeigt der Bildschirm eine große Kreisbahn, die aus vielen kleinen Kreisen zusammengesetzt ist. Sie werden durch eine helle Umrahmung dargestellt und sind – wie auch der Bildschirmhintergrund – in der Mitte dunkel. Die kleinen Kreise leuchten nacheinander im Uhrzeigersinn hell auf, so dass ein Bewegungseindruck entsteht. Der aufleuchtende Punkt springt um einen Schritt zum nächsten Kreis weiter, wobei jedoch manchmal ein Kreis übersprungen wird (kritisches Ereignis). Aufgabe der Probandinnen/Probanden ist es, auf diese Sprünge zu reagieren, indem sie eine Taste am Gerät drücken. Jeder Tastendruck wird mit einem leisen Piepton bestätigt. Damit Veränderungen der Leistung während des Testverlaufs innerhalb festgelegter Abschnitte miteinander verglichen werden können, ist der Test in Teilzeiten eingeteilt.

Der Standardparameterblock S1 beinhaltet folgende Parameter:

- Schrittdauer: 1,5 Sekunden (der Punkt bewegt sich in Schritten von 1,5 Sekunden im Uhrzeigersinn weiter)
- Anzahl der Teilzeiten im Gesamtttest: 20
- Anzahl der Sprünge: fünf kritische Reize (Sprünge) pro Teilzeit, d.h. im Gesamtttest 100 Sprünge
- Anzahl der Schritte: 50 pro Teilzeit, d.h. der Gesamtttest beinhaltet 1000 Schritte
- Testdauer: Eine Teilzeit dauert 75 Sekunden, der Gesamtttest 25 Minuten

2. *Version Muggenburg* (Standardparameterblöcke S2 und S3):

Bei dieser Version springt ebenfalls ein leuchtender Punkt im Uhrzeigersinn entlang einer Kreisbahn, allerdings ist diese Bahn im Gegensatz zur Version „Quatember-Maly“ nicht durch Kreise dargestellt, so dass die Probandinnen/Probanden die Sprünge schätzen müssen. Die Bearbeitung dieser Form erfordert nach Schuhfried (1995, S. 6) einen höheren Konzentrationsaufwand und beinhaltet einen höheren Schwierigkeitsgrad. Damit Veränderungen der Leistung während des Testverlaufs innerhalb festgelegter Abschnitte miteinander verglichen werden können, ist der Test ebenfalls in Teilzeiten eingeteilt. S3 entspricht der Form S2, allerdings mit einer Testdauer von 42 statt 33 Minuten.

Die Standardparameterblöcke beinhalten folgende Parameter:

Standardparameterblock S2 – Müggenburg 33 („33“ für 33 Minuten Testdauer):

- Schrittdauer: 2 Sekunden
 - Anzahl der Teilzeiten: 8
 - Anzahl der Sprünge: vier kritische Reize (Sprünge) pro Teilzeit (d.h. im Gesamttest 32 Sprünge)
 - Anzahl der Schritte: 125 pro Teilzeit (d.h. der Gesamttest beinhaltet 1000 Schritte)
- Testdauer: Eine Teilzeit dauert 4 Minuten 10 Sekunden, der Gesamttest ca. 33 Minuten.

Standardparameterblock S3 – Müggenburg 42 („42“ für 42 Minuten Testdauer):

- Schrittdauer: 2 Sekunden
 - Anzahl der Teilzeiten: 10
 - Anzahl der Sprünge: vier kritische Reize (Sprünge) pro Teilzeit (d.h. im Gesamttest 40 Sprünge)
 - Anzahl der Schritte: 125 pro Teilzeit (d.h. der Gesamttest beinhaltet 1000 Schritte)
- Testdauer: Eine Teilzeit dauert 4 Minuten 10 Sekunden, der Gesamttest ca. 42 Minuten.

Der Testablauf gliedert sich in eine Übungsphase und in die eigentliche Testphase. Während der Übungsphase wird das Programm automatisch nach dem vierten Fehler abgebrochen. Während der Testphase läuft das gesamte Programm ungeachtet der Anzahl der Fehler ab.

Die Auswertung erfolgt computergestützt. Es werden folgende Variablen ausgewertet:

- Anstieg der Richtigen b.R. (in % pro Stunde): Auf Itemniveau wird die Regression der richtigen Antworten auf kritische Reize berechnet und der Anstieg der Regressionsgeraden ausgegeben.
- Anstieg der Reaktionszeiten b.RT (in Sekunden pro Stunde): Es wird die Regression der Reaktionszeiten von der kritischen Reizvorgabe bis zum Tastendruck berechnet.
- Bestimmtheitsmaß der Regression der Richtigen B.R: Gibt die Güte der Anpassung der Regressionsgeraden an die Kurve des tatsächlichen Leistungsverlaufs wieder. Die Zahl liegt zwischen 0 und 1. Je näher der Wert bei 1 liegt, umso kleiner sind die Abweichungen der beobachteten Werte von der Regressionsgeraden.
- Bestimmtheitsmaß der Regression der Reaktionszeiten B.RT
- Anzahl der Richtigen *R*
- Anzahl der Falschen *F*
- Anzahl der Ausgelassenen *Ausg*
- Mittelwert der Reaktionszeiten *Mw.RT*
- Mittelwert der Richtigen *Mw.R* (über die Anzahl der Teilzeiten)
- Mittelwert der Falschen *Mw.F* (über die Anzahl der Teilzeiten)
- Mittelwert der Ausgelassenen *Mw.Ausg.* (über die Anzahl der Teilzeiten)
- Standardabweichung der Reaktionszeiten *St.Rt*
- Standardabweichung der Richtigen über die Teilzeiten *St.R.*

Für jede Teilzeit werden weiter folgende Variablen ausgewertet:

- Anzahl der Richtigen
- Anzahl der Falschen
- Anzahl der Ausgelassenen
- Mittelwert der Reaktionszeiten

Altersgruppe: ab 6;0 Jahren

Hinweise zur Durchführung

An jedem Gerät kann jeweils **nur eine Person** getestet werden. Es liegen drei Standardparameterblöcke vor, auf die bereits Bezug genommen wurde: S1 Quatember-Maly, S2 Müggenburg 33 und S3 Müggenburg 42.

Zur Untersuchung spezifischer Fragestellungen kann die Einstellung der Parameter der Testvorgabe über das Programm „Parametererstellung“ des zentralen Testleitprogramms verändert werden (Schuhfried, 1995, S. 23 f.).

Die Durchführungszeit richtet sich nach dem gewählten Standardparameterblock:

- S1: Eine Teilzeit dauert 75 Sekunden, der Gesamttest 25 Minuten.
- S2: Eine Teilzeit dauert 4 Minuten 10 Sekunden, der Gesamttest ca. 33 Minuten.
- S3: Eine Teilzeit dauert 4 Minuten 10 Sekunden, der Gesamttest ca. 42 Minuten.

Vor der eigentlichen Testphase läuft computergesteuert eine ausführliche Instruktions- und Übungsphase mit Feedback für die Testperson ab. Die Übungsphase wird automatisch nach dem vierten Fehler abgebrochen (Schuhfried, 1995, S. 16).

Die Testleiterin/der Testleiter sollte mit den Grundsätzen der Anwendung psychologischer Testverfahren vertraut sein. Sinnvollerweise sollte sie/er sich darüber hinaus mit dem computergesteuerten Testablauf vertraut gemacht haben. Vor Beginn der Testphase sollte sich die Testleiterin/der Testleiter vergewissern, ob die Instruktion von der Probandin/dem Probanden verstanden wurde. Zur Bedienung des Testvorgabeprogramms stehen Standardfunktionen zur Verfügung. Durch Bildschirminformationen wird die Testleiterin/der Testleiter über jeden Schritt des Testverfahrens auf dem Laufenden gehalten (Schuhfried, 1995, S. 16 f.).

Hinweise zu Auswertung und Interpretation

Mit dem Standardparameterblock S1 wurde von Teichmann, Ricken, Bader und Hofmann (nach Schuhfried, 1995) eine Stichprobe von $n = 120$ Studierenden untersucht sowie von Rossmann zwi-

schen 1990 und 1994 eine Stichprobe von $n = 141$ Kindern im Alter von 7–12 Jahren. Mittelwerte und Standardabweichungen für das Personenmerkmal Alter sowie für die Leistungsmerkmale R, F und Mw.ST sind angeführt bei Schuhfried (1995, S. 14). Als Normwerte wurden für die Normierungsstichproben, für die Gesamtstichproben sowie für die zwei Altersgruppen der Kinderstichprobe (7–9, 10–12 Jahre), Prozentränge und T-Werte berechnet (Schuhfried, 2004, Anhang).

Zur Normierung des Standardparameterblocks S2 wurde von Sturm und Büssing (1990) eine geschichtete Stichprobe von $n = 200$ Personen aus der Normalbevölkerung untersucht. Die Schichtung erfolgte nach Alter und Bildungsgrad. Mittelwerte und Standardabweichungen des Leistungsverlaufs (R und F) für die Gesamtgruppe und für vier Altersgruppen sind angeführt bei Schuhfried (2004, S. 15). Normen wurden für die Gesamtgruppe sowie für zwei Altersgruppen ermittelt. Die Vergleichswerte beziehen sich auf die Anzahl richtiger Reaktionen. Die Normen (Prozentränge und T-Werte) sind angeführt bei Schuhfried (2004, Anhang).

Des Weiteren wurden für den Standardparameterblock S2 von Pfister-Oswald von einer Vergleichsstichprobe von Patientinnen/Patienten eines psychiatrischen Krankenhauses ($n = 111$) Normen für die Gesamtstichprobe sowie für vier Altersgruppen zur Verfügung gestellt (Schuhfried, 2004, Anhang).

Die Auswertung wird im Zusammenhang mit der Verfahrensdurchführung beschrieben. Für die Interpretation sind die Rohwerte anhand von Normtabellen in T-Werte und Prozentrangwerte umzuwandeln. Die Normtabellen sind angeführt bei Schuhfried (2004, Anhang). Hinweise zur Interpretation der Testergebnisse finden sich ebenfalls bei Schuhfried (2004, S. 19–22).

Die Auswertung erfolgt computergestützt. Der Ergebnisausdruck beinhaltet die Testergebnisse, eine graphische Darstellung der Regressionsgeraden und ein Testprotokoll (Schuhfried, 2004, S. 18). Daneben kann auch eine Kurzauswertung mit Kopf und Testergebnissen erstellt werden. Beispiele für die Bildschirmanzeige der Auswertung, eines Auswertungsausdrucks und eines Kurzauswertungsausdrucks sind angeführt bei Schuhfried (2004, Anhang). Druckoptionen zur Veränderung des Umfangs des Ausdrucks können über das zentrale Testleitprogramm eingestellt werden (Schuhfried, 2004, S. 25). Die Probandinnen-/Probandendaten können für die Weiterverar-



beitung, z.B. mit Statistikprogrammen, aufbereitet werden (Konvertierung der Testergebnisse ins ASCII-Format).

Die computerisierte Auswertung erfordert einen minimalen Zeitaufwand.

Testgüte

Objektivität

Da die Testvorgabe, die Registrierung und die Auswertung der Resultate standardisiert und computergesteuert erfolgen, ist von Objektivität in Durchführung und Auswertung auszugehen.

Reliabilität

Zur Bestimmung der Reliabilität des Verfahrens wurden an verschiedenen Stichproben und Standardparameterblöcken hinsichtlich der Variablen R, F und Mw.RT Cronbachs α und die Split-half-Reliabilität berechnet, wobei die Split-half-Reliabilität nach der Odd-even-Methode berechnet wurde, d.h. eine Testhälfte bestand aus den Items mit ungeraden Zahlen, die andere aus den Items mit geraden Ordnungszahlen (siehe Tab. 7).

Reliabilitäten für die Variablen R, F, und Mw.RT (Schuhfried, 1995, S. 10)

Tab. 7: Reliabilitäten für die Variablen R, F, und Mw.RT nach Schuhfried (1995, S. 10)

| Stichprobe | | | Reliabilität | | |
|------------|----------------|-----------------|--------------|-----|-------|
| Testform | Personengruppe | Alter in Jahren | R | F | Mw.RT |
| S1 | Kinder | 7–9 | .95 | .99 | .94 |
| S1 | Kinder | 10–12 | .89 | .96 | .91 |
| S1 | Studierende | | .77 | .39 | .97 |
| S2 | Erwachsene | 20–49 | .81 | .85 | |
| S2 | Erwachsene | 50–75 | .84 | .98 | |

Validität

Das Verfahren verfügt über Augenscheinvalidität. Die inhaltliche Validität ist nach Schuhfried (2004, S. 11) gegeben, da der Test alle Kriterien zur Operationalisierung der Vigilanz erfüllt.

Von Sturm und Büssing (1990) wurde ein Vergleich von links- bzw. rechtshemisphärisch geschädigten Patientinnen und Patienten vorgenommen. Es ergab sich ein deutlicher Zusammenhang zwischen der mit dem Parameterblock S2 ermittelten Aufmerksamkeitsleistung und anderen Aufmerksamkeitsfunktionen bei Patientinnen und Patienten mit Läsionen der rechten Hirnhemisphäre. Dagegen wurden für rechtshemisphärisch geschädigte Patientinnen und Patienten nur in sehr geringem Maß Korrelationen zu Testergebnissen gefunden, die überwiegend sprach-

liche Funktionen erfassen. Nach Sturm und Büssing (1990) ist daraus zu folgern, dass der Vigilanztest typische Aufmerksamkeitsleistungen erfasst, die besonders von der rechten Hirnhemisphäre kontrolliert werden.

Als weiteren Beleg für die Validität des Verfahrens führt Schuhfried (2004, S. 12) die Interkorrelationen der Testvariablen R, F, Mw.RT und St.RT, ermittelt anhand des Parameterblocks S1, für die Stichproben Kinder, Studierende und Sportler/innen an (siehe Tab. 8).

Tab. 8: Interkorrelationen der Variablen R, F, Mw.RT und St.RT für den Standardparameterblock S1 und die Stichproben Kinder, Studierende, Sportler/innen (Schuhfried, 2004, S. 12)

| | F | Mw.RT | St.RT |
|-----------------------|----------|--------------|--------------|
| Kinder | | | |
| R | -.18* | -.50** | -.43** |
| F | | .14 | .43** |
| Mw.RT | | | .58** |
| Studierende | | | |
| R | -.61** | -.44** | -.70** |
| F | | .10 | .57** |
| Mw.RT | | | .54** |
| Sportler/innen | | | |
| R | -.26 | -.24 | -.36** |
| F | | .12 | .42** |
| Mw.RT | | | .59** |

Anmerkungen: * $p < .01$, ** $p < .001$.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Der vorliegende computergesteuerte Vigilanztest dient der Erfassung von Vigilanz im Sinne von Daueraufmerksamkeit in reizarmen Beobachtungssituationen. Diese Fähigkeit ist in vielen Bereichen des täglichen und beruflichen Alltags (z.B. bei Überwachungsaufgaben) von großer Bedeutung. Aufgrund der automatischen Testdurchführung, Überwachung und sofortigen computergesteuerten Auswertung ist die Ökonomie des Verfahrens – als wichtige Voraussetzung einer praktikablen Anwendung in der diagnostischen Praxis – gewährleistet. Die zu den Gütekriterien mitgeteilten Resultate weisen das Verfahren insgesamt als brauchbar für seinen angestrebten Anwendungsbereich aus.



An der Darstellung des Verfahrens ist allerdings die unvollständige Auflistung der Literaturhinweise zu bemängeln – ein Punkt, der bei einer Überarbeitung des Manuals unbedingt behoben werden sollte.

Autorinnenkommentar

Das VIGIL-Verfahren ist bei Kindern und Jugendlichen aufgrund der langen Aufmerksamkeits-spanne bei verhältnismäßig wenigen Ereignissen eher unbeliebt. Es ermöglicht allerdings eine gute Abschätzung und Einteilung der Probandin/des Probanden in die 4-Felder-Tafel „schnell/langsam

x genau/ungenau“ und die genaue Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“, welche mit den herkömmlichen Papier-Bleistift-Tests nicht valide erfasst werden kann.

Literatur

- Arnold, W. (1980). *Lexikon der Psychologie*. Freiburg: Herder.
- Buckner, D. N. & McGrath, J. J. (1963). *Vigilance: A symposium*. New York.
- Fraser, D. C. (1957). *A study of vigilance and fatigue*. (Doctoral dissertation, University of Edinburgh).
- Mackworth, N. H. (1950). *Researches on the measurement of human performance*. Medical Research Council, Special Report, No. 268.
- Müggenburg, U. (1981). *Zur Beurteilung verschiedener Testverfahren und Einsatzmöglichkeiten des Vigilanzgerätes nach Quatember und Maly*. Unveröffentlichte Dissertation, Technische Hochschule Aachen, Medizinische Fakultät.
- Neumann-Zilke, L. (2009). *Vigilanz (VIGIL). Untertest aus dem Wiener Testsystem (WTS)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 99–109). Göttingen: Hogrefe.
- Roth, E. (1989). *Enzyklopädie der Psychologie* (S. 324 ff.; Kap. Verarbeitungskapazität, Kap. Flaschenhalstheorie, Kap. Selektive Aufmerksamkeit, Kap. Informationsverarbeitungskapazität). Göttingen: Hogrefe.
- Schuhfried, G. (1995). *Vigilanz (Version, 5.00)*. Mödling: Schuhfried.
- Schuhfried, G. (2004). *Vigilanz (Version 25.00)*. Mödling: Schuhfried.
- Sturm, W. & Büssing, A. (1990). *Normierungs- und Reliabilitätsuntersuchungen zum Vigilanzgerät nach Quatember und Maly*. *Diagnostica*, 36 (1), 50–59.

3.22. WAF; Wahrnehmungs- und Aufmerksamkeitsfunktionen

Sturm, W. (2006). WAF. Wahrnehmungs- und Aufmerksamkeitsfunktionen [Manual und Computerprogramm]. Mödling: Schuhfried.

Beschreibung

Die Testbatterie ermöglicht die Erfassung von Teilfunktionen der Aufmerksamkeit. Wichtige Einsatzgebiete liegen in der Neuropsychologie, klinischen Psychologie und Gesundheitspsychologie, Flugspsychologie und Sportpsychologie.

Mit dem von van Zomeren und Brouwer (1994) vorgeschlagenen Modell lassen sich moderne Auffassungen zur Dimensionalität von Aufmerksamkeit zusammenfassen. Zu den zentralen Aspekten gehört nach diesem Modell zunächst die Unterscheidung von Intensitäts- und Selektivitätsaspekten der Aufmerksamkeit, die jeweils in spezifischere Komponenten zu differenzieren sind.

Der Intensitätsaspekt der Aufmerksamkeit umfasst die Komponenten Alertness und Vigilanz, basale Prozesse der kurzfristigen und der längerfristigen Aufmerksamkeitsaktivierung bzw. der Aufrechterhaltung einer solchen Aktivierung. Unter dem Aspekt der Selektivität der Aufmerksamkeitsprozesse unterscheidet das Modell zwischen fokussierter bzw. selektiver und geteilter Aufmerksamkeit.

