

Interdisciplinaria

ISSN: 0325-8203

interdisciplinaria@fibercorp.com.ar

Centro Interamericano de Investigaciones
Psicológicas y Ciencias Afines
Argentina

MERINO SOTO, CÉSAR; VILCHERREZ PIZARRO, LIZETH KATHERINE
EVALUACIÓN PSICOMÉTRICA DE UN INSTRUMENTO DE DESEABILIDAD SOCIAL PARA NIÑOS
Interdisciplinaria, vol. 31, núm. 1, 2014, pp. 107-120
Centro Interamericano de Investigaciones Psicológicas y Ciencias Afines
Buenos Aires, Argentina

Disponible en: <http://www.redalyc.org/articulo.oa?id=18031545007>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

EVALUACIÓN PSICOMÉTRICA DE UN INSTRUMENTO DE DESEABILIDAD SOCIAL PARA NIÑOS

PSYCHOMETRIC ASSESSMENT OF A MEASURE FOR SOCIAL DESIRABILITY FOR CHILDREN

CÉSAR MERINO SOTO* Y LIZETH KATHERINE VILCHERREZ PIZARRO**

*Magister en Psicología. Docente investigador del Instituto de Investigación de Psicología de la Universidad de San Martín de Porres. Lima - Perú. E-Mail: sikayax@yahoo.com.ar

**Licenciada en Psicología. Práctica Privada. E-Mail: karin_14_2005@hotmail.com

RESUMEN

El objetivo del estudio realizado fue evaluar las propiedades psicométricas de una medida de *deseabilidad social* (8 ítemes) para *niños*, elaborado en habla hispana (Lemos, 2005), y que se distingue por su brevedad, respaldo de la literatura científica, novedad y originalmente construido en habla hispana. El instrumento es particularmente relevante para la investigación que usa el auto-reporte como medida de recolección de datos en niños, la cual es muy frecuente en niños y adultos. Sin embargo, hay apenas algunos estudios en niños con este constructo en habla hispana, especialmente con el desarrollo de *test* para su medición. La validación original muestra propiedades psicométricas que posiblemente no son replicables por cuestiones metodológicas. Debido a la importancia de describir y cuantificar este constructo en población infantil, se evaluó su estructura interna (análisis factorial confirmatorio de grupo múltiple), *validez* de constructo (diferencias de edad y sexo) y confiabilidad (consistencia interna, *KR-20* con modificación Horst) en una muestra peruana de 202 niños de nivel primario de una institución educativa pública, de educación regular. Los análisis mostraron que solo una dimensión retiene mejor la variancia de los ítemes (contrastando con la bidimensionalidad original), algunos ítemes tienen bajas correlaciones inter-ítem y cargas factoriales, la consistencia

interna es baja, y que las correlaciones con la edad no siguieron el patrón esperado. Se discuten los resultados estructurales en el marco del sesgo producido por el fraseo de los ítemes y las diferencias metodológicas del estudio que se informa, aspectos que pueden influir en las propiedades psicométricas del instrumento.

Palabras clave: Deseabilidad social; Validez; Niños; Evaluación; Agradabilidad.

ABSTRACT

The importance of the effect of *social desirability* on self-reported measures in *children* has been analyzed for several decades, with the development of instruments for children (eg, Crandall, V.C., Crandall, V.J. & Katkovsky, 1965). This type of research is still ongoing and developing basic research (Haghghat, 2007), especially in the creation of short measures (Carifio, 1992; Haghghat, 2007) with good discrimination and masked within test battery for use in group research in children. However, research with this construct is more abundant in adults, and less common in children, especially in Latin America. May be the limitations in reading comprehension, development aspects in the recognition introspective cognitions and emotions, among other things, do not allow a deeper analysis and safety compared with adults. Also, in Spanish

language, it is rare to find independent instruments that describe social desirability in children, but rather integrated into a multidimensional test (i.e., Eysenck, S.B.G. & Eysenck, H.J., 1973). One of the most important measures for social desirability speaking is developed in Hispanic context is the Children's Social Desirability Scale (EDESI - Lemos, 2005). This instrument has a particular relevance to psychological research, which calls attention to the impact of social desirability on reports that children make their own psychological attributes using questionnaires. Since EDESI can make an important psychometric contribution to quantify the social desirability in Hispanic children, ongoing evaluation of its psychometric properties have not only local value (i.e., relevant to the geographical region in which it was built) but also in the international context Hispanic. This study assessed the replicability of results Lemos (2005), with this potential instrument of social desirability. Although some features of the study design Lemos not repeated here (e.g., distribution in different socioeconomic levels), the evidence obtained provides additional evidence in a rather partial replication framework.

The sample consisted of 202 students from grade 4 to 6 (55% male and 45% female) primary level from a public educational institution of Lima, Perú, between 8 and 12 years old. For analysis, exploratory factor analysis in ULS method, parallel analysis and Scree Test for number of factors, polychoric correlations and Promax rotation ($k=4$) was used. The estimate of reliability was made with the KR-20 coefficient and confidence intervals, and Horst's (1953) correction was used to reduce the impact of the dispersion of item difficulty (Merino & Charter, 2010).

