

Nueva versión reducida del test de denominación de Boston para mayores de 65 años: aproximación desde la teoría de respuesta al ítem

Miguel A. Fernández-Blázquez, José M. Ruiz-Sánchez de León, José A. López-Pina, Marcos Llanero-Luque, Mercedes Montenegro-Peña, Pedro Montejo-Carrasco

Introducción. Uno de los tests más utilizados para la evaluación de la afasia en la práctica clínica es el test de denominación de Boston (BNT), una prueba clásica en la que se presentan 60 láminas con dibujos en blanco y negro que valoran la capacidad de los sujetos para denominar dichos dibujos. Pese a su bondad psicométrica, resulta necesario reducir el número de ítems de la prueba con el fin de disminuir su tiempo de aplicación.

Sujetos y métodos. Se reclutó una muestra de 547 sujetos mayores de 65 años a los que se les aplicó un protocolo de evaluación neuropsicológica, que incluía el BNT, para determinar sus estados cognitivos; 405 sujetos no presentaron alteraciones cognitivas relevantes, frente a 142 sujetos que fueron diagnosticados de deterioro cognitivo leve.

Resultados. La reducción del número de ítems se realizó de acuerdo con los supuestos de la teoría de respuesta al ítem. Puesto que tanto la edad como el nivel educativo mostraron un efecto significativo en el rendimiento de la prueba, ambas variables fueron utilizadas para hallar los puntos de corte de la nueva versión reducida. Ésta presentó una adecuada fiabilidad ($\alpha = 0,765$) y una elevada correlación con la prueba original ($r = 0,876$).

Conclusiones. La nueva versión reducida consta de 15 ítems ordenados en función de su dificultad. Se trata de una tarea con un elevado poder discriminativo de utilidad en la clínica diaria para la detección de alteraciones del lenguaje en personas mayores.

Palabras clave. Afasia. Anomia. Demencia. Denominación. Deterioro cognitivo leve. Envejecimiento. Evaluación neuropsicológica. Teoría de respuesta al ítem. Test de denominación de Boston.

Introducción

Los pacientes afásicos presentan una disminución del repertorio de palabras que usan al hablar y un aumento del tiempo que necesitan para recuperarlas, si bien esta alteración puede no afectar demasiado a su fluidez en el discurso espontáneo y su conversación puede parecer prácticamente normal [1,2]. No obstante, algunos pacientes no sólo presentan esa restricción general del vocabulario, sino que, además, presentan una pérdida selectiva de la capacidad para evocar palabras específicas, que se conoce como anomia.

Este déficit en la producción verbal, también llamado afasia anómica [2,3], afasia nominal [4] o afasia amnésica [5], constituye una de las alteraciones más frecuentemente asociadas al envejecimiento normal [6,7], la enfermedad de Alzheimer [8-11] y otros cuadros neurodegenerativos [8,12,13], y es, además, la alteración del lenguaje más frecuente en las afasias secundarias al daño cerebral [14-16]. Esta alteración en la producción del lenguaje puede

describirse y analizarse cualitativamente en el discurso espontáneo. No obstante, su valoración cuantitativa suele realizarse con tests en los que se solicita al individuo que denomine, utilizando el término más preciso posible, un conjunto de láminas que contienen dibujos.

La habilidad para denominar estímulos visuales requiere la integridad de un conjunto de componentes del sistema cognitivo, entre los cuales cabe destacar el reconocimiento perceptivo, la memoria semántica y el almacén lexicofonológico de salida (Fig. 1). La alteración de alguno de estos componentes puede conllevar que el individuo no denomine el estímulo presentado, aunque, en cada caso, la razón por la que no lo haga sea diferente [17,18]. Esquematizando al máximo el número de procesos implicados, para que el individuo pueda denominar un dibujo debe realizar lo siguiente:

- Reconocer perceptivamente el estímulo; como esta representación inicial está carente de información semántica, si resulta errónea el individuo no reconocerá lo que está viendo. Quizá pueda

Departamento de Psicología Básica y Metodología; Universidad de Murcia (J.A. López-Pina). Centro de Prevención del Deterioro Cognitivo; Madrid Salud; Ayuntamiento de Madrid (M.A. Fernández-Blázquez, J.M. Ruiz-Sánchez de León, M. Llanero-Luque, M. Montenegro-Peña, P. Montejo-Carrasco); Madrid, España.

Correspondencia:

Dr. Miguel Ángel Fernández Blázquez. Centro de Prevención del Deterioro Cognitivo. Montesa, 22. Edif. B. E-28006 Madrid.

Fax:

+34 915 886 793.

E-mail:

ma.fdez.blazquez@gmail.com

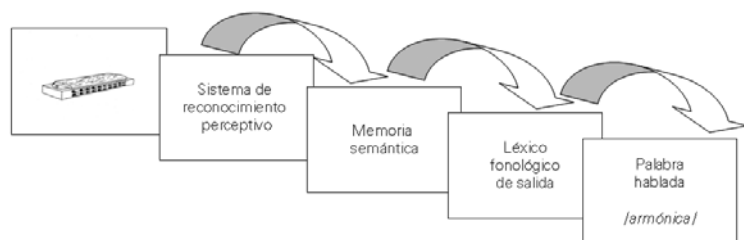
Aceptado tras revisión externa:

29.08.12.

Cómo citar este artículo:

Fernández-Blázquez MA, Ruiz-Sánchez de León JM, López-Pina JA, Llanero-Luque M, Montenegro-Peña M, Montejo-Carrasco P. Nueva versión reducida del test de denominación de Boston para mayores de 65 años: aproximación desde la teoría de respuesta al ítem. Rev Neurol 2012; 55: 399-407.

© 2012 Revista de Neurología

Figura 1. Esquema simplificado de los procesos implicados en la denominación de dibujos.

describir los aspectos básicos del estímulo, como su forma, pero presentará errores importantes en la denominación (por ejemplo, puede confundir el estímulo de la figura 1 con una caja de lápices, un aparato de aire acondicionado o una jaula para pájaros). En este caso, es recomendable valorar la presencia de agnosia visual mediante algún test específico [19].

