

***PSICOLOGÍA
IBEROAMERICANA***

Psicología Iberoamericana

ISSN: 1405-0943

psicología.iberoamericana@uia.mx

Universidad Iberoamericana, Ciudad de
México
México

Liht Sigall, José

El Análisis de Factores y la Validez de las Medidas Psicométricas

Psicología Iberoamericana, vol. 13, núm. 1, 2005, pp. 2-4

Universidad Iberoamericana, Ciudad de México

Distrito Federal, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=133926982002>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Editorial

El Análisis de Factores y la Validez de las Medidas Psicométricas

Factor Analysis and Validity of Psychometric Measurements

José Liht Sigall*

UNIVERSIDAD IBEROAMERICANA, CIUDAD DE MÉXICO

Resumen

En el presente editorial se discuten algunos malos hábitos y errores comunes en la elaboración de instrumentos psicométricos. El concepto de *validez* y su necesidad en la disciplina de la psicología se clarifican y se elaboran en los subtipos principales de la misma. Asimismo, se considera brevemente la técnica estadística del análisis factorial y su relación con el establecimiento de la validez.

Descriptores: psicometría, validez de constructo, validez de criterio, validez de contenido, análisis factorial

Abstract

Some common bad habits and errors in the construction of psychometric instruments are discussed. The concept of validity and its necessity in psychology is argued for as well as its subtypes. The statistical technique of factor analysis is briefly considered in its relationship with establishing the validity of a psychological test.

Keywords: psychometrics, construct validity, criterion validity, content validity, factor analysis

A lo largo de mi labor editorial y docente me ha llamado la atención la poca seriedad con que se elaboran instrumentos para medir constructos psicológicos. Pareciera que dicha actividad se ha vuelto la preferida de autores y tesisistas sin importar que la gran mayoría de dichos instrumentos no tengan la menor trascendencia para la disciplina. A menudo se cree que se aporta algo al elaborar instrumentos nuevos para constructos viejos o cuando se pretende innovar proponiendo constructos novedosos pero inútiles en su capacidad para lograr inferencias valiosas. En este triste panorama de la disciplina, la carencia que encuentro con mayor frecuencia es que en los trabajos donde se elaboran instrumentos de medición, independientemente de la viabilidad o necesidad de un nuevo instrumento, no se propone establecer la validez de criterio y de constructo de los mismos derivadas de teorías psicológicas sólidas, entendiendo como

teoría sólida a un conjunto reducido de principios que explica o predice un importante número de observaciones, y se propone, en cambio, sustentar la validez por medio de documentar la homogeneidad de los reactivos o la confirmación de cierta estructura factorial esperada. Sin embargo, se ignora flagrantemente el sustentar con datos empíricos la utilidad del instrumento para predecir o indicar conducta externa a una situación de medición psicométrica o que pretenda vincular el mismo con una formulación teórica del constructo que se pretende medir. En consecuencia, en el presente trabajo propongo profundizar en el concepto de *validez* y en los errores o, en el mejor de los casos, omisiones en las que a menudo caen los investigadores al ignorar la teoría como sustento de cualquier esfuerzo psicométrico serio y al errar en lo que se considera evidencia de la validez de un instrumento.

* Dirigir correspondencia a: Prol. Paseo de la Reforma 880, México D.F., C.P. 01210. Correo electrónico: jose.sigall@uia.mx

