

Inventarios de estilos de aprendizaje: una actualización de la matriz de fiabilidad y validez de Coffield, Moseley, Hall, y Ecclestone

Francisco José Álvarez-Montero¹, María Guadalupe Leyva-Cruz¹, Flérida Moreno-Alcaraz¹

¹ Facultad de Ciencias de la Educación, Universidad Autónoma de Sinaloa

México

Correspondencia: Francisco J. Álvarez Montero. Facultad de Ciencias de la Educación, Ave. Cedros y Calle Sauces s/n Fracc. Los Fresnos, C.P. 80034. Culiacán, Sinaloa, México.
E-mail: francisco_alvarez_montero@uas.edu.mx

© Universidad de Almería e Ilustre Colegio Oficial de la Psicología de Andalucía Oriental (Spain)

Resumen

Introducción. Una de las muchas nociones que se han introducido en educación en los últimos 50 años, es la de los estilos de aprendizaje. Esta idea es muy popular y goza de buena aceptación. Sin embargo, desde una perspectiva científica la idea es muy controversial. En este sentido, Coffield, Moseley, Hall, y Ecclestone, realizaron la primera revisión sistemática sobre el tema, y mostraron que los instrumentos psicométricos de estilos de aprendizaje son no tienen niveles adecuados de fiabilidad y validez. Sin embargo, los resultados de su revisión no han sido actualizados desde 2004. Esta contribución presenta una revisión sistemática de estos aspectos para el periodo 2005-2010, con el objetivo de actualizar la información sobre el tema.

Método. Se examinaron 58 documentos que contenían suficiente información sobre fiabilidad y validez acerca de los siguientes inventarios: el LSI de Kolb, el ILS de Felder, el TSI de Sternberg, los inventarios sensoriales tipo VAK-VAKT-VARK, los inventarios del modelo de Entwistle (ASSIST, LASSI y RASI), el SPQ de Biggs, el LSQ de Honey y Mumford y el Delineador de Estilos de Gregorc (GSD).

Resultados. El análisis de las propiedades psicométricas señala que: (1) más de la mitad de los inventarios escrutados carece de consistencia interna y validez predictiva, (2) alrededor del 31% no tiene fiabilidad test-retest, (3) no hay evidencia acerca de por lo menos un criterio de fiabilidad y validez para cerca del 31% de ellos, (4) el perfilador de estilos de aprendizaje de Jackson (LSP) sigue siendo un inventario sin evidencia en todos los indicadores.

Discusión y conclusiones. A pesar de la popularidad y aceptación de los estilos de aprendizaje, la recomendación inicial hecha hace 14 años por Coffield, Moseley, Hall, y Ecclestone, de no basar las intervenciones pedagógicas tomando como referencia únicamente a alguno de los instrumentos de estilos de aprendizaje sigue siendo válida.

Palabras Clave: Estilos de aprendizaje, inventarios, fiabilidad test-retest, validez convergente, validez factorial, consistencia interna, revisión sistemática.

Abstract

Introduction. One of the many notions that have been introduced in education in the last 50 years is that of learning styles. This idea is very popular and enjoys good acceptance. However, from a scientific perspective the idea is very controversial. In this sense, Coffield, Moseley, Hall, and Ecclestone, made the first systematic review on the subject, and showed that the psychometric instruments of learning styles do not have adequate levels of reliability and validity. However, the results of its review have not been updated since 2004. This contribution presents a systematic review of these aspects for the 2005-2010 period, with the aim of updating the information on the subject.

Method. Fifty-eight documents were examined, which contained enough information about the reliability and validity of the following inventories: Kolb's LSI, Felder's ILS, Sternberg's TSI, the VAK-VAKT-VARK sensory inventories, the inventories based on the model proposed by Entwistle (ASSIST, LASSI and RASI), Biggs' SQP, Honey and Mumford's LSQ and Gregorc's Style Delineator (GSD).

Results. The analysis of the psychometric properties indicates that: (1) more than half of the inventories surveyed lack internal consistency and predictive validity, (2) around 31% have no test-retest reliability, (3) there is no evidence on at least one reliability and validity criteria for close to 31 % of them (4) Jackson's Learning Styles Profiler (LSP) remains an inventory without evidence on all reliability and validity criteria.

Discussion and Conclusion. Despite the popularity and acceptance of learning styles, the initial recommendation made 14 years ago by Coffield, Moseley, Hall, and Ecclestone, of not basing pedagogical interventions solely on any of the learning styles instruments is still valid.

Keywords: Learning styles, instruments, test-retest reliability, convergent validity, factorial validity, internal consistency, systematic review.

Introducción

Una de las muchas nociones que se han introducido en educación en los últimos 50 años, es la de los estilos de aprendizaje. Esta idea es muy popular y goza de buena aceptación. En una revisión de publicaciones científicas Lilienfeld, Lynn, Ruscio, y Beyerstein (2010) contabilizaron, sólo en el año de 2008, 1,984 artículos en revistas arbitradas, 919 presentaciones en conferencias y 701 libros o capítulos de libro sobre estilos de aprendizaje. Un análisis posterior de 220 artículos, indexados en las bases de datos de investigación de ERIC y PubMed en el periodo 2013-2015, detectó que más del 85% de la literatura comienza y termina con una visión positiva de los estilos de aprendizaje (Newton, 2015).

Sin embargo, como se ha demostrado en diversas áreas (Gottfredson, 2009, Nirenburg, McShane, y Beale, 2004, Rao y Andrade, 2011), un gran número de citas, así como la popularidad y aceptación de ideas, métodos, constructos e instrumentos no implica que sean científicamente válidos y presenten resultados positivos. En este sentido, Coffield, Moseley, Hall, y Ecclestone (2004) realizaron la primera revisión sistemática acerca de la validez científica de los estilos de aprendizaje. Esta revisión encontró que no hay evidencia en la literatura que apoye la hipótesis de que emparejar la estrategia de enseñanza con los estilos de los alumnos mejore el aprendizaje, es decir no se observaron interacciones aptitud-tratamiento (ATI). Además, la misma revisión no recomienda utilizar los inventarios de estilos debido a que carecen de fiabilidad y validez.

Objetivos e hipótesis

Análisis más recientes (Cuevas, 2015; Pashler, McDaniel, Rohrer, y Bjork., 2008) han confirmado los resultados iniciales con respecto a las ATI. Sin embargo, la información sobre fiabilidad y validez no ha sido actualizada desde 2004. Las revisiones sistemáticas y metanálisis deben actualizarse periódicamente para evitar que conclusiones científicas desactualizadas influyan en la toma de decisiones, así como en intervenciones futuras (Lakens, Hilgard, Stakes, 2016). Consecuentemente, esta investigación presenta un esfuerzo para actualizar la matriz de fiabilidad y validez de Coffield et al. (2004), para artículos publicados entre 2005 y 2010, de la misma manera que Cuevas (2015) actualizó la revisión de Pashler et al. (2008) sobre el emparejamiento entre estrategias de enseñanza y estilos de aprendizaje. La pregunta de investigación que se intenta responder es la siguiente: ¿se han corregido las deficiencias de fiabilidad y validez en los inventarios de estilos de aprendizaje?

El resto de este documento está organizado de la siguiente forma. Primero, se introducen las reglas para la interpretación de resultados estadísticos sobre fiabilidad y validez. Segundo, se describe el método de búsqueda y selección de la literatura. Tercero, se presentan los resultados del análisis, instrumento por instrumento. Por último, se discuten estos resultados y se hacen algunas conclusiones.

Método

Procedimiento

Para la búsqueda de la literatura se utilizó Google Académico (GA). Se puede argumentar que GA no es una base de datos bibliográfica o de literatura científica; sin embargo, los resultados de estudios recientes (Haddaway, Collins, Coughlin, y Kirk, 2015; Harzing, y Alakangas, 2016), indican que GA está al mismo nivel en términos de cobertura, y recuperación de literatura relevante que Scopus, Web of Science y PubMed. El texto utilizado para la búsqueda fue el siguiente: “*learning AND (style OR styles) AND (inventory OR inventories) AND (validity OR reliability)*”. Se revisaron los primeros 160 resultados de GA año por año. En total, 99 documentos fueron descargados para revisión. Se excluyeron del análisis final aquellos manuscritos que no contuvieran información sobre fiabilidad y validez, los que no hacían referencia a instrumentos de estilos de aprendizaje, los que no presentasen suficiente información estadística, así como los inventarios que tuviesen menos de 3 estudios. La figura 1 muestra este proceso.

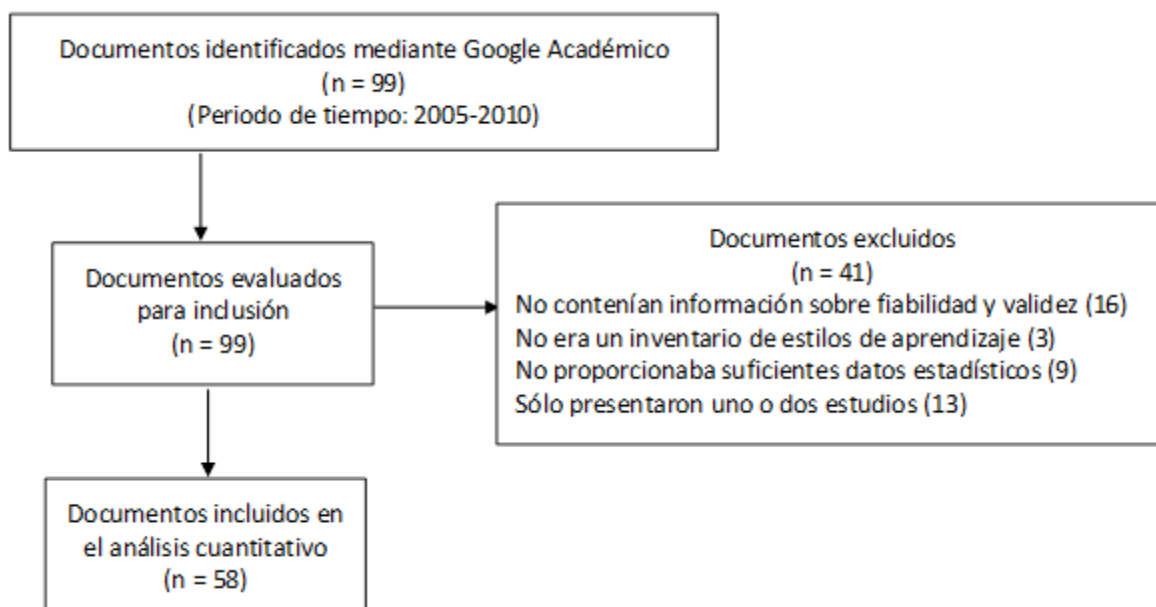


Figura 1. Proceso de búsqueda y selección de literatura.

