

Nuevos métodos estadísticos para la investigación en evaluación de la personalidad

María Cristina Richaud de Minzi¹

Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

En este trabajo se realiza una revisión de las nuevas técnicas estadísticas y de su utilidad para la investigación en personalidad. Los nuevos modelos y los avances en la medición de la personalidad y la psicopatología sugieren que la investigación en este campo y en su evaluación han entrado en un estadio avanzado de desarrollo. En las dos últimas décadas se han producido importantes desarrollos en estadística y medición. El refinamiento de las técnicas de análisis multivariado ha sido fundamental en la evaluación de la personalidad debido a la complejidad de las relaciones entre sus variables. Los procedimientos de análisis multivariado proveen la oportunidad de examinar la complejidad de esas interacciones a través de métodos de análisis para variables múltiples. Por otra parte, se han desarrollado los modelos de ecuaciones estructurales y técnicas multivariadas para analizar variables categóricas. Los últimos desarrollos corresponden al escalamiento multidimensional y a la teoría de la respuesta al ítem.

Palabras clave: personalidad, medición, evaluación, métodos multivariados.

New statistical methods for research in personality assessment

In the present work a review of the new multivariate techniques and why they appear especially suited to the personality research is presented. Emerging models of personality and advances in the measurement of personality and psychopathology suggest that research in this field has entered a stage of advanced development. The past two decades have shown important developments in statistics and measurement. Refinement of multivariate statistics has been especially important in personality assessment because of the complexity of relations among personality variables. Multivariate procedures provide the opportunity to examine the complexity of these interactions by providing methods of analysis for multiple variables. On the other hand, structural equation modeling and multivariate techniques for analyzing categorical variables have been developed. Multidimensional scaling and item response theory are the last developments.

Key words: personality, measurement, assessment, multivariate methods.

¹ Doctora en Filosofía y Letras con mención en Psicología. Directora del Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental (CIIPME) en Argentina. Se desempeña como Investigadora Principal del CONICET. Es Profesora Invitada de la Universidad Nacional de San Luis, Argentina. Sus áreas de interés son: estrés en niños y adolescentes, teoría y medición de la personalidad. Dirección postal: Tte. Gral Perón 2158, 1040 Buenos Aires, Argentina. Correo electrónico: mrichaud@conicet.gov.ar o minzi@ciudad.com.ar.

Los nuevos modelos y los avances en la medición de la personalidad y la psicopatología sugieren que la investigación en este campo y en su evaluación han entrado en un estadio avanzado de desarrollo. El progreso en cualquier campo científico está marcado por etapas. Las primeras etapas se caracterizan generalmente por el descubrimiento del fenómeno de interés, la construcción de la teoría y la invención de métodos, procedimientos y análisis adecuados para explorarlo. Los estadios más tardíos se caracterizan por la creciente complejidad de los modelos subyacentes, teorías e hipótesis de investigación. Se hace necesario entonces diseñar métodos de investigación y procedimientos más complejos para el análisis de los datos con el objeto de obtener pruebas más rigurosas de las nuevas teorías. Reis y Stiller (1992) llaman a esta etapa investigación de segunda generación e indican que refleja un crecimiento en la madurez en el campo científico (Schinka, LaLone & Broeckel, 1997).

Si nos referimos a los métodos de análisis de la personalidad y de su evaluación, es indudable que en las dos últimas décadas se han producido importantes desarrollos en estadística y medición. El refinamiento de las técnicas de análisis multivariado ha sido fundamental en la evaluación de la personalidad debido a la complejidad de las relaciones entre sus variables. Los procedimientos de análisis multivariado proveen la oportunidad de examinar la complejidad de esas relaciones.

Por otra parte, se han desarrollado los modelos de ecuaciones estructurales (SEM, por sus siglas en inglés) (Bentler, 1985; Jöreskog & Sörbom, 1979) y técnicas multivariadas para analizar variables categóricas como la regresión logística (Aldrich & Nelson, 1984). Los últimos desarrollos corresponden al escalamiento multidimensional (Schiffman, Reynolds & Young, 1981) y a la teoría de la respuesta al

ítem para respuestas no dicotómicas (Hoskens & De Boeck, 1995, 2001; Lord, 1980).

