

Empirische Sonderpädagogik, 2011, Nr. 1, S. 3-19

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)

Thomas Dreisörner¹, Janine Georgiadis²

¹Universität Frankfurt am Main, ²Hildesheim

Theoretischer Hintergrund: Die besonders häufig gestellte Diagnose ADHS wird gemäß den Empfehlungen der DGKJPP anhand der Exploration von Eltern und Kind, der störungsspezifischen Entwicklungsgeschichte, Informationen aus der Schule und einer orientierenden Überprüfung der Intelligenz und schulischen Leistungsfähigkeit gestellt. Des Weiteren werden als Möglichkeit zur Verbesserung der Diagnosestellung Aufmerksamkeitstests genannt. Im deutschsprachigen Raum werden derzeit unter anderem die Testbatterie zur Aufmerksamkeitsprüfung (TAP) und die Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP) eingesetzt. Wir untersuchten die Sensitivität und Spezifität beider Verfahren.

Methode: In vorliegender Studie wurden 24 Kinder mit der Diagnose ADHS und 24 Kontrollkinder mit der TAP und der KITAP untersucht. Mittels ROC-Analysen und logistischer Regressionen wurden die Vorhersagegüte, Sensitivität und Spezifität der Parameter berechnet.

Ergebnisse: Die Kontrollgruppe erbrachte mehrheitlich signifikant bessere Leistungen in den Parametern der Untertests als die klinische Stichprobe. Die TAP wies bei der Diagnose von ADHS eine mäßig höhere Sensitivität bzw. Spezifität auf als die KITAP. In den logistischen Regressionen ließ sich im Modell mit den TAP-Parametern eine höhere Varianz (32.5%) aufklären als mit den KITAP-Parametern (21.6%), wobei sich in der schrittweisen Regression die Varianzaufklärung der KITAP-Parameter durch die Hinzunahme der TAP-Parameter signifikant verbessern ließ. Insgesamt ließen sich durch die TAP 75% der Kinder korrekt klassifizieren (72.9% durch die KITAP).

Fazit: Die in dieser Studie untersuchten computergestützten Verfahren zur Aufmerksamkeitsmessung können einen Beitrag im diagnostischen Prozess bei fraglicher ADHS liefern, sollten jedoch aufgrund der nicht zufriedenstellenden diagnostischen Zuordnung und fehlender Cut-off-Werte für die klinische Praxis allenfalls als Ergänzung zu Interviews und Fragebögen eingesetzt werden.

Schlüsselwörter: ADHS, Aufmerksamkeit, computergestützte Tests, Kinder

Diagnostic Accuracy of Computer based Tests for Attention-Deficit Hyperactivity Disorder: A Receiver-Operating Characteristic Analysis

Objectives: Attention-Deficit Hyperactivity Disorder (ADHD) is one of the most common neurobehavioral diagnosis affecting children today. Guidelines of child psychiatry encourage clinicians to employ clinical interviews and questionnaires, learning and performance tests and behavior obser-

vation in making the diagnosis. Special attention tests can be conducted additionally. Computer-based attention tests constitute a promising method to do so as they appeal to children. In clinical practise in Germany two test batteries are commonly used: "Testbatterie zur Aufmerksamkeitsprüfung" (TAP) and "Testbatterie zur Aufmerksamkeitsprüfung für Kinder" (KITAP). The latter having been adapted to the motivational needs of children. This study aims at examining the diagnostic accuracy of the TAP and KITAP.

Methods: ROC-analysis of TAP- and KITAP parameters and logistic regressions were conducted in a sample of 24 children with and without ADHD.

Results: Both TAP and KITAP proved to differentiate between children with and without ADHD. TAP parameters possessed a higher discriminating power for ADHD. In logistic regression analysis TAP parameters led to an amount of variance of 32.5% (KITAP: 21.6%) and correct classification rate of 75% (KITAP: 72.9%). Adding KITAP parameters did not improve the amount of variance significantly whereas adding TAP parameters to the KITAP data did result in a significant increase.

Conclusion: These findings suggest that TAP and KITAP could serve as useful additional test instruments to help identify cases likely to meet criteria for ADHD in clinical settings. For individual diagnosis TAP and KITAP lack cut-off criteria and substantial classification rates.

Key words: ADHD, attention, computer based test battery, children

Hyperkinetische Störungen gehören zu den häufigsten psychischen Störungen des Kindes- und Jugendalters. Sie verlaufen meist chronisch und sind Anlass zu einer langfristigen Behandlung (Döpfner, Frölich & Lehmkuhl, 2000; Schmidt & Petermann, 2008). In internationalen Studien wurden Prävalenzraten zwischen 3.6 und 6.7% nachgewiesen (Döpfner, Banaschewski & Sonuga-Barke, 2008). Werden die strengeren ICD-10-Kriterien angewandt, liegt in Deutschland die Prävalenz zwischen 1 und 3.4% (Döpfner et al., 2008).

Bis heute ist die eindeutige Diagnosestellung von ADHS schwierig. Die Diagnostiker gründen ihr Urteil auf zahlreiche Basisverfahren der multiplen Verhaltens- und Psychodiagnostik sowie der störungsspezifischen Verhaltens- und Interaktionsdiagnostik (Döpfner et al., 2000). Erstere geben ein umfassendes Bild über Verhaltens- und emotionale Auffälligkeiten des Kindes auf der Basis von Eltern-, Lehrer- und Selbstbeurteilungen, die im Interview oder durch Fragebögen erhoben werden können. Für die störungsspezifische Diagnostik können zum einen spezielle Diagnosechecklisten eingesetzt werden, die sich

an den Kriterien der Klassifikationssysteme orientieren. Zum anderen können Verhaltensbeobachtungen des Kindes und/oder der familiären bzw. schulischen Interaktion durch die/den Diagnostiker/in durchgeführt werden. Für die Differentialdiagnostik (z.B. Entwicklungsstörungen, geistige Behinderung) sind Intelligenz- und Leistungsmessungen unverzichtbar. Daneben geben eine Reihe von Aufmerksamkeits-, Reaktions- und Konzentrationstests Aufschluss über die Ausprägung der Kernsymptomatik, die aber für die Diagnosestellung laut Leitlinien nur optional einzusetzen sind (Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie, 2007). Eine apparative Diagnostik der Aufmerksamkeits- und exekutiver Funktionen kann die Diagnose jedoch erhärten, wobei zu bedenken ist, dass die Korrelation zwischen den Testleistungen und dem Verhalten in natürlichen Situationen, z.B. in der Schule, gering ist.

Es gibt einige Konzentrationstests, die im Kinderbereich einsetzbar sind, z.B. der Differentielle Leistungstest – KG (DL-KG; Kleber, Kleber & Hans, 1999), die Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK; Kurth

& Büttner, 1999) oder der Aufmerksamkeits-Belastungs-Test d2 (Brickenkamp, 2002). Die Objektivität der genannten Verfahren ist hoch, da sie standardisierte Instruktionen und klare Auswertungsrichtlinien haben. Die Reliabilitäten sind z.T. sehr hoch (interne Konsistenz beim d2 zwischen $r = .95$ und $r = .98$). Vorteile des d2 sind die zuverlässige Messung der Konzentrationsfähigkeit und die kurze Durchführungsdauer. Der Test misst aber nur das kurzzeitige Konzentrationsvermögen (5 Minuten) und lässt sich leicht trainieren. Der DL-KG dauert relativ lange (45 Minuten) und wird von den Kindern als anstrengend und langweilig empfunden. Bei der TPK werden Aufmerksamkeitsleistungen in Abhängigkeit vom Vorhandensein schulischer Grundfertigkeiten erfasst. Weitere Angaben zu diesen und anderen Konzentrationstests finden sich in Brickenkamp (2002) und Büttner und Schmidt-Atzert (2004).

