

Sobre el reporte de confiabilidad del CLARP-TDAH, de Salamanca (2010)

SERGIO DOMÍNGUEZ-LARA

Universidad Inca Garcilaso de la Vega, Perú.

CÉSAR MERINO-SOTO

Universidad de San Martín de Porres, Perú.

Sr. Editor,

El importante esfuerzo de Salamanca (2010) en la construcción del cuestionario CLARP-TDAH responde directamente a la creciente necesidad en el ámbito colombiano para tener métodos de evaluación culturalmente sensibles y adecuados a los estándares internacionales (concretamente el CIF). En conjunto, todo esto permitiría valorar aspectos vinculados con la discapacidad, la cual se entiende actualmente más en función de la interacción con el contexto (limitaciones potenciales a la actividad de la persona) y menos como producto de una deficiencia propia de la estructura corporal (o mental) de la persona. Por lo tanto, contar con un instrumento con evidencias de validez y confiabilidad es importante para el psicólogo en su práctica profesional, ya que tendrá un mejor respaldo para hacer decisiones diagnósticas. Sin embargo, los resultados preliminares de Salamanca (2010) no parecen favorables, y ese es el punto que abortamos en esta carta.

La autora reporta la confiabilidad de las puntuaciones por el método de consistencia interna, específicamente el coeficiente α . Brevemente, la confiabilidad expresada por este método está influenciada principalmente por el número de ítems y por la proporción de la varianza total del test debida a la covariación de los ítems; esto implica que cuanto más covaríen los ítems y más ítems contenga el instrumento, mayor será la confiabilidad del test. Además, el coeficiente que hace referencia, una vez que se transforma dicho número a un porcentaje, al monto de varianza verdadera que contiene el puntaje obtenido y, en consecuencia, cuánto error de medición contiene el mismo puntaje.

Los resultados obtenidos por Salamanca respecto a la consistencia interna son predominantemente inaceptables para aplicaciones clínicas, en incluso para propósitos de despistaje e investigación. Exceptuando el puntaje total ($\alpha = 0.74$) y de las sub-áreas Tareas y demandas generales ($\alpha = 0.7$) y Autocuidado ($\alpha = 0.81$), el resto muestra niveles de error de medición que difícilmente pueden ser considerados aceptables y adecuados para cualquier propósito profesional y científico: Aprendizaje y aplicación del conocimiento ($\alpha = 0.49$), Comunicación ($\alpha = 0.48$), Vida doméstica ($\alpha = 0.6$), Interacciones y relaciones interpersonales ($\alpha = 0.47$), y Áreas principales de la vida y vida comunitaria social y cívica ($\alpha = 0.45$). Es cierto que no puede afirmarse que hay un criterio unívoco para la valoración de dichos coeficientes, ya que esto depende de varios aspectos y sobre todo del uso que se le brindará al instrumento (investigación, apoyo al diagnóstico, etc.). Sin embargo, sin ánimo de ser exhaustivos, a lo largo de los años algunas propuestas metodológicas (Cicchetti, 1994, Kaplan & Sacuzzo, 2008, Murphy & Davidsholder, 1988, Merino-Soto, Navarro & García, 2014, Nunnally & Bernstein, 1995) han reconocido que los niveles aceptables de confiabilidad están entre las magnitudes de .60 y .70. Incluso, los puntajes con mejores resultados de confiabilidad no podrán ser considerados óptimos para su potencial generalización a la población de referencia, pues si se estiman sus intervalos de confianza con el método Fisher, como recomiendan Romano, Kromrey y Hibbard (2010) y Romano, Kromrey, Owens y Scott (2011), los posibles resultados aún son más bajos. Por ejemplo: el puntaje total (IC

95%: 0.56, 0.85) y Tareas y demandas generales (IC 95%: 0.50, 0.82) muestran que su intervalo inferior está debajo de 0.60.

Este re-análisis pone en duda la posible aplicación del instrumento en el ámbito aplicado, ya sea en decisiones clínicas o en apoyo al diagnóstico. Aunque la autora plantea un "... valor aceptable de la consistencia interna..." (p.1126), su afirmación en sí misma no puede considerarse aceptable en base a la evidencia previa. En una de las afirmaciones finales del manuscrito, se resalta que el instrumento aún necesita estudios adicionales para fortalecerlo, pero en el último párrafo se recomienda su uso por los profesionales implicados (médicos, psicólogos, etc.) en el tratamiento de dicho trastorno. Estas afirmaciones que parecen contradictorias pueden elevar más el riesgo de su uso inapropiado, pues el CLARP-TDAH aún esta fase experimental, presenta inestabilidad en la estimación de sus puntajes y puede llevar a clasificar de forma errónea a aquellas personas evaluadas por medio de este.

Lista de referencias

- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, pp. 284-290.
- Kaplan, R. & Saccuzzo, D. (2008). *Psychological testing: Principles, applications and issues (7a Edición)*. Belmont: Thomson/Wadsworth.
- Merino, C., Navarro, J. & García, W. (2014). Revisión de la consistencia interna del Inventario de Inteligencia Emocional de Bar-On, EQ-I: YV. *Revista Peruana de Psicología y Trabajo Social*, 3 (1), pp. 141-154.
- Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing: Principles and applications*. Englewood Cliffs, New Jersey: Prentice Hall.
- Romano, J. L., Kromrey, J. D. & Hibbard, S. T. (2010). A Monte Carlo study of eight confidence interval methods for coefficient alpha. *Educational and Psychological Measurement*, 70 (3), pp. 376-393.
- Romano, J. L., Kromrey, J. D., Owens, C. M. & Scott, H. M. (2011). Confidence interval methods for coefficient alpha on the basis of discrete, ordinal response items: Which one, if any, is the best? *The Journal of Experimental Education*, 79 (4), pp. 382-403.
- Salamanca, L. (2010). Construcción, validación y confiabilidad de un cuestionario sobre niños y niñas con TDAH. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 8 (2), pp. 1117-1129.