

Originalartikel

Skala zur Erfassung von Aufmerksamkeitsdefiziten (SEA)

Erste psychometrische Evaluation mit einer Rasch-Analyse

Eftychia Volz-Sidiropoulou¹, Maren Böcker¹, Hendrik Niemann²,
Caroline Privou³, Petra Zimmermann³ und Siegfried Gauggel¹

¹Institut für Medizinische Psychologie und Medizinische Soziologie, Universitätsklinik der RWTH Aachen, ²Neurologisches Rehabilitationszentrum Leipzig, Bennewitz, ³Reha-Zentrum Reichshof, Fachklinik für Neurologie, Pneumologie und Onkologie, Reichshof-Eckenhagen

Zusammenfassung. Die Skala zur Erfassung von Aufmerksamkeitsproblemen (SEA) ist ein Selbst- und Fremdbeurteilungsverfahren zur Erfassung von Beeinträchtigungen der Aufmerksamkeit. Zur Untersuchung der psychometrischen Eigenschaften der SEA wurden 202 Patienten mit neurologischen Erkrankungen gebeten, die SEA auszufüllen. Eine Fremdbeurteilung erfolgte durch Neuropsychologen/Ergotherapeuten bzw. eine psychologisch-technische Assistentin. Die Rasch-Analyse ergab, dass mit den SEA-Items eindimensional Aufmerksamkeit erfasst werden können. In der Selbstbeurteilung form zeichneten sich die Items durch hohe Differenzierungskraft und Zuverlässigkeit aus (Itemtrennung/Reliabilität = 4.10/.94) und konnten gut zwischen mindestens drei Beeinträchtigungsgruppen unterscheiden (Personentrennung/Reliabilität = 2.86/.89). Die entsprechenden psychometrischen Kennwerte fielen bei der Fremdbeurteilung noch höher aus (Personentrennung/Rel. = 3.40/.92, Itemtrennung/Rel. = 5.37/.97). Insgesamt war aber das Item-Set für die untersuchte Patientenstichprobe zu leicht, so dass Patienten mit geringen Beeinträchtigungen weniger gut differenziert werden können. Die Übereinstimmung zwischen Selbst- und Fremdurteilen war erwartungsgemäß eher gering (Intraklassen-Korrelationskoeffizient, ICC = .34). Die ersten psychometrischen Eigenschaften der SEA sind ermutigend, allerdings müssen weitere Untersuchungen und Optimierungen (z. B. Zufügen einiger schwieriger Items) folgen.

Schlüsselwörter: Aufmerksamkeit, Diagnostik, Aufmerksamkeitsfragebogen, Rasch-Analyse, Selbstbeurteilung

The Attention Deficits Questionnaire (ADQ): Initial Psychometric Assessment

Abstract. The objective of this study was to assess the psychometric properties of the Attention Deficits Questionnaire (ADQ), a new Rating scale for the assessment of attention deficits. Two-hundred and two patients with neurological disorders (e.g., traumatic brain injury, stroke) filled out the self rating version of ADQ (ADQ-S). The proxy rating version of the ADQ (ADQ-P) was provided from psychologists, occupational therapists or psychological technical assistants. Rasch analysis was applied and indicated that the ADQ measures a single dimension and demonstrated good person separation (ADQ-S: 2.86/ADQ-P: 3.40) and reliability (ADQ-S: .89/ADQ-P: .92) as well as a very good item separation (ADQ-S: 4.10/ADQ-P: 5.37) and reliability (ADQ-S: .94/ADQ-P: .97). However, as a whole the items of the ADQ were relatively easy for the included patients. In order to enhance the difficulty of the scale, so that it becomes more appropriate to assess a broader range of attention problems, new items have to be added. Intraclass correlation coefficient between self- and proxy- reports was low (ICC = .34). In summary, initial psychometric properties suggested that the ADQ is a promising rating scale for the assessment of attention deficits.

Keywords: attention, attention deficits, assessment, rating scale, Rasch analysis

Aufmerksamkeitsstörungen stellen eine der häufigsten kognitiven Störungen nach einer erworbenen Hirnschädigung dar. Die unterschiedlichen Arten von Aufmerksamkeitsdefiziten sowie ihre Koexistenz mit anderen Störungen (z. B. Gedächtnisstörung) stellen hohe Anforderungen an den diagnostischen Prozess. Es bedarf daher psychomet-

risch gut abgesicherter Verfahren, um diesen Anforderungen zu genügen (Niemann & Gauggel, 2005; Sturm, 2005).

Die Diagnostik von Aufmerksamkeitsstörungen erfolgt in der Regel durch «objektive» Tests, die im Papier-Bleistift-Format (z. B. Zahlenverbindungstest, d2-Test, Konzentrationsverlaufs-Test) oder computergestützt (Auf-

merksamkeits-Netzwerk-Test (ANT), Testbatterie zur Aufmerksamkeitsprüfung (TAP), Wiener Determinationsgerät) durchgeführt werden.

Bei einer aktuellen Befragung nordamerikanischer Neuropsychologen haben Rabin und Kollegen festgestellt, dass bei einer solchen Aufmerksamkeitsdiagnostik am häufigsten der «Trail Making Test» (Lewis & Rennick, 1979), die «Digit Span» (Wechsler 1997a,b), der «Paced Auditory Serial Addition Test» (PASAT; Gronwall, 1977), der «Stroop Test» (Jensen & Rohwer, 1966) und der «Continuous Performance Test» (CPT/CPT-II; Conners & Multi-Health Systems Staff, 1995) eingesetzt werden (Rabin, Barr & Burton, 2005).

Obwohl Testverfahren zur Routinediagnostik gehören und in der klinischen Praxis unverzichtbar sind, liefern diese Verfahren nur ein eingeschränktes Bild der Störung. Nicht berücksichtigt werden die subjektiv erlebten Defizite sowie die Auswirkungen der Aufmerksamkeitsstörung auf den Alltag des betroffenen Patienten (Kinsella, 1998). Neben der Feststellung von Aufmerksamkeitsstörungen durch «objektive» Tests, ist daher die Erhebung der subjektiv erlebten bzw. der im Alltag beobachtbaren Beeinträchtigungen unabdingbar; denn dadurch wird eine Verbindung zwischen der Test- und der Alltagssituation hergestellt (Sturm, 2005). Nur wenige Verfahren stehen jedoch bislang für die klinische Praxis zur Verfügung, mit denen eine differenzierte Diagnostik dieser Aspekte erfolgen kann.

Im angloamerikanischen Raum werden Aufmerksamkeitsprobleme teilweise mit Interviews (z. B. Severe Traumatic Brain Injury Outcome Interview (STBIOI) von Brown et al., 2001) oder Breitband-Skalen (z. B. Key Behaviors Change Inventory (KBCI) von Kolitz et al., 2003) anhand einiger weniger Items dokumentiert. Solche Verfahren erfassen allerdings nicht nur Aufmerksamkeitsstörungen, sondern auch andere kognitive Defizite.

Neben diesen Verfahren finden sich auch eigens zur Erfassung von Aufmerksamkeitsdefiziten konzipierte Fragebögen. Diese Fragebögen erfassen Störungen einzelner oder mehrerer Aspekte der Aufmerksamkeit. Beispielsweise ist der «Divided Attention Questionnaire» (DAQ, Tun & Wingsfield, 1995) ein Selbstbeurteilungs-Fragebogen, mit dem primär Störungen der geteilten Aufmerksamkeit gemessen werden. Bei der «Rating Scale of Attentional Behaviour» (RSAB, Ponsford & Kinsella, 1991) und die «Moss Attention Rating Scale» (MARS, Hart et al., 2006; Whyte et al., 2003) handelt es sich dagegen um Fremdbeurteilungs-Fragebögen, mit denen Störungen in mehreren Aufmerksamkeitskomponenten erfasst werden sollen. Die RSAB basiert auf Posner's und Peterson's Aufmerksamkeitstheorie (1990) und erfasst anhand von 14 Items Beeinträchtigungen im Bereich der Alertness, der selektiven und der Daueraufmerksamkeit (Ponsford & Kinsella, 1991). Die MARS geht auf Mirsky's Theorie der Aufmerksamkeit zurück und soll mit der Kurzversion von 22 Items Aspekte wie «Unruhe/Ablenkbarkeit», «Daueraufmerksamkeit» und «Handlungsinitiierung» erfassen.