A one-dimensional structure as the most suitable was found to represent the construct. Internal consistency was slightly lower compared to that found in the original study, but still be considered equivalent. The results presented in this study suggest that previous findings (Lemos, 2003, 2005) are limited in their replicability, and probably the methodological choices for analyzing the internal relations of its items are the explanations for this. The instrument may require items to be scaled in three or four response choices and a slightly higher number of items if one want to improve efficiency in mass applications.

Additionally, a check to verify the factor structure of the scale settings as a first step (Merino & Arndt, 2009) compared to more advanced methods (e.g., structural equations), should be conducted. According to results in this study, should be considered not just a two-dimensional model, but include a one-dimensional model that seems to compete more successfully with the original two-factor model. An important aspect is the phrasing of the items as grouping factor of the items. This problem is recognized in self-report instruments that have phrased items on direct and opposite direction to construct. The prospect of research on this topic, and with this instrument should include structural hypotheses as bifactor model and method factor as plausible explanations of relations between items. Also, assess the measurement model for get the best estimation of reliability. It can also anticipate constructing items with Hispanic multicultural relevance.

Key words: Social desirability; Validity; Children; Assessment; Agreeableness.

INTRODUCCIÓN

El objetivo de la investigación que se informa fue evaluar la replicabilidad del estudio de construcción de la Escala de Deseabilidad Social Infantil (EDESI - Lemos, 2005). Este instrumento posee una particular relevancia para la investigación psicológica porque invoca la atención hacia el impacto de la deseabilidad social en los reportes que los niños hacen de sus propios atributos psicológicos, usando cuestionarios de evaluación. Esta área es aparentemente poco aplicada en los trabajos empíricos latinoamericanos que utilizan el método de auto-reporte como vía para obtener información.

En una de las más recientes revisiones efectuada por Lemos (2003, 2005, 2006) sobre la deseabilidad social en la evaluación de la personalidad infantil, se indica que esta variable es de importancia crítica para evaluar la validez de las respuestas que dan los niños durante el reporte de sus propias experiencias

psicológicas, debido a que pretenden mostrar una imagen mejorada de ellos mismos para agradar a los demás. Sin embargo, esta interpretación puede depender de la situación que enfrenta el niño; por ejemplo, una moderada magnitud de este fenómeno en niños podría sugerir problemas para comprender introspectivamente las propias reacciones en situaciones de la vida diaria (Eysenck, S.B.G., Nias & Eysenck, H.J., 1971).

Esta tendencia parece también derivarse del constructo de agradabilidad (basado en el enfoque de los Cinco Grandes Factores de Personalidad - Lemos, 2003), en el que es característico preocuparse por el mantenimiento de las relaciones sociales y que influye en la presentación socialmente deseable (Lemos, 2003). El hallazgo de la autora del instrumento de personalidad (y en el que la agradabilidad es un componente) fue interesante, porque reveló las diferencias de este constructo entre niños con riesgo por pobreza extrema y niños sin estas características; específicamente, los niños de bajos recursos económicos presentaban puntajes más altos. Pero estas investigaciones iniciales aún no aportan claras conclusiones sobre la validez de la EDESI en relación a criterios relevantes de discriminación y convergencia con otras medidas, o su capacidad discriminante para identificar niños con este atributo en situaciones en que la deseabilidad razonablemente incrementa. Esta diferenciación de constructo es necesaria si se toma en cuenta que puede revelarse el componente original de un constructo (Lemos, 2003; Paulhus, 1991) y no únicamente un factor de sesgo de las respuestas (Crandall, V.C., Crandall, V.J. & Katkovsky, 1965), como frecuentemente se lo considera. Estos aspectos, por lo tanto, deben ser aclarados desde fuentes internas (por ejemplo, consistencia y estructura interna) y externas de validez (por ejemplo, correlaciones convergentes) si se desea tener un juicio fundamentado sobre el uso aplicado o de investigación de la EDESI.

La importancia del efecto de la deseabilidad social sobre las medidas de auto-reporte en niños es analizada desde hace varias décadas, con la creación de instrumentos clási-

cos para niños (por ejemplo, Crandall et al., 1965), y aún continúa siendo desarrollada en la investigación básica (Haghigat, 2007), especialmente para crear versiones breves (Carifio, 1992; Haghigat, 2007) que sean suficientemente discriminativas y enmascaradas dentro de una batería de medidas aplicables a grandes grupos de niños. Sin embargo en habla hispana, es poco frecuente hallar instrumentos independientes que describan la deseabilidad social en niños, sino más bien integrados a un test multidimensional (por ejemplo Eysenck, S.B.G. & Eysenck, H.J., 1973), e interpretados como la preferencia por la propia presentación favorable. En cualquier situación, las medidas de deseabilidad social generalmente no se incluyen como covariables de control de la validez de las respuestas, un aspecto que puede poner en cuestionamiento las correlaciones entre las variables que legítimamente pueden ser afectadas por esta conducta.