Análisis de datos

Los valores umbrales que indican consistencia interna, fiabilidad test-retest, validez predictiva y convergente aceptables, se muestran en la tabla 1 de acuerdo con las recomendaciones de Henson (2001), Lezak, Howieson, Bigler, y Tranel (2012), Mukaka (2012) y Carlson y Herdman (2012) respectivamente. En cuanto a la validez estructural o factorial, la tabla 2 presenta los valores que según Iacobucci (2010) y Bagozzi y Yi (2012), pueden tomar los índices de bondad de ajuste, para ser considerados suficientemente buenos.

Tabla 1. *Valores umbrales para consistencia interna, fiabilidad test-retest, validez predictiva y validez convergente*

Consistencia interna	Fiabilidad test-retest	Validez predictiva	Validez convergente
$\alpha \geq .800$	$r \geq .700$	$r \geq .300$	$r \geq .700$

Tabla 2. *Valores umbrales para índices de bondad de ajuste de la validez estructural*

Ji al cuadrado relativa ($\chi^2/g.l.$)	Raíz cuadrada media residual estandarizada (SRMR)	Raíz cuadrada media de error de aproximación (RMSEA)	Índice de ajuste normativo (NNFI)	Índice de ajuste comparativo (CFI)
$\chi^2/g.l. \leq 3$	$SRMR \leq .070$	$RMSEA \leq .070$	$NNFI \geq .920$	$CFI \geq .930$

Para que un instrumento tenga una consistencia o fiabilidad aceptable todas sus escalas deben presentar coeficientes con valores mayores iguales a los de la tabla 1. Un inventario tiene validez factorial adecuada si por lo menos 2 de los índices concurren en ello (Bagozzi y Yi, 2012). Realizar un análisis de componentes principales (PCA), para determinar la validez estructural de un instrumento, se juzgará como falta de evidencia. El PCA, no es un método de análisis factorial, y sus múltiples defectos indican que su uso debe evitarse (Costello y Osborne, 2005; Schmitt, 2011). En su lugar se debe realizar un análisis factorial exploratorio (EFA), o bien, un confirmatorio (CFA). En caso de que un estudio presente menos de 2 índices de ajuste, se considerará que no hay evidencia de validez estructural. Existiera validez de constructo, si el instrumento cumple con los criterios de validez convergente y factorial. Cuando haya evidencia sólo de validez convergente o factorial, ya sea a favor o en contra, se marcará como un indicador sin resolver.

Como parte del análisis, se calculará la estadística omega cuadrada parcial (ω_p^2) para la estadística F de las pruebas ANOVA (Lakens, 2013), y se convertirá en coeficiente de correlación. Los resultados de las pruebas χ^2 , también se transformarán en correlaciones utili-

zando la fórmula estándar (Rosenberg, 2010). Cuando sea posible, las correlaciones se presentarán con sus intervalos de confianza del 95%.

Resultados

Este apartado describe la información obtenida de 8 instrumentos. Se presentan en orden descendente, del inventario con más estudios al de menos estudios.

El inventario de estilos de aprendizaje de Kolb (LSI)

Kayes (2005) analizó la consistencia interna y la validez estructural del LSI ($N=221$). Sólo 2 de las dimensiones mostraron una consistencia adecuada: CE ($\alpha=.77$), AC ($\alpha=.76$), RO ($\alpha=.82$) y AE ($\alpha=.82$). La validez factorial se intentó determinar mediante un PCA.

Miller (2005) realizó dos estudios ($N=36$ y $N=34$), para determinar la validez predictiva del LSI. En el primero, se determinó que no hay efecto de los estilos de aprendizaje en la cantidad de material aprendido, $F(3, 32)=.704$, $p=.557$, $\omega_p^2 =-.025$, $r=-.158$, 95% IC [-.463,.180], ni en el resultado de la evaluación final, $F(3,32)=.821$, $p=.492$, $\omega_p^2 =-.015$, $r=-.123$, 95% IC [-.443,.225]. Ninguno de los estilos tuvo una asociación significativa con el rendimiento académico, con correlaciones que oscilaron entre $r=-.290$, $p>.05$, 95% IC [-.565,.043] y $r=.160$, $p>.05$, 95% IC [-.178,.464]). El segundo experimento señala que sólo hay efectos triviales no significativos de los estilos en la cantidad de material aprendido, $F(3, 30)=1.40$, $p=.262$, $\omega_p^2 =.034$, $r=.184$, 95% IC [-.164,.492] y con el resultado de la evaluación final, $F(3,30) =.969$, $p=.420$, $\omega_p^2 =-.003$, $r=-.055$ 95% IC [-.386,.289]. En el segundo estudio, la dimensión AC tuvo una correlación negativa significativa con la cantidad de material aprendido, $r=-.370$, $p<.05$, 95% IC [-.629,-.036]. El resto de los estilos sólo presentó asociaciones triviales no significativas que fluctuaron entre $r=-.240$, $p>.05$, 95% IC [-.535,.107] y $r=.290$, $p>.05$, 95% IC [-.552,.053].

Cooper Brathwaite (2006) midió la validez predictiva del LSI ($N=76$). Los resultados indican que los estilos no fueron un predictor significativo del nivel de conocimiento o del grado de competencia cultural, mostrando correlaciones que variaron entre $r=.055$, $p=.640$, 95%, IC [-.173,.277] y $r=.141$, $p=.220$, 95% IC [-.087,.355].

Wang, Wang, Wang, y Huang (2006) probaron la validez predictiva del LSI ($N=455$). El análisis señala que los estilos tuvieron una asociación trivial significativa con el rendimiento académico, $F(3, 442)=6.81$, $p=.0002$, $\omega_p^2 =.038$, $r=.195$, 95% IC [.105,.282].

Chapman y Calhoun (2006) examinaron la validez estructural, así como la convergen- te entre el LSI y el inventario de Preferencias de Aprendizaje (LPI) de Rezler y Rezmovic($N=94$). Sólo las escalas abstractas de los instrumentos (i.e., LSI-AC y LPI-AB) mostraron una correlación positiva pero insuficiente entre ellas, $r=.296$, $p<.05$, 95% IC [.099,.470]. La validez factorial se intentó establecer mediante un PCA.

Demirbas y Demirkan (2007) aplicaron el LSI a 3 grupos distintos ($N=111$, $N=88$ y $N=74$), para examinar su consistencia interna y validez predictiva. No se observó una consistencia aceptable en ninguno de los grupos. Los coeficientes oscilaron entre $\alpha=.51$ y $\alpha=.73$. Aunque uno de los objetivos del estudio fue probar la validez predictiva de los estilos en relación con el rendimiento académico, no se reporta el resultado de ninguna prueba estadística, para analizar la asociación estilos-rendimiento.

Yukselturk y Bulut (2007) investigaron la validez predictiva del LSI($N=80$). Ningún estilo fue un predictor significativo del rendimiento académico. Las correlaciones fluctuaron entre $r=-.080$, $p>.05$, 95% IC [-.295,.142] y $r=.150$, $p>.05$, 95 % IC [-.072,.358].

Lu, Jia, Gong y Clark (2007) aplicaron el LSI($N=40$) para evaluar su validez predictiva. No se encontró una asociación significativa entre estilos de aprendizaje y rendimiento académico, $\chi^2(3,N=40)=2.707$, $p=.538$, $r=.260$, 95% IC [-.056,.529].

Akkoyunlu y Yilmaz-Soylu (2008) estudiaron la validez predictiva del LSI($N=34$). Los resultados indican que no hay asociación entre estilos y rendimiento académico, $F(1,32)=.389$, $p=.537$, $\omega_p^2=-.018$, $r=-.134$ 95% IC [-.452,.214].

Sun, Lin, y Yu (2008) examinaron la validez predictiva del instrumento de Kolb($N=56$). Se encontró una asociación trivial no significativa entre estilos y rendimiento académico, $F(3,51)=1.532$, $p=.217$, $\omega_p^2=.028$, $r=.167$ 95% IC [-.100,.412].

An y Yoo (2008) investigaron la validez predictiva del inventario ($N=742$). Se estableció que los estilos tienen una relación significativa trivial con el pensamiento crítico, $r=.219$, $p<.001$ 95% IC [.149,.286].

Metallidou y Platsidou (2008) analizaron la consistencia interna y validez estructural del LSI($N=338$). Las dimensiones no mostraron una consistencia adecuada: CE($\alpha=.77$), RO($\alpha=.78$), AC($\alpha=.73$), AE($\alpha=.78$). La validez factorial se intentó examinar mediante un PCA.

Andreou, Andreou y Vlachos (2008) aplicaron el LSI para analizar su consistencia interna y validez predictiva ($N=452$). Las variables dependientes fueron los aspectos fonológicos, sintácticos y semánticos de la fluidez verbal. Se observó un nivel aceptable de consistencia en todas las dimensiones: CE($\alpha=.83$), RO($\alpha=.89$), AC($\alpha=.86$), AE($\alpha=.82$). Las correlaciones entre estilos y el aspecto fonológico fluctuaron entre $r=-.021$, $p>.05$ 95% IC [-.113,.071] y $r=.246$, $p<.01$, 95% IC [.157,.331]. Dos estilos tuvieron una asociación significativa, pero trivial con el aspecto sintáctico. El estilo Accomodative(AE+CE), $r=.286$, $p<.01$, 95% IC [.199,.369], y el Convergent(AC+AE), $r=.250$, $p<.01$, 95% IC [.161,.335]). La regresión múltiple corroboró estos resultados para los aspectos fonológico, $F(4,447)=2.58$, $p=.037$, $\omega_p^2=.014$, $r=.118$ 95% IC [.026,.208], sintáctico, $F(4,447)=4.54$, $p=.001$, $\omega_p^2=.030$, $r=.173$ 95% IC [.082,.261], y semántico, $F(4,447)=8.04$, $p=0.000003$, $\omega_p^2=.059$, $r=.243$, 95% IC [.154,.328].

Brittan-Powell, Legum, y Taylor (2008) investigaron la validez predictiva del LSI ($N=102$). Se encontró un efecto trivial no significativo entre estilos y rendimiento académico, $F(3,100)=1.44$, $p=.47$, $\omega_p^2=.013$, $r=.114$ 95% IC [-.082,.302].

