

## Dos versiones breves del Big Five Inventory en universitarios peruanos: BFI-15p y BFI-10p

### Two brief versions of the Big Five Inventory in Peruvian university students: BFI-15p and BFI-10p

Sergio Dominguez-Lara<sup>a,\*</sup>, César Merino-Soto<sup>a</sup>

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

Recibido: 29 de enero de 2018

Aceptado: 02 de mayo de 2018

#### Resumen

El presente estudio tuvo como objetivo obtener dos versiones breves (de 15 y 10 ítems) del *Big Five Inventory* (BFI) con bases empíricas y conceptuales. Participaron 332 universitarios peruanos (82.83% mujeres) entre 16 y 48 años ( $M = 20.39$ ). El análisis principal fue realizado mediante el método de mínimos cuadrados no ponderados y rotación procusteana oblicua con una matriz completamente especificada. En cuanto a los resultados, las dos versiones breves mostraron propiedades psicométricas robustas: estructura factorial adecuada y evidencias de equivalencia empírica con la versión extensa. Finalmente, las dos versiones (15 y 10 ítems) mostraron índices de confiabilidad aceptables ( $\alpha_{\text{promedio}} > .60$ ). En conclusión, las versiones estudiadas evidencian indicadores psicométricos favorables que posibilitarían su uso en evaluaciones de grupo e investigación básica.

**Palabras clave:** Cinco Grandes Factores (BFI), rotación procusteana, estudiantes universitarios, validez, confiabilidad.

#### Abstract

The aim of this study was to create two short versions (15 and 10 items) of the Big Five Inventory (BFI) with empirical and theoretical bases. Participants included 332 Peruvian university students (82.83% women), aged between 16 and 48 years ( $M = 20.39$ ). The main analysis was performed using the unweighted least squares method with oblique Procrustes rotation and fully specified matrix. The analysis results showed that both brief versions had robust psychometric properties: an adequate factor structure and empirical equivalence with the original BFI. Finally, the two versions displayed acceptable reliability indexes as well ( $\alpha_{\text{average}} > .60$ ). In conclusion, both versions show favorable psychometric indicators that would enable their use in basic and group research.

**Keywords:** Big Five Inventory (BFI), Procrustes rotation, university students, validity, reliability.

Para citar este artículo:

Dominguez-Lara, S., & Merino-Soto, C. (2018). Dos versiones breves del Big Five Inventory en universitarios peruanos: BFI-15p y BFI-10p. *Liberabit*, 24(1), 81-96. doi: 10.24265/liberabit.2018.v24n1.06

Este es un artículo Open Access bajo la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0



## Introducción

El *Big Five Inventory* (BFI; John, Donahue, & Kentle, 1991) es un instrumento creado para evaluar los cinco componentes del modelo que le da el nombre (Cinco Grandes factores - 5GF) y es uno de los más usados por su brevedad y por las evidencias de estructura interna que se replican satisfactoriamente en distintos contextos culturales (Benet-Martínez & John, 1998; Carciofo, Yang, Song, Du, & Zhang, 2016; Denissen, Greenen, van Aken, Gosling, & Potter, 2008; Domínguez-Lara, Merino-Soto, Zamudio, & Guevara, en prensa; Fossati, Borroni, Marchione, & Maffei, 2011; Gurven, von Rueden, Massenkoff, Kaplan, & Lero Vie, 2013; Kim et al., 2010; Lovik, Verbeke, & Molenberghs, 2017; Pejić, Tenjović, & Knežević, 2014; Reyes, Álvarez, Peredo, Miranda, & Rebolledo, 2014; Ubbiali, Chiorri, Hampton, & Donati, 2013; Worrell & Cross, 2004). La relativa predominancia de su uso puede explicarse debido a su parsimoniosa extensión y fraseo, por ejemplo, al redactarse los ítems del BFI con frases elaboradas y contextualizadas (e.g., ítem 3: «Es minucioso en el trabajo», es más probable obtener respuestas consistentes que mostrando solo adjetivos (e.g., «Minucioso»)) (Goldberg & Kilkowski, 1985). Este último, podría interpretarse libremente en ausencia de mayor información.

El BFI también posee utilidad descriptiva predominantemente intercultural (McCrae & Terracciano, 2005; Schmitt, Allik, McCrae, & Benet-Martínez, 2007; Yamagata et al., 2006), aunque enfatizada en sociedades con características comunes como el nivel educativo o de ingresos (Henrich, Heine, & Norenzayan, 2010; Rammstedt, Kemper, Klein, Beierlein, & Kovaleva, 2013). Por otro lado, esta característica es también su limitación porque en otros contextos la replicación no es exitosa y requiere modificaciones que alteran no solo el instrumento, sino la comprensión de la conducta evaluada (Fetvadjiev & van der Vijver, 2015; Schmitt et al., 2007). Una consecuencia es que se han planteado modelos alternativos que respondan mejor a la idiosincrasia

cultural (Cheung, van der Vijver, & Leong, 2011; Gurven et al., 2013).

El BFI tiene ventajas como su relativa brevedad (44 ítems), con respecto a otros instrumentos como el NEO-PI-R y NEO-FFI (Costa & McCrae, 1999), su distribución libre (no genera un costo monetario al usuario) y su versión hispana neutral (sin regionalismos), cuyos ítems se comprenden independientemente del lugar de procedencia de la persona. Sin embargo, existen contextos en los cuales se requieren medidas aún más breves, como evaluaciones masivas en encuestas nacionales, procesos de selección masiva y despistaje de perfiles de conducta específicos. En este caso, es necesario considerar la posibilidad de extraer versiones cortas para facilitar el uso del modelo 5GF y así completar descripciones de conducta que no posean restricciones comerciales (Goldberg et al., 2006).

Otras ventajas adicionales de las escalas breves sobre las extensas es que reduce la fatiga del evaluado, es útil en contextos en los que no se cuenta con mucho tiempo para evaluar (e.g., empresas) y sirve a su vez para fines de investigación cuyo objetivo principal no es la personalidad. Del mismo modo, puede usarse en docencia para la demostración de propiedades psicométricas y el tiempo de calificación es reducido. Sin embargo, hay dos limitaciones inherentes: la reducción de la confiabilidad en comparación con la versión extensa del mismo instrumento (Saucier, 1994) y la reducción de amplitud del dominio de contenido del constructo. Ambas limitaciones pueden restringir su uso cuando se requieren decisiones con medidas altamente confiables y una más amplia representación del constructo.