Kinder mit ADHS zeigen verschiedene Defizite in exekutiven Funktionen, was in computerbasierten Testverfahren zur Überprüfung von Aufmerksamkeitsleistungen in der Regel zu schlechten Ergebnissen führt. Fraglich ist allerdings, ob diese Defizite Ursache der Aufmerksamkeitsproblematik sind. Die ungenügende Verhaltenssteuerung bei den exekutiven Funktionen äußert sich in verzögerten Reaktionszeiten sowie Schwierigkeiten bei der Reaktionshemmung. Es sind Auffälligkeiten in der Daueraufmerksamkeit und der Steuerung von Motorik sowie voreilige Reaktionen beobachtbar (Eisert, 1995; Gawrilow, Schmitt & Rauch, 2011). Dem neuropsychologischen Modell von Barkley (1997) zufolge liegt bei Kindern mit ADHS-Mischtyp (nach DSM-IV) ein globales Defizit in der Verhaltenshemmung vor. In einer Studie von Hampel, Petermann und Desmann (2009) konnte dies bei Jungen belegt werden. Sie leiten aus ihrer Studie ein globales Defizit in den exekutiven Funktionen ab und gehen von einer Entwicklungsabweichung bei Jungen mit ADHS aus.

Neben Defiziten in exekutiven Prozessen könnten auch motivationale Zustände für diese Symptome verantwortlich sein. Bei Aufgaben, die von den betroffenen Kindern als langweilig oder anstrengend erlebt werden, können sie sich nur kurze Zeit konzentrieren (Stadler & Döpfner, 2004). Es gibt Hinweise darauf, dass Kinder mit ADHS besonders auf den Entzug von erwarteter Verstärkung reagieren und ihre Leistung derjenigen von Kontrollkindern gleicht, wenn sie angemessen motiviert sind. Sie zeigen sich besonders abhängig von sofortiger und positiver Verstärkung und benötigen höhere Anreize (Douglas & Parry, 1994; Slusarek, Velling, Bunk & Eggers, 2001). Es ist demnach notwendig, zwischen Fähigkeit und Performanz bei Aufgaben zur Aufmerksamkeitsmessung zu unterscheiden.

Im Rahmen dieser Studie soll überprüft werden, ob die in der klinischen Praxis häufig verwendeten computergestützten Diagnoseinstrumente „Testbatterie zur Aufmerksamkeitsprüfung“ (TAP, Zimmermann & Fimm, 1993) und „Testbatterie zur Aufmerksamkeitsprüfung für Kinder“ (KITAP, Zimmermann, Gondan & Fimm, 2002) einen substanziellen Beitrag zur Absicherung der Diagnose ADHS leisten können. Die Testbatterien wurden ausgewählt, weil sie wesentliche Aspekte von Aufmerksamkeitsleistungen messen: die Reaktionsgeschwindigkeit, die Sorgfalt der Reaktionen und die Kontinuität bzw. Schwankungen der Aufmerksamkeit. Beides sind mehrdimensionale Verfahren zur Aufmerksamkeitsmessung, die einfache Reaktionsaufgaben beinhalten, bei denen leicht diskriminierbare, sprachfreie Reize mit einem Tastendruck zu beantworten sind. Die TAP wurde für neuropsychologische Untersuchungen im Erwachsenenbereich konzipiert und später auch für die Diagnostik bei Kindern normiert, während die KITAP für den Einsatz im Kinderbereich entwickelt wurde. So war es vorrangiges Ziel, die Motivation der Kinder zu steigern, indem die Aufgaben in eine kindgerechte Geschichte mit dem

Thema „Geisterschloss“ integriert wurden. Ein großer Vorteil beider Verfahren ist die Messung im Millisekundenbereich, wodurch selbst kleinste Abweichungen festgestellt werden können. Anhand der Normierung kann die Leistungsgüte der Kinder identifiziert werden. Für die zu untersuchenden Kinder sind die Verfahren attraktiv, weil sie gern mit dem Medium Computer arbeiten. Es stellt sich jedoch die Frage, ob die Attraktivität von Computern bei der computergestützten Testdiagnostik von Nachteil ist. Bei Tätigkeiten, die vermehrt Aufmerksamkeit benötigen, wie das Spielen eines Computerspiels, können die Anzeichen der Störung ausbleiben, wenn es sich dabei um eine beliebte Beschäftigung des Kindes handelt (Döpfner et al., 2000; Stadler & Döpfner, 2004). Ausgehend von der Beobachtung, dass sich Kinder mit ADHS sehr gut am Computer und insbesondere bei Computer- sowie Konsolenspielen konzentrieren können, untersuchten Shaw, Grayson und Lewis (2007) deren Leistungen bei verschiedenen Computerspielen und standardisierten Aufgaben am Computer zur Messung von Aufmerksamkeit. Wenn das Aufgabenformat einem gewöhnlichen Computerspiel ähnelt, zeigen Experimental- und Kontrollgruppe vergleichbare Leistungen. Das heißt, bei Computerspielen werden die Faktoren Interesse, Anstrengung und Motivation aktiviert und somit die Leistungsunterschiede nivelliert. Von Bedeutung ist die Unterscheidung der Reizanforderung bei den Aufgaben. Wenn nur eine direkte Reaktion auf einen Reiz erforderlich ist, z.B. bei „Jump-and-Run-Spielen“ oder so genannten „Ballerspielen“, erzielen die Kinder mit ADHS ähnlich gute Leistungen wie Kontrollkinder. Wenn aber bei der Aufgabenbewältigung ein hohes Maß an Vorausplanungen, d.h. metakognitiven Kompetenzen, Reflexivität und dauerhafter Konzentration erforderlich ist, schneiden sie schlechter ab.

Drechsler, Rizzo und Steinhausen (2009) kommen in der Bewertung ihrer Untersuchung, in der ebenfalls Kinder mit ADHS mit

Kontrollkindern verglichen wurden, zu dem Schluss, dass die KITAP zwar zwischen Kindern mit und ohne ADHS differenzieren kann, sie jedoch für eine diagnostische Zuordnung anhand von Normwerten nicht geeignet ist (die Bewertung einer Leistung eines untersuchten Kindes kann anhand der im Manual der KITAP aufgeführten Normwerte vorgenommen werden). In der Studie von Drechsler et al. ist bemerkenswert, dass die prozentuale Häufigkeit der Kinder der ADHS-Gruppe mit einem Prozentrang ≤ 16 in keinem Parameter über 50% liegt. Die Unterschiede in der Verteilungshäufigkeit im unterdurchschnittlichen, durchschnittlichen und überdurchschnittlichen Bereich von ADHS- und Kontrollgruppe waren gering.

Obgleich die Ergebnisse o.g. Studie nahelegen, dass die KITAP keinen essentiellen Beitrag zur Verbesserung der Diagnosestellung von ADHS liefert, soll in vorliegender Studie für beide Testbatterien geprüft werden, ob von Auffälligkeiten in der TAP bzw. KITAP auf das Vorliegen von ADHS geschlossen werden kann. Ermittelt werden soll, ob ein Verfahren besser differenziert als das andere und ob einzelne Subtests der TAP bzw. KITAP zwischen Kindern mit und ohne ADHS besonders gut unterscheiden.

Wir stellen folgende Hypothesen auf:

1. Mit beiden Verfahren lässt sich die Diagnose ADHS absichern.
2. Das Design der KITAP mit der Einbettung der Aufgaben in eine Schlossgeschichte motiviert die Kinder stark, wodurch Aufmerksamkeitsprobleme schlechter entdeckt werden können als durch die TAP.

Methode

Bei dieser Untersuchung handelt es sich um eine quasi-experimentelle Querschnittstudie im Kontrollgruppendesign. Zur Überprüfung der Sensitivität und Spezifität der TAP und der KITAP wurden die Testleistungen von Kin-

den mit ADHS (klinisches Expertenurteil) mit einer Gruppe von unauffälligen Kontrollkindern verglichen. Jeder Proband absolvierte die TAP und die KITAP. Weiterhin wurde bei allen Probanden die kognitive Leistungsfähigkeit mit der Grundintelligenztest-Skala 2 (CFT 20) von Weiß (1998) gemessen. Dies ermöglichte die Kontrolle des Einflusses der Intelligenz auf die Leistungen in den Aufmerksamkeitsverfahren.

Beschreibung der Stichprobe

Die vorliegende Untersuchung wurde in drei klinisch-psychiatrischen Einrichtungen (Klinik für Kinder- und Jugendpsychiatrie und -psychotherapie am Ökumenischen Hainich Klinikum in Mühlhausen/Thüringen, Poliklinische Institutsambulanz des Psychologischen Instituts der Universität Göttingen, Kinderärztliche Praxis in Lehrte) durchgeführt. Es nahmen 24 Kinder mit der Diagnose ADHS und 24 Kontrollgruppenkinder im Alter von 8 bis 10 Jahren an der Studie teil. Sieben Kinder der Stichprobe hatten die Diagnose Aufmerksamkeitsstörung ohne Hyperaktivität (F98.8), 15 Kinder die Diagnose einfache Aktivitäts- und Aufmerksamkeitsstörung (F90.0) und ein weiteres Kind die Diagnose Hyperkinetische Störung des Sozialverhaltens (F90.1) erhalten.

