



ARTÍCULO DE INVESTIGACIÓN

Estudio psicométrico del Inventario Balanceado de Respuesta Deseable



Isabel María Mikulic^{a,*}, Melina Crespi^{a,b} y Romina Caballero^a

^a Facultad de Psicología, Universidad de Buenos Aires, Buenos Aires, Argentina

^b Consejo Nacional de Investigaciones Científicas y Tecnológicas-CONICET, Buenos Aires, Argentina

Recibido el 16 de diciembre de 2015; aceptado el 27 de junio de 2016

Disponible en Internet el 20 de septiembre de 2016

PALABRAS CLAVE

Deseabilidad social;
Inventario
Balanceado de
Respuesta Deseable
(BIDR-6);
Confiabilidad;
Validez

Resumen La deseabilidad social representa la tendencia psicológica a autoatribuirse cualidades socialmente deseables, rechazando aquellas indeseables. El Inventario Balanceado de Respuesta Deseable (BIDR-6) constituye uno de los instrumentos más utilizados para medirla. El presente estudio indagó la validez factorial y externa a partir de la comparación de grupos, la consistencia interna de las escalas y la existencia de correlaciones con variables sociodemográficas. Se aplicó el instrumento a una muestra intencional y no probabilística de 406 participantes adultos, 188 hombres y 218 mujeres de Buenos Aires, Argentina. Ante la discusión actual sobre la naturaleza y dimensionalidad del constructo se realizaron análisis factoriales confirmatorios de distintos modelos alternativos y se compararon los valores de ajuste y las estructuras factoriales obtenidas, encontrando mayor apoyo empírico para un modelo unidimensional de deseabilidad social con 18 indicadores. Los resultados de esta investigación permiten apoyar la validez de constructo y fiabilidad de la adaptación del BIRD-6 en su versión reducida. También se han obtenido evidencias favorables sobre su validez externa, al comparar la deseabilidad social entre infractores de la ley penal y no infractores. Respecto a la relación entre deseabilidad social y las variables sociodemográficas estudiadas los resultados son consistentes con la literatura previa. Las mujeres presentan valores más elevados de deseabilidad social. A mayor edad de los participantes, mayor la tendencia a mostrar una imagen socialmente aceptable de sí mismos. Mientras que cuanto mayor es el nivel de escolaridad, más bajas son las puntuaciones en deseabilidad social.

© 2016 Publicado por Elsevier España, S.L.U. en nombre de Universitat de Barcelona.

* Autor para correspondencia.

Correo electrónico: mikulic@psi.uba.ar (I.M. Mikulic).

KEYWORDS

Social desirability;
Balanced Inventory of
Desirable Responding
(BIDR-6);
Fiability;
Validity

Psychometric study of Balanced Inventory of Desirable Responding

Abstract Social desirability represents the psychological tendency to self-attributed socially desirable qualities, rejecting those undesirable. Balanced Inventory of Desirable Responding (BIDR-6) is one of the most used instruments to measure it. The present study investigated the factorial, convergent and external validity, internal consistency of the scales and correlations with demographic variables. The instrument was applied to an intentional and non-probabilistic sample of 406 adult participants, 188 men and 218 women residing in Buenos Aires, Argentina. For the current discussion about the nature and dimensionality of the construct, confirmatory factor analysis of different models were performed and the setting values and the factor structures obtained were compared, finding greater empirical support for a one-dimensional model of social desirability with 18 indicators. The results of this research bring support to construct validity and reliability of adaptation BIRD-6 in its reduced version. Also, it has been obtained favorable evidence of external validity, comparing the social desirability between offenders and non-offenders. Regarding the relationship among social desirability and socio-demographic variables studied the results are consistent with previous literature. Women have higher social desirability values. The older the participants are, the higher tendency they show in socially acceptable self-image. On the other hand, higher educational levels tend to report lower scores in social desirability.

© 2016 Published by Elsevier España, S.L.U. on behalf of Universitat de Barcelona.

Introducción

El estudio del sesgo de respuesta es un tema relevante en el campo de la evaluación psicológica. Cronbach (1946, p. 476) señaló su importancia y lo conceptualizó como «cualquier tendencia que causa que el sujeto consistentemente dé una respuesta diferente de la que daría a un mismo contenido si las preguntas se hubieran presentado de diferente manera». En este sentido, la validez y la confiabilidad de una prueba psicológica puede verse afectada, en virtud de la influencia de otras variables diferentes de las medidas.

Dentro de los llamados sesgos de respuesta se han clasificado 3 fenómenos de mayor relevancia: a) aquiescencia, entendida como la tendencia consistente a estar de acuerdo con una afirmación independientemente de su contenido (Cronbach, 1946); b) respuestas extremas, las cuales comprenden la tendencia a responder solamente en los extremos de un continuo psicológico (Cronbach, 1946), y c) deseabilidad social (DS), definida comúnmente como una tendencia a dar autodescripciones positivas (Paulhus, 2002) o bien, de manera más específica, como la necesidad de aprobación y aceptación social y la creencia que esta puede ser alcanzada a través de comportamientos culturalmente aceptados (Marlowe y Crowne, 1961).