La utilización de estas nuevas técnicas en la evaluación de la personalidad es importante porque los diseños que emplean múltiples variables son generalmente óptimos para el estudio de las hipótesis de investigación en personalidad.

Los dos mayores argumentos para recomendar el uso de técnicas multivariadas son: a) múltiples tests estadísticos univariados aumentan la tasa de error, y b) la investigación en ciencias del comportamiento no es univariada por naturaleza y esta realidad debe ser res-petada utilizando técnicas multivariadas apropiadas (Crowley & Fan, 1997).

Los científicos que tienen que ver con el comportamiento... raramente pueden darse el lujo de trabajar con experimentos bivariados simples, en los cuales una variable independiente simple es manipulada y sus consecuencias son observadas en una variable dependiente simple. Aún aquellos científicos que piensan que hacen esto están a menudo equivocados. Las variables que ellos manipulan y observan directamente no son las que tienen real interés teórico sino que son sólo algunas variables convenientes que actúan como derivadas... Un análisis experimental completo debería ser multivariado, con un número de manipulaciones experimentales alternativas por un lado y un número de medidas de respuestas alternativas por el otro (Loehlin, 1992, p.1).

Esto no significa que estos procedimientos deban utilizarse sólo porque son potentes y sofisticados, sino que en diseños claros, dirigidos a un problema bien definido, permiten explorar posibilidades que otros procedimientos no consiguen. Como dice Weiner (1991), las implicaciones de los resultados de investigación siempre tienen preponderancia sobre la elegancia del diseño.

Regresión logística

Para comenzar con técnicas relativamente nuevas para el estudio de variables categoriales, en los años setenta se empezaron a utilizar las regresiones logísticas, pero sobre todo en el campo de la bioestadística, epidemiología y economía. Esta técnica que constituye una estadística mucho más ajustada para predecir, a partir de variables cualitativas o cuantitativas, sucesos clasificados como cualitativos, no era aún mencionada en los ochenta en psicología por autores de la importancia de Cohen o Pedhazur (Davis & Offord, 1997). Anteriormente al desarrollo de esta técnica se utilizaba para estos casos la regresión múltiple lineal con una probable violación de los supuestos básicos de esta medida.

En la investigación en personalidad y psicopatología, muchas variables de interés son categóricas. Las tipologías de personalidad y la presencia o ausencia de enfermedad como objetos de interés son casos en los que corresponde el análisis de regresión logística (Davis & Offord, 1997).

En la elección de las técnicas de análisis es indispensable revisar si se respetan los supuestos subyacentes a cada una de ellas. Por ejemplo, si se utiliza regresión lineal la distribución del error en la variable dependiente debe ser normal, lo cual no es posible cuando la misma es dicotómica, lo que implica una distribución binomial. Si se utiliza regresión múltiple lineal con una variable dependiente categorial, los tests de significación y el intervalo de confianza no serán válidos. En primer lugar, la función logística binaria está limitada por 0 y 1; cuando la variable dependiente es codificada como 0 (ausencia) y 1 (presencia), produce una probabilidad estimada de ocurrencia del resultado, dados los diferentes valores de los predictores. Segundo, la regresión logística utiliza la razón de los residuales. Dado un cierto valor de los predictores, pueden calcularse los residuos que favorecen cierto resultado. También pueden estudiarse

los cambios en las razones estimadas de los residuales cuando se cambia uno o más predictores. Tercero, la regresión lineal múltiple implica una función lineal recta, mientras que la regresión logística es una función curvilínea. El incremento en la probabilidad estimada del resultado por unidad de la variable dependiente no es constante a través de todos los valores de la variable independiente. Cuarto, cuando las variables predictoras son categoriales, el análisis de los datos usando regresión logística es muy similar al análisis de la tabla de contingencia (Davis & Offord, 1997).