- Si el reconocimiento perceptivo es adecuado, posteriormente debe activar en el sistema semántico el conocimiento que ha adquirido a lo largo de su vida sobre el concepto en cuestión, almacenado como un juego de rasgos definitorios y característicos (por ejemplo, ‘es’ un instrumento musical de viento, ‘es’ pequeño, ‘puede’ sonar, ‘tiene’ lengüetas, etc.). Si se produce un error a este nivel semántico, el individuo puede no distinguir entre conceptos con los que comparta un buen número de atributos (por ejemplo, la flauta o cualquier otro instrumento musical en función de la gravedad del cuadro). De ser así, sería adecuado valorar la presencia de alteraciones mediante algún test de abstracción verbal [20].
- Si la activación en la memoria semántica ha sido correcta, por último, debe activar el programa fonológico que contiene la información motora acerca de cómo se articula ese concepto previamente activado (/armónica/). Si aparecen alteraciones en este nivel, el individuo puede presentar problemas para denominarlo correctamente, produciendo parafasias fonológicas o, simplemente, experimentar el fenómeno de tener la palabra ‘en la punta de la lengua’ [21]. No obstante, podrá describir el estímulo sin problemas, realizar gestos acerca de cómo se usa, comentar múltiples aspectos relacionados con él, e incluso evocar experiencias autobiográficas pasadas (por ejemplo, ‘se sopla por los agujeros y suena’, ‘tenía una cuando era un niño’, ‘se mueve así’, etc.).

El test de denominación de Boston (BNT) es una prueba concebida específicamente para valorar la capacidad para la denominación de estímulos visuales. En su primera versión, el BNT constaba de un total de 85 láminas con dibujos en blanco y negro [22]. Años después, el número de láminas de la prueba se redujo hasta obtener la versión definitiva con 60 dibujos [23], y fue incluido en la batería de evaluación de la afasia y de trastornos relacionados [24]. Recientemente, algunos de los dibujos originales han sido cambiados con motivo de una versión más moderna de dicha batería [25].

La tarea consiste en denominar cada una de las 60 láminas, presentadas en orden de dificultad creciente, en un tiempo máximo de 20 segundos. En caso de que el individuo no ofrezca una respuesta correcta de forma espontánea, el examinador puede proporcionar una pista semántica (para estudiar la posibilidad de que se haya producido un error en el reconocimiento del dibujo) o fonológica (cuando la pista semántica aún no resulta suficiente para evocar una respuesta o si se ha producido un error en la respuesta espontánea que no es de tipo perceptivo). La puntuación total de la prueba resulta de sumar el número de respuestas espontáneas correctas más el número de láminas denominadas con ayuda de la clave semántica. Las respuestas correctas tras la clave fonológica se tienen en cuenta como indicador del tipo de dificultad para denominar dibujos [17,18].

Diversos trabajos han mostrado una relación consistente entre la puntuación total del BNT y variables sociodemográficas como la edad y el nivel educativo. De este modo, se ha observado que el rendimiento en denominación disminuye a medida que aumenta la edad [26-29] y descienden los años de escolaridad [29-31].

Durante los últimos años han surgido diferentes versiones reducidas del BNT con el fin de minimizar el tiempo de aplicación de la prueba [32]. Estas versiones del BNT difieren tanto en el modo en que se han obtenido como en el número de ítems que plantean [31-38] (Tabla I). Así, algunas versiones optan por preservar la esencia de la prueba original, manteniendo un criterio de frecuencia léxica y dificultad de los ítems [31,33], mientras que otras adoptan un criterio clínico para identificar los dibujos que mejor discriminan entre mayores sanos y personas con enfermedad de Alzheimer [34], o entre sujetos con y sin alteraciones en la denominación, independientemente del diagnóstico [32]. Con respecto al número de láminas, las versiones estudiadas también son diferentes, pudiendo contar generalmente con un máximo de 30 y un mínimo de 15 dibujos. En cualquier caso, existen pocos datos

Tabla I. Versiones reducidas del test de denominación de Boston.

	N.º de ítems	Criterio y muestra	Fiabilidad	Correlación
Williams et al [33]	Una forma de 30 ítems	Criterio clínico: estudio con 15 controles y 15 pacientes con EA	$\alpha = 0,95$ con muestra clínica $\alpha = 0,74$ con muestra control	$r = 0,99$ con muestra clínica $r = 0,96$ con muestra control
Mack et al [34]	Cuatro formas de 15 ítems	Criterio de frecuencia léxica: 26 controles y 26 pacientes con EA	$\alpha = 0,84$ a $0,87$ con muestra clínica $\alpha = 0,31$ a $0,52$ con muestra control	$r = 0,95$ a $0,97$ con muestra clínica $r = 0,72$ a $0,82$ con muestra control
Fastenau et al [35]	Dos formas (A y B) de 30 ítems	Construidos a partir de las formas 1 + 2 y 3 + 4 de [39]: 108 mayores sanos	$\alpha_A = 0,51$ $\alpha_B = 0,75$	$r_A = 0,76$ $r_B = 0,76$
Lansing et al [36]	Una forma de 15 ítems	Criterio clínico: análisis discriminante <i>stepwise</i> en 717 controles y 237 pacientes con EA	–	–
Saxton et al [31]	Cuatro formas de 30 ítems	Criterio de dificultad de los ítems: 314 mayores sanos	$\alpha = 0,83$	$r = 0,82$
Graves et al [32]	Una forma de 30 ítems Una forma de 15 ítems Una forma adaptativa de 15/30 ítems	Criterio clínico: 206 sujetos mayores en dos grupos con/sin alteración de la denominación	$\alpha_{30} = 0,90$ $\alpha_{15} = 0,84$	$r_{30} = -$ $r_{15} = -$
Nebreda et al [37]	Una forma de 11 ítems	Criterio clínico: 102 pacientes (43 con demencia) y 78 controles	–	$r = -$
Hobson et al [38]	Dos formas de 30 ítems (pares e impares del test de denominación de Boston) y una forma de 15 ítems (CERAD)	Criterio clínico: 120 pacientes mayores con demencia y 29 controles	–	$r_{30 \text{ par}} \text{ estimada} = 0,97$ $r_{30 \text{ impar}} \text{ estimada} = 0,98$ $r_{15} \text{ estimada} = 0,93$

