

Validación de Prueba de Asociación Semántica mediante la Teoría de la Respuesta al Ítem

Validação de Teste de Associação Semântica por meio da Teoria da Resposta ao Item
Validation d'un test d'association sémantique à l'aide de la théorie de la réponse à l'item
Validation of Semantic Association Test using Item Response Theory

Jonatan Ferrer Aragón¹, Víctor Manuel Patiño Torrealva¹ y
Adela Hernández Galván¹

1. Centro de Investigación Transdisciplinar en Psicología, Universidad Autónoma del Estado de Morelos, México.

Resumen

La mayoría de las tareas empleadas para la evaluación del conocimiento conceptual emplean algún tipo de procesamiento no semántico adicional, además, no controlan el grado de dificultad o demanda cognitiva de los ensayos ni consideran el efecto de diversas variables psicolingüísticas. El objetivo fue validar una tarea de asociación semántica desde el modelo de dos parámetros de la teoría de respuesta al ítem. La tarea fue diseñada bajo un formato de elección forzada, con diferentes demandas de control semántico y empleando palabras concretas y abstractas. Se emplearon 392 palabras para conformar 112 ítems. La tarea se aplicó en 149 personas sanas (38 hombres/111 mujeres). Se estimaron los parámetros de dificultad (b) y discriminación (a). No hubo diferencias significativas en el número de letras, frecuencia léxica, familiaridad, imaginabilidad y concreción entre ítems de baja y alta demanda semántica. Se seleccionaron 12 ítems con los mayores parámetros de discriminación y que mostraran un adecuado grado de ajuste global al modelo. A modo de conclusión, la tarea es válida y sensible, de modo que se puede emplear de forma experimental y clínica para identificar déficits del conocimiento conceptual en pacientes con daño cerebral, así como para estudiar los diferentes tipos de procesamiento semántico.

Palabras clave: teoría de respuesta al ítem, cognición, semántica, pruebas de lenguaje, neuropsicología.

Resumo

A maioria das tarefas utilizadas para a avaliação do conhecimento conceitual emprega algum tipo de processamento não semântico adicional, além, não controlar o grau de dificuldade ou a demanda cognitiva dos ensaios, nem considerar o efeito de diversas variáveis psicolinguísticas. O objetivo foi validar uma tarefa de associação semântica a partir do modelo de dois parâmetros da Teoria de Resposta ao Item. A tarefa foi elaborada em um formato de escolha forçada, com diferentes demandas de controle semântico e utilizando palavras concretas e abstratas. Foram empregadas 392 palavras para compor 112 itens. A tarefa foi aplicada em 149 pessoas saudáveis (38 homens e 111 mulheres). Estimaram-se os parâmetros de dificuldade (b) e discriminação (a). Não houve diferenças significativas no número de letras, frequência lexical, familiaridade, imaginabilidade e concretude entre itens de baixa e alta demanda semântica. Foram selecionados 12 itens com os maiores parâmetros de discriminação e que apresentaram um grau adequado de ajuste global ao modelo. Em conclusão, a tarefa é válida e sensível, podendo ser utilizada de forma experimental e clínica para identificar déficits no conhecimento conceitual em pacientes com lesão cerebral, bem como para estudar diferentes tipos de processamento semântico.

Palavras-chave: teoria de resposta ao item, cognição, semântica, testes de linguagem, neuropsicologia.

Artículo recibido: 15/07/2024; Artículo aceptado: 15/08/2025.

Correspondencias relacionadas con este artículo deben ser enviadas a Jonatan Ferrer Aragón, Centro de Investigación Transdisciplinar en Psicología, Universidad Autónoma del Estado de Morelos –(CITPsi), Universidad Autónoma del Estado de Morelos, Pico de Orizaba 1, Colonia Los Volcanes, 62350, Cuernavaca, Morelos – Morelos.

E-mail: jonatan.ferrer@docentes.uaem.edu.mx

DOI:10.5579/rnl.2025.0888

Résumé

La majorité des tâches utilisées pour l'évaluation des connaissances conceptuelles impliquent un certain type de traitement non sémantique supplémentaire, sans, autant contrôler le degré de difficulté ou la charge cognitive des essais, ni tenir compte de l'effet de diverses variables psycholinguistiques. L'objectif de cette étude était de valider une tâche d'association sémantique selon le modèle à deux paramètres de la théorie de la réponse à l'item. La tâche a été conçue sous un format de choix forcé, avec différentes exigences de contrôle sémantique, et en utilisant des mots concrets et abstraits. Un total de 392 mots a été utilisé pour constituer 112 items. La tâche a été administrée à 149 personnes en bonne santé (38 hommes / 111 femmes). Les paramètres de difficulté (b) et de discrimination (a) ont été estimés. Aucune différence significative n'a été trouvée concernant le nombre de lettres, la fréquence lexicale, la familiarité, l'imaginabilité et la concrétude entre les items à faible et à forte demande sémantique. Douze items présentant les paramètres de discrimination les plus élevés et un ajustement global satisfaisant au modèle ont été sélectionnés. En conclusion, la tâche s'avère valide et sensible, ce qui permet son utilisation à la fois dans des contextes expérimentaux et cliniques pour identifier les déficits des connaissances conceptuelles chez des patients présentant une lésion cérébrale, ainsi que pour étudier les différents types de traitement sémantique.

Mots-clés : théorie de la réponse à l'item, cognition, sémantique, tests de langage, neuropsychologie.

Abstract

Most tasks used for the evaluation of conceptual knowledge involve some type of additional non-semantic processing and do not control for the level of difficulty or cognitive demand of the trials, nor do they consider the effect of various psycholinguistic variables. The objective was to validate a semantic association task using the two-parameter model of Item Response Theory. The task was designed in a forced-choice format, with different levels of semantic control demands, and used both concrete and abstract words. A total of 392 words were used to create 112 items. The task was administered to 149 healthy individuals (38 men / 111 women). Difficulty (b) and discrimination (a) parameters were estimated. There were no significant differences in the number of letters, lexical frequency, familiarity, imageability, and concreteness between low and high semantic demand items. Twelve items were selected with the highest discrimination parameters and that showed an adequate degree of overall model fit. In conclusion, the task is valid and sensitive, and can be used experimentally and clinically to identify conceptual knowledge deficits in patients with brain damage, as well as to study different types of semantic processing.

Keywords: item response theory, cognition, semantics, language tests, neuropsychology.

1. INTRODUCCIÓN

Puede considerarse que prácticamente todas las formas de actividad humana se llevan a cabo involucrando activamente los conocimientos que la persona posee sobre la realidad. La actividad perceptiva, la realización de actos aprendidos, la organización y regulación del comportamiento, el lenguaje y la actividad mnésica, por mencionar algunas, se llevan a cabo con la participación determinante de dichos conocimientos. Tulving (1972) distinguió por vez primera un sistema de memoria humana en el que se almacenan los conocimientos que la persona tiene sobre los significados de las palabras, sobre las personas y lugares que conoce, los objetos, eventos y sobre la realidad en general. Este sistema de memoria se conoce desde entonces como *Memoria Semántica* (MS) y es reconocido como una de las bases sobre las que se desarrollan procesos intelectuales, tales como el razonamiento (Béghin & Markovits, 2021; Brisson & Markovits, 2020; Markovits, 2024; Markovits, 2024).

La MS se ha estudiado por la Neuropsicología y otras disciplinas científicas. Así, Warrington (1975) publicó uno de los primeros reportes documentando alteraciones selectivas de la MS en pacientes con secuelas por daño cerebral. Posteriormente se publicaron diversos estudios que evidenciaron patrones distinguibles de alteración en el conocimiento conceptual ante lesiones de distinta localización cerebral (Caramazza & Shelton, 1988; Damasio, 1990; McCarthy & Warrington, 1988; Warrington & Shallice, 1984; Warrington & McCarthy, 1987;). Por otra parte, Snowden et al. (1989), así como Hodges et al. (1992), identificaron a la *demencia semántica* como una entidad diagnóstica, caracterizada por un inicio insidioso en el que se presentan alteraciones más o menos aisladas en la MS a partir de la degeneración cortical inicial de los sectores anteriores de los lóbulos temporales. A partir de la evidencia clínica y los estudios en pacientes con lesiones cerebrales, mucha investigación se ha desarrollado respecto a los sustratos

cerebrales del procesamiento semántico mediante estudios con neuroimagen funcional (Binder & Desai, 2011; Binder et al., 2009; Frisby et al., 2023; Pescatore et al., 2025; Pestilli, 2018; Pescatore et al., 2025).

