



## CARTA AL EDITOR

### Sobre carretas, caballos, y otras cuestiones psicométricas

*About carts, horses, and other psychometric issues*

Sergio Alexis Dominguez-Lara <sup>1\*</sup>

<sup>1</sup> Universidad de San Martín de Porres, Perú

\* Dirección de correspondencia: Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Av. Tomás Marsano 242 (5to piso), Lima 34 – Perú; Teléfono: 0051988053909. Correo: [sdominguezmpcs@gmail.com](mailto:sdominguezmpcs@gmail.com), [sdominguezl@usmp.pe](mailto:sdominguezl@usmp.pe)

Recibido: 21 de agosto de 2018; Revisado: 30 de agosto de 2018; Aceptado: 30 de agosto de 2018; Publicado Online: 01 de septiembre de 2018.

#### CITARLO COMO:

Dominguez-Lara, S. (2018). Sobre carretas, caballos, y otras cuestiones psicométricas. *Interacciones*, 4(3), 199-200. doi: 10.24016/2017.v4n3.152

#### Sr. Editor,

Hace un tiempo le comentaba a un colega cómo algunos avances en el ámbito de la psicometría, concretamente en el análisis factorial, permitían esclarecer (y a veces cuestionar) la configuración de algunos instrumentos de evaluación psicológica ampliamente conocidos y usados en nuestro medio. Al finalizar, lo que atiné a decirme, sin quitar mérito a tales avances, fue que *los métodos y sus resultados deben subordinarse a la teoría*. Cualquier psicólogo, sea de base cuantitativa o no, estaría inicialmente de acuerdo con ello. Sin embargo, hay algunos detalles a precisar, sobre todo lo que concierne a las interpretaciones derivadas de los análisis.

En primer lugar, que los instrumentos de evaluación psicológica buscan cuantificar la *presencia* de un constructo, sea simple (unidimensional) o complejo (multidimensional) mediante una serie de enunciados (ítems) que fueron elaborados siguiendo un proceso riguroso. Ello implica que, si se trata de un constructo simple, los ítems deben ser esencialmente unidimensionales, es decir, recibir influencia significativa de un solo constructo (o variable latente), y no

de dos o tres. Por el contrario, si se trata de uno multidimensional, por ejemplo, de cuatro dimensiones, se espera que exista un grupo de ítems que representen adecuadamente alguna de esas dimensiones, y así sucesivamente hasta que se refleje de forma empírica lo planteado teóricamente, es decir, cuatro factores, y no cinco o seis. Claro está, es un panorama ideal.

Sin embargo, es probable que cuando se analicen los ítems de un instrumento, la configuración no sea la esperada. Esto es algo frecuente, ya que luego de procesar un conjunto de ítems es probable que estos no se agrupen según lo *teóricamente esperado* (si se trata de un análisis factorial exploratorio) o que la configuración predeterminada no alcance índices de ajuste satisfactorios o presente malas especificaciones (si es un análisis factorial confirmatorio). En la mayoría de esos casos no se puede *culpar* al método, asumiendo que la elección del proceso analítico fue adecuada, ya que lo que en última instancia influye sobre la configuración de un instrumento es *la forma* cómo los ítems fueron respondidos. Por ejemplo, si predomina un estilo de respuesta positivo o negativo, es natural que se configure un solo factor (denominado, *factor de dificultad*).



Además, y pese a ser un tema poco abordado al menos por este lado del mundo, influye sobremanera el hecho de que muchos de los instrumentos utilizados son oriundos de otros contextos culturales. Por ello, es *natural* que al ser empleados y analizados desde un enfoque analítico-factorial su estructura no converja con la original (repito, creada en otro contexto). Esto no quiere decir que el instrumento *funcione mal*, o que *la teoría está mal hecha*. Simplemente, que los ítems *no capturan el constructo* del mismo modo en una cultura distinta a la original. En ese caso, el método no está *en contra* de la teoría; solo ayuda a comprender mejor la estructura de un instrumento de medición. Esta situación puede derivar en configuraciones factoriales distintas, ítems eliminados, o reducción de factores, que está asociado a una falta de diferenciación empírica. Eso último se vincula al siguiente punto.