CERAD: Consortium to Establish a Registry for Alzheimer's Disease; EA: enfermedad de Alzheimer.

de fiabilidad de todas estas formas reducidas, si bien la mayor parte de los estudios muestran unos coeficientes de correlación entre las versiones y la prueba original por encima de 0,8 [32].

Con respecto al procedimiento para obtener una versión reducida del BNT, hay que tener en cuenta dos consideraciones importantes. En primer lugar, la fiabilidad de una prueba disminuye a medida que se reduce su número de ítems [39]. Por esta razón, las versiones de 15 láminas corren el riesgo de perder fiabilidad frente a las de 30. En segundo lugar, la obtención de los ítems como resultado de un análisis discriminante basado en un criterio diagnóstico (por ejemplo, participantes controles comparados con participantes con demencia) no parece la más adecuada y ha sido duramente criticada. El elevado poder discriminativo exhibido por este método se restringe a la muestra utilizada, por lo que la generalización al resto de la población se encuentra sesgada [40]. Un procedimiento alternativo para reducir el número de láminas consiste en aplicar los supuestos de la teoría de respuesta al ítem, de forma que se seleccionen aquellos ítems que forman una

prueba homogénea y que mejor discriminan las dificultades en denominación [32,37].

El objetivo del presente trabajo consiste en obtener una nueva versión reducida del BNT que permita disminuir su tiempo de aplicación sin sacrificar por ello su bondad psicométrica. Para ello, se adoptará un criterio clínico para la selección de dos grupos de sujetos ('sin alteración en la denominación' y 'con alteración en la denominación'), y se aplicarán los supuestos de la teoría de respuesta al ítem. Además, se estudiarán las características psicométricas de la nueva versión y se proporcionarán sus datos normativos y puntos de corte para su aplicación en la consulta clínica.

Sujetos y métodos

Muestra

Se obtuvo de forma consecutiva una muestra de 547 sujetos mayores de 65 años que acudieron voluntariamente al Centro de Prevención del Deterioro Cog-

Tabla II. Pasos en el análisis de los ítems del test de denominación de Boston.

	Paso 1: modelo clásico		Paso 2: análisis 2p			Paso 3: análisis Rasch	
	IH	ID	Slope (DE)	χ^2	p	χ^2	p
Ítem 1	0	0	–	–	–	–	–
Ítem 2	0	0	–	–	–	–	–
Ítem 3	0,14	0,624	–	–	–	–	–
Ítem 4	0	0	–	–	–	–	–
Ítem 5	0	0	–	–	–	–	–
Ítem 6	–0,031	–0,205	–	–	–	–	–
Ítem 7	0,135	0,693	–	–	–	–	–
Ítem 8	0,123	0,493	–	–	–	–	–
Ítem 9	0,136	0,386	–	–	–	–	–
Ítem 10	0,367	0,653	1,39 (0,21)	0,4	0,996	2,3	0,801
Ítem 11	0,282	0,578	–	–	–	–	–
Ítem 12	0,092	0,3	–	–	–	–	–
Ítem 13	0,397	0,678	1,42 (0,19)	1,6	0,901	5,9	0,315
Ítem 14	0,263	0,605	–	–	–	–	–
Ítem 15	0,072	0,22	–	–	–	–	–
Ítem 16	0,238	0,499	–	–	–	–	–
Ítem 17	0,246	0,466	–	–	–	–	–
Ítem 18	0,362	0,512	1,19 (0,51)	1,6	0,954	3,9	0,692
Ítem 19	0,403	0,539	0,97 (0,12)	7,9	0,338	0,8	0,993
Ítem 20	0,113	0,42	–	–	–	–	–
Ítem 21	0,264	0,803	–	–	–	–	–
Ítem 22	0,308	0,876	–	–	–	–	–
Ítem 23	0,555	0,774	2,25 (0,26)	9	0,061	–	–
Ítem 24	0,325	0,594	1,33 (0,22)	5,3	0,377	0,4	0,994
Ítem 25	0,499	0,625	1,39 (0,14)	4,3	0,63	8,3	0,138
Ítem 26	0,297	0,379	–	–	–	–	–
Ítem 27	0,304	0,456	0,85 (0,13)	1,9	0,964	7,5	0,274
Ítem 28	0,24	0,683	–	–	–	–	–
Ítem 29	–0,072	–0,093	–	–	–	–	–
Ítem 30	0,474	0,598	1,27 (0,14)	7	0,318	22,5	0,001

nitivo de Madrid manifestando quejas subjetivas de memoria. Tras la aplicación sistemática de un protocolo de evaluación neuropsicológica, todos los participantes recibieron un diagnóstico clínico en función de si cumplían o no los criterios para el diagnóstico de deterioro cognitivo leve [41]. En este sentido, 405 sujetos no presentaron signos de alteración cognitiva significativa, por lo que formaron parte del grupo control (edad: $75,98 \pm 6,51$; años de escolaridad: $10,07 \pm 4,77$; 71,4%: mujeres), frente a 142 sujetos que fueron diagnosticados de deterioro cognitivo leve (edad: $76,32 \pm 5,5$; años de escolaridad: $7,91 \pm 5,01$; 71%: mujeres). Todos los participantes con sospecha de demencia, sintomatología depresiva, accidente cerebrovascular, traumatismo craneoencefálico, encefalitis, hidrocefalia normotensiva, antecedentes neuroquirúrgicos, enfermedades sistémicas, historial de abuso de drogas o puntuaciones que pudieran interpretarse como sospechosas de simulación fueron excluidos del estudio. La lengua materna de todos los participantes en el estudio era el español.

