

Validez.

4.2. Fuentes de evidencia de validez.

Es un aspecto esencial de la medición psicológica y se relaciona con la investigación del significado teórico de las puntuaciones obtenidas por medio de un test (Oltiden, 2003).

Según Oltiden (2003), existen fuentes de evidencia internas y externas:

Fuentes internas de evidencia.
Según Oltiden (2003), existen fuentes de evidencia internas y externas. La primera categoría se relaciona con el test y sus componentes (Items) en sí mismos. La lógica implícita de las evidencias incluidas en esta categoría se relaciona con el primer objetivo del proceso de validación de las puntuaciones de un test, es decir, medir un constructo con un significado unívoco, estrictamente definido.

Evidencia basada en el contenido del test.
Este tipo de evidencia se obtiene demostrando que el contenido (Items) del test es una muestra representativa del constructo o dominio respecto del cual se desea hacer alguna inferencia.

Evidencia basada en la estructura interna del test.
Las evidencias relacionadas con la estructura interna de un test indican si las relaciones entre los items y las dimensiones (factores, escalas) permiten confirmar la existencia de los constructos que el test pretende medir. El marco conceptual de un test puede proponer una dimensión unitaria de comportamiento o varios factores.

Fuentes externas de evidencia.
El análisis de las relaciones de las puntuaciones del test con variables externas al mismo test es otra fuente importante de evidencia. Las variables externas pueden ser las medidas de algún criterio que el test intenta predecir, así como las puntuaciones de otros tests que miden constructos semejantes o diferentes.

Evidencia convergente-discriminante.
Al construir un test (una escala para medir inestabilidad emocional, por ejemplo) se deben comparar los puntajes obtenidos, tanto con otros tests elaborados para medir el mismo atributo (la escala Neuroticismo de un inventario de personalidad, por ejemplo) como con los diseñados para medir otros atributos (la escala Amabilidad del mismo test, por ejemplo).
La lógica de ambos procedimientos complementarios es evidenciar que el test en cuestión mide realmente el constructo que se propone medir, al correlacionarse con otros tests reconocidos que miden el mismo constructo, y no correlacionarse con tests que miden constructos diferentes.

Evidencia de las relaciones entre las puntuaciones del test y criterios externos.
Este tipo de evidencia es especialmente importante en los contextos aplicados de la psicología en los que se busca predecir de manera precisa un determinado comportamiento o desempeño a partir de las puntuaciones de un test.
En función del número y tipo de variables predictoras y dependientes (criterios), pueden utilizarse diferentes análisis estadísticos dentro del modelo lineal general.

La validez de las puntuaciones de un test significa la efectividad con que se puede predecir el desempeño de una muestra en una situación real (laboral o académica, por ejemplo) o criterio diferente del test en sí mismo.

El estadístico que se emplea habitualmente para corroborar esa utilidad predictiva es la correlación entre la puntuación obtenida en la prueba y la puntuación en algún criterio que representa los desempeños externos al test.

Correlación bivariada con un criterio.

El diseño del experimento básico que se emplea para verificar si las puntuaciones de un test se relacionan con un criterio es, según Murat (1985), el siguiente:

- a. De una población determinada se extrae una muestra grande.
- b. Los individuos de la muestra son evaluados tanto con el test como con el criterio que interesa predecir.
- c. Con los datos obtenidos del paso b se estima un coeficiente de correlación. Si éste es significativamente distinto de cero podrá aseverarse que el test es un predictor del criterio para cualquier individuo que pertenezca a esa población.

Diferencias de grupo.

Se refieren a variables que pueden influir (moderar) en las correlaciones entre un test y un criterio, tales como sexo, edad o nivel socioeconómico. En general, se recomienda emplear una muestra lo más heterogénea posible.

Por otra parte, existen factores que pueden afectar las relaciones entre los puntajes de un test y las medidas del criterio. Tres de los más importantes son:

Variabilidad de las respuestas al test y al criterio:

En este caso, los puntajes de un test extenso poseen varianzas mayores y las puntuaciones obtenidas en pruebas breves tienen varianzas más reducidas y, por consiguiente, coeficientes de correlación inferiores con las medidas del criterio. Estas afirmaciones deben entenderse atendiendo a los condicionantes formulados en el capítulo de confiabilidad, en el sentido de que la calidad de los items es un factor más relevante que la mera cantidad.

Confiabilidad del test y del criterio:

Puesto que en la medida en que el test predictor y el criterio aumenten su confiabilidad se incrementará el coeficiente de validez test-criterio.

Correlación múltiple con un criterio.

Los métodos multivariados son numerosos (análisis factorial, regresión múltiple, análisis de senderos, análisis discriminante, entre otros) y se emplean con diferentes finalidades, tales como predecir la pertenencia a un grupo, explicar la variabilidad de una variable dependiente o verificar la estructura de un constructo medido.

En el contexto de la validez de los tests, mediante métodos multivariados como el análisis de regresión múltiple podemos estimar en cuánto se incrementa la precisión de la predicción de un criterio cuando un test se incluye en una batería de tests en comparación con las ocasiones en que no se lo incluye (Aiken, 2003).

El análisis de regresión múltiple permite ponderar la contribución independiente realizada por cada variable predictor para la explicación de un criterio determinado, así como estimar la contribución conjunta de un conjunto de predictores a la explicación del criterio (Aron y Aron, 2001).

Los estadísticos fundamentales del análisis de regresión múltiple son:

- a) el coeficiente de regresión estandarizado o beta (β), que indica cual es la importancia relativa de cada variable independiente en la predicción de la variable dependiente.
- b) el coeficiente de correlación múltiple (R) que expresa el grado de asociación entre dos o más variables independientes (predictoras), en conjunto, con una variable dependiente.
- c) el coeficiente de correlación múltiple al cuadrado (R^2), que permite determinar el porcentaje de varianza explicada de la variable dependiente por parte del conjunto de predictores de la ecuación.
- d) el cambio en R que indica el porcentaje de varianza de la variable dependiente explicada independientemente por cada uno de los predictores.

Evidencia de las consecuencias de la aplicación de test.

La evidencia relacionada con las consecuencias de la aplicación de tests sólo es relevante para la validez cuando se relaciona con alguna fuente de invalidez, tales como una pobre representación del constructo o la existencia de componentes no relevantes para el constructo. Al igual que acontece con la evidencia de proceso de respuesta al test, la evidencia de consecuencias es poco investigada en la práctica.