

La Escala de Deseabilidad Social de Marlowe-Crowne: baremos para la población general española y desarrollo de una versión breve

Sara Gutiérrez, Jesús Sanz*, Regina Espinosa, Clara Gesteira y María Paz García-Vera

Universidad Complutense de Madrid (España).

Resumen: Las adaptaciones españolas de la Escala de Deseabilidad Social de Marlowe-Crowne (M-C SDS) han sido baremadas con estudiantes universitarios y adultos evaluados en contextos de selección/promoción laboral. Dada la influencia de la edad y el contexto en la M-C SDS, estos baremos no son apropiados para adultos voluntarios. Se presentan por primera vez datos normativos y de fiabilidad y una versión breve de la adaptación de la M-C SDS de Ávila Espada y Tomé Rodríguez obtenidos con adultos voluntarios españoles ($N = 575$). Los índices de consistencia interna de la M-C SDS fueron adecuados ($\alpha = .72-.80$). Las diferencias estandarizadas en sus puntuaciones entre adultos voluntarios y estudiantes universitarios o adultos de selección/promoción alcanzaron tamaños del efecto medianos ($d = 0.46$ y 0.70), lo que justifica la creación de baremos específicos para adultos voluntarios. En dos submuestras independientes de adultos voluntarios, una versión de 18 ítems obtuvo índices de consistencia interna para sus ítems mejores que los de los 33 ítems de la escala completa y que no implicaba una pérdida significativa de fiabilidad ($\alpha = .72$ y $.80$) ni de validez nomológica. Se ofrecen baremos globales y separados por grupos de edad para las versiones breve y completa de la M-C SDS.

Palabras clave: Deseabilidad social; medidas de personalidad; baremos de test; versiones breves de test.

Title: The Marlowe-Crowne Social Desirability Scale: norms for the Spanish general population and development of a short version

Abstract: The Spanish adaptations of the Marlowe-Crowne Social Desirability Scale (M-C SDS) were normed on samples of university students or of adults assessed in personnel selection/promotion settings. Given the influence of age and assessment setting on the M-C SDS, these norms are inadequate for adult volunteers. This is the first study that provides normative and reliability data and a short form of Ávila Espada and Tomé Rodríguez's Spanish adaptation of the M-C SDS with Spanish adult volunteers ($N = 575$). Internal consistency estimates for the M-C SDS were adequate ($\alpha = .72-.80$). The standardized differences in M-C SDS scores between the adult volunteer sample and university student or personnel selection/promotion samples reached values of medium size ($d = 0.46$ and 0.70). This result supports the development of specific norms for adult volunteers. In two independent subsamples of adult volunteers, a short form of 18-items showed internal consistency indices for its items that were higher than those for the 33 items of the complete scale. This short form did not involve a significant loss of reliability ($\alpha = .72$ and $.80$) or nomological validity. Global norms and norms for age groups are provided for the short and complete forms of the M-C SDS.

Key words: Social desirability; personality measures; test norms; short test forms.

Introducción

Hace más de 50 años, Crowne y Marlowe (1960) desarrollaron la Escala de Deseabilidad Social de Marlowe-Crowne (M-C SDS) como una medida de deseabilidad social entendida “en términos de la necesidad de los sujetos de responder de una manera culturalmente sancionada” y basada en “la definición de una población de los comportamientos culturalmente aceptables y aprobados y que son, al mismo tiempo, relativamente poco probable que ocurran” (p. 354). Desde entonces, la M-C SDS se ha convertido en una de las escalas de evaluación psicológica más utilizadas. Una búsqueda realizada el 29 de octubre de 2013 en PsycINFO con el término “Marlowe Crowne Social Desirability Scale” en el campo “test y medidas” reveló que desde 2010 se habían publicado 541 trabajos científicos sobre dicho instrumento, 175 solo en 2012.

La M-C SDS también ha sido utilizada en más de una veintena de estudios en España y, de hecho, existen dos adaptaciones españolas de la M-C SDS, una realizada por Ávila Espada y Tomé Rodríguez (1987, citado en Ávila Espada y Tomé Rodríguez, 1989) y otra por Ferrando y Chico (2000), de las cuales la más popular quizás sea la primera. Una búsqueda en PsycINFO parecida a la anterior, pero añadiendo el término “Spain”, localizó 14 estudios que ha-

bían utilizado la versión de Ávila Espada y Tomé Rodríguez (p. ej., Avia, Sanz, Sánchez, Martínez-Arias, Silva y Graña, 1995; Calvo y Eysenck, 2000; Cano Vindel, Sirgo y Pérez Manga, 1994; Nuevo, Montorio, Márquez-González, Cabrera, Izal y Pérez-Rojo, 2009), ocho estudios que habían utilizado la versión de Ferrando y Chico (p. ej., Fernández-González, O’Leary y Muñoz-Rivas, 2013; Luciano y Algarebel, 2006) y dos estudios que habían utilizado traducciones propias o no indicaban la versión española empleada (p. ej., Gracia, García y Lila, 2011).

Tanto en los estudios españoles como en los realizados en otros países, la M-C SDS ha sido utilizada fundamentalmente para dos objetivos. Primero, controlar si las respuestas a un test psicológico pueden estar sesgadas por una tendencia a responder de forma socialmente deseable. De hecho, la M-C SDS se ha convertido en el instrumento más utilizado para medir el sesgo de deseabilidad social (Barger, 2002; Beretvas, Meyers y Leite, 2002). Segundo, la M-C SDS también se ha convertido en el instrumento estándar para definir el estilo de afrontamiento represivo (Mund y Mitte, 2012; Weinberger, 1990; Weinberger, Schwartz y Davison, 1979), de manera que se supone que las personas con bajos niveles de ansiedad rasgo medidos, por ejemplo, con el Cuestionario de Ansiedad Estado-Rasgo (STAI) de Spielberger, Gorsuch y Lushene (1970), y con altos niveles de deseabilidad social medidos por la M-C SDS y entendida como rasgo de personalidad, también denominado defensividad, utilizan un estilo de afrontamiento represivo (“represores”), lo cual puede tener importantes consecuencias para su salud física (Mund y Mitte, 2012; Myers, 2010).

*** Dirección para correspondencia [Correspondence address]:**

Jesús Sanz, Departamento de Personalidad, Evaluación y Psicología Clínica, Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas. CP 28223, Madrid (España).
E-mail: jsanz@psi.ucm.es

Para ambos usos, el disponer de baremos de la M-C SDS adecuados para la población objeto de estudio es una cuestión clave. Así, por ejemplo, como instrumento para controlar el sesgo de deseabilidad social, se han propuesto puntuaciones de corte en la M-C SDS basadas en la combinación de estadísticos de tendencia central y de dispersión (p. ej., una o dos unidades de desviación típica por encima o por debajo de la media) o en las puntuaciones percentiles (p. ej., el percentil 75 y el percentil 25) para detectar a las personas que presentan una imagen socialmente deseable o una imagen socialmente indeseable de sí mismas y que, por tanto, podrían estar distorsionando sus respuestas de forma positiva o negativa para tratar de causar una buena impresión o sugerir que tienen problemas o trastornos psicológicos (Ávila Espada y Tomé Rodríguez, 1989; Salgado, 2005; Tatman, Swogger, Love y Cook, 2009). Como instrumento estándar para definir el estilo de afrontamiento represivo, se han propuesto puntuaciones de corte en la M-C SDS basadas en los estadísticos de tendencia central (p. ej., por encima de la mediana), en las puntuaciones percentiles (p. ej., por encima del tercer cuartil o percentil 75), en la combinación de estadísticos de tendencia central y dispersión (p. ej., un cuarto de desviación típica por encima de la media) o en la distribución de las puntuaciones (p. ej., por encima del tercio superior) para identificar a las personas con altos niveles de defensividad y para considerarlas “represoras” si además presentan una puntuación baja en un instrumento de ansiedad rasgo (p. ej., Calvo y Eysenck, 2000; Cano Vindel et al., 1994; Myers, 2010; Weinberger, 1990; Weinberger et al., 1979).

La adaptación española de la M-C SDS de Ávila Espada y Tomé Rodríguez cuenta con baremos (medias, desviaciones típicas y percentiles) obtenidos con muestras de estudiantes universitarios y de candidatos en procesos de selección de personal o de promoción laboral (Ávila Espada y Tomé Rodríguez, 1989), mientras que la adaptación española de Ferrando y Chico cuenta con baremos (media y desviación típica) obtenidos con una muestra de estudiantes universitarios (Ferrando y Chico, 2000), pero ninguna cuenta con baremos obtenidos con muestras de adultos de la población general en condiciones voluntarias de participación. Sin embargo, la investigación previa sugiere que estos últimos baremos pueden diferir significativamente de los anteriores.

