

El efecto de generación en la producción de falsas memorias

Marisol Martínez Miramontes^{1*}, Carlos F. A. Gomes¹, Ambrocio Mojardín-Heráldez² y Lillian Milnitsky Stein^{1*}

¹ Pontificia Universidade Católica do Rio Grande do Sul (Brasil)

² Universidad Autónoma de Sinaloa (México)

Resumen: Efecto de Generación (EG) es el hecho de recordar mejor información que fue generada por nosotros, en comparación a aquella información que solo leímos. Las Falsas Memorias (FM), se refieren a recordar un hecho que no ocurrió u ocurrió de manera distinta de cómo es recordado. Podríamos suponer que el EG podría producir altos índices de FM, dado que EG y FM son explicados como debido a un procesamiento semántico. No obstante, algunos estudios han demostrado la ventaja que implica el EG para la memoria sin un incremento en FM. El objetivo general del presente estudio fue explorar el impacto del EG en Memorias Verdaderas (MV) y FM, además conocer su persistencia al tiempo, el efecto de mera-evaluación y los índices de certeza en el EG. Los resultados indicaron que el EG mostró su efecto clásico en el incremento de índices en MV, sin ir acompañados por el aumento de FM. El EG no dio evidencia de ser persistente al tiempo, mostrando altos índices de confianza iniciales mismo que disminuyeron en más de 50% con el tiempo. Además, se observó que la repetición de evaluación de ítems beneficia las MV (mera-evaluación).

Palabras clave: memoria; Efecto de Generación; falsas memorias; mera-evaluación; certidumbre.

Title: The Generation Effect in the production of false memories.

Abstract: The Generation Effect (GE) is the effect of better recalling information that was created by us in comparison to information that was only read. False Memories (FM) refer to recalling information that did not happen or that happened in a different way. We may hypothesize that the GE can produce higher levels of FM, as both the GE and FM have been explained in terms of semantic processing. Nevertheless, some studies have shown that the GE improves memory without increasing FM. The aim of the present study was to explore the impact of the GE on True Memories (TM) and on FM, as well as its influence on the time of testing, the mere-testing effect, and confidence level. The results indicated the classical GE: increased rates of TM without increases in FM. The GE was not persistent across time, showing a decline of more than 50% in confidence level over time. In addition, repetition enhanced TM (mere-testing effect).

Key words: memory; Generation Effect; false memories; mere-testing; confidence rating.

Introducción

En el área de la psicología de la memoria, por más de dos décadas, se ha estado investigando un fenómeno conocido como Efecto de Generación (EG) (Slamecka & Graf, 1978). Este fenómeno consiste en una mejora sustancial del recuerdo para información que es generada por nosotros mismos (e.g., categoría, pájaro: canario – po___), comparada con el recuerdo para información que solo leímos o escuchamos (e.g., categoría, pájaro canario-pombo).

El EG ha sido ampliamente estudiado y es ahora considerado una ventaja para la memoria y para las tareas de aprendizaje en general (DeWinstaley & Bjork, 2004). Las razones básicas son dos: a) Ayuda a retener mayores cantidades de información (Taconnat et al., 2008) y b) La memoria que promueve es de mayor duración en el tiempo (Crutcher & Healey, 1989). Es decir, el EG nos permite retener mayor cantidad de información y por periodos más largos de tiempo.

A principios de la presente década, motivado por el reconocimiento de que el EG se debe en mucho a la promoción del procesamiento semántico de la información, algunos investigadores empezaron a extender sus estudios y a explorar efectos colaterales del EG en las tareas de aprendizaje (e.g. Mojardín, et. al., 2003). Uno de esos efectos, que más inquietud ha presentado, es el de la promoción de Falsas Memorias (FM), es decir la promoción de memorias que no corresponden con lo experimentado (Stein, 2010).

El interés por el desarrollo de FM ligados al EG, encierra preocupaciones teóricas y prácticas de mucho significado (e.g., McCabe & Smith, 2006; Neufeld, Brust & Stein, 2010). Resulta muy negativo que una ventaja cognitiva, como la que supone el EG, sea neutralizada por que se asocia al incremento en los índices de FM. Al mismo tiempo resulta interesante para la psicología de la memoria identificar los mecanismos que producen resultados tan opuestos.

Teóricamente, el EG es producto de la multiplicación de redes de significado que se promueven durante la tarea (McElroy, 1987). Redes, que a la postre se convierten en la razón para la expresión de altos índices de falsos recuerdos (Brainerd & Reyna, 2005).

Algunos estudios experimentales recientes han evaluado esa relación y han encontrado resultados contradictorios. Mientras unos reportan el EG sin asociación con FM (e.g., Gunter, Bodner & Azad, 2007), otros han reportado el EG neutralizado por altos índices de FM (e.g. Mojardín, 2007).

Dado esto nace la inquietud de que posiblemente el EG podría no ser tan buen mecanismo de aprendizaje, porque la ganancia que implica la cantidad de información retenida en memoria podría ser afectada por el incremento de las imprecisiones que suponen las FM. Es decir, el hecho de generar una información a partir de las conexiones semánticas que motiva podría ser el origen de una información, ciertamente consistente en significados, pero errónea en su expresión (e.g. recordar la categoría en lugar del nombre de un integrante de ésta). La posibilidad de obtener dos resultados de la interacción del EG y FM- (a) el EG reduce los índices de FM, (b) el EG aumenta los índices de FM, es factible y teóricamente un reto que necesita ser atendido.

El objetivo general del presente estudio es explorar el impacto del EG en las MV y las FM. En consecuencia, se

* Dirección para correspondencia [Correspondence address]: Pontificia Universidade Católica do Rio grande do Sul, Pós-Graduação em Psicologia. Av. Ipiranga, 6681- Prédio:11- Sala 940, CEP: 90619-900, Porto Alegre/RS, Brasil. E-mail: lilian@pucrs.br

buscó conocer si el EG aumenta las MV sin promover FM, así como identificar el nivel de certidumbre con que se acompañan las respuestas de la prueba de memoria, para cada condición de aprendizaje. El artículo se organiza de la siguiente forma: primeramente, se abordan estudios que fundamenta la idea que el EG puede reducir las FM, en esta misma línea se abordara el supuesto de que el EG puede producir FM. Después de esto se abordan tres características que el EG y FM tienen en común y han sido poco estudiadas: a) persistencia, se refiere a la estabilidad de un recuerdo sobre el tiempo, b) certidumbre, la cual consiste en los índices de confianza que reportan los participantes en sus recuerdos y, c) sensibilidad a la repetición de prueba o efectos de la Mera-Evaluación (*Mere-Testing*) en el EG.

La relación empírica de EG y FM

De la interacción del EG y FM podemos encontrar diferentes hipótesis. (a) El EG puede reducir los índices de FM. Ésta tiene sustento en los postulados de la Teoría Heurística Distintiva (*Distinctiveness Heuristic*, Dodson & Schacter, 2001). Esta teoría propone que el efecto se debe a un proceso metacognitivo en el que la persona tiene la sensación de conocer con seguridad la información. Por ejemplo, si se le mostrara a un participante una imagen y a su vez esta imagen es verbalizada, esto produce una sensación de experiencia (i.e., una metamemoria que nos permite una especie de familiaridad de haber estado en contacto previamente con esa información). Al momento de realizar una prueba de reconocimiento para evaluar la memoria del participante, la sensación de experiencia provocaría que evocara información rica en detalle permitiendo el correcto reconocimiento verdadero, rechazando toda aquella información que carezca de esa sensación o características de la información presentada, reduciendo así el falso reconocimiento.