La DS constituye uno de los estilos de respuesta o sesgo más estudiado en el marco de la evaluación psicológica, considerando la gran cantidad de situaciones y contextos evaluativos en las cuales el sujeto puede intentar responder de una manera que le resulte beneficiosa. Este aspecto es particularmente relevante en ámbitos como el jurídico y el laboral, que implican situaciones en las que el evaluado podría tender a no informar sobre problemas emocionales y psicológicos que efectivamente presenta, con alguna finalidad utilitaria. También el tipo de instrumento de evaluación que se utilice es un factor relevante, siendo principalmente

los autoinformes los que resultan más sensibles a este fenómeno, pues a través de ellos el evaluado tiene que dar su opinión o referirse a características propias, por lo que pueden presentarse distorsiones en las respuestas, ya sean intencionadas o inconscientes (Nunnally y Bernstein, 1995).

Una de las estrategias de investigación para reducir la distorsión de respuesta se basa en el uso de escalas especialmente construidas y validadas para ese fin. Estas escalas suelen componerse de afirmaciones deseables, pero falsas (p. ej., «*siempre digo la verdad*») y/o de afirmaciones no deseables socialmente, pero ciertas (p. ej., «*a veces hablo sobre cosas que conozco poco o nada*»). Las puntuaciones totales en estas escalas son un indicador de la distorsión deliberada de respuesta (Dodaj, 2012). En consecuencia, los resultados sirven para identificar a los candidatos que exageran sus atributos positivos al autodescribirse. Es asumido que la identificación de este sesgo de respuesta mejora la capacidad de predicción de los instrumentos utilizados (Salgado, 2005).

En razón de ello, se construyeron numerosas escalas para evaluar la DS. Un amplio espectro de las mismas fue incluido dentro de inventarios amplios, como controles de confiabilidad de las respuestas. Tal es el caso de la escala Lie del Inventario Multifásico de Personalidad de Minnesota (MMPI) de Hathaway y McKinley (1951) y de la escala Good Impression del Inventario de Personalidad de California de Gough (1952). Por su parte, Edwards (1957), basándose en el MMPI, desarrolló una Escala de Deseabilidad Social que reflejaba la tendencia general de los sujetos a dar respuestas socialmente deseables en sus autodescripciones. Sin embargo, como mencionan Crowne y Marlowe (1960), la escala de Edwards se basaba en un modelo de desviación estadística y tenía contenidos que se confundían con patología, por lo que estos autores decidieron desarrollar su propio modelo de medición apoyándose en un enfoque diferente. Para tal

fin, seleccionaron ítems de pruebas de personalidad que describían comportamientos socialmente deseables, pero poco frecuentes, y comportamientos socialmente indeseables, pero frecuentes en uno mismo, sin tener en cuenta contenido psicopatológico.

Si bien tanto la Escala de Deseabilidad Social de Edwards (1957) como la de Crowne y Marlowe (1960) parten de modelos conceptuales diferentes, comparten la concepción y evaluación de la DS como un constructo unidimensional. No obstante, esta idea de unicidad del constructo comenzó a ser objeto de reflexión a partir de diversos estudios que pusieron de manifiesto bajas correlaciones entre distintas medidas del mismo (Stöber, Dette y Musch, 2002).

Wiggins (1964), al aplicar el método de análisis factorial en diversas escalas de DS, identificó 2 factores ligados a la respuesta socialmente deseable que denominó: alfa (p. ej., evaluación favorable vs. evaluación desfavorable de sí mismo) y gamma (p. ej., mentira). Desde este estudio, diversos autores intentaron interpretarlos y caracterizarlos. Así, Damarin y Messick (1965) sugirieron que el factor alfa podría reflejar un intento inconsciente y no deliberado de presentarse a sí mismo de la mejor forma, mientras que el factor gamma podía implicar un intento consciente o intencional de sesgar las respuestas a favor de uno mismo.

Sackeim y Gur (1978), por su parte, desarrollaron un instrumento para evaluar la DS, en función de 2 dimensiones: autoengaño, comprendida como la tendencia a describirse positivamente a sí mismo de manera poco realista, y el engaño a otros, definido como el esfuerzo deliberado por agradar y mostrarse a los otros de la mejor manera posible.

Basándose en el modelo propuesto por Sackeim y Gur (1978), Paulhus (1984) desarrolló el Inventario Balanceado de Respuestas Socialmente Deseables (BIDR) para medir 2 facetas de la DS: autoengaño y manejo de impresión. La primera de ellas fue diseñada para evaluar el grado en que las personas exageran atributos positivos y ocultan las características negativas de sí mismas. La segunda se diseñó para medir el grado en que los individuos en forma deliberada resaltan comportamientos que se consideran socialmente deseables.

Conjuntamente con la Escala de Deseabilidad Social de Marlowe-Crowne, el BIDR representa el inventario más utilizado para medir DS (Dodaj, 2012). Entre otros aspectos, la importancia de su utilización radica en que valores altos de deseabilidad pueden sesgar los resultados de investigaciones científicas, así como también de los procesos de evaluación psicológica en diversos ámbitos. Desde el trabajo inicial de Paulhus (1984), el BIDR ha pasado por varias revisiones. Actualmente se encuentran publicadas la sexta versión —con fines de investigación en muestras colectivas— y la séptima versión —para contexto clínico con aplicación individual (Paulhus, 2002).