Respecto a las versiones breves del BFI, los estudios previos han señalado propiedades psicométricas adecuadas de una versión de diez ítems, BFI-10 (Rammstedt, 2007; Rammstedt & John, 2007). Este instrumento está conformado, para cada dimensión, por un ítem en escala directa y otro en escala invertida. Esta propuesta enfatizó la evaluación

de los polos opuestos de los constructos; sin embargo, su construcción mostró algunas limitaciones metodológicas y omisiones en el reporte de los resultados. Por un lado, a) no evaluó el problema de los ítems en escala invertida (polo opuesto del factor), generalmente denominado *efecto del método* (Lance, Dawson, Birkelbach, & Hoffman, 2010; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), y parece ubicuo cuando un instrumento contiene ambos tipos de fraseo (dirección positiva y negativa respecto al constructo); por otro, b) la selección de ítems fue complementada con correlaciones bivariadas entre ítems, pero sin controlar la multidimensionalidad de estos; así también, c) la correlación entre el BFI-10 y su versión extensa no fue corregida por errores correlacionados o varianza común espuria ocasionada por los ítems comunes en ambas (Levy, 1967); y, por último, d) no se reporta la confiabilidad de las subescalas del BFI-10, lo cual limita el conocimiento de la precisión de la medida. Estas limitaciones restringen las interpretaciones de validez obtenidas, pues el BFI-10 contiene varianza irrelevante al constructo proveniente de varias fuentes propias del diseño del estudio y con información faltante no se puede evaluar eficazmente la replicabilidad de sus resultados.

El reporte problemático de la confiabilidad en la evaluación del BFI-10 parece extenderse más allá del estudio primigenio. Por ejemplo, en un estudio posterior, Guido, Peluso, Capestro y Miglietta (2015) incorporaron el modelamiento de ecuaciones estructurales en la evaluación psicométrica del BFI-10, encontrando índices de ajuste apropiados y cargas factoriales de moderada magnitud. Sin embargo, respecto a la confiabilidad de las puntuaciones, su estimación se hizo con el coeficiente por mitades (Brown, 1910; Spearman, 1910), un procedimiento apropiado para medidas psicométricamente paralelas (Charter, 2001; Feldt & Charter, 2003) y que produce coeficientes espuriamente elevados en situaciones contrarias. A pesar de ello, parece replicable en otras poblaciones culturales europeas como las citadas por Rammstedt et al. (2013); pero en muestras con

características distintas, como adultos mayores de nacionalidad coreana como las citadas en Kim et al. (2010), la versión completa tuvo indicadores bajos de confiabilidad evaluada con el coeficiente alfa ( $\alpha_{\text{extraversión}} = .52$ ;  $\alpha_{\text{afabilidad}} = .61$ ;  $\alpha_{\text{responsabilidad}} = .75$ ;  $\alpha_{\text{neuroticismo}} = .71$ ;  $\alpha_{\text{apertura}} = .75$ ), por lo que se esperaba una reducción en el BFI-10. Un panorama similar se aprecia en la versión holandesa (Lovik et al., 2017), donde los indicadores fueron bastante bajos ( $\alpha_{\text{extraversión}} = .42$ ;  $\alpha_{\text{afabilidad}} = .09$ ;  $\alpha_{\text{responsabilidad}} = .39$ ;  $\alpha_{\text{neuroticismo}} = .62$ ;  $\alpha_{\text{apertura}} = .28$ ); así como en la versión serbia ( $\alpha_{\text{extraversión}} = .64$ ;  $\alpha_{\text{afabilidad}} = .21$ ;  $\alpha_{\text{responsabilidad}} = .52$ ;  $\alpha_{\text{neuroticismo}} = .57$ ;  $\alpha_{\text{apertura}} = .31$ ) (Pejić et al., 2014). El TIPI (Ten-Item Personality Inventory; Gosling, Rentfrow, & Swann, 2003) es otra medida de 10 ítems, pero no parece ser replicable en culturas de habla no anglosajona (e.g., Carvalho, Nunes, Primi, & Nunes, 2012).

En la línea de la construcción de versiones breves del BFI, existe una de 15 ítems (tres por dimensión) elaborada para la población alemana (Gerlitz & Schupp, 2005) y con criterios empíricos a partir de una base de 25 ítems del BFI (BFI-25). Sin embargo, se utilizó el procedimiento denominado *Little Jiffy* (Análisis de Componentes Principales y rotación Varimax), cuya validez metodológica está más que cuestionada en el desarrollo de medidas en las ciencias conductuales (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014). Sus estimaciones de confiabilidad  $\alpha$  fueron aceptables (entre .50 y .73), considerando que es una medida de amplia cobertura conceptual y de pocos ítems. Sin embargo, las correlaciones con la versión larga (BFI-25) no fueron corregidas por varianza espuria (Levy, 1967), y, por lo tanto, estuvieron sobrestimadas. Posteriormente, esta versión fue validada estructuralmente y con criterios externos en población alemana (Hahn, Gottschling, & Spinath, 2012), obteniendo resultados teóricamente consistentes.

De los estudios anteriores se deriva que cada versión breve construida tiene una fuerte influencia cultural, es decir que los ítems seleccionados identifican

mejor a un factor en una determinada población. En el contexto peruano, la versión extensa desarrollada por Dominguez-Lara et al. (en prensa) ha mostrado evidencias aceptables de validez con respecto a su estructura interna, pero el número desigual de los ítems en sus factores no provee un adecuado y equitativo balance del contenido de los constructos evaluados.

En este sentido, el objetivo del presente reporte fue derivar dos versiones breves (de 15 y 10 ítems) con bases empíricas y conceptuales de la versión BFI trabajada en Perú (Dominguez-Lara et al., en prensa), ya que aparentemente hasta la fecha no existe en español alguna versión abreviada del BFI que sea sensible a las posibles variaciones culturales. Estas variaciones culturales pueden corresponder no solamente al fraseo del contenido, sino también a la definición del constructo, específicamente en la configuración factorial y la métrica del mismo (Cheung & Rensvold, 2002). Como ocurrió en otros estudios que identificaron una versión breve (Gerlitz & Schupp, 2005; Rammstedt & John, 2007), estas no poseen ítems similares. Esto toma en cuenta que las versiones abreviadas del BFI poseen un número estable de ítems (dos o tres) por factor, lo que evitaría que alguno de ellos se vea menos representado en función de la cantidad de ítems. Asimismo, de acuerdo a la revisión efectuada de los estudios de las versiones BFI breves, el presente estudio buscó incorporar mejoras metodológicas para reducir resultados sesgados. Finalmente, para las dos versiones estudiadas, se realizó un reporte de confiabilidad de sus puntuaciones observadas y variables latentes.

## Método

### Participantes

La población estuvo formada por adultos de Lima Metropolitana (Perú) con instrucción superior. La muestra incidental fue conformada por 332 estudiantes (82.83% mujeres) con edades comprendidas entre 16 y 48 años ( $M = 20.39$ ,  $DE = 3.97$ ) y que cursaban del 1<sup>er</sup> al 7<sup>mo</sup> ciclo de la carrera de Psicología de una

universidad privada ubicada en Lima Metropolitana. Fue usado un muestreo no probabilístico en función del acceso de los investigadores. Todos los participantes mostraron disponibilidad y ninguno rechazó su inclusión en el estudio.

Para desarrollar el presente estudio, se contó con la base de datos de un estudio previo (Dominguez-Lara et al., en prensa), para el cual consignaron los permisos respectivos, tanto a las autoridades, docentes y estudiantes, así como la aprobación del estudio por parte del Instituto de Investigación de Psicología de la universidad. Los participantes fueron informados sobre los objetivos de la investigación y su carácter voluntario; quienes accedieron firmaron el formulario de consentimiento informado. Los estudiantes no fueron recompensados académica ni económicamente. La aplicación se realizó en las aulas de clase en horario regular de estudio.