La investigación científica sobre el procesamiento del conocimiento semántico en el cerebro ha evidenciado la participación de diversos sistemas neuronales distribuidos en extensas zonas de la corteza cerebral. Ralph et al. (2017) propusieron un modelo teórico respecto a la organización general de la cognición semántica en el cerebro, al que denominaron *Modelo de la Cognición Semántica Controlada*. De acuerdo con él, el procesamiento de la cognición semántica requiere de dos sistemas cerebrales distintos, pero que participan de manera coordinada entre sí. Uno de estos sistemas procesa representaciones semánticas (es decir, información tanto de origen perceptual como de tipo proposicional respecto a significados de palabras, estructura y función de objetos, sobre personas y lugares); el otro sistema es de control semántico, y permite la realización de todas las operaciones necesarias para organizar y controlar el procesamiento semántico durante la realización de tareas específicas. Este modelo confirma los resultados de investigaciones experimentales anteriores, de acuerdo con las cuales pueden identificarse sistemas neuronales que se relacionan con el procesamiento del contenido de la MS, así como otros, relacionados con ellos, vinculados con la ejecución de los procesos que tienen lugar durante la realización de tareas semánticas (Binder et al., 2009; Koenig & Grossman, 2007).

De acuerdo con el modelo de Ralph et al. (2017), el sistema cerebral que procesa las representaciones semánticas (contenido de la MS) incluye regiones asociativas multimodales ubicadas en las regiones dorsales y ventrales de los lóbulos temporales, incluyendo ambos polos temporales, así como regiones parieto-temporales, todas en ambos hemisferios. Por su parte, el sistema cerebral que organiza y controla las actividades de procesamiento semántico incluyen

regiones prefrontales dorsolaterales y sus conexiones con las regiones incluidas en el primer sistema arriba señalado. Los polos temporales actúan como regiones de asociación transmodal que integran información proveniente de distintas zonas corticales de ambos sistemas (Chiou et al., 2018; Coutanche & Thompson-Schill, 2015; Ralph et al., 2017).

Las investigaciones realizadas en torno a la organización cerebral de la memoria semántica y sus alteraciones en casos de lesiones cerebrales han arrojado datos, hipótesis y teorías al respecto de diversos problemas específicos. Entre ellos, destaca el problema de la eventual organización del conocimiento en el cerebro por dominios y categorías: por ejemplo, seres vivos vs no vivos (McCarthy & Warrington, 2016). Asimismo, se han distinguido las alteraciones en las representaciones semánticas respecto a aquellas en los procesos de acceso a las mismas. Por otra parte, se ha abordado el problema de las diferencias entre la información semántica de tipo concreto y de tipo abstracto (Mkrtychian et al., 2019).

Desde las primeras investigaciones publicadas por Warrington (1975) en pacientes con lesiones cerebrales, se utilizaron diversas tareas para la obtención de datos para la descripción de las alteraciones selectivas de la MS. Entre ellas, pueden mencionarse las tareas de denominación de objetos, denominación ante definiciones, apareamiento palabra-figura, fluencia semántica y categorización. Algunos estudios han tomado este tipo de tareas de baterías generalmente construidas para la evaluación neuropsicológica general, o de los trastornos afásicos, por lo que no incluyen en su estructura variables que permitan la discriminación específica de distintos tipos de procesamiento semántico. Otros estudios han recurrido al empleo de pruebas específicamente diseñadas para la exploración de la MS, siendo el más conocido el test de Pirámides y Palmeras (Howard & Patterson, 1992), del cual Gudayol-Ferré et al. (2008) publicaron datos normativos preliminares para población de habla hispana. Este test mide el acceso a información semántica detallada mediante el emparejamiento de figuras o palabras referentes a objetos con base en su asociación semántica.

Algunos autores se han dado a la tarea de desarrollar baterías de pruebas para la exploración clínica de la MS en pacientes de habla hispana. Una de ellas es la *Evaluación de la Memoria Semántica en la Demencia tipo Alzheimer* (EMSDA), publicada por Peraita et al. (2000). Esta batería consta de ocho pruebas que evalúan tanto la producción como la comprensión semántica, utilizando, para la presentación de los estímulos, tanto material gráfico (dibujos de objetos) como verbal. Las pruebas que evalúan la producción son: denominación, clasificación (modalidad gráfica), fluidez y definición de categorías (modalidad verbal). Por su parte, las pruebas que evalúan la comprensión son: emparejamiento, analogías, verificación (modalidad gráfica) y reconocimiento de atributos (modalidad verbal). El objetivo particular de esta batería es la evaluación del deterioro de los conocimientos léxico-semánticos de categorías naturales y de objetos (seres vivos y no vivos), así como de sus atributos. Su validación se realizó mediante la aplicación a 75 participantes españoles de 54 a 93 años (30 sanos y 47 diagnosticados con demencia tipo Alzheimer), sin reportarse datos referentes a la escolaridad.

Otra batería para la evaluación de la MS es la batería *Nombela 2.0* (Moreno-Martínez & Rodríguez-Rojo, 2015), que tiene su antecedente en la batería Nombela 1 (Moreno & Cañamón, 2005). Al igual que la EMSDA, esta batería tiene el propósito de evaluar el deterioro en el conocimiento léxico-semántico de seres vivos y no vivos en población de habla hispana. Consta de 5 tareas, a saber: generación de definiciones verbales, definición ante fotografías, definición ante descripciones verbales, fluencia semántica (14 subcategorías) y apareamiento palabra-figura. Se validó mediante su aplicación a 49 participantes (10 con diagnóstico de Deterioro Cognitivo Leve, 20 con diagnóstico de Enfermedad de Alzheimer y 19 participantes sanos). Los grupos se controlaron por años de educación (medias entre 7.4 y 9.8) y por edad (medias entre 69.5 y 75.4).

En América Latina, Martínez-Cuitiño et al. (2009) presentaron la adaptación y validación de la *Batería Semántica 64* en población rioplatense. Consta de 6 pruebas que evalúan el acceso y recuperación de la información semántica de 64 estímulos, pertenecientes a los dominios de seres vivos y objetos inanimados. Estas pruebas son: emparejamiento palabra-dibujo, test de camellos y cactus (acceso al conocimiento semántico), dictado, lectura y denominación por confrontación visual (recuperación del conocimiento semántico). Este instrumento fue validado mediante su aplicación a 50 participantes sanos (34 a 79 años, 17.8 años de escolaridad promedio), 10 pacientes con la variante temporal de la Demencia Frontotemporal (49 a 88 años, escolaridad promedio de 14.4 años) y cinco con Afasia de Wernicke (54 a 72 años y escolaridad promedio de 16 años).

Si bien estos instrumentos han resultado ser de utilidad para la evaluación general del procesamiento semántico en población clínica hispanoparlante, sigue siendo necesario ampliarlos y perfeccionarlos para poder analizar algunos procesos que quedan todavía fuera de su alcance. Desde los primeros estudios publicados en torno al procesamiento semántico (véase, por ejemplo, Warrington & McCarthy, 1987), se desarrolló una hipótesis según la cual la MS estaría organizada por categorías, siendo fundamentales las que se refieren a seres vivos (subdivididas en mamíferos, depredadores, domésticos u otras) y otros grupos de cosas, tales como herramientas, prendas de vestir o vehículos, por mencionar algunos. Los instrumentos de evaluación que se mencionan líneas arriba dejaron de lado otros dominios categoriales, quizás más difíciles de delimitar taxonómicamente, y tampoco han controlado el proceso mismo de la categorización como actividad del individuo.

Adicionalmente, las diversas pruebas incluidas exploran principalmente conceptos concretos (que incluyen representaciones de origen perceptivo como forma, color o tamaño), y no están estructuradas para analizar suficientemente conceptos abstractos (que carecen de representaciones de origen perceptivo y solo se procesan por medios verbales). Además de lo anterior, es conocida la influencia de variables culturales en los estímulos y en el rendimiento en pruebas neuropsicológicas (Ardila, 2007), aún al interior de la propia comunidad hispanoparlante iberoamericana. Esto justifica la búsqueda de estrategias para reducir los sesgos que este hecho puede ocasionar, por ejemplo, los asociados al ítem (cuando un mismo ítem tiene

significados distintos en distintas culturas). Entre estas estrategias se encuentran el desarrollo de instrumentos centrados culturalmente (que tengan significado y validez con respecto a una función cognitiva en una población) o el análisis psicométrico de los ítems con métodos novedosos como la teoría de respuesta al ítem (TRI; Fernández & Abe, 2018; Merkley et al., 2023).

De manera particular, debe señalarse que no en todos estos instrumentos se han controlado todas las variables intervinientes que sesgan las ejecuciones y, por tanto, los datos que se obtienen. Para el caso de los estímulos pictóricos, deben considerarse variables tales como la complejidad visual (Humphreys & Forde, 2001). Asimismo, tanto para éstos como para los estímulos lingüísticos (palabras escritas o emitidas verbalmente), variables que se deben controlar son la frecuencia léxica (Oldfield & Wingfield, 1965), la familiaridad (Funnell & Sheridan, 1992), la tipicidad (Kiran & Thompson, 2003) y la edad de adquisición (Ellis & Morrison, 1998). Estas variables son específicas para cada comunidad lingüística, por lo que deben considerarse para la evaluación de la MS en cada contexto particular.

