

UNIVERSIDAD IBEROAMERICANA

Estudios con Reconocimiento de Validez Oficial por Decreto Presidencial
del 3 de abril de 1981



**“ ESTUDIO PSICOMÉTRICO DE UNA ESCALA ETNOPSICOLÓGICA DE
DESEABILIDAD SOCIAL PARA EL CONTEXTO MEXICANO “**

TESIS

Que para obtener el grado de

DOCTOR EN INVESTIGACIÓN PSICOLÓGICA

Presenta

MIGUEL ANGEL MÉNDEZ GARCÍA

Director: DRA. ALEJANDRA DOMÍNGUEZ ESPINOSA.

Lector: DRA. CAROLINA BRAVO CONTRERAS.

Lector: DRA. GRACIELA POLANCO HERNÁNDEZ.

Índice

Introducción	4
Capítulo I. Deseabilidad Social	11
a. Deseabilidad Social: definición y evolución del constructo	11
b. Importancia de la Deseabilidad Social en la Personalidad.....	14
c. Operacionalización de la Deseabilidad Social.	17
d. Cómo se ha controlado.	20
Capítulo II Teoría Psicométrica.....	23
a. Teoría Clásica del Test (TCT).....	23
b. Teoría de Respuesta al Ítem (TRI).	25
Modelo de Rasch.....	28
Análisis del Funcionamiento Diferencial del Reactivo (DIF).....	30
Análisis Factorial Confirmatorio.....	30
Método.....	32
Justificación y Planteamiento del Problema	32
Objetivo General	33
Objetivos Específicos.....	33
Estudio 1.....	35
Participantes:	35
Instrumento.....	35
Procedimiento	36
Fase 1: Análisis Psicométrico con base en TCT.	37
Análisis de Datos	37
Resultados:.....	37
Condición Estándar	37
Condición Experimental (demanda social).....	40

Discusión	42
Fase 2: Evaluación de los Formatos Likert vs. Dicotómico.....	45
Procedimiento.....	45
Resultados	46
Discusión	47
Fase 3: Análisis Psicométrico utilizando TRI.....	48
Participantes e Instrumento.....	48
Procedimiento.....	48
Resultados	48
Discusión	53
Fase 4: Análisis de DIF	55
Participantes e Instrumento.....	55
Procedimiento.....	55
Resultados	56
Discusión	59
Estudio 2	61
Instrumento.....	61
Participantes.....	61
Discusión	67
Obtención de Baremos para la EPDDS	69
Discusión general	70
Conclusiones	78
Limitaciones	80
Sugerencias	80
Referencias.....	81
Anexo 1. Banco de Reactivos para la Escala de Deseabilidad Social.....	88

Anexo 2. EPDDS.....	92
Anexo 3. Baremos para la EPDDS.....	94

INDICE DE TABLAS

Tabla 1. Solución Factorial con los 4 primeros factores, Versión Likert, Instrucciones Estándar.....	39
Tabla 2. Valores Eigen, Varianza Explicada, versión Likert, Instrucciones Estándar	39
Tabla 3. Solución Factorial con los 4 primeros factores, Versión Likert, Instrucciones Experimentales	41
Tabla 4. Valores Eigen, Varianza Explicada, versión Likert, Instrucciones Estándar	41
Tabla 5. Puntajes promedios de los dos formatos de la EPDDS	46
Tabla 6. Índices de confiabilidad para cada formato de la EPDDS.....	46
Tabla 7. Modelo de efectos fijos entre condiciones y formato de la EPDDS.....	47
Tabla 8. Valores de los parámetros del modelo Rasch para la EPDDS.....	49
Tabla 9. Valores de los parámetros del modelo Rasch para la dimensión 1 (Honradez)...	50
Tabla 10. Valores de los parámetros del modelo Rasch para la dimensión 2 (Respeto) ...	51
Tabla 11. Valores de los parámetros del modelo Rasch para la dimensión 3 (Humildad). 51	
Tabla 12. Valores de los parámetros del modelo Rasch para la dimensión 4 (Sagacidad)51	
Tabla 13. Valores de los parámetros del modelo Rasch para la dimensión MI (1, 2 y 4)...	52
Tabla 14. Valores Eigen de los residuales del Modelo Rasch en las 2 dimensiones.....	53
Tabla 15. Valores de DIF a partir de la Condición.	57
Tabla 16. Valores de DIF a partir del Sexo.	58
Tabla 17. Valores de DIF a partir de la Edad.....	58
Tabla 18. Correlaciones entre los factores del AFC.	63
Tabla 19. Coeficientes Estandarizados para el Factor 1.	64
Tabla 20. Coeficientes Estandarizados para el Factor 2.	64
Tabla 21. Coeficientes Estandarizados para el Factor 3.	64
Tabla 22. Coeficientes Estandarizados para el Factor 4.	65

INDICE DE FIGURAS

<i>Figura 1. Valores de las Dimensiones de Hofstede para Canadá, EU, España y México. ...</i>	8
<i>Figura 2. Coeficientes estandarizados para la EPDDS.....</i>	62
<i>Figura 3. Coeficientes estandarizados para la dimensión MI.....</i>	66
<i>Figura 4. Coeficientes estandarizados para la dimensión AE.....</i>	67

Introducción

El desarrollo de la Psicología como ciencia ha sido posible gracias a los esfuerzos muy diversos de los diferentes teóricos para pasar de un plano abstracto a un plano medible, los diferentes constructos de interés para la disciplina. Desde sus inicios, se han desarrollado diferentes herramientas para dichos propósitos, dentro de las que destacan las escalas, pruebas y cuestionarios psicológicos. Estas herramientas tienen como propósito reflejar el grado o la cantidad de atributo que un individuo posee. Básicamente, la psicología ha desarrollado cantidades innumerables de instrumentos de medición psicológica, desde las primeras pruebas objetivas hasta las baterías de medición actuales, abarcando la medición de los rasgos de personalidad, habilidades, inteligencia, actitudes, etc.

La creación de las pruebas objetivas ha permitido en gran medida la generalización de las teorías a través de diversos escenarios, haciendo que actualmente las pruebas psicológicas estructuradas sean de las herramientas más utilizadas por los psicólogos de todas las corrientes teóricas. Si bien es cierto que las escalas y pruebas psicológicas nos permiten estimar aquello que nos interesa, las mismas no están libres de error como cualquier otro instrumento de medida y esto se debe fundamentalmente a que siempre va a existir un porcentaje de error pudiendo ser grande o pequeño, atribuido al sujeto, al instrumento y a las condiciones de aplicación (ver Nunnally & Bernstein, 1995; Thorndike, 1951).

El primer tipo de error se refiere principalmente a los errores de tipo aleatorios inducidos por las características y estados anímicos del propio individuo que pueden intervenir en su estimación de los propios rasgos, considerándose el estado de ánimo, actitudes hacia la prueba, excitabilidad o ansiedad causada por la prueba, etc. El segundo tipo de error es directamente vinculado a las propias

características del instrumento tales como carecer de confiabilidad, poder discriminativo y validez. El tercer tipo es más vinculado a las especificaciones de aplicación (errores de instrumentación), calificación e interpretación de las puntuaciones (Nunnally & Bernstein, 1995). Cada uno de los tipos de errores mencionados es importante; sin embargo, el que más preocupa a los estudiosos del comportamiento es el segundo, debido al cual existe una subespecialidad dentro de la psicología llamada psicometría.

Una prueba o una escala psicológica que supuestamente mide una variable, puede también estar midiendo algún otro atributo, el cual puede *interferir* o contaminar el puntaje obtenido en esa medición. Esta noción fundamental fue inicialmente desarrollada por Cronbach (1946, p. 476) quien acuñó el término de sesgo de respuesta y lo definió como “*cualquier tendencia que causa que el sujeto consistentemente de una respuesta diferente de la que daría a un mismo contenido si las preguntas se hubieran presentado de diferente manera*”. Así, muchos de los puntajes de las pruebas pueden estar influenciados por otras variables que inintencionadamente se miden, dando lugar a que las personas con el mismo nivel de conocimiento, actitud o rasgo reciban diferentes calificaciones, y por lo mismo ocasionan que disminuya la confiabilidad y validez de una prueba.

Principalmente los denominados autorreportes (aquellos instrumentos en donde el individuo describe sus propios rasgos), son influenciados en gran medida por estas variables contingentes de las que hablaba Cronbach, debido a que las mediciones están supeditadas a confiar en que las respuestas que dan los participantes son genuinas y lo más exactas posibles, cosa que no es posible asegurar del todo. Así, potencialmente, todas las mediciones de los rasgos humanos están afectadas por este tipo de variables contingentes, tales como límites de tiempo, adivinación, *estilos de respuesta*, fatiga, motivación, etc. Estos fenómenos pueden (1) reducir la confiabilidad, (2) introducir fuentes confiables de

diferencias individuales que reducen la validez, o bien (3) producir correlaciones substanciales entre pruebas o diferencias entre las medias de los grupos experimentales, las cuales en realidad son falsas (Megargee, 1971; Rosenthal & Rosnow, 1991). Por esta razón la estandarización de las medidas es necesaria para así minimizar el efecto de las distorsiones ya sean intencionadas o inconscientes (Nunnally & Bernstein, 1995).

Más específicamente se considera un *estilo de respuesta* o sesgo a un efecto importante sobre la medición que es independiente de lo que se busca medir. Dentro de esta categoría se han identificado tres fenómenos: Aquiescencia (tendencia consistente de acuerdo más que desacuerdo en una comunicación, que es independientemente de su contenido), Respuestas Extremas (los sujetos contestan solamente en los extremos de un continuo psicológico) y Deseabilidad Social (DS).

Éste último concepto, DS, es un fenómeno que se presenta comúnmente en las evaluaciones, y es aquella tendencia donde la persona contesta un instrumento o una entrevista dando la imagen más favorable de sí para obtener un beneficio. Esto puede ser un problema ya que en ocasiones esta imagen modificada, puede afectar la medición que se realice y de esta forma las decisiones que se tomen con base en estas pruebas no sean las más idóneas; por ejemplo, en la educación, las expectativas que pueden tener los profesores sobre sus alumnos, lleva a distorsiones que solo ajustan el estilo de respuesta y actuación de estos por tratar de mostrar una imagen favorable al profesor (Dompnier, Pansu, & Bressoux, 2007). En otro campo como lo es la mercadotecnia este fenómeno es una preocupación frecuentemente ya que las personas constantemente responden tergiversadamente y simplemente proveen respuestas que los hacen lucir bien (Steenkamp, De Jong, & Baumgartner, 2010). Es de esta forma que este

sesgo o distorsión, es un elemento que se debe considerar en cualquier medición que involucre a la persona.

La DS ha sido estudiada en el contexto anglosajón, y se han construido diversas escalas con la finalidad de estimarla y por consiguiente controlarla (p. ej. Crowne & Marlowe, 1960; Eysenck & Eysenck, 1964; Paulhus, 1998b; Stöber, 2001) algunas de las cuales han sido traducidas y en menor frecuencia adaptadas a nuestro idioma (Castro, Maya, & Orozco, 1986; Collazo, 2005; Domínguez, Salas, & Reyes- Lagunes, 2008; Ferrando & Chico, 2000; Lara-Cantú, 1990). Se ha señalado en diversos estudios que tanto el contexto y la época en que fueron desarrolladas las pruebas ocasiona que las mismas pueden ya no ser adecuadas para los tiempos actuales o para el ambiente en que nosotros estamos inmersos.

Actualmente nuestro país se encuentra en el ranking mundial de transparencia en el lugar número 98¹ con 3.2 puntos (Transparency International, 2010), sobrepasado, por ejemplo por Chile (lugar 21 con 7.2 puntos), Uruguay (lugar 24 con 6.9 puntos) España, (lugar 30 con 6.1 puntos), Costa Rica (lugar 41 con 5.3 puntos), Cuba (lugar con 3.7 puntos), Colombia (lugar 78 con 3.5), por mencionar solo algunos países de habla española. Refiriéndonos más explícitamente a aquellos países en los cuales se han desarrollado las escalas de DS como son Canadá y Estados Unidos, encontramos que éstos ostentan el lugar 6 (8.9puntos) y 22 (6.9 puntos) respectivamente. ¿Qué implica esto? Los índices anteriores nos representan qué tanto a nivel institución permitimos la mentira, la falsedad y la distorsión adentro e nuestras instituciones y el hecho de nosotros vivir de manera *diferente* a otros países esta distorsión trae consecuencias a nivel de la DS y su consecuente concepción, medición y calificación.

¹ El índice define corrupción como el abuso de los oficiales públicos para ganancia privada y mide

Nuestro país, a diferencia de Canadá, España o Estados Unidos (países que nos exportan las pruebas psicológicas con mayor frecuencia, incluyendo las dos más utilizadas de DS), están diferenciados de nosotros, a nivel cultural en muy diversos aspectos que directamente nos afectan nuestra cosmovisión del mundo y consecuentemente el desarrollo de la personalidad prototípica. Se puede observar en la Figura 1 que México como cultura, tiene rasgos claramente diferenciados con respecto a los otros tres países representados. Se aprecian diferentes niveles de Distancia de Poder (PD), Individualismo (I), Masculinidad (M) y Temor a la Adversidad (TA) (Hofstede, 1980).

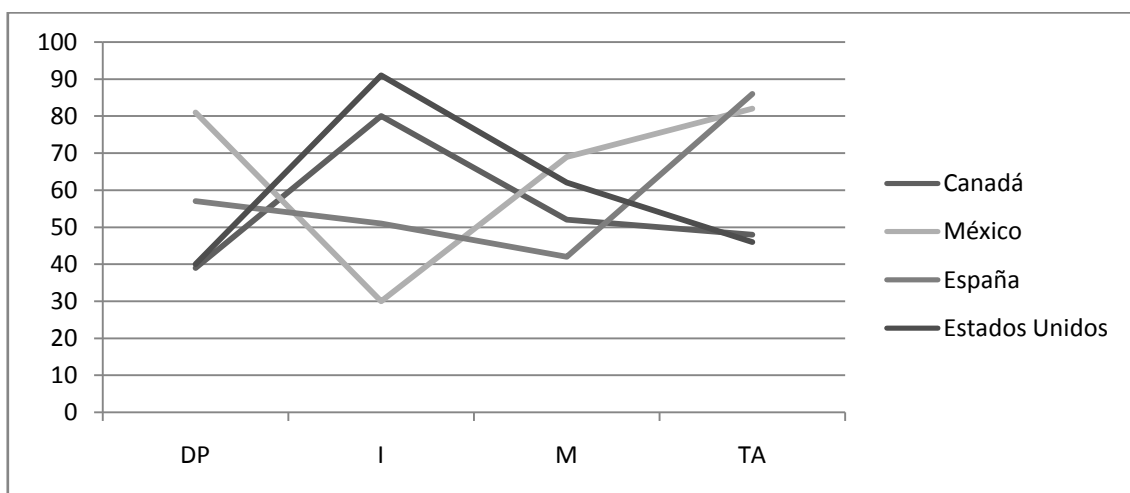


Figura 1. Valores de las Dimensiones de Hofstede para Canadá, EU, España y México.

Observamos que por ejemplo, la DP es más grande para México y por lo mismo el deseo por quedar bien con autoridades es mayor en nuestro contexto. En la dimensión de I en donde el énfasis lo tiene el individuo sobre la colectividad, observamos que México se encuentra muy por debajo de EU y Canadá lo que implica que en México le damos más importancia al grupo, lo que supondría una respuesta más llamativa de necesidad de aprobación por parte de los demás. La masculinidad, relacionada con la distribución de roles, indica que los mismos están todavía más diferenciados en México que en los países y esto nos sugiere que los roles asignados también están en función de la deseabilidad de los mismos. TA tiene que ver con la falta de reglas, leyes, normas estrictas de comportamiento y

observamos que España y México están más altos que los otros dos países, lo que implica que tenemos menos estructura y rigidez en el seguimiento de las leyes, lo cual no nos resulta sorprendente que esto también afecte la DS. Con base en todo lo anterior, podemos asumir que la medición de la DS es un aspecto importante para el área de la medición y que su conceptualización y posible medición están íntimamente relacionadas con la cultura.

Ya se han realizado algunos estudios en nuestro país que abarcan esta temática (Castro, et al., 1986; Domínguez-Espinosa, 1997; Domínguez-Espinosa, Navarro Contreras, García Campos, Gutiérrez Fierros, & Sada Monroy, 2010; Domínguez-Espinosa, Salas-Menotti, & Reyes-Lagunes, 2008; Domínguez-Espinosa, Salas Menotti, & Contreras Bravo, 2010; Lara-Cantú, 1990; Lara-Cantú & Suzan-Reed, 1988), sin embargo, a pesar de su importancia, todavía es un tema tomado con poca seriedad. Domínguez-Espinosa (2008; Domínguez-Espinosa, et al., 2010) y colaboradores han identificado algunos elementos definitorios para la operacionalización de la DS y elaborado un banco de reactivos amplio para la elaboración de una escala que midiera este constructo. Sin embargo, la propuesta no había atravesado por los procedimientos psicométricos pertinentes para verificar sus propiedades, por lo que el presente trabajo tuvo como principal objetivo evaluar el banco de reactivos de DS y someterlo a diferentes estudios y experimentos para dar una propuesta para la constitución de la Escala Domínguez de Deseabilidad Social (EDDS).

La presente tesis pone a prueba varios supuestos psicométricos como son la manipulación experimental de instrucciones, la utilización de los modelos de medición denominados TCT y TRI, la detección de DIF, confirmación factorial y la obtención de baremos. En el capítulo 1 se abordará más específicamente la teoría del constructo de DS y sus correlatos con otras variables. En el capítulo 2 se abordarán datos de los modelos psicométricos utilizados en el presente trabajo.

La presente tesis formó parte de un proyecto mayor con financiamiento de la líder del proyecto Dra. Alejandra Domínguez Espinosa, por parte de la Dirección de Investigación de la Universidad Iberoamericana y del Consejo Nacional de Ciencia y Tecnología, # 82988.

Capítulo I. Deseabilidad Social

a. Deseabilidad Social: definición y evolución del constructo.

Uno de los objetivos principales de los psicómetras ha sido el de proponer y establecer formas de medición, objetivas y adecuadas para explicar los fenómenos de su interés. El uso de las pruebas estandarizadas data del año 1920, y a la fecha su uso es extendido a lo largo del planeta y en muy diversos escenarios.

Cronbach (1946) menciona que estas pruebas son validadas por el contenido de cada reactivo, además de presentarse factores que modifican los puntajes finales, como las condiciones de aplicación y construcción que afectan a la medida en lo que se denomina *estilos de respuesta*. Éstos últimos se refieren a formas en las que una persona contesta de manera consistente una prueba, pudiendo clasificarse en cinco tipos: *a) tendencia a contestar precaución contra riesgo*, la persona responde lo que está completamente seguro; *b) categorías de juicio*, donde las diferencias de personalidad, o comprensión del texto puede dar como resultado un conservadurismo de las respuestas; *c) inclusividad* que se presenta en preguntas abiertas donde la persona puede realizar una lista tan larga que impida realizar una medida objetiva; *d) aquiescencia* esta se presenta cuando una persona contesta a los reactivos sin distinguir contenidos; *e) velocidad contra exactitud* que puede a la vez interferir con respuestas certeras y *f) respuestas debido a pruebas de ensayo* en este caso el realizar ejercicios similares llevará a un aprendizaje de respuestas. Estos estilos de respuesta ocasionan sesgos en los puntajes de las pruebas, estas desviaciones deben ser consideradas o inclusive ser corregidas, ya que darían puntajes inexactos de las respuestas y por lo tanto es necesario identificar la magnitud del efecto.

La Deseabilidad Social (DS) ha sido considerada como un estilo de respuesta (Morf & Jackson, 1972; Paulhus, 2002), además de presentar un componente de rasgo de personalidad (Kuncel & Tellegen, 2009; McCrae & Costa, 1983). La DS define Fordyce (1956), es el consenso de juicios acerca de que sentimientos, actitudes o comportamientos que ganan aprobación social en la sociedad. Edwards (1957) propone su definición como la tendencia de los sujetos que responden los reactivos de una prueba en dirección de presentarse a sí mismos en términos socialmente aceptables para ganar la aprobación de otro.