Por un lado, las muestras de estudiantes universitarios suelen tener un rango de edad mucho menor que las muestras extraídas de la población general y existen varios estudios que indican que las personas de más edad muestran puntuaciones significativamente más altas en la M-C SDS que las de menos edad. Por ejemplo, en ocho de los nueve estudios realizados por Ray (1988) con muestras de la población general australiana y con una edad media de 40 años, se encontraron correlaciones positivas y estadísticamente significativas entre la edad y las puntuaciones en una versión breve de la M-C SDS que oscilaban entre .20 y .42. Más recientemente, Nuevo et al. (2009), utilizando una versión breve de la adaptación de Ávila Espada y Tomé Rodríguez, encon-

traron que las personas mayores de 60 años puntuaban más alto que los estudiantes universitarios, alcanzando la diferencia, en términos del estadístico d de Cohen, de 1.22, lo que representa un tamaño del efecto grande según los estándares de Cohen (1988).

Por otro lado, en los contextos de selección de personal o de promoción laboral se asume que las personas están motivadas para presentar una imagen favorable de sí mismos y, por tanto, los baremos obtenidos en dichos contextos no son apropiados en otros contextos como, por ejemplo, en las investigaciones realizadas con voluntarios, donde se asume que las respuestas a los cuestionarios de personalidad son autoinformes verídicos de los comportamientos de las personas evaluadas ya que éstas no tienen nada que ganar o perder como consecuencia de la evaluación, especialmente cuando tales evaluaciones se hacen con garantías de anonimato o confidencialidad. De hecho, en la adaptación de Ávila Espada y Tomé Rodríguez (1989), los propios baremos para las muestras de estudiantes universitarios presentaban unas medias (16.32, 18.27 y 15.81) más bajas que los baremos para candidatos en procesos de selección/promoción laboral (20.28 y 23.17), de manera que todas las pruebas t de diferencias entre pares de medias fueron significativas con $p < .01$ y cuatro de las seis diferencias fueron grandes ($d = 0.96, 0.98, 1.33$ y 1.61). En la misma línea, Tomé Rodríguez (1985) encontró que la media en una versión previa de la adaptación de Ávila Espada y Tomé Rodríguez de un grupo de funcionarios evaluados de forma voluntaria fue significativamente más baja que la de un grupo de funcionarios evaluados durante un proceso de selección/promoción laboral, 16.99 frente a 22.72, $t(179) = 6.23, p < .001$, y dicha diferencia fue grande ($d = 1.01$).

En conclusión, el primer objetivo del presente estudio era presentar por primera vez datos normativos y de fiabilidad de la adaptación española de la M-C SDS de Ávila Espada y Tomé Rodríguez obtenidos en una muestra heterogénea de adultos voluntarios de la población general española, de forma que tales datos permitan utilizar la escala con ese tipo de participantes en contextos de investigación o en contextos en los que se presume que la deseabilidad social es baja (p. ej., contextos clínicos).

Un segundo objetivo del presente estudio fue desarrollar una versión breve de la M-C SDS. Cualquiera de los dos usos de la M-C SDS mencionados anteriormente presupone que la escala se aplica junto a otros test psicológicos. Aunque la M-C SDS no es muy larga (33 ítems) y su cumplimentación no lleva mucho tiempo, es mucho más larga que muchos cuestionarios de personalidad con los que se aplica como, por ejemplo, la escala de ansiedad rasgo del STAI (20 ítems), y, por tanto, sería deseable tener una versión más breve. Con esta motivación, varios autores han propuesto versiones breves de 10, 11, 12, 13 y 20 ítems (Ballard, 1992; Reynolds, 1982; Strahan y Gerbasi, 1972).

Por otro lado, y tal como afirman Ferrando y Chico (2000), la creación de una versión breve podría resolver uno de los problemas psicométricos que tiene la M-C SDS: los

bajos índices de discriminación de bastantes de sus ítems que se traducen en saturaciones factoriales bajas en las soluciones de un solo factor resultantes de analizar factorialmente sus ítems. Por ejemplo, Strahan y Gerbasi (1972) y Ballard (1992) encontraron que la saturación factorial media de los ítems de la M-C SDS fue de tan sólo .35 y .33, respectivamente, mientras que Reynolds (1982) encontró que tan sólo 11 de sus 33 ítems mostraban saturaciones factoriales iguales o mayores de .40. De hecho, todos estos autores, para desarrollar las versiones breves, utilizaron la misma estrategia: eliminar los ítems con saturaciones factoriales bajas para obtener así una versión con menos ítems, pero más consistente y que apenas perdiera fiabilidad frente a la escala completa. No obstante, según Ferrando y Chico (2000), todas estas versiones breves sufren de una misma crítica metodológica: en ninguno de los trabajos citados se realizó un estudio de validación cruzada de las escalas obtenidas (aunque, en el caso de Strahan y Gerbasi, 1972, sí se hizo respecto al coeficiente de fiabilidad, pero no respecto a las saturaciones factoriales).

En conclusión, el segundo objetivo del presente estudio fue desarrollar por primera vez una versión breve de la adaptación de la M-C SDS de Ávila Espada y Tomé Rodríguez que pudiera cumplir las condiciones anteriormente citadas: que sus ítems fueran más consistentes que los de la escala completa, pero sin perder mucha fiabilidad. Aunque Nuevo et al. (2009) utilizaron una versión breve de la adaptación de Ávila Espada y Tomé Rodríguez, en realidad dicha versión

consistió en utilizar los 10 ítems de una de las versiones reducidas desarrolladas por Strahan y Gerbasi (1972) en una muestra de estudiantes universitarios estadounidenses, sin comprobar si dicha versión era psicométricamente adecuada para los adultos españoles. Además, para solventar la crítica señalada por Ferrando y Chico (2000), en el presente estudio se desarrolló la versión breve de la M-C SDS con una muestra de adultos de la población general española y se validó la misma con otra muestra distinta de la misma población.

Método

Participantes

En este estudio participaron 575 adultos (59.8% mujeres) de la población general de la Comunidad de Madrid con edades entre 18 y 84 años (media = 39.4 años; $DT = 15.1$). Estas personas fueron reclutadas mediante la técnica de la “bola de nieve” por estudiantes universitarios de psicología quienes invitaron a sus familiares y amigos a participar en un estudio sobre personalidad e hipertensión siguiendo unos criterios que aseguraran cierta heterogeneidad de la muestra en cuanto a la edad y sexo (véase la Tabla 1). Más información sobre las características sociodemográficas de la muestra se recoge en la Tabla 2, en la cual se puede apreciar su heterogeneidad también en cuanto a su nivel de estudios, estado civil y profesión o situación laboral.

Tabla 1. Comparación de la distribución por sexo y edad de la muestra de adultos de la población general de este estudio con la distribución de la población española.

Edad	Muestra de la población general			Población de España (INE, 2004)		
	18-25 años (<i>n</i> = 187)	26-49 años (<i>n</i> = 192)	≥ 50 años (<i>n</i> = 196)	18-25 años (<i>N</i> = 4948312)	26-49 años (<i>N</i> = 15101930)	≥ 50 años (<i>N</i> = 13455725)
Varones	16.3 %	12.5 %	11.3%	7.5 %	22.7 %	18.2%
Mujeres	16.2 %	20.9 %	22.8%	7.3 %	22.4 %	21.9%
Subtotal	32.5 %	33.4 %	34.1%	14.8%	45.1%	40.1%

Instrumentos

Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS; Crowne y Marlowe, 1960). La M-C SDS consta de 33 ítems, 18 de los cuales (ítems directos) reflejan comportamientos y rasgos socialmente deseables, pero infrecuentes (p. ej., “Nunca dudo en dejar mis cosas si tengo que ayudar a alguien que lo necesita”), mientras que los restantes 15 ítems (ítems inversos) reflejan comportamientos y rasgos indeseables, pero muy frecuentes (p. ej., “A veces me gusta cotillear”). Sobre estos ítems, la persona evaluada debe indicar si son verdaderos o falsos en lo que respecta a sí misma, de manera que se puntúa con un 1 los ítems directos que la persona considera verdaderos o los ítems inversos que considera falsos, mientras que se puntúa con un 0 las respuestas opuestas. La suma de las puntuaciones a los ítems permite obtener una puntuación total entre 0 y 33 en la que una mayor puntuación indica una mayor deseabilidad social entendida como sesgo de respuesta o rasgo de personalidad (de-

fensividad). En este estudio se utilizó la adaptación española de Ávila Espada y Tomé Rodríguez (1987, citado en Ávila Espada y Tomé Rodríguez, 1989). La M-C SDS ha sido extensamente estudiada en cuanto a sus propiedades psicométricas. Por ejemplo, un metaanálisis realizado por Beretvas et al. (2002) sobre los resultados de 93 estudios reveló que su fiabilidad de consistencia interna promedio es de .797 para las mujeres adultas y de .704 para los varones adultos. La adaptación española presenta índices de fiabilidad semejantes, con coeficientes alfa de consistencia interna que oscilan entre .75 y .80 (Ávila Espada y Tomé Rodríguez, 1989).