Si bien Paulhus (1984) propuso un modelo de 2 factores medidos por el BIDR, la evidencia empírica existente en relación con tal estructura bifactorial ha resultado contradictoria e incluso no ha sido apoyada a través de métodos de análisis factorial confirmatorio (AFC). Por tal motivo es que en 1991 reestructura el modelo de DS propuesto, manteniendo la escala original de manejo de impresión, pero desdoblado la dimensión de autoengaño en 2 factores: mejora autoengañosa o autoengaño y negación autoengañosa o autonegación. El primero refiere a la medida

en que los individuos inadvertidamente exageran sus cualidades deseables, mientras que el segundo comprende el grado en que los mismos ocultan o minimizan sus cualidades indeseables (Paulhus y Reid, 1991).

Sobre la base de lo expuesto, resulta particularmente útil poner a prueba el modelo de DS que subyace al instrumento BIDR-6, ya que esto podría tener importantes implicaciones en lo que respecta a la evaluación e interpretación de la DS.

Por otro lado, si bien se cuenta actualmente con diversas adaptaciones y validaciones del BIDR-6 (Hart, Ritchie, Hepper y Gebauer, 2015; D'Amours-Raymond, 2011; Kroner y Weekes, 1996a, b; Moral de la Rubia, García-Cadena y Antona-Casas, 2012), resultan escasos los estudios reanalizados con poblaciones de habla hispana, no hallándose antecedentes de estudios psicométricos de adaptación alguna en Argentina. Con vistas al presente trabajo, es de interés citar que existe una versión del BIDR-6 en español desarrollada en México por Moral de la Rubia et al. (2012). Tras la exhaustiva revisión de dicha versión se pudo comprobar que debido a que contenía gran cantidad de modismos y expresiones poco habituales en nuestro medio, la misma no ofrecía garantías para su utilización. Se comprende la limitación que ello representa dada la importancia de contar con instrumentos validados ecológicamente para la evaluación psicológica y la investigación (Muñiz y Hambleton, 1996; van de Vijver y Hambleton, 1996). Por esta razón, se decidió llevar a cabo la adaptación directamente de la versión original.

El presente estudio tiene por objetivo general evaluar las propiedades psicométricas del BIDR-6 (Paulhus, 1988) adaptado para su uso en población de Buenos Aires, Argentina. Como objetivos específicos se plantean: a) validar la estructura factorial del inventario; b) analizar su validez externa; c) explorar la consistencia interna, y d) establecer correlaciones entre variables sociodemográficas (género, edad, escolaridad) y los puntajes de las subescalas de DS y del total en la muestra estudiada.

Método

Participantes

Se aplicó el instrumento a una muestra intencional y no probabilística de 406 participantes adultos, 188 hombres (46,3%) y 218 mujeres (53,7%), con un rango de edad entre los 18 y los 45 años ($M = 26,6$; $SD = 5,7$) residentes en ciudad y provincia de Buenos Aires. Respecto al nivel educativo, el 4,4% tiene primario incompleto, el 28,3% primario, el 25,4% secundario, el 32% universitario incompleto y el 9,9% universitario.

Por otro lado, para el estudio de la validez externa se evaluó a una submuestra intencional conformada por 72 sujetos adultos hombres infractores de la ley penal, pertenecientes a distintas Unidades Penitenciarias y comisarías de la provincia de Buenos Aires. Se controlaron las variables: sexo, edad y tipo de delito (p. ej., delitos contra la propiedad). Dentro del rango considerado (18 a 45 años) la media de edad fue de 28 años ($SD = 7$). El mayor porcentaje se concentró en la categoría de secundario incompleto (35%), seguido de primario completo (33%). Asimismo, se seleccionó a un grupo

de 77 sujetos hombres no delinquentes, homologando las variables edad y nivel educativo.

Instrumentos

El BIDR-6 (Paulhus, 1988) consta de 40 ítems redactados como proposiciones afirmativas, la mitad en sentido de respuestas socialmente deseables (directos) y la otra mitad en sentido opuesto (inversos). Se responden utilizando una escala tipo Likert que va de 1 (*no verdadero*) a 7 (*muy verdadero*). Está compuesto por 2 factores: autoengaño y manejo de impresión.

Se ha encontrado consistencia interna, por el alfa de Cronbach, que varía de 0,68 a 0,80 para la escala de autoengaño; de 0,75 a 0,86 para la escala de manejo de impresión, y de 0,81 a 0,85 para los 40 ítems (Paulhus y Reid, 1991).

Procedimiento

El BIDR-6 fue traducido mediante el método de traducción directa. Cuatro traductores con conocimientos académicos en Psicología y en Psicometría realizaron la traducción en forma independiente del instrumento original. Con el objetivo de establecer la equivalencia conceptual, las 4 traducciones obtenidas fueron evaluadas por 5 jueces. Cada uno de los miembros del jurado de forma independiente comparó la idoneidad de la adaptación (del inglés al castellano) de cada uno de los ítems. El jurado evaluó la equivalencia de los ítems considerando una escala de 4 puntos, cuyos valores fueron: 1 (*idéntico*), 2 (*bastante similar*), 3 (*bastante diferente*) y 4 (*diferente*). Se obtuvieron los porcentajes de acuerdo entre los distintos miembros del jurado para cada ítem. Se seleccionaron los reactivos que tenían un porcentaje de acuerdo mayor del 80% en la categoría «idéntico». Luego, 2 jueces expertos realizaron una revisión de los ítems conflictivos, efectuando los cambios necesarios en los ítems que así lo requerían, como por ejemplo aquellos que presentaban doble negación o que resultaban ambiguos.