### Instrumentos

Se utilizó el *Big Five Inventory* (BFI; John et al., 1991). Fue utilizada la versión adaptada de Dominguez-Lara et al. (en prensa), basada en la de Benet-Martínez y John (1998). Esta es una medida de autoinforme que mediante 37 ítems evalúa los 5GF: Extraversión (8 ítems), Afabilidad (4 ítems), Responsabilidad (8 ítems), Neuroticismo (8 ítems) y Apertura (9 ítems). Sus puntuaciones se obtienen por la suma simple de los ítems: a mayor puntuación, mayor presencia del atributo en el evaluado. Cada ítem está escalado en un formato ordinal de cinco puntos que va de *Muy en desacuerdo* hasta *Muy de acuerdo* y en las instrucciones se solicita al evaluado indicar el grado de acuerdo con cada uno de los ítems.

El *Big Five Inventory-15* (Gerlitz & Schupp, 2005) y el *Big Five Inventory-10* (Rammstedt & John, 2007) fueron construidos a partir de los ítems del BFI. Estos evalúan cada uno de los Cinco Grandes factores con tres y dos ítems, respectivamente.

## Procedimiento

**Estrategia de análisis.** Fue utilizada la estrategia analítica del trabajo de Dominguez-Lara et al. (en prensa). El método de análisis factorial fue el de Mínimos Cuadrados no Ponderados (MCNP) con base en la matriz de correlaciones policóricas inter-ítem (Dominguez-Lara, 2014) y la rotación procusteana completamente especificada oblicua (Browne, 1972). El método requiere construir una *matriz objetivo* (TM, por su significado en inglés, *target matrix*) que defina la relación hipotetizada de los ítems con los factores. Para esto, se consideró una carga factorial por ítem de .90 en el factor correspondiente y de .15 en el factor secundario (no correspondiente). La elección de esta configuración de la TM estuvo guiada por el análisis ejecutado en la versión extensa (Dominguez-Lara et al., en prensa). Para el análisis, se aplicó el programa FACTOR (Lorenzo-Seva & Ferrando, 2013). Para determinar la aproximación de la TM con la matriz factorial resultante, se usó el coeficiente de congruencia factorial  $\Phi$  (Lorenzo-Seva & ten Berge, 2006; Tucker, 1951), considerando magnitudes entre .80 (ten Berge, 1986) y .95 (Lorenzo-Seva & ten Berge, 2006; Ortet et al., 2007), como aceptables. Asimismo, se implementó el *Índice de Simplicidad Factorial* (ISF; Fleming & Merino, 2005; Kaiser, 1974) para valorar el grado en que un ítem es explicado por su constructo principal (Merino & Grimaldo, 2011).

Mediante estos procedimientos, fueron evaluadas las versiones BFI-15 (Gerlitz & Schupp, 2005) y BFI-10 (Rammstedt & John, 2007). De acuerdo a los resultados, fueron propuestas dos versiones breves, BFI-10p y BFI-15p, de dos y tres ítems por factor, respectivamente. Los criterios utilizados fueron:

**Dos y tres ítems por factor.** La literatura indica que el número mínimo de ítems para estudios analítico factoriales es de tres ítems por factor cuando el tamaño muestral es de al menos 200 sujetos (Fabrigar, Wegener, MacCallum, & Strahan 1999; Ferrando & Anguiano-Carrasco, 2010; Lloret-Segura et al., 2014), pero existen medidas ultrabreves que evalúan con dos

ítems cada uno de los 5GF (Gosling et al., 2003; Rammstedt & John, 2007). Por lo tanto, a pesar de sus limitaciones inherentes, es viable postular una versión de dos ítems por factor.

**Los ítems deben estar en redacción directa al constructo (a mayor puntuación, mayor presencia del constructo).** Con una cantidad pequeña de ítems y con la mitad de ellos en escala invertida, es probable que se introduzca varianza irrelevante al constructo, específicamente debido a la inclusión de ítems invertidos (Lance et al., 2010; Podsakoff et al., 2003). En consecuencia, es preferible que los ítems incorporados estén redactados en escala directa.

**Debe poseer un coeficiente de congruencia adecuado ( $\Phi > .80$ ) en el estudio previo del BFI.** Al especificarse de forma preliminar una TM, se modela hipotéticamente la organización de los ítems con base a un modelo teórico, en esta situación, los 5GF (Goldberg, 1990). Por lo tanto, como una medida de ajuste, si el ítem presenta un  $\Phi$  elevado, se deduce que se aproxima al modelo hipotético inicial.

**Debe presentar fuerte simplicidad factorial (ISF  $> .80$ ) en el estudio previo del BFI.** Aun cuando los factores del 5GF retienen algún monto de varianza compartida, se espera que los ítems resultantes de la versión breve sean *marcadores* del factor que pretenden evaluar, es decir que presenten una carga factorial más elevada que la de los otros ítems en el mismo factor. Por lo tanto, la simplicidad factorial de los ítems debe indicar, aunque valores cercanos a la unidad en el ISF no sean realistas, valores mínimamente elevados que aseguren que las cargas factoriales de los ítems sean de mayor magnitud en el propio factor y menor que en los demás factores.

De este modo, en primera instancia fueron extraídos tres ítems por cada factor siguiendo los criterios antes mencionados. Entre estos, fueron elegidos dos de los ítems factorialmente más representativos (cargas factoriales más fuertes) para configurar la versión de dos ítems por factor. No

obstante, dado que el programa FACTOR no permite un análisis factorial con dos ítems por factor, este se llevó a cabo en el programa IBM SPSS versión 23.

En este último caso, fue especificada la selección de cinco factores con rotación oblicua Promax, a fin de conservar los criterios de oblicuidad delimitados al principio del manuscrito. Asimismo, el factor de oblicuidad ( $k$ ) elegido para Promax fue 4, ya que tiene mayor respaldo en estudios de simulación (Tataryn, Wood, & Gorsuch, 1999).

**Confiabilidad por consistencia interna.** Se utilizó el coeficiente  $\alpha$  y sus intervalos de confianza (IC; Domínguez-Lara & Merino-Soto, 2015) con el método de Fisher (1950). Este método es el que brinda valores más estables independientes del número de ítems (Romano, Kromrey, & Hibbard, 2010; Romano, Kromrey, Owens, & Scott, 2011). Para las medidas de dos ítems se empleó el coeficiente de Angoff-Feldt ( $r_{AF}$ ; Domínguez-Lara, Merino-Soto, & Navarro-Loli, 2016; Feldt & Brennan, 1989), dado que reduce el impacto de la heterogeneidad de varianzas de los ítems. Se espera una magnitud  $\geq .70$  (Merino-Soto, Navarro-Loli, & García, 2014) y  $\geq .60$  para medidas breves y ultrabreves (Ponterotto & Charter, 2009). Del mismo modo, se calculó el  $\alpha_{\text{ordinal}}$  (Domínguez-Lara, 2018; Elosua & Zumbo, 2008), el coeficiente  $\omega$  (McDonald, 1999) para conocer la confiabilidad de los constructos (Domínguez-Lara, 2016) y el coeficiente  $H$  para estimar la replicabilidad del constructo (Hancock & Mueller, 2001). Este conjunto de estimaciones permite evaluar la confiabilidad en el nivel de variables latentes ( $\alpha_{\text{ordinal}}$ ,  $\omega$  y  $H$ ) y la confiabilidad de las puntuaciones de medidas congénicas de dos ítems ( $r_{AF}$ ; Domínguez-Lara et al., 2016) para la versión de 10 ítems.