Supongamos un estudio donde se quisiera conocer cuán bien predicen las variables edad, sexo y valores obtenidos en una prueba de satisfacción corporal, la presencia o ausencia de trastornos en la alimentación. Sería un caso típico de aplicación de la regresión logística, que nos daría información acerca de cuánto mejora el modelo propuesto la predicción al puro azar y por lo tanto cuán bueno es para predecir los trastornos de alimentación. Por otra parte, a partir de los valores de significación puede determinarse cuáles de las variables aumentan la probabilidad de sufrir trastornos de alimentación y cuáles disminuyen esa probabilidad. Sin esta herramienta estadística se habría realizado una regresión lineal múltiple, con la comentada violación de los supuestos básicos de la misma, o un Ji cuadrado que significaría una parcialización del estudio, ya que habría que analizar las variables de a dos.

Ecuaciones estructurales

En los años recientes, el modelo de ecuaciones estructurales (SEM) se ha vuelto muy útil en las ciencias sociales y del comportamiento porque permite estimar y poner a prueba interrelaciones hipotéticas entre un conjunto de variables significativas.

El SEM expresa relaciones entre varias variables, que pueden ser observadas o hipotéticas o no observables. Muchas técnicas estadísticas

ampliamente utilizadas pueden considerarse como casos especiales del SEM, incluyendo el análisis de regresión, las correlaciones canónicas, el análisis factorial confirmatorio (AFC) y el análisis de senderos (*path analysis*) (Bagozzi, Fornell & Larcker, 1981; Bentler, 1992; Jöreskog & Sörbom, 1989).

Sin embargo, hay diferentes aspectos del SEM que lo distinguen cualitativamente de las técnicas univariadas (por ejemplo, análisis de regresión múltiple y análisis de variancia) y de las multivariadas (análisis multivariado de variancia y análisis factorial exploratorio).

Las técnicas univariadas están limitadas metodológicamente para estudiar las relaciones entre múltiples variables dependientes, debido a que puede examinarse sólo una variable por vez, dejando ciertas relaciones o interacciones completamente inexplicadas. Comparado con la mayoría de las otras técnicas multivariadas, el SEM tiene un enfoque confirmatorio más que exploratorio del análisis de los datos (Byrne, 1994) ya que el patrón de relaciones entre variables es especificado a priori basándose en hipótesis teóricas. La característica distintiva del SEM lo hace especialmente apropiado para poner a prueba modelos teóricos. En cambio, la mayoría de las técnicas multivariadas son descriptivas y exploratorias por naturaleza, lo que las hace menos apropiadas para poner a prueba modelos.

Diferencia entre el análisis factorial exploratorio y el análisis factorial confirmatorio

El análisis factorial confirmatorio (CFA) está guiado por teorías previas mientras que el análisis factorial exploratorio (EFA) es una técnica para el descubrimiento de la probable estructura subyacente a una muestra de datos (Bollen, 1989; Jöreskog & Sörbom, 1989; Pedhazur & Schmelkin, 1991). Es decir que el EFA se aplica cuando deseamos explorar los datos para ver qué características, rasgos y relaciones existen.

En este caso generalmente no tenemos ningún modelo hipotético sobre los datos y todas las variables pesan en todos los factores. El EFA es muy útil en las primeras etapas en la construcción de instrumentos de medición cuando estamos generando un modelo teórico o hipótesis que luego se pondrán a prueba empíricamente (Gorsuch, 1983). El CFA comienza, en cambio, con un modelo teóricamente plausible construido con base en información a priori acerca de la naturaleza de la estructura de los datos o en teorías importantes del campo de estudio. Las variables están limitadas a pesar sólo en un factor o en pocos de ellos. El propósito de utilizar CFA es poner a prueba la hipótesis de que el modelo teórico propuesto se ajusta a los datos empíricos. Aunque el EFA puede ser muy útil en la generación de hipótesis, es necesario poner a prueba dichas hipótesis a través del CFA.

Finalmente, el EFA se diferencia del CFA en términos del efecto de las fluctuaciones de muestreo en los resultados. El EFA tiende a recibir la influencia de la idiosincracia de cada muestra particular. A menos que se lleven a cabo procedimientos de validación cruzada o de replicación, la estructura derivada del EFA debe ser analizada con cuidado, ya que los errores de muestreo y las características propias de la muestra tienen mucha influencia en este enfoque.

