



Artículo de revisión

Revisión de dos instrumentos de simulación neurocognitiva y su utilidad en el contexto forense

Review of the neurocognitive malingering instrument and its useful in a forensic context

Álvaro Aliaga^{1,2*}, Mila Arch³, Alberto García-Molina⁴, Adolfo Jarne³

1 Área de Salud Mental, Servicio Médico Legal. Santiago, Chile.

2 Facultad de Psicología, Universidad del Desarrollo. Santiago, Chile.

3 Departamento de Personalidad, Evaluación y Tratamiento Psicológico. Universidad de Barcelona. Barcelona, España.

4 Institut Universitari de Neurorehabilitació Guttmann, Universitat Autònoma de Barcelona, Badalona, Barcelona, España.

Resumen

Gran parte de la precisión de la evaluación neuropsicológica, depende que los instrumentos que utiliza, sean medidas estandarizadas, válidas y confiables. Sin embargo, la simulación, la exageración y el bajo esfuerzo son aspectos que pueden interferir en los resultados de los test. El objetivo de esta revisión pretende describir y analizar dos instrumentos en el contexto de la Neuropsicología Forense que evalúan simulación de síntomas cognitivos: el *Test of Memory Malingering* (TOMM) y el *Victoria Symptoms Validity Test* (VSVT). Se realizó una búsqueda dirigida no exhaustiva, en diversas bases de datos y libros afines. El criterio de inclusión fue la utilización o revisión de los test. Se seleccionaron 68 trabajos publicados. Cada uno de los cuales fue analizado en base a las características técnicas de los instrumentos. La revisión ha demostrado que el TOMM y el VSVT son instrumentos pertinentes para valorar simulación y exageración de sintomatología cognitiva. Estos instrumentos cuentan con un cúmulo de investigaciones que avalan sus propiedades, pero también las limitaciones y restricciones de su utilización.

Palabras clave: evaluación, neuropsicología clínica, neuropsicología forense, simulación, síntomas cognitivos, test de validez del síntoma

Abstract

Most of the accuracy in neuropsychological assessment depends that the instruments used are standardized, valid and reliable measures. However, malingering, exaggeration and poor effort are aspects that can interfere with the interpretation of the results. The aim of this review was to describe and analyze two instruments in the context of forensic neuropsychology that assess malingering of cognitive symptoms: *Test of Memory Malingering* (TOMM) and *Victoria Symptoms Validity Test* (VSVT). Several scientific databases and related books were searched non-exhaustively. The criterion for inclusion in this review was the use of test. We selected 68 documents, each one was analyzed based on the technical characteristics of the instruments. Review showed that the TOMM and VSVT are relevant instruments for assessing malingering and exaggeration of cognitive symptoms. These instruments have several evidence that support their psychometric properties but also limitations and restrictions on their use.

Keywords: assessment, clinical neuropsychology, cognitive symptoms, forensic neuropsychology, malingering, symptoms validity test

Introducción

La "simulación" es definida según el DSM-IV-R como la producción intencionada y voluntaria de síntomas, tanto físicos como psicológicos, falsos o exagerados con el fin de alcanzar u obtener incentivos de tipo externo (American Psychiatric Association, 2000). Sin embargo, en la actualidad se cuenta con criterios más precisos para diagnosticar esta problemática que los ofrecidos originalmente por el DSM IV-R (Slick, Sherman, & Iverson, 1999). La prevalencia de este fenómeno se estima entre 13-66%, dependiendo del contexto de evaluación y el tipo de población estudiada (Duncan, 2005; Mittenberg, Patton, Canyock, & Condit, 2002; Greiffenstein, Baker, & Gola, 1994), siendo los déficit cognitivos –especialmente la memoria-, el trastorno más frecuentemente simulado (Surh, Tranel, Wefel, & Barrash, 1997), aunque también es bastante usual la simulación de patologías psiquiátricas (Brandt, 1988). Esta elevada tasa se relaciona con el incremento de los incentivos externos que puede generar la simulación y con la dificultad diagnóstica de validar dichos síntomas (Larrabee, 2005; Binder & Rohling, 1996).

Estos hechos han conducido a la creación de metodo-logías (Verdejo-García, Alcázar-Córcoles, Gómez-Jarabo, & Pérez-García, 2004; Iverson & Franzen, 1996; Ellwanger, Rosenfeld, Sweet, & Bhatt, 1999) y técnicas que permitan de manera específica y sensible aumentar la precisión con la que se detecta la simulación, a través de dos líneas de investigación diferenciadas. La primera, centrada en el análisis de los patrones de ejecución más habituales de sujetos simuladores en algunos de los instrumentos de evaluación neuropsicológica tradicionales (Ashendorf, O'Bryant, & McCaffrey, 2003; Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995). La segunda, ha centrado su interés en la creación de instrumentos capaces de detectar específicamente casos de simulación (Slick, Hopp, Strauss, & Thompson, 1997; Green, Allen & Astner, 1996; Tomabaugh, 1996; Hiscock & Hiscock, 1989). Sin embargo, si bien existe una gran variedad de instrumentos, no todos cuentan con los estándares técnicos necesarios para ser utilizados como pruebas en procesos judiciales. De este modo, el propósito de la siguiente revisión es, describir y analizar dos instrumentos en el contexto de la Neuropsicología Forense que evalúan simulación de síntomas cognitivos: el *Test of Memory Malingering* (TOMM), y el *Victoria Symptoms Validity Test* (VSVT). Se realizó una búsqueda dirigida no exhaustiva, de acuerdo a la revisión en diversas bases de datos de consultas bibliográficas

* Correspondencia: aaliaga@sm.l.cl. Área de Salud Mental, Servicio Médico Legal, Chile. Dirección: Av. La Paz 1012, Independencia.
Recibido: 31-10-12. Revisión desde: 12-04-13. Aceptado: 21-07-13.

(*Medline, PsycInfo, Scopus, PsycArticles*). Complementariamente se efectuaron búsquedas manuales en algunas fuentes bibliográficas del ámbito de la neuropsicología en espera de encontrar aproximaciones al tema que no estuvieran registradas en las bases bibliográficas habituales. El criterio de inclusión para la búsqueda, fue la utilización o revisión en los textos consultados de los instrumentos: el *Test of Memory Malinger* (TOMM) y el *Victoria Symptoms Validity Test (VSVT)*. Se excluyeron trabajos con diseños de caso único. En total se incluyeron 66 trabajos publicados, los cuales fueron clasificados en dos grupos: TOMM ($n=43$) y VSVT ($n=23$). Cada grupo fue analizado en función de las características técnicas de los instrumentos, estudiando aspectos como la (1) descripción de los instrumentos, (2) criterios de corrección, (3) normalización, (4) fiabilidad, (5) validez, (6) efectos demográficos (7) sensibilidad y especificidad.