Las investigaciones de tipo experimental sobre la MS, bien sean de carácter conductual o mediante el registro de la actividad fisiológica, requieren de tareas que permitan medir procesos específicos del procesamiento semántico. Una gran cantidad de reportes de investigación describen tareas organizadas exprofeso, pero que no controlan suficientemente las variables intervinientes arriba mencionadas (Binder et al., 2009 y Moreno-Martínez & Rodríguez-Rojo, 2015).

El propósito de este estudio es el de proporcionar los resultados de una investigación mediante la cual se validó una *tarea de asociación semántica* desde la TRI que pueda utilizarse de manera clínica y experimental. La TRI ofrece ventajas sobre la Teoría Clásica de los Test (TCT) al permitir estimaciones más precisas del rasgo latente en distintos niveles de habilidad (Embretson & Reise, 2000). A diferencia

de la TCT, la TRI facilita el análisis de curvas características del ítem, esencial en contextos transculturales (Boone et al., 2014; De Ayala, 2009). Además, permite el desarrollo de pruebas computarizadas, incrementando la eficiencia y validez de la medición (Van der Linden & Glas, 2010). De acuerdo con el Modelo de la Cognición Semántica Controlada (Lambon-Ralph et al., 2017), resulta importante controlar el grado de dificultad o demanda cognitiva de los ítems, pues mientras mayor sea aquella se infiere una mayor participación del sistema de control semántico. Además de esto, esta tarea incluye estímulos concretos y abstractos de manera que se puedan observar efectos diferenciales entre ellos. Finalmente, los estímulos se encuentran balanceados entre los dominios de seres vivos (por ejemplo, *personas, animales, plantas*) y objetos inanimados (por ejemplo, *herramientas, muebles, vehículos*), así como de un conjunto de categorías semánticas de cada uno de ellos, incluyendo también conceptos abstractos como *virtud, emoción o estado mental*.

2. MÉTODO

2.1 Participantes

En el estudio participaron 149 personas sanas (38 hombres /111 mujeres), divididas en tres grupos etarios: jóvenes (20 a 34 años), adultos (35 a 59 años) y adultos mayores (más de 60 años). Los tres grupos no difirieron estadísticamente con respecto al nivel educativo. Ningún participante mostró depresión ni ninguna otra enfermedad médica, neurológica y/o psiquiátrica que afectara el desempeño cognitivo. No hubo diferencias significativas en la prueba de escrutinio cognitivo, ni en las puntuaciones tipificadas de los inventarios de depresión (Tabla 1).

Tabla 1.

Características sociodemográficas de los participantes

	Jóvenes		Adultos		Adultos mayores		F	p	n ²
	Media	DE	Media	DE	Media	DE			
Sexo (H/M)	14/28		11/28		13/55				
Edad	25.40	4.89	46.28	5.54	67.69	5.19	871.75	<0.01	0.92
Escolaridad	13.74	3.68	12.95	4.14	12.21	3.95	1.61	0.20	0.02
MoCA	26.10	2.13	26.36	1.74	25.70	1.96	1.16	0.32	0.02
IDBz*/GDSz*	0.03	0.99	-0.04	1.02	0.00	1.00	0.05	0.96	<0.01

Nota. IDBz= Puntuación tipificada del Inventario de Depresión de Beck; GDSz= Puntuación tipificada de la Escala de Depresión Geriátrica de Yesavage

2.2 Instrumentos

Se recogió la información sociodemográfica y clínica de todos los participantes por medio de una entrevista estructurada. Se aplicó un protocolo breve de exploración constituido por una prueba de escrutinio cognitivo, la Evaluación Cognitiva de Montreal (MoCA por sus siglas en inglés) (Aguilar-Navarro et al., 2018; Nasreddine et al., 2005) y dos escalas para la evaluación de los síntomas depresivos para los tres grupos etarios: Escala de Depresión Geriátrica de Yesavage (GDS; Salgado et al., 2024; Yesavage et al., 1983)

para los adultos mayores y el Inventario de Depresión de Beck (IDB) (Jurado et al., 1998) para el grupo de jóvenes y de adultos.

2.3 Diseño y construcción de tarea experimental

a. Elaboración de ítems

La tarea de asociación semántica fue diseñada bajo un formato de elección forzada, donde cada ítem estaba conformado por una palabra de prueba que aparecía en la parte superior de la pantalla y una fila de tres palabras

potenciales que se presentaban en la parte inferior. La consigna era decidir cuál de las tres palabras era más similar por su significado a la palabra de prueba. La tarea estaba conformada originalmente por 112 ítems. Para ello se emplearon 392 palabras en español, restringiéndose únicamente a sustantivos, cuya mitad eran palabras concretas y la otra mitad palabras abstractas. Las palabras concretas incluían seres vivos y objetos, mientras que las palabras abstractas hacían referencia a defectos o cualidades. De la totalidad de palabras, 56 sirvieron como palabras de prueba, 112 como palabras blanco y 224 como distractores. Se obtuvieron los valores de frecuencia léxica, familiaridad, concreción e imaginabilidad de cada palabra en la base *EsPal* (Duchon et al., 2013). Se seleccionaron como palabras de prueba aquellas que tuvieran los valores más altos en familiaridad y frecuencia léxica para reducir el efecto cultural. Asimismo, las palabras concretas y abstractas se clasificaron con base en los valores de concreción e imaginabilidad.

Los ítems se dividieron también en baja y alta demanda a partir del nivel de similitud semántica entre la palabra de prueba y las tres palabras potenciales. En el caso de las palabras concretas, la similitud semántica podía basarse en relaciones taxonómicas (pertenencia a la misma categoría) o temáticas (compartición de propiedades funcionales, perceptuales y/o espaciotemporales). En cambio, en las palabras abstractas, la similitud podía establecerse principalmente a través de relaciones de sinonimia, es decir, por la identidad o semejanza de significados. La similitud semántica de los ítems se midió mediante un cuestionario que se aplicó a 30 adultos sanos que no participaron en el estudio, quienes calificaron la relación entre la palabra de prueba y el resto de las palabras a partir de una escala que va desde 6 puntos (muy similar) hasta 1 (menos similar).

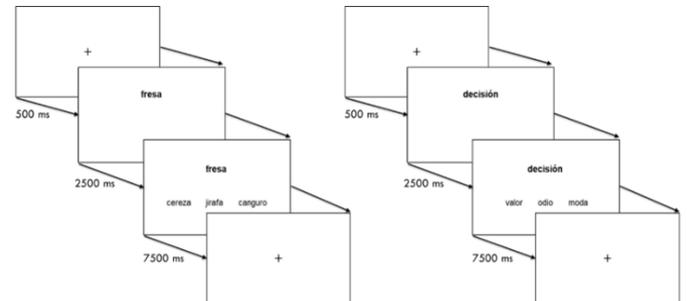
Para los ítems de baja demanda, se seleccionó como estímulo blanco aquella palabra con el valor promedio más alto de similitud conceptual respecto a la palabra de prueba, y se eligieron como distractores las dos opciones con los valores más bajos de similitud conceptual. De este modo, la palabra blanco y la palabra de prueba presentaban una alta similitud semántica, ya fuera por pertenecer a la misma categoría semántica, por compartir múltiples propiedades o por sinonimia. En contraste, los distractores no guardaban relación semántica con la palabra de prueba. Así, la demanda de control semántico era mínima, ya que la respuesta correcta podía identificarse de forma rápida y eficiente, sin una competencia significativa por parte de los distractores (véase Figura 1).

En la condición de alta demanda, se seleccionó como estímulo blanco la palabra con el segundo valor promedio más alto de similitud conceptual con respecto a la palabra de prueba, y como distractores se eligieron aquellas con el tercer y cuarto valores más altos. En estos ítems, existía una mayor competencia entre la palabra blanco y los distractores, ya que todas compartían la misma categoría semántica o hacían referencia a un tipo similar de cualidad o defecto. No obstante, la palabra blanco mantenía una mayor similitud conceptual con la palabra de prueba en comparación con los distractores, ya fuera por compartir un mayor número de propiedades o por una mayor cercanía en significado. Esta configuración incrementaba la demanda de control semántico, ya que se requerían recursos ejecutivos adicionales para guiar

la búsqueda e identificación de la representación semántica adecuada (véase Figura 2).

Figura 1.

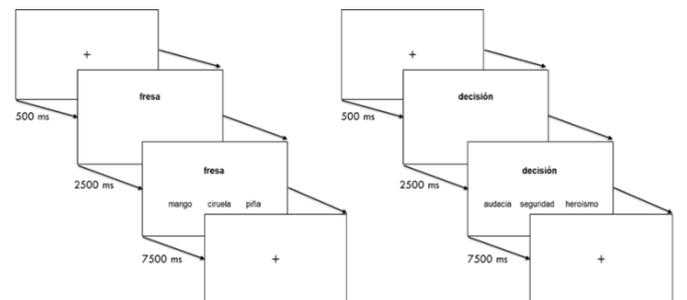
Estructura de la tarea de asociación semántica en la condición de baja demanda



Cabe señalar que, en ambas condiciones, se aseguró que existieran diferencias significativas en los niveles de similitud conceptual entre la palabra blanco y los distractores, con el fin de manipular de forma sistemática el grado de competencia semántica implicado en cada ítem.

