

# Entwicklung einer deutschen Kurzform des Boston Naming Test

Th. Merten

*Vivantes Netzwerk für Gesundheit, Klinikum im Friedrichshain, Berlin*

## Zusammenfassung

Die 60-Item-Originalversion des Boston Naming Test (BNT) wird auf ihre Anwendbarkeit im deutschen Sprachraum hin empirisch überprüft. Grundlage sind die Testprotokolle von 200 neurologischen Patienten mit einem Durchschnittsalter von 57,9 Jahren ( $SD = 16,7$ ), die zur neuropsychologischen Diagnostik vorgestellt wurden. Der Gesamtwert betrug 48,3 ( $SD = 11,3$ ). – Auf der Grundlage einer detaillierten Itemanalyse und nach Eliminierung zweifelhafter Items wird eine 30-Item-Kurzform für die Verwendung im deutschen Sprachraum vorgeschlagen. Reliabilitätsschätzungen für beide Formen fallen gut aus (Split-Half jeweils 0,96); Rangkorrelationen mit zahlreichen anderen Tests werden dargestellt. Die höchste Korrelation (0,86) erscheint zwischen der 60-Item-Form des BNT und dem Semantischen Altgedächtnisinventar von *Schmidtke* und *Vollmer-Schmolck* [32].

Die Reliabilitäts- und Validitätsvorhersagen für die Kurzform müssen anhand einer weiteren Stichprobe überprüft werden, wobei nur die Items der vorgeschlagenen Kurzform in der ermittelten Schwierigkeitsrangfolge dargeboten werden.

**Schlagwörter:** Boston Naming Test, Neuropsychologische Diagnostik, Tests, Aphasie

## Development of a German Short Version of the Boston Naming Test

Th. Merten

### Abstract

An item analysis of the Boston Naming Test (BNT) was performed using the test protocols of 200 German-speaking neurological patients with an average age of 57.9 years ( $SD = 16.7$ ) who had been referred for neuropsychological assessment. The sample yielded a mean BNT test score of 48.3 ( $SD = 11.3$ ). – On the basis of an item analysis and the elimination of doubtful items, a German 30-item short version of the BNT is proposed. The reliability analysis yielded a Spearman-Brown corrected odd-even correlation of 0.96 for both versions. Rank correlations with numerous other neuropsychological instruments were computed. The highest correlation (0.86) was obtained between the 60-item test version and the Semantic Memory Test by *Schmidtke* and *Vollmer-Schmolck* [32]. The short version still has to be tested in its own right, in a separate sample and with the item ranking corresponding to the order of difficulty.

**Key words:** Boston Naming Test, neuropsychological assessment, tests, aphasia

© Hippocampus Verlag 2004

## Einleitung

Der Boston Naming Test [14] gehört zu jenen wenigen Tests, die sich einen wichtigen Platz im »Handbook of Normative Data« von *Mitrushina*, *Boone* und *D'Elia* [25] sichern konnten. Damit wurde die große Bedeutung dieses Tests in der nordamerikanischen und internationalen klinischen Neuropsychologie gefestigt. In der Tat ist in der Gegenwart neben einer breiten klinischen Anwendung, die bei weitem nicht auf die Untersuchung aphasischer Benennungsstörungen beschränkt ist, eine große Anzahl von Untersuchungen mit diesem Test zu verzeichnen, darunter

auch Arbeiten, die auf fremdsprachigen Testversionen beruhen [4, 17, 20]. Dabei ist die Übertragung eines solchen Tests, der die Wortbenennung und damit direkt Sprachfunktionen prüft, nicht unbesehen als unkompliziert einzuschätzen (vgl. [24]). So musste beispielsweise das Testmaterial für die koreanische Version grundlegend verändert werden, zahlreiche Items erwiesen sich für diesen Kulturkreis als nicht angemessen, zahlreiche andere schnitten in Itemanalysen gut ab und wurden in den Test aufgenommen [17]. Aber bereits innerhalb des englischen Sprachraums wurden rasch einzelne Probleme deutlich, etwa in englischsprachigen Kulturen außerhalb der USA (z. B. schlugen *Worrall*,

*Yiu, Hickson* und *Barnett* 1995 einige Itemersetzungen für die Anwendung in Australien vor [38]), und auch die Anwendbarkeit bei nationalen Minderheiten ist wiederholt untersucht worden (z. B. [1, 11]). Andererseits stellten *Mariën* et al. [20] eine überraschend gute Korrespondenz der beobachteten Gruppenmittelwerte von englisch- und flämischsprachigen älteren Probanden fest.

In jedem Falle scheint der Test wichtige praktische Bedürfnisse klinischer Neuropsychologen auch außerhalb des englischen Sprachraums, den deutschen inbegriffen (z. B. [18]), zu erfüllen. Drei wesentliche Gründe können dafür geltend gemacht werden: Erstens wird dadurch die große Bedeutung der nordamerikanischen Psychologie für die Gegenwart dokumentiert, die auf vielen Gebieten Entwicklungen wesentlich bestimmt. Durch den Einsatz von äquivalenten Testverfahren wird zweitens eine Vielzahl von vorhandenen Forschungsergebnissen nutzbar gemacht, drittens werden aber eigene Forschungsbeiträge einfacher international kommunizier- und publizierbar. Über Nachteile oder Einschränkungen, die bei der Adaptation ausländischer Verfahren in Kauf zu nehmen sind, wird in vielen Kontexten dagegen kaum reflektiert.