Se realizó una aplicación piloto a una muestra pequeña, antes de administrar el instrumento a la muestra mayor de la población de Buenos Aires.

Los participantes fueron contactados personalmente por los investigadores. Se los invitó a colaborar en el estudio de forma voluntaria y se garantizó la confidencialidad en el tratamiento de los datos.

Análisis de datos

Se llevó a cabo el análisis psicométrico de la versión adaptada del BIDR-6 (de acuerdo con el procedimiento reseñado anteriormente), para lo cual:

- Se realizó un análisis descriptivo de los ítems examinando media, desviación estándar, asimetría y curtosis de cada uno de ellos. Como criterio para evaluar los índices de asimetría y la curtosis se consideraron como excelente valores entre $\pm 1,00$ y adecuados valores inferiores a $\pm 2,00$ (George y Mallery, 2011).

- Se utilizó AFC para verificar el ajuste a los datos del modelo teórico que subyace a la versión adaptada del BIDR-6. La normalidad multivariante se analizó por medio de índices de curtosis multivariada. Para el AFC se seleccionó el método de mínimos cuadrados no ponderados (ULS). Para evaluar el ajuste del modelo se consideraron múltiples indicadores, de acuerdo con las recomendaciones actuales (Byrne, 2009; Hu y Bentler, 1999). Se utilizaron: Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Normal Fit Index (NFI), Relative Fit Index (RFI) y Root Mean Square Residual (RMR). Los índices GFI, AGFI, NFI y RFI pueden variar entre 0 y 1, con 0 indicando ausencia de ajuste y 1 ajuste óptimo. Valores de entre 0,90 y 0,95 indican un ajuste aceptable del modelo a los datos y superiores a 0,95 un ajuste excelente. El índice RMR es considerado aceptable cuando sus valores son inferiores a 0,08 (Hu y Bentler, 1999).
- Se verificó la consistencia interna de los factores obtenidos mediante el cálculo del coeficiente α de Cronbach.
- Se analizó la validez externa a través de la comparación de DS entre infractores de la ley penal y no infractores.
- Se estudió la relación entre DS y las variables sociodemográficas reseñadas en los objetivos del presente estudio.

Los datos fueron analizados mediante los programas informáticos SPSS 19.0. y AMOS 22.0.

Resultados

Estructura factorial confirmatoria del Inventario Balanceado de Respuestas Socialmente Deseables

Se obtuvieron las estadísticas descriptivas de todos los ítems. Para comprobar los supuestos de normalidad de la muestra se realizaron análisis de asimetría y curtosis para cada reactivo. Estos resultados permitieron concluir que del total de ítems, 26 presentaban valores de asimetría y curtosis inferiores a $\pm 1,00$ y 14 inferiores a $\pm 2,00$.

Al indagar la normalidad multivariante se encontraron valores de curtosis multivariada altos (Kline, 2010), indicando incumplimiento del supuesto de normalidad. Debido a ello se seleccionó el método de ULS para estimar el ajuste del modelo, dado que no hace suposiciones acerca de la distribución y es adecuado para variables ordinales (Forero, Maydeu-Olivares y Gallardo-Pujol, 2009; González Álvarez, Abad González y Levi Mangin, 2006; Ximénez, 2007).

Se contrastó en primera instancia el modelo original de 2 factores correlacionados: autoengaño y manejo de impresión con 20 indicadores cada uno (Paulhus, 1984). Los índices de ajuste en general fueron bajos (GFI=0,87; AGFI=0,86; NFI=0,71; RFI=0,70; RMR=0,07), mostrando valores inferiores al corte recomendado de 0,90. Los pesajes de regresión del modelo resultante arrojaron valores aceptables entre 0,31 a 0,72 en 26 ítems y valores bajos (inferiores a 0,30) en los 14 restantes.

En segundo lugar, se contrastó el modelo de 3 factores correlacionados propuesto por Paulhus y Reid (1991): mejora autoengañoso, negación autoengañoso y manejo de impresión. Los índices de ajuste en general fueron más bajos aún que en el modelo de 2 factores (GFI=0,87; AGFI=0,85; NFI=0,69; RFI=0,65; RMR=0,07). Con respecto a los pesajes

Tabla 1 Análisis factorial confirmatorio del BIDR-6: índices de bondad de ajuste de los modelos estimados

Índices	Modelo 2 factores	Modelo 3 factores	Modelo abreviado 2 factores	Modelo abreviado 1 factor
GFI	0,87	0,87	0,95	0,96
AGFI	0,86	0,85	0,94	0,95
NFI	0,71	0,69	0,91	0,93
RFI	0,70	0,65	0,91	0,92
RMR	0,07	0,07	0,06	0,06

AGFI, Adjusted Goodness of Fit Index; GFI, Goodness of Fit Index; NFI, Normal Fit Index; RFI, Relative Fit Index, y RMR, Root Mean Square Residual.

de regresión del modelo resultante, se obtuvieron valores aceptables entre 0,30 a 0,65 en 24 indicadores y valores bajos en los 16 restantes.

Dado que la mayoría de los índices de bondad de ajuste no alcanzaron los valores recomendados y, considerando que muchos ítems presentaron pesajes muy bajos, sugiriendo una débil relación con los factores correspondientes, se puso a prueba un modelo alternativo. Se planteó un modelo abreviado de 18 indicadores agrupados en los 2 factores centrales propuestos por Paulhus: autonegación y manejo de impresión. Como se presenta en la [tabla 1](#), los índices de ajuste fueron mejores que en los modelos precedentes.