En un grado considerable el CFA evita este problema ajustando un modelo teórico previamente especificado a una muestra de datos. Dado que es una técnica que implica una teoría, la construcción del modelo no se ve afectada por una muestra particular de datos y la probabilidad de ser influida por fluctuaciones de muestreo e idiosincracias de la muestra están muy reducidas. Por ello, el CFA es considerado generalmente como más ventajoso que el exploratorio en la construcción y puesta a prueba de teoría.

El CFA es especialmente apropiado para estudiar la validez constructiva de un instrumento de evaluación de la personalidad de dos maneras diferentes: a) estudiando si los ítems particulares se

ajustan a determinadas subescalas o dimensiones, y b) estudiando si las subescalas dentro o entre instrumentos de medición realmente evalúan los constructos latentes supuestos. Por ejemplo, en un estudio supuesto se desea poner a prueba la validez constructiva de un test que operacionalice un modelo de personalidad con cuatro factores (F1, F2, F3, F4) que covarían entre sí y cada uno con dos facetas (X1, X2, X3, X4, Y1, Y2, Y3, Y4). En un primer momento se pueden construir los ítems siguiendo el modelo teórico propuesto y luego realizar un análisis factorial exploratorio para observar cómo se pesan los ítems en los factores de primer orden (facetas) propuestas. Luego, se podrá hacer un análisis factorial exploratorio de segundo orden una vez más para estudiar cómo se distribuyen las facetas en los cinco factores hipotetizados.

Una vez obtenido este modelo empíricamente, para estudiar la validez constructiva del mismo, la técnica más apropiada sería el CFA, donde se analizaría si efectivamente determinadas facetas se correlacionan con determinados factores y si determinados ítems se correlacionan con facetas predeterminadas, si existe o no ortogonalidad entre las facetas y entre los factores.

La forma más fácil de presentar el SEM es gráficamente. Por convención, un círculo o un óvalo representan las variables latentes, que constituyen los constructos hipotéticos inobservables, y un cuadrado o rectángulo representan las variables observadas medidas. Una variable latente puede tener una o preferentemente más variables observadas como indicadores. Una línea recta con una flecha en el extremo representa un efecto hipotético de una variable sobre otra, mientras que una línea curva con una flecha en ambos extremos representa una correlación hipotética entre dos variables sin que implique una relación causal.

En la terminología del SEM, las variables latentes u observadas que solamente ejercen un efecto sobre otras variables se llaman variables

exógenas, mientras que las que reciben efecto de otras se llaman endógenas. El SEM puede ser visto como formado por tres componentes: dos modelos de medida y uno estructural. El primer modelo de medida describe la relación entre variables latentes exógenas y sus indicadores. El segundo modelo de medida describe la relación entre las variables latentes endógenas y sus indicadores. El modelo estructural por otro lado describe la relación entre las variables latentes. Es decir que el modelo de medida implica relaciones entre variables empíricas y variables latentes y el modelo estructural describe las relaciones entre variables latentes. En los modelos de medida las e representan la variancia en la variable observada que no está explicada por los constructos latentes o variancia de error. En las aplicaciones del SEM no siempre es necesario implementar el modelo completo.

La Figura 1 representa el modelo de personalidad postulado en el ejemplo:

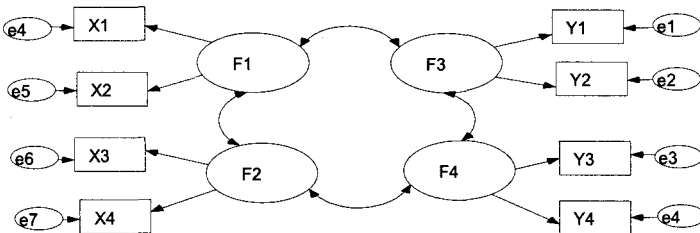


Figura 1. CFA de un modelo de personalidad con cuatro factores con dos facetas cada uno.

Se pone entonces a prueba si los datos ajustan al modelo, analizando así la validez constructiva de la prueba a través de un análisis de tipo confirmatorio, mucho más preciso que el que proporciona un EFA.