Material y procedimiento

Se recogió la información sociodemográfica y clínica de todos los participantes por medio de una entrevista estructurada. Además, se aplicó un protocolo breve de exploración cognitiva constituido por dos pruebas de evaluación cognitiva global –*Minimal State Examination* y test de los 7 minutos– [42,43], una tarea para evaluar la memoria episódica (lista de palabras de la escala de memoria de Weschler) [44], una escala abreviada para la evaluación de los síntomas depresivos (*Geriatric Depression Scale*) [45] y otra escala, cumplimentada por un informador fiable, para la valoración de las actividades instrumentales de la vida diaria (cuestionario de actividad funcional de Pfeffer) [46]. El protocolo de evaluación se aplicó en una sola sesión de aproximadamente una hora de duración, manteniendo unas condiciones similares para todos los sujetos. Durante la realización de las tareas de evaluación se informó a los participantes de la posibilidad de hacer un descanso para evitar el efecto de la fatiga sobre su rendimiento cognitivo. El BNT se aplicó a todos los sujetos de la muestra siguiendo las instrucciones del manual del test [23]. Todas las respuestas se recogieron en las hojas de anotación y después se codificaron en una base de datos para su posterior análisis estadístico.

Análisis estadístico

Una condición esencial para la aplicación del modelo de Rasch es el estudio previo del ajuste de los

ítems y del test completo al modelo especificado. El modelo de Rasch dispone de una gran variedad de estadísticos de ajuste, destacando aquéllos basados en los residuales entre la respuesta dada y la probabilidad esperada. Si el método de estimación de parámetros converge y los estadísticos de ajuste confirman que las respuestas de las personas pueden explicarse en función de la dificultad de los ítems, entonces las estimaciones de los parámetros de las personas y de los ítems son independientes, es decir, la habilidad estimada de cada persona no depende del número y tipo de ítems que ha contestado [47,48]. En ciencias de la salud, el modelo de Rasch se ha convertido en un estándar precisamente porque la mera suma de ítems acertados es un estimador suficiente del parámetro de habilidad.

En primer lugar, se realizó un primer cribado de los ítems, eliminando aquéllos que no cumplieron determinados criterios de dificultad y homogeneidad de acuerdo con el modelo clásico de construcción de tests. Posteriormente, para estudiar la estructura dimensional de las láminas del BNT seleccionadas, se empleó el programa TestFact. Este programa permite determinar la dimensionalidad de una matriz de correlaciones tetracóricas resultante de la estructura correlacional de los ítems dicotómicos del BNT.

Una vez determinada la estructura unidimensional de la escala resultante, se empleó el programa BILOG-MG para realizar una estimación de parámetros de acuerdo con el modelo logístico apropiado. Asimismo, se obtuvieron los parámetros de discriminación (*slope*) de cada lámina. En el modelo de Rasch, se postula que dichos parámetros son iguales en todos los ítems e iguales a 1. Por tanto, aquellos ítems que tengan parámetros de discriminación muy elevados (> 1,7) provocarán problemas de ajuste y, por tanto, deberían ser eliminados del análisis. Se utilizó el programa estadístico SPSS v. 17.0 para calcular la consistencia interna de la nueva versión reducida y para hallar su correlación con el BNT original. Por último, se obtuvieron los puntos de corte de la prueba para detectar a los sujetos con alteraciones en la denominación de dibujos por confrontación visual.

Resultados

Variables sociodemográficas y rendimiento en el BNT

Se estudió la posible influencia de las variables sociodemográficas sobre el rendimiento del BNT en los 405 sujetos del grupo control. Se observó que

Tabla II. Pasos en el análisis de los ítems del test de denominación de Boston (*cont.*).

	Paso 1: modelo clásico		Paso 2: análisis 2p			Paso 3: análisis Rasch	
	IH	ID	<i>Slope</i> (DE)	χ^2	<i>p</i>	χ^2	<i>p</i>
Ítem 31	0,443	0,63	1,23 (0,16)	2,9	0,716	6,4	0,38
Ítem 32	0,275	0,391	–	–	–	–	–
Ítem 33	0,597	0,753	2,82 (0,33)	1,9	0,755	–	–
Ítem 34	0,478	0,601	1,35 (0,14)	1,9	0,927	5,4	0,374
Ítem 35	0,303	0,553	–	–	–	–	–
Ítem 36	0,507	0,668	1,47 (0,16)	4	0,67	8,1	0,153
Ítem 37	0,154	0,414	–	–	–	–	–
Ítem 38	0,5	0,679	1,69 (0,18)	3	0,804	–	–
Ítem 39	0,484	0,606	1,5 (0,15)	4,2	0,652	–	–
Ítem 40	0,489	0,714	1,81 (0,2)	3,1	0,796	–	–
Ítem 41	0,542	0,68	1,83 (0,18)	2,4	0,88	–	–
Ítem 42	0,434	0,588	1,64 (0,19)	3,7	0,592	–	–
Ítem 43	0,563	0,813	2,36 (0,25)	6,4	0,171	–	–
Ítem 44	0,361	0,576	1,25 (0,17)	8,3	0,216	7,2	0,307
Ítem 45	0,535	0,686	2,17 (0,23)	2,4	0,665	–	–
Ítem 46	0,247	0,445	–	–	–	–	–
Ítem 47	0,223	0,361	–	–	–	–	–
Ítem 48	0,298	0,377	–	–	–	–	–
Ítem 49	0,301	0,501	0,96 (0,16)	3,1	0,8	4,1	0,539
Ítem 50	0,521	0,683	1,58 (0,17)	2,1	0,908	–	–
Ítem 51	0,122	0,405	–	–	–	–	–
Ítem 52	0,563	0,712	2,01 (0,21)	1,5	0,915	–	–
Ítem 53	0,336	0,423	0,77 (0,11)	4,1	0,664	8,7	0,123
Ítem 54	0,363	0,562	1,17 (0,16)	5,4	0,497	2	0,847
Ítem 55	0,53	0,672	1,83 (0,18)	3,3	0,654	–	–
Ítem 56	0,168	0,21	–	–	–	–	–
Ítem 57	0,204	0,468	–	–	–	–	–
Ítem 58	0,544	0,691	1,75 (0,18)	3,3	0,654	–	–
Ítem 59	0,325	0,564	1,78 (0,29)	0,9	0,926	–	–
Ítem 60	0,473	0,687	2,66 (0,41)	2,4	0,499	–	–

