

Editorial

¿Es necesario que los test sean equitativos?

Los profesionales de la psicología, así como otros profesionales de la salud, estamos acostumbrados a utilizar test para estimar el nivel verdadero que tiene una persona o grupo de personas en un determinado constructo o variable latente a partir de sus respuestas a una serie de ítems (Abad, Olea, Ponsoda, & Garcia, 2011). Cuando en un test determinados grupos de personas obtienen puntuaciones distintas a las de sus propios méritos o condiciones debido a aspectos no relevantes para lo que pretende medir el test, que además están relacionadas con la pertenencia a determinados grupos, es cuando podemos estar delante de un test inequitativo (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Gómez-Benito, Hidalgo, & Guilera, 2010). Por consiguiente, que un test sea inequitativo puede acarrear importantes consecuencias tanto individuales como grupales. Imaginemos, por ejemplo, que en un proceso de selección de personal un determinado grupo social obtiene mejores puntuaciones en el test de selección que otro grupo, y que estas diferencias se deban a un ítem que funciona de forma diferencial entre grupos (Abad et al., 2011) debido a un aspecto no relacionado con lo que quiere medir el test (por ejemplo el idioma o algún otro aspecto cultural). Las personas del grupo más favorecido tendrían, de forma sistemática, muchas más probabilidades de conseguir el empleo que el grupo menos favorecido.

Entonces, las preguntas que nos plantearíamos serían las siguientes: ¿podemos evitar que los test sean inequitativos? y en caso afirmativo, ¿en qué momento podemos evitarlo? Para contestar a estas preguntas, podemos decir en primer lugar, que en el momento de desarrollar un test, este debería construirse en el marco de un diseño universal. Es decir, el test debería estar pensado para que pueda ser respondido sin desventajas, no relacionadas con lo que el test pretende medir, por el máximo número de personas de la población a la que va dirigido el test, independientemente de aspectos como el género, edad, antecedentes culturales o lingüísticos, nivel socioeconómico o discapacidad. Sin embargo, no en todos los casos se puede maximizar la equidad a partir de un diseño universal. En estas situaciones, es útil recurrir a adaptaciones. Es fácil entenderlo si nos imaginamos determinados colectivos, como personas con dificultades visuales, personas con un nivel insuficiente del idioma, etc.

Las adaptaciones son cambios en el diseño o en la administración de la prueba que pretenden incrementar el acceso al test a estos determinados grupos que podrían ser evaluados de forma injusta por el mero hecho de pertenecer a un sector desfavorecido. De este modo los *standards for educational and psychological testing* señalan dos importantes tipos de adaptaciones: las acomodaciones y las modificaciones del test (American Educational Research Association et al., 2014).

Acomodación: Las acomodaciones son aquellos cambios pequeños en la presentación, el formato y/o la administración del test con el objetivo de mantener el constructo original y asegurar que los resultados sean comparables con el test original. Por ejemplo, a las personas que tienen un nivel insuficiente del idioma se les puede dejar un diccionario o a las personas con problemas de visión se les puede administrar el test de forma oral.

Modificación: Las modificaciones son aquellas adaptaciones del test, que para solventar un problema de inequidad suponen un cambio en el constructo. Por ejemplo, imaginemos a personas con discalculia (dificultad para comprender y realizar cálculos matemáticos) a los que se les tiene que evaluar sus conocimientos matemáticos. Si a estas personas les dejamos una calculadora podrían ser evaluadas independientemente de su enfermedad. Sin embargo, al introducir la calculadora,

podríamos evaluar su capacidad de resolver problemas pero no su capacidad computacional con lo que al final evaluaríamos un constructo distinto al original. En el caso que hagamos una modificación del test, al evaluar otro constructo, es lógico pensar que se necesitaran estudiar las pruebas psicométricas (validez, fiabilidad y establecimiento de criterios para la interpretación de sus puntuaciones) (American Educational Research Association et al., 2014).

En conclusión, cuando se desarrolla o se adapta un test, se pretende que este pueda ser respondido sin problemas por toda la población para la que ha sido diseñado (población estándar). Dentro de esta población, nos encontramos determinados subgrupos de personas que pueden tener distintas características pero no por eso necesitan de una acomodación o una modificación del test. Sin embargo, alguno de estos subgrupos podría presentar alguna desventaja que hiciera necesaria la introducción de algún pequeño cambio en el test (acomodación) o algún cambio que llegara a afectar al constructo (modificación). Lo ideal sería que en el momento de desarrollar el test ya se pudieran identificar y afrontar los problemas de inequidad para determinados grupos. Es decir, todas estas posibles modificaciones y/o adaptaciones ya se hubieran podido plantear y analizar desde el principio. Sin embargo, estas adaptaciones nos las podemos plantear en cada una de las fases del test, ya sea en el desarrollo, en una posible adaptación y/o en un test ya publicado. Existen varias formas para determinar si un test presenta inequidad (Espelt et al., 2014) y hemos visto que para solucionar dicho problema podemos realizar acomodaciones o modificaciones. De este modo, y respondiendo a la pregunta formulada en el título de esta editorial, existiendo las herramientas que nos permiten detectar y solucionar la inequidad de un test, debería ser indiscutible que es necesario poner el máximo de esfuerzo para asegurar que los test sean equitativos y de este modo solventar estas diferencias, innecesarias, indeseadas, evitables e injustas que pueden producirse en las puntuaciones de un test entre grupos de personas por el mero hecho de pertenecer grupos distintos.

Espelt A

Agència de Salut Pública de Barcelona, Plaça Lesseps 1 – 08023 Barcelona;
 Departament de Psicobiologia i de Metodologia de les Ciències de la Salut,
 Universitat Autònoma de Barcelona – 08193 Bellaterra;
 CIBER de epidemiologia y Salud Pública, Instituto Carlos III,
 Av. Monforte de Lemos, 5 – 28029 Madrid.

Doval E

Departament de Psicobiologia i de Metodologia de les Ciències de la Salut;
 y Grupo de Investigación en Estrés y Salud (GIES).
 Universitat Autònoma de Barcelona – 08193 Bellaterra.

Viladrich C

Departament de Psicobiologia i de Metodologia de les Ciències de la Salut,
 Universitat Autònoma de Barcelona – 08193 Bellaterra.

REFERENCIAS

- Abad, F., Olea, J., Ponsoda, V., & Garcia, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Editorial Síntesis.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Espelt, A., Viladrich, C., Doval, E., Aliaga, J., García-Rueda, R., & Tárrega, S. (2014). [Fair use of tests in health sciences.]. *Gaceta Sanitaria/S.E.S.P.A.S.*
- Gómez-Benito, J., Hidalgo, M. D., & Guilera, G. (2010). El sesgo de los instrumentos de medición. Tests justos. *Papeles Del Psicólogo*, 31(1), 75-84.