Análisis de senderos, análisis factorial confirmatorio y modelo de ecuaciones estructurales

El análisis de senderos intenta analizar los efectos causales hipotéticos entre un conjunto de variables a través del análisis de regresión de varios pasos, descomponiendo las correlaciones entre las variables en cuatro componentes: efecto directo, efecto indirecto, no analizado (debido a causas correlacionadas) y espurio (debido a causas comunes). Las mayores debilidades del análisis de senderos son que: a) se asume que todas las variables se han medido sin error y b) faltan mecanismos estadísticos para probar el ajuste entre el modelo hipotetizado y los datos empíricos. La suposición de que todas las variables son medidas sin error no es realista, ya que las medidas psicológicas invariablemente contienen una gran cantidad de error de medición. Además, el análisis de senderos describe relaciones causales entre variables empíricas pero no es capaz de estudiar situaciones en las cuales varias variables son indicadores hipotéticos del mismo constructo latente.

El CFA y el análisis de senderos pueden ser incluidos en el SEM. La diferencia entre el análisis de senderos, el CFA y el SEM consiste en que el análisis de senderos examina las relaciones causales entre variables observadas, el CFA examina las relaciones entre variables observadas y constructos latentes (factores), mientras que un SEM completo examina las relaciones causales entre variables observadas y constructos hipotéticos y entre constructos latentes entre sí.

Un ejemplo de análisis de senderos para el estudio de la personalidad se refiere a una investigación realizada para analizar la influencia de la percepción de la relación con los padres sobre la percepción de la relación con los pares y de ambos sobre el estilo de afrontamiento de la amenaza en 1500 adolescentes argentinos (Richaud de Minzi, Sacchi & Moreno, 2003). El diagrama correspondiente aparece en la Figura 2.

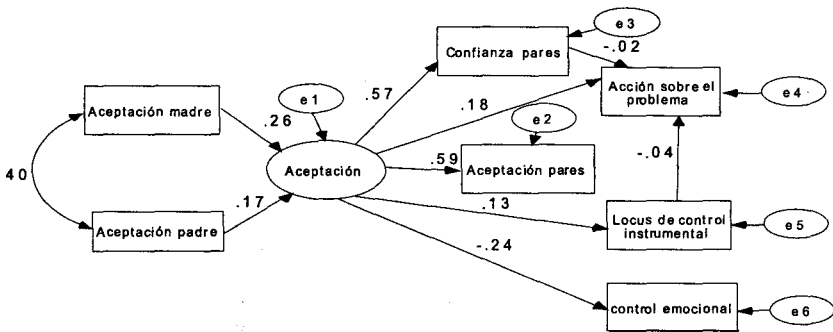


Figura 2. Análisis de senderos de un estudio de variables de personalidad con 1500 adolescentes argentinos.

$$\chi^2 = 17,29 \quad gl = 10 \quad p = .07$$

El χ^2 indica que los datos ajustan al modelo pudiéndose observar la percepción de la aceptación de los padres como predictora de la aceptación y la confianza en los pares y, aunque en menor medida, de estilos adaptativos de afrontamiento. El estudio de estos modelos multicausales no es posible sino a través del SEM. Asimismo, estos resultados confieren validez constructiva a las pruebas desarrolladas para medir percepción de la relación con los padres y percepción de la relación con los pares.

Teoría de la respuesta al ítem (TRI)

Este modelo se aplica explícitamente al caso en que se supone que una dimensión simple subyacente (por ejemplo la “soledad”) genera un conjunto de respuestas (Hambleton, 1989). El modelo, que puede verse como un modelo logístico multivariado con un predictor inobservable simple (o factor), tiene un foco primario en las propiedades del ítem y un foco secundario en la estimación de la persona o valores a lo largo de la dimensión atributo subyacente, llamada θ .

La probabilidad de garantizar que un determinado ítem pertenece a un conjunto, por ejemplo indicando que un ítem de “soledad” es verdadero o auto-descriptivo, es una función de dos componentes: a) características del ítem o parámetros (para un determinado ítem), y b) un parámetro simple (para una determinada persona) que refleja la ubicación que esta persona tiene a lo largo de las dimensiones del atributo subyacente (por ejemplo, “soledad”). Cada ítem de la escala (por ejemplo “no tengo amigos”) puede estar representado gráficamente con una curva del ítem donde la forma de la curva y la ubicación caracteriza las propiedades específicas del ítem. La curva es una función no lineal que toma la forma general de la función cuadrática, siendo las más típicas la curva normal acumulativa (ojiva normal) o funciones logísticas. Originariamente esta teoría ha sido aplicada con los tests de habilidades cognitivas, donde las dimensiones latentes subyacentes q han sido llamadas *habilidades*, tales como verbal, espacial, numérica, etc. En el caso de la personalidad, los mismos conceptos generales de la TRI se pueden aplicar para describir cómo los individuos responden a los ítems de los inventarios, las características de esos ítems y la estimación que cada individuo tiene en las dimensiones de personalidad subyacentes (Panter, Swygart, Dahlstrom & Tanaka, 1997).