Die räumliche Ausrichtung der Aufmerksamkeit stellt eine zusätzliche, eigenständige Dimension dar, welche im o.a. Modell noch nicht berücksichtigt, in neuere Taxonomien aber aufgenommen wurde. Posner und Raichle (1994) unterscheiden drei Aufmerksamkeitsnetzwerke:

1. Orientierung (entspricht dem Netzwerk der räumlichen Aufmerksamkeitsausrichtung)
2. Vigilance (entspricht der Intensitäts-Dimension)
3. Executive Attention (entspricht in etwa der Selektivitäts-Dimension)

Die computergestützte Testbatterie WAF (Sturm, 2006) besteht aus sechs Testverfahren, die unabhängig voneinander oder in beliebiger Kombination als Testbatterie vorgegeben werden können:

1. WAFA – Alertness
2. WAFV – Vigilanz, Daueraufmerksamkeit
3. WAFS – Selektive Aufmerksamkeit
4. WAFF – Fokussierte Aufmerksamkeit
5. WAFG – Geteilte Aufmerksamkeit
6. WAFR – Räumliche Aufmerksamkeit und Gesichtsfeld/Extinktion-Neglect

Zusätzlich können mit dem Vortest WAFW sensorische Beeinträchtigungen in den Wahrnehmungsfunktionen abgeklärt werden. Ausgewertet werden die Reaktionszeiten bzw. verschiedene Fehlertypen.

Die nach der Probabilistischen Testtheorie konstruierte Testbatterie greift auf Taxonomien der Aufmerksamkeit zurück, die von van Zomeren und Brouwer (1994) sowie Posner und Raichle (1994) vorgeschlagen wurden. Zu den zentralen Aspekten der Aufmerksamkeit gehört nach diesen Modellen die Unterscheidung von Intensitäts- und Selektivitätsaspekten der Aufmerksamkeit, die jeweils in spezifischere Komponenten zu differenzieren sind.

Der Intensitätsaspekt der Aufmerksamkeit umfasst die Komponenten Alertness und Vigilanz, basale Prozesse der kurzfristigen und der längerfristigen Aufmerksamkeitsaktivierung bzw. der Aufrechterhaltung einer solchen Aktivierung. Unter dem Aspekt der Selektivität der Aufmerksamkeitsprozesse unterscheidet das Modell zwischen fokussierter bzw. selektiver und geteilter Aufmerksamkeit. Zusätzlich wurde die noch nicht im Modell berücksichtigte Komponente „Räumliche Ausrichtung der Aufmerksamkeit“ in die Testbatterie aufgenommen.

Altersbereich: ab 7;0 Jahren

Hinweise zur Durchführung

Die Testbatterie WAF ist als **Einzeltest** am Computer auszuführen.

Für jeden der WAF-Tests liegen verschiedene Testformen vor, die es ermöglichen, Dimensionen der Aufmerksamkeit unter verschiedenen Vorgabemodalitäten zu erfassen – so umfassen die WAF-Tests konsequent Subtests für visuelle, auditive und crossmodale Testvorgaben. In einigen Untertests der WAF-Testbatterie werden automatisierte und kontrollierte Aufmerksamkeitsaspekte separat erfasst, indem die Reize entweder durch Intensitätsanhebung stärker

Tab. 9: Die Durchführungsdauer der einzelnen WAF-Tests

| WAF-Test | Durchführungsdauer |
|----------|---|
| WAFW | pro Vortest etwa 2 min |
| WAFV | jeweils etwa 5 min pro Subtest |
| WAFR | je nach Testform zwischen 15 min und 30 min |
| WAFS | pro Testform etwa 5 min |
| WAFG | jeweils etwa 10 min pro Subtest |
| WAFB | jeweils etwa 8 min pro Subtest |
| WAFD | jeweils etwa 12 min pro Subtest |

hervortreten („Popping Out“) oder durch Intensitätsabschwächung schwächer hervortreten und damit kognitiv kontrollierte „top-down“-Prozesse notwendig werden. Beide Aufmerksamkeitsprozesse sind im Alltag relevant, beide können interagieren und beide können selektiv, z.B. durch eine Hirnschädigung, beeinträchtigt werden, da ihnen unterschiedliche zerebrale Netzwerke zugrunde liegen.

WAFW:

Um auszuschließen, dass nicht schon Wahrnehmungsbeeinträchtigungen die Verarbeitung der bei der WAF verwendeten Stimuli beeinflussen und somit eine zuverlässige Aufmerksamkeitsdiagnostik verhindern, kann mit WAFW vor Beginn einer Untersuchung abgeklärt werden, ob die Wahrnehmungsleistung der Probandin/des Probanden für die Durchführung der WAF-Tests ausreicht.

WAFV:

Bei WAFV wird die Reaktionszeit auf einfaches visuelles bzw. auditives Reizmaterial erfasst. Die Reizdarbietung erfolgt ohne bzw. mit einem in der gleichen oder der jeweils anderen Reizmodalität präsentierten Warnreiz (intrinsische vs. phasische Alertness). Eine spezielle Normierung erlaubt es, Ermüdungs- bzw. Belastbarkeitsparameter zu erfassen.

WAFR:

Bei WAFR werden der Probandin/dem Probanden visuelle bzw. auditive Reize vorgegeben, die in seltenen Fällen etwas an Intensität verlieren. Auf diese seltenen Fälle, die in der Durchführungsbedingung Daueraufmerksamkeit etwa 25 % der Reize und in der Durchführungsbedingung Vigilanz etwa 5 % der Reize ausmachen, soll die Probandin/der Proband reagieren.

WAFS:

Die räumliche Ausrichtung der Aufmerksamkeit wird an 4 bzw. 8 räumlichen Positionen in einer Aufgabenstellung ähnlich dem Posner-Paradigma erfasst. Es werden periphere (exogene) und zentrale (endogene) räumliche Hinweisreize verwendet. Bei der Neglect-Prüfung werden Stimuli an verschiedenen Positionen im rechten oder linken visuellen Feld oder simultan an äquivalenten Positionen in beiden Gesichtsfeldhälften dargeboten (Extinktionsbedingung).

WAFI:

Der Probandin/dem Probanden werden – je nach Subtest – visuell oder auditiv relevante Reize vor ablenkendem Reizmaterial vorgegeben. Wenn zwei zuvor definierte Veränderungen der relevanten Reize hintereinander auftreten, soll die Probandin/der Proband reagieren, das übrige Reizmaterial soll ignoriert werden.

WAFD:

Die Probandin/der Proband erhält relevante und irrelevante Reize in einer oder beiden Vorgabemodalitäten und soll auf Veränderungen der relevanten Reize reagieren und die irrelevanten Reize dabei ignorieren.

WAFG:

Die Probandin/der Proband erhält Reizmaterial auf zwei visuellen bzw. einem visuellen und einem auditiven Kanal. Sie/er soll dabei stets beide Kanäle dahingehend überwachen, ob sich einer der Reize zweimal hintereinander verändert hat.

Die Durchführungsdauer ist für die einzelnen WAF-Tests verhältnismäßig gering (vgl. Tab. 9). Es ist daher möglich, auch Testbatterien für komplexere Fragestellungen zusammenzustellen, ohne die Probandin/den Probanden zeitlich bzw. motivational übermäßig zu belasten. In der Regel wird es nicht notwendig sein, jeden Test in allen Reizmodalitäten durchzuführen. Dies muss die Testleiterin/der Testleiter unter Berücksichtigung von Vorinformationen zum Störungsbild einer Patientin/eines Patienten entscheiden.

Für die Durchführung der auditiven und crossmodalen Subtests der WAF-Tests wird ein standardisiertes USB-Headset benötigt.

Hinweise zu Auswertung und Interpretation

Die Testergebnisse sind nur dann vorbehaltlos interpretierbar, wenn die Testperson die sensorischen und motorischen Voraussetzungen für eine einwandfreie Testdurchführung aufweist.

Es stehen bevölkerungsrepräsentative Normen (Prozentränge und T-Werte) im Umfang von $n = 295$ Personen (Altersbereich: 16–77 Jahre) zur Verfügung, die im Zeitraum von Dezember 2005 bis April 2006 im Forschungslabor der Schuhfried GmbH erhoben wurden (Häusler & Sturm, 2009). Die Normen liegen sowohl für die Gesamtstichprobe als auch getrennt nach Bildungsgruppen vor. Darüber hinaus bieten alle WAF-Tests von Alterseffekten bereinigte Rohwerte für die Hauptvariablen an, was eine besonders effiziente Methode der Altersnormierung darstellt. Zusätzlich existieren Kinder- und Jugendlichennormen ($n = 270$) für den Altersbereich 7 bis 17 Jahre. Für alle WAF-Tests liegen Auswertungen der Reaktionszeiten bzw. der verschiedenen Fehlertypen vor. Für die meisten Variablen wird zusätzlich ein Normvergleich (Prozentränge und T-Werte) durchgeführt.

Testgüte

Objektivität

Da die Testvorgabe, die Registrierung und die Auswertung der Resultate standardisiert und computergesteuert erfolgen, ist von Objektivität in Durchführung und Auswertung auszugehen.

Reliabilität

Für die Hauptvariablen der WAF-Tests ergeben sich sehr gute Reliabilitäten (Cronbachs α). Die Werte liegen je nach Subtest bzw. Testform und Version zwischen $\alpha = .88$ und $\alpha = .99$. Die Ein-

dimensionalität der Subtests im Sinne der Probabilistischen Testtheorie wurde in Modelltests nach dem Latency-Modell von Scheiblechner (1979, 1985) bestätigt (Häusler & Sturm, 2009).

Validität

Eine Prüfung der Konstruktvalidität an einer Stichprobe von $n = 256$ erwachsenen Probandinnen/Probanden (Häusler & Sturm, 2009) konnte das theoretische Modell, das der Testbatterie WAF zugrunde liegt, in konfirmatorischen Faktorenanalysen empirisch bestätigen und gegenüber alternativen Modellen abgrenzen: In dem linearen Strukturgleichungsmodell mit der höchsten Anpassungsgüte spiegeln sich sowohl die Unterscheidung der Aufmerksamkeitsaspekte „Intensität“ und „Selektivität“ als auch die unterschiedlichen Darbietungsmodalitäten (visuell, auditiv, crossmodal) wider. An der gleichen Stichprobe ergab eine gemeinsame exploratorische Faktorenanalyse der WAF-Tests mit weiteren Tests, dass Verfahren, die zur konvergenten Validierung verwendet wurden (Cognitrone, vgl. Wagner & Karner, 2001; Determinationstest, vgl. Schuhfried, 1998; Reaktionstest, vgl. Schuhfried & Prieler, 1997), hohe Ladungen auf einem Faktor aufwiesen, der die Selektivität der Aufmerksamkeit abbildet. Die Standard Progressive Matrices (SPM, vgl. Raven, Raven & Court, 1997) luden erwartungsgemäß auf keinem der Aufmerksamkeitsfaktoren, d.h. im Sinne diskriminanter Validität konnte gezeigt werden, dass die Testbatterie WAF verschiedene Aufmerksamkeitsfunktionen unabhängig von der allgemeinen Intelligenz abbildet.

Kommentar der Psyndex-Testreviewerinnen und -reviewer

Auf Basis einer theoriegeleiteten Testkonstruktion ermöglichen die Testverfahren der WAF eine differenzierte Untersuchung nahezu aller heute als relevant erachteter Teilfunktionen der Aufmerksamkeit. Erste Reliabilitäts- und Validitätsbefunde sprechen für eine hohe psychometrische Brauchbarkeit des Verfahrens; eine Prüfung der Retest-Reliabilität und der klinischen Validität steht jedoch noch aus (Günther, 2009).

Autorinnenkommentar

Die WAF eignet sich sehr gut für eine umfassende Abklärung eines möglichen Aufmerksamkeitsdefizits bei Jugendlichen und mit Ein-



schränkungen auch bei Kindern. Problematisch kann bei den Jugendlichen oft eine mangelnde Arbeitseinstellung oder fehlende Motivation sein, die hier zu Fehleinschätzungen führen kann. Für kleinere Kinder bietet sich alternativ die KiTAP an. Es empfiehlt sich, die WAF-Batterie in mehrere Sitzungen zu unterteilen, um Müdigkeitseffekte oder ein Absinken der Motivation zu vermeiden.

Literatur

- Günther, T. (2009). *Wahrnehmungs- und Aufmerksamkeitsfunktionen (WAF). Untertest aus dem Wiener Testsystem (WTS)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 320–330). Göttingen: Hogrefe.
- Häusler, J. & Sturm, W. (2009). *Konstruktvalidierung einer neuen Testbatterie für Wahrnehmungs- und Aufmerksamkeitsfunktionen (WAF)*. *Zeitschrift für Neuropsychologie*, 20 (4), 327–339.
- Posner, M. I. & Raichle, M. E. (1994). *Images of mind*. New York, NY: Scientific American Books.
- Raven, J., Raven, J. C. & Court, J. H. (1997). *Handanweisung Standard Progressive Matrices Plus (SPM Plus)*. Mödling: Schuhfried.
- Scheiblechner, H. (1979). *Specifically objective stochastic latency mechanisms*. *Journal of Mathematical Psychology*, 19, 18–38.
- Scheiblechner, H. (1985). *Psychometric models for speed-test construction: The linear exponential model*. In S. E. Embretson (Hrsg.), *Test design: Developments in psychology and psychometrics* (S. 219–244). New York, NY: Academic Press.
- Schuhfried, G. (1998). *Wiener Determinationstest*. Mödling: Schuhfried.
- Schuhfried, G. & Prieler, J. (1997). *Wiener Reaktionstest*. Mödling: Schuhfried.
- Sturm, W. (2006). *WAF. Wahrnehmungs- und Aufmerksamkeitsfunktionen*. Mödling: Schuhfried.
- Van Zomeren, A. H. & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. Oxford, UK: Oxford University Press.
- Wagner, M. & Karner, T. (2001). *Cognitrone*. Mödling: Schuhfried.

Zusammenfassung

Die vorgestellten Verfahren erlauben in Kombination mit Explorationsgesprächen, anamnestischen Daten, Informationen aus Fremd- und Selbstbeobachtungen, psychodiagnostischen Gesprächen, Fragebögen und Arbeitsproben eine umfassende diagnostische Abklärung eines möglichen Aufmerksamkeitsdefizits.

Problematisch ist bei der Aufmerksamkeitsdiagnostik, dass es keine Aufgaben gibt, die ausschließlich Aufmerksamkeit oder Konzentration verlangen. Die Tests verwenden daher meist solche Aufgaben, die den Einfluss von speziellen Fähigkeiten und Fertigkeiten auf die Testleistung minimieren. Testet man Probandinnen/Probanden mit einer Teilleistungsstörung, kann es aber trotzdem sein, dass diese stark benachteiligt sind, insbesondere dann, wenn Rechenfähigkeit, Lesegeschwindigkeit oder Merkfähigkeit Einfluss auf das Testergebnis haben.

Bei Konzentrationsaufgaben ist zudem die aufgebrachte Leistungsmotivation von großer Bedeutung. Es lässt sich insbesondere bei Jugendlichen nicht immer feststellen, ob die Testperson nicht besser (schneller) arbeiten konnte oder wollte.

Weiters zeigt die Erfahrung, dass Kinder in der Testsituation manchmal unauffällig sind, im Schulalltag aber durch Konzentrationsprobleme auffällig werden. Umso wichtiger ist es daher, in der Diagnostik dieser Komplexität Rechnung zu tragen und Lehrer/innen und Eltern zu befragen sowie Arbeitsproben zu analysieren, also in Anlehnung an Cattell (1973) „Life-, Questionnaire- und Test-Daten“ zu erheben.

Besonders im Vorschulbereich besteht noch Bedarf an kindgerechten Verfahren, um die Frühdiagnostik zu erleichtern und so zur Prävention beitragen zu können.

Ein weiteres Entwicklungspotenzial besteht im Bereich der computergestützten Diagnostik (Reaktionszeitanalysen, dynamische Vorgaben, adaptives Testen, auditive Vorgaben).

Gesamt-Literaturverzeichnis



- Abels, D. (1954). *Konzentrations-Verlaufs-Test K-V-T*. Göttingen: Hogrefe.
- Abels, D. (1961). *Konzentrations-Verlaufs-Test K-V-T (überarbeitete Version)*. Göttingen: Hogrefe.
- Abels, D. (1974). *Konzentrations-Verlaufs-Test KVT (2., verbesserte Auflage)*. Göttingen: Hogrefe.
- Amthauer, R. (1970). *Intelligenz-Struktur-Test (I-S-T 70, 3. erw. Auflage)*. Göttingen: Hogrefe.
- Amthauer, R. (1973). *Intelligenz-Struktur-Test (I-S-T 70, 4. Auflage)*. Göttingen: Hogrefe.
- Arnold, K., Stender, B. & Marschner, G. (1976). *Revisions-Test (Rev.T.)*. *Diagnostica*, 22, 138–139.
- Arnold, W. (1958). *Neue Erfahrungen mit dem Pauli-Test. Eine kritische Erwiderung*. *Zeitschrift für experimentelle und angewandte Psychologie*, 5, 534–541.
- Arnold, W. (1961). *Beiträge zur Faktorenanalyse des Pauli-Tests*. *Psychologische Beiträge*, 5, 312–327.
- Arnold, W. (1975). *Der Pauli-Test. Anweisung zur sachgemässen Durchführung, Auswertung und Anwendung des Kräpelinischen Arbeitsversuchs (5., korrigierte Auflage)*. Berlin: Springer.
- Arnold, W. (1980). *Lexikon der Psychologie*. Freiburg: Herder.
- Atkinson, R. C. & Shiffrin, R. M. (1971). *The control of short-term memory*. *Scientific American*, 225, 82–90.
- Barchmann, H., Ettrich, K. U. & Kinze, W. (1987). *Methodische Aspekte zur Diagnostik der Konzentrationsfähigkeit im Kindesalter*. *Psychologie für die Praxis*, 2, 156–160.
- Bartenwerfer, H. (1963). *Mitteilungen zur Frage der Reliabilität dreier Merkmale des Pauli-Tests*. *Diagnostica*, 9, 77–79.
- Bartenwerfer, H. (1964). *Allgemeine Leistungstests*. In R. Heiss (Hrsg.), *Handbuch der Psychologie. Band 6, Psychologische Diagnostik* (S. 385–410). Göttingen: Hogrefe.
- Bartenwerfer, H. (1982). *Allgemeine Leistungsdiagnostik*. In K.-J. Grottmann & L. Michel (Hrsg.), *Intelligenz- und Leistungsdiagnostik (Enzyklopädie der Psychologie, Themenbereich B, Serie II, Band 3. S. 482–512)*. Göttingen: Hogrefe.
- Barth, A. (2002). *Der Differentielle Leistungstest-KE (DL-KE)*. In E. Brähler, H. Holling, D. Leutner & F. Petermann (Hrsg.), *Brickenkamp Handbuch psychologischer und pädagogischer Tests (3., vollständig überarbeitete und erweiterte Auflage, Band 1, S. 270–272)*. Göttingen: Hogrefe.
- Bäumler, G. (1964). *Zur Faktorenstruktur der Pauli-Testleistung unter besonderer Berücksichtigung des sogenannten numerischen Faktors*. *Diagnostica*, 10, 107–119.
- Bäumler, G. (1985). *Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop*. *Handanweisung*. Göttingen: Hogrefe.
- Bäumler, G. & Breitenbach, W. (1970). *Zusammenhänge zwischen Intelligenz, Konzentration, Angst und Leistungsmotivation bei einer studentischen Stichprobe*. *Psychologie und Praxis*, 14, 37–40.
- Bäumler, G. & Weiss, R. (1967). *Eine Zweifaktorentheorie der nach der TAT-Methode gemessenen Leistungsmotivation (Heckhausen)*. *Psychologie und Praxis*, 11, 23–45.
- Beck, L., Heusinger, A., Boecker, M., Niemann, H. & Gauggel, S. (2008). *Convergent and predictive validity of two computerized attention tests in brain-damaged patients*. *Zeitschrift für Neurologie*, 19 (4), 213–222.
- Becker, M., Sturm, W., Willmes, K. & Zimmermann, P. (1996). *Normierungsstudie zur Aufmerksamkeitstestbatterie (TAP) von Zimmermann und Fimm*. *Zeitschrift für Neuropsychologie*, 7 (1), 3–15.
- Bernstein, J. & Kuenzel, R. (1976). *Vergleich der Intelligenz jugendlicher Cerebralparetiker und jugendlicher Körperbehinderter ohne Hirnschädigung*. *Sonderpädagogik*, 6, 119–128.
- Birke, W. & Köhne, H. (1967). *Zur Kritik des KVT (nach Abels)*. *Diagnostica*, 13, 130–133.