Marlowe y Crowne (1960) definen que la DS son conductas culturalmente aprobadas o sancionadas pero con baja probabilidad de ocurrencia, posteriormente al tener la escala validada definieron la DS como la necesidad de obtener aprobación respondiendo de una manera aceptable y apropiada culturalmente y consideraban al constructo como unidimensional.

Damarin y Messick en 1965 efectúan una interpretación teórica del constructo y establecen que la DS, presenta dos factores, un factor asociado a la “respuesta autista del autoestima” que tendría que ver con una defensa de la imagen del self, y otra que sería “Respuesta propagandística”, cuyo motor sería el promover una reputación pública deseable (Paulhus, 2002).

Sackeim & Gur (1979) determinan que la DS está asociada a dos constructos: “el autoengaño” donde la respuesta a reactivos con contenido ofensivo estarían asociados a esta dimensión, mientras que “el engaño a otros” estaría definido por reactivos con conductas deseables públicas. Paulhus, denomina Respuesta Socialmente Deseable (RSD) como la tendencia a dar una autodescripción positiva, para hacer una integración de las ideas de Damarin, Messick, Sackeim y

Gur (Paulhus, 2002), para proponer que la DS está en función de una respuesta solicitada a un autoreporte.

Paulhus (1984) indica que la DS tiene dos componentes, a partir de que el *self*, aceptará o negará esas condiciones deseables o indeseables, una de manera consciente y otra no tan consciente. Inicialmente el autor propone denominar a la primera como *Manejo de Impresión(MI)* se refiere a las estrategias conscientes dirigidas a elaborar una impresión positiva en los otros, mientras que a la segunda como *Auto-Engaño(AE)*, que se refiere a la autopromoción inconsciente y narcisista (Robins & Paulhus, 2001).

Conforme se fue desarrollando más investigación al respecto. Paulhus y su grupo de investigación sugieren que el AE agrupa a su vez dos dimensiones, que son una *positiva (Enhancement)* que promueve cualidades aceptadas o de realce, mientras que el otro es de *negación (Denial)* de cualidades desaprobadas (Paulhus & Reid, 1991). Asimismo se ha llegado a establecer que la DS tiene un componente en relación con las condiciones de aplicación (Paulhus, 2002), de tal manera que se presentan dos factores con dos dimensiones cada uno, el primero denominado *Respuesta Egoísta (Egoistic Bias)* donde se presenta una exageración a mostrar el estatus social e intelectual, formado por el *AE positivo* y la *atribución de cualidades*, que representarían características de superhéroe e inclusive pueden llegar al narcisismo. En el segundo factor se denomina *Respuesta Moralista (Moralistic Bias)*, formada por el *AE negativo* y el *manejo de comunión*, estas explicarían atributos de agradabilidad e inclusive de mostrarse “santo” o “como si lo fuera”.

Lönnqvist, Paunonen, Tuulio-Henriksson, Lönnqvist, y Verkasalo (2007) definen a la DS como la tendencia a dar una autodescripción demasiado positiva, que

muestra dos componentes uno asociado al MI para parecer “santo” y uno autodescripción exagerada que tendría que ver con la tendencia a mostrarse como “superhéroe”.

DS se refiere a la capacidad individual para satisfacer las motivaciones de los miembros de un grupo social determinado, y por lo tanto el grado en que se es aceptado (Darnon, Dompnier, Delmas, Pulfrey, & Butera, 2009). La DS definida por Kuncel & Tellegen (2009) establece que es la respuesta de comportamiento consistente con lo que se percibe así como aquello que se desee por otros.

b. Importancia de la Deseabilidad Social en la Personalidad.

La Personalidad es definida como la organización sistemática en el interior del individuo de los sistemas psicofísicos que determinan su conducta y su pensamiento característicos (Allport, 1965). La personalidad no es algo acabado, es el resultado de la interacción de una serie de elementos que le permiten funcionar al individuo de forma óptima de acuerdo al medio en que se desarrolla.

Todo hombre comparte ciertas características: igualdad entre sí que son características generales, diferencias entre los comunes debidas en una parte a la sociedad así como su entorno y diferencias individuales debidas a nuestra naturaleza (Kluchhohn, Murray, & Schneider, 1953).

Más específicamente, Edwards (1964) estudio los efectos de la DS en el MMPI, para lo cual diseñó las escalas de corrección, donde personas que presentarían DS, son aquellas que contesten a los reactivos asignados con alta deseabilidad social como verdadero, mientras que contestarían falso a aquellos con baja deseabilidad social, contestar de una manera u otra en forma que los puntajes sean elevados mostrarían que la persona muestra una alta defensividad, o disimulo.

Eysenck en su modelo de personalidad establece dos dimensiones importantes, Extraversión-Introversión y Neuroticismo-control (Sánchez-Turet & Cuadras-Avellana, 1972), estas dimensiones aunadas a una escala de mentira, al hacer el estudio de la DS y una manipulación de las instrucciones de la prueba muestran que en la demanda social de exagerar la propia percepción “manipulación positiva”, se presentan valores altos de extraversión, mientras que el solicitar que hagan una descripción desfavorable de sí mismo “manipulación negativa” muestra valores altos en el puntaje de Neuroticismo (Dunnett, Koun, & Barber, 1981), a los valores altos de extraversión se considera DS, mientras que los valores altos de Neuroticismo como socialmente indeseables. La extraversión (E) se asocia como características el ser sociable, correspondiente al *estado*, mientras que la introversión está descrita como personas pasivas y contenidas que referirían el *rasgo*.

Catell y Mead (2008) en su modelo de los 16 factores de personalidad de la que se desprenden 5 escalas globales que serían el antecedente de la escala de los cinco grandes factores (Extraversión, Ansiedad/Neuroticismo, Mente abierta, Independencia y Autocontrol), identifico que el MI se presenta como resultado de la DS.

Paulhus, Fridhandler & Hayes (1997), han establecido que los mecanismos de defensa (usualmente definidos como procesos mentales que operan inconscientemente para reducir el dolor de una emoción, o con resultados de mala adaptación), no solamente evitan el dolor, sino como mecanismos adaptativos que pueden servir al self y en este sentido extienden la idea de que la negación de atributos negativos o el AE podrían ser mecanismos adaptativos de la especie humana (Robins & Paulhus, 2001).

La DS se presenta, no porque la persona quiera mentir en el mayor de los casos, si no que está empapado por todas las características de la persona misma que incluye desde factores sociales, personales, culturales, es por ello que es importante que dentro de los estudios de personalidad se considere como un elemento que afecta a los autoreportes y que tendría que ser escalado para así determinar si los resultados del mismo son confiables, por ejemplo, se ha observado que la distorsión por DS puede introducir hasta 38 unidades de desviación en contextos de selección (Salgado, 2005).

En los estudios que se han realizado al respecto muestran que la DS no es un problema en sí mismo, si no que dependiendo de la situación en la que se efectúe el estudio o la aplicación de un instrumento, la persona responderá con cierto grado de defensividad o apertura (Pauls, Wacker, & Crost, 2005). Es por ello que la DS se presenta de manera natural en el individuo como una necesidad de exaltar sus propias características que lo lleven a obtener un resultado esperado. Aunque por otro lado, algunos autores consideran que su efecto no es importante en la medición de la personalidad ya que se “espera que las personas se ajusten a una demanda situacional específica” (McCrae & Costa, 1983).

Es así que los efectos de la DS se han estudiado de forma paralela con algunas pruebas de personalidad, validadas y confiables, por ejemplo en el caso del inventario de Eysenck, se observan correlaciones estadísticamente significativas entre extraversión/neuroticismo y la extraversión/escala de mentiras, una persona con puntajes altos en Extroversión habla de sujetos con alta deseabilidad social, mientras que personas que muestran altos niveles de Neuroticismo son sujetos que presentan conductas socialmente indeseables (Dunnett, et al., 1981; Meleddu & Guicciardi, 1998).

En situaciones de demanda la persona que no presenta psicopatología responde generalmente mostrando un ajuste a las normas y reglas (Downey, Godfrey, Hansen, & Stough, 2006; Hopwood, Talbert, Morey, & Rogers, 2008; Meleddu & Guicciardi, 1998; Robson, Jones, & Abraham, 2008), que sería un rasgo adaptativo referente a la Agradabilidad, situación contraria al Narcisismo, simulación donde se muestra una relación inversa (Hopwood, et al., 2008; Rogers, et al., 2002).

Es de esta forma que la DS se ha analizado desde el punto de vista del estado y del rasgo de personalidad; en el primero los puntajes de DS muestran correlaciones estadísticamente significativas en Neuroticismo, mientras que en el segundo las correlaciones se presentan con la Agradabilidad (Hofstee & Hendriks, 1997; Lönnqvist, et al., 2007; Paulhus, 1998a). Por otro lado la persona cuando da una descripción de sí mismo, debe contar con un autoconocimiento que le permita saber quién es, lo que desea, qué quiere hacer o ser, Meleddu y Guicciardi (1998) denominan a esta como *amor propio*, necesidad vital del ser humano relacionado con el adecuado manejo de la conciencia.

c. Operacionalización de la Deseabilidad Social.

Edwards desarrollo una escala que pudiera determinar si un sujeto estaba mintiendo al aplicársele la prueba del MMPI (Meehl & Hathaway, 1946). Basándose en los reactivos de las escalas F, L, K del MMPI y de la Escala de Ansiedad Manifiesta de Taylor, desarrolló la primera escala conocida de DS. Lo que Edward obtuvo fue un sistema de equivalencia del grado de DS que presentaba cada uno de los reactivos de las pruebas de personalidad y los apareo con base en la probabilidad de dar una Respuesta Socialmente Deseable (RSD).

Finalmente obtuvo una escala en donde cada par de opciones de respuesta estaba equiparado en DS. Esta escala, sin embargo y en opinión de Crowne y Marlowe (1960) al haberse desarrollado con base en reactivos de personalidad, media no solamente DS sino personalidad patológica, por lo que ellos mismos desarrollaron un segundo modelo de medición para la DS que abordaba su definición desde una perspectiva más contextual. Consideraron entonces elaborar reactivos que reflejaran conductas aprobadas socialmente pero infrecuentes y conductas desaprobadas socialmente pero muy frecuentes. Su banco de reactivos lo sometieron a un juicio por parte de alumnos y profesores de la Universidad de Ohio y es actualmente una de las propuestas métricas más utilizadas, a pesar de haberse desarrollado en la década de los 60's.

Originalmente su formato es dicotómico y se han realizado infinidad de estudios con ella. Desde muy temprano su publicación la escala generó mucha investigación, ya que las expectativas por contar con una escala que pudiera controlar la DS era muy esperada, sin embargo, con el transcurso de los años se han encontrado algunas dificultades, tanto de formato como de contenido (Greenwald & Clausen, 1970; Ones, Viswesvaran, & Reiss, 1996; Stöber, 2001; Uziel, 2010; Wiggins, 1959; Ziegler, 2009).

Paulhus preocupado por las ambigüedades teóricas y metodológicas de la DS, decidió llevar a cabo un estudio para evaluar su dimensionalidad utilizando las seis escalas, que desde su perspectiva eran las más importantes para evaluar la DS: La escala de DS de Edwards (1957) con 39 reactivos; la escala L del MMPI (Meehl & Hathaway, 1946) con 15 reactivos; la escala de DS de Wiggins (1959) con 40 reactivos; la escala de DS de Marlowe y Crowne (Crowne & Marlowe, 1960) con 33 reactivos; y los cuestionarios de Autoengaño y Engaño a Otros (Sackeim & Gur, 1979) cada uno con 20 reactivos. A partir de la aplicación de las escalas a una muestra total de 611 estudiantes de la Universidad de Vancouver, Canadá,

realizó un procedimiento de reducción de datos y obtuvo que del total de reactivos, 40 se distribuyeron en dos dimensiones, las cuáles denominó Autoengaño y Manejo de la Impresión, constituyéndose así el Inventario Balanceado de Respuestas Deseables (BIRD por sus siglas en inglés). Actualmente este inventario es la segunda herramienta más utilizada a nivel mundial para medir la DS.

En México Domínguez-Espinosa (1997) desarrollo una escala de DS en población mexicana, formada por 18 reactivos y en estudios posteriores se demostró su validez convergente (Domínguez-Espinosa , Salas-Monotti, & Reyes-Lagunes, 2008) con la escala Marlowe-Crowne. En Alemania Stöber (2001) desarrollo una escala de 17 reactivos, que presentó valores de confiabilidad y validez consistentes, hallando una dimensión relacionada con el MI de Paulhus, pero no así la escala de AE.

Es así que las escalas de DS han tomado una gran importancia en el contexto de la medición, por ello se han adaptado a otros contextos culturales y sociales diferentes para las que fueron diseñadas las escalas principales (Cosentino & Castro-Solano, 2008; Ferrando & Chico, 2000; Lück & Timaeus, 1969), a pesar de encontrar valores aceptables de confiabilidad y validez en estas adaptaciones, existe la necesidad de hacer instrumentos para cada población debido a las diferentes normas culturales, sociales, individuales, que ayudarán a tener mejores porcentajes de explicación de varianza acerca del impacto que existe en la DS tanto a nivel individual como a nivel grupal y de grupo cultural.

d. Cómo se ha controlado.

Las escalas de DS generalmente no se aplican como un instrumento individual, sino que son herramientas concurrentes a otras pruebas para determinar el grado de distorsión que por este fenómeno pueda invalidar a la prueba principal.

Cronbach (1946) propone que las distorsiones debidas al estilo de respuesta que se presentan cuando se contesta una prueba escrita afectan a la confiabilidad y validez empírica, y es por ello que el diseño del instrumento es muy importante para establecer de qué manera se evaluarán las respuestas de interés. Cuando se somete a un individuo a una situación de demanda social, éste variará su respuesta y es por eso que se sugiere que dichas aplicaciones vayan acompañadas de otras técnicas de recolección de datos como son la observación y entrevista, que ayudarían a darle una mejor interpretación a los resultados (Azrin, Holz, & Goldiamond, 1961). Arnold y Feldman (1981) proponen que la relación entre la medición y la DS, constituye una evidencia de la evocación de la demanda de la respuesta, y que más que corregirla se debe analizar el origen de esta desviación.

El control de la DS se ha analizado desde tres puntos de vista: *el modelo de observaciones espurias, modelo de supresión y el modelo moderador* (Ganster, Hennesey, & Luthans, 1983). El primer modelo, *de las observaciones espuria* indica que la DS provoca que se presenten asociaciones significativas entre la variable independiente y dependiente de interés sin que sean reales. El segundo modelo de *supresión*, afirma por el contrario que la DS inhibe correlaciones existentes. El tercer modelo, *de moderación*, indica que la DS solamente reduce el tamaño del efecto (ya sea de una correlación, de una diferencia, de una predicción, etc.) y que por lo mismo es posible que atenúe los resultados y de igual manera afectar en la interpretaciones de los hallazgos. Algunos

investigadores efectivamente han logrado observar este fenómeno cuando se ha presentado en asociaciones con otras escalas de personalidad, por ejemplo el NEO-FFi y han observado que en particular hay escalas más susceptibles a la distorsión que otras (p.ej. Borkenau & Ostendorf, 1992).

En una amplia gama de estudios, lo que se ha desarrollado es la experimentación con las instrucciones de las pruebas, para ver de qué manera se puede disminuir o aumentar el efecto de DS sobre los puntajes (Ellingson, Sackett, & Hough, 1999; Holden, 2007, 2008; Mueller-Hanson, Heggstad, & Thornton III, 2003; Robson, et al., 2008; Rogers, et al., 2002). La *experimentación* se establece de dos formas: la primera consiste en pedirle al sujeto que muestre una imagen favorable de sí mismo, y se conoce como “manipulación positiva”. La segunda implica solicitarle al individuo que se muestre con un detrimento de sí mismo o si tuviera problemas psicológicos o de conducta, y se conoce como “manipulación negativa” (Holden, 2008). Estas manipulaciones experimentales a su vez han sido reportadas como un indicador de consistencia interna ya que al solicitar que se exagere de manera deliberada las respuestas, los puntajes tienden hacia una dirección que muestra el nivel en el que la prueba puede distorsionarse y probar que tanto es posible que la prueba se mantenga estable ante dicha distorsión (Holden, 2007). La manipulación de instrucciones en la construcción y validación de escalas, permite prever que reactivos pueden ser susceptibles de distorsión para tomar decisiones de inclusión o eliminación dentro de un instrumento (Ellingson, et al., 1999; Méndez-García & Domínguez-Espinosa, 2010; Steenkamp, et al., 2010) ayudando así a corregir las medidas de los instrumentos de personalidad.

Salgado (2005) sugiere algunas estrategias para reducir los efectos de la DS en el uso de autoreportes como por ejemplo el empleo de escalas de elección forzosa equiparadas en DS, la aplicación concurrente de escalas de DS, ajuste de los puntajes eliminando puntuaciones extremas, e informar al evaluado de las

consecuencias de presentar una distorsión. Estas estrategias no son infalibles y por lo mismo se sugiere ocupar más de una. Existen actualmente modelos más complejos de corrección con base en aleatorización de preguntas y correcciones matemáticas alternas; sin embargo es todavía una alternativa poco estudiada para sacar conclusiones contundentes (p. ej. De Jong, Pieters, & Fox, 2010).

Adicionalmente, la aplicación de instrumentos por medios electrónicos ha demostrado reducir los efectos de la DS (Moon, 1998) debido a la distancia entre el aplicador y el aplicado, incluyendo también el anonimato. Adicionalmente la DS ha sido estudiada también en la administración vía Internet, observando que se evoca menos DS al estar atrás de una computadora (Risko, Quilty, & Oakman, 2006).

Finalmente, se puede deducir que la DS ha sido un tema de interés para el área de medición y para el área de la personalidad. Es concebida como una fuente de distorsión, basada en un rasgo característico de la persona y va a ser *disparada en situaciones* de demanda social. Es un constructo vinculado explícitamente a la cultura y a otros atributos propios del individuo, por lo que las adaptaciones y construcción de escalas y baremos para cada sociedad son necesarias, para la comprensión de la conceptualización de la DS.

Capítulo II Teoría Psicométrica.

El objetivo del presente capítulo es dar una breve introducción a los dos modelos de medición vigentes en Psicología y Educación, ya que los análisis de la presente tesis retoman cada uno para los estudios de validación psicométrica de la escala de DS etno-psicológicamente desarrollada. Asimismo se mencionan generalidades del DIF y Análisis Factorial Confirmatorio para un mejor entendimiento de los hallazgos empíricos.

a. Teoría Clásica del Test (TCT).

Uno de los objetivos de las ciencias es explicar y predecir los fenómenos que ocurren a su alrededor, las ciencias naturales a lo largo de su desarrollo han definido que medir es comparar una magnitud con otra tomando una como referencia o patrón (Pérez, 2004). Las fuentes de medición en Psicología generalmente provienen del llamado mundo psicofísico y de pruebas mentales o tests (Reidl-Martínez, 2007), estas pruebas son la herramienta más usada por los psicómetras, y siguen ciertas reglas que aseguran que su construcción es la adecuada, así como la conceptualización del atributo dando como resultado validez y confiabilidad de la medida.

En toda medición el error es parte inherente al proceso, aunado a la limitación de repetir la medida, una prueba debe acercarse a la puntuación verdadera del constructo que se consigue por modelos formales de análisis de los puntajes obtenidos ante una prueba, la Teoría Clásica del Test (TCT) *“es una forma que describe la influencia de los errores de medida en las puntuaciones observadas y sus relaciones con las puntuaciones verdaderas”* (Martínez, 1996, p. 58).