Die theoretische und psychometrische Fundierung der

beiden Fragebögen ist sehr unterschiedlich und nicht immer befriedigend. Bei der RSAB ist der Validitätsbereich eingeschränkt und bezieht sich nur auf die Gruppe der schwerbeeinträchtigten Hirngeschädigten in der Rehabilitationsphase. Da die Ergebnisse beurteilerabhängig sind, muss sich die Interpretation an den jeweiligen Beurteilern und ihrem Kontext orientieren (Kinsella, 1998).

Die MARS befindet sich noch in der Entwicklungsphase. Insbesondere ist die faktorielle Struktur noch nicht eindeutig geklärt. Die Autoren gehen von einer drei-Faktoren-Lösung aus, allerdings korrelieren zwei Faktoren (Daueraufmerksamkeit und Ablenkbarkeit) sehr hoch miteinander ($r = .79$), was ihre Interpretierbarkeit als getrennte Faktoren erschwert. Die einzelnen Faktoren liefern auch keine ausreichend reliablen Personenmessungen (Hart et al., 2006).

Zusätzlich zu den bisher vorgestellten Verfahren gibt es Fragebögen, mit denen Konzentrationsfehler (durch Störungen der Aufmerksamkeit bedingte Handlungsfehler) erfasst werden. Als Vorreiter ist der «Cognitive-Failures-Questionnaire» (CFQ, Broadbent et al., 1982, deutsche Bearbeitung von Klumb, 1995) zu erwähnen, welcher die Ausgangsbasis für die Entwicklung weiterer vergleichbarer Instrumente darstellte (z. B. «Short Inventory of Minor Lapses» (SIML) von Reason, 1993, «Attention-Related Cognitive Errors Scale» (ARCES) von Cheyne et al., 2006). Wie bei den Aufmerksamkeitsfragebögen ist der Stand der psychometrischen Absicherung auch hier noch nicht befriedigend. Der CFQ scheint eine gute Vorhersagevalidität im Bereich der Verkehrsdiagnostik zu besitzen. Menschen, die viele Aufmerksamkeitsfehler im CFQ angaben, hatten auch ein erhöhtes Autounfallrisiko (Larson et al., 1997). Allerdings scheint der CFQ nicht nur aufmerksamkeitsbezogene Fehler zu messen, da mehrere Studien zu sehr unterschiedlichen Faktorenlösungen kommen (Pollina et al., 1992; Matthews, Coyle und Craig, 1990; Larson et al., 1997; Wallace, Kass & Stanny, 2002).

Im deutschsprachigen Raum gibt es bislang keine vergleichbare Entwicklung wie in den angloamerikanischen Ländern. Häufig werden die angloamerikanischen Skalen einfach nur übersetzt und eingesetzt, ohne dass eine nochmalige psychometrische Evaluation stattgefunden hätte. Eine Ausnahme stellt der «Fragebogen erlebter Defizite der Aufmerksamkeit», (FEDA; Zimmermann et al., 1991) dar. Dieser Mehrkomponenten-Fragebogen wurde schon relativ früh im Zusammenhang mit der TAP entwickelt. Der FEDA besteht aus 27 Items, welche drei Aufmerksamkeitskomponenten zuzuordnen sind: «Ablenkbarkeit und Verlangsamung bei geistigen Prozessen», «Ermüdung und Verlangsamung bei praktischen Tätigkeiten» und «Antriebsminderung». Es gibt Hinweise für eine differenzielle Validität des FEDA bei Parkinson-Patienten (Müller et al., 1997) und bei Alzheimer-Patienten (Thöml & Romero, 2001). In der Studie von Suslow et al. (1998) konnte allerdings kein Unterschied zwischen schizophrenen Patienten und Personen ohne psychische Störung gefunden werden. In der Studie von Eidenmüller et al. (2001) stellte sich der FEDA nicht als ausreichend veränderungssensitiv dar, um

bedeutsame Veränderungen im Ausmaß der erlebten Aufmerksamkeitsdefizite im Verlauf eines Aufmerksamkeitstrainings zu erfassen.

Die bislang publizierten psychometrischen Studien zu den einzelnen Verfahren machen deutlich, dass es momentan im deutschsprachigen Raum kein Verfahren gibt, mit dem reliabel und valide Aufmerksamkeitsstörungen im Alltag erfasst werden können. Am häufigsten tritt das Problem der Konstruktvalidität auf: Welche und wie viele latente kognitive Prozesse werden gemessen? Auch lassen sich die theoriegeleitet postulierten Aufmerksamkeitskomponenten häufig nicht in einer Faktorenanalyse wieder finden. So stellt sich eine a posteriori Suche nach Erklärungsmodellen über die tatsächlich zugrunde liegenden Dimensionen ein. Das variierte Antwortformat erschwert die Vergleichbarkeit der Ergebnisse: manche Verfahren fordern katamnestische Aussagen von Patienten, deren Validität und Reliabilität noch zu überprüfen ist; andere verwenden 4- bis 5-stufige Likert-Skalen, welche das Ausmaß an Aufmerksamkeitsstörungen (z. B. MARS) erfassen sollen. Auch der Geltungsbereich der vorhandenen Verfahren variiert. So sind manche Verfahren speziell für hirngeschädigte Patienten entwickelt worden (z. B. MARS, RSAB) und dementsprechend nur eingeschränkt bei anderen Krankheitsbildern (z. B. ADHS, Asperger-Syndrom, Schizophrenie) einsetzbar, bei denen auch Aufmerksamkeitsstörungen auftreten können.

Mit der Entwicklung der Skala zur Erfassung von Aufmerksamkeitsdefiziten (SEA) soll daher ein psychometrisch fundiertes Instrument entwickelt werden, welches sich zur reliablen und validen Erfassung im Alltag erlebter Aufmerksamkeitsbeeinträchtigung eignet. Bei der Konstruktion wurde darauf geachtet, dass ein breiteres Spektrum an Symptomen einer Aufmerksamkeitsstörung krankheitsunspezifisch erfasst wird. Ferner wird das Ziel verfolgt, ein änderungssensitives Instrument zu entwickeln, welches bei Patienten im gesamten Krankheitsverlauf anwendbar ist.

Die vorliegende Studie zielt darauf ab, herauszufinden, inwieweit die für den Fragebogen ausgewählten Items gute Indikatoren des Konstruktes «Aufmerksamkeit» sind und inwieweit sie in der Lage sind, reliabel zwischen Patienten zu differenzieren. In Folge dieser ersten psychometrischen Ergebnisse soll eine Optimierung der SEA eingeleitet werden.

Methodik

Stichprobe

Die Stichprobe bestand aus 202 Patienten mit einem Durchschnittsalter von 53.3 Jahre ($SD = 14.3$). Die Mehrheit davon waren Männer ($N = 147$) und besaßen einen Hauptschulabschluss ($N = 101$). In Tabelle 1 sind die wich-

Tabelle 1
Stichprobencharakteristika

Alter M (SD)	53.3 (14.3)
Geschlecht m/w	237/55
Schulabschluss (N)	
Hauptschule	101
Mittlere Reife	65
Abitur	34
kein Abschluss	1
Diagnose (N)	
Ischämischer/Hämorrhagischer Insult	125
Schädel-Hirn-Trauma	25
Hirntumor	21
Weitere neurologische Erkrankungen*	18
Multiple Sklerose	5
Hypoxie	4
Morbus Parkinson	4
Dauer der Erkrankung in Wochen	
M (SD)	39.2 (124.9)
Median	7
Min.–Max.	0–804

*z. B. Hydrocephalus, inkompletter Querschnitt, zervikale Myelopathie

tigsten demographischen und klinischen Daten der untersuchten Stichprobe zusammengestellt.

