



**UNIVERSIDAD DE CIENCIAS  
EMPRESARIALES Y SOCIALES**  
[www.uces.edu.ar](http://www.uces.edu.ar)

**INSTITUTO DE ALTOS ESTUDIOS EN PSICOLOGÍA Y CIENCIAS  
SOCIALES (IAEPCIS) "David Maldavsky"**  
**Doctorado en Psicología**  
**Departamento de Investigaciones**

**Sábado 22 de julio de 2023**  
**XIX Jornadas Internacionales de Investigación en  
Psicología UCES 2023**  
**XXI Jornadas Internacionales de Actualización del  
Algoritmo David Liberman**

**Título:** Aplicación de los Estándares para Pruebas Psicológicas en el Ejercicio Físico.

**Autor/es:** Ursino, Damián Javier y Attorresi, Horacio Félix.

**E-mail:** [damian.ursino@gmail.com](mailto:damian.ursino@gmail.com)

### **Introducción**

El propósito de los Estándares para Pruebas Educativas y Psicológicas (American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education [NCME], 2014) es proporcionar un conjunto de criterios (estándares) para el desarrollo, evaluación y uso de las pruebas con el objetivo de evaluar la validez de las interpretaciones de sus puntajes en función del uso previsto. La validez es un concepto unitario y se refiere al grado en que la evidencia y teoría apoyan dichas interpretaciones comprometiendo una cuidadosa atención del contenido, formato y administración de la prueba. Su proceso implica la acumulación de evidencia relevante para proporcionar una base científica sólida para la prueba. Este proceso constituye la construcción y evaluación de argumentos a favor y en contra de la interpretación, uso y relevancia de la prueba (AERA et al., 2014). Las siguientes secciones describen las fuentes de evidencia de validez que la literatura ha reportado para evaluar las necesidades psicológicas básicas en el ejercicio físico. Dependiendo del contexto y el propósito de la prueba, algunas fuentes de validez serán más destacadas que otras. Para el desarrollo de este escrito y en función del

instrumento seleccionado para ejemplificar, se considerarán las pruebas basadas en evidencia de contenido, estructura interna y relación con otras variables.

### **Desarrollo**

La evidencia basada en el contenido de la prueba refiere a la redacción y formato de los ítems generados a partir de las definiciones conceptuales del constructo a medir con la prueba. Constituye principalmente evaluar si el contenido refleja adecuadamente los conceptos establecidos a partir de la concordancia entre jueces expertos. Aquí también se pone en juego las equivalencias lingüísticas e idiomáticas y aspectos de la validez aparente. Los análisis basados en evidencia de la estructura interna indican el grado en que los elementos de la prueba se ajustan al constructo en el que se basan sus interpretaciones. Se puede diferenciar entre la dimensionalidad, confiabilidad e invarianza en la medición. El estudio de la dimensionalidad determina si las interrelaciones entre los ítems apoyan las puntuaciones previstas para la prueba. Para ello el análisis factorial es el método estadístico más utilizado proporcionando pruebas del número de variables latentes subyacentes (factores) que constituyen al instrumento (unidimensional o multidimensional) (Rios Y Wells, 2014). A su vez, utilizando la información provista por los parámetros estimados del modelo (cargas factoriales y covarianzas de las variables latentes) se puede evaluar la validez convergente y discriminante de las subdimensiones implicadas. En tal caso, la validez convergente presenta evidencia cuando los ítems tienen una fuerte relación con la respectiva variable latente, mientras que la validez discriminante implica una asociación entre las distintas variables latentes de pequeña a moderada. Por su parte los estudios de invarianza en la medición proporcionan pruebas de que el instrumento a nivel de la escala (AFC de grupos múltiples) y del ítem (funcionamiento diferencial de los ítems) presentan interpretaciones equivalentes y comparables entre diferentes grupos (Zumbo, 2003).

La tercera fuente de evidencia de validez tratada en este escrito refiere a las relaciones del constructo a medir con la prueba y otras variables proporcionando así evidencia sobre el grado en que estas relaciones son coherentes con la teoría de partida. El marco referencial de la prueba describe cómo son las distinciones y relaciones con estas variables externas. Específicamente, los análisis de las relaciones entre los puntajes de la prueba y otras medidas que evalúan constructos similares proporcionan evidencia convergente, mientras que, relaciones con medidas de constructos teóricamente diferentes proporcionan evidencia discriminante. En cambio, si la prueba es una medida definida como un potencial predictor de alguna variable criterio, se puede evaluar relaciones test-criterio. Su estudio puede diferenciarse según el momento de su administración. Un estudio predictivo indica la fuerza de

relación entre un puntaje de la prueba y de un criterio obtenido en un momento posterior, mientras que un estudio concurrente obtiene las medidas aproximadamente al mismo tiempo. También es aplicable la evidencia sobre relaciones con otras variables para investigar preguntas de diferencias entre grupos (AERA et al., 2014; Arias y Sireci, 2021).