En el estudio presentado por Lemos (2005) sobre la EDESI se muestra evidencia de validez de la estructura interna y de constructo; sin embargo, se pueden reconocer algunas limitaciones en sus métodos que podrían ocasionar que los parámetros calculados para los ítems sean poco replicables en otras muestras. Por ejemplo, el método de extracción de factores utilizado fue por componentes principales y correlaciones *Pearson* entre los ítems, mas no se aclara la forma de determinación del número de factores; y en este punto, se puede asumir que se aplicó una combinación de juicio y procedimiento técnicos. Por otro lado, la consistencia interna no tuvo en cuenta la dispersión razonable de las medias de los ítems y solo se reportó para el puntaje total, mas no para los dos factores encontrados. En resumen, las razones que limitarían seriamente sus resultados se relacionan con: (a) no haber aplicado una adecuada recuperación de factores con mejor vínculo y cuerpo teórico, que debería basarse en los métodos de factores comunes (Gorsuch, 1990; Widaman, 2003), (b) el tratamiento de las variables dicotómicas, ya que su naturaleza categórica y la influencia de su distribución de frecuencia sobre las soluciones fac-

toriales se controla mejor con correlaciones teóricas (Holgado, Chacón, Barbero & Vila, 2010; Lorenzo-Seva & Ferrando, 2006; Morata-Ramírez & Holgado-Tello, 2013; Muthén, 1993) y (c) la variabilidad de las medias de los ítems, que produce un sesgo negativo sobre el coeficiente *KR20* (Charter, 1995; Horst, 1953; Merino & Charter, 2010). En cuanto a este último aspecto (grado de dispersión de las medias de los ítems), no puede evaluarse en los resultados presentados por Lemos (2005) debido a que los estadísticos descriptivos de los ítems no fueron reportados; por lo tanto, sin esta información no se puede valorar si la estimación de la consistencia tendría un sesgo negativo ni si el conjunto de ítems se aproxima a un modelo equivalente *tau*. Este último aspecto es crítico para el correcto uso del coeficiente *Alpha* (Cronbach, 1951) y por extensión, para el uso del *KR-20*. Es posible pensar que los resultados de Lemos puedan representar relaciones internas de sus ítems que son más bien característicos de la muestra de estudio, situación que debe ser evaluada en otro contexto cultural para justificar su uso.

Ya que la EDESI puede hacer un importante aporte psicométrico para cuantificar la deseabilidad social en niños, la continua evaluación de sus propiedades psicométricas no tiene solo valor local (es decir, relevante para la región geográfica en que se construyó), sino también en el contexto internacional hispano. El estudio que se informa evaluó la replicabilidad de lo hallado por Lemos (2005) con este interesante y potencial instrumento de descripción de la deseabilidad social. Aunque algunas características del diseño del estudio de Lemos no se repiten aquí (por ejemplo, la distribución en niveles socioeconómicos distintos), la evidencia que se obtendrá provee de evidencia adicional en un marco más bien parcial de replicabilidad.

MÉTODO

PARTICIPANTES

Contando con la autorización de la institución educativa y de los padres de familia,

fueron evaluados todos los niños que se encontraban presentes el día de la administración de las pruebas. Si bien el colegio está ubicado en una zona principalmente urbanizada, las familias provienen de zonas semi-urbanas y el nivel socioeconómico estimado y asociado a las familias de los niños es medio bajo o menor.

La muestra estuvo compuesta por 202 estudiantes de cuarto a sexto grado (55% varones y 45% mujeres), de nivel primario y de una institución educativa pública de Lima Metropolitana, de 8 a 12 años ($M = 10.4$, $DE = 1.99$).

La selección de la muestra fue no probabilística. Durante el tiempo de evaluación no se recibió información sobre la presencia de alumnos con discapacidad física, cognitiva o neurológica observables que pudieran comprometer los resultados.

INSTRUMENTO

ESCALA DE DESEABILIDAD SOCIAL INFANTIL (EDESI - LEMOS, 2005)

Este instrumento permite evaluar la deseabilidad social como una característica de la personalidad infantil en niños de 9 a 12 años. La versión final de 8 ítems resultó del análisis de la estructura interna de los 22 ítems iniciales. Lemos (2005) obtuvo evidencias de validez de la estructura interna y de constructo (en relación con la edad) y evidencias de confiabilidad para el puntaje total ($\alpha = .70$). La validez de la estructura interna halló dos dimensiones (de cuatro ítems cada una) relacionadas con aspectos socialmente aceptados y aspectos opuestos a esta; no se reportaron estimaciones de consistencia interna para estos dos componentes.

PROCEDIMIENTO

La EDESI se aplicó grupalmente en el aula y en horario regular de clase.

Se analizaron descriptivamente los ítems (respuesta promedio o tendencia de respon-