Test of Memory Malinger

Descripción del instrumento

El Test of Memory Malinger (TOMM) es un test de reconocimiento visual utilizado para evaluar déficits de memoria fingidos o exagerados (Tombaugh, 1996). Es el test de validación del sintoma más utilizado por los neuropsicólogos que son expertos en casos de indemnización por lesiones (Sharland & Gfeller, 2007; Slick, Tan, Strauss, & Huktsch, 2004). Puede ser utilizado en personas mayores de 5 años, así como en población con baja escolaridad y de distintas culturas. Su utilidad se fundamenta en que la memoria de reconocimiento de imágenes es una capacidad que muestra una baja tasa de afectación tanto en sujetos adultos sanos como en una variedad de trastornos neurológicos (Rees & Tombaugh, 1998).

Se compone por dos ensayos de aprendizaje y una tarea de retención. En cada ensayo el examinado tiene que recordar 50 dibujos de objetos cotidianos, capacidad que es examinada a través del reconocimiento de los objetos aprendidos previamente en ítems de elección forzada con dos alternativas. Cabe señalar que después de cada respuesta de evaluado, el evaluador informa si el resultado es correcto o erróneo. Por su parte, la retención de los estímulos también evaluada mediante una tarea reconocimiento, pero sin un ensayo de aprendizaje y 15 minutos después de haber administrado el segundo ensayo. Como criterio de corrección, se otorga un punto por cada respuesta correcta en cada ensayo de reconocimiento y el de retención. La puntuación máxima del test es de 50 puntos para cada una de las fases (Strauss, Sherman & Spreen, 2006).

Tombaugh (1996) sostiene que las dos pruebas de reconocimiento son suficientes para evaluar simulación. Sin embargo, el uso de la tarea de retención permite corroborar los resultados obtenidos. Ahora bien, Greve y Bianchini (2006) sugieren que es necesario administrar el tercer ensayo, ya que se han observado tasas de hasta del 3% de simuladores que no fueron detectados en los dos ensayos anteriores. Estudios recientes plantean que también es posible detectar simulación a través del primer ensayo (Greve, Bianchini, Black, Heinly, Love, Swift, & Ciota, 2006; O'Bryant, Engel, Kleiner, Vasterling, & Black, 2007; Bauer, O'Bryant, Lynch, McCaffrey, & Fisher, 2007; Greve, Bianchini, & Doane, 2007; Horner, Bedwell, & Duong, 2006).

La versión informatizada del TOMM muestra resultados equivalentes a la versión manual, al menos en el paradigma de la simulación (Strauss, Sherman, & Spreen, 2006). Sin embargo, pueden surgir diferencias en el contexto clínico ante esta versión, ya que presenta una sensibilidad más elevada al bajo esfuerzo. Por otro lado, la versión por ordenador también puede ser preferida en función de evitar errores de imprecisión en la administración del test. En la versión computarizada no solo es posible analizar la proporción de respuestas correctas e incorrectas, sino que también se puede observar la latencia utilizada por los sujetos en cada ensayo. En este sentido, Rees y Tombaugh (1998), encontraron que la latencia para el grupo de simuladores es mayor (1 ó 2 segundos más) en comparación a la latencia observada muestras de pacientes normales.

Una nueva versión del TOMM fue desarrollada por su mismo autor. El *Test of Memory Malinger-2* (TOMM-2) se crea como una prueba suplementaria a la versión original. Esta fue diseñada para parecer ser una prueba más difícil que la versión original, a través de la utilización de 24 figuras geométricas abstractas. Si bien la forma de administración del TOMM-2 es idéntica al TOMM, la capacidad para discriminar entre grupos de simuladores y no simuladores del TOMM-2, no ha sido tan buena como la del TOMM (Strauss, Sherman, & Spreen, 2006).

Características psicométricas

El TOMM fue estandarizado en diversas muestras, con edades que fluctúan entre los 16 y 84 años, así como también, en una muestra de pacientes derivados para evaluación neuropsicológica que incluían diagnósticos de demencia, traumatismos craneoencefálicos, afasias, amnesias, esclerosis múltiple y enfermedad de parkinson ($n=135$). Tombaugh (1996; 1997) señaló que el rendimiento en el ensayo 2 fue muy alto en pacientes con trastornos neurológicos (a excepción de la demencia), observando que más del 95% de los adultos obtuvieron puntajes de 49 ó 50 en el segundo ensayo de reconocimiento. Más aun, los puntajes para las diferentes muestras clínicas mostraron que la mayoría de los pacientes con trastornos neurológicos obtenían puntuaciones casi perfectas en el segundo ensayo. Por el contrario, los simuladores mostraban rendimientos que estaban por debajo de los 45 puntos en el ensayo 2 ó en la tarea de retención. En base a lo anterior, Tombaugh (1996) recomienda que cualquier puntuación que se encuentre por debajo de los 45 en el ensayo 2 ó la tarea de retención puede ser indicativo que el evaluado no está realizando el máximo esfuerzo a la hora de responder al test. Además, el autor propone la utilización de un puntaje de corte de 45, a fin de aumentar la probabilidad de discriminar entre población clínica y personas que simulan un trastorno neurocognitivo (Tombaugh, 1996; 1997).

Constantinou y McCaffrey (2003), aportaron datos normativos para una muestra de 67 niños anglosajones. Los rendimientos observados, incluso en niños con antecedentes psiquiátricos y médicos de relevancia, fue superior a los 45 puntos en el ensayo 2. En el estudio desarrollado por Nagle, Everhart, Durham, McCammon y Walker (2006) observaron que para una muestra de niños ($n=35$), las puntuaciones para el segundo ensayo de aprendizaje ($M=49.67$; $Dt=.84$) también se encontraban por sobre los puntajes de corte propuestos originalmente.

Respecto a los efectos demográficos, Tombaugh (1996) sostiene que el test es poco sensible al efecto de la edad, mientras que otros han encontrado una efecto moderado para el segundo ensayo ($r=-.40$) y para el ensayo de retención ($r=-.50$) en población de adultos mayores (Teichner & Wagner, 2004). En población infantil no se han observado relaciones significativas en el segundo ensayo (Constantinou & McCaffrey, 2003; Nagle et al., 2006; Donders, 2005). Por otro lado, la variable educación, sexo y etnicidad han mostrado una baja relación con las puntuaciones del test tanto en población adulta como infantil (Tombaugh, 1996; Constantinou & McCaffrey, 2003; Teichner & Wagner, 2004; Gervais, Rohling, Gree, & Ford, 2004; Weinborn, Orr, Woods, Conover, & Feix, 2003).

En cuanto a la fiabilidad, de acuerdo a lo planteado por Tombaugh (1996), el instrumento muestra un elevado coeficiente alfa para los tres ensayos ($E_1=.94$, $E_2=.95$ y $E_3=.94$). Sin embargo, no se encontraron antecedentes en relación a fiabilidad test-retest, e interevaluadores.