Eine Kontrollstichprobe bildeten 24 unauffällige Kinder der 3. und 4. Klassen einer Grundschule in Göttingen. Um zu verhindern, dass Kinder mit Aufmerksamkeitsproblemen in die Kontrollstichprobe aufgenommen

werden, nahmen nur solche Schüler an der Studie teil, bei denen die Lehrer keine Auffälligkeiten angaben.

Die klinische Stichprobe unterschied sich nicht signifikant hinsichtlich des Alters, des Geschlechterverhältnisses und der Intelligenz von der Kontrollgruppe. Zur genauen Stichprobenbeschreibung siehe Tabelle 1.

Bei allen medizierten ADHS-Kindern wurden die Medikamente mindestens 48 Stunden vor der Untersuchung abgesetzt. Alle Eltern gaben ihr schriftliches Einverständnis für die Teilnahme an der Studie.

Material und Durchführung

Die Testbatterie zur Aufmerksamkeitsprüfung TAP

Die Testbatterie zur Aufmerksamkeitsprüfung TAP hat sich wie die meisten Verfahren zur Erfassung von Aufmerksamkeitsfunktionen zunächst an der neuropsychologischen Diagnostik orientiert (Zimmermann & Fimm, 1993). Deren Ziel bestand in der Untersuchung spezifischer Ausfälle bei oft multipler Schädigung der Patienten. Aufgrund dessen wurden für die TAP wenig komplexe Aufgaben ausgewählt, die umschriebene Teilleistungen überprüfen, ohne Beeinflussung durch sensorische oder motorische Ausfälle, Gedächtnis- oder Sprachstörungen. Diesen Ansprüchen können einfache Reaktionsaufgaben gerecht werden, bei denen der Proband auf leicht diskriminierbare, sprachfreie Reize mit einem Tastendruck reagiert. In der klinischen Praxis wird die TAP trotz ihrer Ent-

Tab. 1: Beschreibung der Stichprobe

	ADHS (n = 24)	Kontrollen (n = 24)	p
Jungen/Mädchen	19/5	14/10	n.s.
Alter (SD)	9.52 (0.98)	9.68 (0.64)	n.s.
IQ (SD)	102.0 (16.0)	107.3 (12.4)	n.s.

Anmerkungen: ADHS = Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung, p = .05, n.s. = nicht signifikant

wicklung im Erwachsenenbereich regelmäßig in der Aufmerksamkeitsdiagnostik von Kindern eingesetzt (Földényi, 1999). Das Gütekriterium Objektivität sehen die Testautoren selbst als eingeschränkt erfüllt an, da die Testanweisung auf dem Bildschirm allein zur Durchführung nicht ausreicht. Laut Manual soll der Testleiter die Kinder durch eine zusätzliche Geschichte (keine Angaben hierzu im Manual) motivieren. Hierdurch können Versuchsleitereffekte entstehen. Die Split-half-Reliabilität liegt für die Geschwindigkeitsmaße durchweg bei mindestens .97, bei den Fehlern und Auslassungen hingegen sind die Koeffizienten z.T. deutlich niedriger. Zimmermann und Fimm (2004) führen dies darauf zurück, dass bei „gesunden“ Kindern Fehlreaktionen und Auslassungen kaum vorkommen und sich dies zwangsweise auf die Koeffizienten auswirkt. Die Validität wurde für die TAP mittels Faktorenanalyse (Hauptkomponentenanalyse, Varimax-Rotation) bestimmt. Die Mehrdimensionalität der Testbatterie wurde anhand der Faktorenanalyse bestätigt.

Die klinische Validität der TAP bei Kindern mit ADHS konnte in einer Studie mit 20 Jungen mit ADHS und 20 Kontrollkindern von Földényi, Imhof und Steinhausen (2000) bestätigt werden. Sie setzten sechs Untertests der TAP ein. Alle Subtests trennten signifikant zwischen der ADHS- und der Kontrollgruppe.

Die TAP ist ein mehrdimensionales Verfahren bestehend aus zwölf Untertests verschiedener Komplexität. Alle Untertests verfügen über eine standardisierte Instruktion, die vor der Testdurchführung auf dem Bildschirm dargeboten wird. Dabei bekommen die Kinder die Anweisung möglichst schnell und fehlerfrei zu reagieren. Jeder Untertest beginnt mit einem Vortest bzw. einer Beispielaufgabe, um das Aufgabenverständnis sicherzustellen.

Die Testleistung eines Probanden kann bezüglich der Leistungsgüte und der Leistungsgeschwindigkeit beurteilt werden, für

die das Programm verschiedene Kennwerte liefert. Parameter der *Leistungsgüte* sind:

- die Anzahl richtiger Reaktionen (d.h. alle Reaktionen, die nicht falsch und innerhalb der zulässigen Zeitspanne erfolgen);
- die Anzahl falscher Reaktionen (d.h. Reaktionen auf nicht-kritische Reize bzw. Reaktionen, bei denen eine falsche Taste gedrückt wird, wenn ein kritischer Reiz erscheint; sie wird als Zeichen einer mangelnden Reaktionshemmung betrachtet);
- die Anzahl von Auslassungen (fehlende Reaktionen auf kritische Reize; sie stellt einen Indikator für Unaufmerksamkeit/gestörte selektive Aufmerksamkeit dar).

Wichtige Parameter der *Leistungsgeschwindigkeit* sind:

- der Median der Reaktionszeiten, der, wegen der häufig schiefen Verteilung von Reaktionszeiten, als angemessenster Kennwert für die mittlere Reaktionszeit gilt. Die Größe des Medians wird im Vergleich zum arithmetischen Mittel nicht von Ausreißern beeinflusst. Ein erhöhter Median gilt, sofern er sich über alle Verfahren nachweisen lässt, als Maß für eine allgemeine Verlangsamung.
- die Standardabweichung der Reaktionszeiten als Maß für die Regelmäßigkeit (Variabilität) der Antworten. Je geringer sie ist, desto konstanter wurde die Aufmerksamkeit aufrechterhalten.

In dieser Studie wurden sechs Subtests eingesetzt, von denen fünf eine Entsprechung in der KITAP haben. Der Untertest Inkompatibilität der TAP wurde trotz fehlender Entsprechung in der KITAP für die Studie ausgewählt, weil er in der faktorenanalytischen Berechnung der Autoren einen substanziellen Beitrag liefern konnte.

Folgende Subtests wurden eingesetzt:

Alertness. Mit dem Untertest werden die tonische und phasische *Alertness* erfasst. Die

phatische Alertness ist die Fähigkeit, in Erwartung eines kritischen Reizes die Aufmerksamkeit zu erhöhen. Sie wird durch die Verkürzung der Reaktionszeit auf einen angekündigten im Vergleich zu einem unangekündigten Reiz untersucht. Der visuelle Reiz besteht in einem weißen „x“, das kurz auf dem schwarzen Bildschirm aufleuchtet. Die einfache Reaktionszeit auf den visuellen Reiz ohne Warnsignal gibt Auskunft über die tonische Alertness bzw. eine allgemeine Reaktionsverlangsamung. Zusammenhang zu ADHS: Gemessen wird der Zustand allgemeiner Wachheit.

Go/Nogo. Mit diesem Subtest soll die spezifische Fähigkeit zur Unterdrückung einer Reaktion auf irrelevante Reize überprüft werden. Hierfür wird eine Abfolge visueller Reize (ein „+“ und ein „x“) präsentiert, wobei so schnell wie möglich bei nur einer Reizart zu reagieren ist. Auf neuropsychologischer Ebene werden also Bahnungs- („x“) und Hemmprozesse („+“) untersucht. Zusammenhang zu ADHS: misst die Fähigkeit zur Reaktionshemmung und die Ausprägung von Impulsivität.