Inventario de Personalidad NEO Revisado (NEO PI-R; Costa y McCrae, 1999). El NEO PI-R consta de 240 ítems de formato de respuesta de tipo Likert de 5 puntos (de 0 a 4) y fue diseñado para evaluar la personalidad según el modelo de los Cinco Grandes. El NEO PI-R tiene cinco escalas básicas (Neuroticismo, Extraversión, Apertura a la experiencia, Amabilidad y Responsabilidad) que se corresponden con las dimensiones básicas del mismo nombre de los Cinco Gran-

des y 30 escalas específicas, seis por cada escala básica, que pretenden medir las facetas que según el modelo de cinco factores de Costa y McCrae conforman los Cinco Grandes. La adaptación española del NEO PI-R muestra índices de fiabilidad y validez en general adecuados y similares a los que muestra la versión original (Costa y McCrae, 1999; Sanz y García-Vera, 2009). En la muestra de participantes de este estudio los índices de consistencia interna (alfa de Cronbach) fueron adecuados ($\geq .70$) o adecuados con limitaciones ($\geq .60$) para 29 de las 35 escalas del NEO PI-R ($\geq .80$ para las cinco escalas básicas y la de Depresión, $\geq .70$ para otras diez escalas y $\geq .60$ para otras 13), pero no para las seis restantes (.48 para Impulsividad y Valores, .51 para Sensibilidad a los demás, .56 para Acciones, .57 para Sentimientos y .58 para Actitud conciliadora).

Tabla 2. Características sociodemográficas de la muestra de la población general española.

Característica sociodemográfica	Porcentaje (N = 575)
Estudios	
Sin estudios	0.3
Primarios, EGB o equivalentes	16.9
Secundarios, bachillerato, FP o equivalentes	29.5
Universitarios o de especialización	52.3
Otros estudios no reglados	0.2
No contesta/valores perdidos	0.7
Estado civil	
Casado/a o convivencia estable con pareja	46.3
Soltero/a	44.3
Separado/a o divorciado/a	5.4
Viudo/a	3.3
No contesta/valores perdidos	0.7
Profesión o situación laboral	
Obrero, trabajador agrícola o no agrícola no cualificado	1.2
Obrero, trabajador agrícola o no agrícola cualificado	2.1
Agricultor o ganadero	0.5
Personal del sector servicios	13.9
Personal administrativo	10.8
Empresario	2.4
Directivo o gerente de empresa o administración	1.0
Profesional o técnico, profesor o mando intermedio de empresa o administración	12.0
Profesional de las fuerzas armadas	0.7
Sus labores	15.3
Estudiante	27.0
Jubilado o pensionista	4.9
Parado	4.2
Otro trabajo u ocupación	3.8
No contesta/valores perdidos	0.2

Inventario de Expresión de Ira Estado-Rasgo 2 (STAXI-2; Spielberger, Miguel-Tobal, Casado y Cano-Vindel, 2001). El STAXI-2 consta de 49 ítems con formato de respuesta de tipo Likert de 4 puntos (de 0 a 3) que mide diferentes facetas del constructo de ira. En el presente estudio se utilizaron las siete escalas del STAXI-2 que miden la ira rasgo y dos de sus componentes así como cuatro estilos diferentes de expresión

y control de la ira. La adaptación española del STAXI-2 tiene índices de fiabilidad y validez adecuados y parecidos a los que muestra la versión original (Spielberger et al., 2001). En la muestra de participantes del presente estudio los índices de consistencia interna (alfa de Cronbach) de seis de las siete escalas utilizadas fueron adecuados, con valores de .81 (Ira rasgo), .83 (Temperamento de ira), .75 (Reacción de ira), .73 (Expresión externa de ira), .82 (Control externo de ira) y .79 (Control interno de ira), mientras que para la escala restante fue adecuado con limitaciones (.66 para Expresión interna de ira).

Procedimiento

Todos los participantes fueron evaluados de forma individual en una única sesión en la que, además de la M-C SDS, completaron otros cuestionarios de personalidad y se tomaron distintas medidas relacionadas con factores biológicos y psicosociales de riesgo cardiovascular. El protocolo de evaluación se inició explicando a los participantes los objetivos del estudio y obteniendo su consentimiento informado. A continuación, durante 20-25 minutos, se realizó una entrevista de riesgo cardiovascular y se tomaron medidas antropométricas y de presión arterial. Posteriormente, se administró el NEO PI-R y, tras un análisis de parámetros sanguíneos de 5-10 minutos mediante punción digital, se aplicó el STAXI-2 y luego la M-C SDS (aunque por razones del estudio general en el que se encuadraba el presente trabajo, en 233 participantes se prescindió del análisis de parámetros sanguíneos y del STAXI-2).

Análisis estadísticos

Salvo que se indique lo contrario, todos los análisis estadísticos se realizaron con el programa SPSS versión 19. Sobre las puntuaciones totales de la M-C SDS de la muestra total, se calcularon estadísticos de distribución (media, desviación típica, curtosis y simetría), baremos en percentiles y el coeficiente alfa de Cronbach para estimar su fiabilidad de consistencia interna. Para examinar la influencia del sexo y la edad sobre las puntuaciones totales de la M-C SDS, se realizó un análisis de regresión múltiple sobre dichas puntuaciones con el sexo y la edad como variables predictoras, y se examinaron las diferencias de medias entre varones y mujeres y entre pares de grupos de edad (18-25 años, 26-49 años y 50 años o más) mediante pruebas *t* de Student. Para el desarrollo de la versión breve de la M-C SDS, se realizó un análisis factorial sobre las puntuaciones de sus ítems para una primera submuestra obtenida aleatoriamente mediante el SPSS de aproximadamente el 50% de los casos de la muestra total, quedando una segunda submuestra con los casos restantes para la validación cruzada. Puesto que los análisis factoriales realizados sobre la matriz de correlaciones de Pearson entre las puntuaciones de ítems dicotómicos corren el riesgo de agrupar los ítems según su distribución y no según su contenido, proporcionando resultados falsos (Bernstein y

Teng, 1989), se realizó un análisis factorial sobre la matriz de correlaciones tetracóricas mediante el programa FACTOR, versión 8.1 (Lorenzo-Seva y Ferrando, 2006), extrayéndose un factor mediante el procedimiento de mínimos cuadrados no ponderados y seleccionándose los ítems con una saturación igual o mayor de .30, siguiendo así el criterio de Ferrando y Chico (2000). La escala breve resultante fue validada mediante un análisis factorial similar realizado sobre la segunda submuestra, y en la cual se examinó el número de factores según el análisis paralelo optimizado (Timmerman y Lorenzo-Seva, 2011) y el test de la correlación parcial media mínima (MAP) de Velicer (1976) realizados ambos con el programa FACTOR, y mediante un análisis de consistencia interna con el SPSS (coeficiente alfa de Cronbach y correlación corregida ítem-test). Estos análisis sobre la escala breve también se realizaron con la primera submuestra de participantes. La solución factorial para la versión breve encontrada en la primera submuestra también fue objeto de análisis factorial confirmatorio en la segunda submuestra. Este análisis se realizó mediante el programa AMOS versión 19 y utilizando el método de estimación de máxima verosimilitud (ML) y los siguientes seis indicadores de bondad de ajuste (y los correspondientes criterios para considerar el ajuste adecuado) (Hooper, Coughlan y Mullen, 2008): (1) cociente de chi-cuadrado por grados de libertad o χ^2/gl (≤ 3); (2) raíz del residuo cuadrático promedio de aproximación o RMSEA ($\leq .08$); (3) raíz del residuo cuadrático promedio estandarizado o SRMR ($\leq .08$); (4) índice de ajuste comparativo de Bentler o CFI ($\geq .95$), (5) índice de bondad de ajuste corregido o AGFI ($\geq .95$), y (6) índice de ajuste normalizado co-

regido por parsimonia o PNFI (no hay criterio consensuado, pero serviría para comparar modelos; mejor cuanto más próximo a 1). Dado que los ítems de la M-C SDS son dicotómicos, el análisis factorial confirmatorio también fue realizado utilizando el método de estimación de mínimos cuadrados no ponderados (ULS), ya que el mismo no asume el supuesto de distribución normal multivariante de los datos. Como algunos de los índices de bondad de ajuste anteriormente citados no son aplicables a este método, solo se utilizaron SRMR, AGFI y PNFI. Finalmente, para la muestra total, se calcularon estadísticos de distribución de las puntuaciones totales y baremos en percentiles de la versión breve de la M-C SDS, y, para examinar la validez nomológica, las correlaciones de las versiones breve y completa con las medidas del NEO PI-R y del STAXI-2.