Respecto a su validez, el TOMM es un test diseñado para ser sensible a la simulación de memoria pero insensible a los déficits de memoria de base orgánica (Strauss, Sherman, & Spreen, 2006; Lezak, Howieson, & Loring, 2004). Por ejemplo, Tombaugh (1996) administró una batería de tests, incluido entre ellos el TOMM, a un grupo de estudiantes universitarios a los que se les pidió simular síntomas cognitivos después de una lesión cerebral y tomo otro grupo como control. En el ensayo 2, todos los miembros del grupo control obtuvieron puntajes de 49 ó mayores (100% especificidad), mientras que el 93% del grupo de simuladores puntuó por debajo de 49 puntos (93% de sensibilidad). Análisis post-hoc revelaron que los miembros de este grupo, no fueron capaces de distinguir al TOMM como un instrumento capaz de detectar simulación de memoria.

En estudios posteriores (Tombaugh, 1997), los rendimientos obtenidos por un grupo de pacientes con TCE en procesos de litigio ($n=11$) fueron contrastados con un grupo de TCE no involucrados en procesos legales ($n=17$), un grupo de sujetos cognitivamente normales ($n=11$) y un grupo de pacientes con alteraciones neuropsicológicas focales ($n=12$). En el ensayo 1, los pacientes con TCE pero "sin riesgo" de simulación rindieron levemente por debajo del grupo de personas sin alteraciones y aquellos pacientes con lesiones focales. Sin embargo, en el segundo ensayo y la fase de retención, los rendimientos fueron similares al de los dos grupos anteriores. Ahora bien, el grupo "con riesgo" de simulación evidenció rendimientos más bajos que los otros tres grupos en los tres ensayos, diferencias que fueron significativas. Estos hallazgos también han sido corroborados en investigaciones posteriores (Greve & Bianchini, 2006; Vilar-López, Santiago-Ramajo, Gómez-Río, Verdejo-García, Llamas, & Pérez-García, 2007; Gavett, O'Bryant, Fisher, & McCaffrey, 2005).

En un intento por replicar el proceso de validación del TOMM, Haber y Fichtenberg (2006), recogieron una muestra ($n=50$) de pacientes con traumatismos craneoencefálicos (leves a graves), así como pacientes que buscaban algún tipo de compensación económica. Los resultados reprodujeron y ampliaron la evidencia encontrada por Tombaugh.

Utilizando puntajes de corte de 45 para el segundo ensayo, los estudios de Cragar, Berry, Fakhoury, Cibula, Schmitt (2006), Rees y Tombaugh (1997), Tombaugh (1996, 1997) mostraron una especificidad superior a 90% considerando poblaciones con variados déficit cognitivos. Sin embargo en pacientes con demencia, el porcentaje de clasificación correcta para el segundo ensayo fue aproximadamente del 27%. De manera similar, los hallazgos de Teichner y Wagner (2004) señalan que en poblaciones de adultos mayores (incluidos trastornos cognitivos leves), la mayoría de los sujetos obtiene en el ensayo 2 y la fase de retención puntajes casi perfectos, pero que los errores de clasificación eran elevados (> 70%), especialmente en pacientes con demencia. Evidencia similar ha sido reportada por Merten, Bossink, Schmand (2007) y Greve y Bianchini (2006) quienes han postulado que déficits cognitivos pueden interferir significativamente en los resultados de distintos test de simulación.

En este sentido, Colby (2001) propone que los falsos positivos pueden reducirse en pacientes con y sin demencia cuando se utiliza como puntaje de corte valores mayores a 15 errores en el ensayo 2 y la fase de retención. Por su parte Weinborn, et al. (2003) y Hurlley y Deal (2006) entregan evidencias para la utilización de estos puntajes de corte entre pacientes con retardo mental leve. Sin embargo, Simon (2007) utilizando las puntuaciones de corte originales, observó en un grupo (n= 22) de pacientes forenses diagnosticados con retardo mental leve, que obtenían en su mayoría puntuaciones que iban de 42 a 50 (M= 48.7, Dt= 1.9) para el ensayo 2, y puntuaciones que oscilaban entre los 46 a los 50 (M= 49.4, Dt= 1.12) para el ensayo de retención. Además mostró que la especificidad del TOMM en el ensayo 2 era de 95.2% y para la tarea de retención una tasa del 100%.

Al evaluar el efecto del feedback sobre los resultados del instrumento, no se han encontrado asociaciones significativas (Bolan, Foster, Schmand, & Bolan, 2002).

Pese a que la mayor parte de la evidencia ha centrado su esfuerzo en identificar bajos niveles de esfuerzo en poblaciones adultas. También existen hallazgos que prueban que el TOMM puede ser un instrumento útil para identificar niños que no realizan un esfuerzo óptimo (Constantinou & McCaffrey, 2003; Donders, 2005).

Sin embargo, Tombaugh (2002) sostiene que el TOMM no es un instrumento apropiado para evaluar si una persona está fingiendo un trastorno psiquiátrico. A favor de lo anterior, hay evidencia que sostiene que las escalas de validez del MMPI-2 miden un constructo diferente a lo que persiguen los test cognitivos de validación del síntoma (Nelson, Sweet, Berry, Bryant, & Granacher, 2007; Greiffenstein & Baker, 2007). Ahora bien, la prueba puede ser de alguna utilidad para detectar trastornos motivacionales en contextos psiquiátricos forenses (Weinborn et al., 2003; Simon, 2007; Tombaugh, 2002; Gierok, Dickinson & Cole, 2005). Por ejemplo, individuos acusados de homicidio, que fingen una incapacidad para enfrentar el juicio, muestran una tendencia a producir bajos puntajes en el TOMM (Weinborn et al., 2003; Heinze & Purisch, 2001). Utilizando los puntajes de cortes recomendados por su autor en contextos psiquiátricos forenses la especificidad llega a ser bastante elevada (90% o más), aunque su sensibilidad es baja (entre 60% y 67%).

El instrumento también ha demostrado su efectividad para diferenciar simulación entre sujetos que refieren alteraciones cognitivas producto de exposición a sustancias neurotóxicas (Greve & Bianchini, 2006; Greve, et al., 2006). Los resultados de esta investigación muestran tasas de especificidad del 100% pero con una pobre sensibilidad (50%).

Por otro lado, tanto la depresión (Duncan, 2005; Yanez, Fremouw, Tennant, Strunk, & Coker, 2006; Ashendorf, Constantinou, & McCaffrey, 2004; Rees, Tombaugh, & Boulay, 2001) como la ansiedad (Ashendorf, Constantinou, & McCaffrey, 2004), las alteraciones cognitivas asociadas con trastornos psicóticos (Duncan, 2005; Weinborn et al., 2003) o el dolor crónico (Gervais et al., 2004; Etherton, Bianchini, Greve, & Ciota, 2005), el entrenamiento (Rose, Hall, & Szalda-Petree, 1995) no muestran efectos interferentes sobre los resultados en el TOMM. Por ejemplo, Iverson, Le Page, Koehler, Shojania y Badii (2007), observaron una muestra (n= 54) de pacientes con el diagnóstico de fibromialgia a quienes se les administró el TOMM, el *Beck Depression Inventory-Second Edition* (BDI-II), el *Multidimensional Pain Inventory-Version 1* (MPI-1), el *Oswestry Disability Index-2.0* (OSI-2), el *British Columbia Cognitive Complaints Inventory* (BCCCI), y el *Fibromyalgia Impact Questionnaire* (FIQ). La mayoría de los sujetos evaluados reportó tener de síntomas depresivos leves (72 %), en tanto que el 22 % informó poseer niveles "severos" de la depresión. Las puntuaciones promedio en el TOMM eran 48.8 (Dt = 1.9) para el Ensayo 1, 49.8 (Dt = 0.5) para el Ensayo 2, y 49.6 (Dt = 0.9) para la fase de Retención. Los resultados muestran que a pesar de los niveles relativamente elevados de síntomas depresivos, dolor crónico e incapacidad, estos se correlacionan pobremente con los resultados en el TOMM.

