

# *Elaboración de la versión breve del Cuestionario de Personalidad IPIP-Revisado: Control del sesgo de aquiescencia*

MARCOS CUPANI\*

Universidad Nacional de Córdoba, Córdoba, Argentina  
ORCID: <https://orcid.org/0000-0003-2132-5552>

URBANO LORENZO-SEVA

Universitat Rovira i Virgili, Tarragona, España  
ORCID: <https://orcid.org/0000-0001-5369-3099>

CELINA GRACIELA KORZENIOWSKI

Universidad del Aconcagua, Mendoza, Argentina  
ORCID: <https://orcid.org/0000-0001-6678-3472>

ANA ESTEFANÍA AZPILICUETA

Universidad Nacional de Córdoba, Córdoba, Argentina  
ORCID: <https://orcid.org/0000-0002-6221-3162>

**How to quote this article:** Cupani, M., Lorenzo-Seva, U., Korzeniowski, C.G., Azpilicueta, A.E. (2019). Elaboración de la Versión Breve del Cuestionario de Personalidad IPIP- Revisado: Control del sesgo de respuesta aquiescencia. *Acta Colombiana de Psicología*, 22(1), 248-260. doi: <http://www.dx.doi.org/10.14718/ACP.2019.22.1.12>

## Resumen

En los últimos años se han desarrollado medidas breves de los cinco factores de personalidad, sin embargo, la ganancia práctica de tiempo provista por las formas breves puede implicar propiedades psicométricas más débiles de los instrumentos. En la construcción de escalas breves, para mantener propiedades psicométricas adecuadas se debe emplear criterios teóricos y empíricos en la selección de los ítems y minimizar los sesgos de respuesta, como el de la aquiescencia (AC), que hace referencia a la tendencia de las personas a estar de acuerdo con afirmaciones positivas independientemente del contenido de la afirmación. Teniendo esto en cuenta, el objetivo principal del presente estudio fue desarrollar un instrumento breve (30 ítems), de dominio público, para medir los cinco factores de personalidad en población latina, controlando el sesgo de respuesta AC. La muestra estuvo compuesta por 910 participantes, 543 de sexo femenino (59.6 %) y 367 de sexo masculino (40.3 %), con edades comprendidas entre los 15 y los 80 años ( $M = 29.52$ ;  $DT = 12.25$ ), pertenecientes a la ciudad de Córdoba, Argentina. Para el proceso de validación se propuso realizar un estudio de convergencia con las cinco escalas del NEO-FFI, un análisis de diferencia de grupos según el sexo y la edad de los participantes, y un estudio de validez predictiva respecto a algunas actividades recreacionales (uso de drogas, irresponsabilidad, amistad, erudición/creatividad y comunicación). Los resultados indican que el IPIP-R-30 presenta una estructura factorial de cinco factores, índices de confiabilidad adecuados tanto de consistencia interna como de estabilidad temporal, evidencia de validez convergente con las escalas del NEO-FFI, evidencia de diferencia de grupos según sexo y edad, y validez predictiva de la frecuencia de diferentes categorías de actividades específicas. De esta manera, se puede concluir que el IPIP-R-30 constituye una herramienta válida de evaluación de los rasgos de personalidad en nuestro medio, con puntuaciones libres del sesgo de AC. *Palabras clave:* cinco grandes factores de la personalidad, inventario, IPIP, aquiescencia, sesgo de respuesta.

## *Development of a Brief Version of the Personality Inventory IPIP-Revised: Control of the Acquiescence Response Bias*

### Abstract

In recent years, brief measures of the five personality factors have been developed; however, the practical gain of time provided by the brief versions may imply weaker psychometric properties of the instruments. To maintain adequate psychometric properties in the construction of brief scales, theoretical and empirical criteria should be used in the selection of items, and response biases such as acquiescence (AC) should be minimized. The term AC refers to people's tendency to agree with positive statements, regardless of their content. The main purpose of the present study is to develop a brief public

\* Instituto de Investigaciones Psicológicas (IIPsi-CONICET), Facultad de Psicología, Universidad Nacional de Córdoba, Ciudad Universitaria, Córdoba 5000, Argentina. Tel./Fax: +54-351-4334104. [marcoscup@gmail.com](mailto:marcoscup@gmail.com)

Recibido, julio 4/2017; Concepto de evaluación, noviembre 17/2017; Aceptado, enero 31/2018

domain instrument (30 items) to measure the five personality factors in the Latin American population, controlling the AC response bias. The sample consisted of 910 participants, 543 women (59.6 %) and 367 men (40.3 %), age range 15-80 years ( $M = 29.52$ ;  $DT = 12.25$ ) from the city of Córdoba, Argentina. For the validation process, a convergence study with the five NEO-FFI scales, an analysis of group differences according to the participants' sex and age, and a predictive validity study regarding recreational activities (drug use, irresponsibility, friendship, erudition/creativity, and communication) were performed. The results indicate that the IPIP-R-30 presents a five-factor factorial structure, adequate reliability indices of both internal consistency and temporal stability, evidence of convergent validity with the NEO-FFI scales, evidence of group differences as regards sex and age, and frequency predictive validity of different categories of specific activities. Thus, it can be concluded that the IPP-R-30 is a valid tool for assessing personality factors in our environment, with scores free of AC bias.

*Keywords:* big five personality factors, inventory, IPIP, acquiescence, response bias.

## *Elaboración de la versión breve del Inventario de Personalidad IPIP-Revisado: control de los sesgos de aquiescencia*

### Resumo

Nos últimos años, se han desarrollado medidas breves de los cinco factores de personalidad; sin embargo, el tiempo previsto por las formas breves puede implicar propiedades psicométricas más débiles de los instrumentos. En la construcción de escalas breves, para mantener propiedades psicométricas adecuadas, deben ser empleados criterios teóricos y empíricos en la selección de los ítems, y deben ser minimizados los sesgos de respuesta, como el de la aquiescencia (AC). Este concepto hace referencia a la tendencia de las personas que están de acuerdo con afirmaciones positivas, independientemente del contenido de la afirmación. En este sentido, el objetivo principal de este estudio fue desarrollar un instrumento breve (30 ítems), de dominio público, para medir los cinco factores de personalidad en la población latina, controlando el sesgo de respuesta AC. La muestra estuvo compuesta por 910 participantes, 543 de sexo femenino (59.6 %) y 367 de sexo masculino (40.3 %), entre 15 y 80 años de edad ( $M = 29.52$ ;  $DP = 12.25$ ), pertenecientes a la ciudad de Córdoba, Argentina. Para el proceso de validación, se propuso realizar un estudio de convergencia con las cinco escalas del NEO-FFI, un análisis de diferencias de grupos según el sexo y la edad de los participantes, y un estudio de validez predictiva con respecto a actividades recreativas (uso de drogas, irresponsabilidad, amistad, erudición/creatividad y comunicación). Los resultados indican que el IPIP-R-30 presenta una estructura factorial de cinco factores, índices de confiabilidad adecuados tanto de consistencia interna como de estabilidad temporal, evidencia de validez convergente con las escalas del NEO-FFI, evidencia de diferencias de grupos según el sexo y la edad, y validez predictiva de la frecuencia de diferentes categorías de actividades específicas. De esta manera, puede concluirse que el IPIP-R-30 constituye una herramienta válida de evaluación de los rasgos de personalidad en nuestro medio, con puntuaciones libres de los sesgos de AC.

*Palabras-clave:* cinco grandes factores de personalidad, inventario, IPIP, aquiescencia, sesgo de respuesta.

### Introducción

En los últimos años se han desarrollado medidas breves de los cinco factores de personalidad con el fin de solucionar una de las desventajas prácticas que poseen los cuestionarios tradicionales de personalidad, a saber, el tiempo requerido para completarlos (Sibley, 2012). Dentro de los más reconocidos se encuentran el Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow & Swann, 2003), el Mini-IPIP (Donnellan, Oswald, Baird & Lucas, 2006), el BFI-2-S y BFI-2-XS (Soto & John, 2017), el Abridged Big Five (Langford, 2003), el Five-Item Measure of the Big Five (Aronson, Reilly & Lynn, 2006), y el cuestionario Big Five Inventory-10 (BFI-10; Rammstedt & John, 2007).