Die Stichprobe wurde aus zwei neurologischen Rehabilitationszentren rekrutiert, wobei die Patienten im Vorfeld über die Ziele der Studie aufgeklärt wurden und ihr Einverständnis zur Teilnahme gaben. Patienten mit Sprachstörungen oder schweren kognitiven Einschränkungen wurden nicht in die Studie einbezogen. Genauso wurden Patienten mit einer beginnenden demenziellen Erkrankung nicht berücksichtigt.

Erfassung von Aufmerksamkeitsdefiziten: die Entwicklung der SEA

Bei der Erstellung des Itempools der SEA wurden 54 Items aus vorhandenen deutsch- und englischsprachigen Fragebögen ausgewählt, die von zwei Experten als besonders relevant zur Erfassung von Aufmerksamkeitsproblemen eingeschätzt wurden. Anschließend wurden die ausgewählten Items weiteren Experten vorgelegt, welche sie nach der erfassten Aufmerksamkeitskomponente, ihrer Relevanz im Behandlungsprozess und ihrer verhaltensnahen Formulierung beurteilten. Anhand dieser Beurteilungen wurden schließlich 33 Items für den Fragebogen ausgewählt. Da die SEA für die Verlaufsdiagnostik geeignet und damit sensitiv für Veränderungen sein soll, wurde ein mehrstufiges Antwortformat mit insgesamt fünf Antwortkategorien gewählt: «0 = kein Problem», «1 = gelegentlich, weniger als

einmal in der Woche», «2 = manchmal (1–3mal in der Woche)», «3 = häufig (problematisch an den meisten Tagen)», «4 = ein ständiges Problem (beeinträchtigt fast alle Aktivitäten jeden Tag)».

Um neben der Selbstbeurteilung auch Fremddangaben erheben zu können, wurde zusätzlich eine Fremdbeurteilungsversion erstellt, die 30 Items umfasst. Durch den Vergleich von Selbst- und Fremdurteilen können Angaben zum Störungsbewusstsein gemacht werden (Gauggel, Peleska & Bode, 2000).

Durchführung

Die Fremdbeurteilung erfolgte durch Neuropsychologen/Ergotherapeuten und in einem Fall durch eine psychologisch-technische Assistentin in der Regel zwei Wochen nach der Klinikaufnahme. Alle Rater wurden vor der Studie über deren Ziel und den eingesetzten Fragebogen informiert. Von den 202 Patienten konnte nur bei 163 Patienten eine Fremdbeurteilung erhoben werden, da bei den restlichen Patienten keine ausreichend reliable und valide Fremdbeurteilung möglich war (z. B. weil es nur wenige Kontakte mit dem Patienten gab).

Datenanalyse

Die Rasch-Analyse wurde mit dem Programm WINSTEPS 3.57.4 (Linarc, 2007), alle anderen statistischen Analysen wurden mit dem Statistikprogramm SPSS durchgeführt.

Die Rasch-Analyse wurde hier eingesetzt, weil sie bestimmte Vorteile gegenüber der klassischen Testtheorie aufweist und sie deshalb zunehmend Verwendung im klinischen Bereich findet. Neben der Gewährleistung eines Intervallskalenniveaus stellt die «spezifische Objektivität» einen wesentlichen Vorteil des Rasch-Modells dar. Diese Eigenschaft besagt, dass Unterschiede in der untersuchten latenten Eigenschaftsdimension zwischen verschiedenen Personen unabhängig von den verwendeten Items (z. B. unabhängig davon, ob leichte oder schwierige Items für den Vergleich herangezogen wurden) festgestellt werden können. Umgekehrt können auch Schwierigkeitsunterschiede zwischen Items unabhängig davon festgestellt werden, ob Personen mit einer niedrigen oder einer hohen Merkmalsausprägung untersucht wurden (Gauggel et al., 2004).

Die psychometrische Analyse der SEA erfolgte anhand des Ratingskalen-Modells (Andrich, 1982). Die Ergebnisse der Rasch Analyse werden grundsätzlich in logits ausgedrückt. Die psychometrische Qualität wird bei der Rasch-Analyse anhand von vier Kriterien beurteilt: (a) der Eindimensionalität, (b) der Reliabilität, (c) der Güte des Antwortformats und (d) des Grades der Zusammenpassung zwischen Stichprobe und Items («targeting»).

Zu (a). Die Eindimensionalität ist die zentrale Annahme des Rasch-Modells und wird anhand der Modellanpassung

der einzelnen Items überprüft. Hierzu werden Item-Fitstatistiken durch Aufsummierung und Quadrierung der Abweichungen der beobachteten von den erwarteten Werten berechnet. Deren Mittelwert (Mnsq-Wert) gibt die Item-Fitstatistik wieder. Item-Fitstatistiken größer als 1.0 deuten eine Abweichung von der Eindimensionalität an, während Mnsq-Werte kleiner 1.0 auf Redundanz und eine potenzielle Verletzung der lokal-stochastischen Unabhängigkeit hinweisen. Bei vergleichsweise kleinen Stichproben (weniger als 1000 Personen) schlagen Wolfe & Smith (2007) als Richtwert für die Verletzung der Eindimensionalität eines Items einen Mnsq-Wert von maximal 2.00 vor, wobei Fitstatistiken zwischen 0.5 bis 1.5 als gut für ein mehrstufiges Antwortformat angesehen werden. Zur besseren Beurteilung der Modellanpassung eines Items wurde zusätzlich der standardisierte Residualwert mit berücksichtigt (Smith, 2000). In der vorliegenden Studie galten Items mit einer Fitstatistik größer als 1.5 und einem standardisierten Residualwert (zstd) > 2.0 als modellunangepasst (misfit). Darüber hinaus wurde die Annahme der Eindimensionalität durch eine Faktorenanalyse der Residuen untermauert.

Zu (b). Die Reliabilität der SEA wurde anhand von zwei Kriterien beurteilt, die sowohl für die Personen- als auch für die Itemparameter ermittelt wurden. Das erste Kriterium heißt «Trennung» (separation) und wird als das Verhältnis der Varianz der Messungen zu ihrem Standard-schätzfehler definiert. Das Kriterium für eine gute Trennung ist 2.0 und höher. Eine Trennung gleich oder höher 2.0 weist darauf hin, dass der Fragebogen in der Lage ist, zwischen mindestens drei Ausprägungsebenen in der Personenfähigkeit (Personen mit einer höheren, einer mittleren und einer niedrigen Fähigkeit) bzw. in der Itemschwierigkeit (hohe – mittlere – niedrige Itemschwierigkeit) zu unterscheiden, was als Eigenschaft eines guten Fragebogens gewertet wird. Auf der Basis der Trennung kann auch eine Trennungsreliabilität berechnet werden, welche sich ähnlich wie Cronbach's α interpretieren lässt. Das Kriterium für eine gute Trennungsreliabilität liegt bei .80 oder höher und stellt hier das zweite Beurteilungs-Kriterium dar.

