

Suboptimales Leistungsverhalten

Risiko und Chance für die klinische Neuropsychologie

Matthias Henry¹ und Thomas Merten²

¹ Otto-von-Guericke-Universität, Magdeburg, Klinik für Neurologie

² Klinikum im Friedrichshain, Berlin, Klinik für Neurologie

psychoneuro 2005; 31 (10): 519–522

P sychologischen Testergebnissen haftet der Nimbus an, objektive Leistungsmaße zu sein. Nicht selten liest man in einem psychologischen Befund, dass bei einem Patienten Leistungsausfälle mit Hilfe von Tests „objektiviert“ worden seien. Sodann ist auch der Schluss, diese vermeintlich objektivierten kognitiven Störungen auf eine hirnanorganische Ursache zu attribuieren, nicht fern, selbst wenn es keinerlei unzweifelhafte Hinweise aus spezieller Anamnese oder Bildgebung dafür gibt.

Wie objektiv aber sind Testergebnisse wirklich? Es ist nicht schwer, sich den Fall eines Probanden vorzustellen, der in einer Begutachtungssituation an einer vollen Entfaltung seiner Leistungsmöglichkeiten nicht interessiert ist. Dieser Proband müsste dann einfach in einer Gedächtnisaufgabe weniger Worte wiederholen, als er wirklich behalten hat, in einer Konzentrationsaufgabe langsamer sein, als er tatsächlich ist, und in einer visuell-räumlichen Aufgabe Fehler begehen, die er eigentlich nicht machen würde. Nach einer Stunde Untersuchungszeit könnte er Erschöpfung signalisieren, obwohl er in Wahrheit noch mühelos zwei weitere Stunden zu arbeiten fähig wäre.

Suboptimales Leistungsverhalten kann derzeit als eine der größten Gefahrenquellen für die Validität neuropsychologischer Testdaten gesehen werden. International hat sich seit geraumer Zeit die Erkenntnis durchgesetzt, dass die Messung der Leistungsmotivation ein notwendiger Bestandteil neuropsychologischer Untersuchungen ist. Vor allem in den englischsprachigen Ländern wurden standardisierte Testverfahren zur Erfassung der Leistungsmotivation entwickelt, von denen sich besonders die sogenannten Symptomvalidierungstests bewährt haben. Im deutschsprachigen Raum ist dagegen bislang ein auffälliger Mangel an relevanten Publikationen und Bemühungen zur Entwicklung bzw. Adaptation von standardisierten Testverfahren zu verzeichnen. Es ist allerdings zu erwarten, dass sich diese Situation in den kommenden Jahren grundlegend ändern wird.

In Analogie dazu könnten bei einem psychiatrischen Patienten bedeutsame Motivationseinschränkungen Teil der Psychopathologie selbst sein. Resultat wäre wiederum eine Verzerrung von Testergebnissen in einem unkalkulierbaren Ausmaß.

In beiden Fällen müsste man vom Vorliegen eines suboptimalen Leistungsverhaltens bzw. negativer Antwortverzerrungen (negative response bias) sprechen. Damit wird beschrieben, dass ein Proband – aus welchen Gründen auch immer – in einem Test schlechtere Ergebnisse erzielt, als es ihm aufgrund seiner Leistungsvoraussetzungen eigentlich möglich gewesen wäre.

Suboptimales Leistungsverhalten kann im Rahmen einer Simulation oder Aggravation, bei somatoformen Störungen, bei artifiziellen Störungen und zahlreichen weiteren

psychiatrischen Erkrankungen auftreten. Eine besonders hohe Auftretenswahrscheinlichkeit wird bei Patienten mit chronischem Müdigkeitssyndrom, Fibromyalgie, sog. Schleudertrauma der Halswirbelsäule und neurotoxischen Erkrankungen beschrieben. Ein Diagnostiker, der das Auftreten einer eingeschränkten Leistungsmotivation nicht erkennt, wertet invalide Testergebnisse fälschlicherweise als valide und läuft Gefahr, einen Motivationsmangel fälschlicherweise als zerebrale Fehlfunktion zu interpretieren.