Spezifisch für diesen Test, der vor dem Hintergrund des sogenannten Boston Process Approach [37] entwickelt wurde, einem einflussreichen methodischen Zugang zur neuropsychologischen Diagnostik, fehlt eine vergleichbare eigene Entwicklung im deutschen Sprachraum. Der Untertest »Benennen« des Aachener Aphasie Tests (AAT: [12]) umfasst neben der Benennung von Gegenständen auch die von Farben und die Beschreibung von Situationen; er zielt, im Gegensatz zum BNT, spezifisch auf die Erfassung von Benennungsstörungen bei aphasischen Standardsyndromen ab, wird im Kontext des gesamten AAT ausgewertet und kann deshalb im Regelfall nicht als äquivalente Anforderung für den deutschen Sprachraum herangezogen werden.

Dennoch ist wegen der besonderen Bedeutung von Störungen des Benennens im Rahmen aphasischer Syndrome der Boston Naming Test in die Boston Diagnostic Aphasia Examination (BDAE: [8]) integriert worden. Für den BNT wird eine interne Konsistenz (Cronbachs Alpha) von 0,98 angegeben. Das Testmaterial enthält auch eine 15-Item-Kurzform, die ursprünglich von *Mack, Freed, Williams* und *Henderson* [19] entwickelt wurde und die mit der Langform in Höhe von 0,97 korreliert.

Der BNT in seiner heute verwandten 60-Item-Version [14] geht auf eine experimentelle 85-Item-Fassung [13] zurück. Wegen seiner unzureichenden Eignung im Rahmen zeitökonomischer Reihenuntersuchungen, beispielsweise zur Demenzdiagnostik, sind verschiedene Kurzformen vorgeschlagen worden. Eine Überblicksarbeit von *Kent* und *Luszcz* [16] nennt insgesamt sieben 30- und sechs 15-Item-Kurzformen im englischen Sprachraum.

Normative Daten verschiedener 15-Item-Versionen wurden für das höhere Lebensalter von *Kent* und *Luszcz* [16] erhoben. *Fisher, Tierney, Snow* und *Szalai* [6] untersuchten zwei 30-Item-Kurzformen des BNT, die aus den Items mit

geraden vs. ungeraden Itemnummern gebildet wurden, und fanden eine gute Äquivalenz der Formen.

Weit verbreitet ist eine 15-Item-Kurzform, die Bestandteil der Testbatterie CERAD (Consortium to Establish a Registry for Alzheimer's Disease) ist und darüber auch ins Deutsche Eingang gefunden hat ([35], vgl. [29]). Die CERAD-Version liefert aufgrund ihrer spezifischen Itemzusammensetzung keine ausreichende Variation in Populationen, die nicht aphasisch sind oder nicht wenigstens eine leichte Demenz aufweisen. Überdies ist die deutsche CERAD-Version des BNT ohne jegliche Adaptation aus dem Amerikanischen übernommen worden, die Itemreihenfolge entspricht für das Deutsche nicht der Schwierigkeitsrangfolge, und wenigstens für das Item »Kanu« sind kulturspezifische Unterschiede als wahrscheinlich anzunehmen. Die CERAD-Version ist aufgrund ihrer Kürze und einer Auswahl nicht zu schwerer Items kaum geeignet, diskrete Wortfindungsstörungen adäquat abzubilden oder gar reliabel im Gesundenbereich das Merkmal Benennungsleistung zu erfassen. Ihr Einsatz muss im wesentlichen auf die Siebtestfunktion im Rahmen der Demenzdiagnostik beschränkt bleiben.

Ziel der hier vorgelegten Arbeit sollte es sein, anhand einer unausgelesenen Patientenstichprobe eine empirisch begründete Kurzform des BNT zu entwickeln, die (1) eine hinreichende Variation der Itemschwierigkeit beibehält, indem (2) kulturell oder linguistisch problematische Items aus dem Test eliminiert werden, so dass (3) durch die Reduktion auf 30 Items eine ökonomische Testform mit guter Anwendbarkeit im deutschen Sprachraum erhalten wird.

## Methodik

### Stichprobe

Es wurde eine retrograde Datenanalyse über alle Testprotokolle des BNT vorgenommen, die im Zeitraum von Februar 2002 bis Juni 2004 vorlagen. Die Daten stammen damit von einer unausgelesenen, heterogenen Stichprobe von 200 deutschsprachigen neurologischen Patienten, die zur neuropsychologischen Diagnostik vorgestellt worden waren. Es handelte sich um 118 männliche und 82 weibliche Patienten im Alter von 15 bis 88 Jahren (Durchschnittsalter: 57,9 Jahre; SD = 16,7) und mit durchschnittlich 14,0 Bildungsjahren (SD = 4,2). Bezüglich ihrer primären klinischen Diagnosen wies die Stichprobe folgende Zusammensetzung auf: 58,5 % zerebrovaskuläre Erkrankungen, 11 % Schädel-Hirn-Verletzungen, 10 % degenerative Erkrankungen, 9 % entzündliche Erkrankungen, 4 % toxische Ätiologien, 2,5 % Tumoren des Gehirns oder seiner Häute sowie Metastasierungen hirnfremder Tumoren, 5 % Sonstige.