La ganancia práctica de tiempo provista por las formas breves, sin embargo, puede implicar en muchos casos propiedades psicométricas más débiles de los instrumentos (Credé, Harms, Niehorster & Gaye-Valentine, 2012). Por ejemplo, en comparación con una escala de 60 ítems, una

de 30 ítems registró una pérdida en la fiabilidad y validez estimada en un 10 %, mientras que en una escala de 15 ítems la pérdida ascendió al 20 % (Soto & John, 2017). Específicamente, la consistencia interna de los cuestionarios suele ser la más afectada, aunque también se registran disminuciones en la estabilidad temporal y la confiabilidad inter-examinador (Baldasaro, Shanahan & Bauer, 2013; Sibley, 2012), al igual que en la validez convergente, discriminante y predictiva en comparación con las versiones extendidas (Baldasaro et al., 2013). Asimismo, con respecto a la validez de constructo, no hay evidencia concluyente sobre si los inventarios breves se ajustan a la estructura factorial de los cinco factores, ya que algunas investigaciones informan un ajuste de adecuado a bueno (Baldasaro et al., 2013; Donnellan et al., 2006; Soto & John, 2017), mientras que otras reportan una falta de ajuste (Gosling et al., 2003).

En concordancia, se ha documentado que el uso de las formas muy cortas (uno o dos ítems por dominio) se relaciona con el aumento de los errores Tipo 1 y Tipo 2

(Credé et al., 2012; Milojev, Osborne, Greaves, Barlow & Sibley 2013; Krueger, Emons & Sijtsma, 2013). Por ejemplo, el TIPI brinda puntajes que no presentan valores adecuados de fiabilidad y ha sido criticado por su falta de amplitud en la evaluación del constructo (Baldaraso et al., 2013), pero, por otra parte, como es el caso del Mini IPIP (Donnellan et al., 2011; Sibley, 2012), con cuatro ítems por dominio la consistencia interna mejora, aunque sigue siendo relativamente baja, y la validez convergente y de constructo parecen no afectarse significativamente (De Vries, 2013).

Además de esto, la pérdida en la consistencia interna es uno de los problemas que más preocupa a los investigadores y los sitúa frente a la disyuntiva de decidir entre consistencia o amplitud en la cobertura del constructo (Baldaraso et al., 2013). En caso de priorizar el contenido, el resultado es una escala heterogénea, pero poco fiable; y, en sentido opuesto, al priorizar la consistencia, pueden producirse problemas tautológicos por incluir ítems que están fuertemente correlacionados entre sí y con el dominio a evaluar (Soto & John, 2017).

En suma, en la construcción de escalas breves se debe balancear el deseo de minimizar el número de ítems necesarios para evaluar cada dominio de personalidad y el propósito de mantener la cobertura del contenido con buenas propiedades psicométricas (Milojev et al., 2013), y un modo de alcanzar este propósito es emplear criterios teóricos y empíricos para seleccionar los ítems, así como minimizar los sesgos de respuesta.

Con respecto a esto último, el concepto de aquiescencia (AC) ha llamado últimamente la atención de los estudiosos en el tema, ya que se define como un sesgo de respuesta que hace referencia a la tendencia de las personas a estar de acuerdo con afirmaciones positivas, independientemente del contenido de la afirmación (Cronbach, 1942). En algunas investigaciones se ha evidenciado que existen diferencias de sexo, edad, nivel de inteligencia, nivel educativo y variaciones culturales en este sesgo de respuesta (Javeline, 1999; Johnson, Kulesa, Cho & Shavitt, 2005; Vazsonyi, Ksinan, Mikuška & Jiskrova, 2015), y que la falta de control de dichas diferencias puede distorsionar la estructura factorial de la escala y sus asociaciones con criterios externos. De este modo, algunos autores consideran que la AC es una fuente de variación indeseable que debe ser suprimida (Hofstee, ten Berge & Hendriks, 1998), mientras que otros sugieren que cuando el sesgo de AC se controla, la estructura factorial de cinco factores se ajusta correctamente (Rammstedt, Kemper & Borg, 2013).

En función de lo anterior, resulta indispensable controlar este sesgo a la hora de producir escalas breves, por lo que, para llevar a cabo este tipo de estudios, es importante diseñar un método en el momento en el que se está desarrollando

el inventario (Cupani & Lorenzo-Seva, 2016), como lo es, por ejemplo, el procedimiento propuesto por Lorenzo-Seva y Ferrando (2009) para escalas parcialmente balanceadas.

En las últimas décadas se ha registrado un incremento en las publicaciones sobre la aplicación, adaptación y validación de las escalas IPIP para la evaluación de los cinco factores de personalidad (Goldberg, 2001). De hecho, el IPIP mismo resultó de un esfuerzo colaborativo a nivel mundial para desarrollar un set de ítems de personalidad de dominio público, que, por lo tanto, pudieran ser utilizados libremente por investigadores de cualquier lugar del mundo tanto para propósitos científicos como comerciales. Con esta iniciativa, en la actualidad se puede contar con un *pool* resultante de 2000 ítems IPIP (Goldberg et al., 2005), los cuales han sido traducidos a más de 10 idiomas, por lo que se observa un aumento sostenido en la adaptación y construcción de escalas IPIP en diferentes culturas y países (Cupani & Lorenzo-Seva, 2016). Por ejemplo, se han validado escalas IPIP de 50 y 100 ítems en EE. UU. (Goldberg et al., 2005), Escocia (Gow, Whiteman, Pattie & Deary, 2005), Nueva Zelanda (Guenole & Chernyshenko, 2005), Croacia (Mlačić & Goldberg, 2007) y China (Zheng et al., 2008), aunque en Latinoamérica son escasos los trabajos en el área (Cupani, 2009; Gross, Zalazar Jaime, Piccolo & Cupani, 2012).

En nuestro contexto, el *Cuestionario de Personalidad IPIP-Revisado* (Cupani & Lorenzo-Seva, 2016), cuyos 50 ítems fueron seleccionados de los 100 ítems propuestos por Goldberg (1999), y con control del efecto de la aquiescencia, presenta una estructura factorial simple, ortogonal y adecuadas propiedades psicométricas (Cupani & Lorenzo-Seva, 2016).

Ahora bien, considerando que son escasos los estudios que han ponderado el efecto de la aquiescencia en la construcción de escalas breves de personalidad, en la presente investigación se propone desarrollar un instrumento breve (30 ítems), de dominio público, para medir los cinco factores en población latina, que controle el sesgo de respuesta y cuyos ítems sean representativos del contexto sociocultural. En este sentido, con el empleo de criterios teóricos y empíricos para seleccionar los ítems se planeó obtener una herramienta de evaluación más precisa, libre del sesgo de respuesta AC y con una estructura factorial más estable con respecto a las diferencias de grupo (edad y sexo).

Para la construcción del inventario se partió del conjunto de 100 ítems propuestos por Goldberg (1999) en el IPIP. Al seleccionar los ítems se tuvo en cuenta que cada uno de ellos estuviera relacionado con una faceta psicológica de cada una de las cinco dimensiones de personalidad que evalúa el test. De este modo, se propuso seleccionar seis indicadores por factor con los objetivos de minimizar

los errores Tipo 1 y Tipo 2 asociados a las escalas breves y equilibrar la reducción de los ítems con la cobertura del contenido y con buenas propiedades psicométricas (Baldasaro et al., 2013).

Para el proceso de validación se propuso realizar: (a) un estudio de convergencia con las cinco escalas del NEO-FFI (Costa & McCrae, 1992), debido a que este es considerado un instrumento emblemático en la evaluación del FFM (Chamorro-Premuzic & Furnham, 2009); (b) un análisis de diferencia de grupos según el sexo y la edad de los participantes; y (c) un estudio de validez predictiva con respecto a actividades recreacionales, como el uso de drogas, la irresponsabilidad, la amistad, la erudición/creatividad y la comunicación (Grucza & Goldberg, 2007), por medio del Inventario de Frecuencia de Actividades de la Vida Cotidiana (Goldberg, 1999), instrumento que también forma parte del IPIP.

De este modo, se espera encontrar diferencias de grupo en cuanto al sexo de los participantes, tales como mayores niveles de neuroticismo y amabilidad en las mujeres (McCrae & Terracciano, 2005; Cupani & Lorenzo-Seva, 2016); y, en cuanto a la edad de los participantes, una disminución de los factores extraversión, neuroticismo y apertura, así como un incremento de los factores amabilidad y responsabilidad con el paso del tiempo (Srivastava, John, Gosling & Potter, 2003; Cupani & Lorenzo-Seva, 2016). Con respecto a la validez predictiva de las actividades recreacionales se espera replicar lo reportado en estudios previos (Grucza & Goldberg, 2007; Cupani & Lorenzo-Seva, 2016), es decir, que el uso de drogas correlacione de manera positiva con la extraversión y de manera negativa con la responsabilidad; y que la irresponsabilidad correlacione de manera negativa con la responsabilidad. También, se esperan correlaciones positivas en la actividad de amistad con los factores de amabilidad/extraversión, erudición/creatividad, así como entre intelecto y comunicación, y entre intelecto y amabilidad.