En base a la validez concurrente, las puntuaciones del TOMM evidencian modestas asociaciones con medidas de aprendizaje y memoria e duración del coma (Tombaugh, 1996; Donders, 2003). En tanto que Merten et al. (2007) al estudiar distintas enfermedades neurológicas (con síntomas cognitivos asociados), a través de una extensa batería de test neuropsicológicos, entre los que se encontraban cuatro test de validación del síntoma como el TOMM, el WMT, el *Bremer Symptomvalidierung* (BSV), y el *Amsterdam Short-Term Memory Test* (ASTM). Encontraron que las correlaciones revelaron asociaciones disímiles entre las distintas pruebas de validación del síntoma y los test neuropsicológicos comunes. Mientras que TOMM y WMT correlacionó principalmente con pruebas de la memoria declarativa, el BSV mostró correlación con pruebas de atención, y el ASTM con pruebas de memoria de trabajo. Por otro lado, las intercorrelaciones entre medidas de validez de síntoma eran relativamente bajas.

Tombaugh (1996) advierte que la interpretación del TOMM involucra muchos factores, y que algunos pacientes pueden tender a fallar más en determinados test de validación del síntoma que en otros, debido a la naturaleza de los estímulos que componen el instrumento. Estos datos han sido corroborados por McCaffrey, O'Bryant, Ashendorf y Fisher, (2003) y Farkas, Rosenfeld, Robbins y van Gorp (2006). Por ejemplo, McCaffrey et al. (2003) señalan que aunque los puntajes del TOMM y el *Rey 15- Items Test* (Rey 15- Item) mostraron una moderada correlación (entre .89 y .78), el TOMM fue capaz de identificar a dos veces más litigantes (27%) como posibles simuladores que el Rey 15- Item (11%).

Aun cuando el TOMM muestra ser más eficiente que el test de Rey para detectar bajo esfuerzo, otros test evidencian indicadores más poderosos que el TOMM (Gervais et al., 2004; Vilar-López et al., 2007; Tan, Slick, Strauss, & Hultsch, 2002). Por ejemplo, Vilar-López et al. (2007) al comparar los resultados obtenidos en el VSVT y el TOMM para tres muestras de sujetos: una con síndrome postconmocional en litigio (n= 14), una con síndrome postconmocional sin estar en litigio (n= 12) y una de simuladores entrenados (n = 35). Encontraron que el TOMM permitía clasificar correctamente al 86% del grupo de simuladores entrenados y 41% del grupo de pacientes en litigio, a diferencia del VSVT que permitió clasificar 97.1% del grupo de simuladores y al 50% de los pacientes en litigio. En la investigación de O'Bryant y Lucas (2006), cuando el WMT fue usada como "*gold standar*" para comparar los resultados del TOMM. Este último presentó un VPP muy alto (.98) y VPN aceptable (.78).

Victoria Symptom Validity Test

Descripción del instrumento

El Victoria Symptom Validity Test (VSVT) es un test informatizado de reconocimiento visual de dígitos, utilizado para evaluar fingimiento o exageración de defectos de memoria. Esta recomendado su uso en contextos forense y clínicos, en poblaciones mayores de 18 años y personas de distintas culturas (Strauss et al., 2006).

El instrumento consta de 48 estímulos, que son presentados en tres bloques de 16 ítems. En esta tarea, se presenta al sujeto en la pantalla del ordenador, cinco dígitos durante unos segundos, seguido por un breve lapso de latencia (en el que la pantalla está en blanco), luego de lo cual se muestran dos alternativas, una que contiene el estímulo previamente presentado y la otra con un distractor. Además con el propósito de incrementar la sensibilidad del instrumento para detectar simulación, el VSVT utiliza dos procedimientos que incrementan la apariencia de dificultad de la prueba. El primero, modificando el período de latencia entre la exposición del estímulo y la fase de reconocimiento, el que va aumentando progresivamente de 10 a 15 segundos (Slick, Hopp, Strauss, & Spellacy, 1996). El otro es dividiendo los ítems en fáciles y difíciles (de acuerdo a la tarea de reconocimiento). Los fáciles no muestran ninguna similitud con los dígitos presentados en la fase de aprendizaje. Por el contrario, los ítems difíciles están compuestos por los mismos dígitos, aunque hay dos de ellos que están cambiados de posición. Cada uno de los tres bloques que componen el test, contiene el mismo número de ítem fáciles y difíciles. La distribución de los ítems según su dificultad y la posición (izquierda o derecha) en el que se presentan, se encuentra contrabalanceada y randomizada.

Debido a que todos los ítems de reconocimiento tienen dos alternativas, la tasa de error debiera ser cercana al 50%, bajo el supuesto que el sujeto responde al azar. Por tanto estos resultados pueden implicar dos cosas: a) que el sujeto tiene severas alteraciones de atención y/o memoria, o b) sea la manifestación de síntomas de exageración. En base a lo anterior, los simuladores son identificados por realizar ejecuciones muy pobres, ya que estos tienden a suponer que el test resulta ser difícil para personas con

problemas de memoria, más aun si va aumentando el intervalo de tiempo entre la exposición y el reconocimiento (Lezak et al., 2004).

La puntuación del test es calculada por el propio programa y puede ser impresa para cada uno de los bloques. La información obtenida indica: a) el número de respuestas correctas por cada bloque y dividida por el nivel de dificultad, b) el puntaje z derivado de una curva de probabilidad binomial, marcando el rendimiento que se obtiene al azar (50%), c) el sesgo es una medida que evalúa la tendencia a usar una tecla más que la otra, y d) el tiempo de latencia media y su desviación estándar, ante los estímulos fáciles y difíciles (Strauss et al., 2006).

A partir de los datos obtenidos, los rendimientos que se encuentra por debajo del azar como $p < .05$ son catalogados inequívocamente como inválidos o simulados, mientras que puntuaciones que están por sobre el azar como $p > .05$ se consideran inequívocamente como válidos. Una tercera categoría son etiquetados como resultados cuestionables, que corresponden a los puntajes que caen con un 90% confianza en el intervalo del azar (Slick et al., 1996; Loring, Lee, & Meador, 2005).