Inkompatibilität. Der Subtest prüft die Fähigkeit zur Fokussierung der Aufmerksamkeit, d.h. inwieweit irrelevante und unter Umständen automatisch verarbeitete Reizaspekte unterdrückt werden können. Dem Probanden werden hierzu unterschiedlich gerichtete Pfeile dargeboten. Aufgabe ist die Reaktion auf die Pfeilrichtung unabhängig davon, auf welcher Seite des Bildschirms die Pfeile präsentiert werden. Zusammenhang zu ADHS: prüft die Fähigkeit zur Fokussierung der Aufmerksamkeit.

Geteilte Aufmerksamkeit. Bei der hier gestellten „dual-task“ Aufgabe sind gleichzeitig zwei Reizdarbietungen aus verschiedenen Informationskanälen (optisch/akustisch) zu bearbeiten. Der erste Teil der Aufgabe besteht darin, auf eine bestimmte Konstellation opti-

scher Reize (vier Kreuze, die ein kleines Quadrat bilden) per Tastendruck zu reagieren, während simultan auf Unregelmäßigkeiten in einer Tonfolge zu achten ist. Zusammenhang zu ADHS: Geprüft wird die Fähigkeit, mit konkurrierenden Informationen umzugehen und mehrere Dinge gleichzeitig zu bewältigen.

Reaktionswechsel. Alternierend soll auf runde und eckige Reizformen reagiert werden, die jeweils als Paar dargeboten werden. Zusammenhang zu ADHS: prüft die Fähigkeit, sich schnell auf einen neue Bedingung einzustellen.

Visuelles Scanning. Bei diesem Subtest soll der Proband eine 5x5-Matrix mit Quadraten, die jeweils nach einer Seite offen sind, nach einem kritischen Reiz absuchen. Es wird die Fähigkeit zum visuellen Abtasten des Gesichtsfeldes geprüft. Zusammenhang zu ADHS: Prüfung der Ausdauer und der Fähigkeit, eine bestimmte Strategie über einen längeren Zeitraum aufrechtzuerhalten.

Die Testbatterie zur Aufmerksamkeitsprüfung für Kinder KITAP

Zimmermann, Gondan und Fimm (2002) entwickelten die KITAP auf Basis der theoretischen Annahmen der TAP zur Erfassung verschiedener Aufmerksamkeitsfunktionen bei Kindern im Alter von sechs bis zehn Jahren. Aufgrund faktorenanalytischer Berechnungen wurden fünf Untertests der TAP für die Konstruktion der KITAP ausgewählt und aus theoretischen und praktischen Gründen durch drei weitere Untertests ergänzt. „Das Schloss der Geister“ bildet den grafischen und thematischen Rahmen, in den die acht Untertests integriert wurden. Die Durchführungsobjektivität ist vermindert (siehe Abschnitt Instruktionsprobleme). Die Testautoren geben im Manual die Splithalf-Reliabilität der Untertests mit .55 bis .97 an. Die niedrigen Reliabilitätskoeffizienten führen sie auf Artefakte durch geringe Anzahlen von Fehlern und

Auslassungen bei einigen Untertests zurück. Die Validität wurde für die KITAP mittels Faktorenanalyse (Hauptkomponentenanalyse, Varimax-Rotation) für Kinder von sechs bis sieben Jahren und für Kinder von 8 bis 10 Jahren bestimmt. Die Mehrdimensionalität der Testbatterie wurde anhand der Faktorenanalyse mit vier bzw. fünf Faktoren bestätigt.

Wie bei der TAP werden verschiedene Kriterien zur Beurteilung der Testleistung herangezogen. Hierzu zählen die Mediane der Reaktionszeiten, die Standardabweichung der Reaktionszeiten sowie die Anzahl an Fehlern und Auslassungen. Analog zur TAP kommen den Kriterien in den verschiedenen Untertests unterschiedliche Bedeutungen zu. Bei der KITAP werden nur für die Untertests Alertness, Flexibilität und Scanning normierte Standardabweichungen angegeben und deshalb in dieser Studie benutzt. In den anderen Subtests sind die SDs nicht normiert.

In dieser Studie wurden sechs Subtests eingesetzt. Neben den fünf zur TAP äquivalenten Untertests wird aus der KITAP zusätzlich der Untertest Daueraufmerksamkeit in die Studie miteinbezogen. Die Testautoren vermuten für diesen Untertest eine höhere prädiktive Validität als für Vigilanztests, weil sie den Anforderungen des Alltags eher entsprechen. Der Untertest Ablenkbarkeit wurde nicht in die Studie einbezogen, weil laut Testautoren bei der Interpretation nicht sicher davon ausgegangen werden kann, dass die Probanden die Aufgabe richtig verstanden haben.

Folgende Tests kamen zum Einsatz:

Alertness („Die Hexe“). Die Probanden sollen so schnell wie möglich eine Hexe, die im Fenster des Schlosses erscheint, mit einem Tastendruck „vertreiben“. Zusammenhang zu ADHS: Gemessen wird die allgemeine Wachheit.

Go/Nogo („Die Fledermaus“). Der Zielreiz ist eine Fledermaus, die mit einem Tastendruck vertrieben werden soll, damit die

Katze keine Angst haben muss. Fledermaus und Katze werden in zufälliger Reihenfolge dargeboten. Beim Erscheinen der Katze soll die Taste nicht gedrückt werden. Zusammenhang zu ADHS: misst die Fähigkeit zur Reaktionshemmung und Impulsivität.

Geteilte Aufmerksamkeit („Die Eulen“). Die optische Aufgabe besteht in der Beobachtung einer Eule, die beim Schließen ihrer Augen mit einem Tastendruck geweckt werden soll. Die akustische Aufgabe besteht im Verfolgen einer alternierenden Tonfolge, in der das gegenseitige Zurufen zweier Eulen zu hören ist. Ist eine Eule zweimal hintereinander zu hören, muss ebenfalls per Tastendruck reagiert werden. Zusammenhang zu ADHS: Geprüft wird die Fähigkeit, mit konkurrierenden Informationen umzugehen und mehrere Dinge gleichzeitig zu bewältigen.

Flexibilität („Das Haus der Drachen“). Die beiden konkurrierenden Reize sind ein grüner und ein blauer Drache, die so schnell wie möglich durch Tastendruck ins Haus gelassen werden sollen. Damit es keinen Streit gibt, müssen immer abwechselnd ein blauer und ein grüner Drache hineingelassen werden. Die Position der Drachen wechselt dabei zufällig. Zusammenhang zu ADHS: prüft die Fähigkeit, sich schnell auf eine neue Bedingung einzustellen.

Visuelles Scanning („Der Ausflug der Hexen“). Eine 5x5-Matrix beinhaltet 25 Hexen. Der Proband soll mit einem Tastendruck auf die linke Taste reagieren, wenn eine Hexe in die „falsche“ Richtung fliegt, bzw. auf die rechte Taste drücken, wenn alle Hexen in die gleiche Richtung fliegen. Zusammenhang zu ADHS: Prüfung der Ausdauer und der Fähigkeit, eine bestimmte Strategie über einen längeren Zeitraum aufrechtzuerhalten.

Daueraufmerksamkeit („Der Tanz der Geister“). Die Aufgabe verlangt, jeden Reiz mit dem nächsten zu vergleichen. Reize sind

hierbei Gespenster verschiedener Farben, die nacheinander in unterschiedlichen Fenstern einer Schlossruine erscheinen. Sobald die Farben von zwei aufeinanderfolgenden Geistern übereinstimmen, muss reagiert werden. Zusammenhang zu ADHS: prüft, ob ein Kind die Aufmerksamkeit über eine längere Zeitspanne kontinuierlich aufrechterhalten kann.

Verfahren zur Bestimmung der ADHS-Diagnose und der Intelligenz

Die Diagnosestellung der ADHS-Stichprobe erfolgte in den zuweisenden Einrichtungen durch klinische Interviews und die nachfolgend dargestellten Fragebögen.

Der **Fremdbeurteilungsbogen-Hyperkinetische Störungen (FBB-HKS)** ist Bestandteil des Diagnostik-Systems für psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV (DISYPS-KJ) von Döpfner und Lehmkuhl (2000). Hierbei erfolgt die klinische Beurteilung mittels Diagnose-Checklisten durch Eltern und Lehrer.