## Método

### *Tipo de estudio*

De acuerdo con la clasificación realizada por Montero y León (2002), el presente trabajo se considera un estudio instrumental, debido a que se encamina al desarrollo de pruebas, incluyendo tanto su diseño como su adaptación.

### *Participantes*

La muestra estuvo compuesta por 910 participantes, 543 de sexo femenino (59.6 %) y 367 de sexo masculino (40.03 %), con edades comprendidas entre los 15 y 80 años ( $M = 29.52$ ;

$DT = 12.25$ ), pertenecientes a la ciudad de Córdoba Capital, Argentina. Con respecto al nivel de instrucción de los participantes, el 6.6 % estaba cursando una carrera de posgrado, el 19.6 % contaba con estudios universitarios completos, el 51.3 % tenía estudios universitarios incompletos, el 12.6 % tenía secundaria completa, el 7.4 %, secundaria incompleta, y el 3 % restante, primaria completa. Solo un caso no había finalizado los estudios primarios.

En cuanto a la ocupación actual de los participantes, la conformación de la muestra fue la siguiente: 40 % estudiantes (28.80 % universitarios y 11.20 % de nivel secundario), 32.3 % empleados, 17 % profesionales (10 % en relación de dependencia y 7 % independientes sin empleados a cargo), 4 % jubilados/pensionados y trabajadores informales (2 % de cada categoría) y el 4 % restante se completó entre dueños o socios de empresas, técnicos independientes o en relación de dependencia, autónomos especializados o no calificados, y obreros. Solo un caso en el momento de la toma se encontraba sin ocupación.

El muestreo fue de tipo accidental (Kumar, 2005), porque el estudio se desarrolló en establecimientos donde se obtuvo el permiso de las autoridades. Con respecto a la administración de los instrumentos, todos los participantes ( $n = 910$ ) respondieron el IPIP Five-Factor Domain Scale (Goldberg, 1999), algunos de estos ( $n = 229$ ) respondieron nuevamente este inventario un mes más tarde para evaluar la estabilidad de las puntuaciones (test-retest), 209 participantes respondieron el inventario NEO-FFI para evaluar la validez convergente, y 402 contestaron las seis escalas del Inventario de Frecuencia de Actividades de la Vida Cotidiana para llevar a cabo el estudio de validez predictiva.

### *Instrumentos*

#### *IPIP Five-Factor Domain Scale (Goldberg, 1999).*

Este instrumento consta de 100 ítems que definen cinco dominios de personalidad: Estabilidad Emocional (EE), Extraversión (E), Intelecto (I), Amabilidad (A) y Responsabilidad (C). Cada dominio es medido por 20 ítems y está redactado en forma de frases que describen comportamientos típicos de las personas. En su aplicación, se le solicita al sujeto que evalúe el grado de precisión con que cada oración lo describe por medio de una escala de cinco opciones de repuesta (desde *muy en desacuerdo* con esta descripción de mí mismo, hasta *muy de acuerdo* con esta descripción de mí mismo). Goldberg (1999) reporta valores alfa de Cronbach de .88 a .91 para las escalas IPIP y una media de la correlación entre las escalas IPIP y los 100 marcadores unipolares de .70, o .78 cuando se utilizó la correlación por la atenuación.



*Inventario de Personalidad NEO-FFI (Costa & McCrae, 1992).*

Este inventario se compone por 60 ítems que ofrecen una medida rápida y general de los cinco grandes factores de la personalidad. Cada escala está compuesta por 12 elementos que describen comportamientos típicos de una persona y permiten medir cada uno de los factores. El formato de respuesta es una escala tipo Likert con cinco opciones de respuesta, que van desde *total desacuerdo* con esa frase hasta *totalmente de acuerdo* con ella (Costa & McCrae, 1992). Los índices de confiabilidad en el presente estudio fueron de  $\alpha = .82$  para Extraversión,  $\alpha = .70$  para Amabilidad,  $\alpha = .80$  para Responsabilidad,  $\alpha = .83$  para Neuroticismo, y  $\alpha = .74$  para Apertura a la experiencia.

*Inventario de Frecuencia de Actividades de la Vida Cotidiana (Goldberg, 1999).*

Para el presente trabajo se utilizaron seis escalas de este inventario, compuestas por 54 ítems que miden seis tipos de actividades recreativas, distribuidas de la siguiente forma: dos actividades relativamente indeseables –uso de drogas (14 ítems) e irresponsabilidad (7 ítems)–, dos actividades deseables –creatividad (11 ítems) y amistad (8 ítems)–, y dos actividades relativamente neutras –comunicación (8 ítems) y erudición (6 ítems)–. Para su aplicación, se le solicita al participante que mencione con qué frecuencia ha efectuado ciertas tareas (por ejemplo, asistir a una biblioteca pública) durante el último año, por medio de una escala tipo Likert de cinco opciones de respuesta que va desde 1 = nunca en mi vida, 2 = no en el último año, 3 = una vez o dos veces en el último año, 4 = tres o más veces en el último año, pero no más que 15 veces, hasta 5 = más de 15 veces en el último año. Los valores de fiabilidad de estas escalas varían desde .70 a .89 según Grucza y Goldberg (2007), y en esta investigación variaron entre  $\alpha = .67$  (comunicación) y  $\alpha = .87$  (uso de drogas).

### *Procedimiento*

La recolección de los datos se realizó tanto de manera colectiva como individual. En la forma colectiva, estudiantes de nivel secundario y de diferentes carreras de la Universidad Nacional de Córdoba (UNC) completaron los instrumentos en las aulas de dictado de clases, con previa autorización de los profesores y autoridades correspondientes, y con el consentimiento informado de los participantes; y, de manera individual, mediante el contacto con amigos, familiares, compañeros de trabajo y conocidos de los investigadores que quisieran participar en el estudio. A cada uno se le explicó previamente el objetivo del trabajo y el tiempo estimado para completar los instrumentos, que

podía variar entre una, dos y tres semanas. Asimismo, se les informó el tiempo aproximado que implicaba responder a cada instrumento y que una vez empezaran a responder un test concreto, se esperaba que lo terminaran de responder sin hacer descansos; después de aceptar, se les entregaba en sobre cerrado un consentimiento informado y los instrumentos correspondientes.

### *Análisis de datos*

Se utilizó la rutina de Análisis de los Valores Perdidos del software SPSS, versión 19.0, para evaluar el patrón de los valores perdidos (Tabachnick & Fidell, 2007). Posteriormente, se dividió la muestra en dos partes de manera aleatoria: la primera mitad ( $n = 455$ ) se seleccionó para identificar los mejores marcadores de los cinco factores, mientras que la segunda mitad ( $n = 455$ ) se reservó para conducir el análisis factorial confirmatorio no restringido de los 30 ítems seleccionados en la primera fase (Ferrando & Lorenzo-Seva, 2000).

Para identificar los 30 ítems principales, se analizaron los 100 ítems del IPIP Five-Factor Domain Scale (Goldberg, 1999) por medio de un análisis factorial exploratorio (AFE), habiendo controlado previamente la varianza debida a la aquiescencia (Ferrando, Lorenzo-Seva & Chico, 2009), con el programa MATLAB y FACTOR 9.3 (Lorenzo-Seva & Ferrando, 2013). También, se consideró como criterio externo la correlación entre los ítems de una dimensión y los puntajes totales de la misma dimensión del NEO-FFI. Como último criterio, se procuró que cada escala estuviera representada por diferentes contenidos. Con la segunda mitad de la muestra se parcializó la varianza debida a la aquiescencia y se factorizó la matriz de correlaciones residual mediante el LISREL 8.5.

Posteriormente, se realizó un AFE con la muestra total ( $N = 910$ ) a partir del cual se estimaron las saturaciones factoriales de los ítems y los pesos factoriales, para luego determinar las puntuaciones factoriales. Además, se calculó la correlación corregida entre la respuesta al ítem y las puntuaciones factoriales. Al calcular las puntuaciones factoriales se informa de la puntuación factorial en los cinco factores, más la puntuación en aquiescencia. Para poder controlar la aquiescencia fue necesario que el test contuviera ítems revertidos.

Cabe indicar que la interpretación de las puntuaciones libres de aquiescencia se consigue calculando las puntuaciones factoriales (en vez de las puntuaciones obtenidas como la mera suma de respuestas a los ítems). De esta manera, las puntuaciones en contenido están libres de aquiescencia. Por tanto, se puede interpretar que los individuos con

puntuaciones extremas en aquiescencia no han respondido con seriedad el test.