Zu (c). Zur Evidenz einer hohen psychometrischen Qualität des Fragebogens trägt die konsistente Benutzung des Antwortformats. Dabei sollen die Antwortkategorien der Beurteilungsskala nach ihrer Kalibrierung im Fähigkeitskontinuum einen monotonen Anstieg aufweisen. Bei der Kategorien-Kalibrierung werden die Schnittpunkte der jeweils benachbarten Kategorienfunktionen geschätzt. Wichtig ist, dass die Antwortkategorien so «geordnet» sind, dass der Schnittpunkt zwischen zwei «leichteren» Kategorien (z. B. Antwortkategorie «0» und «1») vor dem Schnittpunkt von zwei «schwierigeren» Kategorien (z. B. Antwortkategorien «2» und «3») liegt. Steigt die Kategorien-Kalibrierung über die einzelnen Antwortkategorien monoton an, bedeutet dies, dass mit wachsender Beeinträchtigung der Aufmerksamkeit die Übergangswahrscheinlichkeit von einer «leichteren» Kategorie (z. B. «2 = manchmal») bis hin zu den «schwierige-

Tabelle 2
Rasch Kennwerte

	Stichprobe	Item	Trennungsindex/Reliabilität*		Item-Misfit**	Was wurde gemacht?
	N	n	Items	Personen	N	
SEA-S	202	33	417/47	319/89	2 Items > 1.50	23 Misfit Personen eliminiert
	179	33	444/95	341/91	2 Items > 1.50	2 Items eliminiert
	179	28	49/19	319/45	0 Items > 1.50	
	163	30				
SEA-F			5.37/97	3.40/92	3 items > 1.50	

*Trennungsindex: das Verhältnis der Varianz der Messungen zu ihrem Standardfehler, Richtwert > 2.0; Reliabilität: Anteil der reproduzierbaren beobachteten Varianz, Richtwert > .80. **Item-Misfit: Items, die den Modellerwartungen nicht entsprechen.

ren» Kategorien (z. B. «4 = ein ständiges Problem») ansteigt. (d) Schließlich wird mit dem Targeting ersichtlich, inwieweit die Items für die Stichprobe angemessen sind. Dabei werden die Items nach ihrem Schwierigkeitsgrad und die Personen gemäß der Ausprägung ihrer Aufmerksamkeitsstörung zusammen auf demselben Kontinuum angeordnet.

Da für die Items und die Personenfähigkeit dieselbe Metrik benutzt wird, können ihre Mittelwerte unmittelbar miteinander verglichen werden. Je enger die zwei Mittelwerte zueinander liegen, desto geeigneter sind die Items für diese Stichprobe.

Darüber hinaus wurde durch einen Intraklassen-Korrelationskoeffizienten (ICC) die Übereinstimmung zwischen den Selbst- und den Fremdurteilen berechnet. Der Einfluss möglicher Moderator Variablen auf die Urteilsbildung wurde durch Varianzanalysen überprüft.

Ergebnisse

Ergebnisse der Selbstbeurteilung (SEA-S)

Wichtige Kennwerte der Rasch Analyse sind in Tabelle 2 aufgeführt.

Eindimensionalität

Das gesetzte Kriterium der Modellanpassung eines Items ($\text{Fit-Mnsq} < 1.50$ und $\text{zstd} < 2.00$) konnten von den insgesamt 33 Items 2 Items nicht erreichen. Am deutlichsten wich das Item «Ich bin ruhelos /zappelig, wenn ich nichts tue» mit einem Infit-Mnsq Wert von 1.78 ($\text{zstd} = 6.0$) von den Modellerwartungen ab, gefolgt von dem Item «Meine Leistungen sind nur morgens, oder nach einer Pause am besten», welches sich als ein Ausreißer erwiesen hat (Outfit-Mnsq Wert = 1.74; $\text{zstd} = 3.9$). Die Faktorenanalyse der Residuen ergab eine einheitliche Messdimension, durch die 69.4 % der Gesamtvarianz erklärt werden. Der erste Residualfaktor erklärt 2.8 % der Varianz der Residuen. Die Messdimension hat sich als 25mal stärker erwiesen als der erste Residualfaktor.

Reliabilität

Die 33 Items erlauben eine reliable Messung der Aufmerksamkeit und die Unterscheidung von mindestens drei Patientengruppen mit unterschiedlicher Defizitprägung: Patienten mit einer niedrigen, mit einer moderaten und mit einer hohen Beeinträchtigung der Aufmerksamkeit (Personen-Trennungs-Index = 2.86, Reliabilität = .89). Weiterhin deckten sie ein breites Schwierigkeitsspektrum mit einer sehr hohen Reliabilität im Sinne des Rasch-Modells (= die Wahrscheinlichkeit der Reproduzierbarkeit der Rangordnung der Itemschwierigkeiten, wenn derselbe Itemset einer anderen Patientengruppe vorgegeben wird), welches sich in einem Item-Trennungs-Index von 4.10 und einer Reliabilität von .94 bemerkbar machte. Von den 202 Patienten haben 23 den Modellerwartungen nicht entsprochen (Infit-Mnsq Wert > 1.50 und $\text{zstd} > 2.00$). Es handelt sich hierbei um Patienten, deren Antwortverhalten durch eine deutliche Präferenz der Extremkategorien der Beurteilungsskala gekennzeichnet war. Ein Ausschluss der nicht-passenden Personen führt zu einer Erhöhung der Reliabilität (s. Tabelle 2).

Targeting

Im Allgemeinen, je näher die durchschnittliche Personenfähigkeitsmessung zu 0 steht, desto besser bilden die Items die Personenfähigkeiten ab. Der Durchschnitt der Personenparameter lag bei -1.33 ($SD = 1.05$, min. max. = -4.28 – $.58$) und damit etwas entfernt vom ideellen Wert 0. Die negative Mittelwertdifferenz der Personenparameter ist ein Indikator dafür, dass der Test vergleichsweise leicht ist. Aufgaben wie «mit einer Aufgabe aufhören, wenn man merkt, dass man Fehler macht» oder «wach bleiben während der Ausführung einer Aufgabe» erwiesen sich als die leichtesten Items. Die meisten Patienten hatten bei diesen Anforderungen keine Beeinträchtigungen. Aufgaben wie beispielsweise «Konzentration beim gleichzeitigen Durchführen von mehreren Aufgaben lässt nach» fielen den Patienten dagegen schwerer (s. Abbildung 1).

Tabelle 3
Schwellenparameter

Antwortkategorie	0	1	2	3	4
Durchschnittliches Kategorienmaß*	-1.92	-1.04	-0.56	-0.18	0.02
Schnittpunkte**		-0.70	-0.16	0.09	0.78

*Inferenzstatistisches Maß, welches monoton ansteigen sollte; dies ist dann ein Zeichen dafür, dass die Personen, die in den einzelnen Items einen höheren Beurteilungswert erhalten, auch Diejenigen sind, die einen hohen Gesamtwert über alle Items hinweg erzielen. ** Punkte auf dem Kontinuum der latenten Dimension, an denen der Übergang von einer Kategorie zur nächsten stattfindet (in Logits ausgedrückt).