Entre los estudios que dan evidencia que el EG reduce los índices de FM, se encuentra el trabajo de McCabe y Smith (2006) quienes encontraron EG para MV y reducción en los índices de FM. Es decir, información que fue generada también fue falsamente reconocida en comparación con información que solo fue leída. Estos autores, explican la reducción en FM debido a una heurística distintiva basada en la recuperación. De acuerdo con esta explicación, el hecho de generar una información modifica los criterios de decisión siendo estos criterios los reductores de errores en la recuperación. En otras palabras, el acto de generar convierte una palabra más distintiva en la fase de codificación (e.g., televisión), puesto que se registran detalles específicos que fortalecen posteriormente el acceso a la información, permitiendo de esta manera en una prueba discriminar entre la información que fue presentada anteriormente (e.g. televisión) y una nueva información (e.g., radio), pudiendo rechazar aquella información nueva ya que esta no presenta las mismas características de la información previamente generada (e.g., si se me hubiera presentado anteriormente “radio” la recordaría).

Por otro lado Soraci, et al. (2003) muestran la ventaja que implica el EG para la MV, sin encontrar diferencia en los índices de FM entre los grupos las condiciones control (leer, escuchar) y experimental (generar). Estos investigadores explican el resultado como efecto del aumento en la distintividad inducida y los consecuentes beneficios en la codificación que da la tarea de generar. Es decir, el hecho de generar una información aumenta su peculiaridad, facilita su codificación y hace más eficiente el almacenamiento. Estas condiciones dan como resultado el incremento en la MV y un estancamiento reducción de las FM.

Si bien algunas evidencias han mostrando el beneficio del EG en MV sin ir acompañado por FM, la literatura registra posiciones teóricas sobre la naturaleza de las FM que dan lugar a hipótesis acerca de la posibilidad de que el EG pueda producir altos índices de FM. Una de esas hipótesis es la Hipótesis de Activación Semántica (McElroy & Slamecka, 1982). En esta se afirma que el EG es producto de las conexiones semánticas que hace la persona, de los contenidos aprendidos con los conocimientos previos. Es decir, las redes semánticas que se activan en el momento de la codificación del aprendizaje, durante la fase de prueba, se convierten en la principal fuente de recuperación de la información (Garnier & Hampton, 1985).

Diversos estudios que respaldan la hipótesis de la activación semántica (e.g., McElroy, 1987), como el realizado por Garnier y Hampton (1985) Diversos estudios que respaldan la hipótesis de la activación semántica (e.g., McElroy, 1987), como el realizado por Garnier y Hampton (1985) realizaron una serie de experimentos en el que manipularon abreviatura de palabras (e.g., ET vs EC), número de dígitos (e.g., 32 vs 3,2) y palabras compuestas (e.g., cheesecake vs cheese ketchup) en el cual se varió el grado de familiaridad o significado. A través de estos experimentos se observó que la memoria semántica es fundamental para que el EG emerja.

Mención especial merece el trabajo, que en esta dirección presentaron McElroy y Slamecka (1982). Estos autores manipularon palabras con y sin sentido semántico (e.g., MTYJX). De esta manera, observándose que el EG no se produce con palabras sin sentido pues no se puede abstraer su significado y la memoria semántica es necesaria para el surgimiento del EG. Estudio similar fue realizado por Mulligan (2002) utilizando palabras sin sentido limitando las conexiones semánticas de procesamiento, mostrando que el EG no se aparece en este tipo de material. De esta manera, dando evidencia que el EG se debe a un procesamiento semántico, pues palabras sin sentido no producen ventaja el beneficio de memoria que es observable en el EG pues no se puede extraer su significado para ser conectado con otras infamaciones.

La Teoría de Trazo de la Huella Borrosa (Fuzzy-Trace Theory; FTT) ofrece una de las explicaciones más sólidas acerca del origen de las FM (e.g. Brainerd & Reyna, 2005) y ayuda a comprender las formas en que estas pueden asociarse con el EG. Para esta teoría la memoria no es un sistema unitario, sino que la información es almacenada en dos tipos

de trazos o códigos de memoria: los literales (*verbatim*), que se refieren a información específica, y los de esencia (*gist*), que se refieren al significado o interpretación de la información. Los primeros son menos resistentes al olvido y a la interferencia que los segundos. De acuerdo con esta teoría, el fortalecimiento de los trazos literales y de esencia fortalecen a la MV. Sin embargo, estos trazos trabajan de forma opuesta con las FM ya que los trazos literales las suprimen y los trazos de esencia las facilitan.

Brainerd, Reyna, & Mojardín (1999) proponen que, en pruebas de reconocimiento, la ruta de acceso a la memoria es la fuente mayor para el reporte de las FM. Las personas reportan FM cuando no pueden acceder a los contenidos literales de la información demandada y se ven obligadas a hacer uso de la información de esencia para la recuperación del contenido. Así y en consecuencia, estos argumentos coinciden con los que ofrece Mulligan (2002). Es de esperarse que las manipulaciones, y condiciones experimentales, que llevan al sujeto a depender de los contenidos semánticos para ofrecer sus respuestas de memoria, sean la mayor razón para obtener altos índices de FM (Brainerd & Reyna, 2005). Si los procedimientos asociados al EG son primordialmente promotores de redes semánticas, como lo afirman McElroy y Slamecka (1982), Garnier y Hampton (1985) y Mulligan (2002), es teóricamente lógico que el EG se asocie a altos los índices de FM.

Persistencia y Certeza de Memoria

La memoria producto del EG y las FM son persistentes en el tiempo. Crutcher y Healy (1989) manipularon el EG en tres diferentes intervalos de tiempo, inmediato, 2 días y 7 días después de la presentación de información. Sus resultados indicaron que aunque el EG tiende a disminuir en pruebas diferidas, la ventaja que da a la memoria se mantiene (e.g., experimento 2).

Resultados similares fueron obtenidos por Brainerd y Mojardín (1998). En una serie de experimentos con adultos y niños, estos autores evaluaron la persistencia de las FM. Teniendo como material de aprendizaje oraciones, los participantes recibieron evaluaciones inmediatas y diferidas, que dieron como resultado mayor resistencia al olvido de las FM que de las MV.

En la misma dirección se dio el estudio de Thapar y McDermont (2001). Ellos corrieron una serie de experimentos con listas de palabras asociadas, las cuales fueron evaluadas en prueba de recuerdo libre y reconocimiento inmediatas, diferidas a dos días y diferidas a una semana. En sus resultados se observó una disminución en las MV, pero el falso reconocimiento y el falso recuerdo fueron más persistentes al tiempo, aun cuando estos mostraron una reducción al igual que las MV.

Otra característica común del EG y de las FM son los altos índices de certeza con son acompañadas. Slamecka y Graf (1978; Experimento 1) reportaron que el EG puede producir altos índices de certidumbre en su información y que estos varían dependiendo del tipo de instrucción para el aprendizaje¹. Por otro lado, Slamecka y Graf (1978) reportan resultados con la misma dirección. El hecho de generar una información produce altos índices de confianza al momento de recordarla, en comparación con la condición control.

Al igual que el EG, las FM presentan altos índices de certidumbre. Es decir, los participante recuerdan con altos índices de seguridad una FM. Roediger y McDermont (1995) realizaron una serie de experimentos con palabras asociadas, con el objetivo de explorar los falsos recuerdos y el falso reconocimiento, así como los índices de certidumbre que reportaban los participantes en las pruebas de recuerdo y reconocimiento. Los resultados indicaron que los niveles de falsos recuerdos y falso reconocimiento mostraron altos niveles de certeza (Experimento 2); con un 53% en recordar específicamente con detalles sobre el evento ocurrido en los falsos reconocimientos.