Teoría e instrumentos psicométricos

Cualquiera que sea el instrumento psicológico del que se esté hablando, podrá decirse, gran parte de su valor residirá en qué tanto el mismo permite proponer enunciados ciertos de los cuales se puedan derivar inferencias trascendentes sobre un individuo; o en otras palabras: inferencias válidas. Por ejemplo, los buenos *tests* de inteligencia fluida nos permiten predecir qué tanto éxito laboral tendrá un individuo durante la mayor parte de su vida; el nivel de inteligencia al que sus hijos podrán aspirar; un rango para sus padres; el provecho que el sujeto podrá derivar de una educación formal, así como su habilidad para resolver problemas intelectuales. Debido a que el constructo teórico de inteligencia implica estabilidad durante la mayor parte de la vida, un amplio factor hereditario y eficiencia en la capacidad de abstracción, inducción y deducción, o la capacidad para razonamiento lógico que subyace a la solución de cualquier problema intelectual, cuando asignamos un coeficiente intelectual estaremos afirmando implícitamente una gran cantidad de información trascendente acerca de un individuo. Dicho poder, se logra debido a que la inteligencia es un constructo teórico que logra explicar un gran número de observaciones. El éxito en la elaboración de un instrumento que logra asignar magnitudes individuales de dicho constructo teórico permite utilizar todo el poder inferencial implícito en la teoría de la inteligencia. Consecuentemente, el puntaje que un individuo obtiene en el Test de Matrices Progresivas de Raven (Raven, 1976), excelente para medir la dimensión subyacente a toda conducta inteligente, es interesante no porque nos informa que un sujeto pudo contestar matrices de un grado de dificultad tal, sino por todas las inferencias que podemos hacer a partir de la conexión entre el instrumento y la teoría de la inteligencia. Dichas matrices fueron elaboradas a partir del refinamiento de la teoría de la inteligencia que establece que el elemento de habilidad mental general es la inteligencia fluida, la cual subyace a toda conducta inteligente independientemente de la exposición a oportunidades educacionales o experiencias de vida, la cual es en gran medida genéticamente transmitida.

Por lo tanto, cuando se pretende elaborar un instrumento simplemente para medir un constructo novedoso que no se desprende de una teoría exitosa ni forma parte de la elaboración eventual de alguna, estará limitado en su poder para llevar a cabo inferencias. Una excepción es cuando se encuentra una relación empírica predictiva entre un instrumento y cierto cri-

terio o conducta, como ocurre con el MMPI. A pesar de que éste es ateorico, su valor se fundamenta en el éxito con el cual sus reactivos son capaces de discriminar grupos diagnósticos.

Validez

Con el objeto de sustentar la validez de un instrumento se utilizan tres principales operacionalizaciones de dicho término: a) la validez de contenido, b) la de criterio y c) la de constructo (Pedhazur & Schmelkin, 1991).

La validez de contenido se obtiene cuando nuestro instrumento evalúa la totalidad del área que supone medir sin excluir aspectos o dimensiones. Por ejemplo, una prueba que pretende establecer mi conocimiento sobre la historia de México no podrá incluir preguntas de la época colonial únicamente, ya que estaría ignorando otras dimensiones de lo que pretende medir: mi conocimiento de historia mexicana global. A menudo dicha validez es obtenida mediante el visto bueno de expertos que conocen el área que se está evaluando. En instrumentos psicológicos, frecuentemente se encuentran diferencias de opinión respecto a cuáles y cuántas deben ser las dimensiones de un constructo. Por ejemplo, cuántos rasgos de personalidad existen.

Por otra parte, la validez de criterio se logra cuando un instrumento predice o coincide con observaciones independientes a las hechas mediante instrumentos (o validez convergente). Por ejemplo, el coeficiente intelectual está fuertemente correlacionado con el desempeño que tienen los alumnos y su promedio general. De la misma manera, si los puntajes de un instrumento sobre satisfacción marital predicen con éxito qué parejas buscarán ayuda psicológica en un futuro cercano, se dice que cuenta con validez en un criterio posible.

La validez de constructo se logra cuando mediante un instrumento se sustentan las hipótesis derivadas de un constructo teórico; por ejemplo, en su instrumento sobre personalidad, Eysenck (1982) propuso la existencia de tres dimensiones principales: extraversión, neuroticismo y psicoticismo, y añadió que las mismas tienen un sustento biológico evolutivo y genético. Formidablemente, el instrumento que él elaboró para medir dichas dimensiones coincidió en predecir los niveles registrados de actividad de la formación reticular en el caso de la extroversión, del sistema límbico en el caso del neuroticismo y de niveles de testosterona en cuanto al psicoticismo. Asimismo, se encontró que cada una de estas dimensiones estaba

altamente determinada por factores hereditarios en estudios de gemelos dizigóticos y monoigóticos. Cuando se sustentan constructos teóricos de tal manera se dice que el instrumento cuenta con validez de constructo. Como se puede percibir, dicha validez se establece con base en un esfuerzo de investigación importante.