Los pesajes de regresión del modelo resultante arrojaron valores significativos en todos los indicadores con valores entre 0,32 a 0,72 ([tabla 2](#)). La correlación entre ambos factores fue alta ($r=0,87$).

Cabe señalar que los ítems que aportan en forma significativa tanto al factor de autoengaño como a manejo de impresión corresponden a lo que Paulhus denominó «tendencia a negar en uno mismo aquellas características

socialmente reprobadas». Contrariamente, el aporte de los reactivos ligados a la «tendencia a atribuirse características favorables o socialmente aprobadas» ha demostrado ser baja. Por esta razón, y, fundamentalmente, atendiendo a la elevada correlación que se obtuvo entre los factores autoengaño y manejo de impresión, se procedió a probar un modelo unidimensional de DS.

Como se aprecia en la [tabla 1](#), los índices de ajuste fueron más elevados que en los modelos bi y tridimensionales de Paulhus y levemente mejores que en el modelo abreviado de 2 factores, mostrando todos los indicadores pesajes de regresión adecuados ([tabla 2](#)).

Consistencia interna

Considerando el modelo de 2 y 3 factores de Paulhus, la consistencia interna obtenida fue baja en todas las escalas obteniéndose valores α inferiores a 0,70 en cada caso.

Tabla 2 Análisis factorial confirmatorio. Saturaciones factoriales de los modelos abreviados del BIDR-6

Ítems		Modelo 2 factores Autoengaño	Modelo 1 factor DS
20.	No siempre sé los motivos de por qué hago lo que hago	0,52	0,52
18.	Algunas veces dudé de mi capacidad como amante	0,46	0,44
16.	Pocas veces valoro las críticas	0,58	0,56
14.	Mis padres no siempre fueron justos al castigarme	0,46	0,43
12.	A veces pierdo oportunidades porque no puedo decidirme a tiempo	0,33	0,32
10.	Me es muy difícil dejar de pensar en una idea preocupante	0,32	0,33
6.	Cuando tocan mis emociones no puedo pensar adecuadamente	0,33	0,32
4.	No siempre fui honesto conmigo mismo	0,45	0,43
2.	Me resultaría difícil dejar alguno de mis malos hábitos	0,32	0,32
Manejo de impresión			
39.	Tengo algunos hábitos muy desagradables	0,72	0,73
33.	A veces manejo excediendo el límite de velocidad	0,37	0,38
31.	Cuando era joven, algunas veces robé cosas	0,65	0,66
29.	Recibí más vuelto del que correspondía sin decirle al vendedor	0,64	0,66
27.	Dije cosas malas de un amigo a sus espaldas	0,58	0,58
25.	A veces trato de vengarme en vez de perdonar y olvidar	0,61	0,61
23.	Hubo ocasiones en las que saqué ventaja de alguien	0,65	0,66
21.	A veces digo mentiras si tengo que hacerlo	0,32	0,34
19.	Me da igual si no le caigo bien a la gente	0,36	0,36

Los números de ítems corresponden a la versión de 40 reactivos del BIDR-6.

Tabla 3 Comparación de los valores medios del BIDR-6 (abreviado) entre infractores y no infractores

Variable	No infractores (n = 77)		Infractores (n = 72)		Valores estadísticos		d de Cohen
	M	DE	M	DE	t	p	
DS	78,77	12,33	87,12	10,42	-3,85	0,00*	-0,81

* p < 0,01.

Tabla 4 Comparación de los valores medios del BIDR-6 (abreviado) según género

Variable	Hombres (n = 188)		Mujeres (n = 218)		Valores estadísticos		d de Cohen
	M	DE	M	DE	t	p	
DS	72,18	16,12	77,98	18,33	-3,24	0,00*	-0,34

* p < 0,01.

Respecto al modelo de 2 factores abreviado, se obtuvieron valores de consistencia interna adecuados: autoengaño ($\alpha = 0,79$) y manejo de impresión ($\alpha = 0,80$).

Por último, al considerar el modelo abreviado unidimensional, se registró una alta consistencia interna ($\alpha = 0,86$).

Evidencias de validez externa del Inventario Balanceado de Respuestas Socialmente Deseables abreviado: evaluación de la deseabilidad social en infractores de la ley y no infractores

En la [tabla 3](#) se muestran los resultados de la comparación de los diferentes puntajes obtenidos entre el grupo de infractores de la ley (n = 72) y el de no infractores (n = 77). Se registró una diferencia significativa en el total del BIDR-6 (18 ítems), con un tamaño del efecto elevado ([Cohen, 1988](#)). Los infractores de la ley penal presentaron puntuaciones medias más elevadas, resultado que apoya la hipótesis del autoengaño como elemento asociado a quienes delinquen.

Deseabilidad social y variables sociodemográficas

Diferencias según género en la deseabilidad social

Para evaluar la existencia de diferencias significativas según género en el BIDR-6 (18 ítems) se trabajó con la muestra total (n = 406) descripta anteriormente. Se utilizó la prueba de la t de Student para muestras independientes, utilizando el puntaje total de DS como variable dependiente y el género como variable de agrupamiento.