Notwendigkeit der Diagnostik von suboptimalem Leistungsverhalten

Es ist keine neue Erkenntnis, dass die Trennung von Einflüssen der Leistungsmotivation und der eigentlichen kognitiven Leistungs-

fähigkeit eines Probanden unerlässlich ist. Besonders im nordamerikanischen Raum haben zahlreiche Forschungsergebnisse in den vergangenen beiden Jahrzehnten zu einer breiten Anerkennung dieses Sachverhalts geführt. Im gleichen Zeitraum entstand eine wachsende Zahl von Methoden und standardisierten Testverfahren, die eine Erfassung der Leistungsmotivation ermöglichen. Beide Entwicklungen haben dazu geführt, dass die Messung der Leistungsmotivation in Ländern wie den USA, Kanada oder Australien inzwischen zu einem unverzichtbaren Bestandteil neuropsychologischer Untersuchungen geworden ist.

Drei wichtige Gründe sollen an dieser Stelle angeführt werden, die zu einer stärkeren Beachtung der Motivationsmessung in der klinischen und Gutachtenpraxis geführt haben.

Neuropsychologische Experten sind in der Erkennung von Simulationsversuchen kaum genauer als der Zufall: Dieser erstaunliche Sachverhalt wurde bereits von Heaton, Smith, Lehman und Vogt (4) berichtet. Neuropsychologische Gutachter sollten in dieser Studie beurteilen, ob vorliegende Testergebnisse von echten Patienten mit einer Kopfverletzung oder von experimentellen Simulanten stammten, d.h. gesunden Probanden, die Folgen einer Kopfverletzung vortäuschen sollten. Die Genauigkeit, mit der die Gutachter die Probanden korrekt zuordneten, entsprach dabei der Ratewahrscheinlichkeit (50%) oder lag teilweise etwas darüber. Es spielten weder die klinische Erfahrung noch die Sicherheit der Gutachter bezüglich ihrer eigenen Urteile eine Rolle für die Güte der Klassifikation. Vergleichbare Studien bescheinigten den Experten sogar Klassifikationswerte unterhalb der Ratewahrscheinlichkeit. Bemerkenswert ist, dass Ergebnisse, die die Fähigkeiten von neuropsychologischen Gutachtern in ein besseres Licht stellen würden, bisher anscheinend nicht vorliegen (13).

Die Leistungsmotivation hat einen größeren Einfluss auf Testergebnisse als die Schwere der Beeinträchtigungen eines Patienten:

Diese auf den ersten Blick ebenfalls erstaunlich anmutende Aussage beruht auf Untersuchungen der Forschergruppe um Paul Green (2). Es wurden 904 Patienten im Rahmen von neuropsychologischen Begutachtungen untersucht. Alle Patienten hatten zum Zeitpunkt der Untersuchung Aussicht auf eine finanzielle Kompensation. Die einzelnen Ergebnisse der verschiedenen neuropsychologischen Verfahren wurden zu einem auf z-Werten basierenden Gesamtwert zusammengefasst (Overall Test Battery Mean; OTBM) und anschließend berechnet, welcher Anteil der Varianz dieses Gesamtergebnisses durch Maße der Leistungsmotivation vorhergesagt werden konnte. Dieser Anteil betrug 53%. Andere wichtige Faktoren, wie etwa Bildungsjahre und Alter, hatten im Vergleich dazu einen recht geringen Vorhersagewert (11% bzw. 4% der Gesamtvarianz). Bemerkenswert ist dabei insbesondere, dass auch die Schwere der Verletzungen, die die Patienten erlitten hatten, weit weniger einflussreich war als die Leistungsmotivation. So erzielten Patienten nach leichten zerebralen Verletzungen keineswegs bessere Ergebnisse in den neuropsychologischen Testverfahren als solche nach schweren! Wurden allerdings die Probanden von der Datenanalyse ausgeschlossen, für die ein spezifischer Test zur Diagnostik der Leistungsmotivation (ein so genannter Symptomvalidierungstest) eine unzureichende Leistungsanstrengung ausgewiesen hatte, so entsprach das Ergebnis wieder dem sachlogisch zu erwartenden: Die Gruppe der schwer beeinträchtigten Patienten wies schlechtere Ergebnisse auf als die der leicht beeinträchtigten. Die Verzerrungen aufgrund suboptimaler Leistungsmotivation fielen um einiges größer aus als tatsächliche Leistungsunterschiede aufgrund unterschiedlich schwerer Hirnverletzungen. In Standardabweichungen des OTBM ausgedrückt waren die Effekte in etwa 4,5-mal so groß.