Estos errores fueron modelados desde el siglo XVIII por Laplace, Gauss y Legendre que tomaron como punto de referencia la teoría de mínimos cuadrados, que considera que el error varía en función a un valor verdadero, de acuerdo a la distribución normal. El modelo de TCT se basa en el concepto de considerar que la puntuación observada está formada por dos partes una denominada puntuación verdadera y por un error de medida, del cual se deducen los siguientes postulados (Martínez, 1996):

- a) La suma de los errores y las puntuaciones es una función lineal aditiva.
- b) El valor esperado de la variable aleatoria conocida como error de medida es igual a cero.
- c) La puntuación verdadera y el error para una población de examinados no están correlacionados.
- d) La puntuación error para un test 1 y la puntuación para un test 2, no están correlacionados, que no se cumple muchas veces en la práctica debido a factores como fatiga, práctica, cambios de humor entre otros que lleva a la necesidad de establecer condiciones homogéneas en las aplicaciones.
- e) Los errores de medida de un test no están correlacionados con las puntuaciones verdaderas de otro test.

La TCT ha sido utilizada para la construcción de la mayoría de los test psicológicos vigentes, aunque esto no quiere decir que no hubiese algunos problemas y limitaciones de este enfoque. Muñiz (1998) y Martínez (1996) han señalado que algunas de las limitantes más importantes se basan en la ausencia de invarianza de las mediciones respecto al instrumento utilizado así como también en la dependencia del instrumento de la muestra utilizada para su construcción. Enumerándolas se encuentra que:

1. Los estadísticos del ítem dependen de la tendencia central y de la variabilidad del rasgo en la muestra y en la población.
2. Dependencia de los estadísticos del test de la población y/o muestra en que se calculen.
3. Dificultades prácticas con el concepto de paralelismo de las medidas

4. La puntuación empírica total, base de la mayor parte de las inferencias de la TCT, depende de los ítems concretos que componen el test.
5. La TCT no proporciona un modelo teórico para las respuestas a los ítems.
6. Los tests de la TCT no producen medidas apropiadas ni adecuadas a los sujetos.
7. No proporciona soluciones válidas en la práctica real del test en situaciones como: diseño, tests adaptados a los sujetos, tests computarizados, análisis de los sesgos, equiparación de puntuaciones de tests obtenidas bajo distintas formas.

b. Teoría de Respuesta al Ítem (TRI).

Para solucionar algunas de las limitaciones antes mencionadas han surgido otros modelos de medición como la Teoría de Respuesta al Ítem (TRI), que se basa en la propuesta de Lazarsfeld quien describió el modelo de la ojiva normal de dos parámetros y procedimientos de estimación, demostrando que algunos constructos de la TCT se podían modelar como funciones de la Curva Característica del Ítem (Martínez, 1996). Estos surgen como respuesta de fundamentación probabilística a la medición de constructos latentes tomando como referencia al reactivo de acuerdo al patrón de respuesta que se presenta (Coartada de Kohan, 2004). Rasch (en Martínez, 1996) en los 60's desarrolló otros modelos que fueron empleados en Europa y Estados Unidos principalmente, su uso fue limitado en un inicio por la complejidad de los cálculos, que han sido sustituidos actualmente por programas de cómputo que facilitan su empleo, además de sustituir los primeros modelos por otros más sencillos como la Función de Distribución Logística, o los modelos logísticos.

Los modelos de TRI tienen como características:

1. La existencia de rasgos o aptitudes latentes que permiten predecir o explicar la conducta de un examinado ante un ítem de un test.
2. La relación entre el rendimiento de un examinado en un ítem y el conjunto de rasgos responsables de dicho rendimiento, puede describirse mediante una función llamada CCI.
3. Especificación de la dimensionalidad, los modelos actuales consideran que el espacio del rasgo es unidimensional.
4. Todo sujeto puede situarse mediante un punto en el espacio del rasgo.
5. Posición del ítem en el conjunto de valores de rasgo, que da la dificultad del ítem.
6. En la TCT se asumía una puntuación verdadera del individuo que se asume aquí como aptitud Θ .
7. Las unidades básicas de análisis serán los ítems del test.
8. Todos los modelos se establecen antes de puntuar el test, lo que permitirá transformar las respuestas en puntuaciones.

Las ventajas que tienen los modelos del TRI con respecto a los de TCT es la búsqueda del ajuste de los datos y no solo la verificación empírica, la invarianza de los parámetros del ítem de la muestra así como la consideración de que la aptitud del sujeto es independiente del número de reactivos que responde. Los modelos de TRI suponen que subyacente al rendimiento del examinado en un test se encuentra un rasgo o aptitud única. El conjunto de reactivos miden solo una aptitud o rasgo es unidimensional.

La independencia local que establece que “si se mantienen constantes las aptitudes (Θ) que explican el rendimiento en el test, las respuesta de los examinados a un par de ítems cualquiera, son estadísticamente independientes”.

Existen diferentes modelos de la TRI, siendo los logísticos los más comunes, ya que se utilizan principalmente en pruebas de rendimiento.

1. Modelo logístico de un parámetro, es uno de los más difundidos, también es conocido como el modelo de Rasch, su variación relativa describe una función parabólica, mientras que la CCI es una función exponencial. Su ecuación básica es

$$P_{i\theta} = \frac{e^{(\theta-b)}}{1 + e^{(\theta-b)}} \quad i = 1,2,3, \dots n$$

En donde

$P_{i\theta}$ = es la probabilidad de que un examinado al azar con aptitud θ conteste correctamente el ítem i .

b = parámetro de dificultad del ítem i .

n = número de ítem del test.

e = base de los logaritmos neperianos

2. Modelo Logístico de dos parámetros.

En este modelo se define la dificultad y la discriminación o bien posición y escala.

$$P_{i\theta} = \frac{e^{Da(\theta-b)}}{1 + e^{Da(\theta-b)}} \quad i = 1,2,3, \dots n$$

En donde

$P_{i\theta}$ = es la probabilidad de que un examinado al azar con aptitud θ conteste correctamente el ítem i .

b = parámetro de dificultad del ítem i .

n = número de ítem del test.

e = base de los logaritmos neperianos

D_a = 1.7 valor arbitrario de ajuste a la ojiva normal.

3. Modelo Logístico de dos parámetros.

En este modelo se define además el parámetro del pseudo azar, que representa la probabilidad de que una persona conteste un reactivo relativamente difícil de manera correcta por azar.

$$P_{i\theta} = c + (1 - c) \frac{e^{D_a(\theta-b)}}{1 + e^{D_a(\theta-b)}} \quad i = 1,2,3, \dots n$$

En donde

$P_{i\theta}$ =es la probabilidad de que un examinado al azar con aptitud θ conteste correctamente el ítem i .

b = parámetro de dificultad del ítem i .

c = parámetro de pseudo azar.

n = número de ítem del test.

e = base de los logaritmos neperianos

D_a = 1.7 valor arbitrario de ajuste a la ojiva normal.

Modelo de Rasch.

El análisis de Rasch está basado en un modelo matemático que describe la relación de responder de manera correcta un reactivo y las diferencias entre la habilidad del respondiente y la dificultad del reactivo (González, 2008). Se basa en

la unidimensionalidad e independencia local y es derivado de la función logística que relaciona una probabilidad creciente de respuesta a los reactivos conforme aumenta la habilidad de los respondientes. El modelo es monoparamétrico ya que solo describe la probabilidad de respuesta correcta a un reactivo y la diferencia de habilidad del respondiente y la dificultad del reactivo.

El programa de WINSTEPS ® (Linacre, 2009) es un paquete especializado para realizar el análisis de Rasch, que establece como criterios de bondad de ajuste los parámetros INFIT y OUTFIT. El INFIT se interpreta como la media cuadrática ponderada de los residuales que es sensible a patrones de respuesta irregulares. Este valor muestra los comportamientos de respuestas no esperadas a reactivos calibrados cerca del nivel de habilidad del sustentante. El OUTFIT es el promedio de los residuales estandarizados derivados tanto de sustentantes como de reactivos. Este promedio no es ponderado para producir estimaciones más sensibles a respuestas no esperadas lejos de las habilidades calibradas para los respondientes o para las dificultades de los reactivos. Este valor es sensible a valores extremos y a comportamientos no esperados que afectan a las respuestas del sustentante. El criterio de interpretación para ambos es estar entre los valores de .7 y 1.3 para indicar que existe un ajuste adecuado. En el mismo programa se indican el parámetro MEASURE que es la calibración de dificultad del reactivo expresada en logits, donde los números positivos establecen los reactivos con mayor dificultad y por ende que requieren mayor cantidad del atributo, mientras que los negativos muestran los más sencillos que requieren menor cantidad de atributo.

Los modelos de IRT han sido empleados en diversos estudios, en el caso de la DS, Ferrando & Chico (2000), emplearon el modelo de dos parámetros para la adaptación y análisis psicométrico de la escala de Marlowe-Crowne, encontrando que en general los reactivos de la escala ajustan de buena manera, excepto tres

reactivos. Cervellione, Lee, y Bonanno (2009) emplearon el Modelo de Rasch para determinar cuál es la dimensionalidad de la escala de autoengaño del BIDR, encontrando que esta escala debería ser estudiada con los modelos multidimensionales para un mejor ajuste.

Análisis del Funcionamiento Diferencial del Reactivo (DIF).

Es posible que un instrumento pueda ser contestado de manera diferente según al grupo al que se pertenece. Hay preguntas que claramente son más fáciles de contestar por un grupo que por otro; por ejemplo, las mujeres es más fácil que contesten afirmativamente preguntas que impliquen sentimientos, mientras que los hombres no. Este sesgo induce errores sistemáticos en el proceso de medición (Martínez, 1996) y se le denomina DIF (Differential Item Functioning) a la serie de procedimientos que determinan aquellos reactivos que exhiben un funcionamiento diferencial respecto al resto de la prueba (Martínez, 1996). Dentro de esta serie de procedimientos, se utiliza muy frecuentemente la regresión logística (Zumbo, 1999), en donde la VD es el reactivo y como VI es la variable de agrupación que se asume causa la diferencia entre los puntajes.

Análisis Factorial Confirmatorio.

Este análisis permite a partir del conocimiento teórico o de estudios previos determinar si la estructura establecida por un número de observaciones puede explicar las variables latentes de interés (Martínez, 1996; Skrandal & Rabe-Hesketh, 2004). Los supuestos en que se sustenta son:

1. El número de variables latentes se establece de antemano por el investigador.
2. Se especifican a priori las saturaciones de las variables observadas.

3. La covarianza o correlación entre las variables latentes es especificada en el modelo.

El modelo de AFC corrige las deficiencias del AFE conduciendo a una concreción de las hipótesis a contrastar (Batista-Fouquet, Coenders, & Alonso, 2004). La asignación de indicadores específicos a las dimensiones teóricas es la aportación del AFC, que permiten contrastar su validez ajustándolo a un modelo que asuma y diagnostique su bondad de ajuste. Los resultados pueden diferir de los valores obtenidos en un análisis de componentes principales o bien no ser estadísticamente significativos o tener signo opuesto al pronosticado. El estadístico que se emplea para realizar el contraste es la prueba χ^2 Que es una prueba de bondad de ajuste que supone que existen saturaciones entre los factores diferentes a cero (Batista-Fouquet, et al., 2004).

Método

Justificación y Planteamiento del Problema

Los desarrolladores de pruebas y los evaluadores psicológicos deben anticiparse a los problemas y tomar las acciones apropiadas para la elaboración de los materiales e instrucciones necesarias para la aplicación de pruebas psicológicas en ambientes laborales, escolares, clínicos y de investigación. Asimismo, los evaluadores deben ser sensibles al número de factores relacionados a los estímulos, materiales, procedimientos de administración y estilos de respuesta, que pueden atenuar la validez de las inferencias elaboradas a partir los puntajes obtenidos.

Si bien es cierto que los usuarios de las pruebas conocen todos estos factores, es igualmente cierto que minimizan el papel que juega la DS al momento de aplicar un autorreporte. Se ha revisado que la DS es una posible fuente de invalidez así como una fuente genuina de varianza que se asume tiene su sustrato en un componente de aprobación social. Desde hace varias décadas se han desarrollado diversas medidas para controlarla y/o aislarla para obtener más precisión de las medidas.

Actualmente se cuentan con diversas escalas que miden DS, pero todas ellas se han elaborado en el extranjero y no necesariamente representan lo que es deseable o no para el contexto mexicano. Actualmente se ha trabajado este constructo en la población mexicana, pero no se ha sometido a una verificación psicométrica completa la propuesta de Domínguez-Espinosa (2008), por lo que este trabajo tiene como objetivos los planteados a continuación.

Objetivo General

Someter a validación psicométrica el banco de reactivos para la construcción de una escala de DS etnopsicológicamente desarrollada para la población mexicana.

Objetivos Específicos

1. Probar las cualidades psicométricas de la escala de DS con base en la TCT.
2. Probar el comportamiento de la escala de DS en dos formatos diferentes: Likert y Dicotómico.
3. Comprobar la sensibilidad de los reactivos ante dos situaciones de demanda social.
4. Determinar con ayuda de la TRI las propiedades de la escala.
5. Determinar la presencia de DIF en la escala.
6. Confirmar la estructura factorial de la escala.

Para cumplir con estos objetivos, la investigación se dividió en dos estudios:

- Estudio 1: Se realizó un experimento de campo para la realización de los análisis psicométricos utilizando la TCT y TRI. En este estudio se contemplaron 3 fases: 1) Realización de análisis descriptivos de sesgo y dispersión, análisis de discriminación, análisis de confiabilidad por consistencia interna y realización de un AFE para la identificación de dimensiones. 2) Exploración de las propiedades de la escala a partir de la comparación entre los dos formatos utilizados, Dicotómico vs tipo Likert con cinco intervalos de respuesta, con dos instrucciones experimentales: Estándar vs. Demanda social. 3) Realización del TRI para la escala, para comprobar valores de ajuste (INFIT/OUTFIT) y la dificultad de cada reactivo, así como la dimensionalidad de la escala. Se realizó adicionalmente los análisis de DIF según el sexo, la edad y las condiciones de aplicación.

- Estudio 2: Se realizó un estudio de campo para llevar a cabo un AFC y verificar si la configuración obtenida por el AFE se ajusta al modelo teórico propuesto.

A lo largo de los dos estudios se realizó un muestreo no probabilístico por cuotas, balanceado por sexo. Los hallazgos encontrados son generalizables únicamente a poblaciones de características similares. En todos los estudios se utilizaron muestras con población joven y adulta de la ciudad de México y zona conurbada, recolectada entre los años 2008-2010, principalmente en centros educativos de nivel superior.

Estudio 1.

Participantes:

La muestra total para este estudio estuvo formada por 1074 participantes, jóvenes y adultos de la Ciudad de México, segmentados en cuatro grupos:

- 1) Grupo Control, versión Likert: 404 participantes, 252 hombres y 152 mujeres con un rango de 17 a 78 años ($\mu=26.30$ años, $\sigma=10.13$ años)
- 2) Grupo Experimental, versión Likert: 312 participantes 146 hombres y 166 mujeres con un rango de 17 a 58 años ($\mu=27.43$ años, $\sigma=10.1$ años)
- 3) Grupo Control, versión Dicotómica: 158 participantes 73 hombres y 85 mujeres con un rango de 18 a 54 años ($\mu=28.73$ años, $\sigma=11.15$ años).
- 4) Grupo Experimental, versión Dicotómica: 200 participantes 100 hombres y 100 mujeres con un rango de 18 a 50 años ($\mu=30.26$ años, $\sigma=10.21$ años).

Todos seleccionados de manera no probabilísticamente en centros de enseñanza superior y en oficinas, de la zona metropolitana.

Instrumento

Se utilizó el banco de reactivos desarrollado por Domínguez-Espinosa (2008; Domínguez-Espinosa, et al., 2010), que constó inicialmente de 94 reactivos, desarrollado propositivamente para cubrir los dos factores teóricos especificados por Paulhus (1998b), MI y AE, dispuesto en dos versiones:

- 1) La versión tipo Likert con cinco opciones de respuesta, desde Muy de Acuerdo (5) hasta Muy en desacuerdo (1).

2) La versión dicotómica con formato de respuesta Verdadero/Falso.

Asimismo se manejaron para cada formato dos tipos de instrucciones:

a) *Instrucciones estándar:*

A continuación se encuentra una serie de enunciados referentes a comportamientos cotidianos. Por favor, lea cada uno e indique si lo describen o no a Usted. Teniendo en cuenta que no hay respuestas correctas o incorrectas, solicitamos la mayor honestidad posible.

b) *Instrucciones experimentales:*

Imagine que solicita un puesto importante en el gobierno, donde manejará grandes cantidades de material confidencial. Como parte del proceso de selección, se le pide que llene el siguiente cuestionario. Desea Aumentar la probabilidad de que lo elijan. Por lo tanto, responderá presentando una imagen favorable de su persona. Será indispensable que mienta. Sin embargo, debe hacerlo con sutileza, ya que la prueba puede detectar mentiras demasiado evidentes.”

Para la elaboración de las instrucciones experimentales, se tomaron los estudios de Dunnet et al. (1981) y Ellingson et al. (1999) como referencia.

De los 94 reactivos, 42 se calificaron en sentido inverso. Del total de reactivos, se considero que 48 correspondían a la dimensión AE y 46 para MI (ver Anexo 1).

Procedimiento

Las aplicaciones se llevaron a cabo de manera individual y grupal, en los centros educativos, recreativos, de trabajo, de los participantes. Se les solicitaba su

participación voluntaria, y una vez accediendo, se procedía a la entrega de manera aleatoria una de las versiones de la escala. El tiempo de aplicación fue de 30 minutos aproximadamente. Una vez realizadas las aplicaciones se procedió a la captura y análisis de datos para cada fase.

Fase 1: Análisis Psicométrico con base en TCT.

Se efectuó un análisis del banco considerando únicamente el formato tipo Likert con los dos tipos de instrucciones.

Análisis de Datos

Se realizaron análisis básicos descriptivos de sesgo, de comparaciones entre las instrucciones estándares y experimentales para cada uno de los 94 reactivos, comparación por grupos extremos, análisis de consistencia interna por α de Cronbach y un AFE con componentes principales y modelo ortogonal con rotación varimax, siguiendo las recomendaciones de Reyes-Lagunes y García y Barragan (2008).

Resultados:

Condición Estándar

Comparación de medias mediante la prueba t de muestras independientes: los 94 reactivos fueron sometidos a este análisis, determinando que no existían diferencias estadísticamente significativas en 9 reactivos, que fueron retirados de los análisis subsecuentes.

Confiabilidad de la escala: La confiabilidad del banco total obtuvo un valor de alfa de Cronbach de .919 para los 94 reactivos. Al ver los valores de correlación por reactivo se eliminaron aquellos cuyo valor estaba por debajo de .20, de este análisis 18 reactivos no presentaron una correlación adecuada, por lo que se eliminaron para el AFE.

Validez factorial de la escala: Se obtuvieron en un primer análisis 8 factores, de los cuales se seleccionaron solo 4 que explicaban la mayor cantidad de varianza y que tenían valores eigen superiores a 1, con un total fueron 44 reactivos.

Con estos cuatro factores se realizó un nuevo análisis que en su solución mostró cuatro dimensiones consistentes con 28 reactivos (Ver Tabla 1).

La escala de DS para la condición estándar quedó configurada con cuatro dimensiones, que en su construcción original explicarían el manejo de impresión y el autoengaño, correspondiendo los factores 1, 2 y 4 a la primera y el factor 3 al segundo. Los valores Eigen, de varianza explicada, comunalidad y confiabilidad, para cada una de los cuatros factores se observan en la Tabla 2.

El análisis del factor 1 muestra un valor eigen de 8.95 con un 31.95% de la varianza explicada, y un $\alpha=.95$ este factor se le denominó **Honradez**, que se define conductas que muestran la calidad de prueba, proceder recto e integridad de un individuo (Raluy, 2004). El factor 2 muestra un valor eigen de 3.16 con un porcentaje de varianza explicada de 11.28%, $\alpha=.77$ esta se nombró como **Respeto** que se define como acatamiento que se hace a uno, es un miramiento, que tiene consideración y atención, prevención o respeto. El factor 3 reportó un valor eigen de 2.18 con un porcentaje de varianza explicada del 7.79%, $\alpha=.58$ este

factor en conjunto recibió el nombre de **Humildad**, que es la virtud de reconocer nuestra bajeza y miseria de obra y obrar conforme a ello, además de mostrar sumisión y rendimiento. El factor 4 reporta un valor eigen de 1.46 explicando una varianza de 5.21% $\alpha=.74$, a este factor se le dió el nombre de **Sagacidad** que define a una persona avisada, astuta y prudente, que prevé y previene las cosas.