### Instrumente

Die Durchführung des Boston Naming Test lehnte sich eng an die Vorgaben des Originals an. Es wurde die 60-Item-Testversion unter Verwendung des spiralgebundenen

Testheftes im Format 16 x 16 cm verwendet. Das Testmaterial besteht aus schwarz-weißen Zeichnungen von Objekten oder Lebewesen (z. B. »Biber«, »Mundharmonika«, »Pyramide«), die in der Reihenfolge der Benennungsschwierigkeit im Englischen angeordnet sind und von hochfrequenten, einfachen Wörtern wie »Bett« (Item #1) bis zu selten benutzten wie »Abakus« (Item #60) reichen. Falls eine falsche Antwort gegeben wird, ist der Proband aufzufordern, nach einer geeigneteren Bezeichnung zu suchen. Ist nach 20 Sekunden die richtige Antwort nicht gegeben oder wird zuvor signalisiert, dass das Bild nicht erkannt oder der Name des Objektes nicht gewusst wird, wird eine semantische Hilfe geliefert (z. B. »ein Tier«, »ein Musikinstrument«, »in Ägypten«). Wenn nach weiteren 20 Sekunden die korrekte Lösung noch nicht vorliegt, folgt eine phonematische Hilfe (z. B. »Bi«, »Mu«, »Py«). – Die Antwortabgabe erfolgte mündlich, die Antworten wurden durch den Testleiter protokolliert.

Entsprechend der Vorgaben für die normative Auswertung des Originals gilt eine Antwort als richtig, wenn sie spontan oder nach Liefern des semantischen Hinweises gegeben wurde.

Zu beachten ist, dass bei Erwachsenen im Regelfall (nicht aber beim Vorliegen einer Aphasie) mit der Präsentation des Items #30 (»Mundharmonika«) begonnen wird. Wenn unter den ersten 8 Aufgaben ein Fehler begangen wird, wird mit Item #29 fortgesetzt und dann absteigend weiter, bis 8 Aufgaben in Folge richtig beantwortet sind. Erst dann kann mit den schwierigeren Items fortgefahren werden. Abbruchkriterium sind 6 aufeinanderfolgende falsche Antworten.

Im Kontext der durchgeführten neuropsychologischen Diagnostik waren zahlreiche weitere Daten verfügbar. Deren Menge variierte entsprechend Fragestellung, Art und Ausmaß der neuropsychologischen Defizite sowie der für die Diagnostik zur Verfügung stehenden Zeit. In die vorliegende Analyse wurden Testdaten dann aufgenommen, wenn sie für mindestens 50 Prozent der Patienten verfügbar waren. Dies war für die folgenden standardisierten Tests der Fall:

- den Mini-Mental State (MMS: [7]) in der CERAD-Version [35];
- den Syndrom-Kurztest (SKT: [5]);
- das Zahlen-Nachsprechen und die Blockspanne des Wechsler-Gedächtnistests [9];
- den Mosaik-Test des HAWIE-R [34];
- eine von Merten [22] entwickelte Kurzform des Hooper Visual Organization Test (VOT: [36]);
- die Untertests Alertness und Go/Nogo aus der Testbatterie zur Aufmerksamkeitsprüfung (TAP: [39]);
- den Trail Making Test (TMT: [28]);
- den Memo-Test [30], einen kurzen Wortlisten-Lerntest, der nach dem Prinzip des »selective reminding« von Buschke [3] aufgebaut ist;
- den Complex Figure Test (CFT: [27]) unter Verwendung der quantitativen Auswertung nach Taylor [33] und der qualitativen Auswertung nach Kenkel, Gottwald und Mehdorn [15];

- den Verbalen Lern- und Merkfähigkeitstest (VLMT: [10]);
- die Computerversion des Aufmerksamkeits-Belastungstests d2 [2];
- den Token Test in der Fassung von Orgass [26];
- zwei Aufgaben zur Erfassung der verbalen Produktionsgeschwindigkeit, und zwar formal und semantisch gebunden (FAS, Tiere assoziieren; vgl. [23]);
- den Wortschatztest (WST) von Schmidt und Metzler [31].

Obgleich das genannte Mengenkriterium nicht zutraf und nur 51 Probanden berücksichtigt werden konnten, wurden wegen der besonderen konzeptionellen Verbindung auch die Ergebnisse im Semantischen Altgedächtnisinventar von Schmidtke und Vollmer-Schmolck [32] berücksichtigt.

### Datenanalyse

Die Antworten der Patienten auf die 60 Items des BNT wurden elektronisch transkribiert. Items, die am Testanfang nicht dargeboten werden mussten, wurden entsprechend dem vorgeschriebenen Procedere als richtig gewertet, aber gesondert gekennzeichnet. Gleiches gilt für als falsch gewertete Items nach Erreichen des Abbruchkriteriums von 6 Falschen.

Bei der Bewertung wurden grundsätzlich zweifelsfreie Synonyme als richtig gewertet (z. B. »Helikopter« #11, »Larve« #18, »Rhinozeros« #31, »Dart« #25, »Hörrohr« #42), ferner Diminutiva (z. B. »Häuschen« #4, »Blümchen« #8). Bei Item #24 wurde zwar, obgleich nicht korrekt, »Seepferd« (statt »Seepferdchen«) als richtig bewertet, jedoch ist dieser auch in der Allgemeinbevölkerung wohl nicht ganz seltene Fehler ein gewichtiger Grund, das Item als für das Deutsche ungeeignet auszuschließen. Ferner galten als richtig Unterbegriffe des Zielwortes, sofern sie nicht offenkundig grob falsch waren (z. B. »Laubbaum« #2, »Trillerpfeife« #5, »Gartenblume« #8, »Stroh-, Hexenbesen« #12, »Steinpilz« #14, »Handsäge« oder »Fuchsschwanz« #9, »Parkbank« #20, »Spitzmaulnashorn« #31; nicht aber »Ätna« oder »Vesuv« bei Item #23, »Vulkan«).