Ein beträchtlicher Anteil der Untersuchten in neuropsychologischen Begutachtungen weist negative Antwortverzerrungen auf:

Um die Gefahr abzuschätzen, die negative Antwortverzerrungen für die Genauigkeit und Zuverlässigkeit der neuropsychologischen Diagnostik bedeuten, muss man wissen, wie häufig ein solches Verhalten unter Patienten vorkommt. Anhaltspunkte dazu gibt eine in den USA durchgeführte Erhebung von Mittenberg, Patton, Canyock und Condit (11), in der Grundraten für suboptimales Leistungsverhalten anhand von 33531 Einzelfällen bestimmt wurden. Dabei ergab sich für heterogene Gruppen von Patienten, die im Rahmen zivil-, renten- oder arbeitsrechtlich relevanter Fragestellungen untersucht wurden, zusammenfasst für alle Diagnosegruppen eine Grundrate in der Größenordnung von 28,7% bis 32,7%. Der höchste Anteil negativer Antwortverzerrungen entfiel dabei auf die Gruppe der Patienten nach leichter Kopfverletzung; es folgten Patienten mit Fibromyalgie bzw. chronischem Müdigkeitssyndrom und solche mit chronischen Schmerzen oder einer somatoformen Störung. Im Gegensatz dazu wurde unter medizinischen oder psychiatrischen Patienten ohne offensichtlichen sekundären Krankheitsgewinn eine Grundrate von nur 8,1% gefunden. Larrabee (7) ermittelte in einem Überblick relevanter Studien eine mittlere Grundrate von 40% (548 von 1363 Patienten). Schmand und Kollegen konnten in einer Stichprobe von Patienten nach so genanntem Schleudertrauma der Halswirbelsäule sogar eine Rate von 61% nachweisen (15).

Schlussfolgerung

Angesichts der drei exemplarisch ausgewählten Forschungsergebnisse liegt der Schluss nahe, dass die Beschäftigung mit der Diagnostik von suboptimalem Leistungsverhalten nicht nur sinnvoll, sondern auch in hohem Maße notwendig ist. Viele Aspekte sprechen dafür, die Messung der Leistungsmotivation zum festen Bestandteil neuropsychologischer Untersuchungen zu machen, denn neben finanziellen Motiven für suboptimales Leistungsverhalten sind auch zahlreiche innerpsychische anzunehmen, so etwa: von Verantwortung entbun-

den zu sein, Anerkennung der eigenen Leiden zu erhalten o. Ä. Besonders häufig ist mit negativen Antwortverzerrungen allerdings in Situationen zu rechnen, in denen es für den Untersuchten um die Beurteilung der strafrechtlichen Verantwortung oder um einen materiellen Gewinn geht, wie in einem Rentenverfahren oder bei einer Entschädigungsforderung. Doch wie kann eine Messung der Leistungsmotivation vorgenommen werden?

Möglichkeiten der Erfassung suboptimalen Leistungsverhaltens

Die traditionelle klinische Diagnostik der Leistungsmotivation beruht, abgesehen von dem subjektiven Urteil des Gutachters, im Wesentlichen auf der Erfassung von Inkonsistenzen. Auch Auffälligkeiten bei Antworten, wie etwa die gehäufte Abgabe fast korrekter Lösungen, gelten als Hinweis auf eine mögliche Antwortverzerrung. Ein Überblick über verschiedene Ansätze der neuropsychologischen Diagnostik bei Verdacht auf suboptimales Leistungsverhalten ist bei Merten (8) zu finden.