Por otro lado, Soraci et al. (2003; Experimento 4) exploraron los índices de certeza en la interacción del EG y FM encontrando que fueron altos para la condición generar, en comparación a la condición control (leer). Siempre la memoria para los ítems objetivo fue más alta que para los ítems distractores relacionados.

Efecto de Repetición de Prueba

Al ser el EG pensado como una posible herramienta para el área de la educación existen puntos interesantes para indagar. Uno de ellos es el efecto de mera-evaluación (mere-testing). Este efecto fue estudiado por Brainerd y Mojardín (1998) encontrando que la repetición de una prueba de reconocimiento para una misma información (mera-evaluación) consolida la MV. Sin embargo, también produce un aumento en el reporte de FM. Previamente, este efecto de mera-evaluación había sido abordado por Brainerd y Reyna (1996).

De acuerdo con la FITT, la repetición fortalece los trazos literales de la información objetivo, pero también activa las redes de significado de la información produciendo FM. Si el simple hecho de repetir la evaluación de una información favorece a la memoria, podemos esperar que el EG se beneficie de esta repetición, aumentando tanto MV como FM.

De acuerdo con la FITT (Brainerd, 2005), el EG sería explicado de la siguiente manera, dentro del EG existen dos tipos de instrucción de generación que son: (a) instrucciones semánticas, consiste en generar la información clave a partir de categorías (e.g., un color: blanco), asociación (e.g., limpio-blanco), sinónimos (e.g., claro-blanco) ó antónimos (e.g., ne-

¹ Nos referimos como tipo de instrucción de aprendizaje al tipo de tarea de generación con la que el participante genera la información ya sea, rima, sinónimos, entre otros.

gro-blanco); y (b) instrucciones no-semánticas (preceptuales) consiste en generar la información clave a partir de fragmentación de palabras (e.g., bl_n_o), rimas (e.g., banco-blanco) ó transposición de letras (e.g., nco-bla, blanco). Para la FTT la instrucción de generación semántica fortalecerían los trazos de esencia, mientras que las instrucciones no-semánticas, fortalecerían los trazos literales.

A partir de esta teoría podríamos predecir que, independiente de la instrucción de generación que se utilice, los procedimientos fortalecen las MV que llevan a obtener la ventaja propia del EG. En lo referente a FM bajo EG, desde FTT es de esperar resultados diferentes, dependiendo de la instrucción de generación. Si las instrucciones no-semánticas fortalecen los trazos literales es de esperar un índice bajo en FM, pero si las tareas son de generación semántica entonces se fortalecerían los trazos de esencia y se deben producir mayor índices de FM.

En relación a la persistencia en MV se espera que ésta disminuya más para la información generada por una instrucción no-semántica, en comparación con la que provenga de la instrucción semántica. La razón, los trazos literales tienden a caer más rápidamente que los semánticos con el paso del tiempo. Respecto de la persistencia en las FM, lo que se ha observado es que estas se mantienen o aumentan con el tiempo, por ello es de esperarse que bajo el EG sean más persistentes.

Si una información es evaluada de forma repetida, independiente del trazo que refuerce en el EG, se espera un aumento tanto en MV como en FM, pues la repetición fortalece la MV al igual que las FM.

Método

Diseño

El presente experimento utilizó un diseño factorial de ANOVA, 2 (Instrucción de Aprendizaje: categoría, rima) \times 2 (Tiempo de Prueba: inmediata, diferida a una semana) \times 3 (Tipo de Ítem: objetivo, distractores relacionados, distractores no relacionados) \times 2 (Tipo de Aprendizaje: generar, leer) \times 2 (Mera-Evaluación: ítems evaluados, ítems no evaluados)², siendo los factores Instrucción de Aprendizaje y Orden manipulados entre-participantes. La variable dependiente fue la proporción de aceptación de cada ítem en la prueba de reconocimiento, así como el nivel de certidumbre para cada respuesta.

Participantes

En este estudio participaron 66 estudiantes universitarios (50 mujeres y 16 hombre $M = 20.92$ años $DE = 4.30$). Los participantes fueron asignados aleatoriamente a una de las condiciones de instrucción de aprendizaje (32 instrucción categoría, 34 instrucción rima), el único criterio de exclusión

fue bajo rendimiento en la tarea de generación. Todos ellos aceptaron firmar un permiso de libre consentimiento en el que se declaraba participar de manera voluntaria. Los participantes obtuvieron una hora de actividad complementaria (i.e., actividad académica extracurricular) para su curso de psicología por dicha participación.

Materiales

Se utilizaron materiales distintos para cada una de las cuatro fases del experimento (aprendizaje, tarea interpolada, prueba de reconocimiento inmediata y prueba diferida). Para la fase de aprendizaje se seleccionaron 83 categorías específicas (e.g., moneda, color, planetas) basadas en los bancos de normas de categoría de Battig y Montague (1969), McEvory y Nelson (1982), Janczura (1996), Overschelde, Rawson y Dunlosky (2004). De estas categorías se crearon dos tipos de listas, una para instrucción Categoría y una para instrucción Rima, todo en idioma portugués brasileño. Cada lista fue conformada por pares de palabras, donde una palabra objetivo tendría que ser generada a partir de una palabra pista (e.g. limpio-bla__, limpio-blanco).

El material de aprendizaje fue producto de un estudio preliminar sobre la facilidad con que algunas palabras se podían generar, teniendo como base instrucciones a partir de categorías o rimas. En éste estudio participaron un total de 40 estudiantes universitarios, asignados de manera aleatoria y equilibrada a cada uno de los grupos de instrucción de aprendizaje (categoría – rima). Las categorías se seleccionaron de aquellas para las que, al menos el 70% de los participantes, ofrecieron la misma respuesta. Al final, un total de 36 listas de categorías fueron seleccionadas para ambas instrucciones de aprendizaje (Ver Anexo, A).

En la tarea interpolada se utilizó una hoja con seis problemas matemáticos (e.g., $2368 \times 15 = ___$) que buscaba eliminar los efectos de memoria de corto plazo y el ejercicio de alguna estrategia nemónica que beneficiara sus respuestas.

La prueba de reconocimiento inmediato estuvo compuesta por un total de 22 ítems: 12 ítems Objetivo, palabras que fueron generadas o leídas (e.g., categoría pájaro: canario) durante la fase de aprendizaje; 5 Distractores Relacionados, palabras asociadas, pertenecientes a la misma categoría de la que fueron generadas o leídas (e.g., águila) y; 5 Distractores No-relacionados, palabras que no tenían ninguna relación de significado o rima con las palabras objetivo (e.g., tipo de madera: pino). La prueba diferida consistió en un total de 44 ítems, 24 objetivos, 10 distractores relacionados, y 10 distractores no relacionados; la mitad de ellos habían sido evaluados en la prueba inmediata y la otra mitad que eran evaluados por primera ocasión (Ver, Anexo B).

Procedimiento

Los participantes fueron evaluados en computadoras individuales, en grupos máximos de 15 personas. Cada participante se encontró separado por mamparas para evitar que

² El factor Mera-Evaluación es derivado del factor prueba diferida.

pudieran ver las pantallas de las computadoras de los otros. Para la presentación de estímulos en la fase de estudio y de prueba, se utilizó el programa Super Lab Pro 2.0. Inicialmente, cada participante fue invitado a sentarse frente a la pantalla de la computadora en la que se presentarían los estímulos utilizados en el estudio y a colocarse los auriculares.