DE: desviación estándar; ID: índice de dificultad; IH: índice de homogeneidad.

Tabla III. Orden de presentación de las láminas de la nueva versión reducida del test de denominación de Boston en función de la dificultad de respuesta a los ítems.

1. Pez espada	-1,471	9. Máscara	-0,120
2. Sacapuntas	-1,364	10. Magdalena	0,333
3. Pulpo	-1,187	11. Cactus	0,493
4. Espárrago	-1,051	12. Armónica	1,207
5. Bozal	-0,844	13. Dardo	1,574
6. Pinzas	-0,696	14. Zancos	1,844
7. Globo	-0,507	15. Pergamino	1,950
8. Rinoceronte	-0,161		

tanto la edad ($F_{28,404} = 2,27$; $p < 0,001$; $\eta^2_p = 0,14$) como el nivel educativo ($F_{5,404} = 13,12$; $p < 0,001$; $\eta^2_p = 0,14$) tenían un efecto significativo sobre dicho rendimiento. Estos resultados se tuvieron en cuenta para obtener los datos normativos y los puntos de corte de la nueva versión reducida del BNT.

Reducción de ítems del BNT

En la tabla II se recogen los datos del análisis de las láminas del BNT dentro del contexto del modelo clásico. Un buen punto de partida para realizar una estimación de parámetros con un modelo de respuesta al ítem es comenzar por eliminar ítems cuyos índice de dificultad (ID) e índice de homogeneidad (IH) tienen valores inadecuados [49,50]. En el presente trabajo se ha considerado que dichos valores serían $IH < 0,3$ o $ID > 0,9$, por lo que, de acuerdo con esta regla, se eliminaron 30 láminas del BNT (de la 1 a la 9, 11, 12, de la 14 a la 17, de la 20 a la 22, 26, 28, 29, 32, 35, 37, de la 46 a la 48, 51, 56 y 57). Dado que esta selección puede ser dependiente de la muestra empleada, se llevó a cabo un estudio de validación cruzada para comprobar que no existirían diferencias significativas. Para ello, se dividió la muestra en dos subgrupos aleatorios y se obtuvo que la selección de ítems en cada subgrupo fue muy semejante, por lo que se procedió a continuar con el análisis de las láminas restantes (10, 13, 18, 19, 23, 24, 25, 27, 30, 31, 33, 34, 36, 38, 39, 40, 41, 42, 43, 44, 45, 49, 50, 52, 53, 54, 55, 58, 59 y 60).

Puesto que la unidimensionalidad de la escala es uno de los supuestos del modelo de Rasch, se realizó un análisis factorial de la dimensionalidad de las 30 láminas seleccionadas. Bajo el supuesto de que la matriz de correlaciones tetracóricas se puede ex-

plicar sólo con un factor, se obtuvieron cargas factoriales mayores de 0,4 en el primer factor con un *eigenvalor* de 13,084 frente al *eigenvalor* del segundo factor de 0,926. Como la razón del primer factor sobre el segundo es 14,124, los datos apuntaron hacia una solución unidimensional.

Se obtuvo el grado de ajuste individual de las 30 láminas seleccionadas, así como el de todas ellas en su conjunto, de forma que se observó un elevado desajuste en algunos ítems, así como del test completo ($\chi^2 = 290,9$; $p < 0,0001$). Por este motivo, se procedió a calcular los parámetros de estimación para cada lámina de forma que se eliminaran aquellas con problemas de ajuste al modelo. Se seleccionaron los ítems que mostraron un adecuado grado de ajuste (parámetro de discriminación $< 1,5$), obteniendo un total de 15 láminas: 10, 13, 18, 19, 24, 25, 27, 30, 31, 34, 36, 44, 49, 53 y 54. Todas estas láminas, excepto el ítem 30, presentaron un buen ajuste al modelo. No obstante, dado que también mostraron un adecuado grado de ajuste global ($\chi^2 = 93,5$; $p = 0,1622$), se optó por mantener el ítem 30 dentro de la nueva versión reducida del BNT. En la tabla III se presenta la nueva versión ordenada en función de la dificultad de respuesta a los ítems.

Propiedades psicométricas de la nueva versión reducida

Esta nueva escala se sometió al programa ConQuest para comprobar que todos los ítems se ajustaban al modelo de Rasch. Los resultados fueron positivos, ya que sus medias cuadráticas estuvieron entre 0,7 y 1,3. Asimismo, la fiabilidad de la separación de los parámetros fue de 0,995, lo que indicó que los parámetros estuvieron suficientemente separados para evaluar el continuo de habilidad. Además, la nueva prueba obtuvo una adecuada consistencia interna ($\alpha = 0,765$) y una elevada correlación con la prueba original ($r = 0,876$; $p < 0,001$).