In der gegenwärtigen Forschung dominieren zwei Ansätze zur standardisierten Erfassung der Leistungsmotivation, deren Entwicklung in den letzten Jahren parallel vorangetrieben wurde. Zum einen wurden für eine Reihe von bewährten neuropsychologischen Standardverfahren, wie etwa die Wechsler Adult Intelligence Scale (WAIS), den Wisconsin Card Sorting Test (WCST), den

Trail Making Test (TMT) oder verschiedene verbale Gedächtnistests simulationssensible Maße entwickelt, die Auffälligkeiten im Antwortverhalten eines Probanden ohne größeren diagnostischen Mehraufwand erfassen sollen. Zum anderen ist eine Gruppe von neuen Verfahren entstanden, die spezifisch die Messung der Leistungsmotivation zum Ziel hat. Dies sind die bereits erwähnten Symptomvalidierungstests (SVT).

Für die aus Standardverfahren gewonnenen Parameter konnten bisher nicht immer überzeugende Ergebnisse gefunden werden. Ein Grund dafür könnte darin liegen, dass diese Verfahren aufgrund ihrer eigentlichen Bestimmung sensitiv für kognitive Störungen sind. Das bedeutet, dass ein schlechtes Ergebnis einerseits durch eine kognitive Beeinträchtigung, andererseits aber auch durch suboptimales Leistungsverhalten begründet sein kann. Die klare Trennung dieser beiden konfundierenden Faktoren ist dabei ein schwer zu lösendes Problem.

Demgegenüber sind die für eine erfolgreiche Lösung eines SVT notwendigen kognitiven Anforderungen im Regelfall sehr gering, sodass solche Verfahren auch durch Personen mit tatsächlichen Beeinträchtigungen perfekt oder nahezu perfekt gelöst werden können. Es wird hierbei zugunsten der Messung der Leistungsmotivation auf eine Erhebung der Leistungsfähigkeit weitgehend verzichtet. Dadurch kann eine deutlich bessere Klassifikationsgüte erzielt werden. Eine Vielzahl positiver

empirischer Befunde hat dazu geführt, dass die SVT aktuell als die bestuntersuchten und bestgeeigneten Instrumente zur Diagnostik von negativen Antwortverzerrungen gelten.

Dies kann anhand der Ergebnisse einer kleinen in Deutschland durchgeführten Studie (10) veranschaulicht werden, bei der eine Gruppe von „experimentellen Simulanten“ in einer neuropsychologischen Untersuchung überzeugend Gedächtnisstörungen vortäuschen sollte, während die Probanden der Kontrollgruppe mit optimaler Motivation arbeiten sollten. In Tabelle 1 sind die erhaltenen Klassifikationsergebnisse für drei Symptomvalidierungstests sowie für zwei Simulationsparameter bekannter Leistungstests dargestellt. Während die drei eingesetzten SVT, der Word Memory Test (1), der Amsterdam Short-Term Memory Test (14) und der Word Completion Memory Test (5) eine ausgezeichnete Klassifikationsgüte aufwiesen, war dies für die beiden anderen untersuchten Simulationsparameter nicht der Fall. Angesichts solcher Ergebnisse ist zu empfehlen, Entscheidungsprozeduren über das Vorliegen suboptimalen Leistungsverhaltens evidenzbasiert zu gestalten. Dies bedeutet auch, dass der Anwender die Validierungsansprüche einzelner SVT und Simulationsindikatoren sehr genau prüfen sollte.