Durante la fase de estudio, como se indica en la Figura 1, el investigador instruyó a cada participante acerca de que aparecerían en la pantalla de la computadora una serie de palabras pares que debían atender según la instrucción que recibieran. Algunos de esos pares eran compuestos por dos palabras completas (e.g., outono-verão), mientras que otros

eran compuestos por una palabra completa y otra incompleta (e.g., revista – li___) que debían completar. Así, la tarea de los participantes en esta fase fue: (a) cuando apareciera en la pantalla de la computadora un par compuesto por dos palabras completas ellos tendrían que escribir la segunda palabra en la hoja de respuestas, proporcionada por el investigador y (b) cuando apareciera una palabra completa y otra incompleta, tenían que generar la segunda y escribirla en la hoja de respuestas asignada. Durante el tiempo asignado para escribir la palabra, escuchaban una señal que indicaba cuando se presentaría el siguiente par de palabras.

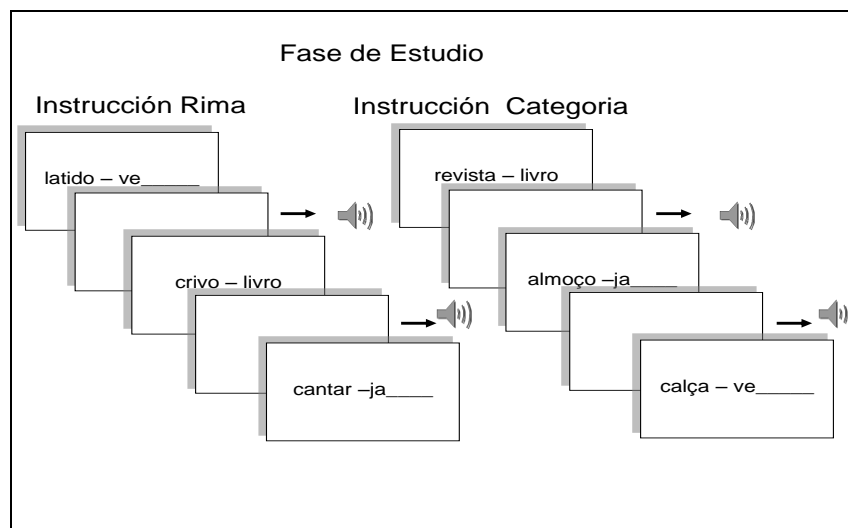


Figura 1. Ejemplo de la presentación del material de la fase de estudio.

La tarea de generación de las palabras fue dividida en instrucción Categoría o instrucción Rima. Al grupo con instrucción de aprendizaje Rima se les informó que se le presentaría una palabra completa y una incompleta (e.g., canário-po___) y que su trabajo era generar la segunda palabra, siguiendo la regla de Rima. Es decir, la segunda palabra que ellos deberían generar tenía que iniciar con las dos letras pistas que se le mostraba y a su vez tenía que rimar con la palabra completa que le acompañaba (e.g., rombo-pombo). Para al grupo Instrucción de aprendizaje Categoría, el participante tenía que generar una palabra que fuera parte de la misma categoría de la palabra que se le proporcionaba, e iniciara con las dos letras que se le mostraban (e.g., pájaro: canário-pombo).

Además, cada participante fue instruido de que cada par de palabras serían presentadas por 3 segundos, después de ese intervalo, aparecería una imagen en blanco con una duración de 5 segundos. En ese tiempo, ellos deberían escribir o generar la segunda palabra atendiendo las instrucciones dadas. Se les informó que al finalizar el tiempo de la tarea de generación escucharían un sonido indicando que un nuevo par de palabras estaba por aparecer en la pantalla y que debían mirarla para proseguir. Este procedimiento se repitió

hasta que todos los pares de palabras fueron mostrados. El orden de presentación en ellas fue completamente aleatorio.

Al término de la fase de estudio se recogían las hojas de respuestas y se les entregaba la hoja con problemas matemáticos básicos ($45638 \times 485 = \underline{\quad}$). Se les pedía que los resolvieran todos los y que para ello tenían 5 minutos.

Al terminar ese tiempo los participantes recibieron prueba de reconocimiento. En esta fase se les indicó que en la pantalla se presentarían una serie de palabras, una a una, y que su trabajo sería identificar si estas fueron parte de la lista de estudio, así como hacer un juicio de confianza de que tan seguros estaban de haberla estudiado (ver, Figura 2). La certeza para cada ítem fue evaluada mediante una escala de confianza de 4 puntos, que iban desde (1) “tengo seguridad de si haber estudiado esa palabra” hasta (4) “tengo la seguridad de no haber estudiado esa palabra”. Los puntos 1 y 2 correspondieron a la respuesta de “si estudié esa palabra”, mientras que los puntos 3 y 4 correspondieron a la respuesta de “no estudié esa palabra”. Sobre el teclado de la computadora fueron marcadas las cuatro opciones: “estoy seguro que si generé o escribí esa palabra” (tecla \), “no estoy muy seguro, pero creo que si generé o escribí esa palabra” (tecla v), “no estoy muy seguro, pero creo que no generé o escribí esa palabra” (tecla m) y “estoy seguro que no generé o escribí

esa palabra" (tecla/). Enseguida se ofrece un esquema que resume el procedimiento.

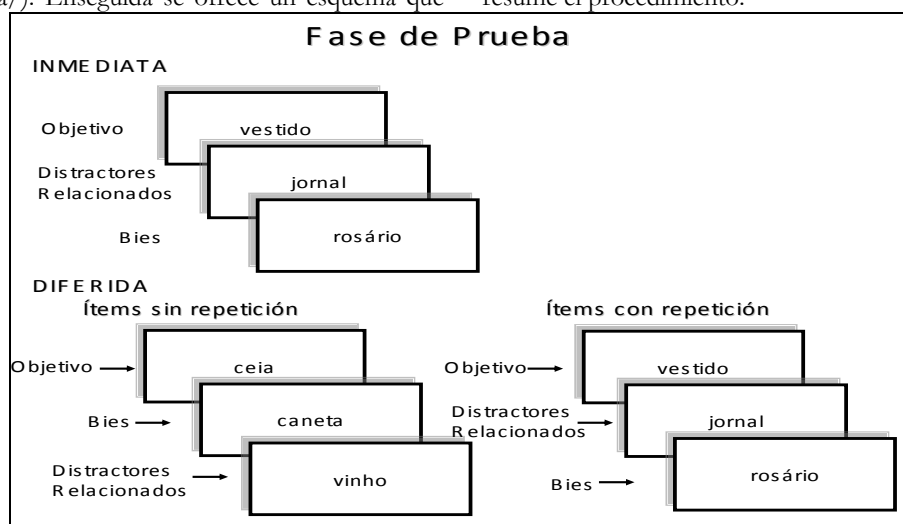


Figura 2. Presentación de la fase de prueba.

Al final del estudio los participantes fueron invitados a regresar en una semana para recibir la prueba diferida, la cual siguió exactamente los mismos procedimientos de la primera.

Resultados

El interés del presente trabajo se concentró en identificar los niveles de asociación entre el EG y las FM. Sin embargo, resulta muy informativo conocer los niveles de confianza que los participantes presentaron para sus respuestas. En primer lugar se analiza cómo se comportaron los niveles de MV y de FM para cada combinación de factores y después se revisa lo referente a la confianza de las respuestas.

Se corrió un análisis de varianza (ANOVA) teniendo como factores Instrucción de Aprendizaje (categoría, rima) \times Tiempo de Prueba (inmediata, diferida a una semana) \times Tipo de Ítem (objetivo, distractores relacionados, distractores no relacionados) \times Tipo de Aprendizaje (generar, leer) \times Mera-Evaluación (ítems evaluados, ítems no evaluados). En general, el porcentaje para las palabras generadas fue de un 90% en la condición Generar, con un 88% y 92% para las instrucciones de aprendizaje Categoría y Rima, respectivamente. La condición Leer obtuvo un 100% en la ejecución de la tarea.