Damit die zu erhaltende Testkurzform bezüglich der Itemschwierigkeit möglichst das gesamte vorhandene Schwierigkeitsspektrum umfasst, wurden für die Itemselektion 15 Itemmengen mit je vier aufeinanderfolgenden Aufgaben gebildet, unter denen zwei auszuwählen und zwei zu verwerfen waren. Auswahlkriterien waren zweifelhafte Bewertungskriterien (z. B. Welche deutschen Synonyma gelten als richtig bei »Spalier«, »Abakus« oder »Stativ«?, Welche Synonyma, eng verwandten Instrumente und umgangssprachlichen Bezeichnungen können bei »Ziehharmonika« noch als richtig gewertet werden?), Schwierigkeiten beim phonematischen Hinweis, zu geringe Item-Gesamtwert-Korrelationen oder inadäquater kultureller Bezug (»Türklopfer«). Das Item #48 (»Schlinge«) wurde aus Pietätsgründen verworfen. Nicht nur in unserem kulturellen Kontext stößt die Darbietung eines Henkerstricks bei einer Reihe von Probanden auf Ablehnung, wie entsprechende Diskussionen

Nr.	Original-Item-Nr.	Bezeichnung	Schwierigkeit (Prozent richtig)	Item-Gesamtwert-Korrelation <sup>1</sup>
01	01	Bett	98,5	0,50
02	07	Kamm	98,5	0,51
03	03	Bleistift	97,5	0,57
04	09	Säge	96,5	0,56
05	21	Schläger	95,0	0,60
06	05	Pfeife	94,5	0,65
07	11	Hubschrauber	94,0	0,62
08	18	Maske	93,5	0,54
09	28	Kranz	92,0	0,54
10	16	Rollstuhl	92,0	0,61
11	19	Brezel	91,5	0,60
12	37	Rolltreppe	90,0	0,66
13	46	Trichter	89,5	0,70
14	13	Krake	87,5	0,67
15	27	Globus	86,5	0,60
16	43	Pyramide	85,5	0,69
17	35	Dominos	83,5	0,54
18	49	Spargel	83,0	0,66
19	30	Mundharmonika	82,0	0,58
20	23	Vulkan	81,0	0,71
21	39	Hängematte	79,5	0,53
22	29	Biber	76,0	0,70
23	33	Iglu	76,0	0,65
24	54	Zange	75,0	0,37
25	58	Palette	71,5	0,59
26	51	Riegel	70,5	0,60
27	45	Einhorn	70,0	0,59
28	41	Pelikan	69,0	0,58
29	59	Winkelmesser	49,5	0,46
30	55	Sphinx	47,5	0,56

**Tab. 1:** Itemliste für die vorgeschlagene deutsche Kurzform des BNT mit den Ergebnissen der Itemanalyse an 200 neurologischen Patienten. <sup>1</sup>korrigiert, d. h. unter Ausschluss des jeweils betrachteten Items.

in Fachforen belegen. So wird immer wieder geäußert, das Item solle auch aus der Originalform des BNT entfernt werden. – Wenn unter den jeweils vier betrachteten Items mehrere als geeignet zur Auswahl standen, wurde mit Hilfe einer Zufallstabelle die Wahl getroffen.

Auf diese Weise konnte eine 30-Item-Kurzform zusammengestellt werden, die aus rationalen und itemanalytischen Gründen zahlreiche Probleme ausschließt, die bei der Verwendung der amerikanischen Testvorlage im Deutschen die Testgüte deutlich einschränken.

**Ergebnisse**

Der Testwert im BNT betrug für die Gesamtgruppe im Durchschnitt 48,3 Punkte (SD = 11,3) und bewegte sich zwischen 0 und 60. Eine Einweg-Varianzanalyse ergab einen signifikanten Einfluss des Geschlechts auf die Testleistung (F = 11,6; FG = 1; p < 0,05), wobei Frauen im Mittel niedrigere Punktwerte (m = 45,1; SD = 13,0) als

Nr.	Bezeichnung	Semantischer Hinweis	phonematischer Hinweis	Ebenfalls als richtig zu wertende Antworten <sup>1</sup>
01	Bett	ein Möbelstück	Be	
02	Kamm	für das Haar	Ka	
03	Bleistift	zum Schreiben	Blei	
04	Säge	benutzt der Zimmermann	Sä	Handsäge, Fuchsschwanz
05	Schläger	für Sport	Schlä	Tennisschläger
06	Pfeife	macht einen Ton	Pfei	Trillerpfeife
07	Hubschrauber	zum Fliegen	Hu	Helikopter
08	Maske	gehört zur Verkleidung	Ma	Larve
09	Kranz	ein Weihnachtsschmuck	Kra	Trauerkranz
10	Rollstuhl	gibt es im Krankenhaus	Ro	Krankenstuhl
11	Brezel	etwas zum Essen	Bre	
12	Rolltreppe	damit fährt man hoch	Ro	
13	Trichter	zum Einfüllen	Tri	
14	Krake	ein Meerestier	Kra	Tintenfisch, Oktopus
15	Globus	für Erdkunde	Glo	
16	Pyramide	in Ägypten	Py	
17	Dominos	ein Spiel	Do	Domino, Dominosteine
18	Spargel	zum Essen	Spa	Spargelstange
19	Mundharmonika	ein Musikinstrument	Mu	
20	Vulkan	eine Art Berg	Vu	
21	Hängematte	da liegt man drin	Hä	
22	Biber	ein Tier	Bi	
23	Iglu	eine Art Haus	I	
24	(Zucker-)Zange	für den Hausgebrauch	Za	Eiszange
25	Palette	vom Maler benutzt	Pa	Farb-, Maler-, Mischpalette
26	Riegel	an der Tür	Rie	Türriegel, Sperrriegel
27	Einhorn	ein Fabeltier	Ei	
28	Pelikan	ein Vogel	Pe	
29	Winkelmesser	für Geometrie	Wi	
30	Sphinx	gibt es in Ägypten	Sf	