También, cabe anotar que para el estudio de consistencia interna se calcularon las fiabilidades de las puntuaciones factoriales (Mislevy & Bock, 1990), mientras que para determinar la estabilidad temporal se utilizó el método test-retest. Con respecto a la estimación de la validez convergente se analizaron las correlaciones entre los puntajes factoriales y los puntajes directos de las cinco escalas del NEO-FFI, teóricamente semejantes.

Posteriormente, se realizó un estudio de evidencia de validez concurrente comparando los puntajes medios de las escalas entre varones, mujeres, jóvenes y adultos, mediante un análisis multivariado 2 (sexo) x 2 (edad) de la varianza (MANOVA). Con este análisis se buscó comparar si las puntuaciones factoriales promedio de los grupos coinciden teóricamente con los perfiles de personalidad identificados por la literatura (Cupani, Pilatti, Urrizaga, Chincolla & de Minzi, 2014). Siguiendo lo propuesto en estudios anteriores (McCrae et al., 2000), se incluyó en el grupo joven a los participantes menores de 30 años.

Por otra parte, para estimar el tamaño del efecto se calculó el coeficiente eta cuadrado ( $\eta^2$ ), y se tomó como criterio lo sugerido por Cohen (1992), es decir, los tamaños del efecto ( $\eta^2$ ) 1 %, 10 % y 25 % son considerados pequeños, medianos y grandes, respectivamente.

Finalmente, se realizó un análisis de regresión múltiple (método *enter*) con el objetivo de evaluar la capacidad predictiva de los cinco factores sobre la variable dependiente "Actividades recreativas", considerando submuestras según sexo y edad de los participantes. En función del tamaño de la muestra, se calculó la potencia estadística de cada modelo mediante el programa G\*Power 3.1 (Faul, Erdfelder, Buchner & Lang, 2009).

## Resultados

A continuación se presentan los resultados de: (a) el análisis de los casos perdidos, (b) el análisis factorial exploratorio y confirmatorio, (c) la fiabilidad y (d) la evidencia de validez convergente y predictiva del instrumento.

### *Análisis de casos perdidos*

El porcentaje de casos perdidos no superó el 5 %, por lo que se decidió imputar los datos faltantes por una medida de tendencia central (moda), debido a que este método proporciona un equilibrio atractivo entre precisión y simplicidad conceptual (Shrive, Stuart, Quan & Ghali, 2006). La elección de imputar por la moda se debió a que se procuró contar

con las cinco opciones (discretas) de respuesta propias de la escala para estimar las correlaciones policóricas.

### *Desarrollo del inventario de 30 ítems*

#### *Análisis factorial exploratorio (AFE).*

En primer lugar, se analizaron los 100 ítems mediante una matriz de correlaciones policóricas, para lo cual se utilizó el método de estimación de mínimos cuadrados no ponderados (ULS) debido a que es un estimador robusto que proporciona estimaciones menos sesgadas y resulta adecuado cuando la muestras no son muy grandes (Ferrando & Lorenzo-Seva, 2013). La medida de adecuación muestral de Káiser-Mayer-Olkin (.883) indicó la factibilidad de realizar el análisis factorial.

Posteriormente, se realizó la extracción de cinco factores y se utilizó la rotación Varimax, ya que se espera que a nivel teórico las dimensiones de personalidad sean independientes en la población (es decir, ortogonales). Para esto, se controló previamente la varianza debida a AC (Lorenzo-Seva & Ferrando, 2009), y, como las escalas de los 100 ítems IPIP son parcialmente equilibradas, se utilizó el procedimiento propuesto por Lorenzo-Seva y Ferrando (2009): (a) se retienen el número esperado de factores según el contenido y un factor adicional relacionado con la AC; (b) se calcula el primer centroide y se toma como una estimación del peso factorial de cada ítem sobre el factor AC; (c) se utiliza este conjunto de estimaciones (una estimación por ítem) como un objetivo en una rotación congruente para calcular la carga factorial de cada contenido del ítem en el factor AC; y (d) una vez que el factor relacionado con AC está disponible, se retira la varianza explicada por este factor y se procede a identificar los factores relacionados con el contenido del instrumento sobre la matriz residual. De este estudio, se observó que las cargas factoriales en el primer factor AC variaron entre .00 y .24, mientras que en los cinco factores de contenidos variaron entre -.03 y -.71 para Amabilidad, entre -.04 y .63 para Extraversión, entre .08 y .62 para Responsabilidad, entre .14 y -.84 para Estabilidad Emocional, y entre .00 y -.50 para Intelecto.

Después de esto, se efectuó una correlación (rango de Spearman) entre los 20 ítems de una dimensión (Estabilidad Emocional, por ejemplo) y el puntaje directo del factor teóricamente semejante del NEO-FFI (Neuroticismo, por ejemplo). Así, las correlaciones ítem-factor variaron entre -.50 y .50 para Extraversión, entre -.50 y .47 para Amabilidad, entre -.66 y .57 para Responsabilidad, entre -.44 y .73 para Estabilidad Emocional, y entre -.45 y .52 para Intelecto.

Por último, se procuró que cada dimensión estuviera representada por diferentes contenidos de facetas de personalidad,

y como resultado de la combinación de estos criterios (ítems que presentaban menor AC, mayor carga factorial y correlación ítem-factor y que representaran las diferentes facetas de personalidad), se seleccionaron los mejores 30 ítems. En función de esto, se observa que el IPIP-R-30, una vez controlado el efecto de la AC, presenta una estructura factorial simple y ortogonal.

#### *Análisis factorial confirmatorio (AFC).*

Con base en el análisis factorial exploratorio, se seleccionaron los principales 30 ítems que combinaron las tres estrategias mencionadas previamente y se utilizaron para el análisis factorial confirmatorio no restringido (Ferrando & Lorenzo-Seva, 2000), haciendo uso de la segunda muestra.

Para esto, se parcializó la varianza debida a la AC y se factorizó la matriz de correlaciones residual mediante LISREL.

Para evaluar el ajuste de los modelos se utilizaron los criterios ya mencionados, y, en función de las saturaciones factoriales obtenidas en el AFE, se seleccionó un ítem como marcador de cada factor. Los estadísticos de bondad de ajuste obtenidos indicaron que el modelo se ajusta de forma adecuada a los datos (GFI = 0.96; CFI = 0.94; RMSEA = 0.05). Después de esto, se utilizó la muestra total (N = 910) para estimar las saturaciones factoriales de los ítems, y los pesos factoriales para estimar las puntuaciones factoriales. De este modo, el AFC sugiere que los 30 ítems del cuestionario presentan un ajuste a los datos adecuado. En la Tabla 1 se pueden observar las cargas factoriales de los factores de contenido.

Tabla 1.  
*Estructura factorial rotada e índices descriptivos*

	Ítem	M	DS	rc	AC	EX	EE	RE	AM	IN
1	IPIP21	3.57	0.93	.49	-.05	.70	-.18	.11	.14	-.09
2	IPIP26	3.70	0.88	.55	.00	.74	-.05	.08	.17	.16
3	IPIP66	2.83	0.88	.40	.10	-.58	-.06	-.11	-.10	-.13
4	IPIP65	2.60	1.03	.35	.11	.56	.10	-.15	-.18	.15
5	IPIP16	2.98	1.09	.38	.29	-.56	-.18	.20	-.05	-.11
6	IPIP76	3.47	0.96	.47	.14	.69	-.13	.09	.16	-.01
7	IPIP14	3.29	1.07	.40	.15	-.03	-.63	.03	-.06	.05
8	IPIP53	2.74	1.06	.56	.26	-.16	.70	-.04	-.03	-.03
9	IPIP29	2.71	1.05	.47	.21	-.05	.66	.01	.05	-.01
10	IPIP69	2.66	1.13	.60	.12	.00	.78	.00	-.15	-.04
11	IPIP89	2.72	1.10	.36	.20	-.05	.52	-.19	.29	.07
12	IPIP99	2.98	1.14	.45	.13	.01	.62	-.10	-.08	-.10
13	IPIP33	2.64	1.09	.33	.24	-.01	.22	-.55	.00	.14
14	IPIP38	2.43	1.03	.45	.24	-.10	.09	-.68	-.10	-.09
15	IPIP62	3.88	0.94	.29	.26	.05	-.02	.47	.13	-.12
16	IPIP82	3.56	0.85	.53	.23	.03	-.09	.72	.05	.10
17	IPIP88	3.26	0.94	.44	.14	-.02	.07	.65	.11	.02
18	IPIP92	3.61	0.83	.42	.13	.05	.04	.64	.09	.14
19	IPIP2	2.22	1.14	.20	.25	.05	-.04	-.12	-.43	-.07
20	IPIP42	2.08	0.97	.46	.07	-.11	.06	-.01	-.75	-.01
21	IPIP51	3.78	1.03	.48	.32	.03	.18	.03	.64	.06
22	IPIP57	3.84	0.85	.53	.16	.09	-.11	.06	.78	.12
23	IPIP81	3.51	0.98	.59	.13	.08	.03	.07	.74	.02
24	IPIP87	3.79	0.85	.55	.12	.13	-.06	.13	.73	-.01
25	IPIP5	3.64	0.88	.32	.18	.10	-.06	.25	-.09	.58
26	IPIP15	2.37	1.02	.39	.28	-.07	-.02	-.10	.03	-.67
27	IPIP25	3.83	0.91	.35	.18	.22	-.05	-.11	.15	.60
28	IPIP70	2.19	0.98	.32	.20	-.18	.04	.01	-.14	-.57
29	IPIP80	2.85	1.08	.32	.04	-.04	-.05	.10	.11	.56
30	IPIP84	3.85	0.87	.25	.08	-.16	-.13	.08	.31	.49