- Bochow, R. (1967). Erfahrungen mit einer Abwandlungsform des Pauli-Tests (20 x 1 Min.). *Zeitschrift für experimentelle und angewandte Psychologie*, 14, 570–599.
- Bodenburg, S., Popp, B. & Kawski, S. (2001). Ergänzende Normdaten zu dem Untertest Alertness aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP) in der Altersgruppe 60+. *Zeitschrift für Neuropsychologie*, 12 (2), 125–130.
- Bonkowsky, D. (1999). *Intelligenz – familiäre und schulische Aspekte*. Unveröffentlichte Zulassungsarbeit, Universität Regensburg.
- Borchert, J., Knopf-Jerchow, H. & Dahbashi, A. (1991). Testdiagnostische Verfahren in Vor-, Sonder- und Regelschulen. Ein kritisches Handbuch für Praktiker (Testkurzdarstellung DL-KG – Differentieller Leistungstest KG: S. 322–323). Heidelberg: Asanger.
- Borchert, J., Knopf-Jerchow, H. & Dahbashi, A. (1991). Testdiagnostische Verfahren in Vor-, Sonder- und Regelschulen. Ein kritisches Handbuch für Praktiker (Testkurzdarstellung DL-KE – Differentieller Leistungstest KE: S. 321–322). Heidelberg: Asanger.
- Borkenau, P. & Ostendorf, F. (1993). NEO-FFI. NEO-Fünf-Faktoren Inventar nach Costa und McCrae. Göttingen: Hogrefe.
- Bourdon, B. (1895). Observations comparatives sur la reconnaissance, la discrimination et l'association. *Revue philosophique*, 40, 153–185.
- Brainerd, C. J. & Reyna, V. F. (1998). Fuzzy-Trace Theory and children's false memories. *Journal of Experimental Child Psychology*, 71, 81–129.
- Brickenkamp, R. (1962). Test d2. Aufmerksamkeits-Belastungs-Test. Göttingen: Hogrefe.
- Brickenkamp, R. (1963). Einige theoretische Grundzüge und praktische Methoden der Zuverlässigkeitsschätzung psychodiagnostischer Verfahren. *Psychologische Rundschau*, 14, 13–25.
- Brickenkamp, R. (1981). Test d2. Aufmerksamkeits-Belastungs-Test (6. Auflage). Göttingen: Hogrefe.
- Brickenkamp, R. (1987). Test d2. Aufmerksamkeits-Belastungs-Test (7. Auflage). Göttingen: Hogrefe.
- Brickenkamp, R. (1991). Fehlinterpretationen von Testleistungen? Anmerkungen zum Beitrag „Konzentrationsleistung ohne Konzentration?“. *Diagnostica*, 37, 52–57.
- Brickenkamp, R. (1994). Test d2. Aufmerksamkeits-Belastungs-Test (8., erweiterte und neu gestaltete Auflage). Göttingen: Hogrefe.
- Brickenkamp, R. (2002). Test d2. Aufmerksamkeits-Belastungs-Test (9., vollständig überarbeitete und neu normierte Auflage). Göttingen: Hogrefe.
- Brickenkamp, R. & Karl, G. A. (1986). Geräte zur Messung von Aufmerksamkeit, Konzentration und Vigilanz. In R. Brickenkamp (Hrsg.), *Handbuch apparativer Verfahren in der Psychologie* (S. 195–211). Göttingen: Hogrefe.
- Brickenkamp, R., Schmidt-Atzert, L. & Liepmann, D. (2010). Test d2 – Revision Aufmerksamkeits- und Konzentrationstest. Göttingen: Hogrefe.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.
- Buckner, D. N. & McGrath, J. J. (1963). *Vigilance: A symposium*. New York.
- Bühner, M. (2009). Alertness. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 59–66). Göttingen: Hogrefe.
- Bühner, M. (2009). Arbeitsgedächtnis. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung

- (TAP Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 797–803). Göttingen: Hogrefe.
- Bühner, M. (2009). Gesichtsfeld-/Neglectprüfung. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 138–143). Göttingen: Hogrefe.
- Bühner, M. (2009). Geteilte Aufmerksamkeit. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Aufmerksamkeit, Gedächtnis und exekutive Funktionen (S. 265–272). Göttingen: Hogrefe.
- Bühner, M. (2009). Go/No-Go-Test. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 224–229). Göttingen: Hogrefe.
- Bühner, M. (2009). Inkompatibilität. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 197–203). Göttingen: Hogrefe.
- Bühner, M. (2009). Intermodaler Vergleich. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 259–264). Göttingen: Hogrefe.
- Bühner, M. (2009). Reaktionswechsel (Flexibilität). Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 289–295). Göttingen: Hogrefe.
- Bühner, M. (2009). Verdeckte Aufmerksamkeitsverschiebung. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 124–129). Göttingen: Hogrefe.
- Bühner, M. (2009). Vigilanztest. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 110–116). Göttingen: Hogrefe.
- Bühner, M. (2009). Visuelles Scanning. Untertest aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP, Version 1.7). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 117–123). Göttingen: Hogrefe.
- Bühner, M., Ziegler, M., Bohnes, B. & Lauterbach, K. (2006). Übungeffekte in den TAP Untertests Test Go/No-Go und Geteilte Aufmerksamkeit sowie dem Aufmerksamkeitsbelastungstest (d2). *Zeitschrift für Neuropsychologie*, 17 (3), 191–199.

- Burgstaller, S. (2001). Validierungsstudien zum PSB-R 6–13. Unveröffentlichte Diplomarbeit, Universität Regensburg.
- Büttner, G. (1994). Aspekte zur Validität eines schulnahen Konzentrationstests. Poster auf dem 39. Kongress der Deutschen Gesellschaft für Psychologie. Hamburg: Universität.
- Büttner, G. & Kurth, E. (1996). Konzentrationsleistungen in der schulnahen Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK) – eine regionale Vergleichsstudie. *Zeitschrift für Pädagogische Psychologie*, 10 (3/4), 187–197.
- Büttner, G. & Schmidt-Atzert, L. (2004, Hrsg.). *Diagnostik von Konzentration und Aufmerksamkeit*. Göttingen: Hogrefe.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Cattell, R. B. (1973). *Die empirische Erforschung der Persönlichkeit*. Weinheim: Beltz.
- Cohen, R. A. (1993). *The neuropsychology of attention*. New York: Plenum Press.
- Dahms, P. & Meissner, A. (1995). Korrelationsstatistische Ergebnisse zur Validität von zwei computergestützten Signal-Entdeckungs-Tests. *Diagnostica*, 41 (2), 95-107.
- Deutsch, J. A. & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review*, 70, 80–90.
- Dirks, H. (1957). Über die Reichweite des KVT (nach Abels) zur Diagnose der Willensstruktur (nach Herwig-Dirks). *Diagnostica*, 3, 57–62.
- Dorsch, F. (1987). *Psychologisches Wörterbuch*. (11., ergänzte Auflage). Bern: Huber.
- Drechsler, R. (2009). Testbatterie zur Aufmerksamkeitsprüfung TAP (Version 2.1). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 296–310). Göttingen: Hogrefe.
- Drechsler, R. (2009). Testbatterie zur Aufmerksamkeitsprüfung TAP-M (Version Mobilität 1.2). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 311–319). Göttingen: Hogrefe.
- Dreisörner, T. & Georgiadis, J. (2011). Sensitivität und Spezifität computergestützter Verfahren zur Diagnostik von Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) im Kindes- und Jugendalter – Die Testbatterie zur Aufmerksamkeitsprüfung (TAP) und Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP). *Empirische Sonderpädagogik*, 3 (1), 3–19.
- Düker, H. & Lienert, G. A. (1965). *Konzentrations-Leistungs-Test. K-L-T*. Göttingen: Hogrefe.
- Düker, H. (1943). Psychopharmakologische Untersuchungen über die Wirkung von Keimdrüsenhormonen auf die geistige Leistungsfähigkeit. *Archiv für experimentelle Pathologie und Pharmakologie*, 202, 262-313.
- Düker, H. (1949). Über ein Verfahren zur Bestimmung der geistigen Leistungsfähigkeit. *Psychologische Forschung*, 23, 10–24.
- Düker, H., Lienert, G. A., Lukesch, H. & Mayrhofer, S. (2001). *KLT-R. Konzentrations-Leistungs-Test – Revidierte Fassung*. Göttingen: Hogrefe.
- Dunkel, B. (1999). Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK). In S. Grubitzsch, *Testtheorie – Testpraxis. Psychologische Tests und Prüfverfahren im kritischen Überblick* (2. unveränderte Auflage der vollständig überarbeiteten und erweiterten Neuausgabe 1991. Kurzanalyse S. 435–437). Eschborn: Klotz.

- Echterhoff, J., Golzarandi, A. G., Morsch, D., Lehmkuhl, G. & Sinzig, J. (2009). Ein Vergleich computergestützter Testverfahren zur neuropsychologischen Diagnostik bei Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung. *Zeitschrift für Neuropsychologie*, 20 (4), 313–325.
- Eikemann, A., Petermann, F. & Daseking, M. (2009). Aufmerksamkeitsstörungen nach Schlaganfällen im Kindesalter. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 36 (6), 419–426.
- Ettrich, C. & Ettrich, K. U. (2004). Entwicklung der Konzentrationsfähigkeit im Vorschulalter und diagnostische Probleme ihrer Erfassung. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 205–231). Göttingen: Hogrefe.
- Ettrich, K. U. & Ettrich, C. (2005). *KHV-VK. Konzentrations-Handlungsverfahren für Vorschulkinder*. Göttingen: Hogrefe.
- Fay, E. (1996). Frankfurter Aufmerksamkeits-Inventar. FAIR (Testrezension). In E. Fay, *Tests unter der Lupe. Aktuelle Leistungstests – kritisch betrachtet* (Band 1, S. 29–40). Heidelberg: Asanger.
- Fay, E. (1996). *Tests unter der Lupe. Aktuelle Leistungstests – kritisch betrachtet* (Band 1. Testrezension INKA, S. 57–65). Heidelberg: Asanger.
- Fimm, B. (1997). Mikroanalyse von Aufmerksamkeitsprozessen. In S. Gauggel & G. Kerkhoff (Hrsg.), *Fallbuch der Klinischen Neuropsychologie. Praxis der Neurorehabilitation* (S. 25–38). Göttingen: Hogrefe.
- Fischer, M. & Ulich, H. (1961). Über die Abhängigkeit einer kurzzeitigen Konzentrationsleistung von der Tageszeit bei Kindern und Jugendlichen verschiedenen Alters. *Zeitschrift für experimentelle und angewandte Psychologie*, 8, 282–296.
- Fitts, P. M. & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46, 199–210.
- Földényi, M., Giovanoli, A., Tagwerker-Neuenschwander, F., Schallberger, U. & Steinhausen, H.-C. (1999). Die Aufmerksamkeitsleistungen von 6–10jährigen Kindern in der TAP. *Zeitschrift für Neuropsychologie*, 10 (2), 87–102.
- Földényi, M., Giovanoli, A., Tagwerker-Neuenschwander, F., Schallberger, U. & Steinhausen, H.-C. (2000). Reliabilität und Retest-Stabilität der Testleistungen von 7–10-jährigen Kindern in der computerunterstützten TAP. *Zeitschrift für Neuropsychologie*, 11 (1), 1–11.
- Földényi, M., Imhof, K. & Steinhausen, H.-C. (2000). Klinische Validität der computerunterstützten TAP bei Kindern mit Aufmerksamkeits-/Hyperaktivitätsstörungen. *Zeitschrift für Neuropsychologie*, 11 (3), 154–167.
- Formann, A. K. & Piswaenger, K. (1979). *Wiener Matrizen-Test*. Weinheim: Beltz.
- Fraser, D. C. (1957). *A study of vigilance and fatigue*. (Doctoral dissertation, University of Edinburgh).
- Freytag, G. (1962). *Experimentalpsychologie und mathematische Typenanalyse bei Depressionszuständen*. *Zeitschrift für Psychologie*, 166, 241–297.
- Gebauer, T. (1965). Vergleichende Untersuchungen über den Voraussagewert von Aufnahmeprüfung und Testuntersuchung für den Erfolg auf weiterführenden Schulen. In *Schülerkonflikt und Schülerhilfe*. Weinheim: Beltz.
- Goldhammer, F. & Moosbrugger, H. (2006). Aufmerksamkeit. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 16–33). Heidelberg: Springer.
- Goldhammer, F., Moosbrugger, H. & Schweizer, K. (2007). On the separability of cognitive abilities related to Posner's attention components and their contributions to conceptually distinct atten-

- tion abilities related to working memory, action theory, and psychometric assessment. *European Psychologist*, 12, 103–118.
- Gottwald, D., Knan, A. & Marsch, E. (1997). Entwicklung einer neuen Variante eines Konzentrations-tests. Unveröffentlichte Forschungspraktikumsarbeit, Universität Regensburg.
- Gray, J. A. & Wedderburn, A. A. (1960). Grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 12, 180–184.
- Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gruhnwald, E. & Ulich, H. (1959). Über die Anwendung des Arbeitsversuchs bei psychisch Beeinträchtigten unter besonderer Berücksichtigung einer Untersuchung an taubstummen Jugendlichen. *Zeitschrift für experimentelle und angewandte Psychologie*, 6, 274–292.
- Guilford, J. P. (1954). *Psychometric Methods*. New York: McGraw-Hill.
- Günther, R. (1986). Anwendung der statistischen Entscheidungstheorie (SDT) bei der visuellen Signalerkennung von Fahrzeugführern. *Verkehrsmedizin*, 33, 237–242.
- Günther, T. (2009). Signal-Detection (SIGNAL). Untertest aus dem Wiener Testsystem (WTS). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 130–137). Göttingen: Hogrefe.
- Günther, T. (2009). Wahrnehmungs- und Aufmerksamkeitsfunktionen (WAF). Untertest aus dem Wiener Testsystem (WTS). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 320–330). Göttingen: Hogrefe.
- Gutewa, J. (1957). Der Wert psychologischer Tests für die Untersuchung der Wirkung von Medikamenten auf psychische Funktionen. *Psychiatria et Neurologia*, 134, 224–235.
- Hagemeister, C. & Westhoff, K. (1998). FAIR Frankfurter Aufmerksamkeits-Inventar. H. Moosbrugger & J. Öhlschlägel. *Diagnostica*, 44 (4), 225–229.
- Hallwachs, M. (1994). Messen FAIR und d2 dasselbe? Unveröffentlichtes Manuskript. Hamburg: Psychologischer Dienst des Berufsbildungswerks.
- Hänze, M. (1997). Mood and the Stroop interference effect. *Psychologische Beiträge*, 39 (3), 229–235.
- Hasselhorn, M. (1996). Kategoriales Organisieren bei Kindern. Göttingen: Hogrefe.
- Hasselhorn, M. & Grube, D. (2006). Gedächtnisentwicklung (Grundlagen). In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung (Enzyklopädie der Psychologie, Serie Entwicklungspsychologie, Band 2, S. 271–325)*. Göttingen: Hogrefe.
- Häusler, J. & Sturm, W. (2009). Konstruktvalidierung einer neuen Testbatterie für Wahrnehmungs- und Aufmerksamkeitsfunktionen (WAF). *Zeitschrift für Neuropsychologie*, 20 (4), 327–339.
- Heber, I. A. (2009). Revisions-Test (Rev. T.). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 187–196). Göttingen: Hogrefe.
- Heck-Möhling, R. (1993). KT 3–4. Konzentrationstest für 3. und 4. Klassen (2. Auflage). Weinheim: Beltz.
- Heinemann, D. (2009). Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren*. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen (S. 971–978). Göttingen: Hogrefe.

- Heinrich, H. C. (1973). Einige Bemerkungen zum d2-Durchstreichtest nach Brickenkamp. *Diagnostica*, 19, 118–124.
- Heller, K. A. & Perleth, C. (1991). Der Pauli-Test. Anweisung zur sachgemäßen Durchführung, Auswertung und Anwendung des Kräpelinischen Arbeitsversuches nach R. Pauli, herausgegeben von W. Arnold. In K. A. Heller (Hrsg.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (S. 184–185). Bern: Huber.
- Heller, K. A. & Perleth, C. (2000). KFT 4-12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz.
- Hellwig, H.-J. (1975). Zur Differenzierung von Intelligenztest- und Konzentrationsleistungen. Unveröffentlichte Dissertation, Technische Universität Berlin.
- Henning, H. (1925). Die Untersuchung der Aufmerksamkeit. In E. Aberhalden (Hrsg.), *Handbuch der biologischen Arbeitsmethoden* (Abt. VI, Teil B. S. 598-802). Berlin: Urban & Schwarzenberg.
- Herwig, B. & Dirks, H. (1951). Strukturtypen des praktischen Handelns. *Psychologische Rundschau*, 2, 20–28.
- Heubrock, D. & Petermann, F. (2001). *Aufmerksamkeitsdiagnostik*. Göttingen: Hogrefe.
- Heyde, G. (1995). *Inventar Komplexer Aufmerksamkeit (INKA)*. Handanweisung. Frankfurt am Main: Swets Test Services.
- Heyde, G. (2004). INKA – Inventar Komplexer Aufmerksamkeit. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 133–142). Göttingen: Hogrefe.
- Heyde, G. (2004). INKA. Inventar Komplexer Aufmerksamkeit. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte Auflage, S. 409–412). Lengerich: Pabst.
- Hitpass, J. (1961). Vergleichende Untersuchung über den Voraussagewert von Aufnahmeprüfung und Testprüfung zur Erfassung der Eignung für die weiterführenden Schulen. *Schule und Psychologie*, 8, 65–71.
- Hitpass, J. (1963). Bericht über eine sechsjährige Bewährungskontrolle von Aufnahmeprüfung und Testprüfung. *Schule und Psychologie*, 10, 211–218.
- Hofgärtner, S. (1998). Zusammenhang zwischen Konzentration und Intelligenz. Unveröffentlichte Zulassungsarbeit, Universität Regensburg.
- Holling, H., Preckel, F. & Vock, M. (2004). *Intelligenzdiagnostik. Kompendien psychologische Diagnostik – Band 6*. Göttingen: Hogrefe.
- Hörmann, H. (1956). Zum Problem der psychischen Starrheit (Rigidität). *Zeitschrift für experimentelle und angewandte Psychologie*, 3, 662–683.
- Horn, W. (1969). *Prüfsystem für Schul- und Bildungsberatung. PSB*. Göttingen: Hogrefe.
- Horn, W. (1983). *Leistungsprüfsystem (LPS)* (2. Auflage). Göttingen: Hogrefe.
- Jäger, A. O. (1960). Zum prognostischen Wert psychologischer Eignungsuntersuchungen. *Psychologische Rundschau*, 11, 160–178.
- Jäger, R. (1973). Bemerkungen zur „Untersuchung über den Zusammenhang zwischen Intelligenz (Wechsler) und Konzentrationsfähigkeit (Test d2 nach Brickenkamp) von W. Wiese und G. Kroj“. *Zeitschrift für experimentelle und angewandte Psychologie*, 20, 572–574.
- Johnston, W. A. & Wilson, J. (1980). Perceptual processing of non-targets in an attention task. *Memory & Cognition*, 8, 372–377.