Resultados

Distribución de puntuaciones y fiabilidad de la M-C SDS en adultos españoles

Las puntuaciones totales de la M-C SDS oscilaron entre 3 y 31, con una media de 18.36 y una desviación típica de 5.25. Como cabría esperar, la media en la muestra de adultos del presente estudio fue significativamente mayor que la media de estudiantes universitarios españoles, pero menor que la de adultos españoles en procesos de selección/promoción laboral. En la Tabla 3 se resumen los estadísticos descriptivos obtenidos en estudios previos con la adaptación española de Ávila Espada y Tomé Rodríguez. En total, en los estudios

Tabla 3. Medias y desviaciones de la adaptación española de la Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS) de Ávila Espada y Tomé Rodríguez en distintos estudios con muestras de participantes españoles.

Población / Referencia	Rango de edad	Media de edad	Sexo	N	Media	DT
Estudiantes universitarios						
Ávila y Tomé (1989; FAC-2)	s.d.	s.d.	s.d.	557	15.81	4.37
Ávila y Tomé (1989; FAC-3)	s.d.	s.d.	s.d.	127	16.32	5.04
Ávila y Tomé (1989; FAR-1)	21-38	23.7	74.8%	115	18.27	4.91
Avía et al. (1995; Muestra 1) ^a	20-53	22.5	85.5%	658	15.86	5.10
Total de estudios previos	21-53	22.7 ^b	83.9%	1457	16.07 ^b	4.81 ^c
Candidatos en selección/promoción						
Ávila y Tomé (1989; MIN-1)	25-50	34.5	29.0%	93	20.28	5.48
Ávila y Tomé (1989; FUN-3)	22-45	33.6	0%	148	23.17	5.22
Total de estudios previos	22-50	33.9 ^b	11.2%	241	22.05 ^b	5.32 ^c
Adultos de la población general						
Este estudio	18-84	39.4	59.8%	575	18.36	5.25

Nota. s.d. = sin datos.

^aDatos obtenidos expresamente para este estudio de la Muestra 1 de participantes de Avía et al. (1995).

^bValor medio ponderado por el número de participantes de cada estudio/muestra.

^cDesviación típica conjunta; $DT_{conjunta} = \sqrt{\{(n_1 - 1)DT_1^2 + (n_2 - 1)DT_2^2 + \dots + (n_k - 1)DT_k^2\} / [(n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1)]}$.

previos se ha administrado dicha adaptación a 1457 estudiantes universitarios españoles y, aunque la desviación típica conjunta fue similar a la obtenida en este estudio (4.81 frente a 5.25), la media ponderada por el número de participantes en cada estudio/muestra fue significativamente más baja que la media del presente estudio (16.07 frente a 18.36), $t(2030) = 9.07$, $p < .0001$, alcanzando la diferencia un tamaño del

efecto casi mediano ($d = 0.46$). En estudios previos (véase la Tabla 3) se ha aplicado la M-C SDS a 241 adultos españoles en procesos de selección/promoción laboral y, aunque la desviación típica conjunta también fue similar a la obtenida en el presente estudio (5.32 frente a 5.25), la media ponderada por el número de participantes fue significativamente más alta que la media del presente estudio (22.05 frente a 18.36),

$t(814) = 9.07, p < .0001$, alcanzando la diferencia un tamaño del efecto entre mediano y grande ($d = 0.70$).

La distribución de las puntuaciones totales de la M-C SDS en la muestra de adultos españoles del presente estudio se ajustaba razonablemente a una distribución normal. Así, la mediana (18) fue igual a la moda (18), ambas muy similares a la media (18.36) y los índices de asimetría (-0.14) y curtosis (-0.35) fueron cercanos a 0.

El coeficiente alfa de fiabilidad de consistencia interna para la muestra total de adultos españoles fue .78, siendo .75 y .80 para las mujeres y los varones, respectivamente, y .72, .75 y .80 para los grupos de edad de 18-25 años, 26-49 años y 50 años o más, respectivamente.

Baremos de la M-C SDS para adultos españoles

En la Tabla 4 se recogen las puntuaciones percentiles de la M-C SDS para la muestra total de adultos. Un análisis de regresión múltiple sobre la puntuación total de la M-C SDS con el sexo y la edad como variables de predicción reveló que ambas variables estaban relacionadas de forma independiente y estadísticamente significativa con dicha puntuación,

aunque la edad presentaba una mayor relación ($\beta = .285, p < .001$) que el sexo ($\beta = .128, p = .001$). De hecho, aunque la puntuación total media en la M-C SDS de las 344 mujeres ($M = 19.11, DT = 5.51$) fue significativamente mayor que la de los 231 varones ($M = 17.25, DT = 19.11$), $t(573) = 4.22, p < .001$, dicha diferencia no llegó a alcanzar un tamaño del efecto mediano ($d = 0.36$). Sin embargo, la diferencia entre el grupo de adultos de 18-25 años y el grupo de 50 años o más, no sólo fue también estadísticamente significativa, $t(381) = 7.21, p < .001$, sino que además dicha diferencia ($d = 0.74$) superó el estándar para un tamaño del efecto mediano ($d = 0.50$) y casi alcanzó un tamaño grande ($d = 0.80$). No obstante, la diferencia entre el grupo de 18-25 años y el de 26-49 años tampoco alcanzó un tamaño del efecto mediano ($d = 0.32$) ni lo hizo la diferencia entre el grupo de 26-49 años y el de 50 años o más ($d = 0.42$), aunque ambas diferencias fueron estadísticamente significativas, $t(377) = 3.08, p = .002$, y $t(386) = 4.12, p < .001$, respectivamente. En definitiva, dada la mayor relevancia de la variable edad, en la Tabla 4 también se presentan las puntuaciones percentiles de la M-C SDS para los tres grupos de edad.

Tabla 4. Puntuaciones percentiles, medias y desviaciones típicas de las versiones completa y breve de la Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS) para la muestra total de adultos de la población general española y para sus submuestras en función de la edad.

Percentil	Puntuación directa en adultos de la población general española								
	M-C SDS completa (33 ítems)				M-C SDS breve (18 ítems)				
	Total	18-25 años	26-49 años	≥ 50 años	Total	18-25 años	26-49 años	≥ 50 años	
1	6	5	6	5	1	1	3	2	
5	10	8	10	10	3	2	4	4	
10	11	11	11	13	5	3	5	6	
15	13	11	13	14	6	5	6	7	
20	14	12	13	16	6	5	6	8	
25	15	14	14	17	7	6	7	9	
30	16	15	15	18	7	6	7	9	
35	16	15	16	18	8	7	8	10	
40	17	16	17	19	9	7	9	10	
45	18	16	18	20	9	8	9	11	
50	18	17	18	21	10	8	10	11	
55	19	17	19	22	10	9	10	12	
60	20	18	20	22	11	9	11	12	
65	21	18	20	23	11	9	11	13	
70	21	19	21	23	12	10	11	13	
75	22	20	22	24	12	11	12	14	
80	23	21	23	25	13	11	12	14	
85	24	21	24	26	14	12	14	15	
90	25	22	25	28	14	13	15	16	
95	27	24	26	28	16	14	15	17	
99	29	27	28	31	17	15	16	18	
Media	18.36	16.58	18.11	20.30	9.59	8.11	9.58	11.02	
DT	5.25	4.60	5.02	5.43	3.68	3.35	3.34	3.75	
<i>n</i>	575	187	192	196	575	187	192	196	

Desarrollo de una versión breve de la M-C SDS

En la Tabla 5 se presentan las saturaciones factoriales de los 33 ítems de la M-C SDS en el factor no rotado extraído tras realizar un análisis factorial con las puntuaciones de la

primera submuestra de participantes ($n = 279$), así como las correlaciones corregidas ítem-total. De los 33 ítems de la M-C SDS, 18 mostraron saturaciones factoriales que, en valor absoluto, fueron iguales o mayores que .30 y, por tanto, dichos ítems fueron seleccionados para la versión breve de la

M-C SDS. Todos los ítems seleccionados también mostraron correlaciones ítem-total iguales o mayores que .30, salvo cuatro que mostraron correlaciones de .28. Además, todos los ítems seleccionados mostraron índices de dificultad adecuados que no excedían los valores extremos de $p \geq .90/\leq .10$ propuestos por algunos autores, aunque dos de ellos (ítems 4 y 10 con $p = 0.18$ y 0.87 , respectivamente) sí lo hacían respecto a los valores no deseables de dificultad de $p \geq .80/\leq .20$ propuestos por otros autores (Miller, McIntire y Lovler, 2011).