**Tab. 2:** Präsentation der Items in der vorgeschlagenen deutschen Kurzform. Bei der Präsentation der phonematischen Hinweise ist zu beachten, dass die Lautstruktur des Zielwortes gegeben wird, also zum Beispiel [do:] bei Domino. <sup>1</sup>Als richtig zu werten sind ebenfalls Diminutiva sowie Unterbegriffe der Zielworte, sofern diese nicht offenkundig grob falsch sind (Beispiele im Text).

Männer (m = 50,5; SD = 9,3) erreichten. Die Maßkorrelation mit der Variablen »Bildung« (Anzahl der Bildungsjahre) betrug 0,19, die mit dem Alter -0,28 (p jeweils < 0,05). Auch wenn eine Korrektur des Einflusses des Geschlechts um die erhaltene Korrelation mit den Variablen »Bildung« und »Alter« vorgenommen wird, bleibt der Unterschied in der Testleistung zwischen Männern und Frauen nachweisbar (Kovarianzanalyse, Haupteffekt: F = 9,59; FG = 1; p < 0,05).

Auch eine andere mögliche Erklärung für den erhaltenen Geschlechtsunterschied lässt sich anhand der Daten nicht belegen: Es könnte sein, dass die vorgestellten Frauen insge-



samt schwerere kognitive Störungen aufweisen, als dies für die Männer der Fall ist. Wenn, trotz aller Beschränkung dieses Vorgehens (vgl. [21]), der Gesamtwert im SKT als globales Maß für die Schwere neuropsychologischer Beeinträchtigungen herangezogen wird und eine weitere Kovarianzanalyse mit dieser Variablen als Kovariablen gerechnet wird, bleibt der Effekt des Faktors Geschlecht auf den BNT-Punktwert bestehen (Haupteffekt:  $F = 14,4$ ;  $FG = 1$ ;  $p < 0,05$ ).

Die Einzelaufgaben der 60-Item-Version korrelieren im Durchschnitt in Höhe von 0,33 miteinander. Durch die Zusammenstellung einer 30-Item-Kurzform wird diese mittlere Korrelation zwischen den Items auf 0,38 gesteigert.

In Tabelle 1 sind die Ergebnisse der Itemselektion wiedergegeben. Die ausgewählten Items sind in der Schwierigkeitsrangfolge dargestellt. In dieser Reihenfolge sollten sie auch in der hier vorgeschlagenen deutschen Kurzform präsentiert werden. In Tabelle 2 sind die zugehörigen semantischen und phonematischen Hinweise, die für die einzelnen Items gegeben werden können, zusammengestellt. Diese Tabelle enthält auch eine Aufstellung von zwar selten genannten, aber ebenfalls als richtig zu wertenden alternativen Antworten.

Die Punktwerte der extrahierten Kurzform korrelieren mit denen der Gesamtform in Höhe von 0,98 ( $p < 0,05$ ). Weitere Ergebnisse zu den psychometrischen Eigenschaften der Lang- und der extrahierten Kurzform sind in Tabelle 3 zusammengefasst.

Die Tabelle listet auch die Korrelationen einer Reihe anderer neuropsychologischer Instrumente mit den Skalen auf. Nicht in der Tabelle enthalten, weil auf einer zu geringen Teilstichprobe beruhend, ist die Korrelation mit dem Semantischen Altgedächtnisinventar von *Schmidtke* und *Vollmer-Schmolck* [32], die hier deshalb Erwähnung finden soll, weil beide Instrumente durch ihren Zugriff auf semantisches Wissen hoch miteinander korrelieren sollten. Tatsächlich wird eine Rangkorrelation des Inventars mit der 60-Item-BNT-Form in Höhe von 0,86 ermittelt. Die Vorhersage für die Kurzform beträgt 0,75 ( $p$  jeweils  $< 0,05$ ).

## Diskussion

Die hier vorgeschlagene Kurzform des Boston Naming Tests versucht in erster Linie, zweifelhafte Items auszuschließen, die eine Übertragung des Tests in den deutschen Kultur- und Sprachraum erschweren. Jeder regelmäßige und kritische Anwender des Tests ist sich einer Anzahl dieser Probleme rasch bewusst geworden. Dennoch ist ganz klar ein Bedarf nach einem Instrument wie dem BNT auch hierzulande vorhanden. Eine Adaptation vorzunehmen, könnte auf sehr verschiedene Weise geschehen. Ideal wäre es, von einem größeren Itempool als dem der 60-Item-Form auszugehen und anhand einer oder mehrerer größerer Stichproben eine rationale und empirische Itemanalyse vorzunehmen, deren Ziel der Erhalt einer adaptierten 60-Item-Version mit gegenüber der Originalform äquivalenten Messeigenschaften wäre. Die Items selbst könnten dann durchaus von den originalen abweichen. So enthält etwa