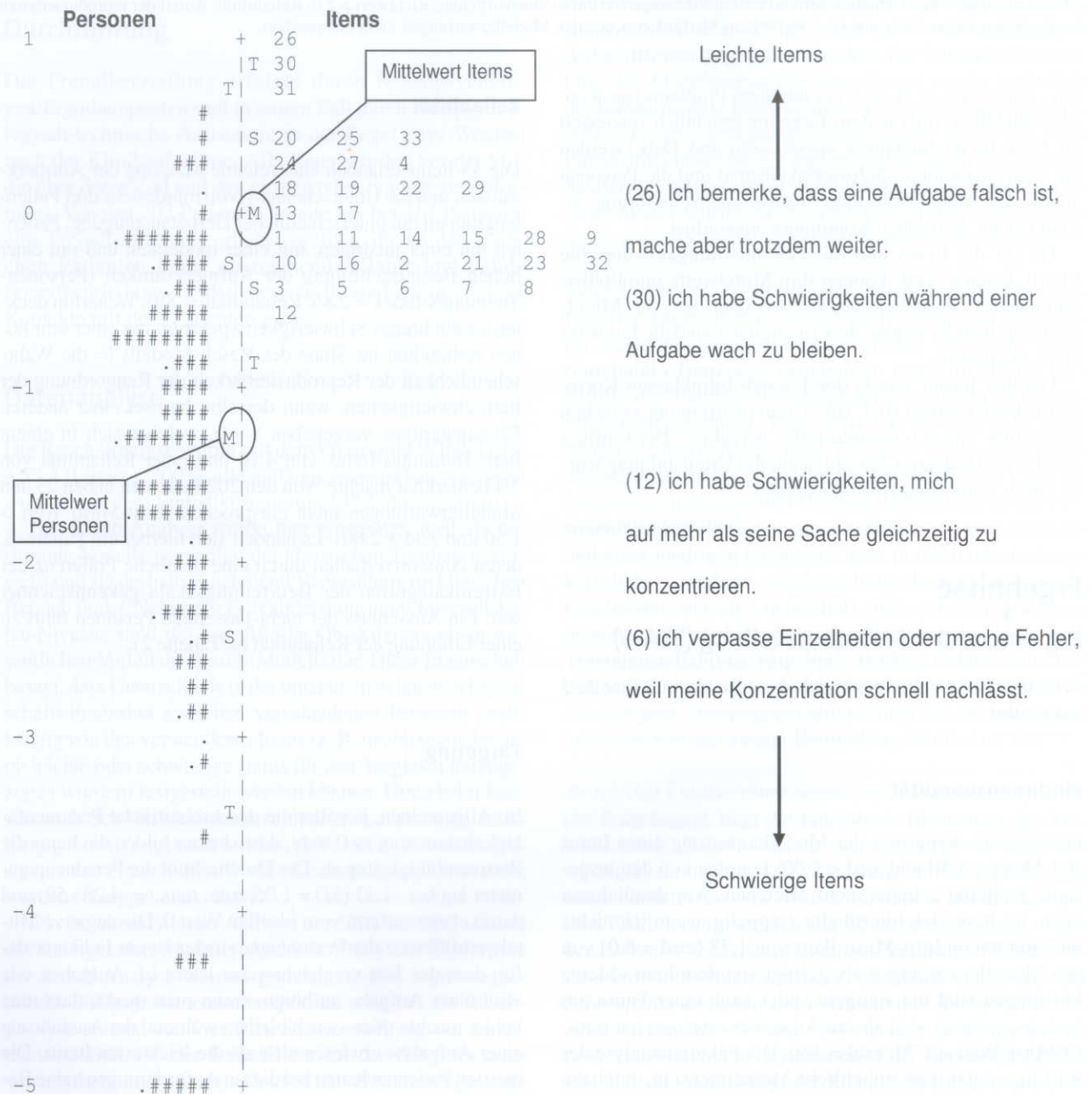


Abbildung 1. Das Targeting-Lineal. Personenfähigkeiten und Itemschwierigkeiten werden auf einer gemeinsamen Metrik angeordnet. Werte entsprechen logits. Jedes «#» bedeutet zwei Personen, jedes «.» bedeutet eine Person, M = Mittelwert, S = Standardabweichung, Z = zwei Standardabweichungen.

Antwortformat

Anhand der Tabelle 3 und zur Veranschaulichung anhand der Abbildung 2 wird ersichtlich, dass die Patienten das Antwortformat konsistent angewandt haben. Das durchschnittliche Kategorienmaß weist einen monotonen Anstieg auf und die Schnittpunkte zwischen den Kategorien überlappen sich nicht. Die Abstände zwischen den einzelnen Antwortkategorien waren jedoch kleiner als der empfohlene Abstand von mindestens 1.4 logits, welcher für eine klare Trennung zwischen den Kategorien spricht; denn nur in diesem Fall ist gesichert, dass die Antwortkategorien tatsächliche Fähigkeitsbereiche auf dem Konstruktkontinuum darstellen und nicht verwechselt werden können. Die häufigste Methode, die Beurteilungsskala zu optimieren, ist das Zusammenfügen von benachbarten Kategorien (Linarce, 2002). Erst die Reduktion der Antwortkategorien auf drei Stufen ergab einen ausreichenden Abstand zwischen den Kategorien (-2.02, 2.02). Dies reduzierte jedoch die Differenzierungskraft der Items (Trennungsindex/Rel. = 3.92/.94). Der anvisierte Einsatz der SEA zur Veränderungsmessung verträgt sich jedoch nicht mit der Komprimierung des Antwortformats, denn damit kann der Krankheits- bzw. Rehabilitationsverlauf nicht sensitiv genug gemessen werden.

Ergebnisse der Fremdbeurteilung (SEA-F)

Die psychometrische Auswertung der SEA-F ergab sehr gute Fitstatistiken, die sogar besser ausfielen als bei der Selbstbeurteilungsform und die Annahme der Eindimensionalität bestätigen. Dabei entsprechen drei Items nicht den Modellerwartungen (Infit-Mnsq > 1.50 und zstd > 2.0): «der Patient ist ruhelos/zappelig» (Infit-Mnsq = 2.16, zstd = 6.9), «der Patient ist ruhelos/zappelig, wenn er mit etwas beschäftigt ist» (Infit-Mnsq = 2.26, zstd = 6.8) und «die Leistungen des Patienten sind nur morgens oder nach einer Pause am besten» (Infit-Mnsq = 1.71, zstd = 5.0). Die SEA-F wies eine hohe Reliabilität und Differenzierungskraft auf – sowohl in Bezug auf die Einschätzung der Aufmerksamkeitsbeeinträchtigung des Patienten (Personentrennung/Rel. = 3.40/.92) als auch hinsichtlich der Itemschwierigkeiten (Itemtrennung/Rel. = 5.37/.97). Das Targeting fiel ähnlich wie bei der Selbstbeurteilungsform aus. Der Mittelwert der Personenparameter lag bei -1.44. Das Antwortformat wurde noch konsequenter von den Fremdbeurteilern als von den Patienten selber angewandt (durchschnittliches Kategorienmaß = -2.41, -1.21, -.49, .00, .26, Schnittpunkte = -1.20, -.51, .24, 1.47).

Selbst- und Fremdübereinstimmung

Das Ausmaß der Übereinstimmung zwischen den Selbst- und den Fremdurteilen fiel niedrig aus (Intraklassenkor-

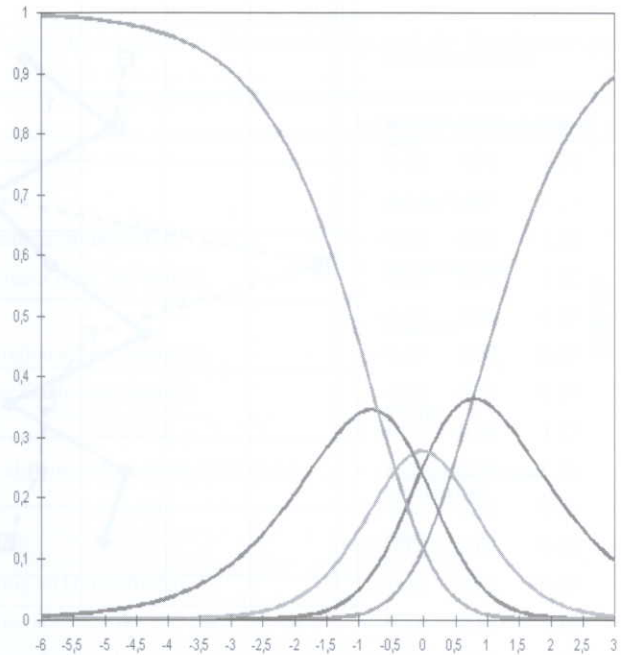
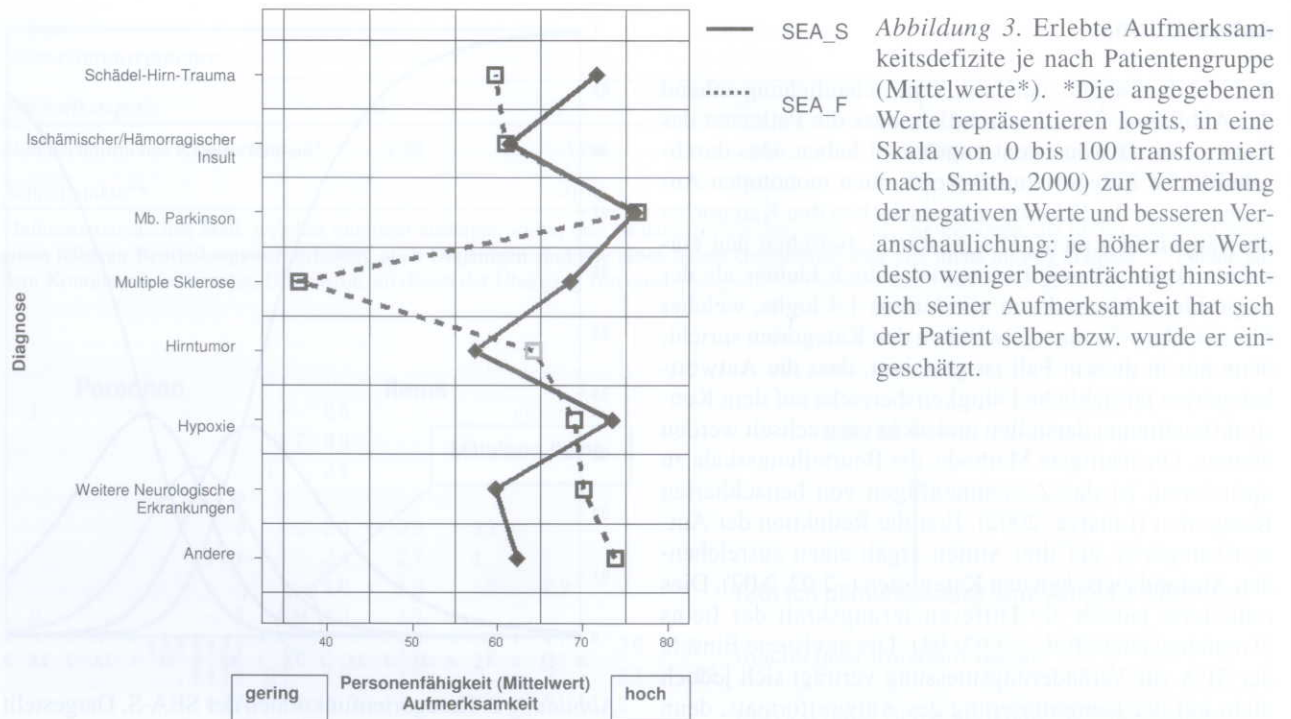