**Equivalencia entre la versión más extensa y la más breve.** Las evidencias de equivalencia fueron evaluadas mediante correlaciones bivariadas entre versiones y comparación de coeficientes  $\alpha$  (estimación del error de medición). Respecto al primer procedimiento, se correlacionó cada dimensión

correspondiente entre ambas versiones (breve y extensa) como una medida fundamental de validez interna. Esta correlación se corrigió por varianza espuria debido a que contienen ítems comunes (Levy, 1967; Bashaw & Anderson, 1967). Las correlaciones obtenidas (antes y después de la corrección) fueron comparadas con el estadístico  $q$  (Cohen, 1992) para estimar la magnitud de sus diferencias ( $\leq .10$  = trivial;  $> .10$  = pequeño;  $> .30$  = moderado;  $> .50$  = grande). En segundo lugar, se compararon estadísticamente los coeficientes de ambas versiones (extensa y breve) mediante el método de Feldt (Feldt, 1980; Feldt, Woodruff, & Salih, 1987) con el programa *AlphaTest* (Lautenschlager & Meade, 2008). Para este último análisis, se usó la correlación corregida por espuriedad.

## Resultados

**BFI-15** (Gerlitz & Schupp, 2005). El índice de Kaiser-Meyer-Olkin (KMO) para la matriz de correlaciones fue aceptable (.717) y el test de esfericidad de Bartlett estadísticamente significativo ( $p < .001$ ). No hubo presencia de multicolinealidad y la magnitud del coeficiente de Mardia (16.037) apoyó el supuesto de normalidad multivariada (Rodríguez & Ruiz, 2008). Los cinco factores explicaron de forma conjunta el 62.43% de la varianza total. Respecto a sus ítems, sus coeficientes de congruencia ( $\Phi$ ) fueron adecuados en su mayoría y marginales a nivel de factor y de la escala completa (.852); por otro lado, los ISF fueron bastante bajos. Finalmente, la confiabilidad indica fuertes magnitudes de error de medición, siendo tres de ellos menores que .50 (Tabla 1). En conjunto, estos resultados no parecen ser satisfactorios.

**BFI-10** (Rammstedt & John, 2007). No fue posible un análisis con rotación procustea debido a limitaciones computacionales. En este caso, el procedimiento se llevó a cabo con el programa IBM SPSS 14. De forma similar al procedimiento previo, la matriz de correlaciones Pearson evidenció un KMO de .556 y el test de Bartlett fue significativo. Los cinco

factores explicaron el 66.68% de la varianza total, pero los coeficientes de congruencia de los ítems no se adecuaron a la estructura hipotetizada y, consecuentemente, no se calcularon sus coeficientes de confiabilidad basados en variables latentes. Finalmente, el reporte de la confiabilidad de las puntuaciones mostró que cuatro de los cinco coeficientes  $\alpha$  fueron  $< .30$  (ver Tabla 1).

De acuerdo con los indicadores psicométricos obtenidos, la configuración de las dos versiones evaluadas (BFI-15 y BFI-10) no fue satisfactoria para ser utilizada en el contexto peruano. Por ello, se propusieron dos nuevas versiones de 15 (BFI-15p) y 10 ítems (BFI-10p) más apropiadas y resultantes de la validación peruana (Dominguez-Lara et al., en prensa).

### Evidencias de estructura interna para nuevas versiones del BFI-15p y BFI-10p

**BFI-15p.** El análisis descriptivo preliminar (Dominguez-Lara et al., en prensa) indicó que la asimetría y curtosis están dentro de lo esperado ( $\pm 1.5$ ; Pérez & Medrano, 2010), aunque la escala Afabilidad presenta exceso de curtosis. La matriz de correlaciones presentó un KMO de .795 (aceptable) y el test de esfericidad de Bartlett, estadísticamente significativo ( $p < .001$ ). No hubo presencia de multicolinealidad y la magnitud del coeficiente de Mardia (21.902) respaldó el supuesto de normalidad multivariada. Los cinco factores explicaron de forma conjunta el 65.44% de la varianza total. Las correlaciones interfactoriales sugieren fuerte independencia entre los componentes del BFI-15p. Los  $\Phi$  de los factores e ítems fueron adecuados en la mayoría ( $> .80$ ) y la congruencia total fue de .95, valor que implica similitud estadística (Lorenzo-Seva & ten Berge, 2006) entre la matriz obtenida y la TM (ver Tabla 3). Comparativamente, los  $\Phi$  y los ISF fueron mayores que la versión completa (ver Tabla 2).

**BFI-10p.** De forma similar al procedimiento previo, la matriz de correlaciones Pearson resultó factorizable (KMO = .737; test de Bartlett

significativo). Los cinco factores explicaron el 75.25% de la varianza total. Los coeficientes de configuración mostraron magnitudes entre moderadas y altas, y los ISF indicaron que los ítems presentan cargas fuertes en el propio factor (ver Tabla 2).

### Equivalencia entre BFI-37, BFI-15p y BFI-10p

Todas las correlaciones no corregidas fueron  $> .70$ , un valor recomendado como mínimamente aceptable para concluir sobre la equivalencia entre formas completas y breves (Petrides, Jackson, Furnham, & Levine, 2003; Putnam & Rothbart, 2006). Luego de aplicar la corrección por espuriedad, estas correlaciones fueron  $< .70$ ; sin embargo, al examinar la diferencia estandarizada entre las correlaciones ( $q$ ), se observa que la reducción fue trivial para el BFI-10p y pequeña para el BFI-15p (ver Tabla 3). Por otro lado, con respecto a la estimación del error de medición, se observa una reducción estadísticamente significativa de la confiabilidad de las puntuaciones al pasar de la versión extensa a las breves (Tabla 2).

### Análisis de Consistencia Interna

En la Tabla 2 aparecen las estimaciones de consistencia interna con los coeficientes  $\alpha$ ,  $\omega$ ,  $\alpha_{\text{ordinal}}$  y  $H$ . Para las dos versiones breves del BFI la confiabilidad de las puntuaciones fue adecuada para propósitos de investigación (Merino-Soto et al., 2014; Ponterotto & Charter, 2009). En todos los casos, los valores de los coeficientes de confiabilidad basados en variables latentes ( $\omega$ ,  $\alpha_{\text{ordinal}}$  y  $H$ ) fueron adecuados (Gadermann, Guhn, & Zumbo, 2012; Hancock & Mueller, 2001), lo que brinda evidencias adicionales de la consistencia de la evaluación.

### Discusión

El objetivo del presente reporte fue derivar dos versiones breves de 15 y 10 ítems de la adaptación peruana del BFI, teniendo como marco la evaluación psicométrica de dos versiones desarrolladas en contextos culturales diferentes, el BFI-10 (Rammstedt & John, 2007) y BFI-15 (Gerlitz & Schupp, 2005).