ta y variabilidad), incluyendo la correlación ítem-test corregida (Henrysson, 1963) como medida de la capacidad discriminativa de los ítems. Asumiendo el modelo equivalente *tau* subyacente a los ítems de la EDESI, se estimó la consistencia interna como evidencia de confiabilidad, aplicando el coeficiente *KR-20* (Kuder & Richardson, 1937) debido a los ítems dicotómicos de la EDESI. También se usó una corrección para mejorar su estimación (*KR-20-Horst*, Charter, 1995; Horst, 1953), suponiendo la presencia de una moderada dispersión de la dificultad de los ítems. Además, se calcularon intervalos de confianza para el coeficiente *KR-20-Horst* por el método de Fisher (1950; Romano, Kromrey & Hibbard, 2010). La validez de las relaciones internas de la EDESI se estudió mediante el análisis factorial exploratorio sobre las correlaciones tetracóricas entre los ítems (algoritmo de Bonnet & Price, 2005), con el programa FACTOR (Lorenzo-Seva & Ferrando, 2006). Este procedimiento asume que las respuestas latentes de los ítems son de naturaleza continua aun cuando el indicador externo se presenta de manera categórica (Holgado et al., 2010; Morata-Ramírez & Holgado-Tello, 2013; Richaud, 2005). El método de extracción factorial usado fue ULS (*Unweighted Least Squares* - McDonald, 1982). La definición de la estructura de la EDESI se inició de acuerdo al planteo original de Lemos (2005), donde se hipotetizó una estructura bifactorial (ítems de aspectos positivos y otros de aspectos negativos); por lo tanto, en la primera fase del análisis factorial se permitió una solución de dos factores. Esto se contrastó con dos métodos exploratorios para determinar el número de dimensiones latente: el análisis paralelo (Horn, 1965) y el Gráfico *Scree Test* (Cattell, 1966). El método *scree* desde hace tiempo ha mostrado un aceptable desempeño (Cattell & Vogelman, 1977; Zwick & Velicer, 1982); por otro lado, el método del análisis paralelo es computacionalmente intensivo, y se usó el Programa FACTOR para generar matrices de correlaciones tetracóricas entre variables aleatorias (Timmerman & Lorenzo-Seva, 2011). La interpretación de la solución bifactorial se facilitó aplicando una ro-

tación Promax (Hendrickson & White, 1964), con el parámetro de oblicuidad *k* puesto en 4 (Tataryn, Wood & Gorsuch, 1999).

Otra evidencia de validez de constructo evaluada en la investigación realizada fue la variabilidad hipotetizada de los puntajes producido por la edad cronológica. Algunos hallazgos han relevado el aumento de la deseabilidad en edades preescolares (Ford, 1970), que difiere de las relaciones reportadas entre ambos atributos en otras investigaciones. Por ejemplo, se ha reportado la disminución de la tendencia de la deseabilidad social con el avance de la edad (Crandall et al., 1965; Lemos, 2005; Porrata, 1995). Esta tendencia parece ser más consistente interculturalmente, incluso en grupos africanos con aculturación occidental, los adultos tendieron a presentar menos puntajes que los grupos adolescentes con los cuales se compararon, sugiriendo la disminución de los puntajes influenciada por la edad (Mwanwenda, 1993, 1995). Por lo tanto, se planteó si la EDESI es sensible a los cambios de la edad, y del modo en que la literatura y Lemos lo han reportado; entonces los puntajes de la EDESI deben disminuir en tendencia lineal decreciente. Esta hipótesis se probó con un contraste a priori (lineal) sobre los puntajes promedio de la EDESI en la aplicación del ANOVA para muestras independientes.

RESULTADOS

ANÁLISIS DE ÍTEMES

Excepto dos ítems (Item 1: “Siempre escucho con atención cuando alguien me habla” e Item 7: “Siempre hago todo lo que mis padres me piden”), la tendencia de respuesta estuvo entre .20 y .90, un rango recomendado para una mejor discriminación del constructo de interés (Anastasi & Urbina, 1998). Las correlaciones tetracóricas inter-ítem mostraron un patrón que parece reconocerse de la siguiente manera: los ítems fraseados hacia la deseabilidad social tendieron a mostrar elevada magnitud comparadas con los ítems fraseados en dirección opuesta (ítems 2, 4, 6

y 8). Por otro lado, las correlaciones con el Item 2 (“Me molesta un poco cuando no están de acuerdo conmigo”) y el Item 8 (“Alguna vez me hice como el que entendía algo, cuando en realidad no entendía nada”) fueron constantemente más bajas que el resto, como así también su correlación ítem-test (ver Tabla 1).

ANÁLISIS FACTORIAL

La dimensionalidad resultante del análisis paralelo y el *Scree Test* resultó que sólo una dimensión latente podría ser suficiente para explicar la variancia total de los ítems. Sin embargo, se probó la extracción de dos factores para indagar la distribución exploratoria de los ítems; estos dos factores corresponderían a los que deberían emerger tomando en cuenta los resultados originales (Lemos, 2005), en que se distinguen dos factores originados por la orientación de los ítems (positivos y negativos). El resultado no fue satisfactorio con los ítems de los aspectos negativos pues las cargas fueron algo bajas y a excepción del Item 4, los demás ítems de este factor obtuvieron principalmente cargas factoriales débiles (entre .25 y .50 - Ximénez & García, 2005). También se halló que el Item 5 (“Siempre dejo mis cosas ordenadas”) se posicionaba en el factor diferente y la existencia de una ligera complejidad factorial (ítems 7 y 4) (ver Tabla 1). Dados estos resultados estadísticos, y que el número de ítems dicotómicos de cada factor no es grande, el mejor resultado empírico, parece corresponder a una sola dimensión latente. En ambas soluciones factoriales se ensayó la eliminación consecutiva de ítems que podría mejorar la validez interna, pero estas mejoras fueron sustanciales y en algunos casos la solución factorial empeoró.