Abbildung 2. Kategorienfunktionen der SEA-S. Dargestellt werden die Antwortkurven der Fünf-Antwortkategorien, welche geordnet verlaufen (von links nach rechts: Kategorie 1 bis Kategorie 5).

relation, ICC = .34), ein in der Literatur immer wieder beobachteter Befund, der vor allem durch das zum Teil eingeschränkte Störungsbewusstsein der Patienten verursacht wird (Gauggel, Peleska & Bode, 2000; Hart et al., 2004).

Moderator-Variablen

Sowohl das Alter als auch der Krankheitsbeginn korrelierten nicht mit dem Gesamtscore der SEA-S (jeweils $r = -.09$ und $r = .07$) und SEA-F (jeweils $r = .07$ und $r = .04$). Das Geschlecht (SEA-S: $F_{(202, 1)} = 3.85$, *ns*) und der Schulabschluss (SEA-S: $F_{(202, 4)} = .57$, *ns*) spielten ebenso keine Rolle bei der Beantwortung beider Formen des SEA. Die Frage, ob es einen Unterschied bei den unterschiedlichen ätiologischen Gruppen gibt, konnte hier aufgrund der zum Teil niedrigen Fallzahl mancher Patientengruppen nicht geklärt werden. Jedoch scheint die SEA Unterschiede bei den Antwortprofilen der Patientengruppen aufzudecken. Abbildung 3 gibt die Gruppenmittelwerte der Personenfähigkeiten sowohl der Selbst- als auch der Fremdbeurteilungsform der SEA wieder. Aus der Grafik wird ersichtlich, dass Aufmerksamkeitsdefizite tendenziell in Abhängigkeit von der Diagnose variieren. Interessant dabei ist, dass die Fremdbeurteilung je nach Störungsbild stark von der Selbstbeurteilung abweichen kann.



Diskussion

Die Entwicklung einer Rating-Skala zur Erfassung von Aufmerksamkeitsdefiziten ist eine Herausforderung, denn viele der Items (Verhaltensweisen), die potenzielle Verhaltens- und Erlebensindikatoren von Aufmerksamkeitsproblemen sind, könnten mit anderen kognitiven Störungen verwechselt werden wie beispielsweise Gedächtnisstörungen (Whyte et al., 2003; Sturm, 2005). Jedoch zeigte sich in der vorliegenden Analyse, dass die 33 Items der Selbstbeurteilungsfom (bzw. 30 Items der Fremdbeurteilungsfom) eine einheitliche Dimension zu messen scheinen und zwar sehr differenziert und reliabel (Trennungindex/Reliabilität 4.24/.95 bei der SEA-S und 5.37/.97 bei der SEA-F). In Bezug auf die Personenparameter ermöglichte der Itempool der SEA eine gute Differenzierung zwischen den Patienten mit einer geringen bis mittleren Beeinträchtigung der Aufmerksamkeitsfunktion.

Zum Vergleich erzielte die Fremdbeurteilungsskala «Moss Attention Rating Scale» (MARS) zwar etwas höhere Personentrennungswerte (Personen-Trennung-Index/Reliabilität: 4.11/.94 bei den Ergotherapeuten- und 3.83/.94 bei den Physiotherapeuten-Urteilen) als die SEA-F (3.40/.92) (Hart et al., 2006). Wenn man aber bedenkt, dass die MARS bis jetzt nur bei hirngeschädigten Patienten in der akuten Phase erprobt wurde, dann ist das Ergebnis der SEA-F bei einer weitaus heterogeneren Gruppe als vergleichsweise gut zu bewerten.

Bei der Konstruktion der SEA wurde darauf geachtet, ein klinisches Instrument zu entwickeln, das über den gesamten Rehabilitationsverlauf eingesetzt werden kann.

Dies setzt einen Itempool voraus, der in der Lage ist, zwischen Patienten in unterschiedlichen Krankheitsphasen zu differenzieren. Die SEA scheint gut geeignet zu sein, um zwischen Patienten mit einer geringen bis mittelschweren Beeinträchtigung der Aufmerksamkeit zu unterscheiden. Um ihren Validitätsbereich auf die Gruppe der Patienten mit einer niedrigen Ausprägung von Aufmerksamkeitsdefiziten auszuweiten, sollte sie um weitere schwierige Items ergänzt werden. Darüber hinaus scheint die SEA in der Lage zu sein, unterschiedliche Antwortprofile je nach Diagnosegruppe aufzudecken, ein Ergebnis, welches durch weitere Studien gesichert werden muss.

Überraschend ist, dass nicht, wie ursprünglich erwartet, mehrere Komponenten der Aufmerksamkeit in der SEA abgebildet werden konnten. Die psychometrischen Analysen weisen auf eine Eindimensionalität der SEA hin. Momentan kann nur spekuliert werden, warum sich nicht mehrere Aufmerksamkeitskomponenten in unserem Fragebogen abbilden ließen. Am wahrscheinlichsten ist, dass die Patienten, aber auch Fremdbeurteiler nur eingeschränkt zwischen den verschiedenen Symptomen einer Aufmerksamkeitsstörung differenzieren können. Zu ähnlichen Ergebnissen kamen auch andere Autoren, die durch die Fragebogen-Methode Aufmerksamkeits-symptome zu erfassen versuchten. Beispielsweise berücksichtigten Whyte et al. (2003) bei der Konstruktion der MARS 11 verschiedene Aspekte der Aufmerksamkeit (Daueraufmerksamkeit, Ablenkbarkeit, Selektivität, Alertness, Geteilte Aufmerksamkeit, Handlungsinitialisierung, Arbeitsgedächtnis, usw.). Empirisch konnten sie mittels Raschmodellierung nur eine Dimension ermitteln (Whyte et al. 2003). Die zweite verkürzte Version der MARS erwies sich ebenso als eindimensional. Mit Hilfe einer Faktorenanalyse

Tabelle 4

Itemkennwerte und -auswahl der SEA-S. Dargestellt sind die Itemschwierigkeiten, Trennschärfen und die Ergebnisse der Fitstatistik in Logits