- Jonkisz, E., Moosbrugger, H. & Mildner, D. (2006). Die „Frankfurt Study“ zur Vorhersage des Studienerfolgs. In B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer & O. Vitouch (Hrsg.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (S. 342–348). Lengerich: Pabst.
- Jost, K. (1992). Die Interferenzneigung Schizophrener in der Version des Stroop-Tests nach G. Bäuml. *Psychologische Beiträge*, 34, 3–15.
- Jung, T. (1992). *Untersuchungen zu Aspekten der Validität psychologischer Konzentrationstests: Ein Vergleich zwischen dem Aufmerksamkeits-Belastungs-Test d2 und dem Frankfurter Aufmerksamkeits-Inventar (FAIR)*. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Kallenbach, K. (1978). Auswirkungen unterschiedlicher Verfahren zum Pauli-Arbeitsversuch. Ein Vergleich von Abwandlungsformen: Kurzform eines Papier-Bleistift-Verfahrens – apparatives Verfahren. *Psychologie und Praxis*, 22, 157–163.
- Karner, T. (2000). Sind verkehrspsychologische Testverfahren geeignete Instrumente, um mögliche Leistungsminderungen alkoholauffälliger Kraftfahrer aufzuzeigen? *Report Psychologie*, 25 (9), 578–583.
- Kastner-Koller, U. & Deimann, P. (1998). *Der Wiener Entwicklungstest (WET). Ein allgemeines Entwicklungstestverfahren für Kinder von 3 bis 6 Jahren*. Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (2002). *WET. Wiener Entwicklungstest. Ein Verfahren zur Erfassung des allgemeinen Entwicklungsstandes bei Kindern von 3 bis 6 Jahren* (2., überarbeitete und neu normierte Auflage). Göttingen: Hogrefe.
- Keller, I. (2009). *Cognitron (COG). Untertest aus dem Wiener Testsystem (WTS)*. In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 237–243). Göttingen: Hogrefe.
- Kleber, E. W. & Kleber, G. (1974). *Differentieller Leistungstest – KE (DL-KE)*. Göttingen: Hogrefe.
- Kleber, E. W., Kleber, G. & Hans, O. (1999). *Differentieller Leistungstest-KG (DL-KG). Test zur Erfassung des Leistungsverhaltens bei konzentrierter Tätigkeit im Grundschulalter* (2., korrigierte Auflage). Göttingen: Hogrefe.
- Koch, J. & Pleissner, S. (1985). *Konzentrations-Handlungs-Verfahren (KHV)*. Berlin: Psychodiagnostisches Zentrum.
- Kohlmann, T. (1954). *Psychologische Untersuchungen mit Rorschach- und Kräpelin-Versuch an vegetativen Neurosen*. *Zeitschrift für diagnostische Psychologie und Persönlichkeitsforschung*, 2, 101–126.
- Krampen, G. (2007). *KKA. Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige*. Göttingen: Hogrefe.
- Krampen, G. (2008). Kognitive Entwicklung bei 3- bis 8-Jährigen. Konzentrationsleistung und Übergang vom vor-operatorischen zum konkret-operatorischen Denken. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40 (2), 79–86.
- Krampen, G. (2009). *Entwicklungsförderung bei Kindergartenkindern mit Konzentrationsschwächen: Vergleichende Untersuchungen unter Einbezug von Entspannungsverfahren*. *Entspannungsverfahren*, 26 (26), 8–25.
- Krowatschek, D., Albrecht, S. & Krowatschek, G. (2004). *Marburger Konzentrationstraining für Kin-*

- dergarten- und Vorschulkinder (MKT). Göttingen: Hogrefe.
- Krumm, S., Schmidt-Atzert, L. & Eschert, S. (2008). Investigating the structure of attention. How do test characteristics of paper-pencil sustained attention tests influence their relationship with other attention tests? *European Journal of Psychological Assessment*, 24 (2), 108–116.
- Küchemann, D. (2002). Erprobung und Validierung einer französischen Version des Frankfurter Aufmerksamkeits-Inventars: Eine interkulturelle Äquivalenzuntersuchung. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Kühne, W. (2009). Konzentrations-Verlaufs-Test (K-V-T). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 92–98). Göttingen: Hogrefe.
- Kunert, H. J., Derichs, G. & Irle, E. (1996). Entwicklung von Aufmerksamkeitsfunktionen im Kindesalter: Ergebnisse einer vorläufigen Normierung der computergestützten Testbatterie zur Aufmerksamkeitsprüfung (TAP) an 9- bis 12jährigen Kindern. *Zeitschrift für Neuropsychologie*, 7 (2), 92–113.
- Kurth, E. (1984). Testreihe zur Prüfung der Konzentrationsfähigkeit TPK (Testmappe mit Manual, je einem Musterrechenblatt und Musterauswertungsbogen, zwei Abschreibetexten und zwei Schablonen). Berlin: Psychodiagnostisches Zentrum.
- Kurth, E. & Büttner, G. (1999). Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK) (2., neu bearbeitete Auflage). Göttingen: Hogrefe.
- Kurth, E. & Büttner, G. (2004). Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK). In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 143-159). Göttingen: Hogrefe.
- Lamberti, G. (1983). Untersuchungen zur Farbe-Wort-Interferenz bei schizophrenen Erkrankungen. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 4, 141–149.
- Lepach, A. C., Gienger, C. & Petermann, F. (2008). Neuropsychologische Befunde zu Merk- und Lernstörungen bei Kindern anhand des BASIC-MLT. *Zeitschrift für Kinder- und Jugendpsychiatrie- und Psychotherapie*, 36 (6), 389–400.
- Lepach, A. C. & Petermann, F. (2008). *Battery for Assessment in Children. BASIC-MLT. Merk- und Lernfähigkeitstest für 6- bis 16-Jährige*. Bern: Huber.
- Lepach, A. C., Petermann, F. & Schmidt, S. (2007). Neuropsychologische Diagnostik von Merk- und Lernstörungen mit der MLT-C. *Kindheit und Entwicklung*, 16, 16–26.
- Lepach, A. C., Petermann, F. & Schmidt, S. (2008). Comparisons of the BASIC-Memory and Learning Test with the WISC-IV under developmental aspects. *Zeitschrift für Psychologie*, 126 (3), 180–186.
- Les Edition du Centre de Psychologie Appliquée. (1986). *Manuel d'Application du Test d'Intelligence de R.B. Cattell (Culture Free Test). Echelle 3*. Paris: Les Edition du Centre de Psychologie Appliquée.
- Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testpraxis*. Weinheim: Beltz.
- Linderkamp, F. (1996). Zur Homogenität des Störungsbildes und die Notwendigkeit zur Subgruppendifferenzierung. *Kindheit und Entwicklung*, 5, 89–92.
- Lingoes, J. C. (1973). *The Guttman-Lingoes Nonmetric Program Series*. Ann Arbor: Mathesis Press.
- Lobeck, A., Frei, M. & Blöching, R. (1990). *Schweizer Rechentest 4.–6. Klasse*. Basel: Beltz.
- Lukesch, H. (2004). KLT-R. Revidierter Konzentrations-Leistungs-Test. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren* (2., überarbeitete und erweiterte

- Auflage, S. 453–457). Lengerich: Pabst.
- Mackworth, N. H. (1950). *Researches on the measurement of human performance*. Medical Research Council, Special Report, No. 268.
- Marsch, E. (2000). *Konvergente und divergente Validität des KLT-R bei Schüler/innen der 4. bis 6. Schulstufe*. Unveröffentlichte Diplomarbeit, Universität Regensburg.
- Marschner, G. (1960). *Zur Diagnostik eines höheren F % im KLT. Eine vergleichende Untersuchung von Konzentrationsleistungstest und Revisions-Test in der betriebspsychologischen Praxis*. *Diagnostica*, 6, 66–72.
- Marschner, G. (1971). *Der Revisionstest als allgemeiner Leistungstest. Beiträge zur Standardisierung und Validierung des Revisions-Tests nach B. Stender*. *Diagnostica*, 17, 14–26.
- Marschner, G. (1972). *Revisions-Test (Rev.T.) nach Dr. Berthold Stender. Ein allgemeiner Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit*. Göttingen: Hogrefe.
- Mayrhofer, S. (2004). *Der Revidierte Konzentrations-Leistungs-Test (KLT-R)*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 119–131). Göttingen: Hogrefe.
- Merz, F. (1961). *Über die individuelle Interferenzneigung*. *Zeitschrift für experimentelle und angewandte Psychologie*, 8, 381–392.
- Moosbrugger, H. & Goldhammer, F. (2007). *FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test II (grundlegend neu bearbeitete und neu normierte 2. Auflage des FAKT von H. Moosbrugger und M. Heyden)*. Bern: Huber.
- Moosbrugger, H. & Heyden, M. (1997). *Frankfurter Adaptiver Konzentrationsleistungs-Test (FAKT)*. Bern: Huber.
- Moosbrugger, H. & Heyden, M. (1997). *Frankfurter Adaptiver Konzentrations-Leistungs-Test (FAKT)*. Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (1996). *FAIR. Frankfurter Aufmerksamkeits-Inventar*. Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (2011). *FAIR-2. Frankfurter Aufmerksamkeits-Inventar 2*. Bern: Huber.
- Moosbrugger, H. & Öhlschlägel, J. (1996). *Frankfurter Aufmerksamkeits-Inventar (FAIR)*. Bern: Huber.
- Moosbrugger, H. & Reiß, S. (2004). *Das Frankfurter Aufmerksamkeits-Inventar FAIR*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 103–118). Göttingen: Hogrefe.
- Moosbrugger, H., Goldhammer, F. & Schweizer, K. (2006). *Latent factors underlying individual differences in attention measures: Perceptual and executive attention*. *European Journal of Psychological Assessment*, 22, 177–188.
- Moray, N. (1970). *Attention: Selective Processes in vision and hearing*. New York: Academic Press.
- Müggenburg, U. (1981). *Zur Beurteilung verschiedener Testverfahren und Einsatzmöglichkeiten des Vigilanzgerätes nach Quatember und Maly*. Unveröffentlichte Dissertation, Technische Hochschule Aachen, Medizinische Fakultät.
- Münsterberg, H. (1912). *Experimentalpsychologie und Berufswahl*. *Zeitschrift für experimentelle Pädagogik und pädagogische Psychologie*, 13, 1–7.
- Neisser, U. (1979). *Kognition und Wirklichkeit. Prinzipien der Implikation der kognitiven Psychologie*. Stuttgart: Klett-Cotta.

- Nell, V. (2003). Konzentrations-Leistungs-Test, revidierte Fassung (KLT-R). In E. Fay (Hrsg.), *Tests unter der Lupe 4. Aktuelle psychologische Testverfahren – kritisch betrachtet* (S. 59–75). Göttingen: Vandenhoeck & Ruprecht.
- Nell, V., Bretz, J. & Sniehotta, F. F. (2001). Bedingungsvariirtes Testen: Eine neue Konzeption zur Sicherung der ökologischen und kriterienbezogenen Validität in der Konzentrations- und Aufmerksamkeitsdiagnostik. In *Deutsche Psychologen Akademie* (Hrsg.), *Psychologie am Puls der Zeit* (S. 348–349). Bonn: Deutscher Psychologen Verlag.
- Nell, V., Bretz, J. & Sniehotta, F. F. (2004). *KT 3–4 R. Konzentrationstest für 3. und 4. Klassen. Revidierte Fassung*. Göttingen: Beltz.
- Neumann, O. (1992). Theorien der Aufmerksamkeit: Von Metaphern zu Mechanismen. *Psychologische Rundschau*, 43, 83–101.
- Neumann, O. (1996). Komponenten der Aufmerksamkeit und ihre Störungen. *Kindheit und Entwicklung*, 5 (2), 75–79.
- Neumann, O. (1996). Theorien der Aufmerksamkeit. In O. Neumann & A. F. Sanders (Hrsg.), *Aufmerksamkeit (Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung, Serie II, Kognition, Band 2, 559–643)*. Göttingen: Hogrefe.
- Neumann-Zilke, L. (2009). Vigilanz (VIGIL). Untertest aus dem Wiener Testsystem (WTS). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 99–109). Göttingen: Hogrefe.
- Neuwirth, W. (2001). Extremgruppenvalidierung verkehrspsychologischer Testverfahren anhand von Zuweisungsgruppen. *Psychologie in Österreich*, 21, 206–212.
- Noterdäme, M., Sitter, S., Mildnerberger, K. & Amorosa, H. (2000). Differenzierung der Aufmerksamkeitsleistung bei sprachentwicklungsgestörten Kindern anhand der Testbatterie zur Aufmerksamkeitsprüfung. In W. Maier, R. R. Engel & H.-J. Möller (Hrsg.), *Methodik von Verlaufs- und Therapiestudien in Psychiatrie und Psychotherapie* (S. 204–207). Göttingen: Hogrefe.
- Oehr, A. (1896). Experimentelle Studien zur Individualpsychologie. *Psychologische Arbeiten*, 1, 92–151.
- Oh, H. (1999). *Aufmerksamkeitsdefizite und jugendpsychiatrische Störungen*. Frankfurt am Main: Peter Lang.
- Oh, H., Moosbrugger, H. & Poustka, F. (2005). Kann eine spezifische Aufmerksamkeitsdiagnostik zur Differentialdiagnostik psychischer Störungen beitragen? *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 33 (3), 181–189.
- Öhlschlägel, J. & Moosbrugger, H. (1991). Konzentrationsleistung ohne Konzentration? Zur Schätzung wahrer Leistungswerte im Aufmerksamkeits-Belastungs-Test d2. *Diagnostica*, 37, 42–51.
- Öhlschlägel, J. & Moosbrugger, H. (1991). Überraschende Validitätsprobleme im Aufmerksamkeits-Belastungs-Test d2. *Report Psychologie*, 16 (9), 16–25.
- Orthmann, D. (1993). *Vergleichende Untersuchungen zu den biosozialen Entwicklungsbedingungen bei leistungsstarken, leistungsschwachen und lernbehinderten Schülern*. Frankfurt: Peter Lang.
- Oswald, W. D. & Roth, E. (1987). *Der Zahlenverbindungstest (ZVT). Ein sprachfreier Intelligenz-Test zur Messung der „kognitiven Leistungsgeschwindigkeit“* (2., überarbeitete und erweiterte Auflage). Göttingen: Hogrefe.
- Pauli, R. (1936). Beiträge zur Kenntnis der Arbeitskurve. *Archiv für die gesamte Psychologie*, 97, 465–532.

- Petermann, F. & Lepach, A. (2006). Neuropsychologische Diagnostik und Therapie von Aufmerksamkeits- und Gedächtnisstörungen im Kindesalter. *Verhaltenstherapie*, 16, 112–120.
- Petermann, F. (2009). Inventar Komplexer Aufmerksamkeit (INKA). In D. Schellig, R. Drechsler, D. Heinemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 217–223). Göttingen: Hogrefe.
- Petermann, F. (2011). Frankfurter Aufmerksamkeits-Inventar 2 (FAIR-2) [Testbesprechung]. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 59 (4), 325–326.
- Plössl, P. (1941). Die Arbeitskurve als diagnostisches Hilfsmittel bei Schwererziehbarkeit. *Zeitschrift für angewandte Psychologie*, 61, 2–69.
- Posner, M. I. & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78, 391–408.
- Posner, M. I. & Rafal, R. D. (1987). Cognitive theories of attention and the rehabilitation of attentional deficits. In M. J. Meier, A. L. Benton & L. Diller (Hrsg.), *Neuropsychological rehabilitation* (S. 182–201). Edinburgh: Churchill Livingstone.
- Posner, M. I. & Raichle, M. E. (1994). *Images of mind*. New York, NY: Scientific American Books.
- Pressley, M., Borkowski, J. G. & Schneider, W. (1989). Good information processing: What it is and what education can do to promote it. *International Journal of Educational Research*, 13, 857–867.
- Prieler, J. (2011). *Computerized Attention and Concentration Tests. Manuals*. Göttingen: Hogrefe.
- Puhr, U. & Wagner, M. (2001). Interferenztest nach Stroop. STROOP (Version 22.00. Computerprogramm mit Manual). Mödling: Schuhfried.
- Raven, J. C. (1962). *Advanced Progressive Matrices*. London: Lewis.
- Raven, J., Raven, J. C. & Court, J. H. (1997). *Handanweisung Standard Progressive Matrices Plus (SPM Plus)*. Mödling: Schuhfried.
- Reuning, H. (1959). Why not artifactor analysis? *Journal of the National Institute for Personnel Research*, 7, 192–200.
- Reuning, H. (1960). *Test administrator's manual for continuous letter checking*. Johannesburg: National Institute for Personnel Research.
- Rodenhausen, T. (2000). Das Cognitron. In E. Fay (Hrsg.), *Tests unter der Lupe III* (S. 45–60). Lengerich: Pabst.
- Roth, E. (1989). *Enzyklopädie der Psychologie* (S. 324 ff.; Kap. Verarbeitungskapazität, Kap. Flaschenhalstheorie, Kap. Selektive Aufmerksamkeit, Kap. Informationsverarbeitungskapazität). Göttingen: Hogrefe.
- Rothaar, P. (1993). *Untersuchung von Interferenz und Übungseffekten bei serieller Anwendung von Parallelformen des Frankfurter Aufmerksamkeits-Inventars (FAIR)*. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Royl, W. (1977). Aus der wissenschaftlichen Begleitung der Gesamtschulversuche in Schleswig-Holstein. Die Rolle der Konzentrationsfähigkeit bei der Bearbeitung von Lernerfolgstests im Fach Deutsch. In W. Arnold (Hrsg.), *Texte zur Schulpsychologie und Bildungsberatung. Band 2*. Braunschweig: Westermann.
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53, 61–97.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Testkurzdarstellung INKA Inventar Komplexer Aufmerksamkeit: S. 297–299)*. Lengerich: Pabst.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Test-*