Figura 2.

Estructura de la tarea de asociación semántica en la condición de alta demanda



b. Estructura de tarea experimental

La tarea se aplicó a los participantes de forma individual en una sesión de 60 minutos y en un espacio libre de distractores. Los ítems se proyectaron en un monitor de computadora, en letras minúsculas de color oscuro sobre un fondo blanco. La presentación de los estímulos y la captura de los datos se realizaron con el software E-Prime v. 1.1. Cada bloque comenzaba con una sesión de práctica de seis ensayos. Enseguida se mostraba una pantalla en blanco con un punto de fijación con duración de 500 ms para indicar el inicio de cada ítem, seguida por la palabra de prueba que se presentaba durante 2500 ms. Después aparecían tres palabras como opciones de respuesta, las cuales se mantenían presentes en la pantalla por un tiempo límite de 7500 ms o cuando los participantes proporcionaban la respuesta. Tan pronto como se producía la respuesta, se pasaba al siguiente ítem, comenzando nuevamente con el punto de fijación. Para elegir la palabra correcta, los participantes debían presionar en el teclado y con su mano derecha uno de tres botones (1, 2 o 3) que corresponden a la posición del elemento de respuesta (izquierda, central, derecha). Se registró el número de aciertos

y errores, así como los tiempos de ejecución de cada uno de los 112 ítems. A la mitad de la tarea había un intervalo de 120 segundos de descanso para evitar un efecto de fatiga sobre el rendimiento cognitivo. Asimismo, se contrabalanceó el orden de presentación de los ítems.

2.4 Validación psicométrica

Se efectuó el análisis psicométrico de la tarea de acuerdo con la teoría de respuesta al ítem (TRI), basándose en el modelo de dos parámetros (M2PL) el cual se emplea para ítems unidimensionales que evalúan rendimiento o habilidad y cuya respuesta sólo admite dos opciones (correcto–incorrecto). Se estimaron los parámetros de dificultad (b) y discriminación (a) mediante el software jMetrik v. 3.0 (Meyer, 2014). Se examinó la bondad de ajuste del modelo a partir del estadístico chi-cuadrado que compara las frecuencias observadas y esperadas por el modelo en distintos niveles del rasgo. La estimación de los parámetros del modelo se efectuó por el método de máxima verosimilitud con un criterio de convergencia de .001. Se descartaron los ítems que tuvieran problemas de ajuste al modelo y se seleccionaron aquellos ítems muy discriminantes con valores de a mayores que 1.34, así como ítems moderadamente discriminantes cuyos valores de a oscilaran entre 0.65 y 1.33 (Baker, 2001). A partir de los ítems seleccionados, se calculó la curva característica del test (CCT) para describir la relación funcional entre la puntuación verdadera y el nivel de aptitud. El proceso culminó al encontrar un subconjunto de ítems cuya función de información alcanzó el máximo ajuste al modelo.

Posteriormente, se analizó la influencia de la edad y escolaridad sobre el rendimiento en esta versión corta de la tarea, mediante un ANOVA mixto, tomando como variables dependientes el número de aciertos y los tiempos de respuesta. Se incluyeron como factores intergrupo el grupo etario (jóvenes, adultos y adultos mayores) y el nivel de escolaridad (alta y baja); para los factores intra-grupo se consideró el tipo de palabra (concretas-abstractas) y la demanda cognitiva (baja demanda-alta demanda). Se tomó como criterio de significación un valor de $\alpha < 0.05$ (95%). Se

Tabla 2.

Medias y desviaciones estándar (DE) de las palabras en las variables psicolingüísticas

	Baja demanda				Alta demanda				<i>F</i>	<i>p</i>	<i>n</i> ²
	Concretas		Abstractas		Concretas		Abstractas				
	Media	DE	Media	DE	Media	DE	Media	DE			
Número de letras	5.82	1.55	7.87	1.94	6.30	1.72	8.11	2.17	2.45	0.12	0.01
Frecuencia léxica	3.27	0.58	3.74	0.48	3.12	0.53	3.66	0.52	3.57	0.06	0.01
Familiaridad	5.49	0.96	5.26	0.75	5.51	0.90	5.12	0.82	0.46	0.49	<0.01
Imaginabilidad	5.99	0.53	3.64	0.80	5.93	0.55	3.68	0.82	0.01	0.99	<0.01
Concreción	5.91	0.58	3.85	0.59	6.00	0.53	3.83	0.49	0.09	0.76	<0.01

3.2 Calibración de ítems

Se realizó un análisis de los 112 ensayos mediante el M2PL de la TRI (ver Anexo 1). Se excluyeron 3 ítems del análisis por no presentar variabilidad en las respuestas. Los parámetros de dificultad de los 109 ítems restantes oscilaron entre $b = -12.70$ y $b = 18.31$, mientras que los parámetros de discriminación fluctuaron entre $a = 0.07$ y $a = 1.85$ (ver Anexo

utilizaron las pruebas post hoc de Tukey para explorar las interacciones significativas. Dichos análisis fueron realizados mediante el programa estadístico SPSS v. 25.0. Finalmente, se obtuvieron los puntos de corte de la tarea por cada grupo etario.

2.5 Consideraciones éticas

El presente trabajo se desarrolló de acuerdo con los principios establecidos en la declaración de Helsinki para investigación con seres humanos. El estudio fue aprobado por el Comité de Ética en Investigación del Instituto Nacional de Neurología y Neurocirugía con número de protocolo 84/19. El estudio no incluía tratamientos o procedimientos invasivos, por lo que no existieron riesgos potenciales asociados al procedimiento de esta investigación. Se respetó el principio de confidencialidad y se garantizó que la información proporcionada por los participantes fuera utilizada con fines científicos. Se comunicó directamente los objetivos del estudio, el procedimiento y las posibles molestias de este a los participantes. Después de tener claro el proceso, se buscó la autorización formal para que participaran en el estudio mediante la firma del consentimiento informado.

3. RESULTADOS

3.1 Elaboración de ítems

Las palabras tenían los mismos valores de longitud, frecuencia léxica, familiaridad, concreción e imaginabilidad entre todas las condiciones de la tarea (Tabla 2). Además, hubo diferencias significativas en la similitud semántica entre las palabras blanco y los distractores tanto en los ítems de baja demanda ($t(166) = 59.47, p < 0.001$) como de alta demanda ($t(166) = 60.18, p < 0.001$). Las palabras concretas mostraron valores más altos de concreción ($t(390) = 39.01, p < 0.001$) e imaginabilidad ($t(384) = 33.98, p < 0.001$) en comparación con las palabras abstractas en todos los ítems.

1). Se seleccionaron los ítems de baja y alta demanda cognitiva que mostraran los parámetros de discriminación más altos y que no presentaran problemas de ajuste al modelo, quedando la tarea conformada por 12 ítems (6 ítems de baja demanda y 6 ítems de alta demanda). En la tabla 3 se presentan los parámetros de dificultad y discriminación de los 12 ítems y en la Figura 4 las curvas características. Todos ellos muestran un adecuado ajuste global al modelo.

Tabla 3.

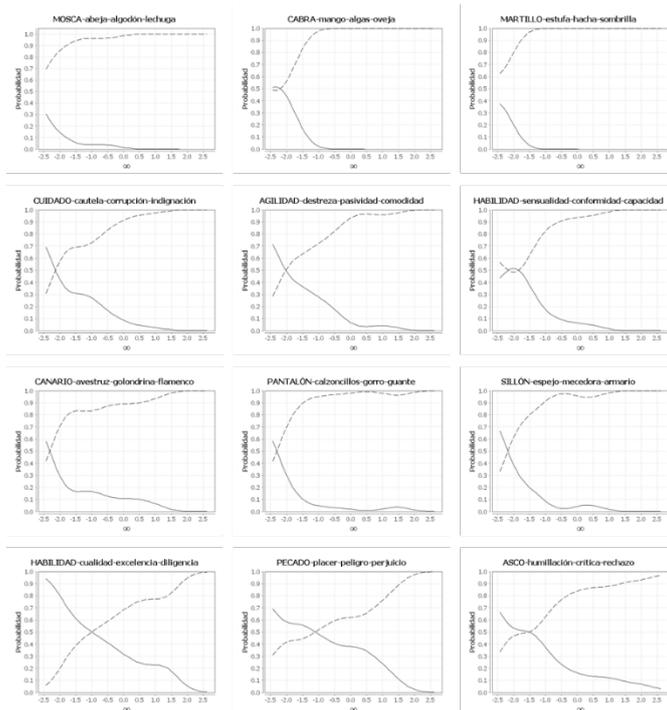
Parámetros de dificultad (b) y discriminación (a) con error estándar (EE) de los 12 ítems