Cagiltay (2008) evaluó la validez predictiva del inventario de Kolb ($N=285$). Los resultados indican una correlación trivial significativa entre estilos y rendimiento académico, $F(3,281)=6.06$, $p=.01$, $\omega_p^2=.051$, $r=.226$ 95% IC [.113,.333].

Erdem (2009) analizó la consistencia interna del instrumento ($N=103$). Todas las dimensiones mostraron una consistencia inadecuada: CE($\alpha=.58$), RO($\alpha=.70$), AC($\alpha=.71$), AE($\alpha=.65$).

Yilmaz-Soylu y Akkoyunlu (2009) escrutaron la validez predictiva del LSI ($N=139$). No se observó una relación entre estilos y rendimiento académico, $F(1,37)=.41$, $p=.526$, $\omega_p^2=-.015$, $r=-.123$, 95% IC [-.422,.200].

Platsidou y Metallidou (2009) investigaron la consistencia interna y validez estructural del inventario ($N=340$). Sólo una de las dimensiones alcanzó una consistencia aceptable: CE($\alpha=.81$), RO($\alpha=.72$), AC($\alpha=.76$) y AE($\alpha=.76$). La validez factorial se analizó mediante un PCA.

Jamieson (2010) midió la validez convergente del instrumento de Kolb ($N=68$), con respecto a 2 inventarios similares construidos con las técnicas de Likert ($N=34$) y de diferencial semántico ($N=34$). Las correlaciones entre el LSI y los otros dos instrumentos no alcanza-

ron el nivel mínimo deseable y oscilarion entre $r=.402$, $p<.05$, 95% IC [.225,.553] y $r=.360$, $p<.05$, 95% IC [.178,.518].

Azevedo y Akdere (2010) aplicaron el LSI para determinar su validez predictiva a un grupo experimental($N=106$) y uno de control($N=96$). Las correlaciones para el grupo experi- mental oscilarion entre $r=-.109$, $p>.05$, 95% IC [-.294,.083] y $r=.098$, $p>.05$, 95% IC [- .095,.283]. Mientras que las del grupo de control fluctuaron de $r=-.059$, $p>.05$, 95% IC [- .256,.143] a $r=.177$, $p>.05$, 95% IC [-.024,.365].

La tabla 3 resume los resultados presentados anteriormente.

Tabla 3. *Resumen de los resultados sobre fiabilidad y validez del LSI*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez pre- dictiva
1. Kayes (2005)	SÍ	SE	SE	SE
2. Miller (2005)	SE	SE	SE	NO
3. Cooper Brathwaite (2006)	SE	SE	SE	NO
4. Wang et al. (2006)	SE	SE	SE	NO
5. Chapman y Cal- houn (2006)	SE	SE	SR	SE
6. Demirbas y De- mirkan (2007)	NO	SE	SE	SE
7. Yukselturk y Bu- lut (2007)	SE	SE	SE	NO
8. Lu et al. (2007)	SE	SE	SE	NO
9. Akkoyunlu y Yil- maz-Soylu (2008)	SE	SE	SE	NO
10. Sun, Lin, y Yu (2008)	SE	SE	SE	NO
11. An y Yoo (2008)	SE	SE	SE	NO
12. Metallidou y Platsidou (2008)	NO	SE	SE	SE
13. Andreou, Andreou y Vlachos (2008)	SE	SE	SE	NO
14. Brittan-Powell, Legum, y Taylor (2008)	SE	SE	SE	NO
15. Cagiltay (2008)	SE	SE	SE	NO
16. Erdem (2009)	NO	SE	SE	SE
17. Yilmaz-Soylu y Akkoyunlu (2009)	SE	SE	SE	NO
18. Platsidou y Meta- llidou (2009)	NO	SE	SE	SE
19. Jamieson (2010)	SE	SE	SR	SE
20. Azevedo y Akdere	SE	SE	SE	NO

(2010)

Nota. SE (Sin Evidencia); SR (Sin Resolver).

El índice de estilos de aprendizaje de Felder (ILS)

Cook (2005) examinó la consistencia interna, la fiabilidad tes-retest, así como la validez convergente del ILS ($N=138$) con respecto a el indicador de tipo de estilo de aprendizaje (LSTI) de Pelley. No se observó una consistencia adecuada: Active-Reflexive ($\alpha=.61$), Sensing-Intuitive ($\alpha=.78$), Visual-Verbal ($\alpha=.70$), Sequential-Global ($\alpha=.67$). Una de las dimensiones no alcanzó una fiabilidad aceptable: Active-Reflective ($r=.75$), Sensing-Intuitive ($r=.81$), Visual-Verbal ($r=.60$), Sequential-Global ($r=.81$). Para la validez convergente, las dimensiones comparadas fueron Active-Reflective (ILS) vs. Extravert-Introvert (LSTI) y Sensing-Intuitive (ILS) vs. Sensing-Intuitive (LSTI). No se alcanzó el grado de validez convergente mínimo aceptable, con correlaciones que fluctuaron entre $r=.60$ y $r=.69$ respectivamente.

Cook y Smith (2006) midieron la consistencia interna, la fiabilidad test-retest, así como la validez convergente del inventario ($N=89$) con respecto al LSI de Kolb, el analizador de estilos de cognitivos (CSA) de Riding y el LSTI de Pelley. Ninguna de las dimensiones mostró suficiente consistencia interna: Active-Reflective ($\alpha=.62$), Sensing-Intuitive ($\alpha=.77$), Visual-Verbal ($\alpha=.72$), Sequential-Global ($\alpha=.65$). Dos de las dimensiones alcanzaron un buen nivel de fiabilidad: Active-Reflexive ($r=.81$), Sensing-Intuitive ($r=.86$), Visual-Verbal ($r=.70$) y Sequential-Global ($r=.68$). Sólo las facetas Active-Reflective y Sensing-Intuitive del LSTI y del ILS, tuvieron una convergencia casi aceptable con correlaciones de $r=.681$ y $r=.684$ respectivamente. El resto de los valores de convergencia entre el ILS, el LSI y el CSA, estuvieron muy por debajo del valor umbral ($r \leq .495$), llegando hasta valores negativos en el caso de la faceta Sequential-Global entre el ILS y el CSA ($r=-.206$).

Galvan (2006) administró el ILS para estudiar su validez predictiva ($N=84$). Ninguno de los estilos fue un predictor significativo del rendimiento académico, con correlaciones que oscilaron entre $r=-.059$, $p=.421$, 95% IC $[-.270, .157]$ y $r=.053$, $p=.398$, 95% IC $[-.163, .264]$.

Strang (2008) investigó la consistencia interna y la validez factorial del ILS ($N=715$). Dos de las dimensiones no alcanzaron un nivel adecuado de consistencia: Active-Reflective ($\alpha=.857$), Sensing-Intuitive ($\alpha=.786$), Visual-Verbal ($\alpha=.827$), Sequential-Global ($\alpha=.798$). La validez estructural se intentó establecer mediante un PCA.

Self y Widmann (2009) analizaron la validez predictiva del ILS($N=86$). Sólo se observaron correlaciones nulas y triviales entre estilos y el nivel de conocimiento conceptual de los alumnos, con asociaciones que fluctuaron entre $r=-.177$, $p=.102$, 95% IC $[-.375,.036]$ y $r=.224$, $p=.038$, 95% IC $[.013,.416]$.

McChlery y Visser (2009) aplicaron el ILS a 2 muestras distintas de participantes ($N=469$ y $N=266$), para medir su consistencia interna. No se observó un nivel de consistencia aceptable en el primer grupo: Active-Reflective($\alpha=.50$), Sensint-Intuitive($\alpha=.60$), Visual-Verbal($\alpha=.63$), Sequential-Global($\alpha=.45$). Tampoco en el segundo: Active-Reflective($\alpha=.53$), Sensint-Intuitive($\alpha=.63$), Visual-Verbal($\alpha=.66$), Sequential-Global($\alpha=.46$).

Choi, Lee, y Kang (2009) investigaron la validez predictiva del ILS($N=59$). No se encontró una asociación entre estilos y rendimiento en la resolución de problemas en ninguna de las dimensiones: Active-Reflective($F(8,106)=1.72$, $p=.10$, $\omega_p^2=.043$, $r=.207$, 95% IC $[-.052,.440]$), Sensing-Intuitive($F(8,106)=.68$, $p=.71$, $\omega_p^2=-.023$, $r=-.152$ 95% IC $[-.393,.108]$), Visual-Verbal ($F(8,106)=.52$, $p=.82$, $\omega_p^2=-.035$. $r=-.187$ 95% IC $[-.073,.423]$), Sequential-Global($F(8,106)=1.01$, $p=.43$, $\omega_p^2=.001$, $r=.032$ 95% IC $[-.226,.286]$).

Ku y Shen (2009) escrutaron la consistencia interna y validez factorial del ILS($N=1034$). El instrumento no mostró una consistencia adecuada: Active-Reflective($\alpha=.48$), Sensing-Intuitive($\alpha=.53$), Visual-Verbal($\alpha=.52$), Sequential-Global($\alpha=.41$). El CFA no pudo validar la estructura de 4 factores bipolares propuesta por Felder, $\chi^2(896,N=1034)= 4636.37$, $p<.001$, $\chi^2/g.l.=5.175$, $CFI=.505$.

Platsidou y Metallidou (2009) indagaron la consistencia interna y validez factorial del inventario($N=340$). No hubo un nivel de consistencia aceptable: Active-Reflective($\alpha=.45$), Sensing-Intuitive($\alpha=.62$), Visual-Verbal($\alpha=.51$) y Sequential-Global($\alpha=.45$). La validez estructural se intentó determinar mediante un PCA.

Zaharias, Andreou, y Vosinakis (2010) analizaron la validez predictiva del ILS ($N=196$). Los resultados indican que hay una relación negativa pero insignificante entre estilos y rendimiento académico, $F(7,188)=.375$, $p=.916$, $\omega_p^2=-.023$, $r=-.152$ 95% IC $[-.286,.012]$.

Alaoutinen (2010) aplicó el ILS para verificar su validez predictiva($N=84$). Sólo la dimensión Active-Refelctive tuvo una asociación no trivial significativa con el rendimiento: Active-Reflective($r=.320$, $p=.05$, 95% IC $[.113,.500]$), Sensing-Intuitive($r=.160$, $p=.320$, 95%

IC [-.056, .362]), Visual-Verbal($r=-.120$, $p=.470$, 95% IC [-.326,.097]), Sequential-Global($r=-.010$, $p=.940$, 95% IC [-.224,.205]).