Análisis de la nueva versión reducida

Dado que la edad y el nivel educativo resultaron significativos para explicar el rendimiento en el BNT, se utilizaron tres grupos de edad (65-72 años, 73-80 años y 81 años y más) y otros dos de nivel educativo ('estudios primarios completos e incompletos' y 'bachiller y estudios superiores'), de forma que se dividió a los 405 sujetos del grupo control en estos seis nuevos grupos. Para cada uno de ellos se obtuvo la puntuación correspondiente a 1,5 desviaciones típicas por debajo de la media, utilizada habitualmente como punto de corte para identificar alteraciones

cognitivas [41]. Teniendo en cuenta dichos puntos de corte para cada uno de los seis grupos, se procedió a clasificar a los 547 sujetos de la muestra, con independencia de su diagnóstico clínico previo, en una nueva variable con dos niveles (sin anomia y con anomia). A partir de esta nueva clasificación, se utilizó al grupo de sujetos sin anomia para obtener los puntos de corte de la nueva versión reducida ajustados por edad y nivel educativo (Tabla IV).

La figura 2 muestra los porcentajes de acierto en los ítems de la nueva versión reducida en sus modalidades de respuesta espontánea, semántica y fonológica. Como puede apreciarse, el grupo sin anomia tiene un porcentaje de aciertos más elevado en todas las condiciones que el grupo con anomia.

Discusión

El propósito de este trabajo es obtener una nueva versión reducida del BNT que permita disminuir su tiempo de aplicación sin sacrificar su bondad psicométrica. Para ello, se han aplicado los supuestos de la teoría de respuesta al ítem con el fin de obtener una prueba unidimensional capaz de discriminar las alteraciones en la denominación por confrontación visual.

Una revisión exhaustiva de la bibliografía al respecto pone de manifiesto la existencia de diferentes versiones reducidas del BNT que han sido obtenidas a partir de procedimientos dispares, que, por regla general, no se basan en criterios psicométricos. Esto supone, a nuestro juicio, un importante sesgo en la selección de las láminas, en tanto que estos criterios psicométricos deberían guiar la construcción de cualquier tipo de escala de medida. El trabajo de Graves et al supuso un punto de inflexión en la medida en que se introdujeron los principios de la teoría de respuesta al ítem en el procedimiento de reducción de láminas del BNT [32]. Nebreda et al también han aplicado dichos principios en nuestro país [37].

Los resultados del presente trabajo demuestran la potencialidad de la teoría de respuesta al ítem para obtener una evaluación exacta de la capacidad de denominación por confrontación visual en personas mayores de 65 años. En la medida en que los datos muestran un buen ajuste al modelo de Rasch, es posible afirmar que los parámetros de los ítems y los parámetros de habilidad de los sujetos son independientes entre sí, condición ampliamente deseada en ciencias de la salud.

La nueva versión reducida del BNT consta de 15 ítems ordenados en función de la dificultad. La me-

Figura 2. Porcentaje de aciertos para los grupos sin anomia y con anomia en denominación en las condiciones de respuesta espontánea (a), clave semántica (b) y clave fonológica (c).

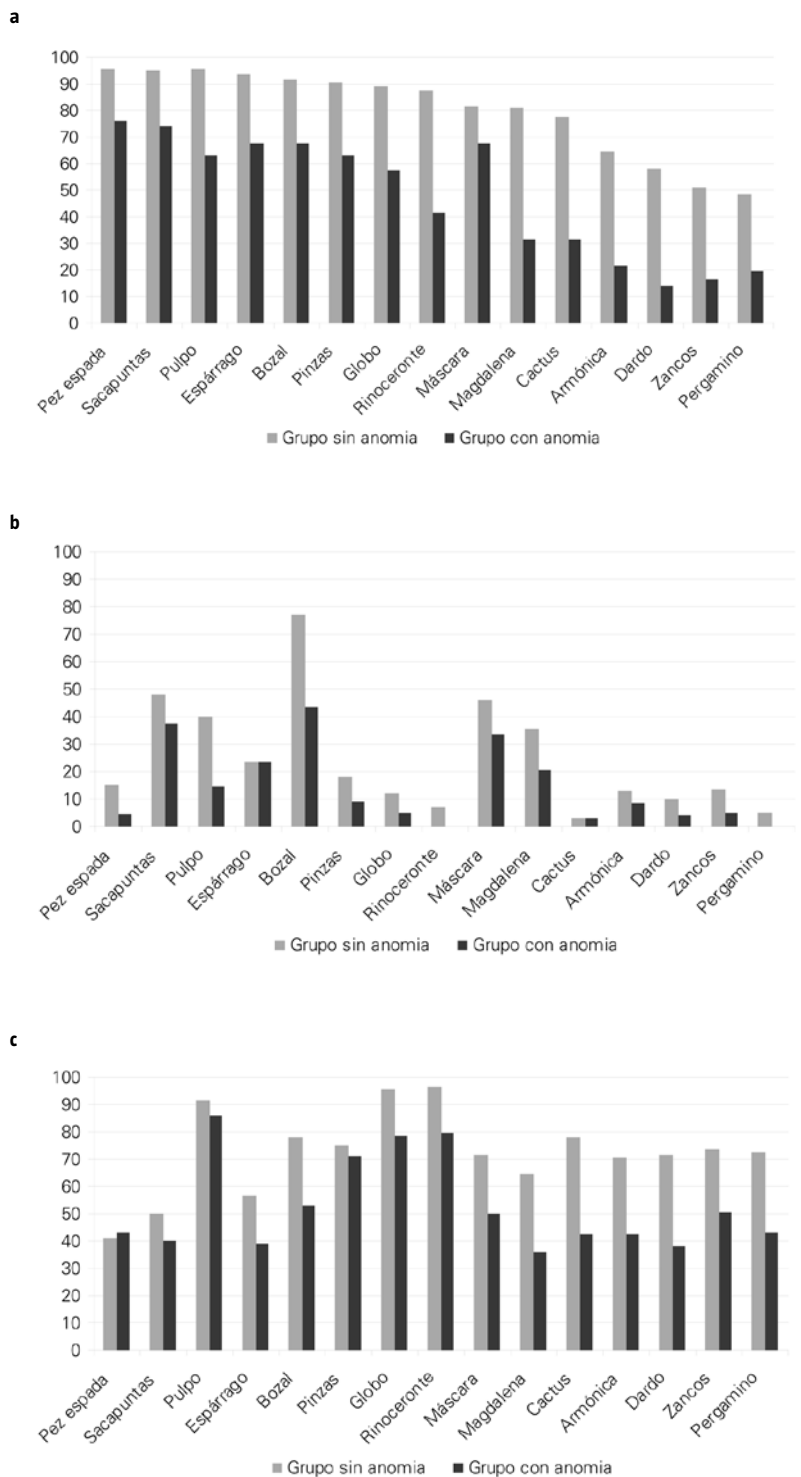


Tabla IV. Puntos de corte de la nueva versión reducida del test de denominación de Boston para su aplicación clínica.