**Tabla 1**  
Análisis estructural y de confiabilidad del BFI-15 (Gerlitz & Schupp, 2005) y BFI-10 (Rammstedt & John, 2007)

Ítems <sup>a</sup>	Versión de 15 ítems (Gerlitz & Schupp, 2005)					Versión de 10 ítems (Rammstedt & John, 2007)								
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	$\Phi_1$	ISF	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	ISF	
BFI1	<b>1.003</b>	.258	.459	.374	.096	.952	.639	BFI6	.152	.048	-.164	.067	-.443	.726
BFI6	<b>.483</b>	-.027	-.095	.034	-.234	.745	.744	BFI43	.917	.006	-.010	.038	-.222	.930
BFI43	<b>.784</b>	.410	.219	.266	.350	.930	.511							
BFI22	.121	<b>.771</b>	.289	.166	-.158	.924	.752	BFI2	-.018	-.130	-.376	.323	.149	.486
BFI28	.079	<b>.171</b>	-.065	.039	.217	.690	.502	BFI24	-.322	.121	.039	.182	-.047	.614
BFI37	.348	<b>.697</b>	.206	.280	.385	.913	.455							
BFI3	.067	-.032	<b>.740</b>	.330	.258	.938	.703	BFI3	-.006	-.003	.660	.144	.164	.878
BFI25	-.027	.228	<b>.280</b>	-.137	-.037	.694	.503	BFI25	-.062	-.026	.128	.720	-.099	.930
BFI29	.203	.208	<b>.495</b>	.172	.450	.844	.410							
BFI9	.086	-.174	-.005	<b>.440</b>	-.197	.715	.659	BFI9	.131	.315	-.087	-.012	.394	.547
BFI26	.264	.308	.491	<b>.718</b>	.103	.897	.486	BFI38	.117	-.759	.023	-.062	.117	.935
BFI38	.045	.322	-.100	<b>.816</b>	.197	.936	.772							
BFI5	.119	.142	.327	.077	<b>.663</b>	.958	.704	BFI20	.354	.015	.329	.018	.174	.507
BFI17	-.106	.130	.146	-.064	<b>.564</b>	.906	.821	BFI44	.009	.317	.116	-.138	.053	.686
BFI20	.196	.166	.206	.080	<b>.625</b>	.978	.719							
$\Phi_F$	.927	.816	.784	.899	.828									
F <sub>1</sub>	1													
F <sub>2</sub>	-.377	1												
F <sub>3</sub>	-.254	.105	1											
F <sub>4</sub>	-.167	-.287	-.371	1										
F <sub>5</sub>	.249	-.242	.132	-.093	1									
$\alpha$	.593	.261	.403	.486	.657									
IC 95%	.518, .659	.158, .359	.309, .489	.399, .564	.591, .714									
$\alpha_{ordinal}$	.785	.517	.483	.685	.648									
$\omega$	.818	.587	.519	.706	.649									

Nota: n = 332; <sup>a</sup> = numeración original; F<sub>1</sub> = extraversión; F<sub>2</sub> = afabilidad; F<sub>3</sub> = responsabilidad; F<sub>4</sub> = neuroticismo; F<sub>5</sub> = apertura (no aplica al BFI-10);  $\Phi_1$  = coeficiente de congruencia factorial del ítem;  $\Phi_F$  = coeficiente de congruencia factorial del factor; ISF = índice de simplicidad factorial;  $\alpha_{ordinal}$  = coeficiente alfa ordinal;  $\omega$  = coeficiente  $\omega$

**Tabla 2**  
Análisis estructural y de confiabilidad del BFI-15p y BFI-10p

Ítems <sup>a</sup>	Análisis factorial: versión de 15 ítems					Análisis factorial: versión de 10 ítems							
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	Φ <sub>1</sub>	ISF	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	ISF
BFI1	.753	.246	.019	.150	.115	.980	.822	.717	.051	-.048	-.030	-.040	.982
BFI11	.579	.201	.226	-.004	.129	.956	.702	-	-	-	-	-	-
BFI43	.806	.320	.122	.219	.211	.974	.700	.807	.028	.016	-.034	-.036	.993
BFI7	.297	.660	.282	.163	.145	.950	.596	.037	.786	-.007	.043	.021	.993
BFI37	.142	.582	.224	.074	.082	.978	.761	-	-	-	-	-	-
BFI41	.234	.755	.252	.212	.190	.977	.677	.174	.480	.038	.006	.116	.801
BFI14	.128	.301	.742	.238	.200	.968	.667	-.080	.022	.896	.070	-.069	.975
BFI21	.228	.167	.705	.160	.348	.954	.620	.153	-.042	.426	-.158	.162	.636
BFI34	.158	.311	.397	.216	.228	.847	.312	-	-	-	-	-	-
BFI4	.049	.147	.167	.784	.185	.992	.847	-	-	-	-	-	-
BFI15	.038	.222	.278	.830	.067	.974	.803	-.214	.178	.009	.489	-.067	.693
BFI30	.300	.106	.143	.814	.243	.977	.737	.141	-.123	.013	.832	.091	.926
BFI17	-.029	.260	.268	.060	.646	.936	.693	-.093	-.017	-.019	-.043	.983	.986
BFI23	.334	-.112	.316	.188	.575	.860	.478	.045	-	-	-	-	-
BFI39	.149	.231	.127	.212	.728	.985	.746	.045	.129	-.004	.099	.426	.833
Φ <sub>F</sub>	.950	.933	.923	.985	.961	-	-	-	-	-	-	-	-
F <sub>1</sub>	1							1					
F <sub>2</sub>	.122	1						.280	1				
F <sub>3</sub>	.112	.065	1					.349	.384	1			
F <sub>4</sub>	-.310	-.413	-.572	1				.048	-.109	-.227	1		
F <sub>5</sub>	-.093	.149	-.067	-.258	1			.557	.411	.506	-.251	1	
α	.763	.700	.567	.602	.592			.739	.686	.608	.517	.604	
IC 95%	.714, .805	.641, .751	.489, .636	.529, .666	.517, .658			.686, .784	.624, .739	.535, .672	.433, .592	.531, .668	
r <sub>AF</sub>	-	-	-	-	-			.739	.686	.611	.519	.609	
α <sub>ordinal</sub>	.752	.703	.634	.851	.685			.733	.548	.553	.578	.590	
ω	.759	.707	.655	.851	.688			.736	.582	.632	.620	.700	
H	.786	.723	.706	.852	.700			.745	.657	.811	.719	.967	

Nota: n = 332; <sup>a</sup> = numeración original; F<sub>1</sub> = extraversión; F<sub>2</sub> = afabilidad; F<sub>3</sub> = responsabilidad; F<sub>4</sub> = neuroticismo; F<sub>5</sub> = apertura; Φ<sub>1</sub> = coeficiente de congruencia factorial del ítem; Φ<sub>F</sub> = coeficiente de congruencia factorial del factor; ISF = índice de simplicidad factorial; r<sub>AF</sub> = coeficiente Angoff-Feildt; α<sub>ordinal</sub> = coeficiente α ordinal; ω = coeficiente ω; H = coeficiente H.