Características psicométricas

Las interpretaciones generadas por el programa computacional a partir de los resultados del VSVT han sido ampliamente analizadas a través del modelo binomial (Slick et al., 1996; Slick, Hopp, Strauss, Hunter, & Pinch, 1994). Esta aproximación provee un indicador acerca del grado de validez de los resultados de los evaluados, respecto de lo que es posible predecir en base al azar.

Es importante señalar que estas puntuaciones de corte no realizan diferencias respecto a las tasas basales de simulación. Sin embargo, el manual lo ofrece tanto para los ítems fáciles y difíciles, considerando diversas tasas basales de esta condición (Strauss et al., 2006).

Slick et al. (1996) encontró que personas con alteraciones pero que no buscaban ningún tipo de ganancia económica, obtenían puntuaciones de 15 ó menos en los ítems difíciles. Investigaciones posteriores (Doss, Chelune, & Naugle, 1999; Grote, Kooker, Garron, Nyenhuis, Smith, & Mattingly, 2000; Strauss, Slick, Levy-Bencheton, Hunter, MacDonald, & Hultsch, 2002) han validado y ampliado los resultados iniciales de Slick y sus colegas, a través de la evaluación de pacientes no-litigantes con significativas alteraciones cognitivas (incluidas severas alteraciones de memoria), pudiendo aportar puntajes de corte que permitan mejorar la sensibilidad del test. Por ejemplo, Grote et al. (2000) encontró que 28/30 (93%) pacientes no litigantes con un diagnóstico de epilepsia refractaria, obtuvieron puntuaciones superiores a 21, en comparación con los 19 puntos (36%) de los 53 pacientes litigantes. Recientemente, Loring et al. (2005) encontró que el 89% de pacientes no-litigantes con el diagnóstico de epilepsia refractaria obtenían puntuaciones de 21/24 en los ítems difíciles del VSVT.

Basándose en estos datos clínicos, las puntuaciones que se encuentran entre 16 y 20, tanto en los ítems fáciles y difíciles, no deben ser consideradas como válidas, sin embargo, pueden ser consideradas como puntuaciones que raramente se observan en pacientes no-litigantes o con daño cerebral. Junto con esto, resulta poco esperable que pacientes no-litigantes obtengan puntuaciones entre 0 y 5 (de un máximo de 8) en cualquiera de los bloques de ítems fáciles o difíciles (Loring et al., 2005).

Otros estudios han sugerido que el tiempo promedio para contestar a cada ítem no debieran ser superiores a 3 segundos en los ítems fáciles y no superior a 4 en los difíciles (Slick, Tan, Strauss, Mateer, Harnadek, & Sherman, 2003). En tanto para los efectos demográficos, no se han observado influencias sobre las puntuaciones del test en relación con la edad (Slick et al., 1996; Strauss et al., 2002) educación (Vilar-Lopez et al., 2007; Strauss et al., 2002), sexo (Slick et al., 1997; Strauss et al., 2002) o etnia (Vilar-Lopez et al., 2007).

Respecto a la fiabilidad mediante la consistencia interna Slick et al. (1996) informó que el alfa para los ítems fáciles, difíciles y el total fue de .82, .87 y .89 respectivamente. Sin embargo no se encontró información para la fiabilidad Inter-evaluadores. Mientras que la fiabilidad test-retest, el VSVT fue administrado en dos oportunidades a 31 participantes sanos y 27 pacientes litigantes (Slick et al., 1997). El intervalo de tiempo para la realización del retest fue de 14 días para el grupo control y 31 días para el grupo de litigantes. Las correlaciones test-retest para algunas medidas específicas fueron bajas (de .53 a .54 en latencia) para el grupo control. Mientras que para el grupo de pacientes litigantes las correlaciones fueron más altas (de .56 para el tiempo de respuestas en ítems fáciles, hasta .84 para la puntuación global). Por otro lado, todos los miembros del grupo control obtuvieron el mismo tipo de clasificación al retest (válidos). Pero para el grupo de paciente en litigio sólo el 86% obtuvo la misma clasificación al retest. Estos

hallazgos sugieren que el VSVT muestra una adecuada confiabilidad y sugiere que los cambios de clasificación entre el test y el retest, probablemente reflejan la sensibilidad del VSVT al cambio en el esfuerzo, aspecto que no es explicado por la varianza de error (Thompson, 2002).

En cuanto a la validez, el VSVT muestra ser un instrumento sensible a los efectos motivacionales (Guilmette, Hart, & Giuliano, 1993). Las investigaciones han mostrado consistentemente como simuladores entrenados y pacientes en situación de litigio que solicitan una compensación monetaria, obtienen puntajes más bajos en esta tarea (especialmente en los ítems difíciles) que aquellas muestras con pacientes controles o pacientes con alteraciones significativas pero que no están demandando una compensación económica (Slick et al., 1996; Thompson, 2002).

Por otro lado, existen datos fiables que permiten sostener que graves trastornos de memoria muestran un bajo impacto sobre los resultados del VSVT (Doss et al., 1999; Slick et al., 2003). Por ejemplo, Slick et al. (1997) encontraron una baja tasa de falsos positivos cuando el test fue aplicado a una muestra de pacientes ($n = 30$) que había sufrido lesiones cerebrales moderadas y graves, pero que no buscaban ningún tipo de compensación económica. De este grupo de pacientes, el 97% obtuvieron puntuaciones válidas en el VSVT, y solamente un paciente obtuvo una puntuación catalogada de cuestionable. Estos datos permiten demostrar que la presencia de puntuaciones que caen en el rango de cuestionable es inusual, incluso para pacientes con severos defectos cognitivos.

En el estudio desarrollado por Macciocchi, Seel, Alderson y Godsall (2006) a fin de evaluar la sensibilidad del VSVT, utilizó una muestra ($N = 71$) homogénea de pacientes agudos que presentaban lesiones cerebrales graves. Los resultados reflejaron que estos pacientes tuvieron un 99% de probabilidad de tener puntuaciones que se encontraron entre el 44.1 y el 46.8.

Loring et al. (2005) administraron el VSVT en 374 pacientes clasificados en tres grupos: con TCE (sin incentivo económico), pacientes sin TCE y un grupo no clínico pero derivado por abogados en casos de compensaciones laborales. La proporción de respuestas que se encontraban por debajo del azar en los ítems difíciles estaba presente únicamente en el 0.3% de los pacientes que no presentaban TCE. A diferencia de lo observado en el grupo no clínico, el 20 % obtuvo puntuaciones que significativamente se encontraban por de bajo del azar. Los datos sugieren que la distorsión intencionada de respuestas, es inusual en pacientes enviados a evaluación neuropsicológica no forense. Por otro lado, se ha observado que la depresión tiene una bajo impacto sobre los resultados del VSVT (Slick et al., 1996; Guilmette, Whelihan, Sparadeo, & Buongiorno, 1994).