Nr.	Iteminhalt	IP	r_{it}	Fit
1	Mir fehlt die mentale Energie, um etwas zu unternehmen.	-0.22	0.52	1.04
2	Ich antworte langsam auf Fragen oder in Gesprächen.	-0.33	0.51	1.24
3	Ich kann meine Gedanken nicht bei einer Sache halten, weil sie ständig abschweifen.	-0.41	0.58	1.21
4	Ich kann meine Gedanken nicht bei einer Sache halten, weil sich mein Kopf leer anfühlt.	0.20	0.50	1.02
5	Ich kann mich nur für sehr kurze Zeit konzentrieren.	-0.40	0.66	0.68
6	Ich verpasse Einzelheiten oder mach Fehler, weil meine Konzentration schnell nachlässt.	-0.49	0.67	0.69
7	Ich verliere den Faden, wenn andere Personen anwesend sind oder Geräusche machen.	-0.38	0.59	0.94
8	Ich werde leicht durch Nebengeräusche abgelenkt.	-0.43	0.58	1.13
9	Ich habe Schwierigkeiten einem Gespräch zu folgen, wenn mehr als eine Person daran beteiligt ist.	-0.15	0.56	0.98
10	Ich verliere leicht den Faden, wenn ich unterbrochen werde.	-0.33	0.60	0.76
11	Ich bin überfordert, wenn eine Aufgabe aus mehreren Teilen besteht.	-0.18	0.62	0.68
12	Ich habe Schwierigkeiten, mich auf mehr als eine Sache gleichzeitig zu konzentrieren.	-0.56	0.65	0.80
13	Selbst leichte Arbeiten muss ich unterbrechen, um mich auszuruhen.	-0.11	0.52	1.18
14	Bei Gesprächen bekomme ich nicht alles mit, weil so schnell gesprochen wird.	-0.18	0.49	1.35
15	Ich fange an, eine Sache zu tun und werde dann von anderen Dingen abgelenkt.	-0.20	0.63	0.67
16	Ich brauche doppelt soviel Zeit wie andere, um eine Arbeit zu erledigen.	-0.30	0.55	1.02
17	Es fällt mir schwer, mit Arbeiten rechtzeitig fertig zu werden.	-0.06	0.54	0.91
18	Im Straßenverkehr fühle ich mich überfordert, weil alles so schnell geht.	0.10	0.45	1.45
19	Schon ein kleines Geräusch kann mich beim Lesen so stören, dass ich den Faden verliere.	0.10	0.48	1.30
20	Meine Familie und Bekannten beklagen sich über meine Zerstretheit.	0.41	0.47	1.02
21	Es kommt vor, dass ich plötzlich nicht mehr weiß, was ich gerade tun wollte.	-0.30	0.58	0.84
22	Ich bemerke, dass ich (tag-)träume, wenn ich eigentlich einer Sache zuhören sollte.	-0.05	0.57	0.66
23	Ich habe das Gefühl, dass mich alltägliche Arbeiten sehr anstrengen	-0.29	0.52	1.06
24	Ich bin zwar anwesend, nehme aber nichts auf.	0.29	0.48	1.11
25	Ich handle anders, als ich es ursprünglich geplant habe.	0.34	0.45	0.89
26	Ich bemerke, dass eine Aufgabe falsch ist, mache aber trotzdem weiter.	1.05	0.33	1.26
27	Ich bin nicht in der Lage, bei einer Aufgabe zu bleiben ohne zu unterbrechen oder abzudriften.	0.28	0.51	0.95
28	Meine Leistungen sind nur morgens oder nach einer Pause am besten.	-0.11	0.46	1.74
29	Bei meinen eigenen Handlungen nehme ich Fehler nicht wahr.	0.12	0.46	1.03
30	Ich habe Schwierigkeiten, während einer Aufgabe wach zu bleiben.	0.80	0.40	1.23
31	Ich höre mitten in einer Aufgabe auf und mache gar nichts mehr.	0.73	0.42	1.11
32	Ich bin ruhelos/zappelig, wenn ich nichts tue.	-0.23	0.36	1.78
33	Ich bin ruhelos/zappelig, wenn ich mit etwas beschäftigt bin.	0.49	0.44	1.01
neu	Ich merke, dass ich immer wieder die gleichen Fehler mache.			
neu	Ich bin schnell müde.			
neu	Ich muss einen Zeitungsartikel mehrmals lesen, damit ich die Zusammenhänge verstehen kann.			
neu	Wenn ich etwas erzähle, sagen mir andere Leute, dass ich nicht auf den Punkt komme.			
neu	Ich fange an, hier mal was und da mal was zu tun, und bringe nichts zu Ende.			
neu	Ich gehe abends früh ins Bett.			
neu	Mir fällt es schwer, Wichtiges vom Unwichtigen zu unterscheiden.			
neu	Ich bin abends so kraftlos, dass ich nichts mehr unternehmen kann.			

IP = Itemparameter = Itemschwierigkeit; r_{it} = Trennschärfe nach dem Rasch Modell

lyse konnten zwar drei Faktoren extrahiert werden, die allerdings bis zu einem r von .79 interkorrelierten. Diese hohe Interkorrelation der Faktoren erschwerte eine distinkte Interpretierbarkeit der Faktoren (Hart et al., 2006). Die Autoren vermuteten, dass es nur einen Hauptfaktor der Aufmerksamkeit mit drei Subkomponenten gibt.

Ein weiteres Problem, das eventuell mit dazu beigetragen haben kann, dass nicht mehrere Aufmerksamkeitsfaktoren gefunden werden konnten, besteht in dem häufig verminderten Störungsbewusstsein hirngeschädigter Patienten (Gauggel, 2007). Die betroffenen Patienten sind in zahlreichen Fällen keine guten Beurteiler und berichten häufig von keinen oder nur sehr geringen Defiziten. Auch Fremdbeurteiler sind von Beurteilungsproblemen nicht verschont. Die Beurteilung hängt ganz wesentlich davon ab, wie oft und in welchen Situationen die Fremdbeurteiler die Patienten beobachten können. Wenn die vorhandenen Aufmerksamkeitsdefizite vom Patienten nicht beklagt werden und im Alltag kaum auffallen, ist es auch für einen Fremdbeurteiler schwer ein reliables Urteil abzugeben.

Eventuell kann aber auch die untersuchte Patientenstichprobe für das Ergebnis verantwortlich sein. Es handelt sich um eine leicht bis mittelgradig schwer beeinträchtigte Patientenstichprobe. Vielleicht würde sich bei schwerer beeinträchtigten Patienten eine klarere Differenzierung bei den verschiedenen Aufmerksamkeitskomponenten zeigen.

Die Entwicklung der SEA ist somit nicht abgeschlossen. Die vorliegende Studie lieferte jedoch erste Hinweise für die Konstruktvalidität der SEA. In einer weiteren Studie von Beck et al. (submitted) wurde die konvergente Validität des SEA überprüft, indem 47 Hirngeschädigte sowohl die SEA als auch zwei computerbasierte Aufmerksamkeits-tests (ANT und TAP) durchgeführt wurden. Beck et al. (submitted) fanden niedrige Korrelationen zwischen den Selbstbeurteilungen und den objektiven Testbefunden ($r = 0$ –.23 mit ANT und $r = -.02$ –.39 mit TAP), jedoch mittlere bis hohe Korrelationen zwischen den Fremdbeurteilungen und den ANT- ($r = .62$ –.70) sowie den TAP-Kennwerten ($r = .42$ –.86). Die hohen Korrelationen zwischen den Fremdbeurteilungen und den objektiven Tests liefern einen Hinweis für die Übereinstimmungsvalidität der SEA. Weitere Validitätsstudien sind jedoch erforderlich, um die diagnostische Qualität der SEA zu untermauern.