Hosford y Siders (2010) examinaron la consistencia interna, fiabilidad test-retest y la validez estructural del ILS($N=358$). Ninguna dimensión alcanzó una consistencia interna aceptable: Active-Reflective($\alpha=.76$), Sensing-Intuitive($\alpha=.62$), Visual-Verbal($\alpha=.64$), Sequential-Global($\alpha=.62$). Las correlaciones test-retest se obtuvieron para 2 y 4 años. Únicamente la dimensión Sensing-Intuitive alcanzó el nivel mínimo deseado: Active-Reflective($r=.58, r=.61$), Sensing-Intuitive($r=.77, r=.72$), Visual-Verbal($r=.65, r=.55$) y Sequential-Global ($r=.57, r=.47$). La validez factorial se intentó establecer mediante un PCA.

Clarke, Lesh, Trocchio, y Wolman (2010) estudiaron la consistencia interna, así como la validez convergente del instrumento($N=86$), con respecto al inventario de estilos de pensamiento (TSI) de Sternberg. Ninguna de las facetas alcanzó un nivel aceptable de consistencia: Active-Reflective($\alpha=.51$), Sensing-Intuitive($\alpha=.79$), Visual-Verbal($\alpha=.75$), Sequential-Global($\alpha=.59$). Sólo las dimensiones Sensing-Intuitive y Sequential-Global tuvieron asociaciones no triviales significativas con el TSI. La primera con los estilos Executive($r=.408$, $p<.001$ 95% IC [.215,.571]), Conservative($r=.414$, $p<.001$ 95% IC [.222,.575]) y Local($r=.444$, $p<.001$, 95% IC [.251, .603]) del TSI. La segunda con Conservative($r=.374$, $p<.001$, 95% IC [.174, .545]) y Local($r=.513$, $p<.001$, 95% IC [.334, .656]) de ese mismo inventario. Sin embargo, ninguna de estas correlaciones alcanzó el nivel mínimo deseable.

La tabla 4 resume los resultados presentados previamente.

Tabla 4. Resumen de resultados de fiabilidad y validez del ILS de Felder

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Cook (2005)	NO	NO	SR	SE
2. Cook y Smith (2006)	NO	NO	SR	SE
3. Galvan (2006)	SE	SE	SE	NO
4. Strang (2008)	NO	SE	SE	SE
5. Self y Widmann (2009)	SE	SE	SE	NO
6. McChlery y Visser (2009)	NO	SE	SE	SE
7. Choi, Lee, y Kang (2009)	SE	SE	SE	NO
8. Ku y Shen (2009)	NO	SE	SR	SE
9. Platsidou y Meta-	NO	SE	SE	SE

llidou (2009)					
10. Zaharias, An- dreou, y Vosinakis (2010)	SE	SE	SE	NO	
11. Alaoutinen (2010)	SE	SE	SE	NO	
12. Hosford y Siders, (2010)	NO	NO	SE	SE	
13. Clarke et al. (2010)	NO	SE	SR	SE	

Nota. SE (Sin Evidencia); SR (Sin Resolver).

Los inventarios de estilos sensoriales (VAK, VAKT, VARK)

A diferencia de los 2 inventarios anteriores, no existe en la literatura un único instrumento de estilos sensoriales de aprendizaje. Diferentes autores proponen diferentes instrumentos. Sin embargo, coinciden por lo menos en los estilos visual y auditivo. En conjunto se les puede identificar por las siglas VAK, VAKT o VARK.

Isemonger y Sheppard (2007) administraron el cuestionario de preferencias de estilos sensoriales (PLSPQ) de Reid, para examinar su consistencia interna y validez factorial ($N=691$). Sólo 2 de 6 escalas, obtuvieron una consistencia aceptable: Visual ($\alpha=.37$), Auditory ($\alpha=.39$), Tactile ($\alpha=.67$), Kinesthetic ($\alpha=.76$), Group ($\alpha=.83$), Individual ($\alpha=.84$). El CFA no confirmó la estructura de 6 factores propuesta por Reid, $SRMR=.09$, $RMSEA=.09$, $CFI=.72$.

Çirkinoglu y Demirci (2007) evaluaron la validez predictiva del inventario VAK de Barsh ($N=573$). Los resultados señalan que los estilos no tienen validez predictiva con respecto al rendimiento académico: Visual ($r=.037$, $p>.05$, 95% IC [-.045,.119]), Auditory ($r=.043$, $p>.05$, 95% IC [-.039,.124]), Kinesthetic ($r=.076$, $p>.05$, 95% IC [-.006,.157]). La prueba ANOVA confirmó estos resultados, $F(3,569)=1.24$, $p=.296$, $\omega_p^2=.001$, $r=.032$ 95% IC [-.050,.114].

Slack y Norwich (2007) investigaron la consistencia interna y la fiabilidad test-retest del inventario VAK de Smith ($N=51$). Ninguna escala alcanzó un grado de consistencia adecuado: Visual ($\alpha=.63$), Auditory ($\alpha=.75$), Kinesthetic ($\alpha=.56$). Sólo una de las dimensiones no tuvo un nivel de fiabilidad aceptable: Visual ($r=.90$), Auditory ($r=.96$), Kinesthetic ($r=.75$). Sin embargo, es importante subrayar que este último criterio se evaluó con una muestra menor ($N=25$).

Peters, Jones, y Peters (2008) aplicaron el PLSPQ de Reid, para analizar su consistencia interna, validez predictiva y factorial ($N=338$). Sólo 2 de 6 escalas obtuvieron una consis-

tencia aceptable: Visual($\alpha=.68$), Auditory($\alpha=.53$), Tactile($\alpha=.67$), Kinesthetic($\alpha=.72$), Group($\alpha=.91$), Individual($\alpha=.90$). Las correlaciones muestran que los estilos tienen muy poco o ningún impacto en el rendimiento académico: Visual($r=.102$, $p>.05$, 95% IC [-.005,.206]), Auditory($r=.048$, $p>.05$, 95% IC [-.059,.154]), Tactile($r=.025$, $p>.05$, 95% IC [-.082,.131]), Kinesthetic($r=.075$, $p>.05$, 95% IC [-.032,.180]), Group($r=-.178$, $p<.01$, 95% IC [-.279,-.073]), Individual ($r=.247$ $p>.001$ 95% IC [.144,.345]). La validez estructural se intentó determinar mediante un PCA.

Yildirim, Acar, Bull, y Sevinc (2008) midieron la consistencia interna y validez predictiva del PLSPQ de Reid($N=746$). Sólo 2 de 6 escalas obtuvieron una consistencia aceptable: Visual($\alpha=.72$), Auditory($\alpha=.63$), Tactile($\alpha=.67$), Kinesthetic($\alpha=.64$), Group($\alpha=.80$), Individual($\alpha=.80$). No hubo interacciones entre los estilos y el rendimiento académico (ver tabla 5).

Tabla 5. Resultados de la prueba F y correlaciones estilo-rendimiento académico

Dimensión	Prueba F	Correlación
Visual	$F(1,1443)=.179$, $p>.05$, $\omega_p^2=-.00057$	$r=-.024$ 95% IC [-.096,.048]
Auditory	$F(1,1443)=.053$, $p>.05$, $\omega_p^2=-.00066$	$r=-.026$ 95% IC [-.098,.046]
Tactile	$F(1,1443)=.039$, $p>.05$, $\omega_p^2=-.00067$	$r=-.026$ 95% IC [-.098,.046]
Kinesthetic	$F(1,1443)=.012$, $p>.05$, $\omega_p^2=-.00068$	$r=-.026$ 95% IC [-.098,.046]
Group	$F(1,1443)=.046$, $p>.05$, $\omega_p^2=-.00066$	$r=-.026$ 95% IC [-.098,.046]
Individual	$F(1,1443)=.087$, $p>.05$, $\omega_p^2=-.00063$	$r=-.025$ 95% IC [-.097,.047]

Dobson (2009) testó la validez predictiva del VARK de Fleming($N=901$). Se observó una relación trivial pero significativa entre estilos y rendimiento académico, $\eta_p^2=.05$, $r=.224$, 95% IC [.161,.285].

Leite, Svinicki, y Shi (2010) examinaron la validez factorial del inventario VARK de Fleming($N=14,211$). Se aplicaron 3 métodos distintos para llevar a cabo el CFA: CTCM, CTCU y CT-C(M-1). Tanto el CTCM ($RMSEA=.047$, $SRMR=.064$) como el CTCU ($RMSEA=.034$, $SRMR=.049$), comprobaron la estructura de 4 factores propuesta por Fleming. Sin embargo, el tercero no pudo, aunque el valor del SRMR estuvo muy cerca del mínimo requerido ($RMSEA=.053$, $SRMR=.074$).

La tabla 6 resume los resultados presentados anteriormente.

Tabla 6. Resumen de resultados de fiabilidad y validez de los inventarios sensoriales

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
---------	----------------------	------------------------	-----------------------	--------------------

1. Isemonger y Sheppard (2007)	NO	SE	SR	SE
2. Çirkinoglu y Demirci (2007)	SE	SE	SE	NO
3. Slack y Norwich (2007)	NO	NO	SE	SE
4. Peters, Jones, y Peters (2008)	NO	SE	SE	NO
5. Yildirim et al. (2008)	NO	SE	SE	NO
6. Dobson (2009)	SE	SE	SE	NO
7. Leite, Svinicki y Shi, Y. (2010)	SE	SE	SR	SE

Nota. SE (Sin Evidencia); SR (Sin Resolver).

Los inventarios de estrategias de aprendizaje y habilidades de estudio (ASSIST, LASSI, RASI)

Al igual que con los inventarios de estilos sensoriales, no hay un único instrumento de estrategias de aprendizaje y habilidades de estudio. Sin embargo, todos están basados en el modelo de Entwistle (Coffield et al., 2004).