Estudios sobre validez divergente también han demostrado una baja correlación entre el VSVT y puntuaciones de test diseñados para medir constructos diferentes. Slick et al. (1996) señalaron que los test que no miden memoria explican solo el 5% de la varianza tanto para ítems fáciles y difíciles del VSVT, lo que implica que este se ve poco afectado por distintos niveles del funcionamiento cognitivo, y por lo tanto, puntuaciones bajas reflejan con mayor propiedad algún grado de exageración o bajo esfuerzo. Sin embargo, los tiempos de respuesta (en ítems fáciles y difíciles) muestran menor divergencia al ser correlacionados con los resultados del *Digit Span* (de .32 a .53) y medidas que evalúan velocidad de procesamiento como el Test de Stroop. De este modo, se recomienda tener precaución a la hora de interpretar el tiempo de respuesta en el VSVT.

Por otro lado, los resultados del VSVT también se han comparado con otros test que miden bajo esfuerzo, a fin de observar su efectividad (Vilar-López et al., 2007). En este sentido, Tan, et al. (2002) usando el modelo de simulación, comparó los resultados del VSVT, el TOMM y el *Word Memory Test (WMT)*. El VSVT mostró ser mejor que el TOMM cuando se intentó detectar el fingimiento de una incapacidad asociada con lesiones cerebrales. Sin embargo, al usar la prueba de retención del TOMM, combinado con el VSVT o el WMT, este mejoró su capacidad de predicción de un modo significativo. Se han observado también que el VSVT al ser comparado con test como el *Key 15 - Items*, el *21 - Items Test*, el *Portland Digit Recognition Test (PDRT)*, ha mostrado una mejor capacidad de clasificación (Vickery, Berry, Inman, Harris & Orey, 2001).

El VSVT (ítems difíciles) también ha mostrado ser mejor predictor al factor de entrenamiento, más que el *Dot Counting Test (DCT)*, *Reliable Digit Span (RDS)* o el índice vocabulario menos el span de dígitos, obtenido a partir de la Escala de Inteligencia para Adultos Weschler (Strauss, Hultsch, Hunter, Slick, Patry, & Levy-Bencheton, 1999). Otras investigaciones han mostrado que en tareas de reconocimiento llega a ser el mejor en la detección de baja colaboración (Rose et al., 1995), aun cuando existen estudios que plantean que si bien muestra una buena capacidad para clasificar a simuladores, sus resultados se encuentran por debajo de los rendimientos que es posible observar en el WMT (Tan et al., 2002).

Las correlaciones entre escalas de validez del MMPI-2 y los puntajes del VSVT tienden a ser medias y bajas (Slick et al., 1996). Para los ítems fáciles del VSVT y la escala F del MMPI-2, la correlación es de $r = -.29$, lo que puede ser interpretado como que elevados puntajes en la escala F se asocian con bajos puntajes en el VSVT para los ítems fáciles. Por otro lado, la escala FBS mostró una asociación más elevada (entre .42 y .43 respectivamente), lo que puede ser explicado ya que la escala FBS es de mayor utilidad en contextos en los que se exagera la sintomatología por un trauma cerebral (ver MMPI-2). Sin embargo, en la clasificación de los participantes muestran una baja tasa de acuerdo. Esto, quizás no resulte sorprendente ya que las tareas difieren considerablemente en su naturaleza y dominio (Greifenstein et al., 1994).

Conclusiones

El TOMM ha mostrado ser el test más utilizado para identificar simulación de problemas cognitivos (Sharland & Gfeller, 2007) demostrando que los puntajes del test se ven poco interferidos por variables como la edad, el sexo, la baja educación, diferentes étnicas, la ansiedad, la depresión, el dolor o diferentes patologías neurológicas. Por ello, tiene la suficiente validez para reunir los criterios de Daubert, sobre admisibilidad como evidencia científica que puede ser presentada ante los tribunales (Vallabhajosula & Van Gorp, 2001). La elevada especificidad del test, reduce la tasa de falsos positivos, pero solo cuando se excluye a poblaciones con demencia moderada o retardo mental moderado. Sin embargo, los puntajes de corte propuestos por Colby (2001), pueden ser de utilidad para evaluar sujetos con demencias leves, psicosis residual y retardos mentales leves.

Sin embargo, en comparación con otros test de validación del síntoma (WMT, VSVT, CARB), el TOMM muestra una menor sensibilidad y una mayor tendencia a realizar falsos positivos, mientras que sus resultados tienden a ser redundantes cuando es acompañado de otros test que evalúan bajo esfuerzo (Farkas et al., 2006).

Por último, recientes investigaciones han mostrado que el TOMM es relativamente insensible a diferentes tipos de entrenamiento (Powell, Gfeller, Hendricks, & Sharland, 2004). Sin embargo, la posibilidad de obtener información través de internet respecto a como responder al test, por ende, engañarlo, implica que hay que tomar en cuenta estos hechos a la hora de aplicar el instrumento y analizar sus resultados (Bauer & McCaffrey, 2006).

De modo similar el TOMM, el VSVT también es ampliamente utilizado (Sharland & Gfeller, 2007) por aquellos profesionales que buscan identificar sujetos que intentan simular o exagerar síntomas cognitivos. Numerosas investigaciones sostienen que pacientes en procesos judiciales que presentan bajas puntuaciones en este instrumento, pueden ser consideradas como un indicador confiable de posible simulación. Esto, debido a que variables como la edad, la escolaridad, la depresión y el grado de sofisticación han mostrado un bajo efecto sobre los resultados en el VSVT, incluso la presencia de trastornos severos de memoria a penas repercute en la ejecución del test.

El hecho que el VSVT pueda ser administrado de manera informatizada permite mejorar su confiabilidad a través de la minimización de cualquier error de varianza (aplicación por más de un evaluador en series repetidas). Si bien en estudios iniciales, los autores utilizaron el modelo de probabilidad binomial con estrictos criterios de corte para guiar la interpretación de los resultados, en la actualidad los estudios sugieren otros puntajes de cortes que han ayudado a mejorar la sensibilidad del instrumento. Por ser un instrumento confiable, sensible, de fácil aplicación y que demanda de poco tiempo para su utilización, Thompson (2002) señala que el VSVT, es probablemente el test que mejor reúne los criterios de Daubert, en cuanto poder ser una prueba científica admisible en un proceso judicial. No obstante, es importante realizar algunas observaciones. En primer lugar, la posición que ocupa el VSVT en una batería de evaluación puede influir en los resultados del test. Guilmette, Whelihan, Hart, Sparadeo y Buongiorno (1996) señalan que muchos errores tienen lugar cuando el test se utiliza al final de la evaluación, después que los pacientes han tenido experiencia con pruebas de memoria más exigentes. En segundo lugar, la prueba es generalmente utilizada en adultos, por lo que se necesitan normas para evaluar a niños, ya que el VSVT no cuenta con tales. Tercero, aunque el VSVT es fácil de administrar, aun cuenta con pocas investigaciones (Hartman, 2002) que permitan respaldar su fortaleza, ya que la mayor parte de estudios al respecto han sido realizados por los propios autores de la prueba. Por último, obtener buenos resultados en el VSVT no descarta la simulación o exageración de déficit cognitivos (Guilmette et al., 1994). Asimismo, aun en casos donde es posible comprobar la existencia de incentivos económicos,

el paciente puede estar legítimamente discapacitado y/o actuando de manera inconsciente (Slick et al., 1999).