Tabla 1. Solución Factorial con los 4 primeros factores, Versión Likert, Instrucciones Estándar

Reactivo	Carga Factorial				Comunalidad
	1	2	3	4	
31Robaría algo si nadie me sorprendiera.	.88				.79
17Tomaría algo que no me pertenece si supiera que le daría mejor uso qu	.82				.68
65Ante ciertos problemas, mi salida más cómoda es sobornar a quien sea	.82				.74
32Consumiría alguna droga si nadie me juzgara por ello.	.77				.61
9Hablo sobre cosas personales de mis amistades sin que ellos se enteren	.75				.58
11He recibido cambio de más en alguna tienda y no he dicho nada.	.74				.57
82He evitado devolver algo que no me pertenece fingiendo olvido.	.73				.59
26Me dejaría sobornar si el beneficio que recibiera fuera muy grande.	.72				.62
91He cobrado de más por un servicio sin merecerlo.	.72				.61
85Suelo juzgar a las personas más por su físico que por sus opiniones.	.71				.56
84Le desearía el mal a otra persona si estuviera justificado.	.68				.48
76Demuestro mi educación sólo en lugares donde se me exige.	.64				.45
64Miento cuando me beneficia.	.62				.73
48He hablado mal de mis amigos sin que ellos lo sepan.	.62				.43
68Respeto a quienes me rodean sin importar su punto de vista.		.72			.58
43Soy amable con todas las personas, sin importar su forma de ser.		.71			.57
42Cuando veo que a alguien se le caen las cosas que está cargando, le a		.70			.52
23Respeto a la gente sin importar su condición social.		.68			.47
41Si alguien me contara un secreto, sería muy difícil que lo divulgara, aun		.63			.40
60Ante cualquier situación, estoy dispuesto a ayudar a la gente.		.60			.45
10Perdono fácilmente a quienes me ofenden.			.78		.61
66Olvido las fallas de otros aunque me lastimen mucho.			.76		.58
47Olvido fácilmente las ofensas que me hacen.			.76		.58
49Constantemente busco reconciliarme con mis enemigos.			.54		.32
75Aunque yo esté triste, soy feliz si los demás lo son.			.53		.32
37Miento si es necesario.				.77	.64
35Omitiría la verdad si me fuera a perjudicar.				.73	.58
53Digo mentiras si sé que no me van a descubrir.				.64	.71

Tabla 2. Valores Eigen, Varianza Explicada, versión Likert, Instrucciones Estándar

Factor	Valor eigen	% de varianza	% acumulativo	α
1	8.95	31.95	31.95	.95
2	3.16	11.28	43.23	.77
3	2.18	7.79	51.03	.58
4	1.46	5.21	56.24	.74

Condición Experimental (demanda social)

Comparación de medias mediante la prueba t de muestras independientes: los 94 reactivos fueron sometidos a este análisis, determinando que no existían diferencias estadísticamente significativas en 2 reactivos, que fueron retirados de los análisis subsecuentes.

Confiabilidad de la escala: Se obtuvo un α de Cronbach de .959 para los 94 reactivos. De este análisis 28 reactivos no presentaron una correlación ítem-total superior a .20, por lo que se eliminaron.

Validez factorial de la escala: Se obtuvieron en un primer análisis 6 factores, de los cuales se seleccionaron 3 que explicaban la mayor cantidad de varianza y que tenían valores eigen superiores a 1, constituyéndose con 39 reactivos. Con estas tres dimensiones se realizó un nuevo segundo análisis factorial y se observó que solamente 33 reactivos tuvieron pesos factoriales superiores a .40, mismos que permanecieron configurados en tres dimensiones (ver Tabla 3). En la tabla 4 se resumen los valores de varianza explicada, comunalidad y confiabilidad.

El primer factor presentó un valor eigen de 15.60 y una varianza explicada de 42.15%, este factor se denominaría **Conveniencia**, *que correlaciona y muestra conformidad entre cosas distintas, es un ajuste con cierto convenio, o búsqueda de provecho*; esta dimensión muestra que tiene componentes de autoengaño y manejo de impresión. El segundo factor tuvo un valor eigen de 3.02 con un porcentaje de varianza de 8.16%, el contenido refiere **Cortesía**, *expresión o demostración de respeto o afecto que tiene una persona por otra*; estas actitudes son referidas por su contenido al manejo de impresión. El tercer factor tiene un valor eigen de 1.13 y un porcentaje de varianza de 3.07%, esta dimensión se

denominó **Tolerancia**, esta se refiere al respeto y consideración hacia las opiniones y prácticas de los demás; actitudes referidas al manejo de impresión.

Tabla 3. Solución Factorial con los 4 primeros factores, Versión Likert, Instrucciones Experimentales

Reactivo	Carga Factorial			Comunalidad
	1	2	3	
31Robaría algo si nadie me sorprendiera.	.88			.80
91He cobrado de más por un servicio sin merecerlo.	.84			.77
26Me dejaría sobornar si el beneficio que recibiera fuera muy grande.	.81			.68
65Ante ciertos problemas, mi salida más cómoda es sobornar a quien sea necesario.	.80			.72
9Hablo sobre cosas personales de mis amistades sin que ellos se enteren.	.77			.65
85Suelo juzgar a las personas más por su físico que por sus opiniones.	.77			.63
84Le desearía el mal a otra persona si estuviera justificado.	.76			.63
32Consumiría alguna droga si nadie me juzgara por ello.	.76			.64
82He evitado devolver algo que no me pertenece fingiendo olvido.	.75			.59
76Demuestro mi educación sólo en lugares donde se me exige.	.74			.67
17Tomaría algo que no me pertenece si supiera que le daría mejor uso que su actual dueño.	.73			.57
80Trabajo menos cuando sé que nadie me está supervisando.	.72			.60
12En mi trabajo, pierdo tiempo en otras actividades si sé que nadie me va a sorprender.	.72			.57
53Digo mentiras si sé que no me van a descubrir.	.71			.71
89He hecho trampa para obtener un beneficio personal.	.71			.68
29Me disgusta que los demás tengan más éxito que yo.	.70			.51
48He hablado mal de mis amigos sin que ellos lo sepan.	.68			.61
54Suelo desear los bienes ajenos.	.68			.51
33Le he faltado al respeto a alguien sin una razón justificable.	.68			.48
64Miento cuando me beneficia.	.67			.73
11He recibido cambio de más en alguna tienda y no he dicho nada.	.66			.55
46Alguna vez he realizado las cosas de manera ilegal o "chueca"	.65			.49
24Me molesta la gente que no piensa como yo.	.63			.46
78Al platicar ciertas anécdotas, las exagero a mi favor.	.62			.43
51He aceptado el crédito por algo que no merecía.	.58			.43
61Si me es posible, realizo solamente las tareas que me corresponden y no busco hacer más.	.53			.35
74Soy igualmente educado con personas de confianza que con personas nuevas.		.78		.68
79Doy las gracias cada vez que pido o recibo algo, aun con gente de confianza.		.77		.62
73Prefiero ocuparme en algo útil que perder el tiempo.		.74		.60
23Respeto a la gente sin importar su condición social.		.52		.51
66Olvido las fallas de otros aunque me lastimen mucho.			.77	.68
67Soy amable a pesar de que los demás no lo sean.			.75	.69
62Respeto a la gente aun cuando ella no me respete a mí.			.75	.61

Tabla 4. Valores Eigen, Varianza Explicada, versión Likert, Instrucciones Estándar

Factor	Valor eigen	% de varianza	% acumulativo	α
1	15.60	42.15	42.15	.97
2	3.02	8.16	50.31	.81
3	1.13	3.07	53.37	.64

Al comparar las soluciones factoriales resultantes de ambas configuraciones, se observa que ambas comparten 17 reactivos, que representan el 60.7% de la solución final de la escala.

Discusión

Las principales diferencias entre las dos escalas obtenidas, radicó en la configuración de los factores, así como en el número de reactivos de cada uno, en la condición estándar se obtuvieron 4 factores, mientras que en la condición experimental se determinaron 3.

Para asignar el nombre a cada factor se tomó en cuenta el contenido de cada reactivo, la mayoría versaba en función de tomar objetos que no le pertenecen, la mentira, entre otros

El factor 1 representa conductas que hablan de la **Honradez** como una conducta socialmente aceptada, donde el abuso o el uso del poder para su propio beneficio es reprobado, si no hay una justificación de por medio, estas conductas que se presentan en la sociedad mexicana hablan de la posibilidad del uso de la deshonestidad en ciertos casos que a pesar de ser vista de forma incorrecta en algunas sociedades, tiene aceptación o justificación según el caso.

El factor 2 fue asignado como **Respeto**, el contenido de los reactivos menciona conductas que expresan lealtad y respeto por el otro, estos valores ante situaciones donde existe la expresión de sentimientos o emociones es esencial, de tal manera que la persona que no lo hace así es considerada como poco confiable.

El factor 3 se denominó **Humildad**, el contenido de los reactivos presenta características personales que son consideradas aceptables por la sociedad, como el perdón por las ofensas así como el llevarse bien con todas las personas, estas cualidades muestran que la persona se puede sobrevalorarse o incluso engañarse.

El factor 4 está formado por 3 reactivos el contenido de los reactivos muestra que hay conductas que en la sociedad mexicana para prevenir malos resultados, a este conjunto se le llamó **Sagacidad**, un ejemplo es aquella actitud de la sociedad que valora la astucia de las personas para salir bien librada de problemas.

La construcción de la escala de DS se realizó tomando en cuenta las dimensiones teóricas de Paulhus (1998b) el contenido de los factores 1, 2, 4 puede considerarse como características pertenecientes al MI, mientras que el factor 3 pertenece a la dimensión del AE, cabe señalar que cada uno de los factores obtenidos tienen un significado individual lo que pone en evidencia que la misma dimensión de MI tiene a su interior otros elementos que lo definen.

La configuración factorial obtenida mediante la manipulación de instrucciones no fue la misma, fue equivalente y aunque permite integrar otros reactivos que en la condición estándar no se incluyeron por el criterio de discriminación y correlación.

La configuración factorial en la condición experimental muestra contenidos diferentes a los de la escala estándar, de tal manera que la manipulación experimental, explica otros factores o elementos que pueden estar presentes en las dimensiones teóricas de MI y AE.

El factor 1 se denominó Conveniencia esta se configuró con 26 reactivos, de los cuáles 12 coinciden con el factor 1 de la escala estándar lo que nos indicaría cierta consistencia interna del factor, aun a pesar de las instrucciones experimentales, los 14 reactivos restantes muestran contenidos que no solo están asociados con la Honradez, si no que se asocian a la **Conveniencia**, que muestra una actitud de la población en la cual el ajuste que tenga a los convencionalismos sociales, como el trabajo, el éxito depende de las circunstancias. El segundo factor denominado **Cortesía**, es una dimensión diferente a las obtenidas en la condición normal, la manipulación hace evidente como dentro de las conductas socialmente aceptadas se encuentra el ser cortés, siendo esta una característica que en la sociedad mexicana tiene un gran peso, de ahí el dicho “lo cortés no quita lo valiente”. El tercer factor se denominó **Tolerancia** ya que el contenido de los tres reactivos que forman esta dimensión mencionan lo importante es saber aceptar al otro a pesar de uno mismo, conducta arraigada en la familia donde el respeto por el otro está antes que el propio.

En este caso los factores no son claros con respecto a las dimensiones de teóricas de MI y AE, por otro lado algo valioso de este procedimiento es la evaluación que hace de los reactivos en cuanto a su importancia como conductas socialmente aceptadas, ya que son los que la gente califica con mayor frecuencia y encuentra diferencias de discriminación y valores más altos de correlación.

La manipulación experimental, muestra que los reactivos que fueron construidos en la dimensión teórica de MI fueron los más susceptibles a distorsión; es decir, la escala varió su configuración a partir de poner al sujeto en una situación de demanda que es acorde a lo que se espera en la literatura (Dunnett, et al., 1981; Holden, 2008).

Ambas versiones capturan un porcentaje de varianza explicada por arriba del 50% valor adecuado para el constructo de DS. Paulhus (2002) en sus últimos estudios plantea la existencia de dimensiones al interior del MI y AE, que en este estudio se corrobora en el hecho del planteamiento de los factores de la escala, por lo que actualmente se considera necesario reconsiderar el constructo por la cercanía a presentarse como un rasgo adaptativo, lejos de la concepción de las primeras ideas del significado de DS (Uziel, 2010).

Fase 2: Evaluación de los Formatos Likert vs. Dicotómico

La presente fase tuvo la finalidad de determinar el comportamiento de la Escala Preliminar Dominguez de Deseabilidad Social (en lo sucesivo denominada EPDDS) en sus dos versiones de calificación: Dicotómica y tipo Likert, y obtener evidencia de cuál de los formatos es susceptible de cambios o distorsión.

Procedimiento

Se efectuó un análisis sobre los 28 reactivos de la EPDDS obtenida en la primera fase empleando para ello los puntajes estandarizados para efectuar las comparaciones entre los valores de cada grupo. Posteriormente, se realizaron análisis de confiabilidad para cada una de las versiones con los reactivos, así

como un análisis de modelo lineal general, para determinar si existía algún efecto entre las condiciones de aplicación y el formato.

Resultados

Los valores descriptivos para cada una de los formatos muestran como los puntajes medios de DS para la escala Likert son mayores en la condición experimental, mientras que para el formato dicotómico la condición estándar su promedio es mayor que la condición experimental (ver Tabla 5)

Tabla 5. Puntajes promedios de los dos formatos de la EPDDS

Condición	Formato	Media	DE	n
Normal	Likert	-0.2491	0.9179	353
	Dicotómico	0.1776	0.9159	149
Experimental	Likert	0.3186	1.011	276
	Dicotómico	-0.1438	1.043	184

Al determinar el valor de confiabilidad de Cronbach, se obtuvo que los puntajes más elevados se encuentran en la escala tipo Likert, con respecto a la escala dicotómica (Ver Tabla 6). Por otro lado la condición de demanda elevó los valores de la confiabilidad en los casos de la situación experimental.

Tabla 6. Índices de confiabilidad para cada formato de la EPDDS

Condición	Tipo de Calificación	
	Tipo Likert	Dicotómica
Estándar	0.905	0.759
Experimental	0.929	0.833

Mediante un análisis de modelo lineal general de efectos fijos, se observó que las diferencias entre la condición de aplicación y el formato de evaluación no son estadísticamente significativos, excepto en la interacción de la condición y el formato que muestran diferencias entre cada una de las condiciones.

Tabla 7. Modelo de efectos fijos entre condiciones y formato de la EPDDS

Fuente	F	Sig.	R ²
Condición	3.465	0.063	3.261
Formato	0.0073	0.787	0.069
Condición/ Formato	45.138	0.000	42.481

Discusión

Los resultados obtenidos en términos de la consistencia interna, nos sugieren que la escala tipo Likert es una mejor alternativa que la escala dicotómica, siendo que presentan mejor ajuste en sus valores de confiabilidad (Cervellione, et al., 2009; Greenwald & O'Connell, 1970; Stöber, Dette, & Musch, 2002).

El análisis de modelo lineal general muestra que de forma independiente la condición de aplicación, aunque en esta comparación no fue significativa, si representa una variación, lo que es coincidente con los hallazgos de la primera fase y de acuerdo con otras investigaciones en donde la manipulación hace una diferencia en los puntajes (Holden, 2007). Respecto al formato, los resultados indican que no ejercen un efecto importante, sino es en conjunto con las instrucciones, por lo que se sugiere mantener el formato tipo Likert ya que es posible convertirlo a dicotómico con solo adjudicarle el puntaje si el sujeto contesta o no en la dirección socialmente deseable.

Fase 3: Análisis Psicométrico utilizando TRI

Este estudio tuvo como finalidad determinar por medio de la TRI, el comportamiento y ajuste de la escala obtenida, para determinar así la consistencia de la misma y la dificultad de cada reactivo.

Participantes e Instrumento

Se consideró solamente para este estudio los datos de las escalas en su formato Likert, tanto la versión estándar como la experimental.

Procedimiento

Se efectuó un análisis de Rasch para la EPDDS, empleando el programa Winsteps® (Linacre, 2009). Obteniendo los parámetros de ajuste para la escala total, así como en base a su distribución teórica según las dimensiones MI y AE por separado y las dimensiones empíricas obtenidas en la fase 1 (mediante el AFE).

Resultados

Al efectuar el análisis de Rasch para la escala completa, se observó que al interior del instrumento no es unidimensional, razón por la cual, algunos de los reactivos mostraron desajuste en sus valores de INFIT y OUTFIT, que indica que existen

valores extremos o que pertenecen a otra escala. los valores de dificultad son muy cercanos e inclusive algunos tiene en el mismo índice (Ver Tabla 8²).

Tabla 8. Valores de los parámetros del modelo Rasch para la EPDDS

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
49 Constantemente busco...	0.36	1.34	1.54
66 Olvido las ...	0.26	1.32	1.45
37 Miento si es...	0.25	1.01	1.09
35 Omitiría la verdad...	0.21	1.1	1.13
11 He recibido...	0.20	0.95	1.02
64 Miento cuando...	0.19	0.71	0.7
53 Digo mentiras...	0.19	0.74	0.71
85 Suelo juzgar...	0.19	0.8	0.81
48 He hablado mal...	0.18	0.89	0.99
75 Aunque yo esté...	0.16	1.42	1.66
9 Hablo sobre...	0.16	0.9	0.92
26 Me dejaría sobornar...	0.15	0.78	0.72
91 He cobrado de más...	0.13	0.74	0.68
17 Tomaría algo que ...	0.13	0.99	1.07
65 Ante ciertos ...	0.13	0.76	0.71
32 Consumiría alguna...	0.11	1	0.96
82 He evitado devolver	0.11	0.86	0.85
47 Olvido fácilmente...	0.11	1.35	1.54
31 Robaría algo si...	0.11	0.84	0.74
10 Perdono fácilmente...	0.09	1.31	1.44
76 Demuestro mi educación	0.08	0.93	0.97
84 Le desearía el mal....	0.08	0.9	0.84
60 Ante cualquier....	-0.38	0.93	1.32
43 Soy amable...	-0.53	0.9	1.14
68 Respeto a quienes...	-0.55	0.81	0.87
41 Si alguien me contara...	-0.55	1.46	2.07
23 Respeto a la gente.....	-0.77	1.31	1.45
42 Cuando veo que	-0.80	1.13	1.28
M	0.00	1.01	1.1
D.E.	0.32	0.23	0.35

Los reactivos que no ajustan a la escala completa por tener valores superiores a 1.3 de INFIT y/o OUTFIT son el 49, 66, 75, 47. 10 que pertenecen al factor número 3, además de los reactivos 41 y 23 que pertenecen al factor 2. La dimensión 3 que se consideró dentro de la dimensión teórica de AE, muestra consistencia con los resultados del AFE, y cabe señalar que el reactivo 23 teóricamente fue construido de acuerdo al autoengaño y al efectuar el AFE, no cargó en esta dimensión si no en la de MI. Estos resultados muestran que existen dimensiones al interior, por lo

² Nota: la numeración de los reactivos en todos los análisis de Rasch pertenecen a los números originales del banco de reactivos.

que se realizó un análisis de ajuste por factor, donde se observó la consistencia de cada escala.

En las tablas 9, 10, 11 y 12 se muestran los valores de ajuste para cada dimensión empírica identificada en la fase 1, respectivamente.

Tabla 9. Valores de los parámetros del modelo Rasch para la dimensión 1 (Honradez).