**Tabla 3***Equivalencia correlacional y de confiabilidad entre BFI-37, BFI-15p y BFI-10p*

	$r_{37-15}$		$q$	Confiabilidad Diferencias $\alpha_{37-15}$ ( $\chi^2$ , gl = 1) <sup>b</sup>	$r_{37-10}$	
	Sin corregir (IC 95%)	Corregido <sup>a</sup> (IC 95%)			Sin corregir (IC 95%)	Corregido <sup>a</sup> (IC 95%)
Extraversión	.877 (.850, .900)	.767 .719, .808	.060	2.692	.845** .811, .873	.754 .704, .797
Afabilidad	.906 (.885, .924)	.665 .600, .721	.138	4.413*	.838** .803, .867	.653 .587, .711
Responsabilidad	.786 (.741, .824)	.598 .524, .663	.121	10.941**	.722** .666, .770	.592 .517, .658
Neuroticismo	.861 (.830, .887)	.663 .598, .719	.116	18.643**	.787** .742, .825	.613 .541, .676
Apertura	.798 (.755, .834)	.598 .524, .663	.128	23.608**	.759** .709, .801	.610 .534, .673

Nota: <sup>a</sup> = Corrección Levy; <sup>b</sup> = Calculada con  $r$  corregida por espuriedad; \* =  $p < .05$ ; \*\* =  $p < .01$ ;  $q = q$  de Cohen.

Los resultados psicométricos obtenidos no apoyaron estas versiones existentes, sugiriendo que la representatividad del constructo con los ítems es plausiblemente condicional a las variaciones culturales de la muestra. Esta falta de consistencia en los resultados se ha observado también en las validaciones internacionales de la versión extensa del BFI (e.g., Gurven et al., 2013), en la cuales existen diferencias entre países respecto a la relevancia de cada ítem dentro de los factores. Un aspecto que razonablemente pudo influir en este problema es la presencia de ítems en escala invertida, pues aunque se asume que estos evalúan el polo opuesto del constructo, generalmente producen un efecto de factores espurios (Lance et al., 2010).

Con relación a las nuevas versiones breves extraídas para la muestra peruana, BFI-10p y BFI-15p, las cargas factoriales de magnitudes entre moderadas y altas sugieren una estructura factorial robusta y de más próxima cercanía con la TM especificada en comparación con la versión extensa (Dominguez-Lara et al., en prensa). Esto se debe a

que fueron elegidos los ítems con mayor  $\Phi$  e ISF para conformar esas versiones, lo que les daría mayor independencia respecto a los otros factores. Quizá es conveniente resaltar que no se pudo emplear la rotación procustea con el BFI-10p, ya que el programa informático utilizado solo realiza análisis con al menos tres ítems por factor. A pesar de dicha limitación, los resultados obtenidos mediante los procedimientos tradicionales (MCNP, Promax) fueron alentadores.

En cuanto a las evidencias de validez del BFI-10p y BFI-15p mediante las relaciones con la versión extensa, fueron halladas correlaciones de magnitud aceptable luego de la corrección por errores correlacionados. Aparentemente, este procedimiento no fue implementado en trabajos preliminares (Rammstedt & John, 2007), por lo cual es necesario resaltarlo para interpretar con precaución tales estimaciones, aunque su uso en otros estudios peruanos orientados a conseguir versiones breves reafirma su utilidad en ese contexto metodológico (e.g., Domínguez-Lara & Merino-Soto, 2017).

Finalmente, estas versiones breves produjeron coeficientes de confiabilidad  $\alpha$  y  $\omega$  bajos, lo cual es un problema habitual en esta reducción (Saucier, 1994); sin embargo, esta disminución es más problemática en el nuevo BFI-15. Considerando que el error estándar de medición puede ser alto y poco tolerable para la descripción individualizada, las versiones breves derivadas en el presente estudio pueden ser aprovechadas en descripciones de grupo y para fines de investigación con el objetivo de enriquecer el conocimiento que se tiene del modelo de los 5GF en Perú y Sudamérica a través de versiones extensas (BFI) o abreviadas (BFI-10p y BFI-15p).

Sobre la confiabilidad, cabe resaltar que las condiciones para aplicar el coeficiente  $\alpha$  no fueron cumplidas en ninguna de las tres versiones, ya que algunos ítems presentaron coeficientes de configuración de mayor magnitud que otros, y en unos casos la diferencia fue notoria. Por ello, fueron consideradas otras alternativas de análisis ( $\alpha_{\text{ordinal}}$ ,  $\omega$  y  $H$ ), cuyas magnitudes fueron superiores. Estos resultados se deben a que los coeficientes basados en variables latentes no requieren la tau-equivalencia como condición necesaria para su cálculo, ya que es suficiente con que se traten de medidas congénicas (Dunn, Baguley, & Brunsten, 2014). Para el BFI-10p y BFI-15p, la diferencia no fue tan abrumadora como los anteriores ya que está compuesto por ítems más homogéneos y el  $r_{AF}$  fue cercano al  $\alpha$ .

Los resultados presentados tienen implicancias teóricas, metodológicas y prácticas. Teóricamente, el modelo evaluado parece mantener la original estructura de las cinco dimensiones del modelo de los 5GF en la muestra evaluada. Esto es importante considerando que este modelo ha sido pocas veces evaluado en la cultura peruana y se desconoce si su estructura dimensional puede ser replicable y si su contenido representa invariablemente las conductas relevantes a sus factores. En cuanto al aspecto metodológico, en el instrumento derivado se obtuvo no solo la validez

de sus ítems, sino también de su distancia frente a los otros factores; esto habitualmente se llama simplicidad factorial. Ambos criterios no están linealmente asociados y la presencia de uno (e.g., fuerte carga factorial del ítem en el propio factor) no significa la ausencia del otro (e.g., carga muy baja o cerca de cero en el resto de factores). Adicionalmente, como no han sido replicadas las versiones breves desarrolladas en otros contextos, entonces se requiere inequívocamente una evaluación dimensional de los instrumentos importados de otros contextos. La acumulación de evidencias de estructura interna, aunque solo es un paso dentro de un conjunto de otras fuentes de validez, debe proveer la mejor evidencia para aceptar inicialmente la utilidad de un instrumento y la generalización de la teoría detrás de ella. El presente estudio concluye en que la versión abreviada obtenida del BFI para evaluar el modelo de los 5GF de personalidad fue parcialmente satisfactoria, y que se requiere de trabajos adicionales que continúen evaluando la dimensionalidad de su estructura propuesta y las relaciones entre dichas dimensiones.

Las limitaciones inherentes a la muestra (e.g., tamaño muestral) fueron las mismas que se presentaron en el estudio anterior (Dominguez-Lara et al., en prensa), por lo que su mención en este espacio sería redundante. No obstante, debe considerarse que las versiones abreviadas no reemplazan a las extensas (Gosling et al., 2003; Rammstedt & John, 2007); solo deben ser empleadas cuando la versión más larga no pueda aplicarse por diversos motivos. Finalmente, cabe la posibilidad que, al tratarse de la misma muestra, los resultados estén sobreajustados. En ese caso, es conveniente replicar el estudio en una muestra distinta, sean otros estudiantes universitarios o adultos de población general.