Der standardisierte **Elternfragebogen über das Verhalten von Kindern und Jugendlichen (CBCL 4–18)** ist die deutschsprachige Version der von Achenbach (1991) entwickelten Child Behavior Checklist (CBCL). Die CBCL wird als Screening-Verfahren zur Erfassung von Kompetenzen und psychischen Auffälligkeiten von Kindern und Jugendlichen eingesetzt. Der **Lehrerfragebogen über das Verhalten von Kindern und Jugendlichen (Teacher's Report Form, TRF)** wurde in Anlehnung an den Elternfragebogen entwickelt (Arbeitsgruppe Deutsche Child Behavior Checklist, 1993).

Die Intelligenz wurde mit dem **Grundintelligenztest Skala 2 (CFT 20)** von Weiß (1998) erhoben. Dies ist ein Verfahren zur Feststellung der grundlegenden geistigen Leistungsfähigkeit im Sinne des g-Faktors. Es soll die Intelligenz unabhängig von soziokulturellen, erziehungsspezifischen oder rassischen Einflüssen feststellen.

Die TAP und die KITAP wurden mit allen Kindern im Einzeltest durchgeführt, der CFT 20 bei der ADHS-Gruppe im Einzelsetting und bei den Kontrollkindern im Klassenverband.

Es wurden alle Kinder in die ADHS-Gruppe aufgenommen, die einen auffälligen Wert in der CBCL- und TRF-Syndromskala „Aufmerksamkeitsstörungen“ hatten, bei denen die DISYPS-Kriterien zutrafen und die gemäß des klinischen Interviews als auffällig galten.

Akzeptanz der Untersuchungsinstrumente

Alle Kinder, die an der Untersuchung teilgenommen haben, zeigten viel Spaß bei der Durchführung der KITAP, wohingegen die TAP häufig Frustration und Abwehr auslöste.

Die größere Durchführungsdauer der KITAP von ca. 60 Minuten stellte jedoch nicht nur für Kinder mit Aufmerksamkeitschwierigkeiten eine hohe Belastung dar. Die kürzere Durchführungszeit der TAP (ca. 45 Minuten) und die kurzen Untertests (mehr Abwechslung) sind positiv im Sinne von Testökonomie und Motivation zu bewerten.

Instruktionsprobleme

Sowohl die TAP als auch die KITAP weisen in der Durchführungsbeschreibung Mängel auf. Zum einen wird die relative Wichtigkeit von Schnelligkeit und Genauigkeit nur unzureichend dargelegt. Dem Versuchsleiter und den Probanden bleibt unklar, ob sie so genau wie möglich oder so schnell wie möglich zu Lasten der Genauigkeit arbeiten sollen. Dadurch sind hohe Fehleranteile nicht eindeutig interpretierbar (wenig sorgsames Arbeiten oder Bemühen um Schnelligkeit). Weiterhin ist unklar, wie der Diagnostiker auf instruktionswidriges Verhalten reagieren oder mit wenig motivierten Kindern umgehen soll. Für die KITAP liegen weder eine allgemeine Ein-

führung im Manual, noch standardisierte Instruktionen und eine festgelegte Reihenfolge der Untertests vor. Die Testautoren begründen die freie Instruktion bei den jeweiligen Untertests (und die damit verbundene Einschränkung der Durchführungsobjektivität) damit, dass so eine gute Motivation der Kinder erreicht werden könne.

Die gesamte Durchführungsdauer wird in der reinen Durchführungszeit für die hier ausgewählten sechs Untertests der KITAP im Manual mit 32 Minuten angegeben. Bei unserer Untersuchung wurden im Mittel etwa 50 bis 60 Minuten von den Kindern benötigt.

Probleme der Konzeption

Neuropsychologische Tests, die für Erwachsene konzipiert sind, bei Kindern einzusetzen, ohne die speziellen Anforderungen der Probanden in dieser Altersgruppe zu berücksichtigen, ist nicht unproblematisch. Die Aufgaben sind oft abstrakt und somit wenig kindgerecht. Sie fordern entsprechend große Anstrengung für Kinder, sodass sie schnell überfordert oder demotiviert sein können. Als Folge können sie möglicherweise nicht ihre tatsächliche Leistungsfähigkeit zeigen (Drechsler, Brandeis, Földényi, Imhof & Steinhausen, 2005). Bei der Interpretation der Testergebnisse gilt es zu beachten, dass die Verfahren bei Erwachsenen und Kindern möglicherweise nicht die gleichen kognitiven Prozesse ansprechen.

Der Einsatz von TAP und KITAP bei der Diagnose von ADHS ist mit weiteren Hindernissen verbunden. Da die Testautoren nicht den Anspruch erheben, spezifische Diagnoseinstrumente für ADHS entwickelt zu haben (Renner & Irblich, 2007), fehlen Angaben darüber, wann ein Testwert als auffällig zu bezeichnen ist. In der klinisch-psychiatrischen Praxis werden für die Beurteilung der Testleistung die Normen eines Verfahrens herangezogen. Wie weit jedoch die Leistungen unter-

halb des Durchschnitts liegen müssen, ist bislang umstritten.

Auswertung

Ziel der vorliegenden Untersuchung war die Beurteilung der Zuordnungsgüte zweier Testverfahren. Für die Bestimmung der Validität eines Tests können die Sensitivität (Fähigkeit, tatsächlich Kranke als krank zu erkennen) und Spezifität (Fähigkeit, tatsächlich Gesunde als gesund zu identifizieren) berechnet werden.

Ein hierfür geeignetes Vorgehen bietet ein Verfahren der Signalerkennungsanalyse. Dabei werden so genannte ROC-Kurven (Receiver Operating Characteristic Curves) erstellt. Diese schätzen die Güte der Zuordnung über das Zusammenspiel von Sensitivität und Spezifität der Testparameter. Die ROC-Analyse beinhaltet die graphische Darstellung von Treffern (richtig positive) und falschen Alarmen (falsch positive diagnostische Entscheidungen) in einer Kurve. Jeder Punkt einer Kurve entspricht einem Paar von Sensitivität und „1-Spezifität“. Der entscheidende Vorteil der ROC-Kurve gegenüber anderen Koeffizienten ist, dass die Fläche unter der Kurve (area under the curve – AUC) den diagnostischen Informationswert des Testverfahrens unabhängig von der Wahl des Trennwertes zur diagnostischen Zuordnung quantifiziert. Darüber hinaus ist die ROC-Analyse robust gegenüber Verletzungen der Normalverteilungs- und Varianzhomogenitätsannahme (McFall & Treat, 1999).

Ergebnisse

Die Betrachtung der Rohwerte beider Versuchsgruppen zeigt, dass die Kontrollgruppe in beiden Testverfahren mehrheitlich bessere Leistungen als die klinische Stichprobe erbracht hat (s. Tab. 2 und 3). Dies betrifft sowohl den Median und die Standardabweichung

Tab. 2: Ergebnisse der Kinder mit ADHS und Kontrollkinder in den TAP-Untertests und Gruppenvergleiche

	ADHS (n = 24) M (SD)	Kontrollen (n = 24) M (SD)	t	Effekt- größe d_{emp}
TAP				
Alertness o.W.				
Median RZ	305.41 (81.87)	246.11 (49.37)	-3.039**	0.96
SD RZ	90.13 (62.37)	38.22 (37.40)	-3.497***	1.06
Alertness m.W.				
Median RZ	274.42 (64.16)	219.08 (32.80)	-3.762***	1.12
SD RZ	76.67 (41.47)	42.50 (28.22)	-3.337**	1.05
Kennwert	.08 (.10)	.08 (.05)	.018	-0.06
Go/Nogo				
Median RZ	555.89 (135.65)	440.13 (87.91)	-3.508***	0.84
SD RZ	171.75 (75.08)	126.01 (39.49)	-2.642*	0.68
Fehler	8 (7)	6 (4)	-1.072	0.31
Auslassungen	2.23 (2.69)	2.61 (3.38)	.426	-0.31
Inkompatibilität				
Median RZ	576.76 (191.87)	502.75 (130.62)	-1.562	0.38
SD RZ	226.25 (141.52)	138.42 (44.79)	-2.898**	0.85
Fehler	17.2 (13.87)	12.81 (11.95)	-1.175	0.33
Geteilte Aufmerksamkeit				
Median RZ	1063.83 (213.14)	961.65 (132.39)	-1.995	0.46
SD RZ	396.76 (151.12)	361.18 (94.21)	-.979	0.13
Fehler	5 (7)	2 (2)	-1.710	0.49
Auslassungen	12.17 (6.54)	7.27 (2.72)	-3.383***	0.91
Reaktionswechsel				
Median RZ	1517.74 (737.03)	854.67 (370.81)	-3.937***	1.50
SD RZ	556.53 (371.83)	218.82 (213.27)	-3.860***	1.33
Fehler	7 (5)	3 (2)	-3.671***	1.06
Visuelles Scanning – nicht kritisch				
Median RZ	6862.31 (3280.56)	7322.81 (1876.66)	.597	-0.13
SD RZ	3244.87 (2266.37)	1841.46 (607.45)	-2.930**	0.85
Fehler	1 (2)	0 (0)	-2.837**	0.82
Visuelles Scanning – kritisch				
Auslassungen	7 (4)	3 (3)	-3.534***	1.02
Zeilenkorrelation	.41 (.38)	.76 (.22)	3.885***	-1.2