Condición	Ensayo	a	EE	b	EE	X ²	gl	p
Baja demanda	MOSCA-abeja-algodón-lechuga	1.10	0.43	-3.82	1.16	3.09	5	0.69
	CABRA-mango-algas-oveja	1.56	0.45	-2.89	0.56	2.66	4	0.62
	MARTILLO-estufa-hacha-sombrilla	1.82	0.55	-3.33	0.64	0.99	4	0.91
	CUIDADO-cautela-corrupción-indignación	1.79	0.36	-1.61	0.22	3.79	5	0.58
	AGILIDAD-destreza-pasividad-comodidad	1.37	0.31	-1.80	0.31	5.98	4	0.20
	HABILIDAD-sensualidad-conformidad-capacidad	1.36	0.33	-2.11	0.38	5.17	4	0.27
Alta demanda	CANARIO-avestruz-golondrina-flamenco	0.85	0.26	-2.74	0.73	0.91	4	0.92
	PANTALÓN-calzoncillos-gorro-guante	1.38	0.40	-2.80	0.58	1.42	4	0.84
	SILLÓN-espejo-mecedora-armario	1.52	0.38	-2.34	0.41	1.89	4	0.76
	HABILIDAD-cualidad-excelencia-diligencia	0.76	0.19	-0.97	0.31	7.54	4	0.11
	PECADO-placer-peligro-perjuicio	0.78	0.19	-0.90	0.30	2.78	5	0.73
	ASCO-humillación-crítica-rechazo	1.05	0.24	-1.51	0.32	6.80	4	0.15

Nota. Palabras prueba en MAYÚSCULAS, palabras blanco en cursiva

Figura 3.

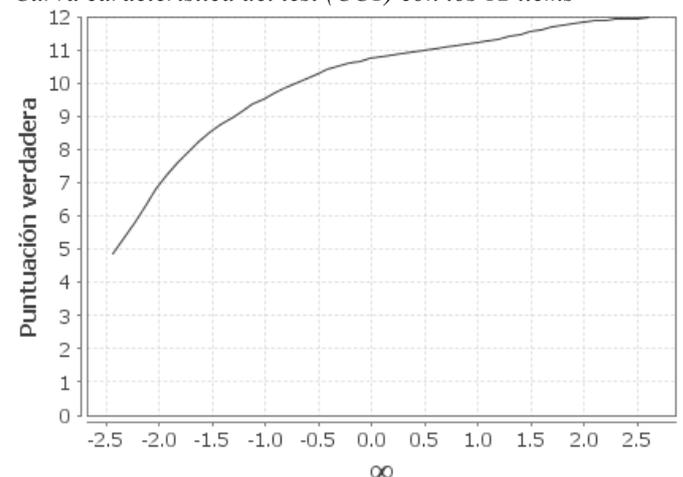
Curvas características de los 12 ítems



La pendiente de la CCT indica el grado en que el valor de la puntuación verdadera depende del nivel de aptitud. Se puede observar que el nivel de dificultad de la tarea se localiza en un nivel de -2.0. De modo que la puntuación total de los 12 ensayos proporciona más información en torno a los niveles de rendimiento de -2.0, es decir, en niveles bajos de habilidad (Figura 4).

Figura 4.

Curva característica del test (CCT) con los 12 ítems

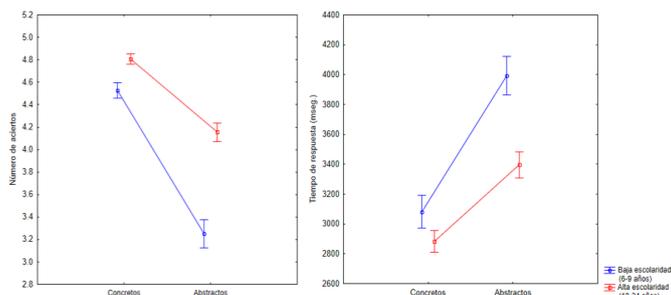


3.3 Análisis psicométrico de la versión reducida

Se estudió la posible influencia de la edad y escolaridad sobre el rendimiento a partir de la aplicación de los 12 ítems en los 149 participantes. Se observó que la edad no tiene un efecto significativo sobre el número de aciertos ($F(2,143) = 2.50; p = 0.08; \eta^2 = 0.03$) ni sobre el tiempo de respuesta ($F(2,143) = 0.25; p = 0.78; \eta^2 = 0.003$). En cambio, hubo un efecto significativo de la escolaridad sobre el número de aciertos ($F(2,143) = 31.99; p < 0.001; \eta^2 = 0.18$) y los tiempos de respuesta ($F(2,143) = 8.87; p = 0.003; \eta^2 = 0.06$). Asimismo, hubo un efecto significativo de interacción entre la escolaridad y tipo de palabra sobre el número de aciertos ($F(2,143) = 23.93; p < 0.001; \eta^2 = 0.14$) y los tiempos de respuesta ($F(2,143) = 13.12; p < 0.001; \eta^2 = 0.08$). Es decir, las personas con baja escolaridad responden menos rápido y con menor precisión ante palabras abstractas en comparación a las concretas (Figura 5).

Figura 5.

Efectos de interacción Tipo*Escolaridad sobre el rendimiento



En la tabla 4 se presentan las medias y desviaciones típicas de la versión reducida para su aplicación clínica, ajustados por grupo de edad (19-34 años, 35-59 años y mayores de 60 años) y nivel educativo (baja escolaridad: 6 a 9 años; alta escolaridad: 10 a 24 años).

Tabla 4.
Medias y desviaciones estándar (DE) de la tarea experimental

	Jóvenes		Adultos		Adultos mayores	
	Media	DE	Media	DE	Media	DE
Baja escolaridad						
Número de aciertos	14.11	1.79	16.00	2.69	16.56	2.49
Tiempo de respuesta	3523	602	3451	724	3560	734
Alta escolaridad						
Número de aciertos	17.97	1.63	18.08	1.72	17.72	2.19
Tiempo de respuesta	3014	606	3285	587	3125	701
Total						
Número de aciertos	17.14	2.73	17.28	2.35	17.29	2.36
Tiempo de respuesta	3559	634	3349	639	3268	739

4. DISCUSIÓN

El objetivo de este estudio fue desarrollar y validar una tarea de asociación semántica que pueda implementarse de manera experimental y clínica. Son escasas las tareas destinadas a la evaluación neuropsicológica general que permitan diferenciar los diferentes tipos de procesamiento semántico, particularmente del control semántico. Asimismo, la mayoría de estas tareas han empleado mayoritariamente conceptos concretos pertenecientes a dos dominios semánticos (seres vivos y objetos inanimados), sin incluir conceptos abstractos ni otros dominios o categorías de conocimiento. El test de Pirámides y Palmeras es un ejemplo de pruebas que pueden ser útiles para investigar el control semántico al emparejar figuras o palabras según su asociación semántica. No obstante, la impronta cultural que poseen los estímulos hace que la prueba no sea adecuada para todas las personas, incluso dentro de la comunidad hispanoparlante iberoamericana y que por ello sea deseable el desarrollo de tareas o pruebas con base en estrategias para disminuir los sesgos culturales asociados a los estímulos, como el análisis

psicométrico de los ítems con métodos novedosos como la TRI (Fernández & Abe, 2018; Merkle et al., 2023).

Los ítems en esta tarea variaban en el nivel de demanda cognitiva, además de que se incluyeron palabras con significado concreto y abstracto de manera que se pueden observar y analizar los efectos diferenciales entre ellos. Asimismo, las palabras fueron seleccionadas en función de diversas variables psicolingüísticas que pueden sesgar las ejecuciones, tales como la frecuencia léxica (Oldfield & Wingfield, 1965), familiaridad (Funnell & Sheridan, 1992), imaginabilidad y concreción (Kroll & Merves, 1986; Paivio, 1991). Además de estas consideraciones, las palabras eran altamente familiares, reduciendo así el efecto que pueda tener la cultura o escolaridad en el procesamiento del significado de las palabras. El control de todas estas variables constituye una de las bondades de la tarea, a diferencia de otras tareas existentes, de modo que se puede aislar el impacto de la degeneración cortical sobre la memoria semántica.

Una de las principales fortalezas de este estudio es la utilización de la TRI como técnica para la validación de la tarea de asociación semántica. A diferencia de los métodos tradicionales basados en la TCT, la TRI permite analizar el comportamiento de cada ítem de manera individual, estimando parámetros como dificultad y discriminación, lo cual aporta mayor precisión y robustez a la medición del rasgo latente. Las curvas características de los ensayos muestran una adecuada capacidad de discriminación en niveles bajos de habilidad, lo cual es conveniente para reducir el posible efecto de piso o para evaluar a sujetos con habilidades limitadas por daño cerebral. Además, una versión corta de esta tarea con 12 ítems resulta benéfica en pacientes con daño cerebral ya que se disminuye la fatiga mental. La TRI se ha utilizado escasamente para validar instrumentos neuropsicológicos, por lo que con este trabajo se continúan los pasos hacia la creación y validación de nuevas herramientas que evalúen la memoria semántica y otras habilidades cognitivas en población mexicana y con base en la TRI.