Para investigar la influencia del EG en las MV y en las FM, se corrió un ANOVA con medidas repetidas de 2 (Tipo de Aprendizaje: generar, leer) \times 2 (Tiempo de Prueba: inmediata, diferida a una semana) \times 2 (Tipo de Ítem: objetivo, distractor relacionado) \times 2 (Instrucción de Aprendizaje: categoría, rima), teniendo como variable dependiente la proporción de aceptación. El ANOVA produjo una interacción entre Tipo de Ítem \times Tipo de Aprendizaje, $F_{(1,64)} = 30.35$, $MSE = 0.043$, $p < .000$. Una inspección de la Tabla 1 muestra que los resultados favorecen la hipótesis (a) el EG no va

acompañado por un incremento en FM., ya que encontramos que los índices de memoria para los ítems Objetivo (MV) fueron mayores en la condición Generar ($M=0.660$ vs. $M=0.441$), los índices para los ítems Relacionados (FM) no lo fueron ($M=0.062$ vs. $M=0.081$). Esto sugiere que el EG facilita la recuperación de trazos literales. Ese padrón de resultados no parece ser influenciado por las demás condiciones experimentales ya que no hubo otras interacciones significativas. Destaca el resultado no significativo entre Tipo de Ítem \times Tipo de Aprendizaje \times Instrucción de Aprendizaje, $p = .708$, sugiriendo que la magnitud del EG no depende del tipo de instrucción utilizada (Categoría, Rima).

Tabla 1. Proporciones de aceptación, según el tipo de ítem y tipo de aprendizaje.

Tipo de Ítem	Leer		Generar	
	Media	DT	Media	DT
Objetivo	0.441	0.214	0.660	0.137
Distractor relacionado	0.062	0.120	0.081	0.131
Distractor no-relacionado	0.108	0.100	0.108	0.100

Es necesario aclarar que los ítems No-relacionados fueron excluidos del análisis, dado que sus índices de aceptación fueron lo suficientemente bajos como confirmar que las respuestas de los participantes estaban siguiendo puntualmente las instrucciones.

Para evaluar si el tiempo de prueba es un factor que influencia el EG y su asociación con las FM, se corrió un ANOVA de 2 (Tiempo de Prueba: inmediata vs. Diferida) \times 2 (Tipo de Aprendizaje: Generar vs. Leer), resultando significativo, $F_{(1,64)} = 5.34$, $MSE = 0.033$, $p = .024$. La prueba Bonferroni demostró que las palabras generadas ($M = .448$), como las palabras leídas ($M = 0.292$) fueron reconocidas en menor proporción durante la prueba diferida ($M_{generadas} = 0.294$ y $M_{leídas} = 0.211$; $ps \leq .007$). La disminución en los índices de reconocimiento que fue mayor en las palabras ge-

neradas ($\Delta M = 0.154$) que en las palabras leídas ($\Delta M = 0.081$) indica que la información generada fue menos persistente al tiempo. Un patrón que aparece en dirección opuesta al esperado.

Para investigar el efecto de la prueba repetida, o de Mera-Evaluación, se corrió un ANOVA en el que solo se sustituyó el factor Tiempo de Prueba por el factor de Mera-Evaluación. El ANOVA con medias repetidas fue 2 (Tipo de Aprendizaje: generar, leer) \times 2 (Mera-Evaluación: ítems evaluados-prueba inmediata vs. ítems no evaluados) \times 2 (Tipo de Ítem: objetivo vs. distractor relacionado) \times 2 (Instrucción de Aprendizaje: categoría vs. rima). Este ANOVA produjo dos resultados de interés: (a) El efecto principal de Mera-Evaluación ($F_{(1, 64)} = 103.19$, $MSE = 0.054$, $p < .000$), sugiere que la repetición de la prueba mejora el recuerdo de la información evaluada. Como se puede apreciar en la figura 3, los beneficios de la repetición de prueba fueron mayores bajo la condición Generar que en la condición Leer. (b) La interacción entre Mera-Evaluación \times Tipo de Aprendizaje, $F_{(1, 64)} = 9.44$, $MSE = 0.043$, $p = .003$, indica que el efecto de Mera-Evaluación favorece más a un tipo de aprendizaje que al otro. El análisis Bonferroni demostró que la repetición de prueba beneficia más a la condición Generar ($\Delta M = 0.261$) que a la condición Leer ($\Delta M = 0.150$) (ver Figura 3).

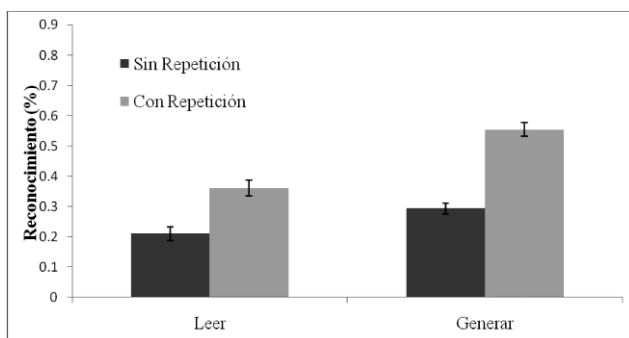


Figura 3. Proporciones de aceptación para el efecto de mera-evaluación en tipos de aprendizaje.

Con el objetivo de obtener una medida de memoria independiente del sesgo de respuesta³ de los participantes, para conseguir una medida de memoria más pura se aplicó el parámetro de d' , propuesto por la Teoría de Detección de Señal (Macmillan & Creelman, 2005). Por cuestiones metodológicas, el parámetro calculado en este trabajo fue llamado $d'_{específico}$ ⁴, ya que se refiere a la capacidad de discriminar únicamente palabras objetivos de palabras distractores relacionados.

Se realizó un ANOVA con medidas repetidas de 2 (Tipo de Aprendizaje: generar, leer) \times 2 (Tiempo de Prueba: inmediata, diferida a una semana) \times 2 (Instrucción de Aprendizaje: categoría, rima), teniendo como variable dependiente

los valores de d' , medida en unidades de desviación estándar. Este análisis produjo efectos principales para el Tipo de Aprendizaje ($F_{(1,64)} = 32.31$, $MSE = 0.565$, $p < .000$) y el Tiempo de Prueba ($F_{(1,64)} = 143.023$, $MSE = 0.441$, $p < .000$).

En la Figura 4 se aprecian los valores de $d'_{específico}$ según el Tipo de Aprendizaje y Tiempo de Prueba. Como se puede ver, la interacción entre Tipo de Aprendizaje y Tiempo de Prueba no fue significativa ($p = .283$), lo que sugiere que, al contrario de lo observado en el análisis con proporción de aceptación, la disminución en el recuerdo de palabras generadas ($\Delta d'_{específico} = 1.068$) no difiere significativamente del recuerdo de palabras leídas ($\Delta d'_{específico} = 0.887$), aun cuando las diferencias entre medias muestren la misma tendencia. La capacidad de discriminar palabras Objetivo de Distractores relacionados disminuyó con el paso del tiempo, independiente del tipo de aprendizaje (generar o leer) utilizado por el participante durante la codificación de esos ítems.

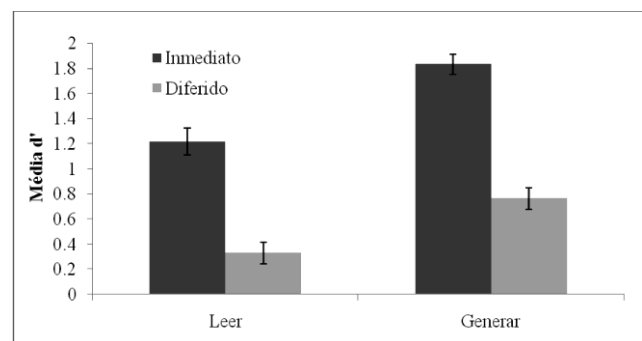


Figura 4. Índices de valores d' de la interacción tipo de aprendizaje y tiempo de prueba.