Como puede observarse en la [tabla 4](#), las mujeres obtuvieron puntuaciones significativamente más elevadas que los hombres en el total de DS, con un tamaño del efecto pequeño ([Cohen, 1988](#)).

Correlaciones entre deseabilidad social, edad y nivel de instrucción

Como se expone en la [tabla 5](#), se encontró correlación significativa de baja magnitud entre las variables edad y DS. Con respecto al nivel educativo, se obtuvieron correlaciones negativas con el total de DS. Niveles altos de escolaridad correlacionaron con puntuaciones más bajas en DS.

Discusión

El objetivo general de este trabajo consistió en evaluar las propiedades psicométricas del BIDR-6 ([Paulhus, 1988](#)) con el fin de obtener una versión validada ecológicamente, capaz de ser utilizada en población de Buenos Aires, Argentina. Para ello, se propuso analizar la estructura factorial y la validez externa del inventario, explorar su consistencia interna y hallar posibles asociaciones con las variables género, edad y escolaridad.

Ante la discusión actual sobre la naturaleza y dimensionalidad del constructo DS ([Barger, 2002](#); [Leite y Beretvas, 2005](#)), los análisis factoriales confirmatorios efectuados sobre base la de la versión adaptada del BIDR-6, no permiten afirmar que el mismo pueda ser comprendido a partir de 2 factores primarios. Estos factores primarios fueron definidos por [Paulhus \(1984\)](#) como «autoengaño» y «manejo de impresión», y están constituidos por 20 indicadores cada uno, que en total conforman la versión de 40 reactivos del BIDR-6.

Si bien en el estudio realizado, este modelo de 2 factores con 20 indicadores cada uno presentó un mejor ajuste que el modelo de 3 factores redefinido por Paulhus y Reid en 1991 (mejora autoengañoso, negación autoengañoso y manejo de impresión); en ambos casos, los índices de ajuste fueron inferiores al corte recomendado de 0,90, presentando además pesajes bajos en diversos ítems y una pobre consistencia interna.

Tabla 5 Correlaciones entre el BIDR-6 (abreviado), edad y nivel de estudios

	Edad	Nivel de estudios
DS	0,24*	-0,47**

* p < 0,05.

** p < 0,01.

De manera que en la adaptación realizada del BIDR-6 no se recomienda el empleo de un factor de manejo de impresión (con 20 indicadores) y otro de autoengaño (con 20 indicadores), así como tampoco el modelo trifactorial de mejora autoengañoso (con 10 indicadores), negación autoengañoso (con 10 indicadores) y manejo de impresión (con 20 indicadores), por presentar problemas de consistencia interna y reproducción factorial.

Tales resultados son consistentes con la adaptación del BIDR-6 realizada en México, en donde la estructura original de 40 indicadores distribuidos en 2 factores (autoengaño y manejo de impresión) no se reproduce ni por análisis factorial exploratorio ni confirmatorio, siendo además baja la consistencia interna registrada en cada factor (Moral de la Rubia et al., 2012). Similares resultados fueron reportados en otros estudios, en los cuales ha sido necesario reducir el instrumento original descartando aquellos reactivos con cargas factoriales bajas y asegurando un mejor ajuste del modelo teórico que subyace al BIDR-6 (p. ej., versión italiana de 16 ítems de Bobbio y Manganelli, 2011; versión del Reino Unido de 18 ítems de Hart et al., 2015; versión francesa de 21 ítems de D'Amours-Raymond, 2011).

Con respecto a la versión abreviada del BIDR-6 (18 reactivos) propuesta en este estudio, los índices de ajuste del modelo de 2 factores evidenciaron valores acordes con el consenso (Byrne, 2009; Hu y Bentler, 1999). Los pesajes de los ítems fueron adecuados y también la consistencia interna de cada factor. Sin embargo, al analizar la conformación de cada uno de esos factores, los reactivos que se conservaron tanto en autoengaño como en manejo de impresión corresponden a la dimensión de «negación» de atributos socialmente reprobados, no quedando representada la dimensión de «atribución» de características socialmente aceptables. Sumado a ello, se advirtió una elevada correlación ($r=0,87$) entre los 2 factores propuestos por Paulhus: autoengaño y manejo de impresión. De allí que se consideró que un modelo de un solo factor con 18 indicadores podría constituir una solución más parsimoniosa. La bondad del ajuste del modelo unidimensional de DS resultó ser adecuado, presentando índices con valores levemente más altos que el modelo abreviado de 2 factores y mayor consistencia interna.

Los resultados obtenidos respecto a la relación entre DS y las variables sociodemográficas estudiadas son consistentes con lo obtenido en la literatura sobre el tema. Existe un perfil diferencial en la respuesta socialmente deseable según género. Las mujeres presentan valores más elevados en DS, aspecto que se corresponde con lo obtenido en diversos estudios previos (Bobbio y Manganelli, 2011; Gravdal y Sandal, 2006; Moral y Ortiz, 2011; Pérez, Labiano y Brusaska, 2010; Riketta, 2005; Stöber et al., 2002).

Respecto a la edad, se obtuvo correlación significativa, aunque baja, con DS, aspecto consistente con investigaciones previas que registraron que a mayor edad de los participantes mayor la tendencia a mostrar una imagen socialmente aceptable de sí mismos ante los demás, rechazando las cualidades individuales consideradas negativas (Bobbio y Manganelli, 2011; Riketta, 2005; Pérez et al., 2010). En consonancia con tales estudios, la correlación entre DS y nivel educativo resultó ser inversa, mostrando que cuanto mayor es el nivel de escolaridad de los participantes, más bajas son las puntuaciones en DS.