	60-Item-Testform (empirische Analyse)	30-Item-Kurzform (Vorhersage)	n
Testwerte an heterogener neurologischer Stichprobe, m (SD)	48,3 (11,3)	24,9 (6,4)	200
Mittlere Inter-Item-Korrelation	0,33	0,38	
<b>Reliabilitätsindices:</b>			
Cronbachs Alpha	0,96	0,94	
Split-Half (oben vs. unten)	0,89	0,84	
Split-Half (gerade vs. ungerade)	0,96	0,96	
<b>Rangkorrelation mit:</b>			
Mini-Mental State (CERAD)	0,56	0,50	199
Syndrom-Kurztest (Gesamtwert)	-0,56	-0,51	197
Syndrom-Kurztest (UT Benennen)	-0,56	-0,55	197
Zahlen Nachsprechen vorwärts	0,31	0,29	194
Zahlen Nachsprechen rückwärts	0,45	0,42	194
Blockspanne vorwärts	0,42	0,37	184
Blockspanne rückwärts	0,46	0,42	184
Mosaik-Test des HAWIE-R	0,59	0,53	197
Hooper VOT (Kurzform)	0,70	0,67	179
TAP Alertness (ohne Warnton, msec)	-0,38	-0,40	192
TAP Go/Nogo, Reaktionszeit	-0,29	-0,29	122
TAP Go/Nogo, Fehlerzahl	-0,34	-0,30	122
Trail Making Test, Teil A	-0,52	-0,46	191
Trail Making Test, Teil B	-0,54	-0,48	184
d2-Computerversion, GZ-F	0,52	0,47	138
d2-Computerversion, Fehler	-0,24	-0,24	138
Kategorienflüssigkeit (Tiere)	0,58	0,54	150
Verbale Produktionsgeschw. (F-A-S)	0,52	0,47	150
Memo-Test, unmittelbares Behalten	0,45	0,42	189
Memo-Test, Durchgänge 1 bis 5	0,52	0,48	189
Memo-Test, verzögerter Abruf	0,29	0,24	189
VLMT, Liste 1, 1. Durchgang	0,43	0,40	125
VLMT, Liste 1, 5. Durchgang	0,42	0,39	125
VLMT, Liste 2	0,44	0,43	125
VLMT, Wiedererkennung	0,03 ns	0,01 ns	125
Rey CFT, Abzeichnen	0,47	0,42	142
Rey CFT, Abruf nach 30 Minuten	0,49	0,46	142
Rey CFT, Planung (qualitativ)	0,46	0,39	125
Token-Test (Fehler, unkorrigiert)	0,56	0,50	131
Wortschatztest	0,54	0,49	172

**Tab. 3:** Skalenkennwerte für die 60-Item-Form des BNT und berechnete Vorhersagen für die vorgeschlagene deutsche Kurzform. Anmerkung: Alle Korrelationen mit Ausnahme der durch ns gekennzeichneten sind signifikant mit  $p < 0,05$ .

die koreanische Version Aufgaben wie »Harpune«, »Drache«, »Hummer« und »Nixe« [17].

Da dieser Weg nicht zur Verfügung stand, wurde, dem zunehmenden Trend entsprechend, eine Kurzform zusammengestellt, die zugleich das Problem einer sehr langen Testzeit bei deutlich beeinträchtigten Patienten reduziert (bei Patienten mit Aphasie oder Demenz kann die Durchführung der Langform bis zu 30 Minuten oder mehr dauern!). Ein ähnliches Vorgehen wurde übrigens auch für die chinesische Form beschrieben [4].

Bei der Betrachtung der angegebenen Korrelationen ist zu beachten, dass diese eine Funktion der Streuung in den Ausgangsvariablen sind und die Streuung der Kurzform durch die Itemreduktion deutlich geringer ausfällt. Insofern erscheint es bemerkenswert, dass die erhaltene Split-Half-Konsistenz für die Kurzform nicht niedriger als für die Langform ausfällt, was unter Berücksichtigung der Skalenslänge eine reale Konsistenzsteigerung signalisiert. Bei der Betrachtung der konkurrierenden Validität fallen besonders die hohen Korrelationen mit dem Semantischen Altgedächtnisinventar von *Schmidtke* und *Vollmer-Schmolck* [32] auf, was sachlogisch durchaus erklärbar ist. Im Gegensatz zum BNT erfolgt in diesem Inventar jedoch kein Benennen von visuell präsentierten Gegenstandsdarstellungen, sondern Wissen wird direkt abgefragt (z.B. Welches Tier liefert Wolle? Wie heißt die Hauptstadt von Österreich? Welcher Beruf trägt als typische Berufskleidung blau?).

Die erhaltene deutsche 30-Item-Kurzform sollte so durchgeführt werden, dass mit dem Item 1 begonnen wird. Als Abbruchkriterium sollte die bisherige Regel von 6 aufeinanderfolgend nicht richtig beantworteten Aufgaben beibehalten werden.

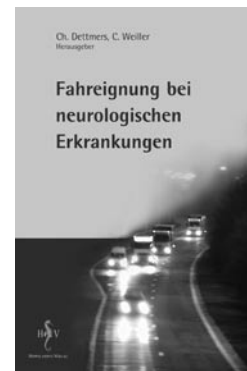
Ein Blick auf die Itemreihfolge nach Schwierigkeit (Tabelle 1) zeigt, dass zahlreiche Items für den deutschen Sprachraum tatsächlich deutlich andere Eigenschaften aufweisen, als dies für das Amerikanische der Fall ist. Für die untersuchte klinische Stichprobe erwies sich das Item #56, »Joch«, mit 15,5% richtigen Antworten als das schwierigste, gefolgt von #60 (»Abakus« oder »Rechenbrett« oder »Zählbrett«, 20,5%). Ähnlich deutliche Abweichungen hatten auch Mariën et al. [20] für eine Stichprobe älterer flämischer Probanden erhalten. Hier waren es Items wie »Brezel«, »Einhorn« oder »Biber«, die klar aus der Schwierigkeitsrangfolge herausstachen, während das im Original schwierigste Item, »Abakus«, von 97,5% der Probanden richtig gelöst wurde.