En segundo lugar, se encuentra el significado de una *correlación entre constructos*. Queda claro que si un investigador construye una escala de 10 ítems para medir dos constructos, espera que los ítems se distribuyan según lo teóricamente esperado. Imaginemos que, siguiendo un análisis factorial confirmatorio, los índices de ajuste son adecuados y las cargas factoriales al menos moderadas, pero la correlación entre los dos factores es .95. ¿Es esto aceptable? Una respuesta que he recibido frecuentemente cuando hice estos cuestionamientos es que *si la teoría lo indica deben ser interpretados de forma independiente*. Es decir, no importa que los dos (o quizás más) factores compartan más del 90% de variabilidad. Esto se configura como un error de interpretación (y de especificación), ya que pese a que ambos factores evalúan (casi) lo mismo, se estarían interpretando por separado *dos cosas* (muy) similares. Por otro lado, desde un análisis factorial exploratorio esto (e.g., extraer dos factores cuando solo es un factor) aparentemente no sería un problema ya que si se conduce apropiadamente debería extraerse solo un factor.

Entonces, ¿el *método* contradice la *teoría*? Al parecer, muy por el contrario, ayuda a entender a cabalidad el funcionamiento del instrumento y sugiere corregir algunos problemas con la redacción de los ítems (e.g., ítems ambiguos). Además, durante muchos años, previo al *boom* de los análisis confirmatorios, fueron preferidos métodos de rotación ortogonales (e.g., varimax), cuya implementación omite el reporte de la correlación interfactorial, de lo que se puede deducir (es una especulación, claro está) que es *una forma de hacer las cosas* que se ha mantenido en el tiempo.

En tercer lugar, que las *puntuaciones totales*, esto es, la sumatoria de todos los ítems de una escala multidimensional, debe estar debidamente justificada. Muchos constructos complejos (e.g., inteligencia emocional) han sido evaluados con escalas multidimensionales, pero que con fines de preservar la naturaleza jerárquica del constructo estas puntuaciones parciales se adicionan para brindar una puntuación general. Si bien últimamente se han re-descubierto métodos orientados a la evaluación de este punto (e.g., modelos jerárquicos), por muchos años esto se ha interiori-

zado en el investigador *senior* o *novel* de modo tal que el lector puede revisar artículos empíricos o instrumentales y notará que con mucha frecuencia se obtiene una puntuación total, aun cuando no exista evidencia empírica de ello. Nuevamente, el *método* empleado ayuda a conocer de forma más profunda la configuración jerárquica de la escala y así evitar interpretaciones basadas en suposiciones sin correlato empírico, por más que estén fundadas teóricamente. Con eso no quiero sonar *antiteoría*, pero es bueno recalcar que si la evidencia indica que, por ejemplo, las cuatro dimensiones de una escala están relacionadas, pero la magnitud de esa relación no permite postular la presencia de un factor general que (supuestamente) es la influencia principal sobre todos los ítems, ¿merece la pena ser obstinado e interpretar de forma arbitraria dicho factor general? Y si se hace, ¿cómo se interpretaría si no recoge la suficiente variabilidad para *vivir por sí mismo*?

Para ir finalizando, la postura que comenté al inicio suena como una versión del conocido refrán *no poner la carreta delante del caballo*, donde la carreta son *los métodos y resultados*, y el *caballo*, la *teoría*. Por recoger uno de los ejemplos previos, su aplicación legitimaría la interpretación de dos factores como independientes, a pesar de que evalúan lo mismo, *porque la teoría así lo indica*. Pero, a la luz de los avances referidos (y que, por cierto, no son los únicos), ¿vale la pena dejar que ese refrán guíe el quehacer psicométrico? La respuesta queda en cada uno, pero está claro que el ritmo metodológico en psicometría y análisis de datos (y creo que en la psicología en general) es más rápido que hace 10 o 20 años. Incluso, el lector interesado puede ver los contenidos de los últimos 10 años de las revistas iberoamericanas de mayor impacto (en SCOPUS) y verá que la complejidad analítica ha crecido exponencialmente, ya sea un estudio empírico o instrumental. No se puede estar ajeno a este avance.

En conclusión, estar alerta a esos avances no implica, necesariamente, perder de vista el aspecto teórico. Todo lo contrario, ya que así pueden ponerse a prueba hipótesis teóricas y de medición de mayor complejidad que antes solo era posible contrastarlas parcialmente, o incluso no contrastarlas, dejando terreno fértil para la especulación o, en el mejor de los casos, para el planteamiento de hipótesis teóricamente razonables pero empíricamente insuficientes. Todo esto, como se puede apreciar, enriquece la *teoría* y, en consecuencia, la ciencia psicológica.

## CONFLICTO DE INTERÉS

El autor declara no tener conflictos de interés al momento de realizar el manuscrito.

## FINANCIAMIENTO

El presente estudio fue autofinanciado.