Referencias

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington D. C., EEUU.
- Ashendorf, L., O'Bryant, S. E., & McCaffrey, R. J. (2003). Specificity of malingering detection strategies in older adults using the CVLT and WCST. *The Clinical Neuropsychologist*, 17, 255-262.
- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology*, 19, 125-130.
- Bauer, L., O'Bryant, S. E., Lynch, J. K., McCaffrey, R. J., & Fisher, J. M. (2007). Examining the Test Of Memory Malingering Trial 1 and Word Memory Test Immediate Recognition as screening tools for insufficient effort. *Assessment*, 14, 215-22.
- Bauer, L., & McCaffrey, R. J. (2006). Coverage of the Test of Memory Malingering, Victoria Symptom Validity Test, and Word Memory Test on the Internet: is test security threatened? *Archives of Clinical Neuropsychology*, 21, 121-126.
- Binder, L. M., & Rohling, M. L. (1996). Money matters: a meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal of Psychiatry*, 153, 7-10.
- Bolan, B., Foster, J. K., Schmand, B., & Bolan, S. (2002). A comparison of three tests to detect feigned amnesia: The effects of feedback and the measurement of response latency. *Journal of Clinical and Experimental Neuropsychology*, 24, 154-167.
- Brandt, J. (1988). Malingered amnesia. En R. Rogers (Ed.), *Clinical assessment of malingering and deception*. New York, EEUU: The Guildford Press.
- Colby, F. (2001). Using the binomial distribution to assess effort: forced-choice testing in neuropsychological settings. *Neurorehabilitation*, 16, 311-312.
- Constantinou, M., & McCaffrey, R. J. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, 9, 81-90.
- Cragar, D. E., Berry, D. T., Fakhoury, T. A., Cibula, J. E., & Schmitt, F. A. (2006). Performance of patients with epilepsy or psychogenic non-epileptic seizures on four measures of effort. *Clinical Neuropsychology*, 20, 552-566.
- Donders, J. (2005). Performance on the test of memory malingering in a mixed pediatric sample. *Child Neuropsychology*, 11, 221-227.
- Doss, R. C., Chelune, G. J., & Naugle, R. I. (1999). Victoria Symptom Validity Test: Compensation-seeking vs. non-compensation-seeking patients in a general clinical setting. *Journal of Forensic Neuropsychology*, 1, 5-20.
- Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the test of memory malingering (TOMM). *Assessment*, 12, 123-129.
- Ellwanger, J., Rosenfeld, J. P., Sweet, J., & Bhatt, M. (1999). Identifying simulators of cognitive deficit through combined use of neuropsychological test performance and event-related potentials. *Journal of Clinical and Experimental Neuropsychology*, 21, 866-879.
- Etherton, J. L., Bianchini, K. J., Greve, K. W., & Ciota, M. A. (2005). Test of Memory Malingering Performance is unaffected by laboratory-induced pain: implications for clinical use. *Archives of Clinical Neuropsychology*, 20, 375-84.
- Farkas, M. R., Rosenfeld, B., Robbins, R., & van Gorp, W. (2006). Do tests of malingering concur? Concordance among malingering measures. *Behavioural Science and Law*, 24, 659-71.
- Gavett, B. E., O'Bryant, S. E., Fisher, J. M., & McCaffrey, R. J. (2005). Hit rates of adequate performance based on the test of memory malingering (TOMM) Trial 1. *Applied Neuropsychology*, 12, 1-4.
- Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology*, 19, 475-487.
- Gierok, S. D., Dickson, A. L., & Cole, J. A. (2005). Performance of forensic and non-forensic adult psychiatric inpatients on the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, 20, 755-60.
- Guilmette, T. J., Hart, K. J., & Giuliano, A. J. O. (1993). Malingering detection: The use of a forced-choice method in identifying organic versus simulated memory impairment. *The Clinical Neuropsychologist*, 7, 59-69.
- Guilmette, T. J., Whelihan, W., Sparadeo, F. R., & Buongiorno, G. (1994). Validity of neuropsychological test results in disability evaluations. *Perceptual and Motor Skills*, 78, 1179-86.