Cano (2006) administró el LASSI (Learning and Study Strategies Inventory) a dos muestras diferentes ($N=527$ y $N=429$), y analizaron su consistencia interna y la validez factorial. Sólo 2 de 10 dimensiones mostraron un nivel de consistencia adecuado: Attitude ($\alpha=.61, \alpha=.53$), Motivation ($\alpha=.72, \alpha=.68$), Time Management ($\alpha=.84, \alpha=.82$), Anxiety ($\alpha=.78, \alpha=.76$), Concentration ($\alpha=.80, \alpha=.79$), Information Processing ($\alpha=.74, \alpha=.73$), Selecting Main Ideas ($\alpha=.66, \alpha=.68$), Study Aids ($\alpha=.62, \alpha=.60$), Self-testing ($\alpha=.63, \alpha=.66$), Test Strategies ($\alpha=.72, \alpha=.73$). Aunque los autores no especifican la muestra que utilizaron, el CFA no pudo corroborar la estructura de 3 factores propuesta por los creadores del inventario, $\chi^2(26)=96.67, p<.01; \chi^2/g.l.=3.718; CFI=.95$. Sin embargo, se pudo determinar una relación no trivial significativa entre estilos y rendimiento académico para el primer ($r=.326, p<.001, 95\% IC [.247, .400]$) y segundo ($r=.427, p<.001, 95\% IC [.346, .501]$) grupo de participantes.

Geertshuis y Fazey (2006) administraron el RASI (Revised Approaches to Study Inventory), para medir su consistencia interna y validez estructural ($N=389$). Ninguna de las dimensiones alcanza alfas adecuados: Deep ($\alpha=.70$), Surface ($\alpha=.73$), Strategic ($\alpha=.76$). La estructura latente del instrumento se intentó determinar mediante un PCA.

Diseth, Pallesen, Hovland, y Larsen (2006) aplicaron el ASSIST (Approaches and Study Skills Inventory for Students), para examinar su validez predictiva y factorial ($N=486$). Todas las dimensiones tuvieron asociaciones insignificantes con el rendimiento académico:

Deep($r=.190$, $p<.01$, 95% IC [.103,.274]), Surface($r=-.180$, $p<.01$, 95% IC [-.265,-.093]), Strategic($r=.240$, $p<.01$, 95% IC [.154, .322]). El CFA comprobó la estructura de 3 factores propuesta por los desarrolladores del instrumento, $\chi^2(45,N=486)=120.86$, $p<.01$, $\chi^2/g.l.=2.689$, $CFI=.96$, $RMSEA=.06$.

Diseth (2007) investigó la consistencia interna, validez predictiva y factorial del ASSIST($N=206$). Únicamente una dimensión no tuvo una consistencia aceptable: Deep($\alpha=.80$), Surface($\alpha=.74$) y Strategic($\alpha=.82$). Sólo el estilo Strategic tuvo una asociación positiva no trivial con el rendimiento académico: Deep($r=.250$, $p<.01$, 95% IC [.117,.374]), Surface($r=-.450$, $p<.01$, 95% IC [-.553,-.334]), Strategic($r=.320$, $p<.01$, 95% IC [.192,.438]). Sin embargo, al controlar el efecto de la dimensión Surface, la contribución de la faceta Strategic al rendimiento se redujo substancialmente, $r=.140$, $p<.01$, 95% IC [.003,.272]. El CFA confirmó la estructura de 3 factores del instrumento, $\chi^2(45,N=206)=83.12$, $p<.05$; $\chi^2/g.l.=1.85$, $CFI=.95$, $RMSEA=.06$.

Speth, Namuth, y Lee (2007) aplicaron el ASSIST para analizar su consistencia interna y validez estructural($N=446$). Ninguna dimensión tuvo una consistencia aceptable: Deep ($\alpha=.65$), Surface($\alpha=.70$) y Strategic($\alpha=.75$). La validez estructural se intentó establecer mediante un PCA.

Valadas, Gonçalves, y Faisca (2010) administraron el ASSIST para evaluar su consistencia interna($N=566$). Dos de las tres dimensiones presentaron una consistencia aceptable: Deep($\alpha=.81$), Surface($\alpha=.86$) y Strategic($\alpha=.79$).

La tabla 7 resume los resultados presentados previamente.

Tabla 7. *Resumen de resultados de fiabilidad y validez de los inventarios de habilidades y estrategias de aprendizaje y estudio*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Cano (2006)	NO	SE	SR	SÍ
2. Geertshuis y Fazey (2006)	NO	SE	SE	SE
3. Diseth et al. (2006)	SE	SE	SR	NO
4. Diseth (2007)	NO	SE	SE	NO
5. Speth, Namuth, y Lee (2007)	NO	SE	SE	SE

6. Valadas, Gonçalves, y Fáisca (2010)	NO	SE	SE	SE
--	----	----	----	----

Nota. SE (Sin Evidencia); SR (Sin Resolver).

El inventario de estilos de pensamiento o razonamiento de Sternberg (TSI)

Liu, Magjuka, y Lee (2008) midieron la validez predictiva del TSI(N=208). Los resultados indican que no hay relación entre estilos y rendimiento académico, $F(2,206)=.447$, $p=.640$, $\omega_p^2=-.005$, $r=-.071$, 95% IC [-.205,.066].

Black y McCoach (2008) analizaron la consistencia interna y la validez factorial del TSI (N=798). De las 12 subescalas que tiene el instrumento, 6 de ellas no alcanzaron un grado de consistencia adecuado: Legislative($\alpha=.81$), Excecutive($\alpha=.80$), Judicial($\alpha=.73$), Monarchic($\alpha=.64$), Hierarchic($\alpha=.80$), Anarchic($\alpha=.67$), Local($\alpha=.67$), Global($\alpha=.71$), Internal($\alpha=.79$), External($\alpha=.82$) Liberal($\alpha=.84$), Conservative($\alpha=.81$). El CFA no pudo validar el modelo de 2 factores propuesto por los creadores del instrumento, $\chi^2(26,N=798)=653.29$; $p<.001$, $\chi^2/g.l.=25.12$, $CFI=.779$; $RMSEA=.175$.

Fan, Zhang, y Watkins (2010) administraron el TSI(N=72), para examinar su validez predictiva. Se aplicaron 11 pruebas distintas de rendimiento académico. Ninguna de las dimensiones fue un predictor del rendimiento, mostrando correlaciones por debajo del mínimo-requerido (ver tabla 8). Sin embargo, al controlar los aspectos de la personalidad y la motivación, los estilos Anarchic, Executive, Judicial y Monarchic, predijeron 6 de las 11 pruebas de rendimiento, con correlaciones que fluctuaron entre $r=.316$, $p<.05$, 95% IC [.091,.510] y $r=.566$, $p<.001$, 95% IC [.385, .705].

Tabla 8. *Intervalo de correlaciones entre estilos de pensamiento y rendimiento académico*

Estilo	Correlación más baja	Correlación más alta
Legislative	$r=-.260$ 95% IC [-.464,-.030]	$r=.020$ 95% IC [-.213,.251]
Executive	$r=-.340$ 95% IC [-.530,-.118]	$r=.010$ 95% IC [-.222,.241]
Judicial	$r=-.310$ 95% IC [-.505,-.084]	$r=.270$ 95% IC [.041,.472]
Monarchic	$r=-.410$ 95% IC [-.586,-.197]	$r=.030$ 95% IC [-.203,.260]
Hierarchic	$r=-.250$ 95% IC [-.455,-.019]	$r=.080$ 95% IC [-.155,.306]
Anarchic	$r=-.240$ 95% IC [-.447,-.009]	$r=.310$ 95% IC [.084, .505]
Local	$r=-.190$ 95% IC [-.404,.044]	$r=.170$ 95% IC [-.064,.386]
Global	$r=-.210$ 95% IC [-.421,.023]	$r=.210$ 95% IC [-.023,.421]
Internal	$r=-.250$ 95% IC [-.455,-.019]	$r=.280$ 95% IC [.052, .480]
External	$r=-.290$ 95% IC [-.489,-.063]	$r=.200$ 95% IC [-.033,.413]
Liberal	$r=-.180$ 95% IC [-.395, .054]	$r=.260$ 95% IC [.030, .464]
Conservative	$r=-.300$ 95% IC [-.497,-.073]	$r=.080$ 95% IC [-.155,.306]
Oligarchic	$r=-.310$ 95% IC [-.505,-.084]	$r=.130$ 95% IC [-.105,.351]

Clarke, Lesh, Trocchio, y Wolman (2010) aplicaron el TSI($N=95$) y analizaron su consistencia interna, así como la validez convergente del instrumento, con respecto al ILS de Felder. Aunque no se reportan coeficientes de Cronbach para todas las escalas, ninguna de las mencionadas presentó una consistencia aceptable: Judicial($\alpha=.57$), Liberal($\alpha=.57$), Executive($\alpha=.79$), Local($\alpha=.79$), Global($\alpha=.63$), Legislative($\alpha=.64$), Conservative($\alpha=.69$). Sólo 3 de los estilos del TSI (i.e., Executive, Conservative y Local), tuvieron correlaciones positivas significativas, no triviales, con las dimensiones del ILS. Sin embargo, las asociaciones no alcanzaron el valor mínimo deseable, oscilando entre $r=.408$, $p<.001$, 95% IC [.215,.571] y $r=.444$, $p<.001$, 95% IC [.251,.603].

La tabla 9 resume los resultados presentados anteriormente.

Tabla 9. *Resumen de resultados de fiabilidad y validez del TSI*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Liu, Magjuka, y Lee (2008)	SE	SE	SE	NO
2. Black y McCoach (2008)	NO	SE	SR	SE
3. Fan, Zhang, y Watkins (2010)	SE	SE	SE	SÍ
4. Clarke et al. (2010)	NO	SE	NO	NO

Nota. SE (Sin Evidencia); SR (Sin Resolver).

El cuestionario de procesos de estudio de Biggs (SPQ)

Furnham, Christopher, Garwood, y Martin (2007) aplicaron el SPQ (Study Process Questionnaire) para examinar su validez predictiva($N=430$). Se encontraron sólo asociaciones nulas o triviales con el nivel de cultura general de los alumnos, que fluctuaron entre $r=-.270$, 95% IC [-.355,-.180] y $r=.150$ 95% IC [.056,.241]. Los resultados de la regresión logística confirman estos resultados, $F(3,373)=9.57$, $\omega_p^2=.064$, $r=.252$, 95% IC [.161,.338]. Sin embargo, cuando se tomaron en cuenta las facetas de motivo y estrategia de estos estilos, el análisis mostró una relación positiva no trivial con el rendimiento, $F(6,367)=7.38$, $\omega_p^2=.093$, $r=.305$, 95% IC [.217,.388].

Chamorro-Premuzic y Furnham (2008) investigaron la consistencia interna y validez predictiva del SPQ(N=158). Ninguna de las tres dimensiones tuvo una consistencia adecuada: Surface($\alpha=.66$), Deep($\alpha=.68$), Achieving($\alpha=.70$). Únicamente el estilo Deep tuvo una asociación positiva significativa baja, con el coeficiente intelectual, $r=.320$, $p<.01$, 95% IC [.172, .453], y el resultado del examen final, $r=.330$, $p<.01$, 95% IC [.183,.462].