Anmerkungen: * $p < .05$, ** $p < .01$, *** $p < .001$; RZ = Reaktionszeit, SD = Standardabweichung, M = Mittelwert. Die SD in der linken Spalte (z.B. SD RZ (Standardabweichung der Reaktionszeiten) beim Subtest Go/Nogo) sind eigenständige Parameter der Untertests. Diese Werte wurden dann über die Kinder der ADHS-Gruppe bzw. Kontrollgruppe gemittelt. Angegeben wird jeweils der Mittelwert (M) und die Standardabweichung (SD) der Gruppe zu diesem Parameter

Tab. 3: Ergebnisse der Kinder mit ADHS und Kontrollkinder in den KITAP-Untertests und Gruppenvergleiche

	ADHS (n = 24) M (SD)	Kontrollen (n = 24) M (SD)	t	Effekt- größe d_{emp}
KITAP Alertness				
Median RZ	338.04 (69.81)	291.33 (49.09)	-2.681**	0.77
SD RZ	78.54 (43.85)	52.54 (25.34)	-2.515*	0.73
Go/Nogo				
Median RZ	477.50 (68.01)	463.92 (49.73)	-.790	0.23
Fehler	3.08 (2.19)	2.54 (2.92)	-.244	0.07
Auslassungen	.75 (1.03)	.13 (.34)	-2.819**	0.81
Geteilte Aufmerksamkeit				
Median RZ	780.33 (125.54)	721.38 (99.76)	-1.801	0.52
Fehler	5.46 (8.35)	5.10 (3.42)	-1.022	0.30
Auslassungen	4.58 (5.03)	2.13 (1.57)	-2.286*	0.66
Flexibilität				
Median RZ	1137.58 (320.30)	848.54 (194.45)	-3.779***	1.09
SD RZ	514.00 (398.74)	265.92 (116.87)	-2.925**	0.84
Fehler	5.04 (3.87)	2.83 (2.58)	-2.324*	0.67
Visuelles Scanning				
Median RZ	7916.92 (3135.95)	8126.13 (1746.2)	.286	-0.08
SD RZ	3072.21 (1642.21)	2310.50 (824.03)	-2.031*	0.59
Fehler	4.04 (7.33)	2.30 (1.58)	-1.568	0.45
Auslassungen	11.46 (9.00)	8.00 (7.25)	-1.466	0.42
Daueraufmerksamkeit				
Median RZ	758.04 (188.26)	641.58 (121.10)	-2.549*	0.74
Fehler	15.79 (27.54)	5.25 (4.59)	-1.850	0.53
Auslassungen	11.83 (9.34)	5.33 (4.48)	-3.074**	0.89

Anmerkung: * $p < .05$, ** $p < .01$, *** $p < .001$. Die SD in der linken Spalte (z.B. SD RZ (Standardabweichung der Reaktionszeiten) beim Subtest Flexibilität) sind eigenständige Parameter der Untertests. Diese Werte wurden dann über die Kinder der ADHS-Gruppe bzw. Kontrollgruppe gemittelt. Angegeben wird jeweils der Mittelwert (M) und die Standardabweichung (SD) der Gruppe zu diesem Parameter

chung der Reaktionszeit (RZ) als auch die Parameter der Leistungsgüte. Diese Unterschiede wurden für jeden Parameter in einem t-Test für unabhängige Stichproben auf Signifikanz geprüft.

Wie aus Tabelle 2 ersichtlich wird, unterschieden sich die Kontrollgruppe und die klinische Stichprobe in 15 der 24 Parameter der TAP signifikant in ihren Leistungen. Diese sig-

nifikanten Unterschiede verteilten sich auf alle sechs Untertests und betrafen sowohl Parameter der Leistungsgüte als auch der Leistungsgeschwindigkeit. Die Effektgrößen waren hierbei überwiegend größer als 0.8 und sind damit als groß zu bezeichnen (Cohen, 1988).

Die Auswertung der t-Tests für die Parameter der KITAP zeigte, dass sich die beiden

Versuchsgruppen auch in diesem Verfahren in ihren Leistungen unterschieden (vgl. Tabelle 3). Bei der Hälfte der insgesamt 18 Parameter war ein signifikanter Leistungsunterschied festzustellen. Diese Parameter verteilten sich auf alle Untertests bis auf das Visuelle Scanning. Die Effektgrößen waren geringer als bei der TAP. Lediglich vier der 18 waren große Effekte, acht waren höher als 0.5 und damit als mittelgroß einzuschätzen (Cohen, 1988).

Die signifikanten Leistungsunterschiede und die Höhe der Effektgrößen lassen vermuten, dass beide Verfahren Kinder mit und ohne ADHS differenzieren können. Ein Vergleich der Effektgrößen in beiden Verfahren legt jedoch nahe, dass die TAP besser differenziert als die KITAP. In den ROC-Analysen der TAP zeigten sich 15 von 24 Parametern als signifikant vom Wert der Zufallsgeraden ($AUC = .5$) verschieden. Die AUCs dieser Parameter der Leistungsgüte und -geschwindigkeit lagen zwischen .694 und .806. Dazu gehörten alle drei Parameter des Untertests Reaktionswechsel, alle Parameter des Untertests Alertness bis auf die physische Alertness, alle Parameter des Untertests Visuelles Scanning bis auf den Median der RZ bei den nicht-kritischen Trials, der Median und die Standardabweichung der RZ im Untertest Go/Nogo, die Anzahl an Auslassungen im Untertest Geteilte Aufmerksamkeit und die Standardabweichung der RZ im Untertest Inkompatibilität.

Die ROC-Analysen der KITAP ergaben einen signifikanten Unterschied von der Zufallsgeraden bei 9 der 18 Parameter. Es handelte sich wie bei der TAP um Parameter der Leistungsgeschwindigkeit und -güte, deren AUCs zwischen .668 und .826 lagen. Dies waren alle drei Parameter des Untertests Flexibilität und des Untertests Daueraufmerksamkeit, der Median und die Standardabweichung der RZ im Untertest Alertness sowie die Anzahl an Auslassungen im Untertest Go/Nogo.

Zusammenfassend schienen die Untertests Flexibilität, Daueraufmerksamkeit und Alertness demnach besonders geeignet zu sein, um zu korrekten diagnostischen Entscheidungen bei ADHS zu führen.

Im Vergleich von TAP und KITAP zeigte sich die TAP mit 15 von 24 signifikanten Parametern als überlegenes Verfahren in der Zuordnungsgüte. Die Größen der AUCs waren jedoch vergleichbar (TAP: $AUC = .694 - .806$; KITAP: $AUC = .668 - .826$). Die 15 in TAP und KITAP äquivalenten Parameter wurden auf signifikante Unterschiede ihrer AUCs überprüft. Lediglich der Median der RZ im Untertest Go/Nogo wies einen signifikanten Unterschied in der Größe seiner AUC zwischen beiden Testverfahren auf. Mit .765 war die AUC in der TAP eindeutig größer als die AUC von .594 in der KITAP. Die weiteren AUCs äquivalenter Parameter unterschieden sich nicht signifikant. Insgesamt ließ sich festhalten, dass es nur eine tendenzielle, statistisch nicht bedeutsame Überlegenheit der TAP gab.