Einschränkungen ergeben sich insbesondere daraus, dass durch die Darbietungs- und Abbruchregeln der 60-Item-Version, die im klinischen Kontext für diese Vollversion durchaus sinnvoll sind, keine vollständigen Datensätze für die Itemanalyse zur Verfügung standen und für die Mehrzahl der Probanden Items als richtig, seltener auch als falsch gewertet wurden, obgleich sie nicht dargeboten wurden. Daraus ergibt sich, dass bei einer Überprüfung der hier vorgeschlagenen BNT-Kurzform selbst in einer äquivalenten klinischen Stichprobe einzelne Item- und Skalenskennwerte von den hier berechneten Prädiktionen relevant abweichen könnten. Damit wird insbesondere die Darbietungsreihenfolge der Items erneut zu überprüfen sein. Dies sollte aber anhand einer Stichprobe erfolgen, die ebenfalls einen Mindestumfang von 200 haben sollte. Alle in dem hier vorgelegten Artikel gelieferten Kennwerte (z.B. Reliabilitätskoeffizienten) sind lediglich als Vorhersagen zu betrachten und beruhen auf einem *Procedere*, bei dem die Items im Kontext und in der Reihenfolge der 60-Item-Version beantwortet wurden.

## Literatur

1. Barker-Collo SL: The 60-item Boston Naming Test: cultural bias and possible adaptations for New Zealand. *Aphasiology* 2002; 15: 85-92
2. Brickenkamp R, Merten T, Hänsgen KD: d2-C – Aufmerksamkeits-Belastungs-Test, Computerversion [Computer-Software]. Hogrefe, Göttingen 1996
3. Buschke H: Selective reminding for analysis of memory and learning. *J Verb Learn Verb Behav* 1973; 12: 543-546
4. Cheung RW, Cheung MC, Chan AS: Confrontation naming in Chinese patients with left, right or bilateral brain damage. *J Internat Neuropsychol Soc* 2004; 10: 46-53
5. Erzigkeit H: SKT Manual. Kurztest zur Erfassung von Gedächtnis- und Aufmerksamkeitsstörungen. 21. Aufl., Geromed, Witten 1996
6. Fisher NJ, Tierney MC, Snow WG, Szalai JP: Odd/even short forms of the Boston Naming Test: preliminary geriatric norms. *Clin Neuropsychol* 1999; 13: 359-364
7. Folstein MF, Folstein SE, McHugh PR: »Mini-Mental State«: A practical method for grading the cognitive state of patients for the clinician. *J Psychiat Res* 1975; 12: 189-198
8. Goodglass H, Kaplan E, Barresi B: The Assessment of Aphasia and Related Disorders. 3rd ed. Lippincott Williams & Wilkins, Philadelphia 2001
9. Härting C, Markowitsch HJ, Neufeld H, Calabrese P, Deisinger K, Kessler J (Hrsg): WMS-R. Wechsler Gedächtnistest – Revidierte Fassung. Deutsche Adaptation der revidierten Fassung der Wechsler Memory Scale. Manual. Huber, Bern 2000
10. Helmstaedter C, Durwen HF: VLMT: Verbaler Lern- und Merkfähigkeitstest. Ein praktikables und differenziertes Instrumentarium zur Prüfung der verbalen Gedächtnisleistungen. *Schweiz Arch Neurol Psychiat* 1990; 141: 21-30
11. Henderson LW, Frank EM, Pigatt T, Abramson RK, Houston M: Race, gender, and educational level effects on Boston Naming Test scores. *Aphasiology* 1998; 12: 901-911
12. Huber W, Poeck K, Weniger D, Willmes K: Aachener Aphasie Test (AAT). Handanweisung. Hogrefe, Göttingen 1983
13. Kaplan E, Goodglass H, Weintraub S: The Boston Naming Test. Experimental edition. Kaplan & Goodglass, Boston 1978
14. Kaplan E, Goodglass H, Weintraub S: Boston Naming Test. 2nd ed. Lippincott Williams & Wilkins, Philadelphia 2001
15. Kenkel S, Gottwald B, Mehdorn HM: Rey-Osterrieth-Figur: Vergleich qualitativer Auswertungssysteme [Abstract]. *Z Neuropsychol* 1997; 8: 97
16. Kent PS, Luszcz MA: A review of the Boston Naming Test and multiple-occasion normative data for older adults on 15-item versions. *Clin Neuropsychol* 2002; 16: 555-574
17. Kim H, Na DL: Normative data on the Korean version of the Boston Naming Test. *J Clin Exp Neuropsychol* 1999; 21: 127-133
18. Konsensusrichtlinien der Österreichischen Parkinson Gesellschaft zur tiefen Gehirnstimulation bei Morbus Parkinson: P-aktuell – Newsletter der Österreichischen Parkinson Gesellschaft, Sonderausgabe 2003; 2: 1-8
19. Mack WJ, Freed DM, Williams BW, Henderson VW: Boston Naming Test: Shortened version for use in Alzheimer's disease. *J Gerontol* 1992; 47: 164-168
20. Mariën P, Mampaey E, Vervaeke A, Scaerens J, de Deyn PP: Normative data for the Boston Naming Test in native Dutch-speaking Belgian elderly. *Brain Lang* 1998; 65: 447-467
21. Merten T: Über den Sinn und Unsinn der Verwendung von Screening-Instrumenten in der neuropsychologischen Diagnostik. *Diagn* 1999; 45: 154-162
22. Merten T: A Short Version of the Hooper Visual Organization Test: Development and Validation. *Clin Neuropsychol* 2002; 16: 136-144
23. Merten T: Flüssigkeitsmaße in der Neuropsychologie: F-A-S, Kategorien, Zeichnungen und der 5-Punkte-Test. *Neurol Rehab* 2003; 9: 226-232
24. Merten T: Eine Kurzform des Hooper Visual Organization Test: Faktorielle Validität und Retest-Reliabilität. *Z Neuropsychol*, im Druck
25. Mitrushina MN, Boone KB, D'Elia LF: Handbook of normative data for neuropsychological assessment. Oxford University Press, New York 1999
26. Orgass B: Token Test. Manual. Deutsche Bearbeitung des Token Tests von E. de Renzi und L.A. Vignolo. Beltz Test, Weinheim 1982
27. Osterrieth PA: Le test de copie d'une figure complexe. Contribution à l'étude de la perception et de la mémoire. *Arch Psychol* 1945; 30: 205-353