Bei der Bestimmung der Validität gilt es zu beachten, dass die Korrelation zwischen Selbstbeurteilungen und objektiven Tests bei Patienten generell kleiner ausfällt als zwischen der Fremdbeurteilung und objektiven Tests (Gauggel, Peleska & Bode, 2000). Dies hängt mit dem häufig anzutreffenden verminderten Störungsbewusstsein hirngeschädigter Patienten zusammen (Whyte et al., 2003; Hart et al., 2004).

Empfehlungen für die neue Version der SEA

Die Fitstatistiken wiesen darauf hin, dass das Itemset eindimensional ist und sehr zuverlässig Aufmerksamkeitsde-

fizite erfassen kann. Allerdings passt das Itemset nicht optimal zur untersuchten Personengruppe. Die Patienten waren nicht so beeinträchtigt, dass sie alle durch die Items erfassten Aufmerksamkeitsdefizite aufzeigen. Eine notwendige Verbesserung des Fragebogens besteht daher darin, nicht-passende Items zu löschen und durch Items zu ergänzen, die geeignet sind, auch zwischen Patienten mit geringeren Beeinträchtigungen der Aufmerksamkeit besser zu diskriminieren (siehe Tabelle 3). Im Anhang ist der Fragebogen mit allen Items dargestellt. Items, die gelöscht bzw. ergänzt wurden, sind markiert. Es liegt somit eine einheitliche Version für die Selbst- und Fremdbeurteilung mit insgesamt 28 Items vor, zu denen fünf neue Items mit höherer Itemschwierigkeit hinzugefügt wurden (s. Tabelle 4).

Literatur

- Andrich, D. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika*, 47, 105–113.
- Beck, L., Heusinger A., Boecker, M., Niemann, H. & Gauggel S. (submitted). Comparison of two computerized attention tests in a sample of brain damaged patients. Manuscript submitted for publication.
- Brown, S. A., McCauley, S. R., Levin, H. S., Boake, C., Goldfader, P. R., McCormick, S. D. et al. (2001). Factor analysis of an outcome interview for use in clinical trials of traumatically brain-injured patients: a preliminary study. *American Journal of Physical Medicine & Rehabilitation*, 80, 196–205.
- Bühner, M., Schmidt-Atzert, L., Richter, S. Grieshaber, E. (2002). Selbstbeurteilungen der Aufmerksamkeit: Ein Vergleich zwischen Hirngeschädigten und Gesunden. *Zeitschrift für Neuropsychologie*, 13, 263–269.
- Conners, C. K. & Multi-Health Systems Staff. (1995). *Conners' Continuous Performance Test*. Toronto: MHS.
- Eidenmüller, A., Kallus, K. W., Fröhlich, H., Bieber, K. & Poimann, H. (2001). Evaluation eines ambulanten Aufmerksamkeitstrainings. *Zeitschrift für Neuropsychologie*, 12, 160–172.
- Gauggel, S. (2007). Anosognosie. In S. Gauggel & M. Herrmann (Hrsg.), *Handbuch der Neuro- und Biopsychologie* (S. 537–544). Göttingen: Hogrefe.
- Gauggel, S., Böcker, M., Zimmermann, P., Privou, C. & Lutz, D. (2004). Item-response-Theorie und deren Anwendung in der Neurologie. *Nervenarzt*, 75, 1179–1186.
- Gauggel, S., Peleska, B. & Bode, R. K. (2000). Relationship between cognitive impairments and rated activity restrictions in stroke patients. *Journal of Head Trauma Rehabilitation*, 15, 710–723.
- Gronwall, D. M. A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367–373.
- Hart, T., Sherer, M., Whyte, J., Polansky, M. & Novack, T. A. (2004). Awareness of behavioral, cognitive, and physical deficits in acute traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 85, 1450–1456.
- Hart, T., Whyte, J., Millis, S., Bode, R., Malec, J., Richardson, R. N. & Hammond, F. (2006). Dimensions of disordered attention in traumatic brain injury: Further validation of the Moss

- Attention Rating Scale. *Archives of Physical Medicine and Rehabilitation*, 87, 647–655.
- Jensen, A.R. & Rohwer, W.D. (1966). The Stroop Color-Word Test: A review. *Acta Psychologica*, 25, 36–93.
- Kinsella, G.J. (1998). Assessment of attention following traumatic brain injury: A review. *Neuropsychological Rehabilitation*, 8, 351–375.
- Kolitz, B.P., Vanderploeg, R.D. & Curtiss, G. (2003). Development of the Key Behaviors Change Inventory: A traumatic brain injury behavioral outcome assessment instrument. *Archives of Physical Medicine and Rehabilitation*, 84, 277–284.
- Klumb, P.L. (1995). Cognitive failures and performance differences: Validation studies of a German version of the cognitive failures questionnaire. *Ergonomics*, 38, 1456–1467.
- Larson, G.E., Alderton, D.L., Neideffer, M. & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *British Journal of Psychology*, 88, 29–38.
- Lewis, R.F. & Rennick, P.M. (1979). *Manual of the repeatable cognitive-perceptual-motor battery*. Grosse Pointe Park, MI: Axon.
- Linarce, J.M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3, 85–106.
- Linarce, J.M. (2007). *WINSTEPS: Rasch measurement computer program, Version 3.57.4* [Computer program]. Chicago: Winsteps.com.
- Matthews, G., Coyle, K. & Craig, A. (1990). Multiple factors of cognitive failure and their relationships with stress vulnerability. *Journal of Psychopathology & Behavioral Assessment*, 12, 49–65.
- Müller, S., Brinkmeier, W. & Schendemann, G. (1997). Erlebte Aufmerksamkeitsdefizite bei Parkinson Patienten – Eine Befragung mit dem FEDA. *Neurologie & Rehabilitation*, 3, 96–102.
- Niemann, H. & Gauggel, S. (2005). Störungen der Aufmerksamkeit. In H.O. Karnath, W. Hartje & W. Ziegler (Hrsg.), *Kognitive Neurologie* (S. 111–125). Stuttgart: Thieme.
- Pollina, L.K., Greene, A.L., Tunick, R.H. & Puckett, J.M. (1992). Dimensions of everyday memory in young adulthood. *British Journal of Psychology*, 83, 305–321.
- Rabin, L.A., Barr, W.B. & Burton, L.A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, 20, 33–65.
- Smith, R.M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, 1, 199–218.
- Sturm, W. (2005). *Aufmerksamkeitsstörungen*. Göttingen: Hogrefe.
- Suslow, T., Arolt, V. & Junghanns, K. (1998). Differentielle Validität des Fragebogens erlebter Defizite der Aufmerksamkeit (FEDA): Konkurrente Validierungsergebnisse bei schizophrenen und depressiven Patienten. *Zeitschrift für Psychologie, Psychiatrie und Psychotherapie*, 46, 152–165.
- Theml, T. & Romero, B. (2001). Selbstbeurteilung von Aufmerksamkeitsdefiziten bei Alzheimer-Kranken mit sehr leichter Demenz. *Zeitschrift für Neuropsychologie*, 12, 151–159.
- Wallace, J.C., Kass, S.J. & Stanny, C.J. (2002). The cognitive failures questionnaire revised: Dimensions and correlates. *The Journal of General Psychology*, 129, 238–256.
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale – Third edition: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale – Third edition: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.
- Whyte, J., Hart, T., Bode, R. & Malec, J. (2003). The Moss Attention Rating Scale for traumatic brain injury: Initial psychometric assessment. *Archives of Physical Medicine and Rehabilitation*, 84, 268–276.
- Zimmermann, P., Messner, C., Poser, U. & Sedelmeier, P. (1991). Ein Fragebogen erlebter Defizite der Aufmerksamkeit (FEDA). Unveröffentlichtes Manuskript, Universität Freiburg.

Dr. Eftychia Volz-Sidiropoulou

Institut für Medizinische Psychologie
und Medizinische Soziologie
Universitätsklinik der RWTH Aachen
Pauwelsstraße 30
D-52074 Aachen
Tel. +49 241 80 89002
Fax +49 241 80 82 487
E-Mail evolz-sidiropoulou@ukaachen.de