Die Anwendung der binären logistischen Regression zur Identifikation besonders geeigneter Vorhersageparameter der TAP führte zu einem Modell, das zwei Parameter umfasste. Die zwei bedeutsamen Prädiktoren für die Diagnose ADHS waren die Standardabweichung der RZ im Untertest Alertness (Trials ohne Warnton) und der Median der RZ im Untertest Reaktionswechsel. Durch die Hinzunahme der zwei Prädiktoren verbesserte sich das Modell signifikant ($\chi^2 = 21.599$, $df = 2$, $p = .000$). Das Modell führte zu 75% korrekten Klassifikationen und erklärte 32.5% der Varianz der Diagnose ADHS. Kein weiterer Prädiktor konnte signifikant zur Vorhersage beitragen.

Die logistische Regression der KITAP-Daten führte zu einem Modell, das neben der Konstante nur eine weitere Variable enthielt. Der bedeutsame Prädiktor für die Diagnose ADHS war der Median der RZ im Untertest Flexibilität. Durch die Hinzunahme dieses Prädiktors verbesserte sich das Modell signifi-

kant ($\chi^2 = 14.382$, $df = 1$, $p = .000$). Dieses Modell führte zu 72.9% korrekten Klassifikationen und erklärte 21.6% der Varianz der Diagnose ADHS. Kein weiterer Prädiktor konnte signifikant zur Vorhersage beitragen.

Die Ergebnisse der binären logistischen Regressionen zeigten, dass mithilfe des Regressionsmodells der TAP tendenziell mehr korrekte Klassifikationen möglich waren als bei der KITAP. Zudem war die Varianzaufklärung der Diagnose bei der TAP größer als bei der KITAP. Bei der Verwendung der KITAP würde es nach dem Regressionsmodell 16.7% falsch negative Diagnosen geben, mit der TAP 18.8%. Die TAP führte nur in 6.3% der Fälle zu einer falsch positiven Diagnose (KITAP: 10.4%).

Der Einschluss der 15 zur TAP äquivalenten Parameter der KITAP in einen zweiten Block bei der binären logistischen Regression führte zu keiner weiteren Varianzaufklärung. Wurde hingegen die TAP zusätzlich zur KITAP eingesetzt, verbesserte sich das Modell signifikant. Das Modell umfasste zwei Prädiktoren, den Median der RZ des Untertests Flexibilität der KITAP sowie im zweiten Block aus der TAP die Fehler im Untertest Reakti-

onswechsel. Durch die Hinzunahme des Prädiktors aus der TAP im zweiten Block verbesserte sich das Modell im Vergleich zum ersten Block signifikant ($\chi^2 = 10.100$, $df = 1$, $p = .001$). Das gesamte Modell führte zu 81.3% richtigen Klassifikationen und erklärte 36.8% der Varianz der Diagnose ADHS, wohingegen das Modell nach dem ersten Block mit den Parametern der KITAP nur 21,6% der Varianz aufklärte.

Tabelle 4 gibt einen Gesamtüberblick über die Ergebnisse der statistischen Analysen der Studie.

Diskussion

Sowohl die ROC-Analyse als auch die binäre logistische Regression bestätigten die Sensitivität und Spezifität beider untersuchter Verfahren. Die statistischen Analysen konnten keine signifikanten Unterschiede zwischen beiden Testverfahren nachweisen, obwohl bei rein deskriptiver Betrachtung die TAP bezüglich der Diagnose von ADHS tendenziell bessere Vorhersagen zu machen scheint als die KITAP. Obwohl die KITAP die Aufgaben

Tab. 4: Zusammenfassender Vergleich der Analysen der TAP und KITAP

	TAP	KITAP
Anzahl der erhobenen Parameter	24	18
T-Test: Parameter mit signifikantem Gruppenunterschied	15 62.5%	9 50%
Parameter mit Effektgröße > 0.8	14 58.3%	4 22.2%
ROC-Analyse		
Parameter mit AUC signifikant > .5 der 15 äquivalenten Parametern	11 61.1%	6 33.3%
Parameter mit AUC signifikant > .5 aller äquivalenter Parameter	15 62.5%	9 50%
Binäre logistische Regression		
Korrekte Klassifikation	75%	72.9%
Varianzaufklärung	32.5%	21.6%

in eine Gespenstergeschichte integriert und die Kinder gern mitmachten, ist die Sensitivität bzw. Spezifität des Verfahrens gegenüber der TAP nicht angestiegen, sondern blieb hinter ihr zurück. Eine mögliche Erklärung der relativ guten Leistungen der ADHS-Kinder in der KITAP könnte in der ansprechenden, kindgerechten Darbietung der Testaufgaben liegen, welche zu einer sehr hohen Motivation der Kinder geführt haben könnte.

Den Parametern (Median und Standardabweichungen der Reaktionszeiten, Anzahl der Fehler und Auslassungen) kommen laut Testautoren in den verschiedenen Untertests unterschiedliche Bedeutungen zu. Dennoch wiesen in dieser Studie nicht alle von den Testautoren als besonders relevant für die Diagnose von ADHS eingestuft Parameter auch tatsächlich eine hohe Diskriminationsfähigkeit auf, obwohl sie theoretisch durchaus begründbar sind. Es ist bspw. inhaltlich nachvollziehbar, dass eine hohe Anzahl an Fehlern im Untertest Go/Nogo auf impulsives Verhalten deutet und damit als Hinweis auf das Vorliegen von ADHS gewertet werden kann. Dennoch wies dieser Parameter in beiden Verfahren keinen hohen diagnostischen Informationswert auf. Dies liegt eventuell daran, dass eben auch die Kinder der Kontrollgruppe Schwierigkeiten hatten, eine dominante Reaktion zu unterdrücken.

Von den 15 Parametern der TAP, denen ihre AUCs eine hohe Zuordnungsgüte in dieser Studie akkreditierten, sind nach Aussagen der Testautoren nur neun diagnostisch besonders aussagekräftige Werte. In der KITAP hingegen sind laut Testautoren nur fünf der neun Parameter mit hoher Zuordnungsgüte auch diagnostisch besonders wertvoll. Würde man den Hinweisen des Manuals folgen, würde die Diskriminationsfähigkeit der gesamten Verfahren demnach nicht vollständig ausgeschöpft.

Es muss vor allem darauf hingewiesen werden, dass in beiden Testverfahren Parameter, die von den Testautoren als bedeutsame Indikatoren bei der Leistungsbewertung

angesehen werden, AUCs aufwiesen, die *nicht* signifikant von .5 verschieden waren. In der TAP und KITAP gehörten hierzu die Anzahl der Fehler im Untertest Go/Nogo und der Median der RZ im Untertest Visuelles Scanning. In der KITAP waren dies weiterhin die Anzahl an Auslassungen im Untertest Geteilte Aufmerksamkeit, der Median der RZ im Untertest Go/Nogo und die Auslassungen, Fehler und Standardabweichung der RZ im Untertest Visuelles Scanning. Den Empfehlungen der Autoren hinsichtlich der relevanten Parameter für die ADHS-Diagnose ist daher nur eingeschränkt zuzustimmen. Weniger theoretische Überlegungen als vielmehr empirische Studien der vorliegenden Art sind notwendig, um die tatsächliche Vorhersagekraft der einzelnen Parameter feststellen und Anhaltspunkte für die Interpretation der Testergebnisse liefern zu können.

Die Zielsetzung bei der Entwicklung der KITAP (kindgerechte Diagnostik) war eine viel versprechende Möglichkeit, entwicklungspsychologischen Aspekten Rechnung zu tragen (Drechsler et al., 2005). Leider fehlt jedoch beiden Verfahren der entwicklungspsychologische Bezug, z.B. zu den Faktoren Temperament, kognitiver Stil, Persönlichkeitsvariablen (z.B. Gewissenhaftigkeit) und Ausdauer. In beiden Verfahren wird die auditive Aufmerksamkeit trotz ihrer offensichtlichen Bedeutung für den schulischen Alltag zu wenig berücksichtigt.

Es ist nicht ungewöhnlich, dass ein Kind mit ADHS während der Exploration und Untersuchung im Testsetting das Geschehen ruhig und aufmerksam verfolgt und keine klassischen Symptome zeigt. Daher und aufgrund der zahlreichen Schwierigkeiten bei der Diagnose von ADHS, gewinnt die Forderung nach geeigneten Testbatterien zur Aufmerksamkeitsprüfung im diagnostischen Rahmen an Bedeutung. Mit klinischen Testverfahren können objektive standardisierte Daten erhoben und mit einer Normalpopulation verglichen werden. Jedoch ist die „künstliche“ testpsychologische Untersuchungssitua-

tion wenig mit realen Alltagsanforderungen vergleichbar.