El análisis factorial

Una técnica importante que puede ser muy útil en la exploración del número de dimensiones que subyacen la conducta de respuesta que evoca un instrumento psicológico, es el análisis factorial. Dicho análisis, aunque no sin complicaciones y ambigüedades en sí mismo, permite explorar las intercorrelaciones entre los reactivos de una prueba y establecer subgrupos o dimensiones subyacentes en las asociaciones entre reactivos. Los más importantes modelos psicométricos de la personalidad y de la inteligencia (los de Cattell, Eysenck y McRae y Costa) han sido establecidos a partir del análisis factorial. Dicha técnica ha servido para descubrir cuántas dimensiones subyacen dichos constructos, por lo que sus más importantes contribuciones han sido la exploratoria y la de construcción teórica y sólo indirectamente se ha utilizado en la búsqueda de evidencia de validez.

El contar con análisis factoriales que repliquen una estructura factorial esperada de acuerdo a la tabla de especificaciones de un instrumento en construcción es un prerrequisito importante en el largo camino para establecer la validez. Sin embargo, esto de ninguna manera implica que el instrumento es eficiente en predecir criterios externos a la situación de prueba. Muchos autores frecuentemente afirman haber establecido

que una prueba es válida a partir de haber reproducido una estructura de cargas factoriales esperada y sin haber llevado a cabo investigación sobre la eficiencia predictiva del mismo fuera de las situaciones de prueba. Dicha posición es idéntica a la de afirmar que la consistencia interna es suficiente como para establecer la validez de un instrumento o que si se cuenta con validez de contenido la validez de criterio y la de constructo estarán dadas.

Como ejemplo, el que se logre elaborar una escala para establecer la apreciación de los animales de manera que todos los reactivos acerca de los mamíferos carguen en un solo factor mientras que los que se refieren al aprecio por los reptiles carguen en otro, no se puede tomar como evidencia para inferir que una persona, con un puntaje alto en dicho instrumento, en una de sus dimensiones tenderá en efecto a querer tener contacto con los animales cuando la oportunidad se le presente. Dicha suposición no puede hacerse por el simple hecho de contar con un número de reactivos que presenten homogeneidad e independencia de los demás factores. Existe amplia evidencia de que grupos de reactivos pueden tener equivalencia semántica y no predecir conducta alguna (Cattell y Kline, 1977); tal fenómeno da cuenta de que la única muestra definitiva de validez es establecer estudios que documenten el valor predictivo de un instrumento y no sólo se asuma al contar con homogeneidad entre sus reactivos. Dicha homogeneidad puede ser artefacto de un factor de deseabilidad social o tendencia a asentir.

Cada uno de los factores homogéneos que se obtengan en un análisis factorial debe ser evaluado en cuanto a su validez de criterio o predictiva (Kline, 1998) si no se quiere correr el riesgo de tener un artefacto que en realidad no prediga conducta alguna.

Referencias

- Butcher, J.N., Grant Dahlstrom, W., Graham, J.R., Tellegen, A. & Kaemmer, B. (1995). *Inventario multifásico de la personalidad Minnesota-2*. México: Manual Moderno.
- Cattell, R.B. & Kline, P. (1977). *The Scientific Analysis of Personality and Motivation*. Nueva York: Academic Press.
- Eysenck, H.J. (1982). *Personality, genetics and behavior, Selected Papers*. Nueva York: Praeger.
- Kline, P. (1998). *The new psychometrics: Science, psychology and measurement*. Londres: Routledge.
- Pedhazur, E.J. & Schmelkin, L. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, Nueva Jersey: Lawrence Earlbaum.
- Raven, J.C. (1976). *Test de matrices progresivas*. México: Paidós.