A modo de ejemplificación se aplicarán las pruebas de evidencia de validez en las pruebas de evaluación psicológica del ámbito del ejercicio físico. En particular, dentro de la Teoría de la Autodeterminación (TAD), se ejemplificará con propuestas instrumentales de las necesidades psicológicas básicas (NPB) (Ryan y Deci, 2017; Vansteenkiste et al., 2020). Las mismas son definidas como nutrientes esenciales para la integridad, bienestar y crecimiento de los seres humanos sin tener en cuenta el género, grupo o cultura. Cuando estas se satisfacen conducen a la integración del comportamiento generando proactividad y funcionamientos óptimos. En cambio, su frustración conlleva malestar y disfuncionalidad. La evidencia basada en el contenido y estructura interna de la prueba respaldan los supuestos teóricos sobre la cantidad, distinción, universalidad y relaciones entre las subescalas de la satisfacción y frustración de las NPB en el ejercicio físico. En referencia a la relación con otras variables, los análisis de correlaciones y modelos de ecuaciones estructurales indican que la satisfacción y frustración de la NPB son constructos diferenciables y que presentan relaciones discrepantes con consecuencias adaptativas y desadaptativas en el ejercicio (Ryan y Deci, 2017).

### **Conclusión**

Cada vez son mayores las exigencias a la que se someten los test psicológicos (Abad et al., 2011) y las mismas se ven reflejadas en los artículos instrumentales. Los estudios instrumentales para las NPB en el ejercicio parecen aportar evidencia del carácter universal, distintivo y correlacionado de sus dimensiones con consistentes pruebas de validez y confiabilidad para la interpretación de sus puntuaciones. Si bien los Estándares son reconocidos para la producción de instrumentos, no se ha identificado su uso extensivo ya que se siguen aplicando diversas terminologías (como por ejemplo validez de constructo, validez interna, validez externa, validez factorial, etc.). En su consecuencia, para incipientes conocedores de la disciplina Psicometría, puede resultar confuso y ambiguo los objetivos de un estudio instrumental. Una solución clara es aprehender respecto a la disciplina aplicada, pero además, identificar las tradiciones que conllevan el uso de los términos y sus implicancias prácticas. En particular, son las instituciones (como la Asociación Americana de Psicología) y otros autores de referencia los que han generado la costumbre y uso de términos que

describen las características de los instrumentos. Su aplicación generalizada en diversas publicaciones y artículos consolidan su práctica olvidando su origen y discusión.

Desde este trabajo, se invita a identificar las fuentes históricas de la disciplina reconociendo sus orígenes y avances para discutir su mejor uso. Además, son crecientes las alternativas de técnicas y métodos que permiten aportar evidencia de validez, por lo que la actualización tanto teórica como empírica es una necesidad que cumplir si se desea mejorar la calidad de las producciones. Conocerlas, comprenderlas y lograr su correcto uso exige una continua formación y la posibilidad de modificar tradiciones arraigadas en la práctica. Finalmente se reconoce como aspecto fundamental la integración de las fuentes de evidencia de validez para brindar una explicación coherente del grado en que la evidencia y teoría respaldan la interpretación de las puntuaciones. Este proceso estará comprendido por diversos y continuados estudios dado que siempre puede incorporarse nueva información para comprender mejor el funcionamiento de la prueba (AERA et al., 2014).

### Referencias

- Abad, F. J., Olea, J., Ponsoda, V. y García, C. (2011). *Medición en Ciencias Sociales y de la Salud*. Síntesis.
- American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education [NCME]. (2014). *Standards for educational and psychological testing*. AERA Publications.
- Arias, A., & Sireci, S. (2021). Validez y Validación para Pruebas Educativas y Psicológicas: Teoría y Recomendaciones. *Revista Iberoamericana De Psicología*, 14(1), 11–22.  
<https://doi.org/10.33881/2027-1786.rip.14102>
- Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, 26(1), 108–116. <https://doi.org/10.7334/psicothema2013.260>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellnes*. Guildford Press.
- Ryan, Richard M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67.  
<https://doi.org/10.1006/ceps.1999.1020>

- Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. In *Motivation and Emotion* (Vol. 44, Issue 1). Springer US. <https://doi.org/10.1007/s11031-019-09818-1>
- Zumbo, B. D. (2003). Does item-level DIF manifest itself in scale-level analyses? Implications for translating language tests. *Language Testing*, 20(2), 136–147. <https://doi.org/10.1191/0265532203lt248oa>