CONSISTENCIA INTERNA

El coeficiente *KR-20* para el puntaje original fue igual a .59, y la correlación interítem (r_{ii}) promedio, igual a .17. La remo-

ción del Item 2 hizo disminuir muy ligeramente el coeficiente *KR-20* (.58) y elevar r_{ii} (.18), evidenciando el efecto del número de ítems sobre el coeficiente. La magnitud del coeficiente puede considerarse baja y cuestionaría el uso del puntaje aún en las aplicaciones para investigación. Sin embargo, la modificación *Horst* del coeficiente *KR-20* (Horst, 1953) fue .66 (95% I.C. = .57, .73), éste último debe considerarse una mejor estimación de la consistencia interna pues desatenía el impacto de las diferencias en las medias de los ítems (Horst, 1953; Merino & Charter, 2010). La comparación de este resultado con el obtenido por Lemos (2005 - $\alpha = .70$; 95% I.C. = .60, .77) sugiere que las diferencias no son estadísticamente significativas y las diferencias observadas entre ellas deben considerarse error de muestreo. Finalmente, el intervalo de confianza reportado indica que el valor poblacional del coeficiente de confiabilidad puede alcanzar un nivel bajo (límite inferior debajo de .60).

DIFERENCIAS ENTRE GRUPOS DE EDAD

Se detectó una muy leve tendencia decreciente de los puntajes medios en el puntaje de la EDESI conforme aumenta la edad, pero esencialmente no fue estadísticamente significativa, $F(1, 199) = .93, p = .33$. Esto indica que las diferencias de edad fueron pequeñas y podrían considerarse como consecuencia del error de muestreo.

DISCUSIÓN

La presente investigación evaluó la replicabilidad de los resultados del estudio de Lemos (2005) con respecto a un instrumento de deseabilidad social para niños. Específicamente, la validación de la generalización psicométrica de la EDESI (Lemos, 2005) se evaluó en una muestra peruana, de mayor tamaño, en un rango más amplio de edad y en un nivel socioeconómico diferente comparado con el estudio original. Los resultados del análisis de esta escala indican que en la

muestra peruana de nivel socioeconómico más bajo es suficiente un solo factor para interpretar los resultados, situación que se contrasta con la solución bifactorial de Lemos (2005). Además, las relaciones internas de los ítems fueron parcialmente satisfactorias, pues los ítems con contenido orientado hacia los aspectos positivos fueron los más intercorrelacionados entre sí, comparados con los de aspectos negativos. Las modificaciones exploratorias realizadas para buscar una mejor solución factorial con base empírica no fueron satisfactorias. Por lo tanto, la falta de relación de alguno de los ítems con el resto de los ítems y con el puntaje parece indicar que su contenido debe ser reevaluado y que se generen nuevos ítems. Introducir nuevos ítems tendrá el conocido beneficio de aumentar la confiabilidad y, asumiendo el contenido relacionado, las correlaciones inter-ítem.

La consistencia interna fue ligeramente baja comparada con la hallada en el estudio original, pero aún pueden considerarse equivalentes. Usando la suma total de todos los ítems como variable de respuesta, la diferencia lineal entre los grupos de edad no fue replicada con respecto a los previos resultados de Lemos. Contrariamente a la tendencia de los estudios interculturales, en los que la edad muestra una asociación lineal negativa con la deseabilidad social en niños hispanos (Porras, 1995; Lemos, 2005) y no hispanos (Mwanwenda, 1993, 1995), la similitud de puntajes en la muestra peruana puede ser un hecho interesante que requiera una exploración adicional. Por extensión, otros aspectos que pueden interaccionar con las diferencias entre grupos es el nivel socioeconómico, aspecto que ha sido identificado como una plausible influencia en las diferencias individuales en la deseabilidad social en niños (Lemos, 2005; Porras, 1993).

Los resultados que se presentan parecen indicar que los hallazgos previos (Lemos, 2003, 2005) son de limitada replicabilidad, y probablemente las elecciones metodológicas para analizar las relaciones internas de sus ítems son las explicaciones para ello. El instrumento puede requerir que los ítems sean escalados en tres o cuatro opciones de