Para aclarar mejor los efectos de Mera-Evaluación en el EG se realizó un ANOVA con los valores de $d'_{específico}$, tan solo sustituyendo el factor Tiempo de Prueba por el factor de Mera-Evaluación (ítems evaluados en la prueba inmediata vs. ítems no evaluados). Este ANOVA produjo apenas efectos principales de Tipo de Aprendizaje ($F_{(1,64)} = 18.08$, $MSE = 0.570$, $p < .000$) y de Mera-Evaluación ($F_{(1,64)} = 29.31$, $MSE = 0.561$, $p < .000$). Anteriormente, en proporción de aceptación se reportó que el hecho de repetir una prueba con una misma información beneficia más a palabras que fueron generadas que a palabras leídas. No obstante, los resultados con $d'_{específico}$ sugieren que esto no sucede cuando se discrimina entre palabras Objetivo y Distractores Relacionados (ver Figura 5). De hecho, el incremento de Mera-Evaluación no indica que la repetición de una misma información beneficie más a la condición de generación ($\Delta d'_{específico} = 0.459$) que a la condición de leer ($\Delta d'_{específico} = 0.539$), cuando de discriminar las MV de las FM se trata. Tal y como se ve en la Figura 5, la repetición de prueba mejora la discriminación, por tanto la precisión.

³ Es la tendencia del participante a dar un tipo de respuesta con cierto grado de favoritismo y no otra, pero esta respuesta no supone una base de memoria.

⁴ $d'_{específico} = z(\text{objetivo}) - z(\text{distractores relacionados})$

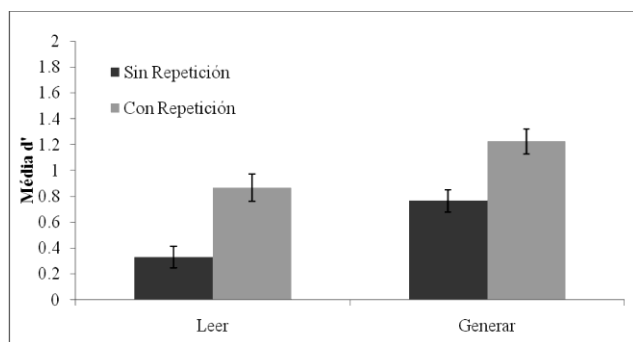


Figura 5. Índices de valores d' del tipo de aprendizaje en mera-evaluación.

Análisis de los juicios de grado de certeza

El análisis de grado de certeza tiene el objetivo de indagar como el EG se experimenta por los participantes. Además, conocer cómo se comportan los índices de certidumbre a lo largo del tiempo y con diferentes tipos de instrucción de aprendizaje. Para atender este objetivo se corrió un ANOVA con medidas repetidas de 2 (Grado de Certeza: alta certeza vs. baja certeza) \times 2 (Tipo de Aprendizaje: generar vs. leer) \times 2 (Tiempo de Prueba: inmediata vs. diferida) \times 2 (Instrucción de Aprendizaje: categoría vs. rima), teniendo como variable dependiente la proporción de aceptación de los ítems evaluados.

Este ANOVA produjo dos interacciones de interés: (a) Tipo de Ítem \times Tiempo de Prueba \times Tipo de Aprendizaje \times Grado de Certeza ($F_{(1,64)} = 12.55$, $MSE = 0.026$, $p < .001$) y (b) Tipo de Ítem \times Tipo de Instrucción de Aprendizaje \times Grado de Certeza ($F_{(1,64)} = 28.10$, $MSE = 0.030$, $p < .000$). En la tabla 2 se presentan las medias de grado de certidumbre por tipo de aprendizaje para cada prueba de las palabras objetivo. Como se puede observar, el EG se asocia a grados de certeza que cambian con el paso del tiempo.

En la prueba de reconocimiento inmediato 77% de las palabras generadas se acompañan por un juicio de alta certeza, mientras que apenas el 7% de éstas palabras se acompañan por un juicio de baja certeza. Estas proporciones fueron significativamente diferentes de las obtenidas para la condición de Leer ($p < .000$). Después de una semana ya no se observó el mismo patrón de certeza ($p = .454$).

Tabla 2. Proporciones de aceptación para palabras objetivos según el tiempo de prueba, tipo de aprendizaje y grado de certeza.

Tipo de Aprendizaje	Baja Certeza		Alta Certeza	
	Media	DT	Media	DT
Prueba Inmediata				
Leer	0.133	0.152	0.436	0.302
Generar	0.067	0.115	0.775	0.176
Prueba Diferida				
Leer	0.201	0.211	0.112	0.173
Generar	0.223	0.201	0.255	0.208

Por último se investigó la interacción entre Tipo de Ítem \times Tipo de Instrucción de Aprendizaje \times Grado de Certeza. Como se dijo arriba, esta interacción indica que el recuerdo de la información bajo instrucción de Categoría parece ser ir acompañada de mayor grado de seguridad que bajo la instrucción Rima (ver Figura 6). El análisis Bonferroni mostró que en instrucción Categoría, 45% del reconocimiento de las palabras Objetivo son acompañadas por un juicio de alta certeza, mientras que en la instrucción Rima solo 34% del reconocimiento de palabras Objetivo es acompañado por un juicio de alta certeza.

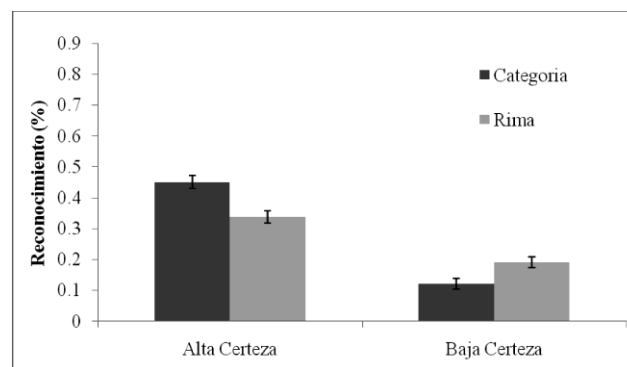


Figura 6. Proporciones de aceptación medias para tipo de aprendizaje y grado de certeza.

Discusión

Los datos del presente estudio parecen corroborar la hipótesis de que el EG no va acompañado por un aumento en las FM. Este resultado niega el respaldo para la propuesta teórica que propone al EG como resultado de la promoción de significados (e.g., Mulligan, 2002). De acuerdo con estos resultados, se pudiera afirmar que el aprendizaje activo, derivado del EG, de una información facilita mayormente la codificación y posterior recuperación de sus trazos literales. Como consecuencia se obtienen altos índices de MV en la información generada, comparada con información leída, con no mayores índices de FM.

En relación al EG y su comportamiento a través del tiempo, lo que se observó fue que la memoria para las palabras generadas no persistió en los índices esperados para la prueba diferida. Las palabras generadas decayeron más que las palabras leídas con el paso del tiempo, dirección contraria a la propuesta por Crutcher & Healy (1989). Ninguna evidencia se encontró acerca de que el aprendizaje activo se asocie a un mecanismo que le mantenga durante la etapa de almacenamiento. Estos resultados tienen sentido con la explicación de que el EG promueve la codificación de trazos literales y facilita su recuperación que proponen la FTT y la Teoría de la Distintividad Heurística.

En relación a los resultados de efecto de Mera-Evaluación estos van al encuentro de resultados obtenidos en otros estudios (Karpicke & Roediger, 2007., Roediger & Karpicke, 2006). Lo que se sugieren en conjunto es que la

repetición de una prueba facilita el recuerdo de la información evaluada en posteriores pruebas y que cuando el aprendizaje es derivado del EG parece verse favorecido.