Se han obtenido evidencias favorables sobre la validez externa del BIDR-6 (18 ítems), corroborándose la hipótesis formulada al comparar la DS en sujetos infractores de la ley penal y no infractores. Los datos son consistentes con los obtenidos en diversos estudios previos. Así, Rogers et al. (2002) demostraron que los delincuentes tendían a reducir sus características negativas, ligadas sobre todo a psicopatía, y se presentaban con un estilo social deseable de respuesta, respecto de los no delincuentes. En igual sentido, Fernández y Echeburúa (2006) advirtieron que el autoinforme de conducta disocial en delincuentes estaba afectado por el manejo de impresión.

Como limitaciones del estudio, debido a que los resultados se basan en una muestra no probabilística, los mismos son generalizables a muestras con características similares.

Será importante en el futuro realizar nuevos estudios que profundicen el análisis de las propiedades psicométricas de la adaptación del BIDR-6. Fundamentalmente, resulta prioritario verificar los resultados obtenidos al poner a prueba el modelo teórico subyacente (unidimensionalidad vs. bi o tridimensionalidad) mediante nuevos análisis factoriales confirmatorios con muestras más grandes y representativas. Asimismo, resulta importante aportar nuevas evidencias sobre la validez externa del BIDR-6 analizando las relaciones del mismo con otras pruebas y criterios externos. Un aspecto de relevancia, de cara al futuro, comprende la posibilidad de aplicar el inventario en condiciones reales de demanda social. Esto permitiría probar la validez del BIDR-6 en situaciones en las que los participantes se sientan motivados a presentarse en forma favorable, como puede ocurrir en un contexto de selección de personal o en el ámbito penitenciario en problemas tales como la concesión de permisos o para la aplicación de medidas penales alternativas a la privación de libertad en unidades de régimen cerrado.

En conclusión, si bien existen diversos estudios psicométricos del BIDR-6, pocas investigaciones se han desarrollado sobre este aspecto en contextos de habla hispana, no hallándose en Argentina adaptaciones y estudios psicométricos del mismo. Por tal razón, este estudio representa un primer avance significativo que requiere de nuevas investigaciones a fin de aportar mayor solidez al modelo de DS puesto a prueba.

Consideramos que el aporte de instrumentos que evalúen sesgos de respuesta como la DS es sumamente relevante, dada su importancia en las evaluaciones psicológicas que se realizan en diversos contextos de aplicación (especialmente clínico, laboral y jurídico) frente a múltiples demandas (selección, predicción, clasificación, diagnóstico). El BIDR-6 abreviado puede constituirse en una herramienta valiosa para ser utilizada tanto en el campo de la evaluación psicológica en los diversos ámbitos reseñados, como en actividades de investigación e intervención.

Bibliografía

- Barger, S. D. (2002). The Marlowe-Crowne affair: Short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, 79, 286–305. <http://dx.doi.org/10.1207/S15327752JPA7902.11>
- Bobbio, A. y Manganelli, A. M. (2011). Measuring social desirability responding. A short version of Paulhus' BIDR 6.