- kurzdarstellung KLT-R Revidierter Konzentrations-Leistungs-Test: S. 331–334). Lengerich: Pabst.
- Sarges, W. & Wottawa, H. (Hrsg.). (2001). *Handbuch wirtschaftspsychologischer Testverfahren (Testkurzdarstellung Signal Detection)*: S. 495–497). Lengerich: Pabst.
- Schäfer, N. & Moosbrugger, H. (1993). Einige Validitätsaspekte des Frankfurter Aufmerksamkeits-Inventars (FAIR) (Arbeiten aus dem Institut für Psychologie der Johann Wolfgang Goethe-Universität, Heft 10/1993). Frankfurt am Main: Johann Wolfgang Goethe-Universität Frankfurt am Main, Institut für Psychologie.
- Scheiblechner, H. (1979). *Specifically objective stochastic latency mechanisms*. *Journal of Mathematical Psychology*, 19, 18–38.
- Scheiblechner, H. (1985). *Psychometric models for speed-test construction: The linear exponential model*. In S. E. Embretson (Hrsg.), *Test design: Developments in psychology and psychometrics* (S. 219–244). New York, NY: Academic Press.
- Schmidt-Atzert, L. & Bühner, M. (2000). *Aufmerksamkeit und Intelligenz*. In K. Schweitzer (Hrsg.), *Intelligenz und Kognition* (S. 125–152). Landau: Verlag Empirische Pädagogik.
- Schmidt-Atzert, L., Bühner, M. & Enders, P. (2006). *Messen Konzentrationstests Konzentration? Eine Analyse von Konzentrationstestleistungen*. *Diagnostica*, 52 (1), 33–44.
- Schmidt-Atzert, L., Büttner, G. & Bühner, M. (2004). *Theoretische Aspekte von Aufmerksamkeits-/Konzentrationsdiagnostik*. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 3–22). Göttingen: Hogrefe.
- Schmidt-Atzert, L. & Ising, M. (1997). *Ein Beitrag zur Konstruktvalidität von d2 und Revisionstest*. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18 (4), 241–250.
- Schoppe, K. J. (1974). *Das MLS-Gerät: Ein neuer Testapparat zur Messung feinmotorischer Leistungen*. *Diagnostica*, 20, 43–47.
- Schuhfried, G. (1993). *Signal Detection (Version 4.00)*. Mödling: Schuhfried.
- Schuhfried, G. (1995). *Vigilanz (Version, 5.00)*. Mödling: Schuhfried.
- Schuhfried, G. (1998). *Wiener Determinationstest*. Mödling: Schuhfried.
- Schuhfried, G. (2001). *Cognitrone. COG (Version 29.00) (Computerprogramm)*. Mödling: Schuhfried.
- Schuhfried, G. (2004). *COG. Cognitrone*. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren (2., überarbeitete und erweiterte Auflage)*, S. 251–255). Lengerich: Pabst.
- Schuhfried, G. (2004). *Signal Detection*. In W. Sarges & H. Wottawa (Hrsg.), *Handbuch wirtschaftspsychologischer Testverfahren (2., überarbeitete und erweiterte Auflage)*, S. 723–726). Lengerich: Pabst.
- Schuhfried, G. (2004). *Vigilanz (Version 25.00)*. Mödling: Schuhfried.
- Schuhfried, G. (2007). *Cognitrone (Version 37.01)*. Mödling: Schuhfried.
- Schuhfried, G. (2007). *Signal-Detection (Version 26.05)*. Mödling: Schuhfried.
- Schuhfried, G. (2012). *Cognitrone (Version 43). Manual*. Verfasst von M. Wagner & T. Karner Mödling: Schuhfried. Abrufbar unter www.schuhfried.at/fileadmin/content/1_Manuale_de/COG.pdf [17.07.2013].
- Schuhfried, G. (2013). *Wiener Testsystem. Computergestützte Verfahren zur Leistungs- und Persönlichkeitsdiagnostik (Katalog)*. Mödling: Schuhfried.
- Schuhfried, G. & Prieler, J. (1997). *Wiener Reaktionstest*. Mödling: Schuhfried.
- Schwalbach, C. (2001). *Laienkonzepte über Aufmerksamkeit und Konzentrationsfähigkeit*. Marburg:

Philipps-Universität Marburg, Fachbereich Psychologie.

- Schwenkmezger, P. (1986). Farbe-Wort-Interferenztest (FWIT) nach J. R. Stroop. und G. Bäumler. *Diagnostica*, 32 (2), 171–173.
- Seitz, W. (1971). Über den Zusammenhang von Leistungen im Pauli-Test und diversen Persönlichkeitsmerkmalen. *Psychologische Beiträge*, 13, 221–238.
- Sharp, S. E. (1899). Individual psychology. *American Journal of Psychology*, 10, 329–391.
- Siebert, U. (2002). Frankfurter Aufmerksamkeits-Inventar (FAIR, Moosbrugger & Öhlschlägel, 1996) als differentialdiagnostisches Instrument bei Kindern mit Aufmerksamkeitsdefizitsyndrom mit oder ohne Hyperaktivität (ADD/ADHD) im Alter von 8–12 Jahren. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität Frankfurt am Main.
- Sprengelmeyer, R., Zimmermann, P., Lange, H. & Hömberg, V. (1994). Attention disorders in extrapyramidal disease: A preliminary report. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 31–40). Tübingen: Gunter Narr.
- Stapf, A. (2010). Differenzialdiagnostik: Hochbegabung und Aufmerksamkeitsstörung (ADHS). In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung. Reihe: Jahrbuch der pädagogisch-psychologischen Diagnostik* (S. 293–318). Tests und Trends. Göttingen: Hogrefe.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Sturm, W. & Büssing, A. (1990). Normierungs- und Reliabilitätsuntersuchungen zum Vigilanzgerät nach Quatember und Maly. *Diagnostica*, 36 (1), 50–59.
- Sturm, W. & Willmes, K. (1994). A normative study on the European Attention Test Battery. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 17–20). Tübingen: Gunter Narr.
- Sturm, W. & Zimmermann, P. (2000). Aufmerksamkeitsstörungen. In W. Zimmermann, M. Herrmann & C. Wallesch (Hrsg.), *Lehrbuch der klinischen Neuropsychologie* (S. 345–365). Lisse, NL: Swets & Zeitlinger.
- Sturm, W. (2006). WAF. Wahrnehmungs- und Aufmerksamkeitsfunktionen. Mödling: Schuhfried.
- Sturm, W., Hartje, W., Orgass, B. & Willmes, K. (1994). Effektivität eines computergestützten Trainings von vier Aufmerksamkeitsfunktionen. *Zeitschrift für Neuropsychologie*, 5 (1), 15–28.
- Thimm, M. (2009). Frankfurter Aufmerksamkeits-Inventar (FAIR). In D. Schellig, R. Drechsler, D. Heineemann & W. Sturm (Hrsg.), *Handbuch neuropsychologischer Testverfahren. Band 1: Aufmerksamkeit, Gedächtnis, exekutive Funktionen* (S. 181–186). Göttingen: Hogrefe.
- Thomae, H. (1951). Experimentelle psychologische Diagnostik. *Fortschritte der Neurologie, Psychiatrie und ihrer Grenzgebiete*, 19, 1–22.
- Toulouse, E. & Pieron, H. (1911). *Technique de psychologie experimentale I*. Paris.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Turß, M. (2002). Frankfurter Aufmerksamkeits-Inventar (FAIR). In U. P. Kanning & H. Holling (Hrsg.), *Handbuch personaldiagnostischer Instrumente* (S. 149–154). Göttingen: Hogrefe.
- Ulich, E. (1958). Neue Erfahrungen mit dem Pauli-Test. *Zeitschrift für experimentelle und angewandte Psychologie*, 5, 108–126.
- Van Zomeran, A. H. & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. Oxford, UK: Ox-

- ford University Press.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13, 37–58.
- Wagner, M. M. (2003). H. Moosbrugger & J. Öhlschlägel (1996). Frankfurter Aufmerksamkeits-Inventar (FAIR) (Testrezension). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24 (1), 81–83.
- Wagner, M. & Karner, T. (2001). *Cognitrone*. Mödling: Schuhfried.
- Wagner-Menghin, M. M. (2003). Konzentrationstests. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der psychologischen Diagnostik* (S. 249-253). Weinheim: Beltz.
- Weis, S. & Nuerk, H.-C. (2011). TBS-TK Rezension: „FAKT-II. Frankfurter Adaptiver Konzentrationsleistungs-Test“. *Psychologische Rundschau*, 62, 139–141.
- Weise, G. (1975). *Psychologische Leistungstests. Ein Handbuch für Studium und Praxis. Band 1: Intelligenz – Konzentration – Spezielle Fähigkeiten (Pauli-Test)*. S. 121–130. Göttingen: Hogrefe.
- Weiss, G. (1950). *Der charakterologische Arbeitsversuch (Pauli-Test) in seiner Anwendung bei Blinden*. Unveröffentlichte Dissertation, Universität München.
- Westhoff, K. (1995). Aufmerksamkeit und Konzentration. In M. Amelang (Hrsg.), *Enzyklopädie der Psychologie, Serie VIII: Differentielle Psychologie und Persönlichkeitsforschung, Band 2: Verhaltens- und Leistungsunterschiede* (S. 375–402). Göttingen: Hogrefe.
- Whipple, G. M. (1910). *Manual of mental and physical tests*. Baltimore: Warwick and York.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York: Academic Press.
- Wieczerkowski, W., Nickel, H., Janowski, A., Fittkau, B. & Rauer, W. (1975). Angstfragebogen für Schüler. AFS. Westermann Test (2. veränderte Auflage). Göttingen: Hogrefe.
- Wiese, W. & Kroj, G. (1972). Untersuchung über den Zusammenhang zwischen Intelligenz (Wechsler) und Konzentrationsfähigkeit (Test d2 nach Brickenkamp). *Zeitschrift für experimentelle und angewandte Psychologie*, 19, 690–699.
- Wothe, K. & Rausch, K. (1988). *Testhandbuch der Eignungs- und Verwendungsprüfung Wehrpflichtiger (TestHB-EVP). Testaufbau und Testanalyse*. In Bundesminister für Verteidigung (Hrsg.), *Fachliche Arbeitsanweisung für den Psychologischen Dienst der Bundeswehr im Wehersatzwesen*. Bonn.
- Wulsten, A.-R. & Ulich, E. (1971). Zur Frage der Wiederholungszuverlässigkeit im Pauli-Test. *Psychologische Rundschau*, 22, 203–212.
- Zimmermann, P. & Fimm, B. (1993). *Testbatterie zur Aufmerksamkeitsprüfung (TAP)*. Freiburg: Psytest.
- Zimmermann, P. & Fimm, B. (2001). *Ergänzungsmanual zur Testbatterie zur Aufmerksamkeitsprüfung (TAP) Version 1.6 – Kinder- und Jugendlichennormen*. Herzogenrath: Psytest.
- Zimmermann, P. & Fimm, B. (2004). Die Testbatterie zur Aufmerksamkeitsprüfung (TAP). In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 177–202). Göttingen: Hogrefe.
- Zimmermann, P., North, P. & Fimm, B. (1994). Diagnosis of attentional deficits: theoretical considerations and presentations of a test battery. In F. J. Stachowiak (Hrsg.), *Developments in the assessment and rehabilitation of brain-damaged patients. Perspectives from a European concerted action* (S. 3-15). Tübingen: Gunter Narr.

Anhang 1 – Musterbefunde und -gutachten



Dieser Abschnitt beinhaltet Mustergutachten bzw. -befunde, die eine Hilfe für die Erstellung von Gutachten im Rahmen der Diagnostik von Aufmerksamkeits- und Konzentrationsfähigkeit bieten sollen.

Auf den nachfolgenden Seiten sind folgende Muster zu finden:

- Untersuchungsbericht zu d2, FAIR
- Klinisch psychologisches Gutachten 1 zu VIGIL, WAF
- Klinisch psychologisches Gutachten 2 zu VIGIL, WAF
- Psychologischer Befund zu DL-KG, WAF

Untersuchungsbericht

d2, FAIR

Die Leistungsmenge und das Arbeitstempo des Probanden, der einerseits mit dem Durchstreichtest von Brinkenkamp (d2), andererseits mit dem Frankfurter Aufmerksamkeitsinventar (FAIR) getestet wurde, waren in beiden Untersuchungsverfahren als überdurchschnittlich einzustufen:

Untersuchungsbericht zum d2:

Herr ***** erfasste die Aufgabenstellung schnell und problemlos.

Auffallend ist, dass in der vierten Testzeile nur insgesamt 33 Zeichen bearbeitet wurden, wohingegen in den restlichen Zeilen jeweils deutlich mehr bearbeitete Zeichen erkennbar sind. Eine Schwankungsbreite von 14 Zeichen ist aber durchaus noch als normal zu werten.

Die Fehlerverteilung ist unauffällig.

Es fällt dennoch auf, dass in den letzten Zeilen keine Fehler mehr gemacht wurden, was möglicherweise auf einen Übungseffekt hindeutet.

Herr ***** erreichte eine Gesamtzahl von 594 bearbeiteten Zeichen, was einem Standardwert von 126 und einem Prozentrang von 99,5 entspricht.

Das bedeutet, dass ungefähr 99 % der Vergleichsgruppe im Test schlechtere Ergebnisse erzielt hätten als er.

Auch der Wert GZ-F, der Gesamttestwert mit einem Prozentrang von 99,9 sagt dasselbe aus.

Herr ***** bearbeitete den Test besser als 99 % seiner Altersgruppe.

Der Fehlerprozentwert von 1,852 ist als unterdurchschnittlich zu werten.

[..]

Die Ergebnisse sprechen für eine überdurchschnittliche Aufmerksamkeit und Leistungsfähigkeit.

Herr ***** zeigte im d2-Test in jedem Bereich im Vergleich zu seiner Altersgruppe überdurchschnittliche Leistungen. [..]

Untersuchungsbericht zum FAIR:

Herr ***** erfasste die Aufgabenstellung schnell und konnte den Test problemlos lösen. Im ersten Teil machte er insgesamt 2 Verspaserfehler, im zweiten Teil keinen mehr, was

möglicherweise auf einen Übungseffekt hindeutet.

Er erreichte bei den konzentriert bearbeiteten Zeichen in Hinblick auf den Leistungswert (L) einen Prozentrang von 97, was einem Stanine-Wert von 9 entspricht und als weit überdurchschnittlich zu werten ist.

Der Anteil konzentriert bearbeiteter Symbole (Q) ergab einen Prozentrang von 94 (Stanine-Wert: 8), was ebenfalls im überdurchschnittlichen Bereich liegt.

Der Kontinuitätswert, der die kontinuierliche Konzentrationsleistung erfasst, liegt bei einem Prozentrang von 98 und damit bei einem Stanine-Wert von 9, was bedeutet, dass 98 % der Gleichaltrigen schlechter oder gleichgut abgeschnitten hätten.

Auch dieser Wert ist weit überdurchschnittlich.

Herr ***** erreichte eine Gesamtzahl von 526 bearbeiteten Zeichen.

Die Ergebnisse sprechen für eine überdurchschnittliche Aufmerksamkeit und Leistungsfähigkeit.

Herr ***** zeigte im FAIR-Test in jedem Bereich im Vergleich zu seiner Altersgruppe überdurchschnittliche Leistungen.

Klinisch-psychologisches Gutachten 1

VIGIL, WAF

An

Salzburg, am *****

Klinisch-psychologisches Gutachten

Aus den erhobenen Daten und den Ergebnissen der psychologischen Untersuchung wird nachstehend Befund und Gutachten über Herrn ***** (geb. am *****) erstellt.

Grundlage der Begutachtung war die Verwertung der Untersuchungsdaten.

Die Fragestellung umfasst folgende Punkte:

- Umfassende Abklärung (Konzentration, Aufmerksamkeit, Ablenkbarkeit, Reaktionsfähigkeit, Auffassungsgabe usw.)

Das psychologische Gutachten basiert auf:

- Psychologische Untersuchung der Testperson mit dem VIGIL-Verfahren aus dem WTS, dem WAFG-Test aus dem WTS, dem WAFV-Test aus dem WTS, dem WAFS-Test aus dem WTS sowie kurzen Testaufgaben zum Wiedererkennen (SKT), zur räumlichen Orientierung, zur Interferenz und zum Gesichter-Gedächtnis und zur Auffassung.

Psychologische Testverfahren

VIGIL

VIGIL ist ein Test zur Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma „low event rate“. Erfasst wird die Aufrechterhaltung der Aufmerksamkeit in einer monotonen Reizsituation.¹

WAFG

Das WAFG-Verfahren testet die Fähigkeit zur geteilten Aufmerksamkeit. Der Proband erhält Reizmaterial auf zwei visuellen bzw. einem visuellen und einem auditiven Kanal. Er soll dabei stets beide Kanäle überwachen und prüfen, ob sich einer der Reize zweimal hintereinander verändert hat.

WAFS

Das WAFS-Verfahren testet die Fähigkeit zur selektiven Aufmerksamkeit. Der Proband erhält relevante und irrelevante Reize in einer oder beiden Vorgabemodalitäten und soll auf Veränderung der relevanten Reize reagieren und irrelevante Reize ignorieren.

¹ Vigilanzanforderungen sind durch folgende Bedingungen bestimmt: Bei längerer Versuchsdauer wird von der Probandin/vom Probanden ununterbrochene Wachsamkeit verlangt. Die zu beachtenden Signale erscheinen unregelmäßig und sie erwecken keine unwillkürliche Aufmerksamkeit. Daraus ergibt sich die Forderung nach relativ geringer Intensität der Reizdarbietung und niedriger Frequenz der kritischen Ereignisse. Der Leistungsabfall bei Vigilanzexperimenten wird mit der Abnahme des Aktivierungsniveaus der Probandin/des Probanden und der damit einhergehenden Zunahme der Reaktionslatenz erklärt. Nach der neurophysiologischen Aktivierungstheorie wird der Kortex infolge Reizarmut vom aufsteigenden retikulären Aktivierungssystem (ARAS) nur unzureichend stimuliert. Die Gehirnrinde erhält für die Aufrechterhaltung bestimmter Tätigkeiten daher nicht die entsprechenden Weckimpulse, was zur psychischen Ermüdung und damit zur Abnahme der Leistungseffizienz führt. In diesem Zusammenhang entstand der Begriff der Überforderung durch Unterforderung.

WAFV

Das WAFV-Verfahren testet die Daueraufmerksamkeit des Probanden. Dem Probanden werden visuelle bzw. auditive Reize vorgegeben, die in einigen Fällen an Intensität verlieren (etwa 25 % beim Test zur Daueraufmerksamkeit im Vergleich zu etwa 5 % beim Test zur Vigilanz), auf die er möglichst rasch reagieren soll.

Aufgaben zur räumlichen Orientierung, zum Gesichtergedächtnis, zur Planungsfähigkeit und zur Realitätssicherheit:

1. *Räumliches Gedächtnis:* Die Person soll sich die Anordnung von Bildern merken und auf der folgenden Seite den Kästchen im Raster zuordnen.
2. *Wiedererkennen von Gesichtern:* Auf einem Gruppenfoto soll die Person ein bzw. zwei Gesichter wiedererkennen, die unmittelbar zuvor dargeboten wurden.
3. *Wiedererkennen:* Die Person soll sich 12 Bilder ansehen und merken und dann aus einer größeren Menge an Bildern herausfinden.
4. *Interferenz:* Die Person soll Interferenzaufgaben lösen (A statt B, Farbe statt Wort...).