Tabla 5. Saturaciones factoriales y correlaciones corregidas ítem-total de los 33 ítems de la Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS) en la primera submuestra de participantes ($n = 279$).

Ítems de la M-C SDS*	Solución unifactorial ^a	$r_{\text{ítem-total}}$
1	-.29	.22
2	-.44	.34
3	.61	.37
4	.51	.34
5	.24	.26
6	-.38	.37
7	.44	.43
8	-.51	.35
9	.21	.20
10	-.34	.28
11	.37	.28
12	-.46	.36
13	.32	.36
14	.46	.37
15	-.17	.18
16	-.11	.22
17	.53	.37
18	-.25	.26
19	.41	.28
20	-.28	.29
21	-.35	.30
22	.21	.17
23	.68	.39
24	-.27	.21
25	-.07	.17
26	.26	.25
27	-.32	.28
28	.40	.30
29	-.19	.17
30	-.46	.34
31	-.01	.00
32	-.06	.10
33	-.07	.14
Media	.32 ^b	.27

Nota. Las saturaciones factoriales iguales o mayores, en valor absoluto, de .30 y las correlaciones ítem-total iguales o mayores de .30 aparecen en negrita.

*Los números corresponden a los de la adaptación de la M-C SDS de Ávila Espada y Tomé Rodríguez (1989).

^aExtracción de un factor mediante el método de mínimos cuadrados no ponderados sobre la matriz de correlación tetracóricas.

^bMedia de los valores absolutos.

Con los ítems seleccionados, se realizó un análisis factorial sobre las puntuaciones de la primera submuestra de participantes. Tanto el test MAP de Velicer (1976) como el análisis paralelo optimizado (Timmerman y Lorenzo-Seva,

2011) revelaron que una solución unifactorial era la más adecuada para la matriz de correlaciones tetracóricas entre los 18 ítems de la versión breve. Por tanto, se extrajo un factor mediante el procedimiento de mínimos cuadrados no ponderados, el cual explicó el 35% de la varianza. Las saturaciones factoriales en ese factor de los 18 ítems de la versión breve se recogen en la Tabla 6, así como sus correlaciones corregidas ítem-total. Todos los ítems de la versión breve mostraron saturaciones factoriales que, en valor absoluto, fueron no sólo iguales o mayores que .30, sino también mayores que .45. Además, todos esos ítems también mostraron correlaciones corregidas ítem-total iguales o mayores que .30, salvo uno cuya correlación fue .26. El coeficiente alfa de consistencia interna de la versión breve para la primera submuestra fue .80.

Estos mismos análisis fueron realizados con las puntuaciones de la segunda submuestra ($n = 296$). De nuevo, tanto el test MAP de Velicer como el análisis paralelo optimizado revelaron que una solución unifactorial era la más adecuada y, por consiguiente, también se extrajo un factor, el cual explicó el 27% de la varianza. Las saturaciones factoriales en ese factor de los 18 ítems de la versión breve para la segunda submuestra de participantes se recogen en la Tabla 6, así como sus correlaciones corregidas ítem-total. Todos los ítems de la versión breve mostraron saturaciones que, en valor absoluto, fueron iguales o mayores que .30. Además, 10 de esos ítems también mostraron correlaciones ítem-total iguales o mayores que .30, mientras que los ocho restantes mostraron correlaciones iguales o mayores que .22. El coeficiente alfa de consistencia interna de la versión breve para la segunda submuestra fue .72.

Dado que la solución unifactorial tan solo explicaba el 27% de la varianza de los datos en la segunda submuestra y el 35% en la primera, cabría la posibilidad de que existiese algún otro factor más en la estructura de la versión breve. Se realizó, por tanto, un análisis factorial exploratorio sobre la primera submuestra con los mismos procedimientos anteriores, pero extrayendo dos factores y rotándolos oblicuamente mediante *promax*. La matriz rotada resultante mostraba que los ítems 2, 3, 4, 7, 13, 14, 17, 19, 23, 27 y 30 conformaban el primer factor (con saturaciones $\geq .30$ en dicho factor), que los ítems 6, 8, 10, 11, 12, 21 y 28 conformaban el segundo factor (con saturaciones $\geq .30$ en este factor) y que ambos factores estaban correlacionados ($r = -.61$). Sin embargo, dicha solución bifactorial no era interpretable psicológicamente. A pesar de este problema, se realizaron sendos análisis factoriales confirmatorios sobre la segunda submuestra poniendo a prueba dicho modelo y el modelo unifactorial. Con el método de estimación ML los índices de bondad de ajuste para el modelo unifactorial fueron $\chi^2/\text{gl} = 2.038$, RMSEA = .059 (IC al 90% = .049–.069), SRMR = .064, CFI = .719, AGFI = .881 y PNFI = .510, mientras que para el bifactorial fueron $\chi^2/\text{gl} = 1.788$, RMSEA = .052 (IC al 90% = .041–.062), SRMR = .060, CFI = .788, AGFI = .899 y PNFI = .554. Para ambos modelos, había índices que indicaban un ajuste adecuado (χ^2/gl , RMSEA y SRMR), pero

también índices que indicaban un pobre ajuste (CFI y AGFI), y los valores de todos esos índices eran muy similares para ambos modelos, aunque superiores muy ligeramente para el modelo bifactorial. Cuando los análisis factoriales confirmatorios se realizaron con el método de estimación ULS los índices de ajuste para el modelo unifactorial fueron SRMR = .064, AGFI = .932 y PNFI = .709, mientras que para el bifactorial fueron SRMR = .059, AGFI = .939 y PNFI = .722. Por tanto, los resultados con el método ULS fueron similares a los obtenidos con ML: para ambos modelos había índices que indicaban un ajuste adecuado (SRMR), pero también índices que indicaban un pobre ajuste (AGFI), y los valores de todos esos índices eran muy parecidos en ambos modelos, aunque superiores muy levemente para el bifactorial.

Tabla 6. Saturaciones factoriales y correlaciones corregidas ítem-total de la versión breve de la Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS) en las submuestras de participantes primera y segunda ($n = 279$ y 296 , respectivamente).

Ítem de la M-C SDS*	Solución unifactorial ^a		$r_{\text{ítem-total}}$	
	Submuestra 1	Submuestra 2	Submuestra 1	Submuestra 2
2	-.57	-.50	.40	.33
3	.69	.47	.48	.31
4	.61	.55	.37	.33
6	-.48	-.50	.34	.33
7	.55	.73	.39	.49
8	-.60	-.39	.40	.23
10	-.47	-.42	.26	.25
11	.49	.37	.34	.24
12	-.58	-.42	.42	.29
13	.47	.64	.32	.42
14	.57	.46	.41	.32
17	.65	.49	.44	.31
19	.53	.35	.34	.22
21	-.47	-.34	.34	.23
23	.72	.55	.47	.32
27	-.46	-.37	.32	.24
28	.51	.44	.30	.25
30	-.57	-.42	.41	.30
Media	.55 ^b	.47 ^b	.37	.30

Nota. Las saturaciones factoriales iguales o mayores, en valor absoluto, de .30 y las correlaciones ítem-total iguales o mayores de .30 aparecen en negrita.

*Los números corresponden a los de la adaptación de la M-C SDS de Ávila Espada y Tomé Rodríguez (1989).

^aExtracción de un factor mediante el método de mínimos cuadrados no ponderados sobre la matriz de correlación tetracóricas.

^bMedia de los valores absolutos.

Fiabilidad, validez nomológica, distribución de puntuaciones y baremos de la versión breve de la M-C SDS en adultos españoles

En la muestra total de participantes, el coeficiente alfa de consistencia interna de la versión breve de la M-C SDS fue .76 y la correlación entre su puntuación total y la puntuación total de la escala completa fue .93.

Respecto a la validez nomológica, en la Tabla 7 se presentan las correlaciones de las versiones completa y breve de

la M-C SDS con las variables de personalidad que mide el NEO PI-R y el STAXI-2. El patrón de correlaciones de ambas versiones fue muy similar. Todas las correlaciones que fueron estadísticamente significativas con la versión completa (33 correlaciones) lo fueron también con la versión breve y en todas ellas la dirección de la correlación fue igual, de manera que la media de las correlaciones estadísticamente significativas (en valor absoluto) para las versiones completa y breve fue muy similar (.29 y .27, respectivamente). Es más,

Tabla 7. Correlaciones de las versiones completa y breve de la Escala de Deseabilidad Social de Marlowe y Crowne (M-C SDS) con distintas variables de personalidad en la muestra total de participantes.