- Guilmette, T. J., Whelihan, W. M., Hart, K. J., Sparadeo, F. R., & Buongiorno, G. (1996). Order effects in the administration of a forced-choice procedure for detection of malingering in disability claimants' evaluations. *Perceptual and Motor Skills*, 83, 1007-1016.
- Greve, K. W., & Bianchini, K. J. (2006). Should the Retention trial of the Test of Memory Malingering be optional? *Archives of Clinical Neuropsychology*, 21, 117-9.
- Greve, K. W., Bianchini, K. J., Black, F. W., Heinly, M. T., Love, J. M., Swift, D. A., & Ciota, M. (2006). Classification accuracy of the Test of Memory Malingering in persons reporting exposure to environmental and industrial toxins: Results of a known-groups analysis. *Archives of Clinical Neuropsychology*, 21, 439-48.
- Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006). Classification accuracy of the test of memory malingering in traumatic brain injury: results of a known-groups analysis. *Journal of Clinical and Experimental Neuropsychology*, 28, 1176-90.
- Green, P., Allen, L., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer-administered forms (US version 1.1)*. Durham, UK: CogniSyst.
- Greiffenstein, M. F., & Baker, W. J. (2007). Validity testing in dually diagnosed post-traumatic stress disorder and mild closed head injury. *The Clinical Neuropsychologist*, 18, 1-18.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingering amnesia measures with a large clinical sample. *Psychological Assessment*, 6, 218-224.
- Grote, C. L., Kooker, E. K., Garron, D. C., Nyenhuis, D. L., Smith, C. A., & Mattingly, M. L. (2000). Performance of compensation seeking and non-compensation seeking samples on the Victoria Symptom Validity Test: Cross-validation and extension of a standardization study. *Journal of Clinical and Experimental Neuropsychology*, 22, 709-719.
- Haber, A. H., & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head trauma sample. *The Clinical Neuropsychologist*, 20, 524-532.
- Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing: neuropsychological malingering and the Word Memory Test. *Archives of Clinical Neuropsychology*, 17, 709-714.
- Heinze, M. C., & Purisch, A. D. (2001). Beneath the mask: Use of psychological tests to detect and subtype malingering of criminal defendants. *Journal of Forensic Psychology Practice*, 1, 23-52.
- Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method of detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, 11, 967-974.
- Horner, M. D., Bedwell, J. S., & Duong, A. (2006). Abbreviated form of the test of memory malingering. *International Journal of Neuroscience*, 116, 1181-1186.
- Hurley, K. E., & Deal, W. P. (2006). Assessment instruments measuring malingering used with individuals who have mental retardation: potential problems and issues. *Mental Retardation*, 44, 112-9.
- Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, 18, 38-51.
- Iverson, G. L., Le Page, J., Koehler, B. E., Shojania, K., & Badii, M. (2007). Test of Memory Malingering (TOMM) scores are not affected by chronic pain or depression in patients with fibromyalgia. *The Clinical Neuropsychologist*, 21, 532-46.
- Larrabee, G. J. (2005). Assessment of malingering. En G. J. Larrabee (Ed.). *Forensic neuropsychology: A scientific approach*. New York, EEUU: Oxford University Press.
- Lezak, M., Howieson, D., & Loring, D. (2004). *Neuropsychological assessment*. New York, EEUU: Oxford University Press.
- Loring, D. W., Lee, G. P., & Meador, K. J. (2005). Victoria symptom validity test performance in non-litigating epilepsy surgery candidates. *Journal of Clinical and Experimental Neuropsychology*, 27, 610-617.
- Maccocchi, S. N., Seel, R. T., Alderson, A., & Godsall, R. (2006). Victoria Symptom Validity Test performance in acute severe traumatic brain injury: implications for test interpretation. *Archives of Clinical Neuropsychology*, 21, 395-404.
- McCaffrey, R. J., O'Bryant, S. E., Ashendorf, L., & Fisher, J. M. (2003). Correlations among the TOMM, Rey-15, and MMPI-2 validity scales in a sample of TBI litigants. *Journal of Forensic Neuropsychology*, 3, 45-53.
- Merten, T., Bossink, L., & Schmand, B. (2007). On the limits of effort testing: symptom validity tests and severity of neurocognitive symptoms in nonlitigant patients. *Journal of Clinical and Experimental Neuropsychology*, 29, 308-318.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24, 1094-1102.
- Mittenberg, W., Theroux-Fichera, S., Zielinski, R. E., & Heilbronner, R. L. (1995). Identification of Malingered Head Injury on the Wechsler Adult Intelligence Scale-Revised. *Professional Psychology: Research and Practice*, 26, 491-498.
- Nagle, A. M., Everhart, D. E., Durham, T. W., McCammon, S. L., & Walker, M. (2006). Deception strategies in children: examination of forced choice recognition and verbal learning and memory techniques. *Archives of Clinical Neuropsychology*, 21, 777-785.
- Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of International Neuropsychology Society*, 13, 440-449.
- O'Bryant, S. E., Engel, L. R., Kleiner, J. S., Vasterling, J. J., & Black, F. W. (2007). Test of memory malingering (TOMM) trial 1 as a screening measure for insufficient effort. *The Clinical Neuropsychologist*, 21, 511-21.
- O'Bryant, S. E., & Lucas, J. A. (2006). Estimating the predictive value of the Test of Memory Malingering: An illustrative example for clinicians. *The Clinical Neuropsychologist*, 20, 533-540.
- Powell, M. R., Gfeller, J. D., Hendricks, B. L., & Sharland, M. (2004). Detecting symptom- and test-coached simulators with the test of memory malingering. *Archives of Clinical Neuropsychology*, 19, 693-702.
- Rees, L., & Tombaugh, T. (1998). Five Validation Experiments of the Tests of Memory Malingering (TOMM). *Psychological Assessment*, 10, 10-20.
- Rees, L., Tombaugh, T., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, 16, 501-506.
- Rose, F. E., Hall, S., & Szalda-Petree, D. (1995). Portland Digit Recognition Test-computerized: Measuring response latency improves the detection of malingering. *The Clinical Neuropsychologist*, 9, 124-134.
- Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, 22, 213-323.
- Simon, M. J. (2007). Performance of mentally retarded forensic patients on the test of memory malingering. *Journal of Clinical Psychology*, 63, 339-44.
- Slick, D. J., Hopp, G., Strauss, E., Hunter, M., & Pinch, D. (1994). Detecting dissimulation: Profiles of simulated malingers, traumatic brain-injury patients, and normal controls on a revised version of Hiscock and Hiscock's Forced-Choice Memory Test. *Journal of Clinical and Experimental Neuropsychology*, 16, 472-481.
- Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria Symptom Validity Test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology*, 18, 911-922.
- Slick, D., Hopp, G., Strauss, E., & Thompson, G. (1997). *The Victoria Symptom Validity Test*. Odessa, EEUU: ParInc.
- Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic Criteria for Malingered Neurocognitive Dysfunction: Proposed Standards for Clinical Practice and Research. *The Clinical Neuropsychologist*, 13, 545-561.
- Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: a survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465-73.
- Slick, D. J., Tan, J. E., Strauss, E., Mateer, C. A., Harnadek, M., & Sherman, E. M. S. (2003). Victoria Symptom Validity Test scores of patients with profound memory impairment: Non-litigant case studies. *The Clinical Neuropsychologist*, 17, 390-394.
- Strauss, E., Hultsch, D. F., Hunter, M., Slick, D. J., Patry, B., & Levy-Bencheson, J. (1999). Using intraindividual variability to detect malingering in cognitive performance. *The Clinical Neuropsychologist*, 13, 420-32.
- Strauss, E., Sherman, E., & Spreen, O. A. (2006). *Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York, EEUU: Oxford University Press.
- Strauss, E., Slick, D. J., Levy-Bencheson, J., Hunter, M., MacDonald, S. W., & Hultsch, D. F. (2002). Intraindividual variability as an indicator of malingering in head injury. *Archives of Clinical Neuropsychology*, 17, 423-444.
- Surh, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology*, 19, 500-514.
- Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, 16, 495-505.

- Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology*, 19, 455-464.
- Tombaugh, T. N. (1996). Test of Memory Malingering. New York, EEUU: MHS.
- Tombaugh, T. (1997). The Test of Memory Malingering (TOMM): Normative Data From Cognitively Intact and Cognitively Impaired Individuals. *Psychological Assessment*, 9, 260-8.
- Tombaugh, T. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. *Journal of Forensic Neuropsychology*, 2, 69-96.
- Thompson, G. B. (2002). The Victoria Symptom Validity Test: An enhanced test of symptom validity. Special Issue: Detection of response bias in forensic neuropsychology. *Journal of Forensic Neuropsychology, Part I*, 2, 43-67.
- Vallabhajosula, B., & van Gorp, W. G. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of American Academy of Psychiatry and Law*, 29, 207-15.
- Verdejo-García, A., Alcázar-Córcoles, M. A., Gómez-Jarabo, G. A., & Pérez-García, M. (2004). Pautas para el desarrollo científico y profesional de la neuropsicología forense. *Revista de Neurología*, 39, 60-73.
- Vickery, C. D., Berry, D. T., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: a meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, 16, 45-73.
- Vilar-López, R., Santiago-Ramajo, S., Gómez-Río, M., Verdejo-García, A., Llamas, J. M., & Pérez-García, M. (2007). Detection of malingering in a Spanish population using three specific malingering tests. *Archives of Clinical Neuropsychology*, 22, 379-388.
- Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the test of memory malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology*, 25, 979-990.
- Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J., & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology*, 21, 161-165.