### Conflicto de intereses

Los autores manifiestan que no existe conflicto de intereses.

## Responsabilidad ética

a) Protección de personas y animales: los autores manifiestan que no fueron realizados experimentos.  
 b) Confidencialidad de datos: el proyecto de investigación fue aprobado por la institución donde fue realizada la recogida de información donde las respuestas a las encuestas fueron completamente anónimas. c) Derecho a la privacidad y consentimiento informado: en el presente estudio no aparecen datos de participantes.

## Referencias

- Bashaw, W. L., & Anderson, H. E. (1967). A correction for replicated error in correlation coefficients. *Psychometrika*, 32, 435-441. doi: 10.1007/BF02289657
- Benet-Martínez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75(3), 729-750. doi: 10.1037/0022-3514.75.3.729
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296-322. doi: 10.1111/j.2044-8295.1910.tb00207.x
- Browne, M. W. (1972). Oblique rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, 25, 207-212.
- Carciofo, R., Yang, J., Song, N., Du, F., & Zhang, K. (2016). Psychometric evaluation of Chinese-Language 44-item and 10-item big five personality inventories, including correlations with chronotype, mindfulness and mind wandering. *PLOS ONE*, 11(2), 1-26. doi: 10.1371/journal.pone.0149963
- Carvalho, L., Nunes, M., Primi, R., & Nunes, C. (2012). Unfavorable evidence for personality assessment with a 10-item instrument. *Paidéia (Ribeirão Preto)*, 22(51), 63-71. doi: 10.1590/S0103-863X2012000100008
- Charter, R. (2001). It is time to bury the Spearman-Brown «prophecy» formula for some common applications. *Educational and Psychological Measurement*, 61(4), 690-696. doi: 10.1177/00131640121971446
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi: 10.1207/S15328007SEM0902\_5
- Cheung, F. M., van de Vijver, F. J., & Leong, F. T. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, 66(7), 593-603. doi: 10.1037/a0022389
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. doi: 10.1037/0033-2909.112.1.155
- Costa, P. T., & McCrae, R. R. (1999). *Manual NEO PI-R, Inventario de Personalidad NEO Revisado y NEO-FFI, Inventario NEO reducido de Cinco Factores*. Madrid: TEA Ediciones, SA.
- Denissen, J. J. A., Greenen, R., van Aken, M. A. G., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment*, 90(2), 151-157. doi: 10.1080/00223890701845229
- Dominguez-Lara, S. (2014). ¿Matrices policóricas/tetracóricas o matrices Pearson? Un estudio metodológico. *Revista Argentina de Ciencias del Comportamiento*, 6(1), 39-48.
- Dominguez-Lara, S. (2016). Evaluación de la confiabilidad del constructo mediante el Coeficiente H: breve revisión conceptual y aplicaciones. *Psychologia: Avances en la Disciplina*, 10(2), 87-94.
- Dominguez-Lara, S. (2018). Fiabilidad y alfa ordinal. *Actas Urológicas Españolas*, 42(2), 140-141. doi: 10.1016/j.acuro.2017.07.002
- Dominguez-Lara, S., & Merino-Soto, C. (2015). ¿Por qué es importante reportar los intervalos de confianza del coeficiente alfa de Cronbach? *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 13(2), 1326-1328.

- Dominguez-Lara, S., & Merino-Soto, C. (2017). Versión breve de la Escala de Afrontamiento ante la Ansiedad e Incertidumbre Preexamen (COPEAU) en universitarios peruanos. *Educación Médica*. doi: 10.1016/j.edume.2017.04.01
- Dominguez-Lara, S., Merino-Soto, C., & Navarro-Loli, J. (2016). Estimación de la confiabilidad en mediciones de dos ítems: el Coeficiente Angoff-Feldt. *Revista Digital de Investigación en Docencia Universitaria*, 10(1), 34-40. doi: 10.19083/ridu.10.463
- Dominguez-Lara, S., Merino-Soto, C., Zamudio, B., & Guevara, C. (en prensa). Big Five Inventory en universitarios peruanos: Resultados preliminares de la validación. *Psykhé*, 27(2). doi: 10.7764/psykhe.27.2.1052
- Dunn, T. J., Baguley, T., & Brunsten, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. doi: 10.1111/bjop.12046
- Elosua, P., & Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20(4), 896-901.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. doi: 10.1037/1082-989X.4.3.272
- Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika*, 45, 99-105. doi: 10.1007/BF02293600
- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.). *Educational Measurement* (pp. 105-146). New York: Macmillan.
- Feldt, L. S., & Charter, R. A. (2003). Estimating the reliability of a test split into two parts of equal or unequal length. *Psychological Methods*, 8(1), 102-109. doi: 10.1037/1082-989X.8.1.102
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11(1), 93-103. doi: 10.1177/014662168701100107
- Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Fetvadjev, V., & van der Vijver, F. J. R. (2015). Universality of the five-factor model of personality. En J. D. Wright (Ed.). *International encyclopedia of social and behavioral sciences* (Vol. 9, pp. 249-253). Oxford: Elsevier.
- Fisher, R. (1950). *Statistical methods for research workers*. Edinburgh, UK: Oliver & Boyd.
- Fleming, J., & Merino, C. (2005). Medidas de simplicidad y ajuste factorial: Un enfoque para la construcción y revisión de escalas derivadas factorialmente. *Revista de Psicología*, 23(2), 252-266.
- Fossati, A., Borroni, S., Marchione, D., & Maffei, C. (2011). The Big Five Inventory (BFI). Reliability and Validity of its Italian Translation in Three Independent Nonclinical Samples. *European Journal of Psychological Assessment*, 27, 50-58. doi: 10.1027/1015-5759/a000043
- Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, & Evaluation*, 17(3). Recuperado de <http://pareonline.net/getvn.asp?v=17&n=3>
- Gerlitz J. Y., & Schupp, J. (2005). *Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. Dokumentation der Instrumentenentwicklung BFI-S auf Basis des SOEP-Pretests 2005*. DIW Research, Notes 4.
- Goldberg, L. R. (1990). An alternative «description of personality»: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216-1229. doi: 10.1037/0022-3514.59.6.1216
- Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, 48(1), 82-98.