Zur Situation:

Herr *****

Befund und Gutachten

Beim VIGIL-Verfahren zur Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“ erreicht er einen Prozentrang von 47 (T = 49) bei den Richtigen und einen Prozentrang von 8 (T = 36) bei den Falschen. Mit der Reaktionszeit ist er mit 0,46 Sekunden (T = 53) im Durchschnittsbereich der Norm. In der Schnelligkeits-/Genauigkeitsanalyse liegt er im Feld schnell und durchschnittlich genau.

Beim WAFG-Verfahren zur geteilten Aufmerksamkeit erreicht er schon unimodal unterdurchschnittliche Reaktionszeiten (PR = 11, T = 38). Es wurden 8 Verpasser registriert, was einem Prozentrang von 26 (T = 44) entspricht und 8 falsche Alarme, was einem Prozentrang von 7 entspricht. D.h., Herr ***** ist hier schlechter als 74 % bzw. 93 % der gleichaltrigen Personen. Wird crossmodal getestet, erreicht er mit T = 41 eine (noch) durchschnittliche Zeit (PR = 19), die Anzahl der Verpasser und falschen Alarme ist mit 14 bzw. 6 überdurchschnittlich hoch. Hier ergeben sich unterdurchschnittliche T-Werte

von 38 bzw. 39, wobei er hier deutlich schlechter bei den auditiven Reizen abschneidet. Bei der unimodalen Vorgabe wird er im Laufe der Zeit langsamer in der Reaktion, bei der crossmodalen Vorgabe bleibt die Reaktionszeit in etwa gleich.

Auch beim Test zur selektiven Aufmerksamkeit ergeben sich unterdurchschnittliche Reaktionszeiten, es ergibt sich ein Prozentrang von 12 ($T = 38$) bei der unimodal visuellen Vorgabe, ein Prozentrang von 15 ($T = 40$) bei der unimodalen auditiven Vorgabe und ebenfalls ein Prozentrang von 15 bei der crossmodalen ($T = 40$) Vorgabe. Die mittlere Reaktionszeit bei der automatischen Verarbeitung ist mit $T < 38$ (abhängig von der Vorgabe) ähnlich zur mittleren Reaktionszeit bei der kontrollierten Verarbeitung ($T < 39$) (abhängig von der Vorgabe).

Beim WAFV zur Daueraufmerksamkeit erreicht er durchwegs Werte im Normbereich und ist unauffällig. Die Reaktionszeit bleibt in etwa gleich im Zeitverlauf.

Bei den Aufgaben zum räumlichen Gedächtnis erreicht Herr ***** einen (noch) durchschnittlichen Wert im Vergleich zu seiner Altersgruppe ($P = 21$), hier sind 21 % der Gleichaltrigen schlechter als er, 79 % besser.

Das Wiedererkennen von Gesichtern fällt ihm bei nur einem Reiz nicht schwer. Soll er sich mehrere Personen zugleich merken, erreicht er (noch) durchschnittliche Werte. Die Wahrscheinlichkeit des Wiedererkennens liegt hier bei ca. 80 %. Bei mehreren Reizen sinkt sie auf unter 50 % ab. In etwa dieselbe Wahrscheinlichkeit des Wiedererkennens ergibt sich auch bei den Bildern des SKT.

Die Planungsfähigkeit ist bei Herrn ***** im unteren Durchschnittsbereich ($T = 42$). Bei der Realitätssicherheit ergeben sich unauffällige Scores ($T = 54$).

Interferenzfähigkeit ist grundsätzlich gegeben. Die Auffassung ist bei Herrn ***** knapp unterdurchschnittlich ($T = 39$).

Zusammenfassung

Es kann gesagt werden, dass die Realitätssicherheit bei Herrn ***** durchschnittlich ausgeprägt ist. Beim räumlichen Gedächtnis, bei den Aufgaben zur Auffassung, bei Interferenzaufgaben und bei der Planungsfähigkeit liegt er im unteren Durchschnittsbereich. Der Vigilanz-Test hinsichtlich der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“ (Vigilanz, Reizdichte 5 %) ergibt bei Herrn ***** einen durchschnittlichen Wert bei den Richtigen und einen unterdurchschnittlichen Wert

bei den Falschen. Die Reaktionszeit ist mit 0,46 Sekunden durchschnittlich. In der Schnelligkeits-/Genauigkeitsanalyse liegt er im Feld schnell und durchschnittlich genau. Beim Test zur Daueraufmerksamkeit (Reizdichte 25 %) erreicht er durchwegs Werte im Normbereich. Die Reaktionszeit bleibt in etwa gleich im Zeitverlauf. Probleme hat er bei der geteilten Aufmerksamkeit. Hier erreicht er schon unimodal unterdurchschnittliche Reaktionszeiten (PR = 11, T = 38) bei überdurchschnittlich vielen Verpassern und falschen Alarmen. Wird crossmodal getestet, erreicht er eine (noch) durchschnittliche Zeit, die Anzahl der Verpasser und falschen Alarme ist auch hier überdurchschnittlich hoch. Bei der unimodalen Vorgabe wird er im Laufe der Zeit langsamer in der Reaktion, bei der crossmodalen Vorgabe bleibt die Reaktionszeit in etwa gleich. Auch beim Test zur selektiven Aufmerksamkeit ergeben sich sowohl unimodal als auch crossmodal unterdurchschnittliche Reaktionszeiten. Die mittlere Reaktionszeit bei der automatischen Verarbeitung ist ähnlich zur mittleren Reaktionszeit bei der kontrollierten Verarbeitung. Untersucht man die Speicherprozesse, so ergibt sich, dass die Sicherheit des Wiedererkennens mit 75 % durchschnittlich ausgeprägt ist. Untersucht man die Gedächtnisleistung explizit, so ergibt sich, dass ihm das Wiederkennen von Gesichtern bei nur einem Reiz nicht schwer fällt, er liegt hier im Durchschnittsbereich. Soll er sich mehrere Personen zugleich merken, erreicht Herr ***** nur noch knapp durchschnittliche Werte. Die Wahrscheinlichkeit des Wiedererkennens liegt bei einem Reiz bei knapp 80 %, bei mehreren Reizen sinkt sie auf unter 50 % ab.

Herr ***** hat sichtbar Probleme mit der Konzentration bei Mehrfachreizen, crossmodalen Aufgabenstellungen und komplexen Fragestellungen. Die Reaktionszeiten sind je nach Vorgabe im (Vigilanz, Daueraufmerksamkeit) oder unter (geteilte Aufmerksamkeit, selektive Aufmerksamkeit) dem Durchschnittsbereich der Normgruppe, was für eine erhöhte Ablenkbarkeit spricht. Herr ***** ist in Stresssituationen oder unter Mehrfachbelastung nicht in der Lage, Aufgaben, die Konzentration erfordern, schnell und genau zu lösen: Die Reaktionszeit und auch die Anzahl der Verpasser/falschen Alarme sind überdurchschnittlich hoch. Wahrnehmung, Interferenzfähigkeit und Auffassungsgabe sind bei Herrn ***** im unteren Durchschnittsbereich.

Klinisch-psychologisches Gutachten 2

VIGIL, WAF

An

Salzburg, am *****

Klinisch-psychologisches Gutachten

Aus den erhobenen Daten und den Ergebnissen der psychologischen Untersuchung wird nachstehend Befund und Gutachten über Frau ***** (geb. am *****) erstellt.

Grundlage der Begutachtung war die Verwertung der Untersuchungsdaten.

Die Fragestellung umfasst folgende Punkte:

- Aufmerksamkeitsstörung
- Interferenz, Ablenkbarkeit

Das psychologische Gutachten basiert auf:

- AD(H)S-*****
- Psychologische Untersuchung der Person mit *****, dem WAFS-Test aus dem WTS, dem WAFR-Test aus dem WTS, dem WAFF-Test aus dem WTS, dem WAFV-Test aus dem WTS und dem VIGIL aus dem WTS.

Psychologische Testverfahren

VIGIL

VIGIL ist ein Test zur Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma „low event rate“. Erfasst wird die Aufrechterhaltung der Aufmerksamkeit in einer monotonen Reizsituation.¹

WAFG

Das WAFG-Verfahren testet die Fähigkeit zur geteilten Aufmerksamkeit. Die Probandin erhält Reizmaterial auf zwei visuellen bzw. einem visuellen und einem auditiven Kanal. Sie soll dabei stets beide Kanäle überwachen und prüfen, ob sich einer der Reize zweimal hintereinander verändert hat.

WAFS

Das WAFS-Verfahren testet die Fähigkeit zur selektiven Aufmerksamkeit ab. Die Probandin erhält relevante und irrelevante Reize in einer oder beiden Vorgabemodalitäten und soll auf Veränderung der relevanten Reize reagieren und irrelevante Reize ignorieren.

¹ Vigilanzanforderungen sind durch folgende Bedingungen bestimmt: Bei längerer Versuchsdauer wird von der Probandin/vom Probanden ununterbrochene Wachsamkeit verlangt. Die zu beachtenden Signale erscheinen unregelmäßig und sie erwecken keine unwillkürliche Aufmerksamkeit. Daraus ergibt sich die Forderung nach relativ geringer Intensität der Reizdarbietung und niedriger Frequenz der kritischen Ereignisse. Der Leistungsabfall bei Vigilanzexperimenten wird mit der Abnahme des Aktivierungsniveaus der Probandin/des Probanden und der damit einhergehenden Zunahme der Reaktionslatenz erklärt. Nach der neurophysiologischen Aktivierungstheorie wird der Kortex infolge Reizarmut vom aufsteigenden retikulären Aktivierungssystem (ARAS) nur unzureichend stimuliert. Die Gehirnrinde erhält für die Aufrechterhaltung bestimmter Tätigkeiten daher nicht die entsprechenden Weckimpulse, was zur psychischen Ermüdung und damit zur Abnahme der Leistungseffizienz führt. In diesem Zusammenhang entstand der Begriff der Überforderung durch Unterforderung.

WAFV

Das WAFV-Verfahren testet die Daueraufmerksamkeit einer Testperson. Der Probandin werden visuelle bzw. auditive Reize vorgegeben, die in einigen Fällen an Intensität verlieren (etwa 25 % beim Test zur Daueraufmerksamkeit im Vergleich zu etwa 5 % beim Test zur Vigilanz), auf die sie möglichst rasch reagieren soll.

WAFR

Das WAFR-Verfahren testet die räumliche Ausrichtung der Aufmerksamkeit an vier bzw. acht räumlichen Positionen in einer Posner-Paradigma-ähnlichen Aufgabenstellung. Es werden periphere (exogene) und zentrale (endogene) räumliche Hinweisreize verwendet. Bei der Neglect-Prüfung werden Stimuli an verschiedenen Positionen im rechten oder linken visuellen Feld oder simultan an äquivalenten Positionen in beiden Gesichtsfeldhälften dargeboten (Extinktionsbedingung).

WAFF

Das WAFF-Verfahren testet die fokussierte Aufmerksamkeit. Der Probandin werden – je nach Subtest – visuelle oder auditive relevante Reize vor ablenkendem Reizmaterial vorgegeben. Wenn zwei zuvor definierte Veränderungen der relevanten Reize hintereinander auftreten, soll die Probandin reagieren, das übrige Reizmaterial soll sie ignorieren.

HASE

Das Verfahren dient der Diagnostik der AD(H)S im Erwachsenenalter und ermöglicht diagnostische Entscheidungen auf verschiedenen Ebenen. Das Instrument leistet eine AD(H)S-Kerndiagnostik, die mit der Beschreibung kindlicher AD(H)S-Phänomene beginnt und über den Nachweis der spezifischen Diagnosekriterien von DSM-IV und der ICD-10-Forschungsversion zu speziell elaborierten AD(H)S-Phänomenen des Erwachsenenalters führt.

Zur Situation:

Befund und Gutachten

Frau *****

Beim VIGIL-Verfahren zur Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“ erreicht die Testperson einen Prozent-

rang von 47 (T = 49) bei den Richtigen und einen Prozentrang von 22 (T=42) bei den Falschen. Mit der Reaktionszeit ist sie mit 0,52 Sekunden (T = 47) im Durchschnittsbereich der Norm. In der Schnelligkeits-/Genauigkeitsanalyse liegt sie im Feld eher langsam und durchschnittlich genau.

Beim Test zur selektiven Aufmerksamkeit (WAFS) ergeben sich durchschnittliche Reaktionszeiten, es ergibt sich ein Prozentrang von 59 (T = 52) bei der unimodal visuellen Vorgabe, ein Prozentrang von 70 (T = 55) bei der unimodalen auditiven Vorgabe und auch von 72 bei der crossmodalen (T = 56) Vorgabe.

Beim WAFV zur Daueraufmerksamkeit erreicht sie durchwegs Werte im Normbereich und ist unauffällig, in der zweiten Testhälfte hat sie keinen einzigen Verpasser und liegt damit über dem Durchschnittsbereich der Norm. Die Reaktionszeit bleibt in etwa gleich im Zeitverlauf.

Beim WAFR-Verfahren zur Erfassung der räumlichen Aufmerksamkeit sind die Reaktionszeiten im Durchschnittsbereich $46 < T < 52$. Sie hat weder bei den unilateralen Reizen rechts noch bei den Reizen links noch bei den bilateralen Reizen einen Verpasser und erreicht Prozentränge im oberen Normbereich $58 < PR < 62$.

Beim Verfahren zur fokussierten Aufmerksamkeit sind die Reaktionszeiten mit T = 60 bzw. T = 58 im oberen Durchschnittsbereich (unimodal auditiv und unimodal visuell) bzw. mit T = 63 über dem Durchschnittsbereich (crossmodal). Visuell können keine Verpasser und 2 falsche Alarme festgestellt werden, auditiv und crossmodal gibt es ebenfalls keine Verpasser und zusätzlich auch keinen falschen Alarm, hier erreicht sie einen Prozentrang von 100 (T = 80) und ist weit über dem Durchschnittsbereich der Norm.

Bei speziellen Aufgaben zu Interferenz und Ablenkbarkeit ist Frau ***** ebenfalls unauffällig.

Zusammenfassung:

Frau ***** ist in allen Bereichen – Vigilanz, Daueraufmerksamkeit, Selektive Aufmerksamkeit, Räumliche Aufmerksamkeit und fokussierte Aufmerksamkeit – unauffällig und liegt im Durchschnittsbereich der Norm oder knapp darüber. Probleme mit Interferenzaufgaben können nicht festgestellt werden.

Eine krankheitswertige Aufmerksamkeitsstörung im Sinne des ICD-10/DSM IV konnte bei Frau ***** nicht festgestellt werden.

Psychologischer Befund

DL-KG, WAF

PSYCHOLOGISCHER BEFUND

für ***** (geb. *****)

***** hat sich einer klinisch-psychologischen Begutachtung zur Abklärung einer möglichen **Aufmerksamkeitsstörung** unterzogen. Die klinisch-psychologische Untersuchung umfasste ein exploratives Gespräch, Arbeitsproben, Daten aus der Verhaltensbeobachtung und die Durchführung von vier psychologischen Testverfahren (DL-KG, Computerbasierter Test zur Erfassung der selektiven Aufmerksamkeit, der räumlichen Aufmerksamkeit und der Vigilanz).

Verwendete Verfahren

Testbatterie zur Erfassung von Aufmerksamkeitsleistungen

Die Testbatterie besteht aus 6 Testverfahren, die unabhängig voneinander oder in beliebiger Kombination als Testbatterie vorgegeben werden können.

Vorgegeben wurden:

- Vigilanz
- Räumliche Aufmerksamkeit
- Selektive Aufmerksamkeit

Die Tests umfassen konsequent Subtests für visuelle, auditive und crossmodale Testvorgaben. In einigen Untertests der Testbatterie werden automatisierte und kontrollierte Aufmerksamkeitsaspekte separat erfasst, indem die Reize entweder durch Intensitätsanhebung stärker hervortreten („Popping Out“) oder durch Intensitätsabschwächung schwächer hervortreten und damit kognitiv kontrollierte, „top down“-Prozesse notwendig werden.

Vigilanz

Der Vigilanztest dient der Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma „low event rate“. Erfasst wird die Aufrechterhaltung der

Aufmerksamkeit in *einer* monotonen Reizsituation.¹

Selektive Aufmerksamkeit

Der Proband erhält beim Test zur Erfassung der selektiven Aufmerksamkeit relevante und irrelevante Reize in einer oder beiden Vorgabemodalitäten und soll auf Veränderung der relevanten Reize reagieren und irrelevante Reize ignorieren.

Räumliche Aufmerksamkeit

Das Testverfahren zur räumlichen Aufmerksamkeit testet die räumliche Ausrichtung der Aufmerksamkeit an vier bzw. acht räumlichen Positionen in einer Posner-Paradigma-ähnlichen Aufgabenstellung. Es werden periphere (exogene) und zentrale (endogene) räumliche Hinweisreize verwendet. Bei der Neglect-Prüfung werden Stimuli an verschiedenen Positionen im rechten oder linken visuellen Feld oder simultan an äquivalenten Positionen in beiden Gesichtsfeldhälften dargeboten (Extinktionsbedingung).

DL-KG

Der Differentielle Leistungstest DL-KG ist ein Konzentrationstest und liefert drei Messwerte für verschiedene Leistungsvariablen (Quantität, Qualität und Gleichmäßigkeit der Leistung). Darüber hinaus ist er als standardisierte Beobachtungssituation der Arbeitshaltung bei konzentrierter Tätigkeit und fremdgesetzter Aufgabenstellung geeignet.

Zusammenfassung der Ergebnisse

Beim Verfahren zur Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“ (Vigilanz) erreicht ***** einen Prozentrang von 7 (T = 35) bei den Richtigen und einen Prozentrang von 4 (T = 32) bei den Falschen. Mit

¹ Vigilanzanforderungen sind durch folgende Bedingungen bestimmt: Bei längerer Versuchsdauer wird von der Probandin/vom Probanden ununterbrochene Wachsamkeit verlangt. Die zu beachtenden Signale erscheinen unregelmäßig und sie erwecken keine unwillkürliche Aufmerksamkeit. Daraus ergibt sich die Forderung nach relativ geringer Intensität der Reizdarbietung und niedriger Frequenz der kritischen Ereignisse. Der Leistungsabfall bei Vigilanzexperimenten wird mit der Abnahme des Aktivierungsniveaus der Probandin/des Probanden und der damit einhergehenden Zunahme der Reaktionslatenz erklärt. Nach der neurophysiologischen Aktivierungstheorie wird der Kortex infolge Reizarmut vom aufsteigenden retikulären Aktivierungssystem (ARAS) nur unzureichend stimuliert. Die Gehirnrinde erhält für die Aufrechterhaltung bestimmter Tätigkeiten daher nicht die entsprechenden Weckimpulse, was zur psychischen Ermüdung und damit zur Abnahme der Leistungseffizienz führt. In diesem Zusammenhang entstand der Begriff der Überforderung durch Unterforderung.

der Reaktionszeit ist er mit 0,87 Sekunden ($T = 20$) unter dem Durchschnittsbereich der Norm. In der Schnelligkeits-/Genauigkeitsanalyse liegt er demnach im Feld durchschnittlich ungenau/langsam.