respuesta y un número ligeramente mayor de ítems si se quiere conservar su eficiencia en aplicaciones masivas. Adicionalmente debería conducirse una comprobación de la estructura factorial para verificar la configuración de la escala como paso inicial (Merino & Arndt, 2009) frente a métodos más avanzados (por ejemplo, ecuaciones estructurales). De acuerdo a lo hallado en el estudio realizado, se debería considerar no solo un modelo de dos dimensiones, sino incluir un modelo de una dimensión que parece competir más satisfactoriamente con el modelo original bifactorial. El modelo de dos factores hallado en el estudio original podría explicarse por un sesgo metodológico debido a la presencia de ítems fraseados en dirección opuesta, problema que ha sido examinado en otros estudios que han empleado escalas de deseabilidad social (Pérez, Labiano & Brusasca, 2011) y también con otros instrumentos con ítems fraseados negativa y positivamente en una misma escala (por ejemplo, Casas & Merino, 2013; Chen, Rendina-Gobioff & Dedrick, 2007; Hankins, 2008; Marsh, Scalas & Nagengast, 2010; Motl & Conroy, 2000; Tomas, Meléndez, Oliver, Navarro & Zaragoza, 2010; Ye, 2009). Es posible que una mejor propuesta de medición de la deseabilidad social en niños deba contener ítems orientados en una misma dirección e incluir ítems de validez para controlar el sesgo de respuesta que se quiere atenuar al presentar ítem en ambas direcciones. Esto puede efectivamente eliminar el efecto del fraseo de los ítems, que es un problema hallado también en niños (Chen et al., 2007) y que puede caracterizar también la complejidad factorial de los ítems (Merino & Grimaldo, 2010). Aunque la evidencia de validez sobre la EDESI aún es escasa en la literatura, algunos experimentos realizados en instrumentos con el problema del fraseo de ítems en ambas direcciones (por ejemplo, la escala de autoestima de Rosenberg, 1965) han demostrado que los cambios hacia el fraseo único (todos los ítems positivos o todos negativos) no afectan seriamente la validez de constructo (Greenberger, Chen, Dmitrieva & Farruggia, 2003).

La elaboración de la EDESI pone en relieve un aspecto poco estudiado entre las investigaciones metodológicas hispanas y debido a ello, debe ser un instrumento psicométrico a ser considerado para su posible uso internacional, pero luego de hacer mejoras en su estructura y posiblemente en su contenido. En este sentido, es necesaria una metodología más apropiada para establecer sus propiedades estructurales, ya que favorecería su replicabilidad psicométrica no solo localmente, sino también internacionalmente. Estos aspectos metodológicos conciernen al método de reducción de variables (análisis de componentes principales vs factores comunes), a la estimación de la confiabilidad (ajuste por dispersión de las medias de los ítems vs no ajuste) y a la elección de la matriz correlacional para factorizar (correlaciones de *Pearson* vs correlaciones tetracóricas) y el apropiado número de variables latentes (*Eigenvalue* mayor a 1 vs análisis paralelo).

Un aspecto que debe discutirse es que el contenido de los ítems actuales de la EDESI contienen alta la variancia específica y que, en consecuencia, hace disminuir la communalidad entre ellos; más aún, puede existir una interacción de sus contenidos con factores culturales en la muestra peruana que no corresponden a la cultura argentina. Particularmente, el contenido sensible a la deseabilidad social puede referirse al reconocimiento de errores, expresión de emociones intensas y negativas, pensamientos socialmente no aceptables, hábitos de trabajo aceptables, etc. Posiblemente esto conduzca a la reconstrucción de una versión modificada de la EDESI que tenga en cuenta estos aspectos para lograr un mejor nivel de generalización de su contenido y de sus propiedades psicométricas. Pero esta es una pregunta abierta que merece ser respondida en el importante marco de evaluar la deseabilidad social en niños.

TABLA 1
ESTADÍSTICOS DESCRIPTIVOS Y CORRELACIONES TETRACÓRICAS DE LOS ÍTEMES DE EDESI

Ítemes	<i>M</i>	<i>DE</i>	R_{itc}	DS1	DS2	DS3	DS4	DS5	DS6	DS7	DS8
DS1	.93	.24	.388	1.0							
DS3	.78	.41	.215	.74	1.0						
DS5	.79	.40	.394	.39	-.02	1.0					
DS7	.91	.27	.337	.73	.45	.52	1.0				
DS2	.49	.50	.224	.10	.06	.37	.21	1.0			
DS4	.50	.50	.411	.53	.34	.58	.29	.25	1.0		
DS6	.70	.45	.295	.24	.03	.38	.13	.26	.35	1.0	
DS8	.56	.49	.219	.19	.15	.26	.24	.05	.16	.34	1.0

Nota:

R_{itc} : correlación ítem-test corregida

TABLA 2
ESTRUCTURA FACTORIAL DE LOS ÍTEMES DEL EDESI

Ítems	Modelo 1 factor		Modelo 2 factores		h^2
	F	h^2	F1	F2	
DS1	.86	.75	.10 (.47)	.95 (.99)	1.0
DS3	.52	.27	-.24 (.10)	.88 (.78)	.66
DS5	.61	.37	.97 (.94)	-.07 (.30)	.89
DS7	.72	.52	.28 (.49)	.54 (.65)	.49
DS2	.30	.09	.41 (.40)	-.03 (.12)	.16
DS4	.64	.41	.49 (.59)	.27 (.46)	.42
DS6	.38	.14	.50 (.49)	-.02 (.17)	.24
DS8	.31	.10	.27 (.31)	.10 (.20)	.10
<i>Eigenvalue</i> (autovalor)	3.21		3.21	1.47	
% de variancia	(40.2)		(40.3)	(18.4)	

Notación:

F: cargas factoriales

Cargas factoriales mayores a .40

h^2 : comunalidad

REFERENCIAS BIBLIOGRÁFICAS

Anastasi, A. & Urbina, S. (1998). *Test psicológicos* [Psychological tests], (7a. ed.). México, DF: Prentice-Hall.