28. Reitan RM: Trail Making Test. Manual for administration and scoring. Reitan Neuropsychology Laboratory, South Tucson 1992
29. Satzger W, Hampel H, Padberg F, Bürger K, Nolde T, Ingrassia G, Engel RR: Zur praktischen Anwendung der CERAD-Testbatterie als neuropsychologisches Demenzscreening. *Nervenarzt* 2001; 72: 196-203
30. Schaaf A, Kessler J, Grond M, Fink GR: Memo-Test. Manual. Beltz Test, Weinheim 1992
31. Schmidt KH, Metzler P: Wortschatztest. Beltz Test, Weinheim 1992
32. Schmidtke K, Vollmer-Scholck H: Autobiographisches Altgedächtnisinterview und semantisches Altgedächtnisinventar. *Z Neuropsychol* 1999; 10: 13-23
33. Taylor EM: Psychological Appraisal of Children with Cerebral Defects. Harvard University Press, Cambridge, MA 1959
34. Tewes U (Hrsg): HAWIE-R: Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision 1991. Handbuch und Testanweisung. 2. korr Aufl. Huber, Bern 1994
35. Thalmann B, Monsch AU, Schneitter M, Ermini-Fünfschilling D, Spiegel R, Stähelin HB: Die CERAD Neuropsychologische Testbatterie. Memory Clinic, Basel 1998
36. Western Psychological Services: Hooper Visual Organization Test (VOT). Manual. WPS, Los Angeles 1983
37. White RF, Rose FE: The Boston Process Approach. A brief history and current practice. In: Goldstein G, Incagnoli TM (eds): Contemporary Approaches to Neuropsychological Assessment. New York: Plenum Press, New York 1997, 171-211
38. Worrall LE, Yiu ML, Hickson LMH, Barnett HM: Normative data for the Boston Naming Test for Australian elderly. *Aphasiology* 1995; 9: 541-551
39. Zimmermann P, Fimm B: Testbatterie zur Aufmerksamkeitsprüfung (TAP). Handbuch Teil 1. Psytest, Würselen 1993



Christian Dettmers, Cornelius Weiller (Hrsg.)

## Fahreignung bei neurologischen Erkrankungen

*Hippocampus Verlag, Bad Honnef 2004,  
172 Seiten, zahlreiche Abbildungen, Broschur  
€ 34,50, ISBN 3-936817-10-3*

Neurologische Erkrankungen haben nicht selten eine Beeinträchtigung der Fahreignung durch motorische und neuropsychologische Funktionsdefizite zur Folge. In diesem Fall stoßen zwei gegensätzliche Interessen aufeinander: Dem Recht des Einzelnen zur Teilnahme am Straßenverkehr steht das Recht aller Verkehrsteilnehmer auf ein möglichst hohes Maß an Verkehrssicherheit gegenüber.

In der Neurologie wird die Frage der Fahreignung immer noch häufig umgangen. Im Studium – und damit auch in den Lehrbüchern – ist sie nicht Gegenstand des Kanons. In der Akutphase der Erkrankung wird sie meist nicht thematisiert, und die Patienten werden zu selten darüber aufgeklärt, dass ihre Erkrankung eine Beeinträchtigung der Fahreignung nach sich ziehen könnte.

Die rechtlichen Bestimmungen versuchen, die körperlichen und geistigen Voraussetzungen zur Teilnahme am Straßenverkehr zu formulieren, lassen jedoch viele Fragen offen. Auch die neuropsychologische Diagnostik hat nur begrenzte Aussagekraft. In der Praxis erweist sich die Zusammenarbeit mit einer spezialisierten Fahrschule in vielen Fällen als hilfreich bzw. unabdingbar.

Die Fahreignung ist für viele Menschen die Voraussetzung zur Teilnahme am sozialen und beruflichen Leben. Die vorliegende Zusammenstellung soll Ärzten und Psychologen in der Akutklinik, in der Rehabilitation und in der Praxis bei der Beurteilung helfen, ob ein Patient die Fahreignung besitzt oder ob seine Gesundheitsstörung diese beeinträchtigt. Auch Juristen, Versicherungsfachleute und Betroffene können von dem Werk profitieren.

### Korrespondenzadresse:

Dr. Thomas Merten  
Vivantes Netzwerk für Gesundheit  
Klinikum im Friedrichshain  
Klinik für Neurologie  
Landsberger Allee 49  
D-10249 Berlin  
e-mail: thomas.merten@vivantes.de