Variables de personalidad	M-C SDS completa	M-C SDS breve
NEO PI-R ($N = 575$)		
Neuroticismo	-.39**	-.35**
Ansiedad	-.25**	-.25**
Hostilidad	-.36**	-.36**
Depresión	-.26**	-.21**
Ansiedad social	-.24**	-.19**
Impulsividad	-.36**	-.33**
Vulnerabilidad	-.33**	-.28**
Extraversión	.02	-.06
Cordialidad	.20**	.17**
Gregarismo	.06	.00
Asertividad	.03	-.04
Actividad	.04	-.05
Búsqueda de emociones	-.26**	-.27**
Emociones positivas	.07	.01
Apertura a la experiencia	-.10*	-.11*
Fantasia	-.26**	-.24**
Estética	.03	.02
Sentimientos	-.05	-.08
Acciones	-.04	-.06
Ideas	.00	-.00
Valores	-.10*	-.11**
Amabilidad	.48**	.50**
Confianza	.33**	.32**
Franqueza	.41**	.42**
Altruismo	.38**	.35**
Actitud conciliadora	.34**	.39**
Modestia	.20**	.24**
Sensibilidad a los demás	.20**	.22**
Responsabilidad	.42**	.30**
Competencia	.30**	.19**
Orden	.18**	.10*
Sentido del deber	.46**	.40**
Necesidad de logro	.26**	.13**
Autodisciplina	.41**	.30**
Deliberación	.33**	.27**
STAXI-2 ($N = 342$)		
Ira rasgo	-.29**	-.29**
Temperamento de ira	-.28**	-.25**
Reacción de ira	-.21**	-.22**
Expresión externa de ira	-.40**	-.42**
Expresión interna de ira	-.16**	-.12*
Control externo de ira	.31**	.32**
Control interno de ira	.19**	.21**

Nota. En negrita las correlaciones que difieren $\geq .10$ en valor absoluto entre ambas versiones del M-C SDS. ** $p < .01$. * $p < .05$.

en la mayoría de las variables de personalidad (30 de 42; 71%) la diferencia en valor absoluto entre la correlación con la versión completa y la correlación con la versión breve fue $\leq .05$, y en tan solo cuatro variables de personalidad dicha diferencia fue $\geq .10$.

Las puntuaciones totales de la versión breve oscilaron entre 0 y 18, con una media de 9.59 y una desviación típica de 3.68, y su distribución se ajustó razonablemente a una distribución normal, de forma que, aunque la moda (7) fue inferior a la mediana (10), esta última fue muy similar a la media (9.59) y los índices de asimetría (-0.06) y curtosis (-0.59) estaban cercanos a 0.

En la Tabla 4 se recogen las puntuaciones percentiles de la versión breve de la M-C SDS para la muestra total de adultos. Un análisis de regresión múltiple reveló que tanto el sexo como la edad estaban relacionados de forma independiente y estadísticamente significativa con la puntuación total de la versión breve, aunque la edad presentaba una mayor relación ($\beta = .327, p < .001$) que el sexo ($\beta = .100, p = .011$). De hecho, aunque la puntuación total media en la versión breve de las 344 mujeres ($M = 10.06, DT = 3.52$) fue significativamente mayor que la de los 231 varones ($M = 8.90, DT = 3.81$), $t(573) = 3.71, p < .001$, dicha diferencia no alcanzó un tamaño del efecto mediano ($d = 0.32$). Sin embargo, la diferencia entre el grupo de adultos de 18-25 años y el de 50 años o más, no sólo fue estadísticamente significativa, $t(381) = 7.99, p < .001$, sino que además superó el estándar para un tamaño del efecto grande ($d = 0.82$). No obstante, la diferencia entre el grupo de 18-25 años y el de 26-49 años tampoco alcanzó un tamaño del efecto mediano ($d = 0.44$) ni lo hizo la diferencia entre el grupo de 26-49 años y el de 50 años o más ($d = 0.41$), siendo ambas diferencias estadísticamente significativas, $t(377) = 4.26, p = .001$, y $t(386) = 3.99, p < .001$, respectivamente. En definitiva, dada la mayor relevancia de la variable edad, en la Tabla 4 también se presentan las puntuaciones percentiles de la versión breve de la M-C SDS para los tres grupos de edad.

Discusión

Tal y como cabría esperar de la literatura científica, los resultados del presente estudio confirman que las puntuaciones en la adaptación de Ávila Espada y Rodríguez-Tomé de la M-C SDS que obtienen los adultos voluntarios españoles difiere notablemente de las puntuaciones que obtienen los estudiantes universitarios españoles (casi media unidad de desviación típica; $d = 0.46$) o los adultos españoles en procesos de selección/promoción laboral (casi tres cuartos de unidad de desviación típica; $d = 0.70$), los cuales son los grupos para los que existen actualmente baremos (Ávila Espada y Rodríguez-Tomé, 1989). Estos resultados implican que la utilización de estos baremos puede distorsionar significativamente la interpretación que se haga de las puntuaciones de un adulto no estudiante universitario que conteste de forma voluntaria a la M-C SDS, presentándolo con niveles más altos o más bajos de discapacidad social o de defensividad o repre-

sión de los que realmente tiene en comparación con la población de referencia. Por consiguiente, esos resultados justifican la elaboración de nuevos baremos para la adaptación española de la M-C SDS que estén basados en una muestra de adultos voluntarios procedentes de la población general española y este precisamente fue el primer objetivo del presente estudio.

La utilidad de estos baremos se puede apreciar, por ejemplo, al analizar bajo su luz los resultados del estudio de Cano Vindel et al. (1994) en el que para examinar la hipótesis de que las personas con cáncer presentan en mayor medida un estilo represivo de afrontamiento, se comparaba un grupo de 31 personas con cáncer con una edad media de 51.7 años y un grupo de control de 35 personas sanas equiparado en la variable sexo y con una edad media de 40.8 años. El grupo de control de personas sanas fue seleccionado mediante muestreo incidental y su puntuación media en la adaptación de Ávila Espada y Rodríguez-Tomé de la M-C SDS fue 14 mientras que la puntuación media del grupo de personas con cáncer fue 23.1, diferencia que fue estadísticamente significativa y era consistente con la hipótesis del estudio al demostrar que las personas con cáncer muestran una discapacidad social más alta (una de las características definitorias del estilo represivo de afrontamiento). Es más, Cano Vindel et al. (1994), en consonancia con la literatura científica, definieron la existencia de un estilo represivo de afrontamiento por la presencia de una puntuación en una medida de ansiedad rasgo inferior a la puntuación media de la muestra total de participantes menos un cuarto de desviación típica y por la presencia de una puntuación en la M-C SDS superior a la puntuación media de la muestra total más un cuarto de desviación típica (> 20.1 en la M-C SDS). Con esta definición, los investigadores encontraron que el 48.4% de los pacientes con cáncer eran "represores" mientras que solo el 5.7% de las personas sanas tenían un estilo represivo de afrontamiento, lo cual corroboraba su hipótesis. Estos resultados podrían haberse entendido y discutido con mayor fundamento teniendo en cuenta los baremos obtenidos en el presente estudio. En este sentido, por ejemplo, parece claro que el procedimiento incidental de selección del grupo de control no fue muy efectivo a la hora de obtener una muestra no sesgada de la población de personas sanas, ya que su media en la M-C SDS fue notablemente más baja que la de los baremos de este estudio (14 frente a 18.4), lo que obviamente afectó a la posibilidad de identificar en dicho grupo a personas "represoras" las cuales, por definición, deben presentar una puntuación alta en la M-C SDS. Por otro lado, teniendo en cuenta los baremos de este estudio, no parece que el nivel de discapacidad social del grupo de pacientes de cáncer fuera tan alto, especialmente si se utiliza el baremo más apropiado para la edad media del grupo de pacientes con cáncer (23.1 frente a 20.3; véase la Tabla 4). Además, dado los baremos de este estudio, no parece que la utilización de una puntuación de corte de 20.1 en la M-C SDS represente una puntuación tan extrema en discapacidad social que permita identificar adecuadamente un estilo represivo de

afrontamiento, ya que la misma tan sólo representa un percentil 60 en los baremos generales del presente estudio o un percentil 50 en los baremos de las personas de 50 años o más. Por tanto, parece más apropiado que tanto en el estudio de Cano Vindel et al. (1994) como en futuros estudios españoles que evalúen la presencia de un estilo represivo de afrontamiento en pacientes con cáncer, los cuales, dada la naturaleza de este grupo de enfermedades, suelen tener como media una edad superior a los 40-50 años, se utilicen *a priori* unos baremos como los aquí presentados para definir la presencia de un estilo represivo de afrontamiento, ya que dichos baremos, en contraste con los actualmente disponibles (estudiantes universitarios y adultos en procesos de selección/promoción laboral), han sido obtenidos con una muestra heterogénea y relativamente amplia ($N = 575$) de adultos voluntarios de la población general española y, aunque no exentos de las limitaciones que se abordarán a continuación, suponen una mejora frente a las limitaciones que implica definir el estilo represivo de afrontamiento a partir del rendimiento en la M-C SDS de un grupo de tan sólo 65 participantes.