	65-72 años (n = 149)		73-80 años (n = 183)		≥ 81 años (n = 123)	
	Estudios primarios	Bachiller/estudios superiores	Estudios primarios	Bachiller/estudios superiores	Estudios primarios	Bachiller/estudios superiores
n	76	73	119	64	66	57
Media ± DE	12,29 ± 1,84	13,41 ± 1,87	11,29 ± 2,38	12,30 ± 1,67	10,83 ± 2,21	12,21 ± 1,85
Punto de corte BNT-15	9	11	8	10	7	9

BNT: test de denominación de Boston; DE: desviación estándar.

todología empleada para obtener esta nueva versión ha permitido atenuar el efecto cultural de los ítems, motivo por el que esta prueba puede aplicarse a cualquier individuo mayor. Además, se presentan los puntos de corte para el diagnóstico de los problemas de denominación ajustados por edad y nivel educativo. Se observa que los sujetos sin anomia muestran un porcentaje de aciertos más elevado que los sujetos con problemas de denominación cuando se les ofrecen claves semánticas y fonológicas, lo que constituye un dato a tener en cuenta en la práctica clínica, puesto que pone de manifiesto que los sujetos con anomia presentan un deterioro semántico asociado que agudiza las dificultades en denominación.

Asimismo, la nueva versión reducida muestra una adecuada fiabilidad en la aplicación de la prueba, así como una elevada correlación con la prueba original. Se trata, por tanto, de una tarea de rápida aplicación que tiene un elevado poder discriminativo para problemas de denominación. Por todo ello, es un instrumento útil en la clínica diaria para detectar de forma rápida alteraciones del lenguaje en personas mayores.

Bibliografía

- Vendrell-Brucet JM. Las afasias: semiología y tipos clínicos. *Rev Neurol* 2001; 32: 980-6.
- Goodglass H, Kaplan E. The assessment of aphasia and related disorders. Philadelphia: Lea & Febiger; 1972.
- Benson DF, Geschwind N. Aphasia and related cortical disturbances. In Baker AB, Baker LH, eds. *Clinical neurology*. New York: Harper & Row; 1971. p. 112-40.
- Head H. Aphasia and kindred disorders of speech. London: Cambridge University Press; 1926.
- Luria AR. Basic problems of neurolinguistics. The Hague: Mouton; 1976.
- Shafiq MA, Burke DM, Stamatakis EA, Tam P, Tyler LK. On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *J Cogn Neurosci* 2007; 19: 2060-70.
- Zec RF, Markwell SJ, Burkett NR, Larsen DL. A longitudinal study of confrontation naming in the 'normal' elderly. *J Int Neuropsychol Soc* 2005; 11: 716-26.
- Reilly J, Peelle JE, Antonucci SM, Grossman M. Anomia as a marker of distinct semantic memory impairments in Alzheimer's disease and semantic dementia. *Neuropsychology* 2011; 25: 413-26.
- Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *J Clin Exp Neuropsychol* 2008; 30: 501-56.
- Moreno-Martínez FJ, Tallón-Barranco A, Frank A. Enfermedad de Alzheimer, deterioro categorial y variables relevantes en la denominación de objetos. *Rev Neurol* 2007; 44: 129-33.
- Fernández-Turrado T, Tejero-Juste C, Santos-Lasaosa S, Pérez-Lázaro C, Piñol-Ripoll G, Mostacero-Miguel E, et al. Lenguaje y deterioro cognitivo: un estudio semiológico en denominación visual. *Rev Neurol* 2006; 42: 578-83.
- Arvanitakis Z. Update on frontotemporal dementia. *Neurologist* 2010; 16: 16-22.
- Donaire-Temprado AM, Gil-Saladié D. Afasia primaria progresiva asociada a degeneración corticobasal. *Rev Neurol* 2001; 32: 1051-4.
- Klebic J, Salihovic N, Softic R, Salihovic D. Aphasia disorders outcome after stroke. *Med Arch* 2011; 65: 283-6.
- Pérez-Andrade MA, Poblano-Luna A. Afasia e hipoacusia secundarias a traumatismo craneoencefálico en niños y adolescentes. *Rev Neurol* 2007; 45: 62-4.
- Piñol-Ripoll G, Pérez-Lázaro C, Beltrán-Marín I, López del Val LJ, Pascual-Millán LF, Mostacero-Miguel E. Afasia como única manifestación de status epiléptico parcial. *Rev Neurol* 2004; 39: 1096-7.
- Cuetos F. Neuropsicología cognitiva del lenguaje. In Cuetos F, ed. *Psicolingüística del español*. Madrid: Trotta; 1999. p. 535-70.
- Cuetos F. Anomia: la dificultad para recordar las palabras. Madrid: TEA Ediciones; 2003.
- Warrington E, Merle J. VOSP: batería de tests para percepción visual objetos y espacio. Madrid: TEA Ediciones; 1994.
- Howard D, Patterson K. Pyramids and palm trees: a test of semantic access from pictures and words. Bury St. Edmunds, UK: Thames Valley Test; 1992.
- Facal-Mayo D, Juncos-Rabadán O, Álvarez M, Pereiro-Rozas AX, Díaz-Fernández F. Efectos del envejecimiento en el acceso al léxico. El fenómeno de la punta de la lengua ante los nombres propios. *Rev Neurol* 2006; 43: 719-23.
- Kaplan EF, Goodglass H, Weintraub S. The Boston Naming Test. Experimental edition. Boston: Kaplan & Goodglass; 1978.
- Kaplan EF, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
- Goodglass H, Kaplan E. Evaluación de la afasia y de trastornos relacionados. Madrid: Editorial Médica Panamericana; 1996.
- Goodglass H. Evaluación de la afasia y trastornos relacionados. Buenos Aires: Panamericana; 2005.