Beim Verfahren zur selektiven Aufmerksamkeit hat ***** v.a. Probleme bei der visuellen Vorgabe, es werden überdurchschnittlich viele Verpasser bei einer unterdurchschnittlichen Reaktionszeit (***** reagiert sehr langsam) registriert. Bei der auditiven Vorgabe ist er im untersten Durchschnittsbereich ($T = 40$). Crossmodal erreicht er weit unterdurchschnittliche Scores.

Im Verfahren zur Überprüfung der räumlichen Aufmerksamkeit erbringt ***** sowohl rechts als auch links unterdurchschnittliche Leistungen ($T = 27$ bzw. $T = 29$). Die Reaktionszeiten sind unterdurchschnittlich (***** reagiert sehr langsam), die Anzahl der Verpasser ist auf beiden Seiten durchschnittlich ($T = 52$). Bei den bilateralen Reizen erreicht ***** ebenfalls einen geringen Prozentrang ($PR = 3$) und ist langsamer als die gleichaltrige Vergleichsgruppe. Die Anzahl der Verpasser liegt auch hier in der Altersnorm.

Bei DL-KG ergibt sich eine unterdurchschnittliche Konzentrationsleistung. Interessant ist die Schwankung der bearbeiteten Zeichen in den Zeitintervallen. Die Leistung ist nicht gleichmäßig, das Leistungsverhalten ist heterogen. Der Quotient aus F % und T deutet auf eine ambivalente individuelle Leistungshaltung hin.

Schlussfolgerung

Es liegt eine Aufmerksamkeitsstörung (F 90.0 (ICD 10)) vor.

Empfohlen wird eine Überprüfung der kognitiven Leistungsfähigkeit und von schulbezogenen Persönlichkeitsmerkmalen mit KLI, AVT, SELMO-S, LAVI und ASF-KJ, um abklären zu können, inwieweit eine lerntherapeutische Intervention sinnvoll sein könnte und inwieweit die schwachen Leistungen durch Motivationsdefizite bedingt sein könnten.

*****, am *****

Anhang 2 – Testübersicht



BASIC-MLT

Battery for Assessment in Children – Merk- und Lernfähigkeitstest für 6- bis 16-Jährige

| | |
|---|--|
| <i>Altersbereich</i> | 6;0–16;11 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Der BASIC-MLT ist ein Merk- und Lernfähigkeitstest, der bei Kindern und Jugendlichen zwischen 6 und 16 Jahren auf Untertest-, Skalen- und Merkquotientenniveau Einblicke in mögliche Defizite des Lern- und Merkvermögens liefert. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte und Prozentrangwerte Reliabilität: mäßig bis gut Validität: geprüft (Konstruktvalidität und klinische Validität) |
| <p>Das BASIC MLT-Verfahren ist für die Diagnostik der Merk- und Lernfähigkeit insbesondere von Kindern der Grundschule und der Sekundarstufe I überaus gut geeignet. Der Test erfasst aber nicht die Konzentrationsfähigkeit, sondern primär die Verarbeitungskapazität, welche hoch mit der Konzentrationsfähigkeit korreliert ist und einen guten Prädiktor für die Konzentrationsfähigkeit darstellt. Will man Aufmerksamkeits- bzw. Konzentrationsfähigkeit prüfen, sollte man zu diesem Test zusätzlich spezifische Verfahren zur Überprüfung der Aufmerksamkeits- bzw. Konzentrationsfähigkeit vorgeben</p> <p><i>Ausführliche Beschreibung ab S. 33.</i></p> | |

COG

Cognitrone (Attention Measure)

| | |
|--|--|
| <i>Altersbereich</i> | ab 4;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Bei Cognitrone handelt es sich um einen computergestützten Test, der Aufmerksamkeit und Konzentration durch Vergleich von Figuren hinsichtlich Kongruenz erfasst. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte und Prozentränge Reliabilität: gut Validität: inhaltliche Gültigkeit, konvergente bzw. diskriminante Validität, Konstruktvalidität und Kriteriumsvalidität gegeben |
| <p>Das Cognitrone-Programm eignet sich gut, um Defizite in der Aufmerksamkeit zu lokalisieren. Probandinnen/Probanden haben mit dem Cognitrone meist keine Probleme und auch kleine Kinder können die Aufgaben des Cognitrone bereits lösen. Die Kennwerte sind verlässlich und zeitlich relativ stabil.</p> <p><i>Ausführliche Beschreibung ab S. 42.</i></p> | |

CompACT

Computerized Attention and Concentration Tests

| | |
|---|---|
| <i>Altersbereich</i> | ab 14;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Die computerisierte Leistungstestbatterie Computerized Attention and Concentration Tests (CompACT) umfasst vier voneinander unabhängige Module, die ein breites Spektrum an Aufmerksamkeits- und Konzentrationsleistungsdimensionen abdecken. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T- und Prozentrangwerte nach Alter und Geschlecht Reliabilität: gut bis sehr gut Validität: Studien zur konvergenten, diskriminanten und faktoriellen Validität sowie Iteminterkorrelationen liegen vor. |
| <p>Das CompACT-Verfahren ähnelt dem WAF-Verfahren und auch der TAP in Aufbau und Durchführung. Die Erfahrungen mit Jugendlichen und Erwachsenen mit derartigen Verfahren sind gut, die Kennwerte sehr gut geeignet, um festzustellen, ob (bei ausreichend guter Motivation, insbesondere bei den Vigilanz- und Daueraufmerksamkeits-Modulen) ein krankheitswertiges Aufmerksamkeitsdefizit vorliegt oder nicht. Bei Schülerinnen und Schülern sollte in jedem Fall ein Verfahren zur Überprüfung der Arbeitshaltung (Arbeitseinstellung, Leistungsmotivation, Anstrengungsvermeidung, Lern- und Arbeitsverhalten etc.) durchgeführt werden.</p> <p><i>Ausführliche Beschreibung ab S. 47.</i></p> | |

DL-KG

Differentieller Leistungstest – KG

| | |
|--|--|
| <i>Altersbereich</i> | 6;0–10;0 Jahre |
| <i>Struktur</i> <i>Kurz-</i> <i>beschreibung</i> | Der Differentielle Leistungstest DL-KG ist ein Durchstreichtest und verzichtet auf den Anspruch, Konzentrationsfähigkeit nur in einem Testwert zu erfassen. Der Test liefert drei Messwerte für verschiedene Leistungsvariablen (Quantität, Qualität und Gleichmäßigkeit der Leistung). Einzel- und Gruppentest |
| <i>Normen</i> | Normen: Prozentränge Reliabilität: hinreichend |
| <i>Testgüte</i> | Validität: Inhaltsvalidität gegeben; Korrelationen mit ähnlichen Verfahren im mittleren Bereich |

Das Instrument ist zur Feststellung spezifischer Leistungsvariablen gut geeignet und ist in leicht abgewandelter Form auch bei jüngeren Probandinnen/Probanden anwendbar (anstatt 14 nur 10 Zeitintervalle). Es ist für Kinder leicht verständlich und die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht sehr gut, aufgrund der Intervalltestung und des relativ langen Testzeitraums, die fokussierte, selektive und die Daueraufmerksamkeit bei Kindern zu überprüfen und in Kombination mit Verfahren zur Überprüfung zur Vigilanz und zur räumlichen Aufmerksamkeit AD(H)S-Kinder mit hoher Wahrscheinlichkeit zu identifizieren. Allerdings ist aufgrund der teilweise fehlenden Normwerte und der Reliabilitätsmängel Vorsicht bei individualdiagnostischer Verwendung geboten.

Ausführliche Beschreibung ab S. 54.

DL-KE

Differentieller Leistungstest – KE

| | |
|--|---|
| <i>Altersbereich</i> | ab 6;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Mit dem DL-KE kann bei Grundschülerinnen und Grundschülern in der Eingangsstufe die Konzentrationsleistung in quantitativer und qualitativer Hinsicht erfasst werden. Damit wird eine differentialdiagnostische Unterscheidung zwischen eindeutig konzentrationsbedingten und neurotischen bzw. organischen Lern- und Leistungsstörungen ermöglicht. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Prozentränge getrennt für Geschlechter Reliabilität: mäßig bis gut Validität: geprüft |
| <p>Das Instrument ist – ebenso wie der DL-KG – zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Es ist für Kinder leicht verständlich und die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht es, die fokussierte, selektive und die Daueraufmerksamkeit bei Kindern zu überprüfen und ist insbesondere in Kombination mit anderen Verfahren wie der KiTAP geeignet, auch Kinder mit AD(H)S zu identifizieren. Allerdings ist aufgrund der schon etwas älteren Normierung Vorsicht geboten.</p> <p><i>Ausführliche Beschreibung ab S. 59.</i></p> | |

d2-R

Aufmerksamkeits- und Konzentrationstest

| | |
|--|---|
| <i>Altersbereich</i> | 9;0–60;0 Jahre |
| <i>Struktur</i> | Der d2-R misst Konzentration bzw. konzentrierte Aufmerksamkeit mittels eines Durchstreichtests mit figuralem Material. Der Test erfordert die Selektion relevanter Reize unter Zeitdruck; er erfasst die Konzentrationsfähigkeit, die Schnelligkeit und Genauigkeit bei der Unterscheidung ähnlicher visueller Reize (Detail-Diskrimination). |
| <i>Kurzbeschreibung</i> | Der Test d2-R stellt eine Weiterentwicklung des bewährten Tests d2 dar. U.a. wurde der Test verlängert und die Auswertung vereinfacht. Die Testaufgabe und die zu bearbeitenden Zeichen blieben bei der Revision unverändert. Einzel- und Gruppentest |
| <i>Normen</i> | Normen: Prozentrang und Standardwertnormen für verschiedene Altersgruppen Reliabilität: gut |
| <i>Testgüte</i> | Validität: Studien zur konvergenten, diskriminanten und faktoriellen Validität liegen für Vorgängerversionen des d2-R vor. |
| <p>Mit dem d2-R kann die fokussierte Aufmerksamkeit überprüft werden. Für Kinder erweist sich eine Testung mit dem DL-KG bzw. DL-KE als vorteilhafter, welche im Grunde dieselben Kompetenzen abprüfen wie der d2-R-Durchstreichtest, mit dem Unterschied, dass sich die Distraktoren beim DL-KG und DL-KE deutlicher unterscheiden. Für Schüler/innen der Sekundarstufen I und II eignet sich dieser Test insbesondere in Kombination mit computerbasierten Verfahren zur geteilten und räumlichen Aufmerksamkeit sehr gut, um Aufmerksamkeitsdefizite festzustellen.</p> <p>Bei schwierigen Fällen hat es sich bewährt, zusätzlich zum d2-R – sofern ausreichend Zeit zur Verfügung steht – auch Verfahren zur Erfassung der Arbeitseinstellung und Motivation durchzuführen, um eine eventuelle Fehldiagnose im Sinne einer Aufmerksamkeitsstörung aufgrund mangelnder Motivation zu vermeiden.</p> <p><i>Ausführliche Beschreibung ab S. 63.</i></p> | |

FAKT-II

Frankfurter Adaptiver Konzentrationsleistungs-Test II

| | |
|--|--|
| <i>Altersbereich</i> | ab 10;0 Jahren (Normen ab 16;0 Jahren) |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Der FAKT-II ist ein computerbasierter, adaptiver Leistungstest der Konzentrationsfähigkeit. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Prozentrang-Werte und Staninewerte Reliabilität: mäßig bis sehr gut Validität: geprüft |
| <p>Der FAKT II ist durch die Möglichkeit des adaptiven Testens eine attraktive Alternative zu den herkömmlichen Durchstreichtests, strengt aber die Testpersonen erfahrungsgemäß sehr an. Es dürfte oft sinnvoller sein, die TAP oder das WAF oder das CompACT-Verfahren durchzuführen, um ein vollständiges Bild der Aufmerksamkeitsleistung zu bekommen. Will man nur Teilaspekte genauer prüfen, so eignet sich dieser Test hervorragend.</p> <p><i>Ausführliche Beschreibung ab S. 68.</i></p> | |

FAIR-2

Frankfurter Aufmerksamkeits-Inventar 2

| | |
|--|--|
| <i>Altersbereich</i> | 9;0–85;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Mit dem FAIR-2 sollen interindividuelle Unterschiede in der gerichteten Aufmerksamkeit abgebildet werden, definiert als Fähigkeit zur konzentrierten (genauen und schnellen) Unterscheidung visuell ähnlicher Zeichen unter Ausblendung aufgabenirrelevanter Information. Das Verfahren kann in der Forschung sowie in allen Praxisfeldern eingesetzt werden, in denen eine Diagnostik von Aufmerksamkeits- und Konzentrationsleistungen erforderlich ist. Einzel- oder Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: altersspezifische Prozentrang- und Staninewerte Reliabilität: mäßig bis sehr gut Validität: Kriteriumsvalidität, konvergente und diskriminante Validität gegeben |
| <p>Das FAIR-2 eignet sich als Alternative zum d2 und zum DL-KG. Kinder im Grundschulalter haben erfahrungsgemäß mehr Spaß am DL-KG. Interessant ist das Profil von Kindern mit AD(H)S im Vergleich zu normalen Kindern. Es zeigt sich, dass Kinder mit AD(H)S im Ausmaß der verfügbaren kognitiven Ressourcen (Testwert L) sowie im Ausmaß der kontinuierlich aufrechterhaltenen Konzentration (Testwert K) signifikant über einer klinisch unauffälligen Kinderstichprobe lagen, im Ausmaß der Selbstkontrollfunktion (Testwert Q) aber tendenziell niedrigere Werte erzielen, das Profil der AD(H)S-Kinder ist also V-förmig im Vergleich zu der klinisch unauffälligen Stichprobe, welches A-förmig ist (Siebert, 2002).</p> <p>Um ein umfassendes Bild der Konzentrationsleistung zu bekommen, bietet es sich an, zusätzlich oder stattdessen computergestützte Verfahren zur räumlichen und geteilten Aufmerksamkeit durchzuführen.</p> <p><i>Ausführliche Beschreibung ab S. 78.</i></p> | |

FWIT

Farbe-Wort-Interferenztest

| | |
|--|---|
| <i>Altersbereich</i> | 10;0–84;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Als objektiver und zuverlässiger, mehrdimensionaler Leistungstest misst der Farbe-Wort-Interferenztest elementare Fähigkeiten der Informationsverarbeitung (Auswahl, Codierung und Decodierung) im optisch-verbale Funktionsbereich. Mit dem Test werden Verarbeitungsgeschwindigkeiten erfasst (Wahrnehmung, begriffliche Umsetzung und verbale Wiedergabe von Reizen durch das Lesen bzw. Benennen). Mittels des Farbe-Wort-Inkongruenzprinzips nach J. R. Stroop wird zudem die Stressbelastung und damit die konzentrierte Beanspruchung variiert. Die mit dem Test messbaren kognitiven Leistungsfunktionen sind: Nomination (Geschwindigkeit der Namenfindung, Benennung), Selektivität (konzentrativer Widerstand gegenüber dominierenden Reaktionstendenzen oder Interferenzneigung), Alertness (Grundgeschwindigkeit der Informationsverarbeitung) sowie – unter bestimmten Voraussetzungen – die Lesegeschwindigkeit. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte nach Alters- und Berufsgruppen Reliabilität: mäßig bis sehr gut Validität: Kriteriumsvalidität und faktorielle Validität geprüft |
| <p>Der FWIT ist sehr gut geeignet, um die Interferenz-Fähigkeit von Personen zu überprüfen. Es handelt sich um ein schnell erklärtes und kurzweiliges Verfahren, mit dem sehr schnell die Interferenzfähigkeit festgestellt werden kann. Will man allerdings eine umfassende Aufmerksamkeitsdiagnostik betreiben, sollte unbedingt ein weiteres Verfahren hinzugezogen werden.</p> <p><i>Ausführliche Beschreibung ab S. 93.</i></p> | |

INKA

Inventar komplexer Aufmerksamkeit

| | |
|--|---|
| <i>Altersbereich</i> | 14;0–39;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Das INKA misst die Konzentrationsfähigkeit über die Kapazität der Informationsverarbeitung. In 18 Buchstabenreihen sollen bestimmte Buchstaben und/oder Buchstabenpärchen gesucht werden. Die Buchstaben müssen zuerst über eine Umwandlungstabelle gefunden werden. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Mittelwerte und Standardabweichungen Reliabilität: ausreichend Validität: Faktorielle Validität und Kriteriumsvalidität geprüft |
| <p>Das INKA hat den Vorteil, dass sich unterschiedliche Ausprägungen individueller Fertigkeiten nicht auf die Bearbeitungsleistung auswirken und dass sich das Inventar in der Praxis gut bewährt hat. AD(H)S kann mit diesem Test nicht festgestellt werden, vielmehr lässt sich eine Aussage über die Konzentrationsfähigkeit und die Arbeitsspeicherkapazität der Testperson treffen. Bei Anwendung des Verfahrens ist es wichtig, die Motivation der Testperson in die Diagnose miteinzubeziehen, da diese beim INKA eine große Rolle spielen kann.</p> <p><i>Ausführliche Beschreibung ab S. 101.</i></p> | |

KKA

Kaseler-Konzentrations-Aufgabe für 3- bis 8-Jährige

| | |
|--|---|
| <i>Altersbereich</i> | 3;0–8;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Die KKA erfasst die kurzzeitige selektive Aufmerksamkeits- und Konzentrationsleistung (Leistungsmenge in vorgegebener Zeit) von Vorschul- und Primarschulkindern im Alter von 3 bis 8 Jahren. Es handelt sich bei der KKA um einen Durchstreichtest, bei dem die Kinder die Aufgabe haben, in mehreren Reihen mit verschiedenen Abbildungen die jeweiligen Zielobjekte anzustreichen, wobei die Bearbeitungszeit pro Testreihe auf 10 Sekunden festgelegt ist. Zeitnahe oder zeitferne Wiederholungsmessungen sind durch 6 Parallelformen der KKA möglich; ergänzend können Leistungsverläufe bei der Anwendung von KKA-Testserien zu einem Testzeitpunkt (maximal sechs KKA-Parallelformen) und die Anzahl von Verwechslungsfehlern zur Interpretation herangezogen werden. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T- und Prozentrangwerte Reliabilität: sehr gut Validität: Umfangreiche Befunde zur konvergenten und diskriminanten Validität |
| <p>Die KKA ähnelt dem DL-KG und dem DL-KE. Sie ist zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Die Kinder haben in der Regel Spaß bei der Durchführung. Der Test ermöglicht sehr gut, die fokussierte und die selektive Aufmerksamkeit bei Kindern zu überprüfen und sollte in Kombination mit weiteren Verfahren zur Überprüfung der Vigilanz und der räumlichen Aufmerksamkeit angewendet werden, sofern ein umfassendes Bild der Konzentrationsleistung gewünscht wird.</p> <p><i>Ausführliche Beschreibung ab S. 106.</i></p> | |