Se han desarrollado modelos de la TRI de uno, dos o tres parámetros. En cada caso se elegirá el modelo que más convenga de acuerdo al tamaño de la muestra, el número de ítems a los que se aplicará el modelo y la interpretabilidad de los parámetros en el contexto de los datos.

Para poder implementar la TRI es necesario tener en cuenta dos supuestos. El primero de ellos es que el conjunto de ítems debe ser unidimensional. En el ejemplo, es de suponer que los treinta ítems de la Escala de Soledad están caracterizados por una sola dimensión dominante. Dada una supuesta dimensión única desde el punto de vista teórico (todos estos ítems tienen un contenido que ha sido tradicionalmente asociado a soledad) se puede estudiar empíricamente

a través de un análisis factorial confirmatorio si esa suposición es admisible.

El segundo supuesto a analizar es el de la independencia local de los ítems, que significa que la dependencia observada en las respuestas al ítem debe basarse solamente en su asociación común con el atributo subyacente.

Las ventajas del modelo son: a) las características de los ítems son independientes de las poblaciones específicas a las que se les administran, b) las estimaciones del atributo subyacente son independientes de los ítems específicos administrados, y c) la precisión de la medición puede ser determinada para todas las estimaciones del atributo a lo largo de las dimensiones subyacentes.

El modelo de un parámetro o modelo de Rasch (Rasch, 1960) es el más simple, pero al mismo tiempo el menos descriptivo de la respuesta real al ítem. En este modelo, un parámetro simple, b , describe la probabilidad de que un ítem dado pertenezca al atributo. Este parámetro llamado umbral o dificultad del ítem refleja el punto de inflexión de la curva característica del ítem en el cual la probabilidad de aceptar o rechazar el ítem es exactamente .50. El parámetro indica el punto de transición entre la probabilidad de rechazar el ítem como no autodescriptivo y de aceptar el ítem como autodescriptivo, mientras que los valores menores reflejan alta aceptación o ítems no dificultosos.

El modelo de dos parámetros (Birbaum, 1968) indica cuál es la probabilidad de que un determinado respondiente seleccionado aleatoriamente responda correctamente al ítem. Esta probabilidad, dado su nivel en la dimensión subyacente, es una función del parámetro umbral antes descrito y del parámetro a . Este segundo parámetro refleja la discriminatividad del ítem que es función de la pendiente de la curva del ítem en el punto umbral.

El modelo de los tres parámetros agrega a los dos primeros un tercer parámetro llamado de asíntota más baja c que denota la probabilidad distinta de cero de que aún respondentes con bajos niveles en el atributo subyacente (personas que no sufran de sentimientos de soledad) puedan aceptar un ítem como verdadero (llamado en las pruebas cognitivas parámetro de respuesta al azar).

La utilidad de la TRI en el estudio de la personalidad se hace evidente en el análisis de cuestiones básicas acerca de las respuestas a ítems de escalas de personalidad con respecto a: a) la dimensionalidad de la escala y las propiedades de los ítems, b) la adecuación de las respuestas observadas, y c) ítems diferenciales que funcionan a través de diferentes submuestras.