Cuando se analiza Mera-Evaluación con d' específico (la discriminación específica entre palabras objetivos y palabras relacionadas) los indicios de que la repetición de prueba mejoran para la información generada no se confirman. Esto difiere de lo encontrado en otros estudios que han reportado que la repetición de una información, si bien favorece a la MV también incrementa las FM (e.g., Brainerd & Reyna, 1996).

Por otro lado, el análisis de grado de certeza nos permite revisar el aspecto cualitativo del recuerdo. Por ejemplo, se espera que un recuerdo bastante vívido sea acompañado por un juicio de certeza alto, mientras que un recuerdo menos vívido sea acompañado por un juicio de certeza bajo. En el análisis se observó que el EG es acompañado por un juicio de alta confianza, así como lo reportado por Slamenka y Graf (1978). Esto nos indica que el EG además de mostrar resultados comportamentales diferentes, el recuerdo de esta información también es experimentado de forma diferente, con mayor intensidad. No obstante, después de una semana se observa que los participantes no mantienen los juicios de certeza de la prueba inmediata. Este último resultado sugiere que a pesar de que el efecto comportamental del EG es persistente después de una semana de la fase de estudio, su aspecto cualitativo se pierde con el paso del tiempo. Si se trazara una relación entre intensidad y grado de exactitud del recuerdo, es posible afirmar que este resultado respalda la conclusión de que el EG se debe a un fortalecimiento en los trazos literales de la información generada, ya que la certeza en el reconocimiento inmediato y la dificultad de recuperación con el tiempo se asocian al comportamiento de la memoria para los trazos literales (Brainerd & Reyna, 2005).

En lo referente al análisis del grado de certeza con respecto al uso de diferentes tipos de instrucción de aprendizaje (e.g., generar una palabra a partir de un sinónimo, antónimo, miembro de una categoría, por asociación ortográfica o semántica, completando fragmentos, solución de anagramas, entre otras), los resultados cuantitativos indicaron que la magnitud del EG es equivalente entre las instrucciones. La manera en cómo se experimenta el EG a través de las instrucciones es diferente, ya que el recuerdo de información aprendida bajo instrucción de Categoría parece ser más vívido que el aprendido bajo instrucciones de Rima. Con esto se

resalta la importancia de la investigación de la influencia de los diferentes tipos de instrucción en los procesos de memoria. Además, las variaciones en el tamaño del efecto asociado a diferentes tipos de instrucción de aprendizaje, como los reportados en estudios de metanálisis (e.g., Bertsch, Pesta, Wiscott y McDaniel, 2007) vienen al encuentro de esta afirmación.

En resumen, nuestros resultados indican que el EG no presenta un incremento en FM y esto es consistente con otros estudios (e.g. Soraci, et al. 2003). Por lo cual se pudiera afirmar que el EG es un beneficio real para el aprendizaje (i.e., EG promueve altos índices de precisión, sin presentar mayores errores de memoria). Además, se confirmó que la repetición de prueba favorece el aprendizaje de la información, sin distinguir del mecanismo de aprendizaje que siga.

El total de resultados obtenidos dejan una deuda pendiente. Los patrones de comportamiento de la memoria en la combinación de los factores evaluados con el EG, no son del todo claros. La mayoría de los resultados específicos enriquecen la discusión del EG y sus asociaciones con las FM, pero solo ofrecen señales débiles acerca de cuáles son las razones específicas de su origen. Mientras que los tipos de instrucción avalaron la hipótesis de que la promoción de significados acompaña al EG, la aplicación de prueba diferida indicó que su beneficio no resiste con el tiempo. En términos de la FITT, estos resultados dan indicios de que la base memorística del EG pueden ser tanto la promoción de codificación de trazos literales, como de trazos semánticos. No obstante la proporción de ellos sigue sin clarificarse. Seguir investigando se hace una prioridad para aumentar la evidencia del grado en que el EG es fundado en la memorización de los trazos literales y de los de esencia.

Como sugerencia a futuros estudios sobre EG y FM con la finalidad de identificar mejor los efectos que varían a través de algunos, sugerimos tener mayor control de las manipulaciones, como el tiempo de exposición de la información a ser generada. Así como verificar la capacidad de discriminación de distractores relacionados y distractores no-relacionados; analizar cómo estos se comportan con el tiempo en los diferentes tipos de aprendizaje. Procedimientos más sofisticados, como los que se derivan de la modelación matemática (e.g. Brainerd, Reyna & Mojardín, 1999; Brainerd, Payne, Wright & Reyna, 2003) se hacen necesarios.

Referencias

- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: a replication and extension of the connecticut category norms. *Journal of Experimental Psychology Monograph*, 80(3), 1-46.
- Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007) The generation effect: A meta-analytic review. *Memory & Cognition*, 35(2), 201-210.
- Brainerd, C. J. (2005). Fuzzy trace theory: memory. En C. Izawa, N. Ohta (Eds.). *Human Learning and Memory: Advances in Theory and Application The 4th Tsukuba International Conference on Memory*. (pp. 219-238). Psychology Press.
- Brainerd, C. J., & Mojardín, A. H. (1998). Children's and adult's spontaneous false memories: long-term persistence and mere-testing effects. *Child Development*, 69(5), 1361-1377.
- Brainerd, C. J., & Reyna, V. F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology*, 32(3), 467-478.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford: Oxford University Press.
- Brainerd, C. J., Payne, D. G., Wright, R., & Reyna, V. F. (2003). Conjoint recall. *Journal of Memory and Language*, 48 (3), 445-467.

- Brainerd, C. J., Reyna, V. F., & Mojardín, A. H. (1999). Conjoint Recognition. *Psychological Review*, 106 (1), 160-179.
- Crutcher, R. J., & Healey, A. F. (1989). Cognitive operations and the generation effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(4), 669-675.
- DeWinstaley, P., & Bjork, E. L. (2004). Processing strategies and the generation effect: implications for making a better reader. *Memory & Cognition*, 32(6), 945-955.
- Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, 8(1), 155-161.
- Garnier, J. M., & Hampton, J. A. (1985). Semantic memory and the generation effect: some tests of the lexical activation hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 732-741.
- Gunter, R. W., Bonerd, G. E., & Azad, T. (2007). Generation and encoding induce a mirror effect in the DRM paradigm. *Memory and Cognition*, 35(5), 1083-1091.
- Janczura, G. A. (1996). Normas associativas para 69 categorías semânticas. *Psicologia: Teoria e Pesquisa*, 12(3), 237-244.
- Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57, 151-162.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A user's guide*. (2ed). New Jersey: Lawrence Erlbaum.
- McCabe, D. P., & Smith, A. D. (2006). The distinctiveness heuristic in false recognition and false recall. *Memory*, 14, 570-583.
- McElroy, L. A. (1987). The generation effect with homographs: evidence for postgeneration processing. *Memory & Cognition*, 15(2), 148-153.
- McElroy, L. A., & Slamecka, N. J. (1982). Memorial consequences of generating nonwords: implications of semantic-memory interpretations of the generation effect. *Journal of Verbal Learning and Verbal Behavior*, 21, 249-259.
- McEvory, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology*, 95(4), 581-634.
- Mojardín, A. H., Velázquez, J. H., Cajiga, S., Apodaca, A., Romero, A. & Alvarado, E. (2003, Noviembre). The generation effect and false memories in sworn testimony. *Paper presented at annual meeting of the SJDM Vancouver, Canada.*
- Mulligan, N. W. (2002). The emergent generation effect and hipermnésia: influences of semantic and no-semantic generation tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 541-554.
- Neufeld, C. B., Brust, P.G., & Stein, L. M. (2010). Compreendendo o fenômeno das falsas memórias. Em L. M. Stein (Eds.), *Falsas Memórias: Fundamentos científicos e aplicações clínicas e jurídicas* (pp. 21-41). Porto Alegre, Brasil: Artmed.
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249-255.
- Roediger, H. L., & McDermont, K. B. (1995). Creating false memories: remembering Words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803-814.
- Slamecka, N. J., & Graf, P. (1978). The generation effect: delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 592-604.
- Soraci, S. A., Carlin, M. T., Togli, M. P., Chechile, R. A., & Neuschatz, J. S. (2003). Generative processing and false memories: when there is no cost. *Journal of Experimental Psychology: Learning, Memory, and cognition*, 29(4), 511-523.
- Stein, L. M. e colaboradores (2010). *Falsas Memórias: Fundamentos científicos e aplicações clínicas e jurídicas*. Porto Alegre: Artmed.
- Taconnat, L., Froger, C., Sacher, M., & Isingrini, M. (2008). Generation and associative encoding in young and old adults. *Experimental Psychology*, 55(1), 23-30.
- Thapar, A., & McDermont, K. R. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, 29(3), 424-432.
- Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004) Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, 50, 289-335.