En el presente estudio se han elaborado tales baremos de forma global y de forma independiente para tres grupos de edad (18-25 años, 26-49 años y 50 años y más) a partir de una muestra incidental de 575 adultos españoles. Dada las limitaciones inherentes al procedimiento utilizado para seleccionar la muestra de baremación, cabría cuestionar su grado de representatividad de la población española. Sin embargo, en relación con variables tan importantes como la edad y el sexo, el perfil de la muestra de baremación respecto a algunos grupos resultantes de la combinación de edad y sexo (mujeres de 26-49 años y mujeres de 50 años y más) fue muy similar al que se encuentra en la población española (véase la Tabla 1). Por otro lado, el tamaño muestral de los baremos se puede considerar, en función de los criterios de Prieto y Muñiz (2000), suficiente para los tres grupos de edad ($150 \leq N < 300$) y moderado para el grupo global ($300 \leq N < 600$). No obstante, la utilización de un muestreo aleatorio de selección de participantes pertenecientes, de forma proporcional, a distintas regiones geográficas españolas hubiera mejorado enormemente la representatividad de los baremos. Es más, sería deseable la realización de nuevos estudios con muestras de voluntarios de la población general que permitieran confirmar la representatividad de los baremos propuestos.

La decisión de calcular baremos independientes para tres grupos de edad vino motivada por la presencia de diferencias significativas entre tales grupos y de una relación significativa, e independiente del sexo, entre edad y puntuación total en la M-C SDS, diferencias y relación que son consistentes con los resultados previos encontrados en muestras de la población general tanto de otros países (Ray, 1988) como de España (Nuevo et al., 2009).

Por otro lado, la utilización de la M-C SDS para evaluar la deseabilidad social, defensividad o represión en adultos españoles que completan el inventario de forma voluntaria

se ve refrendada por la obtención de una distribución prácticamente normal de su puntuación total y de índices adecuados de fiabilidad de consistencia interna. Respecto al ajuste a la curva normal, los datos del presente estudio confirman los resultados de estudios previos con ambas adaptaciones españolas de la M-C SDS (Ávila Espada y Tomé Rodríguez, 1989; Ferrando y Chico, 2000). En cuanto a la fiabilidad, los índices de consistencia interna para la muestra total y para las diferentes submuestras en función del sexo y la edad oscilaron entre .72 y .80. Estos valores no sólo fueron adecuados ($\alpha \geq .70$) de acuerdo con los criterios de Prieto y Muñiz (2000), sino que fueron similares a los obtenidos en estudios previos con muestras españolas de estudiantes universitarios (.75) o de candidatos en procesos de selección/promoción laboral (.79 y .80) (Ávila Espada y Tomé Rodríguez, 1989), incluso con la adaptación española de Ferrando y Chico (.78 en una muestra de estudiantes universitarios; Ferrando y Chico, 2000), y similares a los obtenidos con la versión original del instrumento en otros países con muestras de estudiantes universitarios o de adultos de la población general (.80 y .70 para los varones y mujeres, respectivamente, según el metaanálisis de Beretvas et al., 2002). Sin embargo, es cierto que para valorar el coeficiente de fiabilidad de un test hay que tener en cuenta para qué se va a utilizar, de manera que aunque un coeficiente $\alpha \geq .70$ sería adecuado y suficiente si el test se va a utilizar en investigación (p. ej., para examinar la relación entre defensividad y otras variables de personalidad), dicho coeficiente debería ser mayor si se va a utilizar en ámbitos aplicados: $\geq .80$ para pronósticos y clasificaciones colectivas y $\geq .90$ para diagnósticos y pronósticos individuales (Yela, 1984/2005). Por tanto, en función de los coeficientes de fiabilidad aquí presentados, parece que el uso más adecuado de la M-C SDS sería en ámbitos de investigación o, en todo caso, en ámbitos aplicados, para pronósticos y clasificaciones colectivas (p. ej., para determinar si los pacientes con cáncer tienen un mayor estilo represivo de afrontamiento), pero no debería utilizarse en ámbitos aplicados (o hacerse con muchísima precaución) para realizar diagnósticos y pronósticos individuales (p. ej., para detectar personas que podrían estar distorsionando sus respuestas de forma positiva o negativa e invalidar o cuestionar sus respuestas a otros cuestionarios de personalidad).

El segundo objetivo del presente estudio fue desarrollar por primera vez una versión breve de la adaptación española de Ávila Espada y Tomé Rodríguez cuyos ítems fueran más consistentes que los de la escala completa, pero sin que la misma perdiera mucha fiabilidad respecto a esta última. La versión breve de 18 ítems que se desarrolló cumplía adecuadamente ambas condiciones no sólo en la submuestra de adultos con la que fue creada, sino también en otra submuestra independiente de adultos que sirvió para realizar una validación cruzada. En ambas submuestras, la saturación factorial media de los 18 ítems de la versión breve fue más alta que la de los 33 ítems de la escala completa (.55 y .47 frente a .32) y también más alta que la de los 33 ítems de la escala completa en su versión original (p. ej., .35 y .33 en

Strahan y Gerbasi, 1972, y Ballard, 1992, respectivamente). Es más, en ambas submuestras, la correlación corregida ítem-total media de los 18 ítems de la versión breve fue más alta que la de los 33 ítems de la escala completa (.37 y .30 frente a .27). En consecuencia, en ambas submuestras, el coeficiente alfa de fiabilidad para la versión breve no sólo fue adecuado (alfa \geq .70), sino también similar al obtenido para la escala completa (.80 y .72 frente a .78), por lo que no parece que la reducción en el número de ítems de la versión breve (un 45% menos) implique una disminución significativa de la fiabilidad de consistencia interna.

Por supuesto, esos coeficientes de fiabilidad hacen referencia a la puntuación total de la escala y tienen sentido en la medida en que la versión breve desarrollada mide una sola dimensión. Con algunas limitaciones, los resultados de los análisis factoriales exploratorios y confirmatorios realizados indican que el modelo de un único factor muestra un ajuste razonablemente adecuado, aunque mejorable, a los datos, y que la mejora que se produce al pasar de uno a dos factores es bastante pequeña. Es más, la solución bifactorial encontrada con la presente muestra de participantes es difícil de interpretar psicológicamente y tampoco refleja otras soluciones bifactoriales encontradas en la literatura científica como, por ejemplo, la solución que distingue un factor compuesto por los ítems que puntúan en deseabilidad social (o defensividad) cuando se contestan “verdadero” (factor denominado de “atribución” porque implica la atribución de comportamientos y características de personalidad socialmente deseables) y otro factor compuesto por los ítems que puntúan en deseabilidad social (o defensividad) cuando se contestan “falso” (factor denominado de “negación” porque implica la negación de comportamientos y características socialmente indeseables) (Ramanaiyah y Martin, 1980). Todo ello permite afirmar provisionalmente que las puntuaciones de la versión breve de la M-C SDS miden una sola dimensión, aunque no cabe duda de que el apoyo empírico a la unidimensionalidad de la escala no es todo lo robusto que sería deseable y es una limitación de la misma, y que futuras investigaciones deberían replicar esa unidimensionalidad y analizar la posible existencia de otros factores en cuya identificación la interpretabilidad psicológica debería ser determinante y necesitaría de nuevos desarrollos teóricos sobre los constructos de deseabilidad social, defensividad o represión.

Dado que la elaboración de la versión breve de la M-C SDS se basó en la selección de los ítems más consistente, cabría la posibilidad de que dicha versión estuviera formada por ítems redundantes y hubiera perdido capacidad para representar las diferentes facetas del constructo (deseabilidad social, defensividad o represión) que pretende medir. Es más, aunque esto último no hubiese ocurrido, cabría también la posibilidad de que la simple eliminación del 45% de sus ítems hubiera supuesto una disminución significativa en la validez de la M-C SDS. Los resultados obtenidos sugieren que ambas posibilidades no parecen plausibles, ya que el pa-

trón, signo y tamaño de las correlaciones de la versión breve de la M-C SDS con las 42 variables de personalidad que miden el NEO PI-R y el STAXI-2 eran muy similares a las que presentaba la versión completa de la M-C SDS con dichas variables.