Un ejemplo de análisis de la dimensionalidad de una escala de personalidad y de las propiedades de los ítems es el citado por Panter et al. (1997) que se refiere al estudio presentado por Steinberg y Thissen (1995) usando datos de la Escala de Acción de Kuhl (1985), que describió el análisis separado de las dimensiones Acción y Pensamiento de orientación acción-estado a través del modelo de dos parámetros de la TRI. Se obtuvo dos conjuntos de curvas características para los ítems, cada una de las cuales mostraba la probabilidad de cada ítem en las dimensiones Acción y Pensamiento como una función de los constructos subyacentes. Los parámetros a de discriminación para los ítems de Pensamiento mostraron que algunos de los ítems de la escala tenían menos relación que otros con el constructo subyacente y los parámetros b de umbral mostraron que también los ítems de Acción diferían entre sí. El examinar las curvas características de cada ítem en estas dimensiones permitió a los investigadores ver cuáles eran los ítems útiles para medir cada dimensión y cuales ítems podían ser eliminados sin disminuir la precisión de la medición.

Otra importante función de los modelos TRI es la que cumplen en la evaluación de patrones de respuesta a los ítems. Los índices de

adecuación basados en la TRI están diseñados para evaluar respuestas anormales a los ítems de la escala, que pueden ocurrir por diversas razones, como el estado de los respondentes y el medio en que se administran los ítems. Estas respuestas producen medidas no válidas del atributo y deben identificarse.

Reise y Waller (1993) examinaron la Escala de Intimidad Social del Cuestionario Multidimensional de Personalidad de Teelegen con el índice de adecuación I_z . Este índice (Drasgow, Levine & Williams, 1985) puede emplearse para identificar respuestas anormales en un inventario de personalidad con ítems dicotómicos, siempre que se satisfagan los supuestos de la TRI (Birembaum, 1985). Estos autores encontraron que el I_z identificaba correctamente a las personas cuya respuesta a los ítems no era consistente con la estimación del rasgo en ellas, y concluyeron que en el caso de los respondentes con bajo I_z (que indica respuestas anormales extremas), la escala de Intimidad Social no parecía medir el mismo constructo que en los respondentes con valores normales o altos en I_z .

Además de los usos anteriores, la TRI es también útil para poner a prueba sesgos o invariancia en los parámetros de los ítems a través de distintos grupos (Reise, Widaman & Pugh, 1993).

Para la TRI la estructura del atributo subyacente es unidimensional y todos los ítems del test son indicadores de ese atributo, mientras que en el CFA, por ejemplo, esa estructura debe estar especificada de antemano. Las dos técnicas permiten poner a prueba formalmente hasta qué punto diferentes aspectos del modelo subyacente son invariantes a través de distintas subpoblaciones como género o estatus clínico.

Un caso específico de sesgo es el del funcionamiento diferencial del ítem (dif), en el que los respondentes pertenecientes a dos grupos tienen diferentes probabilidades de responder igualmente al ítem, aún cuando los dos grupos tienen la misma cantidad del atributo subyacente.

Thissen, Steinberg y Gerrard (1986) trabajaron con el Inventario de Culpa Sexual de Elección Forzada y encontraron que no sólo las mujeres y los varones tenían diferentes valores de media, sino que también algunos ítems tenían diferentes umbrales según el género.

Conclusiones

Sin un adecuado análisis de los datos, los esfuerzos para investigar carecen de rigor científico. Si simplemente se relacionan dos conjuntos de conceptos se alcanza un resultado teórico sin validación. Por otra parte, si sólo se relacionan dos conjuntos de conductas se llega a un resultado empírico sin explicación.

Gracias al desarrollo de nuevos procedimientos de análisis se han obtenido importantes progresos en la evaluación de la personalidad. Estos avances metodológicos parecen haber minimizado el riesgo de los investigadores de detenerse demasiado en los conceptos sin una adecuada atención a las conductas, o de focalizarse exclusivamente en las conductas sin tener en cuenta los conceptos.

El nuevo riesgo que aparece de la mano de estos sofisticados métodos es que se conviertan en objetivos en sí mismos debido al uso indiscriminado de la computadora.

Los nuevos avances en metodología bien utilizados pueden impulsar el estudio y evaluación de la personalidad a límites impensados, pero siempre y cuando éstos sean utilizados críticamente y respondiendo a un diseño previamente establecido y planificado por el cual los resultados tengan un claro sentido teórico.