Bonett, D.G. & Price, R.M. (2005). Inferential methods for the tetrachoric correlation coefficient. *Journal of Educational and Behavioral Statistics*, 30, 213-225. <http://dx.doi.org/10.3102/10769986030002213>

Carifio, J. (1992, marzo). *Parallel short forms of the Crandall Social Desirability Test for Children: Shortening instruments for research purposes*. Trabajo presentado en el Annual Meeting of the Eastern Educational Research Association. South Carolina, USA.

Casas, M. & Merino, C. (2013,). *Cuestionario de Salud de Goldberg (GHQ 12) en adolescentes: Una evaluación psicométrica en adolescentes* [Goldberg Health Questionnaire (GHQ-12) in adolescents: An psychometric examination in adolescents]. Trabajo presentado en el I Congreso Regional e Internacional de Psicología. Lima, Perú.

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. http://dx.doi.org/10.1207/s15327906mbr0102_10

Cattell, R.B. & Vogelmann, S.A. (1917). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*, 12, 289-325. http://dx.doi.org/10.1207/s15327906mbr1203_2

Charter, R.A. (1995). The under-representation of Horst's modification of the KR-20 reliability coefficient. *Perceptual and Motor Skills*, 81(3), 770. <http://dx.doi.org/10.2466/pms.1995.81.3.77>

Chen, Y., Rendina-Gobioff, G. & Dedrick, R.F. (2007, noviembre). *Detecting effects of positively and negatively worded items on a self-concept scale for third and sixth grade elementary students*. Trabajo presentado en el Annual Meeting of the Florida Educational Research Association, Tampa, FL.

Crandall, V.C., Crandall, V.J. & Katkovsky, W. (1965). A children's social desirability questionnaire. *Journal of Consulting Psychology*, 29(1), 27-36. <http://dx.doi.org/10.1037/h0020966>

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 22 (3), 297-334. <http://dx.doi.org/10.1007/BF02310555>

Eysenck, S.B.G. & Eysenck, H.J. (1973). Test-retest reliabilities of a new personality questionnaire for children. *British Journal of Educational Psychology*, 43(2), 126-130. <http://dx.doi.org/10.1111/j.2044-8279.1973.tb00748.x>

Eysenck, S.B.G., Nias, D.K.B. & Eysenck, H.J. (1971). The interpretation of children's lie scale scores. *British Journal of Educational Psychology*, 41(1), 23-31. <http://dx.doi.org/10.1111/j.2044-8279.1971.tb00654.x>

Fisher, R.A. (1950). *Statistical methods for research workers*. Edinburgh, UK: Oliver & Boyd.

Ford, L.H. (1970). A social desirability questionnaire for young children. *Journal of Consulting and Clinical Psychology*, 35(2), 195-204. <http://dx.doi.org/10.1037/h0030042>

Greenberger, E., Chen, C., Dmitrieva, J. & Farruggia, S.P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, 35, 1241-1254. [http://dx.doi.org/10.1016/S0191-8869\(02\)00331-8](http://dx.doi.org/10.1016/S0191-8869(02)00331-8)

Gorsuch, R.L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, 25(1), 33-39. http://dx.doi.org/10.1207/s15327906mbr2501_3

Haghigiat, R. (2007). The development of the brief Social Desirability Scale. *Europe's Journal of Psychology*, 3(4). Recuperado el 8 de mayo de 2012 de <http://www.ejop.org/archives/2007/11/the-development-of-the-brief-social-desirability-scale-bsds.html>

Hankins, M. (2008). The reliability of the twelve-item General Health Questionnaire (GHQ-12) under realistic assumptions. *Public Health*, 8, 355.

Hendrickson, A.E. & White, P.O. (1964). Promax: A quick method for rotation to orthogonal oblique structure. *British Journal of Statistical Psychology*, 17, 65-70. <http://dx.doi.org/10.1111/j.2044-8317.1964.tb00244.x>

Henrysson, S. (1963). Correction for item-total correlations in item analysis. *Psychometrika*, 28, 211-218. <http://dx.doi.org/10.1007/BF02289618>

Holgado, F.P., Chacón, S., Barbero, I., & Vila, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity*, 44, 153-166.

Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185. <http://dx.doi.org/10.1007/BF02289447>

Horst, P. (1953). Correcting the Kuder-Richardson reliability formula for dispersion of item difficulties. *Psychological Bulletin*, 50, 371-374. <http://dx.doi.org/10.1037/h0062012>

Klein, E.B., Gould, L.J. & Corey, M. (1969). Social desirability in children: An extension and replication. *Journal of Consulting and Clinical Psychology*, 33(1), 128. <http://dx.doi.org/10.1037/h0027339>

Kuder, G.F. & Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151-160. <http://dx.doi.org/10.1007/BF02288391>

Lemos, V. (2003). *La deseabilidad social en la evaluación de la personalidad infantil* [Social desirability in the evaluation of child personality]. Trabajo presentado en el IV Congreso Iberoamericano de Evaluación Psicológica. AID EP. Lima, Perú.

Lemos, V. (2005). Construcción y validación de una Escala para la Evaluación de la Deseabilidad Social Infantil (EDESI) [Construction and validation of Social Desirability Scale for Children]. *Interdisciplinaria*, 22(1), 77-96.

Lemos, V. (2006). La deseabilidad social en la evaluación de la personalidad infantil [Social desirability in personality assessment of children]. *Suma Psicológica*, 13(1), 7-14.