En la misma dirección, puesto que es deseable que la selección de ítems tenga en cuenta más de un criterio (p. ej., la consistencia de los ítems, pero también su nivel de dificultad) para no provocar sesgos, es importante subrayar también que todos los ítems de la versión breve de la M-C SDS presentaban índices de dificultad adecuados, salvo los ítem 4 y 10. Aunque cabría la posibilidad de eliminar esos dos ítems para tener una versión breve más depurada, su eliminación, aunque no afectaba sustancialmente a los resultados de los análisis factoriales exploratorios y confirmatorios, reducía la fiabilidad de la escala (de .78 a .76) y reducía los coeficientes de correlación con el NEO PI-R y el STAXI-2 (p. ej., el 71% de las correlaciones de la versión breve de 18 ítems se diferenciaban en una magnitud \leq .05 de las correlaciones de la versión completa, pero ese porcentaje se reducía al 62% en el caso de una versión breve de 16 ítems), por lo que se decidió mantenerlos en la versión breve de la M-C SDS.

Además, la utilización de la versión breve de la M-C SDS para evaluar la deseabilidad social o la defensividad o represión en adultos españoles se ve refrendada por la obtención de una distribución prácticamente normal de su puntuación total y de una correlación muy alta entre dicha puntuación y la puntuación total de la escala completa (.93).

Como ocurría con la escala completa, se han encontrado diferencias significativas en la versión breve de la M-C SDS entre tres grupos de edad (18-25 años, 26-49 años y 50 años y más) y una relación significativa, e independiente del sexo, entre su puntuación total y la edad, diferencias y relación que son consistentes con la literatura científica previa con otras versiones breves de la M-C SDS (Nuevo et al., 2009; Ray, 1988). En consecuencia, en el presente estudio, se han calculado baremos de la versión breve para la población general española de manera global y de forma independiente para los tres grupos de edad mencionados.

Finalmente, es importante señalar que los datos del presente estudio avalan la utilización de la versión breve de 18 ítems de la M-C SDS en adultos de la población general española, pero sería conveniente replicar los resultados encontrados no sólo con otras muestras de la población general española sino también con muestras de poblaciones españolas más específicas y para las cuales puede resultar útil la M-C SDS (p. ej., estudiantes universitarios, adultos evaluados en procesos de selección/ promoción laboral, adultos evaluados en contextos forenses, adolescentes).

Agradecimientos.- Esta investigación ha sido parcialmente financiada por dos ayudas del Ministerio de Educación y Ciencia (SEJ2006-02003 y PSI2009-10867).

Referencias

- Avia, M. D., Sanz, J., Sánchez, M. L., Martínez-Arias, R. M., Silva, F., y Graña, J. L. (1995). The five-factor model—II. Relations of the NEO-PI with other personality variables. *Personality and Individual Differences*, *19*, 81-97.
- Ávila Espada, A., y Tomé Rodríguez, M. C. (1989). Evaluación de la deseabilidad social y correlatos defensivos emocionales. Adaptación castellana de la Escala de Crowne y Marlowe. En A. Echevarría y D. Páez (Eds.), *Emociones: perspectivas psicosociales* (pp. 505-514). Madrid: Fundamentos.
- Ballard, R. (1992). Short forms of the Marlowe-Crowne Social Desirability Scale. *Psychological Reports*, *71*, 1155-1160.
- Barger, S. D. (2002). The Marlowe-Crowne affair: short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, *79*, 286-305.
- Beretvas, S. N., Meyers, J. L., y Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, *62*, 570-589.
- Bernstein, I. H., y Teng, G. (1989). Factoring items and factoring scales are different: spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, *105*, 467-477.
- Calvo, M. G., y Eysenck, M. W. (2000). Early vigilance and late avoidance of threat processing: repressive coping versus low/high anxiety. *Cognition & Emotion*, *14*, 763-787.
- Cano Vindel, A., Sirgo, A., y Pérez Manga, G. (1994). Cáncer y estilo represivo de afrontamiento. *Ansiedad y Estrés*, *0*, 101-112.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2ª ed. Hillsdale, NJ: LEA.
- Costa, P. T., Jr., y McCrae, R. R. (1999). *Inventario de Personalidad NEO Revisado (NEO PI R). Inventario NEO reducido de Cinco Factores (NEO FFI)*. Manual. Madrid: TEA.
- Crowne, D. P., y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349-354.
- Fernández-González, L., O'Leary, K. D., y Muñoz-Rivas, M. J. (2013). We are not joking: need for controls in reports of dating violence. *Journal of Interpersonal Violence*, *28*, 602-620.
- Ferrando, P. J., y Chico, E. (2000). Adaptación y análisis psicométrico de la escala de deseabilidad social de Marlowe y Crowne. *Psicothema*, *12*, 383-389.
- Gracia, E., García, F., y Lila, M. (2011). Police attitudes toward policing partner violence against women: Do they correspond to different psychosocial profiles? *Journal of Interpersonal Violence*, *26*, 189-207.
- Hooper, D., Coughlan, J., y Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, *6*, 53-60.
- Instituto Nacional de Estadística (2004). *Censos de población y viviendas 2001. Resultados definitivos*. Documento consultado en Internet: <http://www.ine.es>
- Lorenzo-Seva, U., y Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments and Computers*, *38*, 88-91.
- Luciano, J. V., y Algarabel, S. (2006). Individual differences in self-reported thought control: the role of the repressive coping style. *Psicothema*, *18*, 228-231.
- Miller, L. A., McIntire, S. A., y Lovler, R. L. (2011). *Foundations of psychological testing. A practical approach*, 3ª ed. Thousand Oaks, CA: Sage Publications
- Mund, M., y Mitte, K. (2012). The costs of repression: a meta-analysis on the relation between repressive coping and somatic diseases. *Health Psychology*, *31*, 640-649.
- Myers, L. B. (2010). The importance of the repressive coping style: findings from 30 years of research. *Anxiety, Stress, & Coping*, *23*, 3-17.
- Nuevo, R., Montorio, I., Márquez-González, M., Cabrera, I., Izal, M., y Pérez-Rojo, G. (2009). Diferencias asociadas a la edad en el efecto de la deseabilidad social en el autoinforme del estado emocional. *Revista Española de Geriatria y Gerontología*, *44*, 85-89.
- Prieto, G., y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, *77*, 65-75.
- Ramanaiah, N. V., y Martin, H. J. (1980). On the two-dimensional nature of the Marlowe-Crowne Social Desirability scale. *Journal of Personality Assessment*, *44*, 507-514.
- Ray, J. J. (1988). Lie scales and the elderly. *Personality and Individual Differences*, *9*, 417-418.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *38*, 119-125.
- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles del Psicólogo*, *92*, 115-128.
- Sanz, J., y García-Vera, M. P. (2009). Nuevos baremos para la adaptación española del Inventario de Personalidad NEO Revisado (NEO PI-R): fiabilidad y datos normativos en voluntarios de la población general. *Clinica y Salud*, *20*, 131-144.
- Spielberger, C. D., Gorsuch, R. L., y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., Miguel-Tobal, J. J., Casado, M. I., y Cano-Vindel, A. (2001). *Inventario de Expresión de Ira Estado-Rasgo: STAXI-2*. Madrid: TEA.
- Strahan, R., y Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *28*, 191-193.
- Tomé Rodríguez, M. C. (1985). *Estudio de adaptación de la Escala de Deseabilidad Social de Marlowe y Crowne*. Tesina de licenciatura. Facultad de Psicología. Universidad Complutense de Madrid.
- Tatman, A. W., Swogger, M. T., Love, K., y Cook, M. D. (2009). Psychometric properties of the Marlowe-Crowne Social Desirability Scale with adult male sexual offenders. *Sex Abuse*, *21*, 21-34.
- Timmerman, M. E., y Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*, 209-220.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321-327.
- Weinberger, D. A., Schwartz, G. E., y Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology*, *88*, 369-380.
- Weinberger, D. A. (1990). The construct validity of the repressive coping style. En J. L. Singer (Ed.), *Repression and dissociation: implications for personality theory, psychopathology, and health* (pp. 337-386). Chicago, IL: University of Chicago Press.
- Yela, M. (2005). Psicometría y teoría de los tests. En *L. Aniversario de la creación de la Escuela de Psicología y Psicotecnia de la Universidad de Madrid. In memoriam Mariano Yela. Publicaciones del Spanish Journal of Psychology* (CD-ROM). Madrid: Universidad Complutense de Madrid. (Orig. 1984).

(Artículo recibido: 7-11-2013; revisado: 23-4-2014; aceptado: 15-9-2014)