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
11 He recibido...	0.09	1.14	1.17
64 Miento cuando...	0.07	0.98	1.16
85 Suelo juzgar...	0.06	0.95	1.03
48 He hablado mal...	0.05	1.19	1.40
9 Hablo sobre...	0.03	1.03	1.12
26 Me dejaría sobornar...	0.02	0.93	0.9
91 He cobrado de más....	-0.01	0.82	0.82
17 Tomaría algo que ...	-0.01	1.04	1.08
65 Ante ciertos ...	-0.02	0.77	0.74
82 He evitado devolver	-0.04	1	1.08
32 Consumiría alguna...	-0.04	1.11	1.15
31 Robaría algo si...	-0.04	0.75	0.59
76 Demuestro mi educación	-0.08	1.20	1.51
84 Le desearía el mal....	-0.08	1.12	1.24
M	0.00	1	1.07
D.E.	0.05	0.14	0.24

Para la dimensión 1 solamente desajustan los reactivos 48 y 76 (Tabla 9). En la dimensión 2, el reactivo 41 (Tabla 10). En la tercera (Tabla 11) y cuarta dimensión (Tabla 12) no se presentan desajustes al modelo.

Tabla 10. Valores de los parámetros del modelo Rasch para la dimensión 2 (Respeto)

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
60 Ante cualquier....	0.37	0.97	1.03
43 Soy amable...	0.12	0.81	0.87
68 Respeto a quienes...	0.08	0.83	0.83
41 Si alguien me contara...	0.07	1.45	1.37
23 Respeto a la gente.....	-0.30	1.26	1.14
42 Cuando veo que	-0.34	0.95	0.81
M	0.00	1.04	1.01
D.E.	0.25	0.23	0.2

Tabla 11. Valores de los parámetros del modelo Rasch para la dimensión 3 (Humildad)

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
49 Constantemente busco...	0.24	1.06	1.10
66 Olvido las ...	0.09	0.90	0.90
75 Aunque yo esté...	-0.05	1.17	1.22
47 Olvido fácilmente...	-0.12	0.95	0.90
10 Perdono fácilmente...	-0.16	0.96	0.93
M	0.00	1.01	1.01
D.E.	0.15	0.1	0.13

Tabla 12. Valores de los parámetros del modelo Rasch para la dimensión 4 (Sagacidad)

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
37 Miento si es...	0.07	0.92	0.88
35 Omitiría la verdad...	-0.01	1.08	1.03
53 Digo mentiras...	-0.06	1	0.94
M	0.00	1	0.95
D.E.	0.05	0.06	0.06

A continuación se presentan los análisis respecto a las dimensiones teóricas MI correspondientes al modelo propuesto por Paulhus (1998b). Se puede observar en la tabla 13, la agrupación de las dimensiones 1, 2 y 4 que son consideradas de MI, los reactivos 35, 60, 43, 41, 23 y 42 con sus valores de INFIT y OUTFIT salen del rango permitido.

Tabla 13. Valores de los parámetros del modelo Rasch para la dimensión MI (1, 2 y 4).

Reactivos EDS	Dificultad	INFIT MNSQ	OUTFIT MNSQ
37 Miento si es...	0.32	1.22	1.4
35 Omitiría la verdad...	0.27	1.35	1.49
11 He recibido...	0.27	1.01	1.09
64 Miento cuando...	0.26	0.76	0.76
53 Digo mentiras...	0.25	0.80	0.78
85 Suelo juzgar...	0.25	0.84	0.88
48 He hablado mal...	0.24	0.96	1.01
9 Hablo sobre...	0.22	0.94	0.96
26 Me dejaría sobornar...	0.21	0.80	0.72
91 He cobrado de más....	0.19	0.74	0.68
17 Tomaría algo que ...	0.19	0.99	1.10
65 Ante ciertos ...	0.18	0.73	0.66
32 Consumiría alguna...	0.17	1.02	1.09
82 He evitado devolver	0.17	0.87	0.90
31 Robaría algo si...	0.16	0.78	0.66
76 Demuestro mi educación	0.13	0.99	1.15
84 Le desearía el mal....	0.13	0.96	0.95
60 Ante cualquier....	-0.38	1.11	2.16
43 Soy amable...	-0.53	1.04	1.94
68 Respeto a quienes...	-0.55	0.93	1.15
41 Si alguien me contara...	-0.56	1.59	2.84
23 Respeto a la gente.....	-0.78	1.39	1.79
42 Cuando veo que	-0.81	1.18	1.62
M	0.00	1	1.21
D.E.	0.37	0.22	0.54

Se efectuó el análisis de dimensionalidad de Rasch, y se identificó que efectivamente la escala no es unidimensional (porcentaje de varianza no explicada de 5.1). El estudio de dimensionalidad de acuerdo a la construcción teórica, muestra que los valores de varianza no explicada para la dimensión de manejo de impresión está por arriba del límite (>2), mientras que la dimensión de autoengaño está en los valores permitidos (ver tabla 14)

Tabla 14. Valores Eigen de los residuales del Modelo Rasch en las 2 dimensiones

Dimensión	Varianza Explicada en las Observaciones	Porcentaje	Varianza no explicada	Porcentaje
Manejo Impresión	43.7	100.00%	3.2	7.30%
Autoengaño	8.1	100.00%	1.5	18.20%

Discusión

Al realizar el análisis de Rasch considerando la escala como unidimensional, los valores de ajuste muestran que siete de los 28 reactivos presentan valores de fuera de rango de OUTFIT o INFIT, el rango de dificultad se encuentra (0.36, -0.80) para los 28 reactivos, de estos dieciséis coinciden en sus valores de dificultad por lo que 14 de ellos sería necesario seleccionar uno ya que tienen el mismo grado, y cualquiera de ellos requeriría la misma cantidad de atributo para ser contestado.

Este primer análisis muestra que si se eliminaran los reactivos con igual dificultad, quedarían solamente algunos por lo que será conveniente re-elaborar reactivos para ampliar el espectro que cubre. El análisis por dimensión obtenida de acuerdo al AFE, mostró mejor consistencia en cada uno de los factores, aunque el más sensible a distorsión fue la dimensión 1, en estas hay 6 reactivos que respectivamente igualan el nivel de dificultad, por lo que tres podrían ser eliminados, además de 2 que tienen valores fuera de rango, de lo que de los catorce reactivos se tendrían que eliminar 4 bajo estos criterios, lo niveles de dificultad es muy cercano (0.09, -0.08). Para la dimensión 2 los valores de ajuste son adecuados excepto para el reactivo 41 que también muestra este desajuste en el primer análisis de la escala completa, el rango de dificultad se encuentra entre (0.37, -0.34), por lo que se ubica que existen más reactivos que requieren

mayor atributo (0.37), al igual que de menor cantidad (-0.34). La dimensión 3 considerada de AE, mostró un adecuado nivel de consistencia el nivel de dificultad se encuentra de (0.24, -0.16). La dimensión 4 su ajuste es adecuado con niveles de dificultad entre (0.07, -0.6). De esta forma se puede señalar que las dimensiones 3 y 4 poseen características que de acuerdo al modelo de Rasch muestran su consistencia e independencia con la escala total, que desajusta al considerar a todos los reactivos dentro de una dimensión e incluso sus niveles de dificultad cambian.

Considerando la teoría de Paulhus (1998b) que determina dos dimensiones para la DS, los factores 1, 2, 4 corresponde al MI, por lo que se realizó un análisis donde se observó que el nivel de dificultad se encontró entre (0.32, -0.81), al considerar como una sola dimensión, algunos de los reactivos desajustan siendo constantes en esta situación los reactivos 41 y 23, el reactivo 41 en todos los análisis mostro desajuste, lo que puede indicar que es un reactivo que a pesar de que en el AFE mostro carga factorial y comunalidad dentro de los valores permitidos, no es un reactivo que aporte información a la construcción de la escala mediante este modelo de medición.

Los valores que se muestran a lo largo de la diferentes escalas tiene la utilidad de situar a los individuos respecto a la cantidad de atributo que presentan para responder en la dirección que se considera como socialmente deseable, el instrumento a pesar de no contar con reactivos con un valor alto de dificultad, los valores muestran la cantidad de DS para contestar el reactivo en la dirección esperada. En los presentes análisis se pone en claro que no solamente tenemos dos dimensiones teóricas al interior de la escala, sino cuatro por lo que se deberá hacer la revisión de los reactivos para eliminar aquellos que tienen el mismo grado de dificultad y posiblemente llegar a una solución de dos dimensiones teóricas.

Fase 4: Análisis de DIF

Este estudio tuvo como objetivo determinar si algunos reactivos eran susceptibles de presentar diferencias en la probabilidad de respuesta de acuerdo a posibles fuentes de variación como la condición de aplicación, el sexo y la edad.

Participantes e Instrumento

Se considero solamente para este estudio los datos de las escalas en sus formatos Likert, tanto la versión estándar como la experimental.

Procedimiento

Se efectuaron análisis de regresión logística para cada reactivo, tomando el puntaje de cada uno como variable dependiente y como variables independientes el total del puntaje en la escala y cada una de las variables de agrupación por separado (condición, sexo y edad).

En el análisis por edad se separó la muestra en dos grupos el primero de 17- 30 años y el segundo con los adultos de 31 a 58 años, tomando en cuenta lo señalado por Costa y McCrae (1985) quienes indican que en términos generales, los cambios en la personalidad se llevan a cabo antes de los 30 años.

Resultados

El análisis por la condición experimental determinó que los reactivos 23 y 42 pertenecientes al factor 2 son sensibles a esta situación al igual que los reactivos 35 y 53 que son del factor 4 (ver Tabla 15). En el análisis por sexo se observa DIF en 3 reactivos, de los cuales 1 pertenece al factor 1, mientras que se presentan 2 reactivos pertenecen a los factores 2 y 4 (Ver tabla 16). En el análisis por edad, tres reactivos presentan DIF, siendo de los factores 1, 2, 4 (Ver tabla 17). Es importante señalar que los reactivos del factor 3 no muestran DIF en ninguna de las condiciones.

En la tabla 15 se observa que el reactivo 23. *Respeto a la gente sin importar su condición social*, muestra un valor β -0.439 que indica más probabilidad de ser contestado en la dirección de la respuesta socialmente deseable bajo las condiciones experimentales de 78.5%. El reactivo 42 *Cuando veo que a alguien se le caen las cosas que está cargando, le ayudo a recogerlas*, tiene una β -0.747 con una probabilidad del 79.9% de acierto en la respuesta bajo la situación experimental. Por otro lado el reactivo 35 *Omitiría la verdad si me fuera a perjudicar*, la probabilidad de respuesta es mayor en las personas que contestan en la versión estándar su β 0.526. El reactivo 35 *Digo mentiras si sé que no me van a descubrir*, muestra una mayor probabilidad de respuesta en la condición estándar de 82% ya que su β 0.561.

De esta forma se observa que no hay diferencias de respuesta en 24 de los reactivos de acuerdo a las instrucciones de aplicación excepto en los casos anteriormente señalados que muestran que existen reactivos más sensibles a la situación experimental 23, 42 mientras que los reactivos 35 y 53 lo hacen en la condición contraria, esto puede deberse en gran medida a la demanda que se solicita en cada reactivo ya que en los primeros de acuerdo al estudio de Rasch

previamente realizado son reactivos de baja dificultad, mientras que los dos restantes están en un nivel de mayor demanda social, que requiere mayor DS.

Tabla 15. Valores de DIF a partir de la Condición.

Reactivo		β	DE	Cociente de probabilidad	Estadística de Wald
23 Respeto a la gente sin....	*	-0.439	0.195	0.785	5.053
42 Cuando veo a alguien....	**	-0.747	0.203	0.799	13.533
35 Omitiría la verdad si	**	0.526	0.205	0.820	6.561
53 Digo mentiras si sé	*	0.561	0.245	0.863	5.240

* $P < .05$

** $P < .01$

Según el análisis por sexo (tabla 16), el reactivo *26 Me dejaría sobornar si el beneficio fuera muy grande*, muestra que las mujeres no contestarían en la dirección de la respuesta socialmente aceptada con respecto a los hombres ya que su valor de $\beta = 0.543$, lo que indica que los hombres requieren menos del atributo de la DS para contestar en la dirección de calificación, que puede estar asociada a la forma en que se solucionan problemas en lo cotidiano.

Por otra parte los reactivos *48 He hablado mal de mis amigos sin que ellos lo sepan*, *85 Suelo juzgar a las personas más por su físico que por sus opiniones*, *41 Si alguien me contara un secreto, sería muy difícil que lo divulgara, aunque me presionaran*, y el reactivo *53 Digo mentiras si sé que no me van a descubrir*, la probabilidad de responder de manera socialmente aceptable es mayor en los hombres, ya que los valores de β respectivamente fueron de -1.019, -0.797, -0.408, -.0822 que señalan que las mujeres lo contestarían con mayor facilidad que los hombres, en gran medida por las ideas de considerar que las mujeres son generalmente más emotivas, sensibles con respecto a los hombres, el contenido de estos reactivos esta asociado a emociones.

Tabla 16. Valores de DIF a partir del Sexo.

Reactivo		β	DE	Cociente de probabilidad	Estadística de Wald
26 Me dejaría Sobornar....	*	0.543	0.246	0.860	4.890
48 He hablado mal de	**	-1.019	0.254	0.842	16.137
85 Suelo juzgar a las	**	-0.797	0.256	0.856	9.670
41 Si alguien me contara....	*	-0.408	0.184	0.749	4.901
53 Digo mentiras si se	**	-0.822	0.263	0.870	9.746

* P < .05

** P < .01

El análisis por edad (tabla 17), mostró que el reactivo 23. *Respeto a la gente sin importar su condición social*, tiene un valor de $\beta = -0.553$ que nos indica que los jóvenes menores de 30 años tienen menor probabilidad de contestarlo en la condición socialmente deseable, mientras que los adultos lo responden con mayor frecuencia, esto puede deberse en gran medida al cambio de la valoración entre jóvenes y adultos, donde la persona adulta tiene menos la tendencia a querer quedar bien y por ello el respeto a las personas sin importar su estrato social, es apreciada de cualquier forma. De forma contraria los reactivos 37 *Miento si es necesario*, la probabilidad de ser contestada en la dirección socialmente deseable es mayor en los jóvenes con valor de $\beta = 0.535$, de la misma forma que el reactivo 84 *Le desearía el mal a otra persona si estuviera justificado* con una $\beta = 0.587$ indican la dirección de respuesta más probable en los jóvenes que en los adultos, lo que puede interpretarse como la necesidad de aprobación que es mayor en los adolescentes con respecto a los adultos.

Tabla 17. Valores de DIF a partir de la Edad.

Reactivo		β	DE	Cociente de probabilidad	Estadística de Wald
84 Le desearía el mal...	*	0.587	0.251	0.849	5.472
23 Respeto a la gente sin....	*	-0.553	0.224	0.789	6.081
37 Miento si es	*	0.535	0.226	0.823	5.632

* P < .05

Discusión

Dentro de la literatura de DS, estudios evidencian las diferencias que pueden encontrarse por pertenencia a un grupo dentro de una muestra (Henry & Raju, 2006), en esta investigación se incluyó un Análisis Diferencial del Reactivo (DIF), que permitiera reconocer si hay mayor probabilidad para algunos miembros del grupo de contestar en la dirección socialmente deseable, de acuerdo a estas diferencias de pertenencia a un grupo determinado.

Si consideramos al DIF como un sesgo, podemos establecer que teóricamente para un instrumento de medición no debería presentarse, ya que habla de la desventaja entre los miembros de un grupo, además de modificar su validez (Martínez, 1996; Skrandal & Rabe-Hesketh, 2004; Zumbo, 1999).

El estudio del DIF en esta investigación se realizó tomando como posibles factores: la condición de aplicación, el sexo y la edad de los participantes.

El análisis por condición de aplicación determinó que 24 reactivos no tienen diferencias significativas en la probabilidad de acierto entre ambas situaciones, los cuatro reactivos que son sensibles al DIF, en el análisis de Rasch se clasifican como de baja dificultad, dos con calificación de sentido positivo y dos en sentido inverso, que indica que la baja cantidad del atributo que se requiere para una RSD hace que la probabilidad se invierta con respecto a la dirección de calificación.

Las diferencias por sexo muestran que existe DIF explicada por las diferencias de género, que hace evidente que cinco de los reactivos de la EPDDS, en su

contenido 4 son referidos a actitudes consideradas como masculinas y una con disposición más hacia lo masculino (Caudillo & Cerna, 2007).

El DIF por edad muestra que dos reactivos la probabilidad para los jóvenes son mayores debido a que son actitudes que como inexpertos, inmaduros, no criticados o evaluados tan fuertemente como es el caso de las mentiras, el respeto y el deseo del mal a otros, mientras que en los adultos el respeto hacia los demás es condición que habla de la madurez y crecimiento del individuo.

Según Kolberg (2002) el desarrollo moral del individuo atraviesa diferentes etapas a lo largo de la vida y se establece más permanentemente al entrar a la madurez. El constructo de la DS supone diferencias en la valoración de los eventos en términos de lo apropiado o no, ético o no, moral o no, razón por lo cual es de esperarse que existan diferencias por edad y sexo como lo sugieren algunos autores (p.ej. Costa & McCrae, 1985; Diekman & Goodfriend, 2006; Domínguez, Navarro, García, Gutierrez, & Sada, 2010; Domínguez, Salas, & Contreras, 2010; Johannesen-Schmidt & Eagly, 2002; Terracciano, et al., 2005).

De esta forma se afirma que existen algunas diferencias por sexo, edad y condición de aplicación, que representan entre el 10 y 20 % de la escala original, analizado por factor identifica que el número 4 señalado como **Sagacidad**, es la más sensible a distorsión, que hace suponer que el contenido mismo presenta en su respuesta la prevención de una RSD, por lo que se debe considerar su supresión de la escala o bien debe modificarse para poder ajustar esta respuesta.

Por otro lado el factor 3 sigue siendo estable a todos los estudios, la cual podría emplearse incluso de manera individual su resultado para determinar el grado de AE que puede presentar una persona.

Estudio 2

El presente estudio tiene la finalidad de determinar la consistencia de los factores obtenidos para la EPDDS mediante un AFC.

Instrumento

Se utilizó la versión obtenida en el estudio 1 que consta de 28 reactivos en un formato tipo Likert con cinco intervalos de respuesta donde 17 reactivos se encuentran en sentido inverso (Ver Anexo 2).

Participantes

La muestra estuvo formada por 585 personas, 264 hombres y 317 mujeres con un rango de edad de 18 años a 44 años ($\mu=20.95$ años, $\sigma=.2.54$ años). La muestra se construyó de manera no probabilística y todos los participantes fueron voluntarios. Las aplicaciones se llevaron a cabo en centros educativos, de la zona metropolitana.

Procedimiento

El AFC, se realizó con el paquete estadístico AMOS, mediante el cual se hicieron los correspondientes análisis tomando la escala completa y por dimensión teórica. Con estos valores se determinó la consistencia de la escala, para así poder sugerir nuevos caminos para la constitución definitiva de la misma.

Resultados

La Figura 2 muestra el ajuste de los factores obtenidos en el AFE en la escala completa.