Un hallazgo relevante del presente estudio es que la edad de los participantes no tuvo un efecto significativo en el número de aciertos ni en los tiempos de respuesta. Este resultado sugiere que el desempeño en tareas semánticas se mantiene relativamente estable a lo largo del ciclo vital. En particular, los adultos mayores parecen conservar depósitos sólidos de conocimiento conceptual, a pesar de los cambios neuroanatómicos asociados al envejecimiento, como la reducción del volumen cortical. De hecho, la amplitud del conocimiento semántico acumulado con la edad podría mitigar los efectos de la demanda cognitiva durante el procesamiento lingüístico, lo que se traduce en un menor impacto de la dificultad en las decisiones semánticas. Este hallazgo es coherente con investigaciones previas que han mostrado que la memoria semántica permanece relativamente preservada con el envejecimiento, en contraste con otros sistemas de memoria más vulnerables, como la memoria episódica (Nyberg et al., 2012).

En contraste, la escolaridad sí tuvo un efecto significativo en el desempeño, lo que respalda la idea de que la experiencia educativa formal contribuye de manera importante al procesamiento semántico. La escolaridad podría funcionar como un factor protector, especialmente en tareas

que exigen un acceso ágil y un manejo eficiente del conocimiento conceptual. Además, se observó un efecto de interacción entre la escolaridad y el tipo de concepto (concreto-abstracto). Específicamente, los participantes con menor escolaridad tendieron a responder de forma más lenta y menos precisa ante palabras abstractas, en comparación con aquellos con mayor escolaridad. Este resultado podría deberse a que las personas con un nivel educativo alto tienden a poseer redes semánticas más complejas y enriquecidas, que incluyen tanto conceptos concretos como abstractos. En contraste, las personas con baja escolaridad podrían estar menos expuestas o familiarizadas con términos abstractos, lo que afecta su procesamiento. Estos hallazgos coinciden con estudios previos que subrayan el papel de la estimulación educativa en el rendimiento cognitivo, especialmente cuando se procesan palabras abstractas (Hill et al., 2014; Hoffman & Woollams, 2015; Jockwitz et al., 2019).

En conclusión, la relevancia de este estudio radica en el diseño y validación de una tarea de asociación semántica con base en la TRI, controlando el nivel de demanda cognitiva y cuyos estímulos incluyen estímulos concretos y abstractos, balanceados de manera rigurosa en diferentes variables que influyen sobre el procesamiento semántico. Esta tarea se puede emplear de manera experimental y clínica para identificar déficits del conocimiento conceptual en pacientes con daño cerebral. Para cumplir con este propósito, será necesario el desarrollo de estudios futuros de validez externa de la tarea en distintos grupos clínicos, lo cual, al mismo tiempo representa una desventaja del presente estudio. Con respecto a las posibles limitaciones de la tarea, puede señalarse la validación a partir del modelo de dos parámetros con los parámetros de dificultad y discriminación. En consecuencia, la posibilidad de adivinar la respuesta correcta a un ítem es alta. Este estudio podría extenderse al modelo de tres parámetros y con ello a la estimación del parámetro de adivinación. Otra limitación es que no se consideró el nivel socioeconómico de los participantes como variable en los análisis. Este factor podría influir en el desempeño en tareas lingüísticas o semánticas, especialmente en contextos transculturales, por lo que futuras investigaciones deberían contemplarlo para un análisis más completo.

Referencias

- Aguilar-Navarro, S. G., Mimenza-Alvarado, A. J., Palacios-García, A. A., Samudio-Cruz, A., Gutiérrez-Gutiérrez, L. A., & Ávila-Funes, J. A. (2018). Validez y confiabilidad del MoCA (Montreal Cognitive Assessment) para el tamizaje del deterioro cognoscitivo en México. *Revista Colombiana de Psiquiatría*, 47(4), 237-243. <https://doi.org/10.1016/j.rcp.2017.05.003>
- Ardila, A. (2007). The impact of culture on neuropsychological test performance. En B. P. Uzzell, M. Pontón, & A. Ardila (Eds.), *International handbook of cross-cultural neuropsychology* (pp. 23-44). Mahwah: Lawrence Erlbaum Associates.
- Baker, F. B. (2001) *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation.
- Béghin, G. & Markovitz, H. (2022). Reasoning strategies and prior knowledge effects in contingency learning. *Memory and Cognition*, 50 (6), 1269-1283. <https://doi.org/10.3758/s13421-022-01319-w>
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15(11), 527-536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767-2796. <https://doi.org/10.1093/cercor/bhp055>
- Brisson, J. & Markovitz, H. (2020). Reasoning strategies and semantic memory effects in deductive reasoning. *Memory and Cognition*, 48 (6), 920-930. <https://doi.org/10.3758/s13421-020-01027-3>
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Springer. <http://dx.doi.org/10.1007/978-94-007-6857-4>
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, 10(1), 1-34. <https://doi.org/10.1162/089992998563752>
- Chiou, R., Humphreys, G. F., Jung, J., & Lambon Ralph, M. A. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex; a journal devoted to the study of the nervous system and behavior*, 103, 100-116. <https://doi.org/10.1016/j.cortex.2018.02.018>
- Coutanche, M. N., & Thompson-Schill, S. L. (2015). Creating Concepts from Converging Features in Human Cortex. *Cerebral Cortex*, 25(9), 2584-2593. <https://doi.org/10.1093/cercor/bhu057>
- Damasio A. R. (1990). Category-related recognition defects as a clue to the neural substrates of knowledge. *Trends in Neurosciences*, 13(3), 95-98. [https://doi.org/10.1016/0166-2236\(90\)90184-c](https://doi.org/10.1016/0166-2236(90)90184-c)
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. The Guilford Press.
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: one-stop shopping for Spanish 190m properties. *Behavior research methods*, 45(4), 1246-1258. <https://doi.org/10.3758/s13428-013-0326-1>
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 515-523. <https://doi.org/10.1037/0278-7393.24.2.515>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates Publishers.
- Fernández, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain*, 6, 1-35. <https://doi.org/10.1007/s40167-017-0050-2>
- Frisby, S. L., Halai, A. D., Cox, C. R., Lambon Ralph, M. A. & Rogers, T. T. (2023). Decoding semantic representations in mind and brain. *Trends in Cognitive Sciences* 27 (3), 258-281. <https://doi.org/10.1016/j.tics.2022.12.006>
- Funnell, E., & Sheridan, J. (1992). Categories of knowledge? Unfamiliar aspects of living and nonliving things. *Cognitive Neuropsychology*, 9(2), 135-153. <https://doi.org/10.1080/02643299208252056>
- Gudayol-Ferré, E., Lara, J. P., Herrera-Guzman, I., Böhm, P., Rodés, E., Ansaldó, A. I., & Peña-Casanova, J. (2008). Semantic memory as assessed by the Pyramids and Palm Trees Test: The impact of sociodemographic factors in a Spanish-speaking population. *Journal of the International Neuropsychological Society*, 14(1), 148-151. <https://doi.org/10.1017/S1355617708080168>
- Hill, F., Korhonen, A., & Bentz, C. (2014). A quantitative empirical 190mo f19 of the abstract/concrete distinction. *Cognitive*