*Nota.* M = Media; DS = Desviación Estándar;  $r_c$  = Correlación corregida; AC = Aquiescencia; EX = Extraversión; EE = Estabilidad Emocional; RE = Responsabilidad; AM = Amabilidad; IN = Intelecto. En negrita se presentan las saturaciones superiores a 0.30 en valores absolutos.

*Fiabilidad.*

Los índices de discriminación para los 30 ítems variaron entre .20 a .60 (mediana de .43), de los cuales tres ítems presentaron valores por debajo del punto de corte (véase Tabla 3). Los coeficientes de fiabilidad variaron entre .79 a .86, y para la estabilidad temporal variaron entre  $r = .71$  en el factor Extraversión y  $r = .80$  en el factor Responsabilidad (véase Tabla 2). Si bien la fiabilidad de las escalas fue satisfactoria, tanto en lo que respecta a la consistencia interna como a la estabilidad temporal se puede observar que la versión de 30 ítems, en comparación con la escala de 50 ítems, registra una pérdida, en promedio, de un 4.82 % en la consistencia interna y de un 5.75 % en la estabilidad temporal.

*Evidencias de validez.*

En el estudio de validez de convergencia, la mediana de las correlaciones fue de  $r = .61$ , con un rango de  $r = .45$  para el factor Amabilidad a  $r = .71$  para Responsabilidad (véase Tabla 2). En el estudio de diferencia de grupos se observó un efecto pequeño a mediano, significativo para sexo y edad. Específicamente, las mujeres presentaron valores más altos que los varones en Amabilidad ( $\eta^2 = .09$ ), mientras que los varones presentaron valores más altos en Estabilidad Emocional ( $\eta^2 = .01$ ); y, con respecto a la edad, se observó que los jóvenes presentaron valores superiores a los adultos en Extraversión ( $\eta^2 = .01$ ), mientras que los adultos lo hicieron en Responsabilidad ( $\eta^2 = .01$ ). De esta manera, se observa que los cinco marcadores del IPIP-R-30 evidenciaron una validez de convergencia con las escalas del NEO-FFI menor (de un 5.85 %) en comparación con la versión de 50 ítems.

Por otra parte, para el estudio de validez test-criterio en la submuestra de mujeres jóvenes, los rasgos de personalidad explicaron entre un 10 % y 18 % de la varianza total (véase Tabla 3), donde rasgos que se relacionan con el Uso de drogas son la Responsabilidad ( $\beta = -.19$ ) y la Extraversión ( $\beta = .24$ ); mientras que el la Erudición se relaciona con el Intelecto ( $\beta = .32$ ); la Comunicación con la Amabilidad ( $\beta = .18$ ) y el Intelecto ( $\beta = .28$ ); y la Amistad con la Extraversión ( $\beta = .26$ ), la Amabilidad ( $\beta = .19$ ), la Responsabilidad ( $\beta = .19$ ) y el Intelecto ( $\beta = .19$ ). Asimismo, en la submuestra de mujeres adultas, los rasgos de personalidad explicaron entre un 13 % y 18 % de la varianza, y los rasgos que se relacionaron con el Uso de drogas fueron la Responsabilidad ( $\beta = -.24$ ) y la Extraversión ( $\beta = .21$ ), mientras que la Irresponsabilidad correlacionó con la Responsabilidad ( $\beta = -.35$ ); y la Creatividad y la Amistad con el Intelecto ( $\beta = .30$  y  $\beta = .30$ , respectivamente).

También, en la submuestra de varones jóvenes, los rasgos explicaron entre un 19 % y 29 % de la varianza, y los rasgos que correlacionaron fueron la Erudición con el Intelecto ( $\beta = .40$ ); la Comunicación con la Extraversión ( $\beta = .20$ ) y el Intelecto ( $\beta = .29$ ); y la Amistad con la Extraversión ( $\beta = .32$ ), la Amabilidad ( $\beta = .25$ ) y el Intelecto ( $\beta = .25$ ). Por último, en la submuestra de varones adultos se explica entre un 15 % y 38 % de la varianza, y los rasgos que correlacionan son el Uso de droga con la Extraversión ( $\beta = .27$ ) y la Responsabilidad ( $\beta = -.21$ ); la Irresponsabilidad con la Extraversión ( $\beta = .20$ ) y la Responsabilidad ( $\beta = -.36$ ); la Erudición con la Extraversión ( $\beta = .23$ ), la Amabilidad ( $\beta = .24$ ), la Estabilidad Emocional ( $\beta = .28$ ) y el Intelecto ( $\beta = .39$ ); la Creatividad con la Estabilidad ( $\beta = .39$ ) y el Intelecto ( $\beta = .24$ ); y la Amistad con la Extraversión ( $\beta = .30$ ) y la Amabilidad ( $\beta = .38$ ).

Tabla 2.

*Índices de fiabilidad, estudio de convergencia y diferenciados según sexo y edad*

Cinco factores	Fiabilidad			Sexo						Edad					
				Femenino (n = 543)		Masculino (n = 367)		Jóvenes (n = 579)		Adultos (n = 331)					
	FCR	T-RT	Conv	M	DT	M	DT	$F_{1,906}$	$\eta^2$	M	DT	M	DT	$F_{1,906}$	$\eta^2$
<b>IPIP-30</b>															
Extraversión	.83	.71	.67	49.96	10.08	50.07	9.93	0.32	.00	50.74	10.40	48.72	9.17	7.88	.01
Amabilidad	.84	.76	.45	52.61	9.00	46.11	10.17	90.61	.09	50.40	9.80	49.27	10.34	0.73	.00
Responsabilidad	.83	.80	.71	49.65	9.82	50.52	10.26	1.46	.00	49.45	10.03	50.96	9.89	4.62	.01
Estabilidad Emocional	.86	.75	-.67	49.17	10.05	51.20	9.81	8.35	.01	49.79	10.11	50.34	9.80	0.36	.00
Intelecto	.79	.75	.61	49.50	9.74	50.72	10.34	3.51	.00	49.88	10.06	50.18	9.91	0.15	.00

Nota. N = 910; FCR = fiabilidad con componente rotados; T-RT = test-retest; Conv. = convergencia con las escalas del NEO-FFI. \*  $p \leq .05$ ; \*\*  $p \leq .001$ .