KHV-VK

Konzentrations-Handlungsverfahren für Vorschulkinder

| | |
|--|---|
| <i>Altersbereich</i> | 3;0–6;11 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Die Konzentrationsfähigkeit eines Kindes stellt eine Grundbedingung für das Lernen und das Erbringen von schulischen Leistungen dar. Mögliche Defizite sollten daher schon im Vorschulalter festgestellt werden. Zu diesem Zweck wurde das Konzentrations-Handlungsverfahren für Vorschulkinder (KHV-VK) entwickelt. Als kindgemäßes Sortierverfahren, das die Parameter Fehler (Sorgfaltsleistung) und Zeit (Arbeitstempo) erfasst, ermöglicht es Aussagen zum Niveau konzentrativer Fähigkeiten und stellt auch bei der Diagnostik von AD(H)S im Vorschulalter eine wichtige Hilfe dar. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte Reliabilität: mäßig bis gut Validität: geprüft |
| <p>Das Verfahren ist eines der wenigen, das sich für Kinder im Vorschulalter eignet. Die Kinder haben erfahrungsgemäß Freude an der Durchführung und verstehen die Instruktion gut. Defizite in der Aufmerksamkeitsleistung können somit schon im Vorschulalter festgestellt werden, wodurch frühzeitig Fördermaßnahmen gesetzt werden können. Für eine umfassende AD(H)S-Diagnostik eignet sich der Test nicht. Hierzu braucht man unbedingt Informationen von Kindergartenpädagoginnen und -pädagogen, Eltern und eine diagnostische Verhaltensbeobachtung.</p> <p><i>Ausführliche Beschreibung ab S. 113.</i></p> | |

KLT-R

Konzentrations-Leistungs-Test – Revidierte Fassung

| | |
|---|--|
| <i>Altersbereich</i> | 4.–13. Schulstufe und Erwachsene |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Der KLT-R misst die individuelle Konzentrationsfähigkeit. Im Unterschied zu Tests der Kurzzeitanspannung wird hier die Langzeitanspannung gemessen. Erfasst werden sowohl die Quantität als auch die Qualität der Dauerbeanspruchung sowie der Leistungsverlauf einer Testperson. Die leichtere Form des KLT-R (KLT-R 4–6) ist für die Anwendung bei Schüler/innen der 4.–6. Schulstufe geeignet. Zusätzliche Einsatzmöglichkeiten dieser Form liegen im Bereich der Förderschule. Die schwierigere Form (KLT-R 6–13) ist in der 6.–13. Schulstufe und bei Erwachsenen anwendbar. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Prozentrang-Werte und Standardwerte Reliabilität: gut bis sehr gut Validität: konvergente und divergente Validität geprüft |
| <p>Der KLT eignet sich hervorragend zur Überprüfung der Konzentrationsfähigkeit bei Kindern und Jugendlichen, die mit dem Rechnen keine Probleme haben, nach Möglichkeit gute Kopfrechner/innen sind und bei welchen auch kein Motivationsdefizit bzw. keine unzureichende Arbeitseinstellung vorliegt. Weiters ist die Testleistung nicht unabhängig von der Arbeitsspeicherkapazität. Zur Feststellung von AD(H)S ist dieser Test eher ungeeignet.</p> <p><i>Ausführliche Beschreibung ab S. 116.</i></p> | |

KVT

Konzentrations-Verlaufs-Test

| | |
|---|--|
| <i>Altersbereich</i> | 14;0–60;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Der Konzentrations-Verlaufs-Test erlaubt eine genaue Analyse des Arbeitsverhaltens unter Daueraufmerksamkeitsbedingungen. Die Probandin/der Proband soll einen Kartenstoß nach vier Kriterien durchsehen und dahingehend sortieren, ob die Karten die eine oder andere von zwei vorgegebenen Zahlen enthalten oder alle beide bzw. keine von beiden. Gemessen werden Arbeitszeit und Fehlerzahl; es ergibt sich eine Arbeits-Verlaufskurve. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Standardwerte aus den Jahren 1960 und 1974 Reliabilität: unzureichend Validität: nicht hinreichend |
| <p>Der KVT eignet sich nur bedingt für den Einsatz in der Schulpsychologie. Es ist (ähnlich wie auch bei anderen Verfahren wie dem FAIR-2 oder dem d2) möglich, Arbeitszeit und Fehleranzahl zu überprüfen, also die Schülerin/den Schüler innerhalb der 4-Felder-Tafel „schnell/langsam x genau/ungenau“ einzustufen. Eine AD(H)S-Abklärung kann mit diesem Test nicht gemacht werden. Die Angemessenheit der Normen (aus den Jahren 1960 und 1974) ist fraglich.</p> <p><i>Ausführliche Beschreibung ab S. 124.</i></p> | |

KT 3–4 R

Konzentrationstest für 3. und 4. Klassen – revidierte Fassung

| | |
|--|---|
| <i>Altersbereich</i> | 3. und 4. Schulstufe |
| <i>Struktur</i> <i>Kurz-</i> <i>beschreibung</i> | Der KT 3–4 R dient zur Erfassung der Aufmerksamkeit und Konzentrationsleistung von Kindern der dritten und vierten Klassenstufe. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Rohwertescores getrennt nach Geschlecht Reliabilität: gut bis sehr gut Validität: umfangreiche Befunde |
| <p>Der Test ähnelt vom Prinzip der DL-KG, dem FAIR-2 und der KKA, hat jedoch andere Stimuli. Er ist zur Feststellung von Aufmerksamkeitsaspekten gut geeignet. Besonders Kinder mit einer Begabung im räumlichen Denken haben Spaß bei der Durchführung. Andere Kinder – insbesondere AD(H)S-Kinder – verlieren, wenn sie das Material sehen, oft das Interesse und die Lust, den Test durchzuführen. Der Test ermöglicht es sehr gut, die fokussierte und die selektive Aufmerksamkeit bei Kindern zu überprüfen und sollte in Kombination mit weiteren Verfahren zur Überprüfung der Vigilanz und der räumlichen Aufmerksamkeit angewendet werden, sofern ein umfassendes Bild der Aufmerksamkeitsleistung gewünscht wird. Eine Ausdehnung des Anwendungsbereichs auf Jugendliche könnte sinnvoll sein.</p> <p><i>Ausführliche Beschreibung ab S. 129.</i></p> | |

P-T

Pauli-Test

| | |
|---|---|
| <i>Altersbereich</i> | ab 7;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Beim Pauli-Test handelt es sich um einen Test zur Erfassung der Konzentrationsleistung. Die Probandin/der Proband soll auf einem standardisierten Rechenbogen fortlaufend und möglichst schnell je zwei einstellige Zahlen addieren und zwar derart, dass in jede Lücke rechts die Summe der darüber und darunter stehenden Zahl zu schreiben ist, unter Weglassung der Ziffer eins, wenn das Ergebnis zweistellig ist. Alle drei Minuten muss die Probandin/der Proband eine Markierung (Querstrich) in der Zahlenreihe vornehmen, um eine Verlaufsauswertung zu ermöglichen. Das Ergebnis des Arbeitsversuchs ist dann unter den Gesichtspunkten Menge, Güte und Verlauf der Leistung zu betrachten. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Summenwerte Reliabilität: mäßig bis sehr gut Validität: geprüft |
| <p>Der Pauli-Test bevorzugt Probandinnen/Probanden mit guten Kopfrechenfähigkeiten. Zudem spielt hier die Kapazität des Arbeitsgedächtnisses eine Rolle im Hinblick auf die Anzahl der bearbeiteten Aufgaben. Es kann daher meist nicht differenziert werden, ob mögliche schlechte Leistungen auf mangelnde Konzentrationsfähigkeit oder schlichtweg auf die mangelnde Fähigkeit, im Kopf schnell zu rechnen, zurückzuführen sind. Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind/der Jugendliche mit dem Kopfrechnen Probleme hat oder an einer möglichen Rechenschwäche leidet.</p> <p><i>Ausführliche Beschreibung ab S. 136.</i></p> | |

Rev.T.

Revisions-Test

| | |
|---|--|
| <i>Altersbereich</i> | 9;6–75;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | <p>Im Revisions-Test wird mittels einfacher Kopfrechenaufgaben ermittelt, wie die Konzentrationsfähigkeit anhält. Der Test dient also als allgemeiner Leistungstest zur Untersuchung anhaltender Konzentration bei geistiger Tempoarbeit. Der Test umfasst 15 Testzeilen mit jeweils 44 Items, wobei sich die Items aus zufällig auf den Aufgabenbögen verteilten Additions- und Subtraktionsaufgaben zusammensetzen. Die Testperson hat die Aufgabe, die Ergebnisse auf ihre Richtigkeit zu überprüfen. Die Bewältigung dieser Aufgabe erfolgt unter einem zeitlich vorgegebenen Rahmen. Bewertet wird anhand quantitativer und qualitativer Gesichtspunkte.</p> <p>Einzel- und Gruppentest</p> |
| <i>Normen</i> <i>Testgüte</i> | <p>Normen: Prozentrang- und Standardwerte nach Geschlecht und nach Berufsgruppen</p> <p>Reliabilität: gut bis sehr gut</p> <p>Validität: geprüft</p> |
| <p>Dieser Test führt insbesondere bei Kindern im Alter zwischen 9 und 12 Jahren oft zu einer Verzerrung, da Probandinnen/Probanden mit guten Kopfrechenfähigkeiten hier signifikant besser abschneiden als Kinder/Jugendliche mit Problemen beim Kopfrechnen oder mit Dyskalkulie. Zudem spielt hier die Kapazität des Arbeitsgedächtnisses eine Rolle im Hinblick auf die Anzahl der bearbeiteten Aufgaben. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit oder schlichtweg an der mangelnden Fähigkeit, im Kopf schnell zu rechnen, bzw. einer geringen Kapazität des Arbeitsgedächtnisses liegen. Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind/der Jugendliche mit dem Kopfrechnen Probleme hat oder an einer möglichen Rechenschwäche leidet.</p> <p><i>Ausführliche Beschreibung ab S. 144.</i></p> | |

SIGNAL

Signal-Detection

| | |
|--|---|
| <i>Altersbereich</i> | ab 6;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Mit dem Signal-Detection-Test aus dem Wiener Testsystem liegt ein computergestütztes Verfahren zur Messung von Konzentration und Aufmerksamkeit sowie der visuellen Differenzierungsleistung vor. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte und Prozentränge Reliabilität: hinreichend Validität: konvergente und diskriminante Validität geprüft |
| <p>SIGNAL eignet sich hervorragend als Zusatzinstrument für die Diagnostik von Aufmerksamkeitsdefiziten und kann gut als Screeningverfahren eingesetzt werden. Für eine umfassende AD(H)S-Diagnostik im Schulkontext ist dieses Programm nicht geeignet, liefert aber wertvolle Hinweise auch zu Verarbeitungsgeschwindigkeit und Reaktion.</p> <p><i>Ausführliche Beschreibung ab S. 149.</i></p> | |

TAP

Testbatterie zur Aufmerksamkeitsprüfung

| | |
|--|---|
| <i>Altersbereich</i> | 6;0–90;0 Jahre |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Die Testbatterie zur Aufmerksamkeitsprüfung erfasst mit computergestützten Testverfahren die unterschiedlichen Komponenten von Aufmerksamkeit, u.a. Alertness, selektive Aufmerksamkeit, geteilte Aufmerksamkeit und Vigilanz. Die Verfahren sind von geringer Komplexität, um die Testleistungen nicht durch sensorische oder motorische Ausfälle, Gedächtnisstörungen, Sprachstörungen oder andere Defizite zu beeinflussen. Das Verfahren wird vorzugsweise in der neuropsychologischen Diagnostik eingesetzt, da Beeinträchtigungen der Aufmerksamkeits- und Konzentrationsleistung eine der häufigsten Schädigungsfolgen bei traumatischen Hirnschädigungen darstellen. Es findet aber auch in der Diagnostik von Aufmerksamkeitsstörungen bei Kindern und Jugendlichen Einsatz. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Prozentrang-Werte Reliabilität: gut bis sehr gut (Ausnahme: geteilte Aufmerksamkeit) Validität: belegt |
| <p>Die TAP eignet sich sehr gut für eine umfassende Abklärung eines möglichen Aufmerksamkeitsdefizits bei Jugendlichen und mit Einschränkungen auch bei Kindern. Bei diesem Testverfahren kann bei Jugendlichen eine mangelnde Arbeitseinstellung oder fehlende Motivation zu Fehleinschätzungen führen. Für kleinere Kinder bietet sich alternativ die KiTAP an. Es empfiehlt sich, die KiTAP in mehreren Sitzungen durchzuführen, um Müdigkeitseffekte oder ein Absinken der Motivation zu vermeiden.</p> <p><i>Ausführliche Beschreibung ab S. 155.</i></p> | |

TPK

Testreihe zur Prüfung der Konzentrationsfähigkeit

| | |
|--|---|
| <i>Altersbereich</i> | 2.–6. Schulstufe |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Die TPK ermöglicht bei Schülerinnen und Schülern der Klassen 2–6 eine differenzierte Diagnostik von Konzentrationsstörungen mit dem Ziel, sonderpädagogischen Förderbedarf zu ermitteln bzw. geeignete Trainingsmaßnahmen einzuleiten. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: Centilwerte Reliabilität: gut Validität: geprüft |
| <p>Dieser Test eignet sich hervorragend zur Überprüfung von schulnahen Aspekten der Konzentration. Will man ein Aufmerksamkeitsdefizit feststellen, eignet er sich nur bedingt, da Teil 1, wie auch der Revisions-Test, insbesondere bei Kindern im Alter zwischen 9 und 12 Jahren oft zu einer Verzerrung führt, da Probandinnen/Probanden mit guten Kopfrechenfähigkeiten hier signifikant besser abschneiden als Kinder/Jugendliche mit Problemen beim Kopfrechnen oder mit Dyskalkulie. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit liegen oder schlichtweg an der mangelnden Fähigkeit, im Kopf schnell zu rechnen. Beim zweiten Teil spielt die Rechtschreibfähigkeit zusätzlich zur Konzentrationsfähigkeit eine nicht unbedeutende Rolle, wenn es um die Fehleranzahl und Schnelligkeit beim Text-Abschreiben geht. Es kann daher meist nicht differenziert werden, ob mögliche Verzögerungen an mangelnder Konzentrationsfähigkeit liegen oder an mangelnden Rechtschreibkenntnissen bzw. Legasthenie. Der dritte Teil des TPK eignet sich hervorragend, um Konzentrationsfähigkeit (und Gedächtnis) zu überprüfen und macht den Kindern erfahrungsgemäß großen Spaß. Dieser kann z.B. in Kombination mit anderen Verfahren vorgegeben werden.</p> <p>Will man mit diesem Verfahren arbeiten, sollte unbedingt im Vorfeld abgeklärt werden, inwieweit das Kind/der Jugendliche mit dem Kopfrechnen und/oder mit der Rechtschreibung Probleme hat bzw. an einer möglichen Rechenschwäche bzw. Legasthenie leidet. Unter Umständen kann man auch nur mit dem dritten Teil arbeiten, da dieser sehr aufschlussreiche Zusatzergebnisse im Rahmen einer Aufmerksamkeitsdiagnostik liefern kann.</p> <p><i>Ausführliche Beschreibung ab S. 170.</i></p> | |

VIGIL

Vigilanz

| | |
|--|--|
| <i>Altersbereich</i> | ab 6;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Das VIGIL-Verfahren ermöglicht Aussagen über die Genauigkeit und Geschwindigkeit der Bearbeitung von Vigilanzaufgaben sowie über den Verlauf der Leistung über die Zeit. Einzeltest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte und Prozentränge Reliabilität: mäßig bis sehr gut Validität: geprüft |
| <p>Das VIGIL-Verfahren ist bei Kindern und Jugendlichen aufgrund der langen Aufmerksamkeitsspanne bei verhältnismäßig wenigen Ereignissen eher unbeliebt. Es ermöglicht allerdings eine gute Abschätzung und Einteilung der Probandin/des Probanden in die 4-Felder-Tafel „schnell/langsam x genau/ungenau“ und die genaue Erfassung der Aufmerksamkeitsleistung bei Dauerbeanspruchung unter dem Paradigma der „low event rate“, welche mit den herkömmlichen Papier-Bleistift-Tests nicht valide erfasst werden kann.</p> <p><i>Ausführliche Beschreibung ab S. 177.</i></p> | |

WAF

Wahrnehmungs- und Aufmerksamkeitsfunktionen

| | |
|--|--|
| <i>Altersbereich</i> | ab 7;0 Jahren |
| <i>Struktur</i> <i>Kurz- beschreibung</i> | Das Verfahren ermöglicht die Erfassung von Teilfunktionen der Aufmerksamkeit: Daueraufmerksamkeit, geteilte Aufmerksamkeit, selektive Aufmerksamkeit, räumliche Aufmerksamkeit, fokussierte Aufmerksamkeit und Alertness. Einzel- und Gruppentest |
| <i>Normen</i> <i>Testgüte</i> | Normen: T-Werte und Prozentrang-Werte Reliabilität: gut bis sehr gut Validität: geprüft |
| <p>Die WAF eignet sich sehr gut für eine umfassende Abklärung eines möglichen Aufmerksamkeitsdefizits bei Jugendlichen und mit Einschränkungen auch bei Kindern. Problematisch kann bei den Jugendlichen oft eine mangelnde Arbeitseinstellung oder fehlende Motivation sein, die hier zu Fehleinschätzungen führen kann. Für kleinere Kinder bietet sich alternativ die KiTAP an. Es empfiehlt sich, die WAF-Batterie in mehrere Sitzungen zu unterteilen, um Müdigkeitseffekte oder ein Absinken der Motivation zu vermeiden.</p> <p><i>Ausführliche Beschreibung ab S. 185.</i></p> | |





Ein herzliches Dankeschön an Bali, Christina, Hermine, Katharina, Raphael und Samuel, die sich für die Fotos zur Verfügung stellten und mit Begeisterung, Ausdauer, Konzentration und kreativen Ideen Banane und Co. sezieren...

In dem vorliegenden Handbuch, dem dritten Band der ÖZBF-Publikationsreihe zu Testverfahren, werden nach einem Überblick über Theorien und Modelle zu Aufmerksamkeit und Konzentration und einer Klärung der wichtigsten Begriffe Testverfahren zur Messung von Konzentrationsfähigkeit und zur Diagnose von Aufmerksamkeits(störungen) vorgestellt.

Die Beschreibungen der Testverfahren beinhalten u.a. Angaben zur Durchführung, Informationen zu Auswertung und Interpretation (inklusive Testgütekriterien) sowie Kommentare zu den einzelnen Verfahren. Im Anhang finden sich Muster für Gutachten und Befunde, die eine Hilfestellung im Rahmen der Diagnostik von Aufmerksamkeits- und Konzentrationsfähigkeit bieten sollen.

Eine verantwortungsvolle Begabungsdagnostik begnügt sich nicht mit Messung von Intelligenz (Band 1). Sie schließt verschiedene nicht-kognitive Persönlichkeitsmerkmale mit ein, etwa Motivation, Prüfungsangst, Selbstkonzept oder Stressverarbeitung (Band 2). Ebenso wichtige Faktoren im Kontext von Begabung sind Aufmerksamkeit und Konzentration. Der vorliegende dritte Band hilft, geeignete Testverfahren für die Diagnostik von Aufmerksamkeits- und Konzentrationsfähigkeit auszuwählen.