Chamorro-Premuzic y Furnham (2009) administraron el SPQ, para examinar su consistencia interna(N=852). Ninguna de las tres dimensiones tuvo una consistencia aceptable: Surface-Motive($\alpha=.65$), Deep-Motive($\alpha=.66$), Achieving-Motive($\alpha=.68$), Surface-Strategy($\alpha=.65$), Deep-Strategy($\alpha=.75$), Achieving-Strategy($\alpha=.76$).

Furnham, Monsen, y Ahmetoglu (2009) estudiaron la validez predictiva del SQP (N=212). Los estilos tuvieron relaciones nulas con el nivel de cultura general de los alumnos, con correlaciones que fluctuaron entre $r=-.110$, 95% IC [-.241,.025] y $r=.130$, IC [-.005,.260]. Los mismos resultados se observaron con respecto a las asignaturas de lengua y matemáticas con asociaciones que oscilaron entre, $r=-.020$, 95% IC [-.115,.154] y $r=.180$, 95% IC [.046,.307].

La tabla 10 resume los resultados presentados previamente.

Tabla 10. *Resumen de resultados de fiabilidad y validez del SPQ*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Furnham et al. (2007)	SE	SE	SE	SÍ
2. Chamorro-Premuzic y Furnham (2008)	NO	SE	SE	NO
3. Chamorro-Premuzic y Furnham (2009)	NO	SE	SE	SE
4. Furnham, Monsen, Ahmetoglu (2009)	SE	SE	SE	NO

Nota. SE (Sin Evidencia); SR (Sin Resolver).

El cuestionario de estilos de aprendizaje de Honey y Mumford (LSQ)

Rassool y Rawaf (2008) probaron la validez predictiva del LSQ($N=110$). Los resultados indican una relación nula entre estilos y la adquisición de conocimiento de los estudiantes, $F(2,106)=2.645$, $p=.076$, $\omega_p^2=.029$, $r=.170$, 95% IC [-.018,.346].

Kappe, Boekholt, Den Rooyen, y Van der Flier (2009) examinaron la fiabilidad test-retest y la validez predictiva del LSQ($N=99$). Sólo una de las dimensiones presentó una fiabilidad aceptable: Activists($r=.70$), Reflectors($r=.63$), Theorists($r=.50$), Pragmatists($r=.46$). Se aplicaron 4 medidas diferentes para el rendimiento académico. Todas las asociaciones entre estilos y rendimiento fueron nulas con correlaciones que fluctuaron entre $r=-.070$, $p>.05$, 95% IC [-.264,.129] y $r=.150$, $p>.05$, 95% IC [-.049,.337].

Jackson, Hobman, Jimmieson, y Martin. (2009) analizaron la validez predictiva del LSQ($N=137$). Los estilos sólo tuvieron correlaciones triviales o nulas con el rendimiento en el trabajo, presentando asociaciones que oscilaron entre $r=-.020$, $p>.05$, 95% IC [-.187,.148] y $r=.140$, $p>.05$, 95% IC [-.028,.301]. La regresión múltiple confirmó estos resultados, $R^2=.030$, $r=.173$ 95% IC [.005,.331].

Sharif, Gifford, Morris, y Barber (2010) administraron el LSQ para investigar su validez predictiva($N=256$). Las asociaciones entre estilos y rendimiento fueron nulas o insignificantes, oscilando entre $r=-.200$, $p<.01$, 95% IC [-.315,-.079] y $r=.200$, $p<.01$, 95% IC [.079,.315].

La tabla 11 resume los resultados presentados anteriormente.

Tabla 11. *Resumen de resultados de fiabilidad y validez del LSQ*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Rassool y Rawaf (2008)	SE	SE	SE	NO
2. Kappe et al. (2009)	SE	NO	SE	NO
3. Jackson et al. (2009)	SE	SE	SE	NO
4. Sharif et al. (2010)	SE	SE	SE	NO

Nota. SE (Sin Evidencia); SR (Sin Resolver).

El delineador de estilos de Gregorc (GSD)

Miller (2005) aplicó el GSD a dos grupos diferentes de participantes ($N=36$ y $N=34$) para evaluar su validez predictiva, así como la convergente con respecto al LSI de Kolb. En el primer grupo, sólo la dimensión Concrete-Random(CR), tuvo una asociación positiva no trivial con la evaluación final, $r=.340$, $p<.05$, 95% IC [.013, 0.601]. Las correlaciones entre el GSD y el LSI, para este mismo grupo fueron nulas, fluctuando entre $r=-.300$, $p>.05$, 95% IC [-.572, 0.032] y $r=.320$, $p>.05$, 95% IC [-.010,.587]. En el segundo grupo, de nuevo la dimensión CR tuvo una relación positiva tanto con la cantidad de material aprendido, $r=.420$, $p<.05$, 95% IC [.095,.664] como con la evaluación final, $r=.390$, $p<.05$, 95% IC [.060,.643]. Sólo las dimensiones Concrete-Sequential(CS) y CR del GSD tuvieron relaciones positivas con las facetas CE y RO del LSI. Las correlaciones fueron de, $r=.370$, $p<.05$, 95% IC [.036,.629] entre ellas.

Ruhnau (2006) administró el GSD usando 3 muestras distintas ($N=14$, $N=19$ y $N=9$) para analizar su validez predictiva. Todas las asociaciones estilos-rendimiento académico fueron nulas, tanto para el primer, $F(3,10)=2.280$, $p=.142$, $\omega_p^2=.215$, $r=.464$, 95% IC [-.088,.798], segundo, $F(3,15)=.499$, $p=.689$, $\omega_p^2=-.086$, $r=-.293$ 95% IC [-.659,.186] y tercer, $F(3,15)=.499$, $p=.689$, $\omega_p^2=-.086$, $r=-.293$, 95% IC [-.659,.186], grupo.

Reio y Wiswell (2006) examinaron la consistencia interna y validez estructural del GSD($N=467$). El instrumento no mostró una consistencia aceptable: CS($\alpha=.51$), CR($\alpha=.64$), AR($\alpha=.23$), AS($\alpha=.66$). El CFA no pudo confirmar la estructura de 4 factores propuesta por Gregorc, $\chi^2(51,N=467)=407$, $\chi^2/g.l.=7.99$, $CFI=.51$, $RMSEA=.149$. Adicionalmente, También se probó una estructura bipolar de 2 factores y tampoco se pudo verificar, $\chi^2(49,N=467)=538$, $\chi^2/g.l.=10.98$, $CFI=.48$, $RMSEA=.191$.

La tabla 12 resume los resultados presentados previamente.

Tabla 12. *Resumen de resultados de fiabilidad y validez del GSD*

Estudio	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez pre-dictiva
1. Miller (2005)	SE	SE	SR	NO
2. Ruhnau (2006)	SE	SE	SE	NO
3. Reio y Wiswell (2006)	NO	SE	SR	NO

Nota. SE (Sin Evidencia); SR (Sin Resolver).

Discusión y/o Conclusiones

Tres conclusiones principales se pueden hacer, a partir de la evidencia presentada previamente. La primera, es que la investigación, sobre la fiabilidad y validez de los inventarios de estilos de aprendizaje, está muy dispersa. La mayoría de los estudios abordan, cuando mucho, 2 de los 5 criterios necesarios para determinar si un instrumento es robusto (i.e., consistencia interna, fiabilidad test-retest, validez predictiva, validez convergente y validez factorial). No hay un enfoque sistemático, por parte de los investigadores, para evaluar estos aspectos. Además, se encontró que alrededor del 55% de las investigaciones que pretendían determinar la validez factorial de algún inventario, utilizaron un PCA para ello.

La segunda, es que los 2 aspectos o criterios más estudiados, de los inventarios analizados en este artículo, son la validez predictiva y la consistencia interna. En este sentido, alrededor del 57% de los resultados analizados indican que ninguno de los instrumentos tiene validez predictiva. Mientras que cerca de un 43% señala que carecen de consistencia interna. Sin embargo, este último dato hay interpretarlo con precaución, pues algunos instrumentos no cumplieron con este criterio, debido a que una de sus escalas no alcanzó, el nivel mínimo requerido de consistencia.

La tercera, es que los 2 indicadores menos estudiados fueron el de fiabilidad test-retest y la validez de constructo. El 91% de los estudios revisados no estudiaron el primer aspecto y un 9% indica que los instrumentos carecen de este tipo de fiabilidad. Cerca del 77% de las investigaciones no analizó el segundo criterio y un 21% no pudieron resolverlo pues sólo examinaron la validez convergente, o bien, la factorial. Por ejemplo, Jamieson (2010) estudió la validez convergente del LSI de Kolb con otros 2 inventarios similares, pero no abordó la validez factorial. De la misma manera, Reio y Wiswell (2006) analizaron la validez estructural del GSD de Gregorc, pero no investigaron la validez convergente.

Para propósitos de comparación, las tablas 13 y 14 muestran la matriz de Coffield et al. (2004) y la nueva matriz actualizada y extendida, a partir de los resultados presentados aquí. Sin embargo, hay que precisar que la etiqueta SE de la matriz de Coffield et al. (tabla 11), incluye tanto a la ausencia de evidencia, así como a evidencia parcial. Mientras que en este artículo se decidió añadir una categoría más, SR, que indica que hay evidencia a favor o en contra, pero que el asunto no ha sido resuelto aún. Las etiquetas con asterisco son las que sufrieron modificaciones de una matriz a otra, a partir del análisis realizado.

Tabla 13. *Matriz de fiabilidad y validez para inventarios de estilos de aprendizaje de Coffield et al.*

Instrumento	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Jackson	SE	SE	SE	SE
2. Riding	NO	NO	NO	NO
3. Sternberg	NO	NO	NO	NO
4. Dunn y Dunn	NO	NO	NO	SÍ
5. Gregorc	NO	NO	NO	SÍ*
6. Honey y Mumford	NO	SÍ	NO	NO
7. Kolb	SE*	SÍ	NO	NO
8. Entwistle	SÍ*	SE	SÍ	NO
9. Herrmann	SE	SÍ	SÍ	SE
10. Myer-Briggs	SÍ	SÍ	SÍ	SÍ
11. Apter	SÍ	SÍ	SE	SÍ
12. Vermunt	SÍ	SÍ	SÍ	SE
13. Allison y Hayes	SÍ	SÍ	SÍ	SÍ

Nota. SE (Sin Evidencia).