Tabla 3.  
*Predicción de las actividades recreativas a partir de los cinco factores*

Predictor	Mujeres jóvenes			Mujeres adultas			Hombres jóvenes			Hombres adultos		
	$\beta$	t	p	$\beta$	t	p	$\beta$	t	p	$\beta$	t	p
<b>Uso de drogas (<math>\alpha = .87</math>)</b>												
Extraversión	.24	2.74	.01	.21	2.22	.03	.04	0.35	.72	.27	2.41	.02
Amabilidad	-.04	-0.50	.62	-.17	-1.82	.07	-.05	-0.50	.62	-.18	-1.61	.11
Responsabilidad	-.19	-2.14	.03	-.24	-2.50	.01	-.22	-2.17	.03	-.21	-1.96	.05
Estabilidad Emocional	-.04	-0.50	.61	-.10	-1.02	.31	-.15	-1.46	.15	.02	0.22	.83
Intelecto	.04	0.48	.63	.12	1.25	.21	.01	0.14	.89	.05	0.41	.68
F	(5.122) 2.658*			(5.96) 4.691**			(5.94) 1.302			(5.71) 2.826*		
R <sup>2</sup>	.10			.18			.06			.17		
1- $\beta$	.68			.88			.35			.67		
<b>Irresponsabilidad (<math>\alpha = .77</math>)</b>												
Extraversión	.14	1.62	.11	.10	1.07	.29	.06	0.60	.55	.21	1.98	.05
Amabilidad	-.03	-0.30	.76	-.07	-0.72	.47	-.05	-0.50	.62	.13	1.17	.25
Responsabilidad	-.16	-1.83	.07	-.35	-3.64	.00	-.18	-1.74	.09	-.36	-3.41	.00
Estabilidad Emocional	-.15	-1.74	.08	-.15	-1.54	.13	.01	0.11	.91	-.07	-0.63	.53
Intelecto	-.03	-0.39	.70	.06	0.58	.56	.16	1.58	.12	.18	1.68	.10
F	(5.122) 1.914			(5.96) 4.212***			(5.94) 1.568			(5.71) 4.452**		
R <sup>2</sup>	.10			.18			.08			.24		
1- $\beta$	.68			.88			.46			.88		
<b>Erudición (<math>\alpha = .71</math>)</b>												
Extraversión	-.09	-1.08	.28	.07	0.72	.47	-.11	-1.22	.23	.23	2.36	.02
Amabilidad	-.04	-0.44	.66	.04	0.37	.71	.14	1.52	.13	.24	2.46	.02
Responsabilidad	.08	0.95	.35	-.14	-1.37	.17	-.07	-0.70	.48	.05	0.52	.60
Estabilidad Emocional	.16	1.82	.07	.14	1.45	.15	.01	0.13	.90	.28	2.95	.00
Intelecto	.32	3.79	.00	.25	2.54	.01	.40	4.20	.00	.39	4.14	.00
F	(5.122) 3.672**			(5.96) 2.254			(5.94) 5.077***			(5.71) 8.636***		
R <sup>2</sup>	.13			.11			.21			.38		
1- $\beta$	.83			.59			.93			.99		
<b>Comunicación (<math>\alpha = .67</math>)</b>												
Extraversión	.02	0.27	.79	.17	1.70	.09	.20	2.08	.04	.11	1.01	.32
Amabilidad	.18	2.07	.04	.05	0.49	.62	.18	1.93	.06	.36	3.11	.00
Responsabilidad	.07	0.78	.44	-.17	-1.68	.10	.00	-0.05	.96	.09	0.81	.42
Estabilidad Emocional	.10	1.13	.26	-.05	-0.45	.65	-.07	-0.73	.47	-.08	-0.73	.47
Intelecto	.28	3.31	.00	.20	1.96	.05	.29	2.97	.00	.12	1.08	.29
F	(5.122) 3.570**			(5.96) 2.100			(5.94) 4.431***			(5.71) 2.430*		
R <sup>2</sup>	.13			.10			.19			.15		
1- $\beta$	.83			.56			.89			.6		
<b>Creatividad (<math>\alpha = .76</math>)</b>												
Extraversión	.13	1.45	.15	.13	1.37	.17	.17	1.68	.10	.19	1.79	.08
Amabilidad	.05	0.55	.59	-.10	-1.07	.29	.06	0.61	.54	-.14	-1.28	.20
Responsabilidad	-.01	-0.14	.89	-.12	-1.25	.21	-.09	-0.87	.39	-.20	-1.93	.06
Estabilidad Emocional	.11	1.18	.24	-.08	-0.81	.42	.00	0.03	.98	.39	3.83	.00
Intelecto	.21	2.37	.02	.30	3.04	.00	.17	1.68	.10	.24	2.36	.02
F	(5.122) 1.856			(5.96) 3.326**			(5.94) 1.770			(5.71) 5.560***		
R <sup>2</sup>	.07			.15			.09			.28		
1- $\beta$	.52			.77			.49			.94		
<b>Amistad (<math>\alpha = 0.75</math>)</b>												
Extraversión	.26	3.21	.00	.18	1.81	.07	.32	3.64	.00	.30	2.86	.01
Amabilidad	.19	2.26	.03	.14	1.40	.16	.25	2.81	.01	.38	3.49	.00
Responsabilidad	.19	2.36	.02	-.03	-0.29	.77	.02	0.21	.84	-.03	-0.33	.74
Estabilidad Emocional	.04	0.45	.66	-.02	-0.20	.84	-.13	-1.49	.14	.12	1.14	.26
Intelecto	.19	2.26	.03	.30	3.03	.00	.25	2.71	.01	.15	1.46	.15
F	(5.122) 5.293***			(5.96) 2.939*			(5.94) 7.498***			(5.71) 5.031***		
R <sup>2</sup>	.18			.13			.29			.26		
1- $\beta$	.96			.71			.99			.92		

## Discusión

En concordancia con el creciente aumento de escalas breves de personalidad, como el Mini-IPIP (Donnellan et al., 2006), por ejemplo, en el presente trabajo se examinaron las propiedades psicométricas del Inventario de Personalidad IPIP-R, en su versión reducida de 30 ítems (IPIP-R30 Five-Factor Domain Scale), en una muestra de ciudadanos argentinos. Los resultados indicaron que el IPIP-R-30, una vez controlado el efecto de la AC, presenta una estructura factorial simple y ortogonal, y el AFC sugiere que los 30 ítems del cuestionario presentan un adecuado ajuste a los datos. Estos resultados coinciden con investigaciones previas que demuestran que los inventarios breves presentan un ajuste a la estructura factorial de los cinco factores (Baldasaro et al., 2013; Donnellan et al., 2006; Soto & John, 2017). Del mismo modo, se observó que los pesos factoriales calculados con el AFE fueron, en promedio, un 4.8 % superiores (entre -0.43 y 0.78) a los observados para la versión de 50 ítems propuesta por Cupani y Lorenzo-Seva (2016).

Por otro lado, la fiabilidad de las escalas fue satisfactoria tanto en lo que respecta a la consistencia interna como a la estabilidad temporal. Sin embargo, en esta versión de 30 ítems, en comparación con la escala de 50 ítems, se observa una disminución de la consistencia interna (en un 4.82 %) y de la estabilidad temporal (en un 5.75 %). Estos resultados coinciden con la literatura específica del tema, en la que se reporta que en las versiones breves de los cuestionarios la consistencia interna suele ser el aspecto más afectado, aunque también se registran disminuciones en la estabilidad temporal y la confiabilidad inter-examinador (Baldasaro et al., 2013; Sibley, 2012).

Igualmente, la validez convergente, discriminante y predictiva también disminuyeron en comparación con las versiones extendidas (Baldasaro et al., 2013); y, en efecto, los cinco marcadores del IPIP-R-30 evidenciaron una validez de convergencia con las escalas del NEO-FFI un 5.85 % menor en comparación con la versión de 50 ítems. Específicamente, en este trabajo se observaron asociaciones altas entre los rasgos de Extraversión, Estabilidad Emocional y Responsabilidad, y en menor medida en Intelecto y Amabilidad; resultados semejantes a los encontrados en otras investigaciones (Gow et al., 2005).

También, con respecto a la diferencia de grupos según el sexo, las mujeres se caracterizaron por presentar niveles más altos en Amabilidad y niveles más bajos en Estabilidad Emocional y en Intelecto, en comparación con los hombres, lo cual reafirma lo propuesto por estudios previos (Cupani et al., 2014; McCrae & Terracciano, 2005). Asimismo, en lo que respecta a las diferencias de

grupo en función de la edad, se observa que los niveles medios de la Extraversión disminuyen con la edad, mientras que los de Responsabilidad aumentan; estos cambios, aunque de efecto pequeño, son similares a los observados en estudios previos con otros inventarios de cinco factores (McCrae et al., 2000; Ledesma, Sánchez & Díaz-Lázaro, 2011; Srivastava et al., 2003), lo cual aporta evidencia de validez concurrente a las escalas IPIP-R-30.

Por otra parte, también se realizó un análisis de regresión múltiple para estimar el grado en que cada escala del IPIP-R-30 contribuye, de manera independiente, a explicar la frecuencia en que los individuos realizan algunas actividades recreativas. Este análisis se realizó en diferentes submuestras, considerando la edad y el sexo de los participantes, y se encontró que las cinco escalas permiten explicar entre un 12 % a un 37 % de la varianza de las actividades recreativas. En líneas generales estos resultados coinciden con los reportados por Gruzca y Goldberg (2007), donde el Uso de drogas se relaciona con la Extraversión; la Irresponsabilidad con la Responsabilidad; la Erudición con el Intelecto; la Comunicación con la Amabilidad y el Intelecto; la Creatividad con la Intelecto; y la Amistad con la Extraversión, la Amabilidad y el Intelecto.