Lorenzo-Seva, U. & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments and Computers*, 38(1), 88-91. <http://dx.doi.org/10.3758/BF03192753>

Marsh, H.W., Scalas, L.F. & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, 22 (2), 366-381. <http://dx.doi.org/10.1037/a0019225>

McDonald, R.P. (1982). Linear versus nonlinear models in item response theory. *Applied Psychological Measurement*, 6, 379-396. <http://dx.doi.org/10.1177/014662168200600402>

Merino, C. & Arndt, S. (2009). Un enfoque confirmatorio en la construcción y revisión de escalas. Análisis factorial de grupo múltiple [A confirmatory approach in the construction and review of scales: Multiple group factor analysis]. *Revista de Psicología de la Universidad García de la Vega*, 1(1), 33-44.

Merino, C. & Charter, R.A. (2010). Modificación Horst al coeficiente KR - 20 por dispersión de la dificultad de los ítems [Host modification to KR - 20 for dispersion of item difficulty].

Revista Interamericana de Psicología, 44(2), 274-278.

Merino, C. & Grimaldo, M. (2010). Complejidad factorial de la permisividad moral hacia las conductas morales controvertidas [Factor complexity of Morally Debatable Behaviors Scale]. *Interdisciplinaria*, 27(2), 297-314.

Morata-Ramírez, M. & Holgado-Tello, F. (2013). Construct validity of Likert scales through confirmatory factor analysis: A simulation study comparing different methods of estimation based on Pearson and polychoric correlations. *International Journal of Social Science Studies*, 1, 54-61.

Motl, R.W. & Conroy, D.E. (2000). Validity and factorial invariance of the Social Physique Anxiety Scale. *Medicine & Science in Sports & Exercise*, 32(5), 1007-1017.

Muthén, B. (1993). Goodness of fit with categorical and other nonnormal variables. En K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 205-234). Newbury Park: Sage.

Mwamwenda, T.S. (1993). A comparison of two samples, South Africans and Canadians, on social desirability. *Psychological Reports*, 72, 965-966. <http://dx.doi.org/10.2466/pr0.1993.72.3.965>

Mwanwenda, T.S. (1995). Age differences in social desirability. *Psychological Reports*, 76, 825-826. <http://dx.doi.org/10.2466/pr0.1995.76.3.825>

Paulhus, D.L. (1991). Measurement and control response bias. En J.P. Robinson, P.R. Shaver & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Pérez, M., Labiano, M. & Brusasca, C. (2011). Escala de Deseabilidad Social: Análisis psicométrico en muestra argentina [Social Desirability Scale: Psychometric analysis in Argentinean sample]. *Evaluar*, 10, 53-67.

Porrata, J.L. (1993). Differences in social desirability scores of students in private and public schools. *Psychological Reports*, 72, 967-970. <http://dx.doi.org/10.2466/pr0.1993.72.3.967>

Porrata, J.L. (1995). Deseabilidad social en niños y adultos portorriqueños. *Revista Latinoamericana de Psicología*, 27(2), 305-312.

Richaud, M.C. (2005). Desarrollos del análisis factorial para el estudio de ítems dicotómicos y ordinales [Development of factor analysis for study of dichotomic and ordinal items]. *Interdisciplinaria*, 22(2), 237-251.

Romano, J.L., Kromrey, J.D. & Hibbard, S.T. (2010). A Monte Carlo study of eight confidence interval methods for coefficient alpha. *Educational and Psychological Measurement*, 70(3), 376-393. <http://dx.doi.org/10.1177/0013164409355690>

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton: Princeton University Press.

Tataryn, D.J., Wood, J.M. & Gorsuch, R.L. (1999). Setting the value of k in promax: A Monte Carlo study. *Educational and Psychological Measurement*, 59, 384-391. <http://dx.doi.org/10.1177/00131649921969938>

Timmerman, M.E. & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220. <http://dx.doi.org/10.1037/a0023353>

Tomas, J., Meléndez, J., Oliver, A., Navarro, E. & Zaragoza, G. (2010). Efectos de método en las escalas de Ryff: Un estudio en población de personas mayores [Method effects in Ryff scales: An study in old age population]. *Psicológica*, 31, 383-400.

Widaman, K.F. (2003). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28(3), 263-

311. http://dx.doi.org/10.1207/s15327906mbr2803_1

Ximénez, M.C. & García, A.G. (2005). Comparación de los métodos de estimación de máxima verosimilitud y mínimos cuadrados no ponderados en el análisis factorial confirmatorio mediante simulación Monte Carlo. *Psicothema*, 17, 528-535.

Ye, S. (2009). Factor structure of the General Health Questionnaire (GHQ-12): The role of wording effects. *Personality and Individual Differences*, 46, 197-201. <http://dx.doi.org/10.1016/j.paid.2008.09.027>

Zwick, W.R. & Velicer, W.F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, 17, 253-269. http://dx.doi.org/10.1207/s15327906mbr1702_5

*Instituto de Investigación de Psicología
Universidad de San Martín de Porres
Lima - Perú*

Fecha de recepción: 6 de febrero de 2013
Fecha de aceptación: 29 de enero de 2014