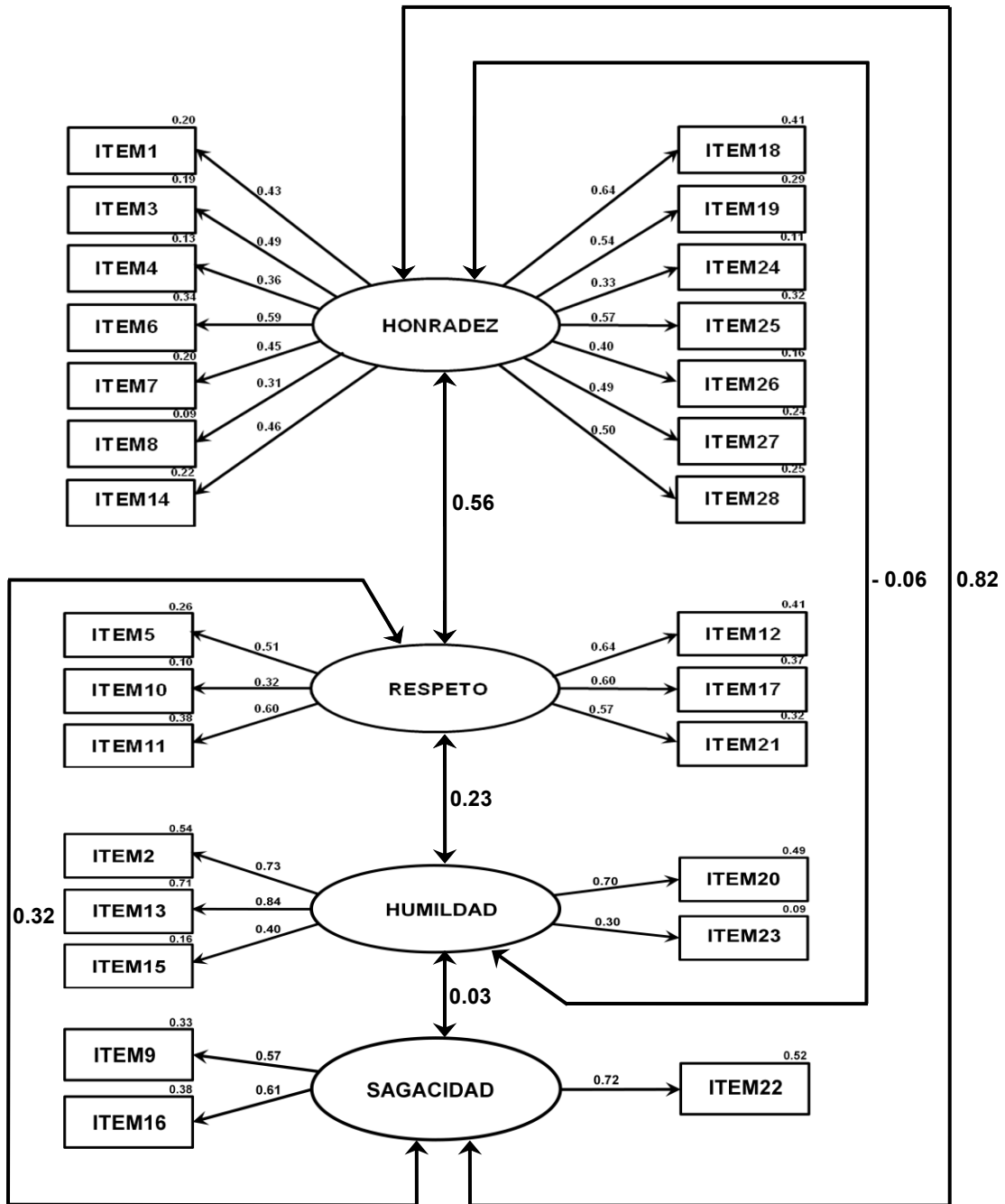


Figura 2. Coeficientes estandarizados para la EPDDS

$\chi^2=740.52$, g.l.=344 P<0.000

En la tabla 18 se presentan las correlaciones entre los cuatro factores, se observa que existen correlaciones de moderadas a altas entre las dimensiones 1, 2, y 4, mismas que se asumen forman el componente teórico MI. Asimismo se observa que la dimensión 3 no correlaciona con las dimensiones 1 y 4 y modestamente con la segunda dimensión.

Tabla 18. Correlaciones entre los factores del AFC.

Dimensión	1	2	3	4
1	1			
2	0.5	1		
3	0.06	0.23	1	
4	0.82	0.32	0.03	1

El valor de la prueba de bondad de ajuste muestra un valor alto de χ^2 y por lo mismo no ajusta el modelo propuesto. Con base en este resultado se decidió realizar AFC para cada una de las dimensiones por separado.

Los valores de carga para cada reactivo dentro del factor 1 (tabla 19), son aceptables, aunque algunos de ellos son inferiores a .40 como es el caso de los reactivos 4, 8, 24, indicando un ajuste pobre. En el factor 2 (tabla 20) el reactivo 10 tiene una carga <.40 y desajusta en el modelo al igual que el reactivo 23 en la dimensión 3 (tabla 21). Finalmente en la dimensión 4 todos los reactivos cuentan con cargas mayores de >.40 (tabla 22).

Tabla 19. Coeficientes Estandarizados para el Factor 1.

Item	β	R^2
1 Hablo sobre cosas personales de mis amistades sin c	0.45	0.20
3 He recibido cambio de más	0.49	0.19
4 Tomaría algo que no me pertenece si supiera que le	0.36	0.13
6 Me dejaría sobornar si el beneficio que recibiera	0.59	0.34
7 Robaría algo si nadie me sorprendiera.	0.45	0.20
8 Consumiría alguna droga si nadie me juzgara por	0.30	0.09
14 He hablado mal de mis amigos sin que ellos lo sepa	0.47	0.22
18 Miento cuando me beneficia.	0.64	0.41
19 Ante ciertos problemas, mi salida más cómoda es sc	0.54	0.29
24 Demuestro mi educación sólo en lugares donde se	0.33	0.11
25 He evitado devolver algo que no me pertenece	0.57	0.32
26 Le desearía el mal a otra persona si estuviera	0.40	0.16
27 Suelo juzgar a las personas más por su físico que	0.49	0.24
28 He cobrado de más por un servicio sin merecerlo.	0.50	0.25

Tabla 20. Coeficientes Estandarizados para el Factor 2.

	β	R^2
5 Respeto a la gente sin importar su condición social.	0.51	0.26
10 Si alguien me contara un secreto, sería muy difícil	0.32	0.10
11 Cuando veo que a alguien se le caen las cosas que	0.60	0.36
12 Soy amable con todas las personas, sin importar su	0.64	0.41
17 Ante cualquier situación, estoy dispuesto a ayudar a	0.60	0.37
21 Respeto a quienes me rodean sin importar su punto	0.57	0.32

Tabla 21. Coeficientes Estandarizados para el Factor 3.

	β	R^2
2 Perdono fácilmente a quienes me ofenden.	0.73	0.54
13 Olvido fácilmente las ofensas que me hacen.	0.84	0.71
15 Constantemente busco reconciliarme con mis enemi	0.40	0.16
20 Olvido las fallas de otros aunque me lastimen much	0.49	0.49
23 Aunque yo esté triste, soy feliz si los demás lo son.	0.30	0.09

Tabla 22. Coeficientes Estandarizados para el Factor 4.

	β	R^2
9 Omitiría la verdad si me fuera a perjudicar.	0.57	0.33
16 Digo mentiras si sé que no me van a descubrir.	0.61	0.38
22 Miento si es necesario.	0.72	0.52

Considerando que los cuatro factores obtenidos en el AFE, tres pertenecen a la dimensión teórica de MI y AE, se hizo un análisis por separado para determinar su ajuste, mismos que se observan en las figuras 3 y 4 respectivamente.

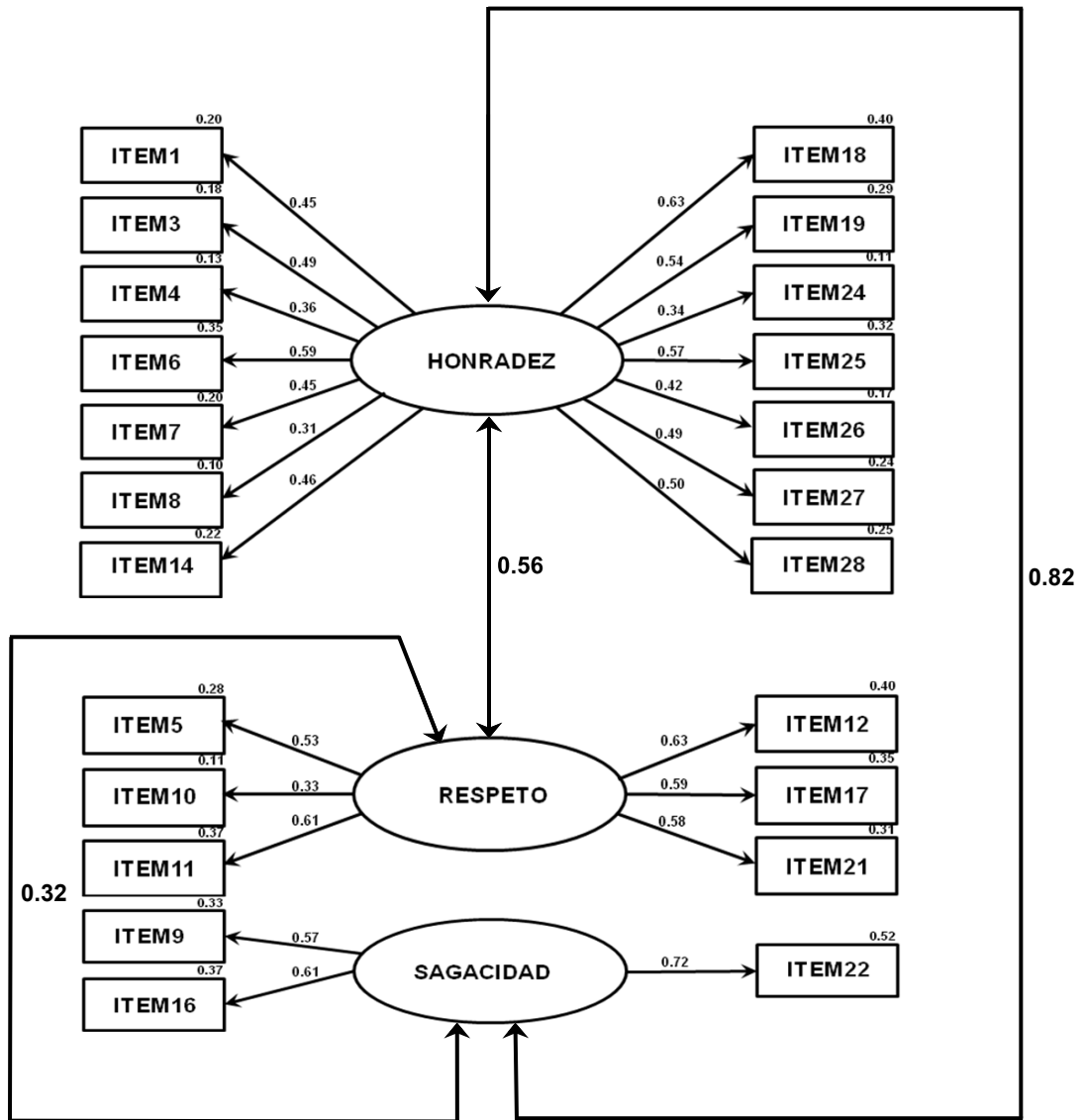


Figura 3. Coeficientes estandarizados para la dimensión MI

$\chi^2=542.326$, g.l.=227 P<0.000



Figura 4. Coeficientes estandarizados para la dimensión AE

$\chi^2=14.242$, g.l.=5 P<0.014

En este análisis el valor de bondad de ajuste es insatisfactoria, así como correlaciones bajas entre el factor de respeto y sagacidad.

Discusión

Los valores de bondad de ajuste de los AFCs llevados a cabo en este estudio nos sugieren que el modelo a cuatro factores no es suficiente. Esto puede deberse a que algunos reactivos pueden compartir variabilidad y por eso los factores de las dimensiones de MI están correlacionando de moderado a alto. Debido a que el estadístico χ^2 requiere independencia entre factores, el hecho de que existan correlaciones moderadas hace que el supuesto de varianza única es poco creíble, lo que sugiere que la teoría propuesta de la dicotomía de la DS (MI y AE) sea insuficiente para capturar toda la variabilidad inmersa en el instrumento (Batista-Fouguet, et al., 2004). Este hallazgo se puede explicar debido a que el instrumento solo captura las cuatro primeras dimensiones que arrojó el AFE, eliminando otros factores no tan claros en términos teóricos y psicométricos.

En el caso de análisis efectuado a la EPDDS, se observa que existen elementos que se deben tomar en cuenta para poder determinar si es consistente o no, ya que los valores obtenidos de bondad de ajuste a pesar de no ser valores

teóricamente aceptables pueden ser tomados como referencia para la toma de decisiones y así determinar si un reactivo será relevante dentro de la escala final.

A pesar de ser el AFC una técnica menos estricta que la TCT autores señalan que pueden presentarse desajustes por las técnicas con las que se hacen los supuestos acerca de la configuración de los reactivos, motivo por el cual se emplean otros modelos de ajuste que consideren problemas de estimación de parámetros que no se detectan de manera rutinaria, por lo que una de las finalidades del AFC no debe ser buscar únicamente un ajuste si no cuantificar el grado de desajuste del modelo (Batista-Fouquet, et al., 2004; Martínez, 1996; Skrondal & Rabe-Hesketh, 2004).

Para hacer la modificación para obtener la escala final en este caso, será necesario tomar en cuenta cada uno de los estudios anteriores, donde la Teoría Clásica del Test, la Teoría de Respuesta al Ítem, el Análisis de Funcionamiento Diferencial, Tipo de Escala de Evaluación y Condición de Aplicación del Instrumento permitan aportar nuevas formas de llegar a conclusiones que ayuden a decidir que reactivos contribuyen a la mejor conceptualización de la variable latente denominada Deseabilidad Social.

Obtención de Baremos para la EPDDS

Para obtener las normas de la prueba se obtuvieron los puntajes totales de cada dimensión factorial de la escala preliminar, así como del valor total de DS.

Se realizaron los análisis de la muestra formada por los participantes del segundo estudio, se determinaron los puntajes para la población general, así como los puntajes segmentados por sexo, inicialmente en valores estandarizados, los cuales se transformaron a puntajes T, mediante la fórmula $T=10Z + 50$ para facilitar su interpretación (Ver Anexo 3).

Para ubicar a cada uno de los sujetos en el continuo de DS, se toman los valores crudos y se ubican en la escala. Para determinar que la persona está contestando el instrumento de manera socialmente deseable es necesario tomar en cuenta aquellos sujetos que en los puntajes Z estén por arriba de 2 y por debajo de -2. Aquellos que estén del lado derecho del valor de $Z=2$ son aquellos que estarán sobreestimando sus características personales, de tal forma que se esta respondiendo de manera socialmente deseable, pueden ser personas que toman mucho en cuenta el qué dirán o las normas establecidas por la sociedad. Por el lado contrario si la persona responde con valores inferiores a $Z= -2$, la persona está respondiendo de manera socialmente indeseable, aparentando o negando características personales o defectos que pueden ser comunes en la población en general.

Discusión general

La DS ha sido estudiada ampliamente en el ámbito anglosajón (Crowne & Marlowe, 1960; Edwards, 1970; Paulhus, 1998b), lo que supondría la existencia de diferencias por cultura y sociedad (Li & Reb, 2009; Stöber, 2001). Elaborar una escala para el contexto mexicano, contribuirá a explicar si estas diferencias culturales existen, además de contribuir a la comprensión del constructo.

Las principales escalas como la de Edwards y Marlowe y Crowne, consideran a la DS como un constructo unidimensional, mientras que la escala propuesta por Paulhus en el BIDR establece 2 dimensiones.

La EPDDS se desarrolló considerando las conductas y actitudes que son percibidas como aceptadas o desaprobadas para el contexto mexicano (Domínguez-Espinosa, 2008; Domínguez-Espinosa, et al., 2010; González-Ruíz, 2010).

Los reactivos fueron redactados tomando en cuenta la teoría propuesta por Paulhus (1998b) que asume que la DS está formada por dos dimensiones MI y AE, estos fueron equilibrados en número, así como en su sentido de calificación en la versión original, formada por 94 reactivos.

El AFE de la EPDDS corroboró que el constructo no es unidimensional, en su primera prueba mostró más de 8 factores de los cuales se descartaron cuatro de acuerdo al escaso ajuste en sus valores de acuerdo a los estándares establecidos en la construcción de instrumentos.

En este sentido los cuatro factores que se obtuvieron empleando la base de reactivos desarrollada por Domínguez y col.(2010), permitió aceptar la propuesta

de Paulhus (2002) que considera que la dimensión de MI a su interior tiene otros factores que están asociados a respuestas moralistas, versus respuestas egoístas en la dimensión del AE.

Los factores denominados en la EPDDS como **Honradez** y **Respeto** son valores que los mexicanos aprecian y consideran que son deseables dentro del comportamiento de las personas, mientras que la **Sagacidad** es una virtud que se debe desarrollar, que explica la valoración que se hace del ingenio y la astucia ante cualquier situación, todas estas características corroboran una parte fundamental de la moral de una persona, y contribuyen a la explicación del MI de Paulhus.

De la misma forma el factor denominado **Humildad** que se obtuvo de la construcción de los reactivos considerados como AE, coincidieron en la asignación inicial como en la carga factorial, de esta forma el contenido de los reactivos permite establecer que el AE, se puede explicar por medio de las actitudes de perdón, olvido y reconciliación consideradas en la sociedad mexicana como socialmente deseables, de ahí que algunas de las expresiones o dichos populares, vayan en este sentido.

Por otra parte la EPDDS se sometió a la manipulación experimental con la finalidad de establecer que tan sensible a la distorsión es ante situaciones de demanda específicas, el incremento en los puntajes de MI ante una demanda positiva mostró diferencias ($M = 86.64$, $DE = 22.77$) con respecto a la condición normal ($M = 75.63$, $DE = 21.02$), lo cual era esperado ya que los valores de DS al existir esta distorsión elevarían los valores totales de esta dimensión teórica (Dunnett, et al., 1981; Holden, 2008; Paulhus, 1998b; Stöber, et al., 2002), por el contrario a los valores de AE. Esta manipulación permitió determinar la validez interna de la EPDDS.

Kerlinger y Lee (2002) establecen que para que un instrumento tenga un valor de validez debe alcanzar valores por arriba del 70% de varianza explicada, en el caso de la EPDDS, la varianza es del 56.24% que es un valor aceptable si tomamos en cuenta que la DS tiene una dificultad teórica ya que a pesar de que se ha investigado y trabajado durante muchos años no tiene una sola explicación, si no por el contrario se está reconceptualizando el constructo (Uziel, 2010).

Las manipulaciones experimentales se han empleado una vez que el instrumento está desarrollado, en esta investigación se utilizó para la construcción misma de la escala, permitiendo corroborar como se mencionó anteriormente que los reactivos que forman la solución preliminar es estable en ambas situaciones, las soluciones factoriales que mostró la EPDDS aunque fueron diferentes en su configuración, comparten un 60% de los reactivos.

Uno de los elementos importantes en el desarrollo de la EPDDS fue aquella que involucra el tipo de formato de calificación del instrumento. Las principales escalas que se emplean actualmente como la de Marlowe y Crowne (1960) emplea el formato dicotómico, mientras que la escala de Paulhus (1998b), propone que ambas calificaciones pueden estimar la DS en su instrumento.

La EPDDS se trabajó con ambos formatos para determinar cuál era el comportamiento de cada una de las escalas y respondía mejor a ambas situaciones, los valores que se obtuvieron mostró que el formato tipo Likert presenta mayor confiabilidad, aunado a la manipulación experimental que proporciona consistencia interna al instrumento, incrementando los valores de α de Cronbach que era lo esperado (Cervellione, et al., 2009; Greenwald & O'Connell, 1970; Stöber, et al., 2002), esta situación se explica por el efecto que tiene la manipulación sobre los puntajes totales en cada dimensión, ya que está

forzando al individuo a responder en un sentido, que en el caso de no presentarse este incremento hablaría de inconsistencia.

Una de las limitaciones que tiene la construcción de instrumentos desarrollados por la TCT, es la dependencia de los puntajes totales obtenidos al aplicar el instrumento para establecer el grado de atributo que presenta un individuo, razón por la cual se empleó la TRI como un modelo que ayudaría a obtener parámetros de ajuste y dificultad en función de cada reactivo, que complementarían los resultados obtenidos de validez y confiabilidad, además de identificar por medio del DIF si existen sesgos de probabilidad en la EPDDS.

El análisis de la TRI seleccionado fue el de un parámetro, también conocido como modelo de Rasch, este mostró que los valores de dificultad que es uno de los parámetros de mayor interés, están muy cercanos al centro de la escala entre 0.36 y -0.80 cuando se le considera unidimensional, la cantidad de atributo para contestar en la dirección de la RSD se encuentra en un rango pequeño, estudios previos muestran valores similares (Cervellione, et al., 2009), con lo que podemos aceptar estos índices dentro del rango que se está empleando en el BIDR.