En líneas generales, los resultados alcanzados sugieren que las escalas del IPIP-R-30 poseen adecuadas cualidades psicométricas, sin embargo, existen algunas limitaciones que no se deben dejar de mencionar, como que, de la misma manera que en otros instrumentos breves de personalidad, existe un costo psicométrico de la versión breve del IPIP-R. En este sentido, en comparación con las medidas de los 50 ítems, la versión breve resultó ser menos fiable, ya que se correlaciona con menos fuerza con las escalas del NEO-FFI, y, a nivel de contenido, los ítems están menos representados por las diferentes facetas de personalidad. Asimismo, es importante mencionar que, con respecto al procedimiento, el periodo de tiempo extenso en el que se produjo la administración de los cuestionarios puede haber incrementado el sesgo AC, aunque no se haya recolectado evidencia al respecto en este estudio.

Estas limitaciones, sin embargo, no restringen la utilidad práctica del estudio (Donnellan et al., 2006), ya que una escala breve elimina la redundancia de temas y, por lo tanto, se reducen la fatiga, la frustración y el aburrimiento asociados a responder preguntas muy similares una y otra vez (Burisch, 1984), entre otras ventajas (Gosling et al., 2003; Saucier & Goldberg, 2002).

### *Implicancias prácticas*

Si bien existe una pequeña disminución de ciertas propiedades psicométricas en la versión IPIP-R-30 (menor al

6 %), en comparación con la versión de 50 ítems, el ahorro de tiempo que implica administrar una escala reducida en ciertos estudios en los cuales minimizar el tiempo de evaluación y la fatiga del encuestado son aspectos vitales que permite compensar esta pérdida. Así, cuando se administra un inventario de 60 ítems se estima que la mayoría de los participantes lo pueden completar en entre cuatro y diez minutos, mientras que uno de 30 ítems requiere entre tres y cinco minutos, y uno de 15 ítems, entre dos y tres minutos (Soto & John, 2017).

Algunos ejemplos de estudios en los cuales resulta de mayor conveniencia la utilización de escalas breves incluyen las encuestas a gran escala, diseñadas para evaluar diferentes constructos; los estudios longitudinales, que requieren que cada participante complete la misma medida de personalidad en diferentes tiempos; los estudios de laboratorio, en los que parte del tiempo se debe destinar a las manipulaciones experimentales y observaciones de la conducta; y los estudios piloto o investigaciones en las cuales algunos factores, como el aburrimiento, la fatiga o la falta de interés, pueden generar un patrón de respuesta descuidado o al azar (Credé et al., 2012; Milojev et al., 2013; Soto & John, 2017; de Vries, 2013).

Por otro lado, algunos estudios han evidenciado que las pruebas de personalidad administradas a muestras de sujetos con bajos niveles educativos, con inteligencia inferior a la media, o de adolescentes, preadolescentes y adultos mayores presentan los peores ajustes al modelo de personalidad de los cinco factores (Meisenberg & Williams, 2008; Rammstedt, Goldberg & Borg, 2010; Ross & Mirowsky, 1984; Soto, John, Gosling & Potter, 2008; Vigil-Colet, Lorenzo-Seva & Morales-Vives, 2015). De hecho, Ramstedt et al. (2010) han demostrado que la estructura factorial del FFM solo se mantiene a través de los diferentes niveles educativos cuando el sesgo de AC es removido. En esta misma línea, Soto et al. (2008) describen resultados similares en cuanto a la diversidad etaria, y, por lo tanto, la validez de las pruebas de personalidad en estas subpoblaciones puede verse comprometida debido a los altos niveles de AC característicos de estos grupos, por lo que resulta relevante que en la evaluación de dichos sujetos se utilicen escalas libres del sesgo de AC, tales como el IPIP-R-30.

En función de lo anterior, se consideró adecuado seleccionar seis ítems por escala, debido a que investigaciones recientes indican que el uso de la forma muy corta (dos ítems, por ejemplo) se asocia con el aumento de los errores Tipo 1 y Tipo 2 (Credé et al., 2012; Krueger et al., 2013). De esta manera, queda conformado un instrumento con una estructura ortogonal simple de cinco factores, índices de fiabilidad (consistencia interna y estabilidad temporal) satisfactorios, evidencia de validez convergente con las

escalas del NEO-FFI, diferencias de grupos según género y edad, y validez predictiva de actividades recreativas.

En síntesis, los resultados del presente estudio sugieren que ambas versiones representan una alternativa válida para medir los rasgos de personalidad en nuestra población, y que las escalas del IPIP-R-30 poseen adecuadas propiedades psicométricas y pueden ser utilizadas con relativa confianza como herramienta auxiliar en investigación y en el trabajo profesional (clínico, laboral, etc.). Especialmente, una de las principales ventajas de este instrumento es que las puntuaciones que se obtienen están libres del sesgo de la AC.

## Referencias

- Aronson, Z. H., Reilly, R. R., & Lynn, G. S. (2006). The impact of leader personality on new product development teamwork and performance: The moderating role of uncertainty. *Journal of Engineering and Technology Management*, 23(3), 221-247. doi: <https://doi.org/10.1016/j.jengtecman.2006.06.003>
- Baldasaro, R. E., Shanahan, M. J., Bauer, D. J. (2013). Psychometric Properties of the Mini-IPIP in a Large, Nationally Representative Sample of Young Adults. *Journal of Personality Assessment*, 95(1), 74-84. doi: 10.1080/00223891.2012.700466
- Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, 39(3), 214. doi: <http://dx.doi.org/10.1037/0003-066X.39.3.214>
- Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, 19(4), 524-529. doi: <https://doi.org/10.1016/j.lindif.2009.06.004>
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155-159. Recuperado de <http://www2.psych.ubc.ca/~schaller/528Readings/Cohen1992.pdf>
- Costa, P., & McCrae, R. R. (1992). NEO PI-R manual profesional. Odessa, FL: *Evaluación Psicológica Resources*, Inc.
- Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology*, 102(4), 874-888.
- Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology*, 33, 401-415. doi: <http://dx.doi.org/10.1037/h0054677>
- Cupani, M. (2009). El Cuestionario de Personalidad IPIP-FFM: Resultados preliminares de una adaptación en una muestra de preadolescentes Argentinos. *Perspectivas en Psicología*, 6, 51-58. Recuperado de [https://www.researchgate.net/profile/Marcos\\_Cupani/publication/274716961\\_The\\_IPIP-FFM\\_Questionnaire\\_of\\_Personality\\_Preliminary\\_re](https://www.researchgate.net/profile/Marcos_Cupani/publication/274716961_The_IPIP-FFM_Questionnaire_of_Personality_Preliminary_re)