Die Befunde dieser Studie unterstützen die Annahme, dass computergestützte Verfahren einen Beitrag im diagnostischen Prozess bei fraglicher ADHS liefern können. Um genauere Aussagen bzgl. der Diagnoseverbesserung für die klinische Praxis machen zu können, sollten auf der Basis einer größeren Stichprobe Cut-off-Werte für TAP und KITAP ermittelt und hinsichtlich deren Sensitivität und Spezifität überprüft werden. Bei der TAP sind es den Ergebnissen dieser Studie folgend insbesondere die Untertests Alertness, Go/Nogo, Reaktionswechsel und Visuelles Scanning, die zu einer guten Vorhersage führen könnten.

Fazit: Es sollten in der klinischen Praxis bei der Diagnostik von ADHS immer Interviews und klinische Fragebögen eingesetzt werden. Aufgrund der nicht zufriedenstellenden diagnostischen Zuordnung und fehlender Cut-off-Werte sollten computergestützte Aufmerksamkeitstests allenfalls ergänzend eingesetzt werden.

Die Entwicklung geeigneter, gut differenzierender und objektiver Verfahren für Kinder und Jugendliche auf fundierter theoretischer Basis bleibt eine äußerst wichtige Forschungsaufgabe der nächsten Jahre, besonders bei einem so häufig diagnostizierten und behandelten Störungsbild wie ADHS.

Literatur

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington. University of Vermont, Department of Psychiatry.
- Arbeitsgruppe Deutsche Child Behavior Checklist (1993). *Lehrerfragebogen über das Verhalten von Kindern und Jugendlichen; deutsche Bearbeitung der Teacher's Report Form der Child Behavior Checklist (TRF)*. Einführung und Anleitung zur Handauswertung, bearbeitet von Döpfner, M. & Melchers, P. Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121, 65- 94.
- Brickenkamp, R. (2002). *Test d2, Aufmerksamkeits-Belastungs-Test*. 9. Auflage. Göttingen: Hogrefe.
- Brähler, E., Holling, H., Leutner, D. & Petermann, F. (2002). *Brickenkamp Handbuch psychologischer und pädagogischer Tests*, Band 1. 3. Auflage. Göttingen: Hogrefe.
- Büttner, G. & Schmidt-Atzert, L. (Hrsg.) (2004). *Diagnostik von Konzentration und Aufmerksamkeit*. Jahrbuch der pädagogisch-psychologischen Diagnostik. Neue Folge, Band 3. Göttingen: Hogrefe.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dt. Ges. f. Kinder- und Jugendpsychiatrie und Psychotherapie u.a. (Hrsg.) (2007). *Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter*. Deutscher Ärzte Verlag, 3. überarbeitete Auflage, S. 239-254.
- Döpfner, M., Frölich, J. & Lehmkuhl, G. (2000). *Hyperkinetische Störungen*. Göttingen: Hogrefe.
- Döpfner, M., Banaschewski, T. & Sonuga-Barke, E.J.S. (2008). *Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS)*. In F. Petermann (Hrsg.), *Lehrbuch der klinischen Kinderpsychologie* (6. Aufl., S. 257-276). Hogrefe Verlag Göttingen.
- Döpfner, M., Breuer, H.D., Wille, N., Erhart, M., Ravens-Sieberer, U., Wittchen, H.U. & Bella, S.G. (2008). How often do children meet ICD-10/DSM-IV criteria of attention deficit/hyperactivity disorder (ADHD) and hyperkinetic disorder (HD)? Parent-based prevalence rates in a national sample – results of the BELLA study. *European Child and Adolescent Psychiatry* (Supplement).
- Döpfner, M. & Lehmkuhl, G. (2000). *DISYPS-KJ: Diagnostiksystem für psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV*. Bern: Huber.
- Douglas, V. I. & Parry, P. A. (1994). Effects of reward and nonreward on frustration and attention in attention deficit disorder. *Journal of Abnormal Child Psychology*, 22, 281-302.
- Drechsler, R., Brandeis, D., Földényi, M., Imhof, K. & Steinhausen, H.-C. (2005). The course of neuropsychological functions in children with attention deficit hyperactivity disorder from

- late childhood to early adolescence. *Journal of Child Psychology and Psychiatry*, 46, 824-836.
- Drechsler, R., Rizzo, P. & Steinhausen, H.-C. (2009). Zur klinischen Validität einer computergestützten Aufmerksamkeits-testbatterie für Kinder (KITAP) bei 7- bis 10-jährigen Kindern mit ADHS. *Kindheit und Entwicklung*, 18 (3), 153-161.
- Eisert, H. G. (1995). Kognitiv-verhaltenstherapeutische Behandlung hyperaktiver Kinder. In H.-C. Steinhausen (Hrsg.), *Hyperkinetische Störungen im Kindes- und Jugendalter* (S. 164-177). Stuttgart: Kohlhammer.
- Földényi, M. (1999). Zur Anwendung der computerunterstützten Testbatterie zur Aufmerksamkeitsprüfung (TAP) und weiterer computerunterstützter Aufmerksamkeits-tests bei Kindern. Dissertation, Universität Zürich.
- 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.
- Gawrilow, C., Schmitt, K. & Rauch, W. (2011). Kognitive Kontrolle und Selbstregulation bei Kindern mit ADHS. *Kindheit und Entwicklung*, 20 (1), 41-48.
- Hampel, P., Petermann, F. & Desmann, C. (2009). Exekutive Funktionen bei Jungen mit Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung im Kindesalter. *Kindheit und Entwicklung*, 18 (3), 144-152.
- Kleber, E., Kleber, G. & Hans, O. (1999). Differentieller Leistungstest – KG: DL-KG; Test zur Erfassung des Leistungsverhaltens bei konzentrierter Tätigkeit im Grundschulalter. 2. korr. Aufl. Göttingen: Hogrefe.
- Kurth, E. & Büttner, G. (1999). Testreihe zur Prüfung der Konzentrationsfähigkeit (TPK). 2. Auflage. Göttingen: Hogrefe.
- McFall, R. M. & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, 50, 215-241.
- Renner, G. & Irblich, D. (2007). Testinformation: Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP) von Peter Zimmermann, Matthias Gondan und Bruno Fimm (2002). *Diagnostica*, 53, 48-55.
- Schmidt, S. & Petermann, F. (2008). Entwicklungspsychopathologie der ADHS. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 56 (4), 265-274.
- Shaw, R., Grayson, A. & Lewis, V. (2007). Inhibition, ADHD, and computer games: The inhibitory performance of children with ADHD on computerized tasks and games. *Journal of Attention Disorders*, 8, 160-168.
- Slusarek, M., Velling, S., Bunk, D. & Eggers, C. (2001). Motivational effects on inhibitory control in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 355-363.
- Stadler, C. & Döpfner, M. (2004). Aufmerksamkeitsdiagnostik bei psychischen Störungen im Kindes- und Jugendalter. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit*. Tests und Trends, N.F. Band 3 (S. 249-270). Göttingen: Hogrefe.
- Weiß, R. H. (1998). Grundintelligenztest Skala 2 (CFT 20). Göttingen: Hogrefe.
- Zimmermann, P. & Fimm, B. (1993). Testbatterie zur Aufmerksamkeitsprüfung TAP. Herzogenrath: Psytest.
- Zimmermann, P. & Fimm, B. (2004). Die Testbatterie zur Aufmerksamkeitsprüfung TAP. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit*. Jahrbuch der pädagogisch-psychologischen Diagnostik. Neue Folge, Band 3. S. 177-202. Göttingen: Hogrefe.
- Zimmermann, P., Gondan, M. & Fimm, B. (2002). Testbatterie zur Aufmerksamkeitsprüfung für Kinder KITAP. Herzogenrath: Psytest.

Anschriften der Autoren:

DR. THOMAS DREISÖRNER
 Institut für Psychologie
 Goethe-Universität Frankfurt am Main
 Senckenberganlage 15
 60325 Frankfurt a.M.
 T.Dreisoerner@paed.psych.uni-frankfurt.de

JANINE GEORGIADIS
 Neisser Straße 11
 31139 Hildesheim
 Jannen23@gmx.net