Al analizar la escala completa como unidimensional muestra 7 reactivos que desajustan en los valores de INFIT y OUTFIT, 8 reactivos diferentes a estos muestran DIF por sexo, condición de aplicación o edad, y 4 que repiten su nivel de dificultad que podrían ser susceptibles de ser eliminados, de tomar esta decisión se contaría con 9 reactivos para la escala de DS, este criterio al ser tan restrictivo aunado al dato de varianza no explicada que está arriba de 5.1 no puede ser criterio de selección, por lo que se consideró como un análisis exploratorio, que pone en evidencia que no es unidimensional por lo cual se realizó un segundo

análisis de las dos dimensiones teóricas donde se agruparon los factores 1, 2 y 4 al MI, y el factor 3 a la dimensión de AE.

Al analizar la dimensión de MI se identificó que de los 23 reactivos que forman la dimensión, 10 de ellos presentan DIF, 2 distintos a estos presentan el mismo nivel de dificultad, y 2 que desajustan en sus valores de INFIT y OUTFIT, al suprimirlos quedaría la dimensión de MI de 9 reactivos, en un intervalo de dificultad de 0.27 hasta -0.55, si se considera que algunas de las escalas de DS son pequeñas (Greenwald & Satow, 1970), esta podría ser una posibilidad que deberá ser analizada mediante AFC, los reactivos seleccionados serían el 11, 64, 9, 17, 65, 32, 31, 76, 68.

Por otro lado tanto el AFE, como el análisis de Rasch muestran que el factor 3 denominado **Humildad**, ajusta en su estructura de manera adecuada en ambos análisis, además de no presentar DIF en ninguno de sus reactivos, los 5 reactivos (49, 66, 75, 47, 10) pueden ser empleados con la propuesta anterior y definir una escala con 14 reactivos.

Debido al alcance que se pretendía en este trabajo no se exploró esta posibilidad por lo que queda como sugerencia para establecer la escala final de DS, en estudios posteriores.

Cabe señalar que el análisis DIF como un parámetro de la TRI, mostró diferencias en algunos reactivos por sexo, esto puede ser explicado por las diferencias de género, donde algunas de las actitudes socialmente aceptadas están en función de creencias, hábitos y costumbres (González-Ruíz, 2010).

Por medio del análisis de Rasch se observó que el factor cuatro **Sagacidad** formado por tres reactivos, tienen un ajuste adecuado de INFIT y OUTFIT, pero en el análisis DIF los tres son susceptibles de sesgo, de tal manera que este factor muestra diferencias en la probabilidad de respuesta por sexo, edad y condición, que lo hace susceptible de eliminación de la escala final.

El factor dos **Respeto** está formado por 6 reactivos, de los cuales 3 muestran DIF por alguna de las condiciones, lo que hace que la mitad de este factor se sugiera eliminar para así evaluar si el ajuste mejora.

Es importante señalar que el DIF muestra que 10 reactivos presentan diferencias significativas de la escala de 28, de esta forma se puede inferir que no existen diferencias por sexo, edad o situación de demanda, y que la DS puede ser estimada con cierta holgura entre los diferentes grupos de la población mexicana.

Para determinar si la EPDDS tiene la configuración mostrada en el AFE, se realizó un AFC con una muestra diferente a la original, el análisis efectuado mostró que en su conjunto el modelo no es consistente a los valores de la prueba de bondad de ajuste de χ^2 que era esperado por los factores al interior de la escala de acuerdo a estudios similares (ej. Loo & Thorpe, 2000), que analizaron la escala de Marlowe y Crowne reportando valores altos de χ^2 estadísticamente significativos, indicando diferencias entre los modelos.

El análisis que se efectuó por dimensión teórica muestra desajustes más grandes en la dimensión de MI, la correlación entre los factores 1 y 4 es alta, mientras que el factor 2 muestra valores de correlación bajos entre las dimensiones 1 y 4, sería necesario probar suprimirlo y ver el ajuste.

La dimensión teórica de AE muestra un ajuste más cercano al teórico, solo uno de los reactivos está en un valor límite que se deberá probar su eliminación para determinar si ajusta de mejor manera.

Los valores de los puntajes en los diferentes grupos a los que se aplicó la escala, determinan que se puede ubicar a los participantes en un continuo de DS, indicando que aquellos puntajes estandarizados obtenidos por arriba de 1 están respondiendo con Deseabilidad Social, mientras que valores por debajo del valor de -1 están respondiendo de manera Socialmente Indeseable.

Cuando se habla de Deseabilidad Social, no se puede asegurar que esta se debe a una distorsión intencionada de la persona, hay que considerar que están presentes una serie de factores que hacen difícil su evaluación y por ende su estimación. Sus elementos no dependen exclusivamente de la psicometría sino de actitudes y valores, por un lado cuenta con características de *Rasgo* que dificultan el poder estimar si la persona contesta con un grado de conciencia y convencimiento de lo que piensa y cree, o bien responde como parte del componente de *Estado* que hace referencia a la necesidad de ser aceptado y sentirse adaptado para poder subsistir en un sistema social demandante de ciertas conductas y comportamientos, que en ocasiones no son propias, pero necesarias para ser incluido en la sociedad.

En el caso de la EPDDS se observa que el contenido de los reactivos debe analizarse con más detalle para poder determinar cuáles de ellos, son conductas que caracterizan a la sociedad mexicana y que no son propiamente distorsión, si no parte de la idiosincrasia que caracteriza a la población, de esta forma no solo se contribuye a la psicometría con un instrumento que escala la DS, si no que a su vez estaría desarrollando la comprensión de la Psicología del Mexicano.

Con todos estos estudios, se puede asegurar que la EPDDS es un instrumento que en términos del análisis psicométrico por TCT y TRI, muestra ajustes adecuados, que requiere mediante análisis de validez concurrente y externa, con otros instrumentos como sería en el BIDR o la escala Marlowe Crowne en el primer caso, mientras que la administración con el NEO-FFI para el segundo, que permitan comprobar su comportamiento y conocer más la escala a partir de los parámetros obtenidos.

Conclusiones

Los hallazgos de la presente investigación contribuyen a la explicación teórica de la existencia de factores al interior de las dimensiones aceptadas por Paulhus. Estos factores que fueron denominados como Honradez, Respeto y Sagacidad en el caso de la dimensión de MI, explican valores que están presentes en el estilo de respuesta de los mexicanos. Por otra parte la virtud de Humildad, es una característica que se aprecia como un ideal dentro de la dimensión de AE. De la misma manera pone de manifiesto esta investigación, que los factores al interior de la dimensión teórica de construcción, deben explorarse por las diferencias que se muestran por edad, sexo y situación de demanda.

La TCT mostró ser una guía efectiva en la configuración de la EPDDS, que se complementó con los estudios de la TRI, incluyendo el DIF y la manipulación experimental, proporcionando información que triangularía la validez de un reactivo.

El modelo de un parámetro de Rasch, fue suficiente para determinar la dificultad de los reactivos, corroborando que el constructo presenta dificultad conceptual que lleva a la construcción de reactivos muy sencillos, y otros con baja dificultad que ponen en evidencia la necesidad de seguir buscando diferentes formas de comprender el constructo, esto en gran parte a la posibilidad de considerar a la DS como un estilo de respuesta o como un rasgo de personalidad.

La EPDDS no tuvo como finalidad hacer la distinción entre estos dos contenidos, sin embargo puede dar una explicación en términos de los puntajes que se obtengan (Paulhus, 1984), puntajes altos en la escala de MI hablarían del estado, puntajes altos en AE se hablaría del rasgo.

La manipulación experimental en este trabajo mostró ser un indicador de consistencia interna de la EPDDS, la cual determinó los reactivos susceptibles de distorsión y que presentarían poca o nula discriminación entre un grupo y otro de aplicación. Este indicador probó ser efectivo en la construcción de la escala con lo cual su empleo en el diseño de instrumentos garantizará que las distorsiones podrán a su vez ser previstas y estimadas.

La EPDDS mostró en su análisis DIF que las diferencias por sexo, edad y condición de aplicación, solo fueron determinadas en un 33% de los reactivos, con lo que se puede asegurar que es un instrumento balanceado para la población de jóvenes y adultos mexicanos.

Todos estos elementos al integrarse permiten tener un modelo rico en explicaciones metodológicas, que permiten disminuir los errores tanto de construcción, aplicación e interpretación del escalamiento de la DS.

De manera práctica la EPDDS es un instrumento válido y confiable, que reúne las características de los instrumentos de DS, corto y conciso, con lo que su aplicación de forma simultánea a otros instrumentos permitirá encontrar correlaciones entre los puntajes de los instrumentos que ayuden a la comprensión de la persona.

Su construcción al ser preliminar deberá contar con la integración de las observaciones señaladas en la presente investigación, de tal manera que las explicaciones que se generarán serán objetivas y tomaran diferentes criterios que enriquecerán el instrumento final.

Limitaciones

Debido a que los resultados presentados se basan en muestras no probabilísticas, los resultados son generalizables solamente a muestras con características similares. El hecho de haberse realizado en su mayoría en instituciones de educación superior sugiere un sesgo en el nivel educativo de los participantes.

La propuesta que se sugiere a partir de la TRI y el DIF no se realizó por no ser parte del alcance de este trabajo, por lo cual en trabajos posteriores se deberá realizar un AFC para determinar el ajuste y corroborar si es una solución factible.

El AFC debe realizarse con otros criterios que permitan determinar si su ajuste es el adecuado. Habrá que aumentar el tamaño de muestra para identificar una solución más estable así como obtener baremos más generales.

Sugerencias

Para poder determinar algunas otras propiedades de la EPDDS es necesario realizar los análisis de correlación con otras escalas similares de DS que permitan obtener su validez concurrente. Asimismo será interesante aplicar la escala en situaciones reales de demanda social como por ejemplo a aspirantes de empleo. Adicionalmente será indispensable realizar estudios con la escala y autorreportes de personalidad en donde se ponga a prueba su capacidad predictiva.

Se deberá realizar aplicaciones a otras culturas semejantes a la mexicana y algunas opuestas completamente que permitan determinar su Operacionalización.

Referencias

- Allport, G. W. (1965). *La personalidad su configuración y desarrollo*. España: Herder.
- Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *The Academy of Management Journal*, 24(2), 377-385.
- Azrin, N. H., Holz, W., & Goldiamond, I. (1961). Response bias in questionnaire reports. *Journal of Consulting Psychology*, 25(3), 324-326.
- Batista-Fouquet, J., Coenders, G., & Alonso, J. (2004). Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Medicina Clínica*, 122(Suplemento1), 21-27.
- Borkenau, P., & Ostendorf, F. (1992). Social Desirability Scales as Moderator and Suppressor Variables. *European Journal of Personality*, 6, 199-214.
- Castro, M. E., Maya, M. A., & Orozco, C. (1986). Normas y estructura factorial de dos escalas: La Escala de Responsabilidad Intelectual Académica y la Escala de Necesidad de Aprobación Social. *Salud Mental*, 9(3), 65-71.
- Catell, H. E. P., & Mead, A. D. (2008). The Sixteen Personality Factor Questionnaire (16PF). In S. P. Ltd (Ed.), *The Sage Handbook of Personality Theory and Assessment* (Vol. 2). NY: Sage.
- Caudillo, C., & Cerna, M. A. (2007). *Sexualidad y vida humana*. México: Universidad Iberoamericana.
- Cervellione, K. L., Lee, Y., & Bonanno, G. A. (2009). Rasch Modeling of the Self-Deception Scale of the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, 69(3), 438-458.
- Coartada de Kohan, N. (2004). Teoría de Respuesta al Ítem: Supuestos Básicos. *Evaluar*, 4(102), 95-110.
- Collazo, A. A. (2005). Translation of the Marlowe-Crowne Social Desirability Scale into equivalent Spanish version. *Educational and Psychological Measurement*, 65(5), 780-806.
- Cosentino, A. C., & Castro-Solano, A. (2008). Adaptación y validación Argentina de la Marlowe y Crowne Social. *Interdisciplinaria*, 25(2), 197-216.
- Costa, P. T., & McCrae, R. R. (1985). *The NEO-PI-R Personality Inventory Manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475-494.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349-354.
- Darnon, C., Dompnier, B., Delmas, F., Pulfrey, C., & Butera, F. (2009). Achievement Goal Promotion at University: Social Desirability and Social Utility of Mastery and Performance Goals. *Journal of Personality and Social Psychology*, 96(1), 119-134.
- De Jong, M. G., Pieters, R., & Fox, J. P. (2010). Reducing Social Desirability Bias Through Item Randomized Response: An Application to Measure Underreported Desires. *Journal of Marketing Research*, 47(1), 14-27.

- Diekman, A. B., & Goodfriend, W. (2006). Rolling with the changes: a role congruity perspective on gender norms. *Psychology of Women Quarterly*, 30(4), 369-383.
- Domínguez-Espinosa, A. (1997). *Construcción de una escala de deseabilidad social en México*. Unpublished Tesis de Licenciatura, Universidad Nacional Autónoma de México, Mexico.
- Domínguez-Espinosa, A. (2008). Evaluación de la Deseabilidad Social y su adaptación teórica al ambiente mexicano. Convocatoria de Investigación Científica Básica 2007 (pp. 1-30). Distrito Federal, México: Consejo Nacional de Ciencia y Tecnología.
- Domínguez-Espinosa, A., Navarro Contreras, G., García Campos, T., Gutiérrez Ferros, E., & Sada Monroy, A. (2010). Operacionalización Etnocultural de la Deseabilidad Social. In S. Rivera-Aragón, R. Díaz-Loving, I. Reyes-Lagunes, R. Sánchez-Aragón & L. M. Cruz-Martínez (Eds.), *La Psicología Social en México* (Vol. 13, pp. 67-74). México: UNAM, AMEPSO, UNISON.
- Domínguez-Espinosa, A., Salas-Menotti, I., & Reyes-Lagunes, I. (2008). Validez Concurrente de la Escala de Deseabilidad Social de Domínguez utilizando la Escala de Deseabilidad Social de Marlowe-Crowne. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 25(1), 125-139.
- Domínguez-Espinosa, A., Salas-Monotti, I., & Reyes-Lagunes, I. (2008). Validez Concurrente de la Escala de Deseabilidad Social de Domínguez utilizando la Escala de Deseabilidad Social de Marlowe-Crowe. *RIDEP*, 1(25), 125-139.
- Domínguez, A. (1997). *Construcción de una escala de deseabilidad social en México*. Unpublished Tesis de Licenciatura, Universidad Nacional Autónoma de México, Mexico.
- Domínguez, A., Navarro, G., García, T., Gutierrez, E., & Sada, A. (2010). *Operacionalización Etnocultural de la Deseabilidad Social* (Vol. XIII). México: AMEPSO.
- Domínguez, A., Salas, I., & Contreras, C. (2010). *Comparación Transcultural de la Deseabilidad Social ¿Variaciones Culturales o Efectos del Instrumento?* (Vol. XIII). México: AMEPSO.
- Domínguez, A., Salas, I., & Reyes-Lagunes, I. (2008). Validez Concurrente de la Escala de Deseabilidad Social de Domínguez utilizando la Escala de Deseabilidad Social de Marlowe-Crowne. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 25(1), 125-139.
- Domínguez Espinosa, A., Salas Menotti, I., & Contreras Bravo, C. (2010). Comparación transcultural de la deseabilidad social, ¿variaciones culturales o efectos del instrumento? In S. Rivera-Aragón, R. Díaz-Loving, I. Reyes-Lagunes, R. Sánchez-Aragón & L. M. Cruz-Martínez (Eds.), *La Psicología Social en México* (Vol. 13, pp. 59-65). México: UNAM, AMEPSO, UNISON.
- Dompnier, B., Pansu, P., & Bressoux, P. (2007). Social utility, social desirability and scholastic judgments: Toward a personological model of academic evaluation. *European Journal Of Psychology of Education*, 12(2), 333-350.
- Downey, L. A., Godfrey, J.-L., Hansen, K., & Stough, C. (2006). The Impact of Social Desirability and Expectation of Feedback on Emotional Intelligence in the Workplace. *Journal of Applied Psychology*, 2(2), 12-18.

- Dunnett, S., Koun, S., & Barber, P. J. (1981). Social desirability in the Eysenck Personality Inventory. *British Journal of Psychology*, 72, 19-26.
- Edwards, A. L. (1957). *Techniques of Attitude Scale Construction*. New York: Appleton- Century- Crofts Inc.
- Edwards, A. L. (1970). *The Measurement of Personality Traits by Scales and Inventories*. New York: Holt, Rinehart and Winston, Inc.
- Edwards, A. L., & Walsh, J. A. (1964). Response sets in standard and experimental personality scales. *American Educational Research Journal*, 1(1), 52-61.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social Desirability Corrections in Personality Measurement: Issues of Applicant Comparison and Construct Validity. *Journal of Applied Psychology*, 84(2), 155-166.
- Eysenck, H. J., & Eysenck, S. B. (1964). *The manual of the Eysenck Personality Inventory*. London: University of London Press.
- Ferrando, P. J., & Chico, E. (2000). Adaptación y análisis psicométrico de la Escala de Deseabilidad Social de Marlowe y Crowne. *Psicothema*, 12(3), 383-389.
- Fordyce, W. E. (1956). Social Desirability in the MMPI. *Journal of Consulting Psychology*, 20(3), 171-175.
- Ganster, D. C., Hennesey, H. W., & Luthans, F. (1983). Social desirability response effects: Three alternative models. *Academy of Management Journal*, 26(2), 321-331.
- González-Ruiz, G. A. (2010). *Representación Social de lo Deseable y lo Indeseable para el Contexto Mexicano, desde un análisis de género*. Unpublished Tesis de Doctorado en Investigación Psicológica, Universidad Iberoamericana, México, D.F.
- González, M. J. (2008). *El análisis de reactivos con el Modelo Rasch Manual Técnico*. Mexico: INNE, UNISON.
- Greenwald, H. J., & Clausen, J. D. (1970). Test of relationship between yeasaying and social desirability. *Psychological Reports*, 27, 139-141.
- Greenwald, H. J., & O'Connell, S. M. (1970). Comparison of dichotomous and likert formats. *Psychological Reports*, 27, 481-482.
- Greenwald, H. J., & Satow, Y. (1970). A short social desirability scale. *Psychological Reports*, 27, 131-135.
- Henry, M. S., & Raju, N. S. (2006). The effects of traited and situational impression management on a personality test: an empirical analysis. *Psychology Science*, 48(3), 247-267.
- Hersh, R. H., Paolitto, D. P., & Reimer, J. (2002). *El Crecimiento Moral de Piaget a Kolberg*. Madrid: Narcea, S.A.
- Hofstede, G. (1980). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Beverly Hills, CA: Sage Publications.
- Hofstee, W. K. B., & Hendriks, A. A. J. (1997). The use of score anchored at the scale midpoint in reporting individual's traits. *European Journal of Personality*, 12, 219-228.