- Goldberg, L., Johnson, J., Eber, H., Hogan, R., Ashton, R., & Robert, C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96. doi: 10.1016/j.jrp.2005.08.007
- Gosling, S., Rentfrow, P., & Swann, W. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality, 37*, 504-528. doi: 10.1016/S0092-6566(03)00046-1
- Guido, G., Peluso, A. M., Capestro, M., & Miglietta, M. (2015). An Italian version of the 10-item Big Five Inventory: An application to hedonic and utilitarian shopping values. *Personality and Individual Differences, 76*, 135-140. doi: 10.1016/j.paid.2014.11.053
- Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager-farmers in the Bolivian Amazon. *Journal of Personality and Social Psychology, 104*(2), 354-370. doi: 10.1037/a0030841
- Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality - Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality, 46*(3), 355-359. doi: 10.1016/j.jrp.2012.03.008
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. H. C. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Past and present. A Festschrift in honor of Karl G. Jöreskog* (pp. 195-261). Chicago: Scientific Software International.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. doi: 10.1017/S0140525X0999152X
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory - Versions 4a and 54*. California: Institute of Personality and Social Research.
- Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-35. doi: 10.1007/BF02291575
- Kim, S. Y., Kim, J. M., Yoo, J. A., Bae K. Y., Kim, S. W., Yang, S. J., ... Yoon, J. S. (2010). Standardization and validation of Big Five Inventory - Korean version (BFI-K) in elders. *Korean Journal of Biological Psychiatry, 17*(1), 15-25.
- Lance, C., Dawson, B., Birkelbach, D., & Hoffman, B. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods, 13*(3), 435-455. doi: 10.1177/1094428109352528
- Lautenschlager, G. J., & Meade, A. W. (2008). AlphaTest: A windows program for tests of hypotheses about coefficient alpha. *Applied Psychological Measurement, 23*, 502-503. doi: 10.1177/0146621607312307
- Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology, 23*, 84-86.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología, 30*(3), 1151-1169. doi: 10.6018/analesps.30.3.199361
- Lorenzo-Seva, U., & Ferrando, P. (2013). FACTOR 9.2 A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT Models. *Applied Psychological Measurement, 37*(6), 497-498. doi: 10.1177/0146621613487794
- Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. *Methodology, 2*, 57-64. doi: 10.1027/1614-2241.2.2.57
- Lovik, A., Verbeke, G., & Molenberghs, G. (2017). Evaluation of a very short test to measure the Big Five personality factors on a Flemish sample. *Journal of Psychological and Educational Research, 25*(2), 7-17.
- McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*(3), 547-561. doi: 10.1037/0022-3514.88.3.547

- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, N.J.: L. Erlbaum Associates.
- Merino, C., & Grimaldo, M. (2011). Complejidad factorial y conductas moralmente controversiales. *Revista Argentina de Ciencias del Comportamiento*, 3(3), 38-43.
- Merino-Soto, C., Navarro-Loli, J., & García, W. (2014). Revisión de la consistencia interna del Inventario de Inteligencia Emocional de Bar-On, EQ-I: YV. *Revista Peruana de Psicología y Trabajo Social*, 3(1), 141-154.
- Ortet, G., Ibáñez, M., Ruipérez, M., Villa, H., Moya, J., & Escrivá, P. (2007). Adaptación para adolescentes de la versión española del NEO PI-R. *Psicothema*, 19(2), 263-268.
- Pejić, M., Tenjović, L., & Knežević, G. (2014). Validacija upitnika ličnosti BFI-10—kratke forme inventara velikih pet. *Primenjena Psihologija*, 7(1), 45-92.
- Pérez, E., & Medrano, L. A. (2010). Análisis factorial exploratorio: bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento*, 2(1), 58-66.
- Petrides, K. V., Jackson, C. J., Furnham, A., & Levine, S. Z. (2003). Exploring issues of personality measurement and structure through the development of a short form of the Eysenck personality profiler. *Journal of Personality Assessment*, 81, 271-280. doi: 10.1207/S15327752JPA8103\_10
- Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. doi: 10.1037/0021-9010.88.5.879
- Ponterotto, J. G., & Charter R. A. (2009). Statistical extensions of Ponterotto and Ruckdeschel's (2007) reliability matrix for estimating the adequacy of internal consistency coefficients. *Perceptual and Motor Skills*, 108, 878-886. doi: 10.2466/pms.108.3.878-886
- Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, 87(1), 102-112. doi: 10.1207/s15327752jpa8701\_09
- Rammstedt, B. (2007). The 10-Item Big Five Inventory (BFI-10): Norm values and investigation of socio-demographic effects based on a German population representative sample. *European Journal of Psychological Assessment*, 23, 193-201. doi: 10.1027/1015-5759.23.3.193
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203-212. doi: 10.1016/j.jrp.2006.02.001
- Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2013). Eine kurze skala zur messung der fünf dimensionen der persönlichkeit. *Methoden, Daten, Analysen*, 7(2), 233-249.
- Reyes, E., Álvarez, C., Peredo, A., Miranda, A., & Rebolledo, I. (2014). Psychometric properties of the Big Five Inventory in a Mexican sample. *Salud Mental*, 37, 491-497.
- Rodríguez, M., & Ruiz, M. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial. *Psicológica*, 29, 205-227.
- Romano, J. L., Kromrey, J. D., & Hibbard, S. T. (2010). A Monte Carlo study of eight confidence interval methods for coefficient alpha. *Educational and Psychological Measurement*, 70(3), 376-393. doi: 10.1177/0013164409355690
- Romano, J. L., Kromrey, J. D., Owens, C. M., & Scott, H. M. (2011). Confidence interval methods for coefficient alpha on the basis of discrete, ordinal response items: Which one, if any, is the best? *The Journal of Experimental Education* 79(4), 382-403. doi: 10.1080/00220973.2010.510859
- Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, 63(3), 506-516. doi: 10.1207/s15327752jpa6303\_8

- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology, 38*(2), 173-212. doi: 10.1177/0022022106297299
- Spearman, C. (1910). Correlation calculated with faulty data. *British Journal of Psychology, 3*, 271-295. doi: 10.1111/j.2044-8295.1910.tb00206.x
- Tataryn, D., Wood, J., & Gorsuch, R. (1999). Setting the value of k in promax: A Monte Carlo study. *Educational and Psychological Measurement, 59*, 384-391. doi: 10.1177/00131649921969938
- ten Berge, J. (1986). Some relationships between descriptive comparisons of components from different studies. *Multivariate Behavioral Research, 21*(1), 29-40. doi: 10.1207/s15327906mbr2101\_2
- Tucker, L. R. (1951). *A method for synthesis of factor analysis studies (Personnel Research Section Report N° 984)*. Washington, D. C.: Department of the Army.
- Ubbiali, A., Chiorri, C., Hampton, P., & Donati, D. (2013). Psychometric properties of the Italian adaptation of the Big Five Inventory (BFI). *Bollettino di Psicologia Applicata, 266*, 37-46.
- Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., ... Jang, K. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology, 90*(6), 987-999. doi: 10.1037/0022-3514.90.6.987
- Worrell, F. C., & Cross, W. E. (2004). The reliability and validity of Big Five Inventory scores with African American college students. *Journal of Multicultural Counseling and Development, 32*(1), 18-32. doi: 10.1002/j.2161-1912.2004.tb00358.x
- 

Sergio Dominguez-Lara

Universidad de San Martín de Porres (Perú)

Investigador en la Universidad de San Martín de Porres, magíster en Psicología, con mención en Psicología Clínica y de la Salud.

ORCID: 0000-0002-2083-4278

Autor corresponsal: sdominguezmpcs@gmail.com sdominguezl@usmp.pe

César Merino-Soto

Universidad de San Martín de Porres (Perú)

Doctorando en la Universidad Autónoma del Estado de Morelos, Magíster en Psicología, con mención en Psicología Educativa.

ORCID: 0000-0002-1407-8306

sikayax@yahoo.com.ar