- Science*, 38(1), 162–177. <https://doi.org/10.1111/cogs.12076>
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain: a journal of neurology*, 115 (Pt 6), 1783–1806. <https://doi.org/10.1093/brain/115.6.1783>
- Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance*, 41(2), 385–402. <https://doi.org/10.1037/a0038995>
- Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees: A test of semantic 20ormo from pictures and words*. Thames Valley Test Company
- Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "category-specific" neuropsychological deficits. *The Behavioral and Brain Sciences*, 24(3), 453–509. <https://pubmed.ncbi.nlm.nih.gov/11682799/>
- Jockwitz, C., Mérillat, S., Liem, F., Oschwald, J., Amunts, K., Caspers, S., & Jäncke, L. (2019). Generalizing age effects on brain structure and cognition: A two-study comparison approach. *Human Brain Mapping*, 40(8), 2305–2319. <https://doi.org/10.1002/hbm.24524>
- Jurado, S., Villegas, M. E., Méndez, L., Rodríguez, F., Loperana, V. & Varela, R. (1998). La estandarización del Inventario de Depresión de Beck para los residentes de la ciudad de México. *Salud Mental*, 21(3), 26-31. https://revistasaludmental.gob.mx/index.php/salud_mental/article/view/706/0
- Kiran, S., & Thompson, C. K. (2003). The Role of Semantic Complexity in Treatment of Naming Deficits: Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research*, 46(4), 773–787. [https://doi.org/10.1044/1092-4388\(2003/061\)](https://doi.org/10.1044/1092-4388(2003/061))
- Koenig, P., & Grossman, M. (2007). Process and content in semantic memory. In J. Hart, Jr. & M. A. Kraut (Eds.), *Neural basis of semantic memory* (pp. 247–264). Cambridge University Press. <https://doi.org/10.1017/CBO9780511544965.011>
- Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(1), 92–107. <https://doi.org/10.1037/0278-7393.12.1.92>
- Markovitz, H. (2024). Are There Two Kinds of Reasoners? *Journal of Intelligence*, 12 (3), 25. <https://doi.org/10.3390/jintelligence12030025>
- Martínez-Cuitiño, M., Barreyro, J., & Jaichenco, V. (2009). Adaptación y validación en español de una herramienta de evaluación semántica: la Batería 64. *Revista Neuropsicología Latinoamericana*, 1(1), 24–31. https://www.neuropsicolatina.org/index.php/Neuropsicologia_Latinoamericana/article/view/5
- McCarthy, R. A., & Warrington, E. K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature*, 334(6181), 428–430. <https://doi.org/10.1038/3344280>
- McCarthy, R. A., & Warrington, E. K. (2016). Past, present, and prospects: Reflections 40 years on from the selective impairment of semantic memory (Warrington, 1975). *Quarterly Journal of Experimental Psychology* (2006), 69(10), 1941–1968. <https://doi.org/10.1080/17470218.2014.980280>
- Meyer, J.P. (2014). *Applied Measurement with jMetrik* (1. Ed.). Routledge. <https://doi.org/10.4324/9780203115190>
- Merkley, T. L., Esopenko, C., Zizak, V. S., Bilder, R. M., Strutt, A. M., Tate, D. F., & Irimia, A. (2023). Challenges and opportunities for harmonization of cross-cultural neuropsychological data. *Neuropsychology*, 37(3), 237–246. <https://doi.org/10.1037/neu0000818>
- Mkrtychian, N., Blagovechtchenski, E., Kurmakava, D., Gnedyk, D., Kostromina, S., & Shtyrov, Y. (2019). Concrete vs. abstract semantics: From mental representations to functional brain mapping. *Frontiers in Human Neuroscience*, 13, Article 267. <https://doi.org/10.3389/fnhum.2019.00267>
- Moreno Martínez, F. J., & Cañamón, S. (2005). Presentación y resultados preliminares de la Batería Nombela (I). Un nuevo instrumento para evaluar el deterioro semántico categorial. *Revista de Psicopatología y Psicología Clínica*, 10(3), 205–219. <https://doi.org/10.5944/rppc.vol.10.num.3.2005.4003>
- Moreno-Martínez, F. J., & Rodríguez-Rojo, I. C. (2015). The Nombela 2.0 semantic battery: an updated Spanish instrument for the study of semantic processing. *Neurocase*, 21(6), 773–785. <https://doi.org/10.1080/13554794.2015.1006644>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, 16(5), 292–305. <https://doi.org/10.1016/j.tics.2012.04.005>
- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology*, 17(4), 273–281. <https://doi.org/10.1080/17470216508416445>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, 45(3), 255–287. <https://doi.org/10.1037/h0084295>
- Peraita Adrados, H., González Labra, M.J., Sánchez Bernardos, M.L. & Galeote Moreno, M.A. (2000). Batería de Evaluación del Deterioro de la Memoria Semántica en Alzheimer. *Psicothema*, 12, 192-200. <https://www.psicothema.com/pdf/276.pdf>
- Pescatore, C. R., Zhang, H., Hadjinicolaou, A. Paulk, A. C., Rolston, J. D., Richardson, R. M., Williams, Z. M., Cai, J., & Cash, S. Decoding semantics from natural speech using human intracranial EEG. (2025). *BioRxiv* <https://doi.org/10.1101/2025.02.10.637051>
- Pestilli F. (2018). Human white matter and knowledge representation. *PLoS biology*, 16(4), e2005758. <https://doi.org/10.1371/journal.pbio.2005758>
- Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature reviews. Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Salgado, Y. K., Andrade P., Hernández-Galván, A., González-Arriata, N. I., Díaz, R. & Velasco, A. E. (2024). Validación de la Escala de Depresión Geriátrica de Yesavage en adultos mayores mexicanos. (2024). *Informes Psicológicos*, 24(1), 123-135. <https://doi.org/10.18566/infpsic.v24n1a08>
- Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioural Neurology*, 2(3), 167–182.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, *Organization of memory*. Academic Press.

- Van der Linden, W. J., & Glas, C. A. W. (Eds.). (2010). *Elements of adaptive testing*. Springer.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *The Quarterly Journal of Experimental Psychology*, 27(4), 635–657. <https://doi.org/10.1080/14640747508400525>
- Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge. Further fractionations and an attempted integration. *Brain: a journal of neurology*, 110 (Pt 5), 1273–1296. <https://doi.org/10.1093/brain/110.5.1273>
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain: a journal of neurology*, 107 (Pt 3), 829–854. <https://doi.org/10.1093/brain/107.3.829>
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M. B., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)

Anexo 1. Parámetros de dificultad (*b*) y discriminación (*a*) con errores estándar (*EE*) de los 109 ensayos

Condición	Ensayo	<i>a</i>	<i>EE</i>	<i>b</i>	<i>EE</i>	χ^2	<i>gl</i>	<i>p</i>
Baja demanda	CONEJO-liebre-amapola-aceituna	0.64	0.55	-8.10	6.60	27.14	19	0.10
	MOSCA-abeja-algodón-lechuga	1.28	0.45	-3.38	0.89	10.58	18	0.91
	CANARIO-cigüeña-naranja-coliflor	0.87	0.28	-2.80	0.79	19.58	19	0.42
	ROSA-rana-clavel-araña	0.74	0.36	-5.58	2.40	28.09	21	0.14
	CABALLO-arbusto-yegua-papa	0.76	0.18	-5.90	0.95	35.90	19	0.01
	VACA-caoba-toro-hiedra	1.29	0.49	-3.61	1.03	13.19	19	0.83
	CABRA-mango-algas-oveja	1.81	0.46	-2.60	0.44	14.69	19	0.74
	MANZANA-paloma-tortuga-pera	1.05	0.59	-5.28	2.51	22.41	21	0.38
	POLLO-jazmín-olivo-gallina	1.18	0.63	-4.81	2.08	10.23	21	0.98
	CEBOLLA-pasa-pato-ajo	0.94	0.40	-3.97	1.41	11.28	20	0.94
	NUEZ-castaña-ballena-iguana	1.02	0.36	-3.38	1.00	16.22	19	0.64
	PINO-ciprés-sapo-carnero	1.00	0.31	-2.81	0.73	14.56	21	0.84
	CAMA-camilla-boina-traje	0.77	0.20	-5.88	1.10	17.92	21	0.65
	FALDA-medias-pala-pico	1.61	0.64	-3.93	1.10	0.59	19	1.00
	SARTÉN-sotana-cazuela-tijeras	1.28	0.49	-3.64	1.05	14.08	19	0.78
	BOTELLA-reloj-jarrón-banca	1.31	0.49	-3.58	1.00	21.58	20	0.36
	MARTILLO-estufa-hacha-sombrilla	1.85	0.54	-3.21	0.61	3.92	19	1.00
	PANTALÓN-lavadero-escopeta-cinturón	0.89	0.42	-4.45	1.81	16.45	20	0.69
	SILLA-llave-bolso-banco	0.89	0.33	-3.42	1.08	15.10	19	0.72
	VASO-avión-coche-copa	0.58	0.09	-5.95	0.51	14.54	19	0.75
	CHAQUETA-regla-hacha-abrigo	0.63	0.15	-5.89	1.06	25.59	20	0.18
	LÁPIZ-goma-bolsa-buzón	1.18	0.63	-4.81	2.08	10.10	20	0.97
	SILLÓN-sofá-pañuelo-bolígrafo	0.86	0.18	-5.96	0.43	27.80	19	0.09
	ZAPATO-bota-ventana-liga	0.34	0.34	-12.74	12.50	16.07	20	0.71
	MESA-escriptorio-esmoquin-automóvil	1.61	0.64	-3.93	1.10	0.59	19	1.00
	ÉXITO-triunfo-ruego-matiz	0.70	0.36	-5.12	2.40	13.33	20	0.86
	FIDELIDAD-amistad-soledad-desorden	0.93	0.29	-2.77	0.75	18.36	20	0.56
	RESPETO-admiración-desgracia-promesa	0.19	0.17	-10.29	8.89	19.51	21	0.55
	CUIDADO-cautela-corrupción-indignación	1.56	0.34	-1.65	0.27	17.15	19	0.58
	CONFIANZA-seguridad-agresividad-desigualdad	0.77	0.19	-5.89	1.05	16.93	20	0.66
	SENCILLEZ-humildad-ganancia-melancolía	0.65	0.25	-5.75	1.92	23.85	20	0.25
	GENEROSIDAD-amabilidad-costumbre-criterio	0.80	0.26	-2.66	0.77	17.54	18	0.49
	AGILIDAD-destreza-pasividad-comodidad	1.48	0.33	-1.66	0.29	18.18	21	0.64
ORGULLO-pasión-honor-modo	0.39	0.17	-0.59	0.49	10.30	20	0.96	
INTELIGENCIA-impaciencia-sabiduría-influencia	1.06	0.40	-3.64	1.14	24.44	21	0.27	
VERDAD-sanción-certeza-placer	0.72	0.23	-2.49	0.76	23.51	21	0.32	
BONDAD-severidad-caridad-vitalidad	0.82	0.24	-2.24	0.61	17.76	19	0.54	
MORAL-dogma-complejidad-ética	0.89	0.23	-1.65	0.42	18.63	21	0.61	
HABILIDAD-sensualidad-conformidad-capacidad	1.40	0.34	-2.01	0.37	19.80	21	0.53	
PECADO-esencia-intuición-lujuria	0.84	0.24	-2.00	0.53	21.90	21	0.41	
EGOÍSMO-coherencia-elegancia-avaricia	0.91	0.28	-2.55	0.67	29.52	20	0.08	
MALDAD-crueldad-juventud-armonía	1.28	0.45	-3.37	0.89	18.45	21	0.62	
PEREZA-cansancio-belleza-entrega	1.39	0.40	-2.72	0.58	21.93	22	0.46	
MANÍA-obsesión-aprecio-encanto	1.08	0.27	-1.92	0.42	14.34	20	0.81	
SOSPECHA-duda-estima-alegría	0.79	0.34	-4.14	1.60	16.74	20	0.67	
MENTIRA-facilidad-falsedad-libertad	1.31	0.36	-2.49	0.52	24.65	19	0.17	