- Testing, *Psychometrics Methodology in Applied Psychology*, 18(2), 117–135. <http://dx.doi.org/10.4473/TPM.18.2.4>
- Byrne, B. M. (2009). *Structural equation modeling with Amos: Basic concepts, applications, and programming*. New York: Taylor & Francis/Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Cronbach, L. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475–494. <http://dx.doi.org/10.1177/001316444600600405>
- Crowne, D. P. y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349–354. <http://dx.doi.org/10.1037/h0047358>
- Damarin, F. y Messick, S. (1965). *Response styles as personality variables: A theoretical integration*. Princeton, NJ: Educational Testing Service.
- D'Amours-Raymond, J. (2011). *Version abregée transculturelle du Balanced Inventory of Desirable Responding (BIDR) [tesis de maestría]*. [consultado 18 Abr 2016]. Disponible en: www.theses.ulaval.ca/2011/28117/28117.pdf.
- Dodaj, A. (2012). Social desirability and self-reports: Testing a content and response-style model of socially desirable responding. *Europe's Journal of Psychology*, 8, 651–666. <http://dx.doi.org/10.5964/ejop.v8i4.462>
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Fernández, J. y Echeburúa, E. (2006). Uso y abuso de los autoinformes en la evaluación de los trastornos de personalidad. *Revista de Psicopatología y Psicología Clínica*, 11(1), 1–12. <http://dx.doi.org/10.5944/rppc.vol.11.num.1.2006.4014>
- Forero, C. G., Maydeu Olivares, A. y Gallardo Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 16, 625–641. <http://dx.doi.org/10.1080/10705510903203573>
- George, D. y Mallery, M. (2011). *SPSS for Windows step by step: A simple guide and reference 18.0 Update* (11st ed.). Boston, MA: Allyn & Bacon.
- González Álvarez, N., Abad González, J. y Leví Mangin, J. P. (2006). Normalidad y otros supuestos en análisis de covarianzas. En J. P. Lévy Mangin y J. Varela (Eds.), *Modelización con estructuras de covarianzas en ciencias sociales* (pp. 31–59). Coruña: Netbiblo.
- Gough, H. G. (1952). On making a good impression. *Journal of Educational Research*, 46, 33–42. <http://dx.doi.org/10.1080/00220671.1952.10881997>
- Gravdal, L. y Sandal, G. M. (2006). The two-factor model of social desirability: Relation to coping and defense, and implications for health. *Personality and Individual Differences*, 40, 1051–1061. <http://dx.doi.org/10.1016/j.paid.2005.11.004>
- Hart, C. M., Ritchie, T. D., Hepper, E. G. y Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open*, 1–9. <http://dx.doi.org/10.1177/2158244015621113>
- Hathaway, S. R. y McKinley, J. C. (1951). *MMPI manual*. New York: Psychological Corporation.
- Hu, L. T. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Kline, R. B. (2010). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.
- Kroner, D. y Weekes, J. (1996). Balanced Inventory of desirable responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences*, 21, 323–333. [http://dx.doi.org/10.1016/0191-8869\(96\)00079-7](http://dx.doi.org/10.1016/0191-8869(96)00079-7)
- Kroner, D. G. y Weekes, J. R. (1996). Balanced inventory of desirable responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences*, 21, 323–333. [http://dx.doi.org/10.1016/0191-8869\(96\)00079-7](http://dx.doi.org/10.1016/0191-8869(96)00079-7)
- Leite, W. L. y Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne social desirability scale and the balanced inventory of desirable responding. *Educational and Psychological Measurement*, 65, 140–154. <http://dx.doi.org/10.1177/0013164404267285>
- Marlowe, D. y Crowne, D. P. (1961). Social desirability and response to perceived situational demands. *Journal of Consulting Psychology*, 25, 109–115. <http://dx.doi.org/10.1037/h0041627>
- Moral de la Rubia, J., García-Cadena, C. H. y Antona-Casas, C. J. (2012). Traducción y validación del Inventario Balanceado de Deseabilidad Social al responder en una muestra probabilística de estudiantes universitarios mexicanos. *Revista de Psicología GEPU*, 3(2), 54–72 [consultado 19 Abr 2016]. Disponible en: <http://revistadepsicologiagepu.es/t/>
- Moral, J. y Ortiz, H. (2011). Un modelo predictivo de conducta disocial por análisis de senderos. *Anuario de Psicología Jurídica*, 21, 44–62. <http://dx.doi.org/10.5093/jr2011v21a3>
- Muñiz, J. y Hambleton, R. K. (1996). Directrices para la traducción y adaptación de tests. *Papeles del Psicólogo*, 66, 63–70 [consultado 10 Mar 2016]. Disponible en: www.cop.es/pdf/dtyatest.pdf.
- Nunnally, J. C. y Bernstein, I. J. (1995). *Teoría psicométrica* (3.ª ed). México, D.F.: Editorial McGraw Hill Latinoamericana.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598–609. <http://dx.doi.org/10.1037/0022-3514.46.3.598>
- Paulhus, D. L. (1988). *Manual for the Balanced Inventory of Desirable Responding (BIDR-6)*. Unpublished manual, University of British Columbia.
- Paulhus, D. L. (2002). Socially desirable responding: the evolution of a construct. En H. I. Braun, D. N. Jackson, y D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 46–69). Mahwah NJ: Lawrence Erlbaum.
- Paulhus, D. L. y Reid, D. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307–317. <http://dx.doi.org/10.1037/0022-3514.60.2.307>
- Pérez, M., Labiano, M. y Brusasca, C. (2010). Escala de deseabilidad social: análisis psicométrico en muestra argentina. *Evaluar*, 10, 53–67 [consultado 18 Mar 2016]. Disponible en: <http://revistas.unc.edu.ar/index.php/revaluar/article/view/458>
- Rickett, M. (2005). Organizational identification: A meta-analysis. *Journal of Vocational Behavior*, 66, 358–384 [consultado 10 Mar 2016]. Disponible en: <http://dx.doi.org/10.1016/j.jvb.2004.05.005>
- Rogers, R., Vitacco, M. J., Jackson, R. L., Martin, M., Collins, M. y Sewell, K. W. (2002). Faking psychopathy? An examination of response styles with antisocial youth. *Journal of Personality Assessment*, 78, 31–46. http://dx.doi.org/10.1207/S15327752JPA7801_03
- Sackeim, H. A. y Gur, R. C. (1978). Self-deception, other-deception and consciousness. En G. E. Schwartz y D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2) (pp. 139–197). New York: Plenum Press.
- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles de Psicólogos*, 26, 115–128 [consultado 12 Mar 2016]. Disponible en: <http://www.papelesdelpsicologo.es/vernumero.asp?id=1252>
- Stöber, J., Dette, D. y Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78, 370–389. http://dx.doi.org/10.1207/S15327752JPA7802_10

- Van de Vijver, F. J. R. y Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89–99. <http://dx.doi.org/10.1027/1016-9040.1.2.89>
- Wiggins, J. S. (1964). Convergences among stylistic response measures from objective personality tests. *Educational and Psychological Measurement*, 24, 551–562. <http://dx.doi.org/10.1177/001316446402400310>
- Ximénez, C. (2007). Effect of variable and subject sampling on recovery of weak factors in CFA. *Methodology*, 3(2), 67–80. <http://dx.doi.org/10.1027/1614-2241.3.2.67>