- sults\_for\_the\_adaptation\_in\_a\_sample\_of\_young\_Argentinean\_adolescents/links/5527c3ec0cf2e089a3a1c57d/The-IPIP-FFM-Questionnaire-of-Personality-Preliminary-results-for-the-adaptation-in-a-sample-of-young-Argentinean-adolescents.pdf
- Cupani, M., & Lorenzo-Seva, U. (2016). The development of an alternative IPIP inventory measuring the Big-Five factor markers in an Argentine sample. *Personality and Individual Differences, 91*, 40-46. doi: <http://dx.doi.org/10.1016/j.paid.2015.11.051>
- Cupani, M., Pilatti, A., Urrizaga, A., Chincolla, A., & de Minzi, M. C. (2014). Inventario de Personalidad IPIP-NEO: estudios preliminares de adaptación al español en estudiantes argentinos. *Revista Mexicana de Investigación en Psicología, 6*(1), 55-73. Recuperado de <http://www.revistamexicanadeinvestigacionenpsicologia.com/article/view/185/98>
- De Vries, R. E. (2013). The 24-item Brief HEXACO Inventory (BHI). *Journal of Research in Personality, 47*, 871-880. doi: <http://dx.doi.org/10.1016/j.jrp.2013.09.003>
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment, 18*(2), 192-203. doi: <http://dx.doi.org/10.1037/1040-3590.18.2.192>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods, 41*(4), 1149-1160. doi: <https://doi.org/10.3758/BRM.41.4.1149>
- Ferrando, P. J., & Lorenzo Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica, 21*(2), 301-323. Recuperado de <https://www.redalyc.org/html/169/16921206/>
- Ferrando, P. J., & Lorenzo-Seva, U. (2013). Unrestricted item factor analysis and some relations with item response theory. Recuperado de <http://psico.fcep.urv.es/utilitats/factor/>
- Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A general factor-analytic procedure for assessing response bias in questionnaire measures. *Structural Equation Modeling, 16*(2), 364-381. doi: <https://doi.org/10.1080/10705510902751374>
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe, 7*(1), 7-28. Recuperado de <http://admin.umt.edu.pk/Media/Site/STD/FileManager/OsamaArticle/26august2015/A%20broad-bandwidth%20inventory.pdf>
- Goldberg, L. R. (2001). International Personality Item Pool. Recuperado de <https://iPIP.ori.org/>
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., ... Cloninger, C. R. (2005). The international personality item pool and the future of public domain personality measures. *Journal of Research in Personality, 40*, 84-96. doi: <https://doi.org/10.1016/j.jrp.2005.08.007>
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr., (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504-528. doi: [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Gow, A.J., Whiteman, M. C., Pattie, A., & Deary, I. J. (2005). Goldberg's "IPIP" Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences, 39*, 317-329. doi: <https://doi.org/10.1016/j.paid.2005.01.011>
- Guenole, N., & Chernyshenko, O. (2005). The suitability of Goldberg's Big-Five IPIP personality markers in New Zealand: A dimensionality, bias, and criterion validity evaluation. *New Zealand Journal of Psychology, 34*, 86-96. Recuperado de [https://www.researchgate.net/profile/Nigel\\_Guenole/publication/290196736\\_The\\_suitability\\_of\\_Goldberg's\\_Big\\_Five\\_IPIP\\_personality\\_markers\\_in\\_New\\_Zealand\\_A\\_dimensionality\\_bias\\_and\\_criterion\\_validity\\_evaluation/links/58e77bd7a6fdcc1fda2b967d/The-suitability-of-Goldbergs-Big-Five-IPIP-personality-markers-in-New-Zealand-A-dimensionality-bias-and-criterion-validity-evaluation.pdf](https://www.researchgate.net/profile/Nigel_Guenole/publication/290196736_The_suitability_of_Goldberg's_Big_Five_IPIP_personality_markers_in_New_Zealand_A_dimensionality_bias_and_criterion_validity_evaluation/links/58e77bd7a6fdcc1fda2b967d/The-suitability-of-Goldbergs-Big-Five-IPIP-personality-markers-in-New-Zealand-A-dimensionality-bias-and-criterion-validity-evaluation.pdf)
- Gross, M. N., Zalazar Jaime, M. F., Piccolo, N. V., & Cupani, M. (2012). Nuevos estudios de validación del cuestionario de personalidad IPIP-FFM. X Congreso Latinoamericano de Sociedades de Estadística, Córdoba, Argentina.
- Gruzca, R. A., & Goldberg, L.R. (2007). The Comparative Validity of 11 Modern Personality Inventories: Predictions of Behavioral Acts, Informant Reports, and Clinical Indicators. *Journal of Personality Assessment, 89*(2), 167-187. doi: <https://doi.org/10.1080/00223890701468568>
- Hofstee, W. K. B., ten Berge, J. M. F., & Hendriks, A.A.J. (1998). How to score questionnaires. *Personality and Individual Differences, 25*, 897-909. doi: [https://doi.org/10.1016/S0191-8869\(98\)00086-5](https://doi.org/10.1016/S0191-8869(98)00086-5)
- Javeline, D. (1999). Response effects in polite cultures: a test of acquiescence in Kazakhstan. *Public Opinion Quarterly, 63*(1), 1-28. Recuperado de <https://www.jstor.org/stable/2991267>
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relationship between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36*, 264-277. doi: <https://doi.org/10.1177/0022022104272905>
- Kumar, R. (2005). *Research Methodology: A Step-by-Step Guide for Beginners* (2.<sup>a</sup> ed.). Thousand Oaks, California: SAGE.
- Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing, 13*, 223-248. doi: <https://doi.org/10.1080/15305058.2012.703734>
- Ledesma, R. D., Sánchez, R., & Díaz-Lázaro, C. M. (2011). Adjective checklist to assess the big five personality factors in the Argentine population. *Journal of Personality Assessment, 93*(1), 46-55. doi: <https://doi.org/10.1080/00223891.2010.513708>



- Lorenzo-Seva, U., & Ferrando, P. J. (2009). Acquiescent responding in partially balanced multidimensional scales. *British Journal of Mathematical and Statistical Psychology*, *62*(2), 319-326. doi: <https://doi.org/10.1348/000711007X265164>
- Lorenzo-Seva, U., & Ferrando, P. J. (2013). Factor 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement*, *37*(2), 497-498. doi: <http://dx.doi.org/10.1177/0146621613487794>
- McCrae, R. R., Costa Jr., P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., ... Saunders, P. R. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of personality and social psychology*, *78*(1), 173. doi: 10.1037//0022-3514.78.1.173
- 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. doi: 10.1037/0022-3514.88.3.547
- Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, *44*(7), 1539-1550. doi: <https://doi.org/10.1016/j.paid.2008.01.010>
- Milojević, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality*, *47*, 936-944. doi: <http://dx.doi.org/10.1016/j.jrp.2013.09.004>
- Mislevy, R. J., & Bock, R. D. (1990). *BILOG 3: Item analysis and test scoring with binary logistic models*. Scientific Software International.
- Mlačić, B., & Goldberg, L. R. (2007). An analysis of a cross-cultural personality inventory: The IPIP Big-Five factor markers in Croatia. *Journal of Personality Assessment*, *88*, 168-177. doi: <https://doi.org/10.1080/00223890701267993>
- Montero, I., & León, O. G. (2002). Clasificación y descripción de las metodologías de investigación en Psicología. *International journal of clinical and health psychology*, *2*(3), 503-508. Recuperado de <https://www.redalyc.org/html/337/33720308/>
- 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*(1), 203-212. doi: <https://doi.org/10.1016/j.jrp.2006.02.001>
- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, *44*, 53-61. doi: <https://doi.org/10.1016/j.jrp.2009.10.005>
- Rammstedt, B., Kemper, C. J., & Borg, I. (2013). Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality*, *27*(1), 71-81. doi: <https://doi.org/10.1002/per.1894>
- Ross, C. E., & Mirowsky, J. (1984). Components of depressed mood in married men and women the center for epidemiologic studies' depression scale. *American Journal of Epidemiology*, *119*(6), 997-1004. doi: <https://doi.org/10.1093/oxfordjournals.aje.a113819>
- Saucier, G., & Goldberg, L. R. (2002). Assessing the big five: Applications of 10 psychometric criteria to the development of marker scales. En B. De Raad, M. Perugini (Eds.), *Big five assessment* (pp. 29-58). Seattle, WA: Hogrefe & Huber.
- Sibley, C. G. (2012). The Mini-IPIP6: Item Response Theory analysis of a short measure of the big-six factors of personality in New Zealand. *New Zealand Journal of Psychology*, *41*(3), 21-31. Recuperado de <http://www.psychology.org.nz/wp-content/uploads/Editor-proof.pdf#page=20>
- Shrive, F. M., Stuart, H., Quan, H., & Ghali, W. A. (2006). Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC medical research methodology*, *6*(1), 57. doi: <https://doi.org/10.1186/1471-2288-6-57>
- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69-81. doi: <http://dx.doi.org/10.1016/j.jrp.2017.02.004>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718-737. doi: 10.1037/0022-3514.94.4.718
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change?. *Journal of personality and social psychology*, *84*(5), 1041. doi: 10.1037/0022-3514.84.5.1041
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5<sup>th</sup> ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Vazsonyi, A. T., Ksinan, A., Mikuška, J., & Jiskrova, G. (2015). The Big Five and adolescent adjustment: An empirical test across six cultures. *Personality and Individual Differences*, *83*, 234-244. doi: <http://dx.doi.org/10.1016/j.paid.2015.03.049>
- Vigil-Colet, A., Lorenzo-Seva, U., & Morales-Vives, F. (2015). The effects of ageing on self-reported aggression measures are partly explained by response bias. *Psicothema*, *27*(3), 209-215. doi: 10.7334/psicothema2015.32
- Zheng, L., Goldberg, L.R., Zheng, Y., Zhao, Y., Tang, Y., & Liu, L. (2008). Reliability and concurrent validation of the IPIP Big-Five Factor markers in China: Consistencies in factor structure between internet-obtained heterosexual and homosexual samples. *Personality and Individual Differences*, *45*(7), 649-654. doi: <http://dx.doi.org/10.1016/j.paid.2008.07.009>