Tabla 14. *Nueva matriz de fiabilidad y validez para inventarios de estilos de aprendizaje*

Instrumento	Consistencia interna	Fiabilidad test-retest	Validez de constructo	Validez predictiva
1. Jackson	SE	SE	SE	SE
2. Riding	NO	NO	NO	NO
3. Sternberg	NO	NO	NO	NO
4. Dunn y Dunn	NO	NO	NO	SÍ
5. Gregorc	NO	NO	NO	NO*
6. Honey y Mumford	NO	SÍ	NO	NO
7. Kolb	NO*	SÍ	NO	NO
8. Entwistle	NO*	SE	SÍ	NO
9. Herrmann	SE	SÍ	SÍ	SE
10. Myer-Briggs	SÍ	SÍ	SÍ	SÍ
11. Apter	SÍ	SÍ	SE	SÍ
12. Vermunt	SÍ	SÍ	SÍ	SE
13. Allison y Hayes	SÍ	SÍ	SÍ	SÍ

14. Felder	NO	NO	SR	NO
15. VAK	NO	SR	SR	NO
16. Biggs	NO	SE	SR	NO

Nota. SE (Sin Evidencia); SR (Sin Resolver).

Como se puede apreciar, la fiabilidad y validez de los inventarios ha cambiado muy poco. De dieciséis instrumentos, diez (63%) carecen de consistencia interna y no se pudo recuperar evidencia para dos (13%). Cinco instrumentos (31%) carecen de fiabilidad test-retest, no hay evidencia para cuatro (25%) y en uno (6%) no está resuelta. Seis inventarios (38%) no tienen validez de constructo, dos carecen de evidencia (13%) al respecto y 3 no han resultado la cuestión (19%). Nueve inventarios (56%) carecen de validez predictiva y tres (19%) no tiene evidencia que la soporte. Sólo el inventario de Allison y Hayes, así como el de Vermunt cumplen con (casi) todos los requisitos. Aunque se obtuvo información sobre estos instrumentos, sólo se recuperó un artículo respectivamente, y no se incluyeron en el análisis.

En general, podemos concluir que los inventarios de estilos de aprendizaje están poco estudiados, y su asociación con las medidas de rendimiento es mayormente nula o trivial. Psicométricamente, la mayoría cojea. Dependiendo del instrumento, son robustos en algunos aspectos, como la consistencia interna, y muy débiles en otros, como la validez predictiva. Consecuentemente, la recomendación original de Coeffield et al. (2004), sobre los inventarios sigue siendo válida: “No aconsejamos intervenciones pedagógica basadas únicamente en los resultados de cualquiera de los instrumentos de estilos de aprendizaje.” (p. 140). Sin embargo, es importante señalar que esta investigación tiene tres limitaciones. Primero, la búsqueda solo incluyó el periodo 2005-2010. Segundo, sólo se revisaron los primeros 160 resultados de GA. Tercero, no se hicieron búsquedas extras, ad-hoc, para inventarios que no aparecieron en la búsqueda inicial.

Finalmente, se puede argumentar que los criterios utilizados en esta investigación, para determinar la fiabilidad y la validez de los inventarios, son muy estrictos. En este sentido, dado que las promesas de adoptar el constructo varían desde ganancias en el rendimiento académico hasta el desarrollo del respeto por uno mismo y por los demás (Dembo y Howard, 2007, Scott, 2010), es útil recordar que las afirmaciones extraordinarias requieren evidencia extraordinaria, basada en los estándares más rigurosos (Voss, Helgen y Jan-sa, 2014, Wagenmakers, Wetzels, Borsboom y van der Maas, 2011). De lo contrario, es fácil confundir las esperanzas y los hechos, y deslizarse hacia la pseudociencia (Sagan, 1997).

Referencias

- *Akkoyunlu, B., & Yilmaz-Soylu, M. (2008). A study of student's perceptions in a blended learning environment based on different learning styles. *Educational Technology & Society*, 11(1), 183-193.
- *Alaoutinen, S. (2010). Effects of learning style and student background on self-assessment and course performance. In *Proceedings of the 10th Koli Calling International Conference on Computing Education Research* (pp. 5-12). ACM. <https://doi.org/10.1145/1930464.1930465>
- *An, G. J., & Yoo, M. S. (2008). Critical thinking and learning styles of nursing students at the baccalaureate nursing program in Korea. *Contemporary Nurse*, 29(1), 100-109. <https://doi.org/10.5172/conu.673.29.1.100>
- *Andreou, E., Andreou, G., & Vlachos, P. (2008). Learning styles and performance in second language tasks. *TESOL Quarterly*, 42(4), 665-674. <https://doi.org/10.1002/j.1545-7249.2008.tb00156.x>
- *Azevedo, R. E., & Akdere, M. (2010). Measuring the effects of learning styles: Is a little knowledge dangerous for excellence in management education? *Journal of Leadership & Organizational Studies*, 17(2), 192-200. <https://doi.org/10.1177/1548051810366713>
- Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the academy of marketing science*, 40(1), 8-34. <https://doi.org/10.1007/s11747-011-0278-x>
- *Black, A. C., & McCoach, D. B. (2008). Validity study of the thinking styles inventory. *Journal for the Education of the Gifted*, 32(2), 180-210. <https://doi.org/10.4219/jeg-2008-849>
- *Brittan-Powell, C., Legum, H., & Taylor, E. (2008). The relationship between student learning style, selection of course delivery format, and academic performance. *International Journal of Instructional Technology and Distance Learning*, 5(5), 41-46.
- *Cagiltay, N. E. (2008). Using learning styles theory in engineering education. *European journal of engineering education*, 33(4), 415-424. <https://doi.org/10.1080/03043790802253541>
- *Cano, F. (2006). An in-depth analysis of the Learning and Study Strategies Inventory (LASSI). *Educational and Psychological Measurement*, 66(6), 1023-1038. <https://doi.org/10.1177/0013164406288167>

- Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, 15(1), 17-32. <https://doi.org/10.1177/1094428110392383>
- *Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and individual differences*, 44(7), 1596-1603. <https://doi.org/10.1016/j.paid.2008.01.003>
- *Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, 19(4), 524-529. <https://doi.org/10.1016/j.lindif.2009.06.004>
- *Chapman, D. M., & Calhoun, J. G. (2006). Validation of learning style measures: implications for medical education practice. *Medical education*, 40(6), 576-583. <https://doi.org/10.1111/j.1365-2929.2006.02476.x>
- *Choi, I., Lee, S. J., & Kang, J. (2009). Implementing a case-based e-learning environment in a lecture-oriented anaesthesiology class: Do learning styles matter in complex problem solving over time? *British Journal of Educational Technology*, 40(5), 933-947. <https://doi.org/10.1111/j.1467-8535.2008.00884.x>
- *Çirkinoglu, A. G., & Demirci, N. (2007). Correlation Between University Students' Kinematic Achievement and Learning Styles. In *Sixth International Conference of the Balkan Physical Union* (Vol. 899, pp. 485-486). <https://doi.org/10.1063/1.2733248>
- *Clarke, T. A., Lesh, J. J., Trocchio, J. S., & Wolman, C. (2010). Thinking styles: teaching and learning styles in graduate education students. *Educational Psychology*, 30(7), 837-848. <https://doi.org/10.1080/01443410.2010.510794>
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: A systematic and critical review*. London: Learning and Skills Research Centre.
- *Cook, D. A. (2005). Reliability and validity of scores from the index of learning styles. *Academic medicine*, 80(10), S97-S101.
- *Cook, D. A., & Smith, A. J. (2006). Validity of index of learning styles scores: multitrait-multimethod comparison with three cognitive/learning style instruments. *Medical Education*, 40(9), 900-907. <https://doi.org/10.1111/j.1365-2929.2006.02542.x>
- *Cooper Brathwaite, A. (2006). Influence of nurse characteristics on the acquisition of cultural competence. *International Journal of Nursing Education Scholarship*, 3(1). <https://doi.org/10.2202/1548-923X.1173>

- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research & evaluation, 10*(7), 1-9.
- Cuevas, J. (2015). Is learning styles-based instruction effective? A comprehensive analysis of recent research on learning styles. *Theory and Research in Education, 13*(3), 308-333. <https://doi.org/10.1177/1477878515606621>
- Dembo, M. H., & Howard, K. (2007). Advice about the use of learning styles: A major myth in education. *Journal of college reading and learning, 37*(2), 101-109. <https://doi.org/10.1080/10790195.2007.10850200>
- *Demirbas, O. O., & Demirkan, H. (2007). Learning styles of design students and the relationship of academic performance and gender in design education. *Learning and instruction, 17*(3), 345-359. <https://doi.org/10.1016/j.learninstruc.2007.02.007>
- *Diseth, Å. (2007). Approaches to learning, course experience and examination grade among undergraduate psychology students: Testing of mediator effects and construct validity. *Studies in Higher Education, 32*(3), 373-388. <https://doi.org/10.1080/03075070701346949>
- *Diseth, Å., Pallesen, S., Hovland, A., & Larsen, S. (2006). Course experience approaches to learning and academic achievement. *Education+ Training, 48*(2/3), 156-169. <https://doi.org/10.1108/00400910610651782>
- *Dobson, J. L. (2009). Learning style preferences and course performance in an undergraduate physiology class. *Advances in Physiology Education, 33*(4), 308-314. <https://doi.org/10.1152/advan.00048.2009>
- *Erdem, M. (2009). Effects of learning style profile of team on quality of materials developed in collaborative learning processes. *Active Learning in Higher Education, 10*(2), 154-171. <https://doi.org/10.1177/1469787409104902>
- *Fan, W., Zhang, L. F., & Watkins, D. (2010). Incremental validity of thinking styles in predicting academic achievements: An experimental study in hypermedia learning environments. *Educational Psychology, 30*(5), 605-623. <https://doi.org/10.1080/01443410.2010.496899>
- *Furnham, A., Christopher, A. N., Garwood, J., & Martin, G. N. (2007). Approaches to learning and the acquisition of general knowledge. *Personality and Individual Differences, 43*(6), 1563-1571. <https://doi.org/10.1016/j.paid.2007.04.013>
- *Furnham, A., Monsen, J., & Ahmetoglu, G. (2009). Typical intellectual engagement, Big Five personality traits, approaches to learning and cognitive ability predictors of aca-