Referencias

- Aldrich, J.H. & Nelson, F.D. (1984). *Linear probability, logit and probit models*. Beverly Hills, CA: Sage.
- Bagozzi, R.P., Fornell, C. & Larcker, D.F. (1981). Canonical correlation analysis as a special case of structural relations model. *Multivariate Behavioral Research*, 16, 437-454.
- Bentler, P.M. (1985). *Theory and implementation of EQS. A structural equations program*. Los Angeles: BMDP Statistical Software.
- Bentler, P.M. (1992). *EQS: Structural equations manual*. Los Angeles: BMDP Statistical Software.
- Birbaum, A. (1968). Some latent trait models and their uses in inferring an examiner's ability. En F.M. Lord & M.R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 392-479). Reading, MA: Addison Wesley.
- Birembaum, M. (1985). Comparing the effectiveness of several IRT based appropriateness measures in detecting unusual response patterns. *Educational and Psychological Measurement*, 45, 523-533.
- Bollen, K.A. (1989). *Structural equations with latent variables*. Nueva York: Wiley.
- Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, application and programming*. Newbury Park, CA: Sage.
- Crowley, S.L. & Fan, X. (1997). Structural equation modeling: Basic concepts and applications in personality assessment research. *Journal of Personality Assessment*, 68 (3), 508-531.
- Davis, L.J. & Offord, K.P. (1997). Logistic regression. *Journal of Personality Assessment*, 68 (3), 497-507.
- Drasgow, F., Levine, M. & Williams, E.A. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, 38, 67-86.

- Gorsuch, R.L. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hambleton, R.K. (1989). Principles and selected applications of item response theory. En R.L. Linn (Ed.), *Educational measurement* (3ra. ed.) (pp. 147-200). Nueva York: Macmillan.
- Hoskens, M. & De Boeck, P. (1995). Componential IRT models for polytomous items. *Journal of Educational Measurement*, 32, 364-384.
- Hoskens, M. & De Boeck, P. (2001). Multidimensional componential IRT models. *Applied Psychological Measurement*, 25, 19-37.
- Jöreskog, K.G. & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: ABT Books.
- Jöreskog, K.G. & Sörbom, D. (1989). *LISREL7: A guide to the program and application*. Chicago: SPSS.
- Kuhl, J. (1985). Volitional mediators of cognition-behavioral consistency. Self-regulatory processes and action versus state orientation. En J. Kuhl & J. Beckman (Eds.), *Action control: From cognition to behavior* (pp. 101-128). Berlin: Springer-Verlag.
- Loehlin, J.C. (1992). *Latent variable models: An introduction to factor, path and structural analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F.M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Panter, A.T., Swygert, K.A., Dahlstrom, W.G. & Tanaka, J.S. (1997). Factor analytic approaches to personality item-level data. *Journal of Personality Assessment*, 68 (3), 561-589.
- Pedhazur, E.J. & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Reis, H.T. & Stiller, J. (1992). Publication trends in JPSP: A three-decade review. *Personality and Social Psychology Bulletin*, 18, 465-472.

- Reise, S.P. & Waller, N.G. (1993). Traitendness and the assessment of response pattern scalability. *Journal of Personality and Social Psychology*, 65, 143-151.
- Reise, S.P., Widaman, K.F. & Pugh, R.H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114, 552-556.
- Richaud de Minzi, M.C., Sacchi, C. & Moreno, E. (2003). *Recursos y afrontamiento de la amenaza en la adolescencia. Línea de investigación*. IX Reunión Nacional de la Asociación Argentina de Ciencias del Comportamiento, Córdoba, Argentina.
- Schiffman, S.S., Reynolds, M.L. & Young, F.W. (1981). *Introduction to multidimensional scaling: Theory, methods and applications*. Nueva York: Academic Press.
- Schinka, J.A., LaLone, L. & Broeckel, J.A. (1997). Statistical methods in personality assessment research. *Journal of Personality Assessment*, 68 (3), 487-496.
- Steinberg, L. & Thissen, D. (1995). Item response theory in personality research. En P.E. Shrout & S. Fiske (Eds.), *Personality research, methods and theory. A festschrift honoring Donald W. Fiske* (pp. 161-181). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thissen, D., Steinberg, L. & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin*, 99, 118-128.
- Weiner, J. B. (1991). Developments in research in personality assessment. *Journal of Personality Assessment*, 56, 370-372.