26. Au R, Joung P, Nicholas M, Obler LK, Kass R, Albert ML. Naming ability across the adult life span. *Aging & Cognition* 1995; 2: 300-11.
27. 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: 555-74.
28. Zec RF, Burkett NR, Markwell SJ, Larsen DL. Normative data for age, education, and gender on the Boston Naming Test. *Clin Neuropsychol* 2007; 21: 617-37.
29. Pineda DA, Mejía-Mag SE, Rosselli M, Ardila A, Romero MG, Pérez C. Variabilidad en la prueba de Boston para el diagnóstico de las afasias en adultos laboralmente activos. *Rev Neurol* 1998; 26: 962-70.
30. Zec RF, Burkett NR, Markwell SJ, Larsen DL. A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *Clin Neuropsychol* 2007; 21: 569-98.
31. Saxton KW, Ratcliff G, Munro CA, Coffey EC, Becker JT, Fried L, et al. Normative data on the Boston Naming Test and two equivalent 30-item short forms. *Clin Neuropsychol* 2000; 14: 526-34.
32. Graves RE, Bezeau SC, Fogarty J, Blair R. Boston Naming Test short forms: a comparison of previous forms with new item response theory based forms. *J Clin Exp Neuropsychol* 2004; 26: 891-902.
33. Williams BW, Mack W, Henderson VW. Boston Naming Test in Alzheimer's disease. *Neuropsychologia* 1989; 27: 1073-9.
34. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: shortened versions for use in Alzheimer's disease. *J Gerontol Psychol Sci* 1992; 47: 154-8.
35. Fastenau PS, Denburg NL, Mauer BA. Parallel short forms for the Boston Naming Test: psychometric properties and norms for older adults. *J Clin Exp Neuropsychol*, 1998; 20: 828-34.
36. Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston Naming Test. *Arch Clin Neuropsychol* 1999; 14: 481-7.
37. Nebreda MC, García-Caballero A, Asensio E, Revilla P, Rodríguez-Gironde M, Mateos R. A short-form version of the Boston Naming Test for language screening in dementia in a bilingual rural community in Galicia (Spain). *Int Psychogeriatr* 2011; 23: 435-41.
38. Hobson VL, Hall JR, Harvey M, Cullum CM, Lacritz L, Massman PJ, et al. An examination of the Boston Naming Test: calculation of 'estimated' 60-item score from 30- and 15-item scores in a cognitively impaired population. *Int J Geriatr Psychiatry* 2011; 26: 351-5.
39. Embretson SE. The new rules of measurement. *Psychol Assess* 1996; 8: 341-9.
40. Thompson B. Stepwise regression and stepwise discriminant analysis need not apply here: a guidelines editorial. *Educ Psychol Meas* 1995; 55: 525-34.
41. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999; 56: 303-8.
42. Folstein M, Folstein SE, McHugh PR. 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatry Res* 1975; 12: 189-98.
43. Solomon PR, Hirschhoff A, Kelly B, Relin M, Brush M, DeVeaux RD, et al. A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease. *Arch Neurol* 1998; 55: 349-55.
44. Wechsler D. Escala de memoria de Wechsler para Adultos (WMS-III). 3 ed. Madrid: TEA Ediciones; 1999.
45. Sheikh JL, Yesavage JA. Geriatric Depression Scale (GDS). Recent evidence and development of a shorter version. In Brink TL, ed. *Clinical gerontology: a guide to assessment and intervention*. New York: Haworth Press; 1986. p. 165-73.
46. Pfeffer RI, Kurosaki TI, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982; 37: 323-9.
47. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Mahwah, NJ: Lea; 2001.
48. De Ayala RJ. The theory and practice of item response theory. New York: Guilford Press; 2009.
49. Crocker L, Algina J. Introduction to classical and modern test theory. New York: Holt, Rinehart & Winston; 1986.
50. Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and Use. 4 ed. New York: Oxford University Press; 2008.

A new shortened version of the Boston Naming Test for those aged over 65: an approach from item response theory

Introduction. One of the tests that is mostly widely used to evaluate aphasia in clinical practice is the Boston Naming Test (BNT), a classic test in which 60 black and white pictures are presented to subjects in order to evaluate their capacity to put a name to such pictures. Despite its psychometric goodness, the number of items in the test has to be reduced in order to lower the time required to apply it.

Subjects and methods. Researchers recruited a sample of 547 subjects over the age of 65, who were then administered a neuropsychological evaluation protocol, including the BNT, to determine their cognitive statuses. No relevant cognitive alterations were observed in 405 subjects versus 142 who were diagnosed with mild cognitive impairment.

Results. The number of items was reduced in accordance with the premises of the item response theory. Since both age and level of schooling were found to have a significant effect on performance in the test, the two variables were used to find the cut-off points of the shortened version. This new version presented an adequate degree of reliability ($\alpha = 0.765$) and a high correlation with the original test ($r = 0.876$).

Conclusions. The new shortened version consists of 15 items that are ordered according to the degree of difficulty. It is a task with a high level of discriminating power that is useful in day-to-day clinical practice for detecting alterations in the language of the elderly.

Key words. Ageing. Anomia. Aphasia. Boston Naming Test. Dementia. Item response theory. Mild cognitive impairment. Naming. Neuropsychological evaluation.