- demic performance. *British Journal of Educational Psychology*, 79(4), 769-782. <https://doi.org/10.1348/978185409X412147>
- *Galvan, M. T. (2006). Can students' learning style preferences explain their academic performance in marketing courses? *Proceedings of the Marketing Management Association*, 117-124.
- *Geertshuis, S. A., & Fazey, J. A. (2006). Approaches to learning in the workplace. *Journal of Workplace Learning*, 18(1), 55-65. <https://doi.org/10.1108/13665620610641319>
- Gottfredson, L. S. (2009). Logical fallacies used to dismiss the evidence on intelligence testing. In R. P. Phelps (Ed.), *Correcting fallacies about educational and psychological testing* (pp. 11-65). Washington, DC, US: American Psychological Association. <http://doi.org/10.1037/11861-001>
- Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PloS one*, 10(9), e0138237. <https://doi.org/10.1371/journal.pone.0138237>
- Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, 106(2), 787-804. <https://doi.org/10.1007/s11192-015-1798-9>
- Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and evaluation in counseling and development*, 34(3), 177.
- *Hosford, C. C., & Siders, W. A. (2010). Felder-Soloman's index of learning styles: Internal consistency, temporal stability, and factor structure. *Teaching and learning in medicine*, 22(4), 298-303. <https://doi.org/10.1080/10401334.2010.512832>
- Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20(1), 90-98. <https://doi.org/10.1016/j.jcps.2009.09.003>
- *Isemonger, I., & Sheppard, C. (2007). A construct-related validity study on a Korean version of the perceptual learning styles preference questionnaire. *Educational and Psychological Measurement*, 67(2), 357-368. <https://doi.org/10.1177/0013164406292073>
- *Jackson, C. J., Hobman, E. V., Jimmieson, N. L., & Martin, R. (2009). Comparing different approach and avoidance models of learning and personality in the prediction of work, university, and leadership outcomes. *British Journal of Psychology*, 100(2), 283-312. <https://doi.org/10.1348/000712608X322900>

- *Jamieson, G. E. (2010). *Investigating the Kolb Learning Style Inventory's ipsative scores using semantic differential and Likert scaling* (Doctoral dissertation, Northern Michigan University).
- *Kappe, F. R., Boekholt, L., Den Rooyen, C., & Van der Flier, H. (2009). A predictive validity study of the Learning Style Questionnaire (LSQ) using multiple, specific learning criteria. *Learning and Individual Differences*, 19(4), 464-467. <https://doi.org/10.1016/j.lindif.2009.04.001>
- *Kayes, D. C. (2005). Internal validity and reliability of Kolb's learning style inventory version 3 (1999). *Journal of Business and Psychology*, 20(2), 249-257. <https://doi.org/10.1007/s10869-005-8262-4>
- *Ku, D. T., & Shen, C. Y. (2009). Reliability, validity, and investigation of the index of learning styles in a Chinese language version for late adolescents of Taiwanese. *Adolescence*, 44(176), 827.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC psychology*, 4(1), 24. <https://doi.org/10.1186/s40359-016-0126-3>
- *Leite, W. L., Svinicki, M., & Shi, Y. (2010). Attempted validation of the scores of the VARK: Learning styles inventory with multitrait-multimethod confirmatory factor analysis models. *Educational and psychological measurement*, 70(2), 323-339. <https://doi.org/10.1177/0013164409344507>
- Lezak, M. D., Howieson, D., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological assessment*. 5th edition. Oxford University Press. *Oxford, New York, ISBN, 10, 9780195395525*.
- Lilienfeld, S.O., Lynn, S.J., Ruscio, J., & Beyerstein B.L. (2010). *50 great myths of popular psychology—Shattering widespread misconceptions about human behavior*. West Sussex, UK: Wiley-Blackwell.
- *Liu, X., Magjuka, R. J., & Lee, S. H. (2008). The effects of cognitive thinking styles, trust, conflict management on online students' learning and virtual team performance. *British Journal of Educational Technology*, 39(5), 829-846. <https://doi.org/10.1111/j.1467-8535.2007.00775.x>

- *Lu, H., Jia, L., Gong, S. H., & Clark, B. (2007). The relationship of Kolb learning styles, online learning behaviors and learning outcomes. *Journal of Educational Technology & Society*, 10(4).
- *McChlery, S., & Visser, S. (2009). A comparative analysis of the learning styles of accounting students in the United Kingdom and South Africa. *Research in Post-Compulsory Education*, 14(3), 299-315. <https://doi.org/10.1080/13596740903139404>
- *Metallidou, P., & Platsidou, M. (2008). Kolb's Learning Style Inventory-1985: Validity issues and relations with metacognitive knowledge about problem-solving strategies. *Learning and Individual Differences*, 18(1), 114-119. <https://doi.org/10.1016/j.lindif.2007.11.001>
- *Miller, L. M. (2005). Using learning styles to evaluate computer-based instruction. *Computers in human behavior*, 21(2), 287-306. <https://doi.org/10.1016/j.chb.2004.02.011>
- Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69-71.
- Nirenburg, S., McShane, M., & Beale, S. (2004). The rationale for building resources expressly for NLP. In *Proc. of the 4th International Conference on Language Resources and Evaluation* (Vol. 218).
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological science in the public interest*, 9(3), 105-119. <https://doi.org/10.1111/j.1539-6053.2009.01038.x>
- *Peters, D., Jones, G., & Peters, J. (2008). Preferred 'learning styles' in students studying sports-related programmes in higher education in the United Kingdom. *Studies in Higher Education*, 33(2), 155-166. <https://doi.org/10.1080/03075070801916005>
- *Platsidou, M., & Metallidou, P. (2009). Validity and Reliability Issues of Two Learning Style Inventories in a Greek Sample: Kolb's Learning Style Inventory and Felder & Soloman's Index of Learning Styles. *International Journal of Teaching and Learning in Higher Education*, 20(3), 324-335.
- Rao, T. S., & Andrade, C. (2011). The MMR vaccine and autism: Sensation, refutation, retraction, and fraud. *Indian journal of psychiatry*, 53(2), 95. <https://doi.org/10.4103/0019-5545.82529>
- *Rassool, G. H., & Rawaf, S. (2008). The influence of learning styles preference of undergraduate nursing students on educational outcomes in substance use education. *Nurse Education in Practice*, 8(5), 306-314. <https://doi.org/10.1016/j.nepr.2008.02.001>

- *Reio Jr, T. G., & Wiswell, A. K. (2006). An examination of the factor structure and construct validity of the Gregorc Style Delineator. *Educational and Psychological Measurement*, 66(3), 489-501. <https://doi.org/10.1177/0013164405282459>
- Rosenberg, M. S. (2010). A generalized formula for converting chi-square tests to effect sizes for meta-analysis. *PloS one*, 5(4), e10059. <https://doi.org/10.1371/journal.pone.0010059>
- *Ruhnau, K. (2006). An Analysis of Learning Outcomes of Adult Students: Learning Styles. *Statistics [NCES]*, 2002, 0-12.
- Sagan, C. (1997). *The demon-haunted world: Science as a candle in the dark*. Random House Digital, Inc.
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. <https://doi.org/10.1177/0734282911406653>
- Scott, C. (2010). The enduring appeal of 'learning styles'. *Australian Journal of Education*, 54(1), 5-17. <https://doi.org/10.1177/000494411005400102>
- *Self, B. P., & Widmann, J. (2009). Work in progress-learning styles and performance on the dynamics concept inventory. In *Frontiers in Education Conference, 2009. FIE'09. 39th IEEE* (pp. 1-2). IEEE. <https://doi.org/10.1109/FIE.2009.5350822>
- *Sharif, S., Gifford, L. A., Morris, G. A., & Barber, J. (2010). The relationship between learning styles, attendance and academic performance of Pharmacy undergraduates. *Pharmacy Education*, 10.
- *Slack, N., & Norwich, B. (2007). Evaluating the reliability and validity of a learning styles inventory: A classroom-based study. *Educational Research*, 49(1), 51-63. <https://doi.org/10.1080/00131880701200765>
- *Speth, C. A., Namuth, D. M., & Lee, D. J. (2007). Using the ASSIST Short Form for Evaluating an Information Technology Application: Validity and Reliability Issues. *Informing Science*, 10.
- *Strang, K. D. (2008). Quantitative online student profiling to forecast academic outcome from learning styles using dendrogram decision models. *Multicultural Education & Technology Journal*, 2(4), 215-242. <https://doi.org/10.1108/17504970810911043>
- *Sun, K. T., Lin, Y. C., & Yu, C. J. (2008). A study on learning effect among different learning styles in a Web-based lab of science for elementary school students. *Computers & Education*, 50(4), 1411-1422. <https://doi.org/10.1016/j.compedu.2007.01.003>

- *Valadas, S. C. A., Gonçalves, F. R., & Faísca, L. M. (2010). Approaches to studying in higher education Portuguese students: a Portuguese version of the approaches and study skills inventory for students. *Higher Education*, 59(3), 259-275. <https://doi.org/10.1007/s10734-009-9246-5>
- Voss, R. S., Helgen, K. M., & Jansa, S. A. (2014). Extraordinary claims require extraordinary evidence: a comment on Cozzuol et al. (2013). *Journal of Mammalogy*, 95(4), 893-898. <https://doi.org/10.1644/14-MAMM-A-054>
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & Van Der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). <https://doi.org/10.1037/a0022790>
- *Wang, K. H., Wang, T. H., Wang, W. L., & Huang, S. C. (2006). Learning styles and formative assessment strategy: enhancing student achievement in Web-based learning. *Journal of Computer Assisted Learning*, 22(3), 207-217. <https://doi.org/10.1111/j.1365-2729.2006.00166.x>
- *Yildirim, O., Acar, A. C., Bull, S., & Sevinc, L. (2008). Relationships between teachers' perceived leadership style, students' learning style, and academic achievement: A study on high school students. *Educational Psychology*, 28(1), 73-81. <https://doi.org/10.1080/01443410701417945>
- *Yilmaz-Soylu, M., & Akkoyunlu, B. (2009). The effect of learning styles on achievement in different learning environments. *TOJET: The Turkish Online Journal of Educational Technology*, 8(4).
- *Yukselturk, E., & Bulut, S. (2007). Predictors for student success in an online course. *Journal of Educational Technology & Society*, 10(2).
- *Zaharias, P., Andreou, I., & Vosinakis, S. (2010). Educational virtual worlds, learning styles and learning effectiveness: An empirical investigation. In *Proceedings of the 7th Pan-Hellenic Conference with International Participation ICT in Education, Korinthos, Greece* (pp. 1-6).