An dieser Stelle soll noch einmal betont werden, dass sich – trotz der gelegentlich verwendeten abgekürzten Formulierungen – „Simulations-tests“ oder „Simulationsforschung“ (malingering research) auf suboptimales Leistungsverhalten in seinen verschiedensten Formen bezieht, von denen die Simulation nur eine spezifische darstellt. Mit Hilfe von Symptomvalidierungstests können Aussagen darüber getroffen werden, ob ein Ergebnis in einem Leistungstest oder ein erhaltenes Testprofil als valide anzusehen oder ob ernsthafte Zweifel an deren Authentizität zu äußern sind. Mit Hilfe der SVT können, von speziellen Fällen abgesehen, keine Aussagen darüber erhalten werden, wie gesteuert oder bewusstseinsnahe die Antwortverzerrungen erfolgten und welche Motivation ihnen

Tab. 1 Klassifikationsgüte von drei Symptomvalidierungstests und zwei Simulationsparametern

Testverfahren	Korrekte Klassifikation	Kontroll- personen	Exp. Simulanten	Gesamt
Symptomvalidierungstests:				
Amsterdam Short-Term Memory Test	90%	100%	100%	95%
Word Memory Test – Parameter 1	100%	100%	100%	100%
Word Memory Test – Parameter 2	100%	100%	100%	100%
Word Memory Test – Parameter 3	100%	100%	100%	100%
Word Completion Memory Test – Parameter 1	100%	90%	100%	95%
Word Completion Memory Test – Parameter 2	100%	100%	100%	100%
Simulationsparameter aus bewährten neuropsychologischen Tests:				
Trail Making Test, Verhältnis TMT-A : TMT-B	90%	20%	100%	55%
Raven Standard Progressive Matrices	88%	33%	100%	56%
				nach (10)

zugrunde lag. Dies muss jedem Anwender eines SVT klar sein und bei der Befunddarstellung sorgfältig berücksichtigt werden (vgl. für eine ausführliche Diskussion dazu [9]).

In ähnlicher Weise ist auch der Begriff „Motivationstest“ (effort test) in genauer Bestimmung nicht ausreichend korrekt, denn die Motivation, ein ungünstiges Testprofil zu erzeugen, kann ausgeprägt hoch sein, sodass ein schlechtes Ergebnis in einem SVT nicht unbesehen mit einem Motivationsdefizit gleichgesetzt werden kann.

■ Simulationsforschung in den deutschsprachigen Ländern

Im Gegensatz zu der nur noch schwer zu überblickenden internationalen Literatur zur neuropsychologischen Diagnostik von Simulation und anderen Formen suboptimalen Leistungsverhaltens ist bislang in den deutschsprachigen Ländern ein nur zögerliches Interesse und ein auffälliger Mangel an relevanten Publikationen zu verzeichnen. Häufig wird zur Motivationsbeurteilung entweder auf das klinische Urteil vertraut oder allein auf den 15-Item-Test von Rey (12) zurückgegriffen, der in einer Gegenüberstellung verschiedener SVT so gut wie keinem relevanten Gütemaßstab entspricht (3).

Wie wenig letztlich das Thema überhaupt im deutschen Sprachraum eine Rolle spielt, wird auch durch eine Recherche in der Newsgroup der Gesellschaft für Neuropsychologie deutlich. Im Zeitraum von drei Jahren (2001 bis 2003) tauchte das Stichwort „Simulation“ gerade 24-mal auf. Das sind weniger Nennungen, als in der Newsgroup NPSYCH des Medical College of Georgia (www.neurolist.com) manchmal in einer Woche zum Stichwort „Malingering“ abzurufen sind. Im Informationssystem PSYINDEX-plus, das deutschsprachige psychologische Veröffentlichungen und solche deutscher Autoren in anderen Sprachen recherchiert, konnten zu diesem Stichwort im Zeitraum von 1986 bis Januar 2004 gerade 21 Artikel ermittelt werden.