- Holden, R. R. (2007). Socially Desirable responding Does Moderate Personality Scale Validity Both in Experimental and in Nonexperimental Contexts. *Canadian Journal of Behavioral Science*, 39(3), 184-201.
- Holden, R. R. (2008). Underestimating the effects of faking on the validity of self-report personality scales. *Personality & Individual Differences*, 44(1), 311-321.
- Hopwood, C. J., Talbert, C. A., Morey, L. C., & Rogers, R. (2008). Testing the Incremental Utility of the Negative Impression-Positive Impression Differential in Detecting Simulated Personality Assessment Inventory Profiles. *Journal of Clinical Psychology*, 64(3), 338-343.
- Johannesen-Schmidt, M. C., & Eagly, A. H. (2002). Another Look at Sex Differences in Preferred Mate Characteristics: The Effects of Endorsing the Traditional Female Gender Role. *Psychology of Women Quarterly*, 26(4), 322-328.
- Kerlinger, F. N., & Lee, H. B. (2002). *Investigación del Comportamiento. Métodos de Investigación en Ciencias Sociales* (4a ed.). México: Mc Graw-Hill.
- Kluchhohn, C., Murray, H. A., & Schneider, D. M. (1953). *Personality in nature, society and culture*. Nueva York: Knopf.
- Kuncel, N. R., & Tellegen, A. (2009). A conceptual and empirical reexamination of the measurement of the social desirability of items: implications for detecting desirable response style and scale development. *Personnel Psychology*, 62(2), 201-228.
- Lara-Cantú, M. A. (1990). Validez y confiabilidad de la Escala de Deseabilidad Social de Marlowe y Crowne en una población de adultos. *Salud Mental*, 13(4), 35-37.
- Lara-Cantú, M. A., & Suzan-Reed, M. (1988). La Escala de Deseabilidad Social de Marlowe y Crowne: un estudio psicométrico. *Salud Mental*, 11(3), 25-28.
- Li, A., & Reb, J. (2009). A cross-nations, cross-cultures, and cross-conditions analysis on the equivalence of the Balanced Inventory of Desirable Responding. *Journal of Cross-Cultural Psychology*, 40(2), 214-233.
- Linacre, J. M. (2009). *Winsteps (Version 3.68.0)*. Chicago: Winsteps.com.
- Lönnqvist, J.-E., Paunonen, S., Tuulio-Henriksson, A., Lönnqvist, J., & Verkasalo, M. (2007). Substance and Style in Socially Desirable Responding. *Journal of Personality*, 75(2), 291-322.
- Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Social Psychology*, 140(5), 628-635.
- Lück, H. E., & Timaeus, E. (1969). Skalen zur Messung Manifester Angst (MAS) und Soziales Wünschbarkeit (SDS-E und SDS-CM) [Scales for measurement of manifest anxiety (MAS) and social desirability (SDS-E und SDS-CM)]. *Diagnostica*, 15, 134-141.
- Martínez, A. R. (1996). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Editorial Síntesis, S.A.
- McCrae, R. R., & Costa, P. T. J. (1983). Social desirability scales: more substance than style. *Journal of Consulting and Clinical Psychology*, 51(6), 882-888.
- Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the MMPI. *Journal of Applied Psychology*, 30, 525-564.

- Megargee, I. (1971). *Métrica de la Personalidad* (Vol. 1). México: Trillas.
- Meleddu, M., & Guicciardi, M. (1998). Self-knowledge and social desirability of personality traits (Vol. 12, pp. 151-168).
- Méndez-García, M. A., & Domínguez-Espinosa, A. (2010). Análisis de la Sensibilidad ante Instrucciones Experimentales y Estándar de la Escala de Deseabilidad Social para el Contexto Mexicano. In S. Rivera-Aragón, R. Díaz-Loving, I. Reyes-Lagunes, R. Sánchez-Aragón & L. M. Cruz-Martínez (Eds.), *La Psicología Social en México* (Vol. 13). México: UNAM, AMEPSO, UNISON.
- Moon, Y. (1998). Impression Management in Computer-Based Interviews: The effects of Input Modality, Output Modality, and Distance. *Public Opinion Quarterly*, 62(4), 610-622.
- Morf, M. E., & Jackson, D. N. (1972). An analysis of two response styles: True responding and item endorsement. *Educational and Psychological Measurement*, 32(2), 329-353.
- Mueller-Hanson, R., Heggstad, E. D., & Thornton III, G. C. (2003). Faking and Selection: Considering the Use of Personality From Select-In and Select-Out Perspectives. *Journal of Applied Psychology*, 88(2), 348-355.
- Muñiz, J. (1998). La medición de lo psicológico. *Psicothema*, 10(1), 1-21.
- Nunnally, J. C., & Bernstein, I. J. (1995). *Teoría Psicométrica* (3a ed.). México: Mc Graw- Hill.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The Red Herring. *Journal of Applied Psychology*, 81(6), 660-679.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality & Social Psychology*, 46(3), 598-609.
- Paulhus, D. L. (1998a). Interpersonal and Intrapyschic Adaptiveness of Trait Self-Enhancement: A Mixed Blessing? *Journal of Personality & Social Psychology*, 74(5), 1197-1208.
- Paulhus, D. L. (1998b). *Manual of the Balanced Inventory of Desirable Responding*. Buffalo: Multi-Health Systems.
- Paulhus, D. L. (2002). Socially Desirable Responding: The Evolution of a Construct. In H. I. Braun, D. N. Jackson & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, N. J.: Erlbaum.
- Paulhus, D. L., Fridhandler, B., & Hayes, S. (1997). Psychological defense. Contemporary theory and research. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of Personality Psychology* (pp. 543-580). San Diego: Academic Press.
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality & Social Psychology*, 60, 307-317.
- Pauls, C. A., Wacker, J., & Crost, N. W. (2005). The Two Componentes of Social Desirability and their Relations to Resting Frontal Brain Asymmetry. *Journal of Individual Differences*, 26(1), 29-42.
- Pérez, H. (2004). *Física General*. México: Publicaciones Cultural.
- Raluy, P. A. (Ed.) (2004) *Diccionario Porrúa de la Lengua Española* (49a ed.). México, DF.

- Reidl-Martínez, L. M. (2007). Medición en Ciencias Sociales. Unpublished Apuntes de Clase. Universidad Nacional Autónoma de México.
- Reyes-Lagunes, I., & García y Barragan, L. F. (2008). Procedimiento de validación psicométrica culturalmente relevante: un ejemplo. In S. Rivera-Aragón, R. Díaz-Loving, R. Sánchez-Aragón & I. Reyes-Lagunes (Eds.), *La Psicología Social en México* (Vol. 12, pp. 625-630). México: AMEPSO, Universidad de Guanajuato, Universidad Nacional Autónoma de México.
- Risko, E. F., Quilty, L. C., & Oakman, J. M. (2006). Socially desirable responding on the Web: Investigating the Candor Hypothesis. *Journal of Personality Assessment*, 87(3), 269-276.
- Robins, R. W., & Paulhus, D. L. (2001). *The character of self-enhancers: Implications for organizations*. Washington, D.C.: American Psychological Association.
- Robson, S. M., Jones, A., & Abraham, J. (2008). Personality, Faking and Convergent Validity: A Warning Concerning Warning Statements. *Human Performance*, 21, 89-106.
- Rogers, R., Vitacco, M. J., Jackson, R. L., Martin, M., Collins, M., & Sewell, K. W. (2002). Faking psychopathy? An examination of response styles with antisocial youth. *Journal of Personality Assessment*, 78(1), 31-46.
- Rosenthal, R., & Rosnow, R. (1991). *Essentials of Behavioral Research: Methods and Data Analysis* (2nd ed.). Boston, MA: Mc Graw-Hill.
- Sackeim, H. A., & Gur, R. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology*, 47(1), 213-215.
- Salgado, J. F. (2005). Personalidad y Deseabilidad Social en Contextos Organizacionales: Implicaciones para la práctica de la Psicología del Trabajo y las organizaciones. *Redalyc*, 26(92), 115-128.
- Sánchez-Turet, M., & Cuadras-Avellana, C. M. (1972). Adaptación española del cuestionario EPI de Eysenck. *Anuario de psicología*, 6, 31-60.
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Interdisciplinary Statistics. Generalized Latent Variable Modeling. Multilevel, Longitudinal, and Structural Equation Models*. US: Chapman & Hall/CRC.
- Steenkamp, J. B. E. M., De Jong, M. G., & Baumgartner, H. (2010). Socially Desirable Response Tendencies in Survey Research. *Journal of Marketing Research*, 47(2), 199-214.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17). Convergent Validity, Discriminant Validity, and Relationship with Age. *European Journal of Psychological Assessment*, 17(3), 222-232.
- Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing Continuous and Dichotomous Scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78(2), 370-389.
- Terracciano, A., Abdel-Khalek, A. M., Adam, N., Adamovova, L., Ahn, C.-k., Ahn, H.-n., et al. (2005). National Character Does Not Reflect Mean Personality Trait Levels in 49 Cultures. *Science*, 310(5745), 96-100.
- Thorndike, R. L. (1951). Reliability. In E. L. Lindquist (Ed.), *Educational Measurement* (pp. 560-620). Washington, D. C.: American Council on Education.

- Transparency International. (2010). The 2010 Transparency International Corruption Perceptions Index. from http://www.transparency.org/policy_research/surveys_indices/cpi/2010/results
- Uziel, L. (2010). Rethinking Social Desirability Scales: From Impression Management to Interpersonally Oriented Self-Control. *Perspectives on Psychological Science*, 5(3), 243–262.
- Wiggins, J. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology*, 23(5), 419-427.
- Ziegler, M. (2009). Modeling Socially Desirable Responding and Its Effects. *Educational and Psychological Measurement*, 69(4), 548-565.
- Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Anexo 1. Banco de Reactivos para la Escala de Deseabilidad Social

		MI	AE
1	Busco tener la mejor...		X
2	Si me topo con un ...	X	
3	Prefiero la productividad...	X	
4	Mis buenos modales...		X
5	Suelo corresponder el cariño...		X
6	Muchas veces hago promesas ...	X	
7	Sigo las reglas aunque...		X
8	Aún estando triste mi ...		X
9	Hablo sobre cosas personales ...	X	
10	Perdono fácilmente...		X
11	He recibido cambio ...	X	
12	En mi trabajo, pierdo ...	X	
13	Si encuentro algo valioso ...	X	
14	Cuando trabajo me interesa ...	X	
15	Al tener una pelea...	X	
16	Prefiero una multa ...	X	
17	Tomaría algo que no ...	X	
18	Prefiero el bien común ...	X	
19	Soy tolerante aunque los...		X
20	Si descompongo algo ...	X	
21	Soy enteramente sincero...		X
22	Estoy tan feliz con mi vida ...		X

23	Respeto a la gente sin importar...		X
24	Me molesta la gente que ...		X
25	Hago mi mejor esfuerzo hasta ...		X
26	Me dejaría sobornar si el beneficio ...	X	
27	Soy más amable cuando me interesa ...	X	
28	Admito muy fácilmente mis...	X	
29	Me disgusta que los demás ...		X
30	Me he enojado exageradamente...		X
31	Robaría algo si nadie me ...	X	
32	Consumiría alguna droga si...	X	
33	Le he faltado al respeto a ...		X
34	Cuando realizo algo ...		X
35	Omitiría la verdad si me...	X	
36	Antes que yo, están ...		X
37	Miento si es necesario.	X	
38	Digo groserías estando ...	X	
39	Digo la verdad cueste...	X	
40	Me comporto de igual manera ...	X	
41	Si alguien me contara un ...	X	
42	Cuando veo que a alguien ...	X	
43	Soy amable con todas las personas...	X	
44	Siento amor por todos mis...		X
45	Realmente le tengo miedo al...	X	
46	Alguna vez he realizado las ...	X	
47	Olvido fácilmente las ofensas...		X

48	He hablado mal de mis amigos ...	X	
49	Constantemente busco ...		X
50	Soy responsable aun en las...		X
51	He aceptado el crédito por ...		X
52	He sentido odio hacia alguna ...	X	
53	Digo mentiras si sé que no me...	X	
54	Suelo desear los bienes...		X
55	Difícilmente rompo.		X
56	Conservo la compostura ...		X
57	Si un amigo tiene una tarea...		X
58	Tengo un verdadero interés ...	X	
59	Me afecta lo que otros piensan...		X
60	Ante cualquier situación, estoy...	X	
61	Si me es posible, realizo solamente ...	X	
62	Respeto a la gente aun cuando ella no...		X
63	Los demás reconocen todas mis ideas...		X
64	Miento cuando me...	X	
65	Ante ciertos problemas...	X	
66	Olvido las fallas de otros ...		X
67	Soy amable a pesar de ...		X
68	Respeto a quienes me rodean ...		X
69	Las críticas a mi persona, aunque...		X
70	Soy muy hábil para ponerme en los...		X
71	Le haría cualquier favor...		X
72	Sólo realizo actividades ...		X

73	Prefiero ocuparme en algo...		X
74	Soy igualmente educado con...	X	
75	Aunque yo esté triste, soy feliz...		X
76	Demuestro mi educación sólo ...	X	
77	Me alegro por el éxito de los ...		X
78	Al platicar ciertas anécdotas...		X
79	Doy las gracias cada vez que ...	X	
80	Trabajo menos cuando sé que ...	X	
81	Llego a tiempo a mis reuniones ...		X
82	He evitado devolver algo que no ...	X	
83	Solamente me enojo por cosas...	X	
84	Le desearía el mal a otra ...		X
85	Suelo juzgar a las personas ...		X
86	Acepto con gusto las decisiones...	X	
87	He demostrado cariño por alguien ...	X	
88	Reconozco sin mayor problema cuando ...		X
89	He hecho trampa para obtener ...	X	
90	A veces fallo en mis ...		X
91	He cobrado de más por...	X	
92	Analizo a profundidad todas...		X
93	Soy un buen...		X
94	Sólo en situaciones extremas...		X

Nota: Para consultar el banco de reactivos completo, dirigirse con la Dra. Alejandra del Carmen Domínguez Espinosa, correo: alejandra.dominguez@uia.mx

Anexo 2. EPDDS.

INSTRUCCIONES

Por favor, lea cuidadosamente todas las instrucciones antes de comenzar. Este inventario consta de frases. Lea cada frase con atención y marque la casilla que mejor refleje su opinión. Asegúrese de que su respuesta se encuentre en la casilla correcta.

Marque:

- TD** si la frase es completamente falsa o si está en **total desacuerdo** con ella.
- D** si la frase es falsa en su mayor parte o si está en **desacuerdo** con ella.
- N** si la frase es tan cierta como falsa, **ni sí, ni no neutral** en relación con ella.
- A** si la frase es generalmente cierta o si está **de acuerdo** con ella.
- TA** si la frase es completamente cierta o si está **totalmente de acuerdo** con ella.

Dé sólo una respuesta para cada frase. Responda a todos los reactivos y asegúrese de haber marcado una X en la casilla con la respuesta correcta. No existen respuestas correctas ni incorrectas, y no se necesita ser un experto para contestar este Cuestionario. Responda de forma sincera y exprese sus opiniones de la manera más precisa posible.

TD En total desacuerdo

D En desacuerdo

N Ni sí, ni no

A De acuerdo

TA Totalmente de acuerdo

Por favor responda a todos los enunciados. ¡Gracias por su cooperación!

		TD	D	N	A	TA
1	Hablo sobre cosas personales ...					
2	Perdono fácilmente a ...					
3	He recibido cambio de...					
4	Tomaría algo que no me...					

5	Respeto a la gente sin ...					
6	Me dejaría sobornar si ...					
7	Robaría algo si nadie me...					
8	Consumiría alguna droga ...					
9	Omitiría la verdad si me ...					
10	Si alguien me contara un...					
11	Cuando veo que a alguien...					
12	Soy amable con todas las ...					
13	Olvido fácilmente las ...					
14	He hablado mal de mis...					
15	Constantemente busco ...					
16	Digo mentiras si sé que ...					
17	Ante cualquier situación...					
18	Miento cuando me...					
19	Ante ciertos problemas...					
20	Olvido las fallas de otros ...					
21	Respeto a quienes me ...					
22	Miento si ...					
23	Aunque yo ...					
24	Demuestro ...					
25	He evitado devolver ...					
26	Le desearía el mal a ...					
27	Suelo juzgar a las ...					
28	He cobrado de más ...					

Nota: Para consultar la escala completa, dirigirse con la Dra. Alejandra del Carmen Domínguez Espinosa, correo: alejandra.dominguez@uia.mx

Anexo 3. Baremos para la EPDDS.

Puntuaciones de la escala de Deseabilidad Social para Varones															
Percentil	Honradez			Respeto			Humildad			Sagacidad			Deseabilidad Social		
	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T
90	1.5	61.1	64.6	1.2	26.0	62.3	1.4	21.0	64.1	1.5	13.0	64.8	1.5	118.4	65.2
80	1.0	55.0	60.4	.7	24.0	57.0	.9	19.0	59.0	.7	11.0	56.8	1.0	106.6	59.6
70	.6	48.3	55.8	.4	23.0	54.4	.4	17.0	53.9	.3	10.0	52.8	.5	97.0	55.1
60	.2	42.4	51.8	.2	22.0	51.7	.1	16.0	51.4	.3	10.0	52.8	.1	88.0	50.8
50	-.1	38.0	48.7	-.1	21.0	49.1	-.1	15.0	48.9	-.1	9.0	48.8	-.2	82.0	48.0
40	-.5	33.0	45.3	-.4	20.0	46.4	-.4	14.0	46.3	-.1	9.0	48.8	-.5	76.0	45.1
30	-.7	29.7	43.0	-.6	19.0	43.8	-.6	13.0	43.8	-.5	8.0	44.7	-.7	72.0	43.2
20	-1.0	26.0	40.5	-.9	18.0	41.2	-.6	13.0	43.8	-.9	7.0	40.7	-.9	68.0	41.3
10	-1.2	22.0	37.7	-1.1	17.0	38.5	-1.1	11.0	38.7	-1.1	6.5	38.7	-1.1	63.2	39.1
Media	39.85			21.35			15.45			9.31			86.27		
DE	14.56			3.79			3.94			2.49			21.09		

Puntuaciones de la escala de Deseabilidad Social para Mujeres															
Percentil	Honradez			Respeto			Humildad			Sagacidad			Deseabilidad Social		
	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T
90	1.2	66.8	62.2	1.5	27.8	65.4	1.3	22.0	63.2	1.6	14.0	65.7	1.3	125.1	62.7
80	1.1	64.0	60.5	.8	25.0	58.0	1.1	21.0	60.7	.8	12.0	57.9	1.1	121.0	60.9
70	.9	61.0	58.7	.5	24.0	55.4	.6	19.0	55.6	.4	11.0	54.0	.8	114.3	58.0
60	.6	56.2	55.7	.3	23.0	52.8	.3	17.8	52.5	.0	10.0	50.1	.5	107.4	54.9
50	.1	48.0	50.7	.0	22.0	50.2	.1	17.0	50.5	.0	10.0	50.1	-.1	94.5	49.2
40	-.5	39.0	45.1	-.2	21.0	47.5	-.2	16.0	48.0	-.4	9.0	46.2	-.5	85.0	45.1
30	-.8	34.6	42.4	-.5	20.0	44.9	-.5	15.0	45.4	-.8	8.0	42.2	-.7	80.7	43.2
20	-1.1	29.4	39.2	-.8	19.0	42.3	-1.0	13.0	40.4	-.8	8.0	42.2	-1.0	73.0	39.8
10	-1.4	24.0	35.9	-1.0	18.0	39.7	-1.2	12.0	37.8	-1.2	7.0	38.3	-1.3	65.9	36.6
Media	39.85			21.35			15.45			9.31			86.27		
DE	14.56			3.79			3.94			2.49			21.09		

Puntuaciones de la escala de Deseabilidad Social Población General															
Percentil	Honradez			Respeto			Humildad			Sagacidad			Deseabilidad Social		
	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T	z	Pc	T
90	1.3	65.0	63.0	1.3	27.0	63.3	1.4	21.0	64.1	1.5	13.0	64.8	1.4	122.6	64.1
80	1.1	60.0	60.5	.8	25.0	58.0	.9	19.0	59.0	.8	12.0	57.9	1.0	114.2	60.3
70	.7	54.0	57.4	.5	23.2	54.6	.4	18.0	53.9	.4	11.0	54.0	.6	103.8	56.4
60	.4	47.0	53.5	.2	22.0	51.7	.1	17.0	51.4	.1	10.0	51.2	.2	92.4	51.7
50	-.1	40.0	49.4	.0	22.0	50.2	-.1	16.0	48.9	-.1	9.0	48.8	-.2	87.0	48.1
40	-.5	35.0	45.3	-.4	20.0	46.4	-.4	15.0	46.3	-.4	9.0	46.2	-.5	80.0	45.1
30	-.7	31.0	42.6	-.6	19.0	43.8	-.6	14.0	43.8	-.5	8.0	44.7	-.7	75.0	43.2
20	-1.0	27.0	39.8	-.9	18.2	41.4	-.9	13.0	41.3	-.9	7.0	40.7	-.9	70.0	41.3
10	-1.3	22.0	37.1	-1.1	17.0	38.5	-1.2	11.0	37.8	-1.2	7.0	38.3	-1.2	64.0	37.7
Media	42.56			21.57			15.97			9.56			90.16		
DE	15.58			3.80			3.99			2.53			22.24		