(Artículo recibido: 17-06-2010, revisión: 19-10-2011, aceptado: 23-10-2011)

ANEXO A. Lista de categorías presentadas en la fase de estudio

Lista de Categorías	Instrucción de Aprendizaje		Palabra Objetivo
	Categoría	Rima	
Artigo de vestuário (artículo de vestuario)	calça (pantalón)	latido (latido)	vestido (vestido)
Bebida alcoólica (bebida alcohólica)	vodka (vodca)	moinho (molino)	vinho (vino)
Calçado (calzado)	sandália (sandalia)	contrato (contrato)	sapato (zapato)
Ciência (ciencia)	física (física)	cronologia (cronología)	biologia (biología)
Cor (color)	preto (negro)	franco (franco)	branco (blanco)
Emoção (emoción)	tristeza (tristeza)	sangria (sangría)	alegria (alegría)
Estação do ano (estación del año)	outono (otoño)	serão (nocturno)	verão (verano)
Fonte de energia (fuente de energía)	hidroelétrica (hidroeléctrica)	motricidade (motricidad)	eletricidade (electricidad)
Forma geométrica (figura geométrica)	círculo (circulo)	mestrado (maestría)	quadrado (cuadrado)
Fruta (fruta)	Maçã (manzana)	granja (granja)	laranja (naranja)
Instrumento cirúrgico (instrumento quirúrgico)	bisturi (bisturí)	caloura (novato)	tesoura (tijeras)
Legume (legumbre)	cenoura (zanahoria)	barata (cucaracha)	batata (papa)
Material de construção (material de construcción)	tijolo (ladrillo)	atendimento (atención)	cimento (cimiento)
Meio de comunicação (medio de comunicación)	telefone (teléfono)	revisão (revisión)	televisão (televisión)
Metal (metal)	aço (acero)	berro (berrido)	ferro (fierro)
Mobília (mobiliario)	cadeira (silla)	chama (llama)	cama (cama)
Modalidade de ginástica (modalidad gimnástica)	olímpica (olímpica)	mímica (mímica)	rítmica (rítmica)
Parentesco (parentesco)	irmão (hermano)	rimo (rimar)	primo (primo)
Parte de um edifício (parte de um edificio)	elevador (elevador)	sentinela (centinela)	janella (ventana)
Pássaro (pájaro)	canário (canario)	rombo (rombo)	pombo (palomo)
Pedra preciosa (piedra preciosa)	esmeralda (esmeralda)	zumbi (zombi)	rubi (rubi)
Tipo de voz (tipo de voz)	grave (grave)	cascuda (cascara)	aguda (aguda)
Unidade de tempo (unidad de tiempo)	minuto (minuto)	piora (peor)	hora (hora)
Um tipo de material de leitura (un material de lectura)	revista (revista)	crivo (cribó)	livro (libro)
Uma construção para serviços religiosos (una construcción para servicios religiosos)	mesquita (mezquita)	deseja (desea)	igreja (iglesia)
Um tipo de moradia humana (un tipo de vivienda humana)	casa (casa)	comportamento (comportamiento)	apartamento (departamento)
Um país (un país)	canadá (Canadá)	cobrança (cobranza)	frança (francia)
Um fenómeno climatológico (um fenómeno climatológico)	furacão (huracán)	alternado (alterno)	tomado (tornado)
Um tipo de moeda (un tipo de moneda)	dólar (dólar)	leal (leal)	real (real)
Um tipo de música (un tipo de música)	rock (rock)	jurássica (jurásica)	clássica (clásica)
Um inseto (un insecto)	mosca (mosca)	urtiga (ortiga)	formiga (hormiga)
Uma cidade (una ciudad)	manaus (manaus)	xerife (sheriff)	recife (recife)
Um estado (un estado)	goiás (goias)	lacre (lacre)	acre (acre)
Ponto cardeal (punto cardinal)	oeste (oeste)	morte (muerte)	norte (norte)
Comida diária (comida diaria)	almoço (comida)	cantar (cantar)	jantar (cenar)
Planeta de nosso sistema solar (planeta de nuestro sistema solar)	marTE (marte)	noturno (nocturno)	saturno (saturno)

*Nota: la traducción en español de la información se encuentra dentro del paréntesis.

ANEXO B. Lista de caterorías presentadas en la fase de prueba

Prueba Inmediata	Ítems	Prueba Diferida	Ítems
Categoría	Objetivos	Categoría	Objetivos
Bebida alcohólica	vinho	Artigo de Vestuário	vestido
Calçado	sapato	Cor	branco
Ciência	biologia	Fonte de energia	eletricidade
Estação do ano	verão	Fruta	laranja
Material de construção	cimento	Legume	batata
Mobília	cama	Meio de comunicação	televisão
Tipo de voz	aguda	Parte de um edifício	janela
Unidade de tempo	hora	Pedra preciosa	rubi
Um tipo de música	clássica	Um tipo de material de leitura	livro
Um inseto	formiga	Um fenômeno climatológico	tornado
Uma cidade	recife	Um tipo de moeda	real
Um estado	acre	Comida diária	jantar
	Distractores relacionados		Distractores relacionados
Forma geométrica	triângulo (triangulo)	Metal	ouro (oro)
Instrumento cirúrgico	pinça (pinza)	Modalidade de ginástica	aeróbica (aeróbica)
Direção cardinal	sul (sur)	Parentesco	filho (hijo)
Pássaro	águia (águila)	Um tipo de moradia humana	barraca (tienda)
Planeta de nosso sistema solar	júpiter (júpiter)	Um país	rússia (rusia)
	Distractores no relacionados		Distractores no relacionados
Tipo de chapéu (tipo de sombrero)	cartola (copa)	Membro da realeza (miembro de la realeza)	conde (conde)
Tipo de madeira (tipo de madera)	pinho (pino)	Erva (hierva)	daninha (daninha)
Um cargo por eleição (un cargo porsenador (senador) elección)		Tipo de noz (tipo de nuez)	castanha (castaña)
Elemento químico (elemento químico)	sódio (sodio)	Roedor (roedor)	castor (castor)
Tipo de pão (tipo de pan)	sírio (sirio)	Um membro do clero (un miembro del clero)	rabino (rabino)

*Nota: con el objetivo de evaluar el efecto de mera-evaluación se incluyeron en la prueba diferida, ítems que había sido evaluados en la prueba inmediata, así como ítems nuevos.

**Nota: la traducción en español de la información se encuentra dentro del paréntesis.