	LIMPIEZA-astucia-orden-esmero	0.34	0.19	-4.83	2.65	24.42	20	0.22
	DEVOCIÓN-cortesía-fe-superación	0.76	0.24	-2.54	0.75	16.82	21	0.72
	VENGANZA-emoción-rencor-acierto	1.14	0.31	-2.30	0.51	25.01	21	0.25
	IRONÍA-exigencia-optimismo-sarcasmo	1.45	0.30	-1.25	0.22	21.47	22	0.49
	TONTERÍA-retórica-insistencia-estupidez	0.74	0.24	-2.58	0.77	18.42	20	0.56
	ERROR-agrado-dicha-fallo	1.11	0.36	-3.02	0.78	18.11	21	0.64
	ASCO-vigor-gozo-desprecio	0.55	0.27	-5.43	2.51	21.04	22	0.52
Alta demanda	TIGRE-jabalí-pantera-elefante	0.52	0.27	-5.02	2.41	15.61	19	0.68
	CONEJO-rata-ratón-ardilla	0.39	0.19	-3.76	1.84	16.46	20	0.69
	MOSCA-hormiga-mariposa-gusano	0.28	0.16	-2.13	1.31	22.35	19	0.27
	CANARIO-avestruz-golondrina-flamenco	0.84	0.27	-2.70	0.76	19.11	22	0.64
	ROSA-lirio-margarita-geranio	0.14	0.12	-2.73	2.51	20.94	22	0.52
	CABALLO-búfalo-ciervo-burro	0.80	0.30	-3.47	1.16	30.14	19	0.05
	PERRO-chivo-cerdo-zorro	0.64	0.22	-2.69	0.90	17.70	20	0.61
	VACA-mono-foca-buey	0.67	0.12	-5.95	0.53	30.16	22	0.11
	CABRA-cordero-coyote-oso	0.80	0.29	-3.23	1.03	19.83	20	0.47
	MANZANA-cereza-zanahoria-plátanos	0.31	0.16	-1.37	0.87	15.97	21	0.77
	POLLO-pavo-cisne-águila	0.52	0.10	-5.92	0.75	23.61	18	0.17
	CEBOLLA-tomate-patata-lentejas	0.26	0.09	-5.77	1.94	16.30	20	0.70
	NUEZ-higos-almendra-garbanzos	0.61	0.29	-4.56	2.02	12.39	21	0.93
	PINO-cactus-roble-bambú	1.04	0.24	-1.30	0.30	30.98	21	0.07
	CAMA-baúl-cuna-aparador	1.26	0.54	-4.01	1.32	16.97	21	0.71
	FALDA-corbata-vestido-sombrero	0.75	0.38	-5.50	2.49	20.94	21	0.46
	SARTÉN-vasija-cuchillo-olla	0.36	0.17	-2.44	1.21	26.71	21	0.18
	BOTELLA-tenedor-servilleta-frasco	1.09	0.61	-5.13	2.40	24.84	21	0.25
	CAMISA-bufanda-bolsillo-chaleco	0.31	0.10	-5.79	1.75	20.11	20	0.45
	MARTILLO-pala-sierra-hoz	0.07	0.08	18.31	19.11	29.21	22	0.14
	PANTALÓN-calzoncillos-gorro-guante	1.46	0.41	-2.63	0.53	28.96	21	0.12
	SILLA-butaca-repisa-estante	0.97	0.43	-4.15	1.54	20.17	21	0.51
	VASO-taza-paño-mantel	0.79	0.29	-5.78	1.76	18.36	21	0.63
	CHAQUETA-jersey-gorra-gafas	0.80	0.21	-0.82	0.28	14.79	21	0.83
	LÁPIZ-tijeras-bolígrafo-regla	0.85	0.25	-2.30	0.61	12.98	21	0.91
	SILLÓN-espejo-mecedora-armario	1.57	0.39	-2.23	0.39	24.12	20	0.24
	ZAPATO-camiseta-calcetín-minifalda	0.80	0.43	-5.21	2.45	17.73	20	0.61
	MESA-ropero-mesilla-vitrina	0.29	0.05	-5.95	0.49	15.99	20	0.72
	ÉXITO-logro-auge-fama	0.39	0.17	-5.61	2.44	26.80	19	0.11
	FIDELIDAD-lealtad-gratitud-compromiso	0.51	0.19	-2.34	0.90	22.14	20	0.33
	RESPETO-obediencia-tolerancia-cordialidad	0.06	0.07	12.70	13.86	25.14	21	0.24
	CUIDADO-precaución-prudencia-discreción	0.31	0.16	-1.76	1.04	16.43	19	0.63
CONFIANZA-objetividad-sinceridad-integridad	0.07	0.03	-5.83	1.63	25.31	21	0.23	
SENCILLEZ-pobreza-modestia-necesidad	1.03	0.26	-1.72	0.39	54.73	20	0.00	
GENEROSIDAD-fraternidad-solidaridad-compasión	0.63	0.19	-0.91	0.37	31.61	19	0.03	
AGILIDAD-creatividad-ingenio-experiencia	0.58	0.18	0.75	0.36	19.65	21	0.54	
ORGULLO-nobleza-vergüenza-prestigio	0.08	0.08	3.26	3.83	26.06	20	0.16	
INTELIGENCIA-calidad-perfección-aptitud	0.15	0.12	-0.26	1.13	18.29	21	0.63	
VERDAD-seguridad-convicción-acierto	0.38	0.16	-0.53	0.50	23.86	21	0.30	
BONDAD-suavidad-benevolencia-generosidad	0.13	0.03	-5.91	0.87	22.38	20	0.32	

MORAL-justicia-conciencia-madurez	0.34	0.16	-0.27	0.51	24.62	20	0.22
HABILIDAD-cualidad-excelencia-diligencia	1.07	0.23	-0.72	0.21	10.91	20	0.95
PECADO-placer-peligro-perjuicio	0.73	0.20	-0.93	0.33	15.26	19	0.71
EGOÍSMO-ambición-deseo-envidia	0.18	0.13	-2.62	2.11	30.32	21	0.09
MALDAD-violencia-malicia-tiranía	0.57	0.18	-0.57	0.35	17.84	19	0.53
PEREZA-placidez-quietud-desinterés	0.32	0.18	-3.87	2.13	19.77	21	0.54
MANÍA-delirio-locura-frenesí	0.26	0.15	1.45	1.04	14.46	21	0.85
SOSPECHA-escepticismo-desconfianza-incredulidad	0.18	0.10	-5.55	3.03	20.35	20	0.44
MENTIRA-farsa-ficción-engaño	0.44	0.18	-2.36	1.02	26.46	20	0.15
LIMPIEZA-disciplina-crueldad-pureza	0.48	0.17	-0.19	0.36	20.02	22	0.58
DEVOCIÓN-sacrificio-creencia-religión	0.11	0.10	-0.10	1.46	20.29	21	0.50
VENGANZA-enojo-traición-odio	0.44	0.17	0.82	0.49	24.96	20	0.20
IRONÍA-burla-humor-comedia	0.24	0.05	-5.90	0.96	20.72	20	0.41
TONTERÍA-broma-ingenuidad-debilidad	0.13	0.11	4.19	3.74	13.37	22	0.92
ERROR-fracaso-frustración-decepción	0.57	0.18	-0.41	0.32	14.58	21	0.84
ASCO-humillación-crítica-rechazo	0.86	0.23	-1.70	0.44	20.94	21	0.46