Dennoch ist für die nächsten Jahre zu erwarten, dass sich mit der besseren Verfügbarkeit sowohl deutschsprachiger Adaptationen in-

ternationaler Verfahren wie auch mit einigen Neuentwicklungen diese Situation grundlegend verändern wird. Der Einsatz spezifischer Verfahren zur Erfassung suboptimalen Leistungsverhaltens bietet dem Diagnostiker, aber auch dem Forscher, die Möglichkeit, eine wesentliche Gefahrenquelle für die Validität neuropsychologischer Daten zu kontrollieren. Erst dadurch wird die inzwischen weithin akzeptierte Feststellung von Iverson und Binder (6) realisiert, dass eine Messung der Testmotivation im Rahmen von Begutachtungen eine notwendige Bedingung für gültige Aussagen darstellt. Damit dürfte der Weg für eine validere neuropsychologische Diagnostik vorgezeichnet sein.

The Assessment of Suboptimal Test Performance in Clinical Neuropsychology: Current Challenges
Negative response bias is presently one of the greatest challenges for the validity of neuropsychological test results. Therefore, the measurement of effort is becoming an integral part of neuropsychological assessment in a number of English-speaking countries. In the past years, standardized methods of effort testing were developed, and Symptom Validity Tests have become the most prominent method in this area. In contrast, the situation in German-speaking countries is marked by a substantial lack of publications and tests for measuring test motivation. However, this situation is likely to change substantially within the years to come.

Key words

malingering – symptom validity testing – neuropsychological assessment – negative response bias – memory

Literatur

1. Green P. Green's Word Memory Test. User's Manual. Edmonton, Canada: Green's Publishing 2003
2. Green P, Rohling ML, Lees-Haley PR, Allen LM. Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Inj* 2001; 15: 1045–1060
3. Hartman DE. The unexamined lie is a lie worth fibbing. *Neuropsychological malingering and the Word Memory Test. Arch Clin Neuropsychol* 2002; 17: 709–714

4. Heaton RK, Smith HH, Lehman RAW, Vogt AT. Prospects for faking believable deficits on neuropsychological testing. *J Consult Clin Psychol* 1978; 46: 892–900
5. Hilsabeck RC, LeCompte DC, Marks AR, Grafman J. The Word Completion Memory Test (WCMT): a new test to detect malingered memory deficits. *Arch Clin Neuropsychol* 2001; 16: 669–677
6. Iverson GL, Binder LM. Detecting exaggeration and malingering in neuropsychological assessment. *J Head Trauma Rehab* 2000; 15: 829–858
7. Larrabee GJ. Detection of malingering using atypical performance patterns on standard neuropsychological tests. *Clin Neuropsychol* 2003; 17: 410–425
8. Merten T. Fragen der neuropsychologischen Diagnostik bei Simulationsverdacht. *Fortschr Neurol Psychiatr* 2002; 70: 126–138
9. Merten T. Der Stellenwert der Symptomvalidierung in der neuropsychologischen Begutachtung. Eine Positionsbestimmung. *Z Neuropsychol* 2005; 16: 29–45
10. Merten T, Henry M, Hilsabeck R. Symptomvalidierungstests in der neuropsychologischen Diagnostik: eine Analogstudie. *Z Neuropsychol* 2004; 15: 81–90
11. Mittenberg W, Patton C, Canyock EM, Condit DC. Base rates of malingering and symptom exaggeration. *J Clin Exp Neuropsychol* 2002; 24: 1094–1102
12. Rey A. L'examen clinique en psychologie. Paris: Presses Universitaires de Paris 1958
13. Rogers R, Harrell EH, Liff CD. Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clin Psychol Rev* 1993; 13: 255–275
14. Schagen S, Schmand B, de Sterke S, Lindeboom J. Amsterdam Short-Term Memory test: A new procedure for the detection of feigned memory deficits. *J Clin Exp Neuropsychol* 1997; 19: 43–51
15. Schmand B, Lindeboom J, Schagen S, Heijt R, Koene T, Hamburger HL. Cognitive complaints in patients after whiplash injury: the impact of malingering. *J Neurol Neurosurg Psychiatr* 1998; 64: 339–343

Korrespondenzadresse:

Dipl.-Psych. Matthias Henry
Otto-von-Guericke-Universität Magdeburg
Klinik für Neurologie
Sektion Neuropsychologie
Leipziger Straße 44
39120 Magdeburg
matthias.henry@medizin.uni-magdeburg.de